*ABA in Multidimensional Paradigms: A Max/Dep-based account

by

Stanislao Zompì

M.A. in Linguistics
Università di Pisa, 2017

Submitted to the Department of Linguistics and Philosophy
in partial fulfillment of the requirements for the degree(s) of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2023

Author . . . .
Department of Linguistics and Philosophy
August, 2023

Certified by . . . .
David Pesetsky
Ferrari P. Ward Professor of Modern Languages and Linguistics
Thesis advisor

Certified by . . . .
Norvin Richards
Professor of Linguistics
Thesis advisor

Certified by . . . .
Sabine Iatridou
David W. Skinner Professor of Linguistics
Thesis advisor

Certified by . . . .
Jonathan David Bobaljik
Professor of Linguistics, Harvard University
Thesis advisor

Accepted by . . . .
Danny Fox
Head of the Department of Linguistics and Philosophy

*ABA in Multidimensional Paradigms: A Max/Dep-based account

Stanislao Zompì

Abstract

The last decade and a half has witnessed intensive research into *ABA universals—generalizations such as "If a nominative and the corresponding dative have the same exponent, then the corresponding accusative has that exponent, too" (Caha 2009; Smith *et al.* 2019). Most existing work on these universals has only focused on one 'paradigm column' at a time, by checking a given paradigm's nominative singular, accusative singular, and dative singular, for example, with no heed to whether any of the relevant exponents would also show up in that paradigm's nominative plural, accusative plural, or dative plural. However, some recent literature has pointed out that inspecting full paradigms is crucial to our understanding of *ABA, because some classic accounts that derive *ABA column-internally turn out to also make predictions as to what may or may not happen across columns, and those predictions are often incorrect (cf., among others, Christopoulos & Zompì 2022). In this dissertation, I review those incorrect predictions and replace them with a novel generalization specifically concerning *ABA-like effects in multidimensional paradigms. I then set out to derive this generalization by setting up an exponent-selection system wherein exponents may both be underspecified and be overspecified with respect to their *exponenda*, with each of these departures from a perfect match being penalized but not necessarily fatal. In particular, I explicitly implement this intuition in optimality-theoretic terms, via a strict-domination ranking of violable Max and Dep constraints (cf. in particular Ackema & Neeleman 2005; Wolf 2008; Müller 2020), and I show that the resulting system, while restrictive enough to derive the desired generalization, is also powerful enough to afford a natural account of some notoriously unnatural ('morphomic') exponent distributions in the inflection of Germanic pronouns and Romance verbs.

Acknowledgements

I couldn't have written this thesis without the support of my advisors: Jonathan Bobaljik, Sabine Iatridou, David Pesetsky, and Norvin Richards. They've all been incredibly generous and served as inspirational role models, each in their own way. I also couldn't have written this without the following three friends: Christos Christopoulos, with whom I first started working on the key problem; Filipe Hisao Kobayashi, who first suggested that I try mixing under- and over-specification; Enrico Flor, who enthusiastically provided the automated implementation that brought a lot of results within my reach. Donca Steriade, Giorgio Magri, and Niels Kühlert also provided feedback at decisive junctures, as did the audience of MorPhun at MIT.

And then there are other kinds of debts of gratitude. To the MIT Linguistics faculty and staff. To my cohort, to my officemates, to GLRG. To my roommates (especially the legendary lockdown squad). To my family. To all my friends — from the department, from places in Boston, from summer schools and conferences, from my college years, from my hometown. I'm not going to name names as I'd be sure to forget way too many. I believe and hope that you know who you are.

I dedicate this thesis to my nieces, in the hopes that someday they'll forgive me for devoting this summer to writing it up rather than to taking them both to the beach.

Contents

Chapter 1

Setting the stage

1.1 Natural and less natural classes

Some morphemes have consistent exponence across all of their contexts of appearance—like the English gerund, for instance, which is always exponed by *-ɪŋ* no matter what verbal head it has as its sister. Morphology textbooks often remark that there would be nothing wrong with a world in which every morpheme behaved in this way, but the actual world appears to be more complicated. In particular, one can frequently observe cases in which what seem to be occurrences of one and the same abstract morpheme are exponed by quite different phonological representations (*allomorphs*) depending on the context where those occurrences find themselves. This is the phenomenon known as *contextual allomorphy*.

Sometimes, the choice between different allomorphs is determined by the context's phonological properties; sometimes, by its morphosyntactic properties. This dissertation is primarily concerned with the latter case. In particular, it tackles an old question in the theory of allomorphy: What are the ways in which morphosyntactically conditioned allomorphy may licitly divide up a paradigm, i.e. a set of contexts where a given abstract morpheme may appear? What are the possible and the impossible distributions of different morphosyntactically conditioned allomorphs of a given morpheme?

Here is a straightforward answer one could have reasonably hoped would prove true: Each allomorph of morpheme M has a certain set of morphosyntactic features as its context specification, and expones M in all and only the contexts that include that set of features.

(1) $$\text{M} \rightsquigarrow \begin{cases} a \mathbin{/} \underline{\quad} \, \varphi, \psi \\ b \mathbin{/} \underline{\quad} \, \chi \\ \vdots \end{cases}$$

Thus, given (1), *a* expones M in all and only the contexts including $\{\varphi, \psi\}$, *b* in all and only those including $\{\chi\}$, etc. Each allomorph therefore picks out a natural class — the set of contexts which include its context specification.

Now suppose, for example, that there exists both a singular-specific number feature $\#_{\text{SG}}$ and a plural-specific number feature $\#_{\text{PL}}$. Then the Turkana paradigm in (2) would seem to be in line with the straightforward answer we were initially hoping for. The root allomorph *-bun-* expones the root $\sqrt{\text{COME}}$ in all and only the contexts including $\{\#_{\text{SG}}\}$, while the root allomorph *-pon-* does in all and only the contexts including $\{\#_{\text{PL}}\}$ — just as (3) would predict.[1]

(2) Turkana 'come' (Dimmendaal 1983: 290)

	SG	PL
1	à-**bun**-it̀	kì-**pon**-it̀
2	ì-**bun**-it̀	ì-**pon**-it-ò
3	è-**bun**-it̀	è-**pon**-it-ò

(3) $\sqrt{\text{COME}} \rightsquigarrow \begin{cases} \textit{bun} \;/\; \_\_\; \#_{\text{SG}} \\ \textit{pon} \;/\; \_\_\; \#_{\text{PL}} \end{cases}$

Many instances of allomorphy, however, do not cut through paradigms so neatly along the seams between natural classes. Rather, they look more like (4) and (5).

(4) Italian 'go', PRES.IND

	SG	PL
1	vado	andiamo
2	vai	andate
3	va	vanno

(5) Languedocien 'go', PRES.IND (Maiden 2018: 195)

	SG	PL
1	vau	anam
2	vas/vai	anatz
3	va	van

If first and second person share a "participant"-specific feature π_{PART} to the exclusion of the third person, then the dark-grey root allomorph *an(d)-* in each of these paradigms does apply to a natural class — the class of all and only the contexts that include $\{\pi_{\text{PART}}, \#_{\text{PL}}\}$. The problem arises, however, when it comes to characterizing the distribution of the light-grey, root allomorph *va-*, as there doesn't seem to be any set of features that 1SG, 2SG, 3SG, and 3PL contexts all share to the exclusion of 1PL and 2PL.

[1] Here I'm not controlling (e.g. via tests based on shared idioms or ellipsis parallelism) for the possibility that these may involve different lexical entries. Let's set the confound aside, for the argument's sake.

(6) $\sqrt{\text{GO}} \rightsquigarrow \begin{cases} \textit{an(d)-} \,/\, \_\_\ \pi_{\text{PART}}, \#_{\text{PL}} \\ \textit{va-} \,/\, \_\_\ ?? \end{cases}$

1.2 Pāṇinian ordering and morphomes

The light-grey set of contexts in (4)–(5) looks rather like one that is derived by subtraction — the set of all contexts that are not participant plurals. And indeed, most approaches to the problem pursue precisely this intuition.

On such approaches, each allomorph of morpheme M still has some set of features as its context specification, and is eligible to expone M in a context if and only if that context includes that set of features. However, an allomorph's eligibility is now a necessary but crucially no longer a sufficient condition for its actual application. What can stand in the way between eligibility and actual application is a mechanism of competition: if a given context is such that multiple allomorphs are all eligible in it, a principle of Pāṇinian ordering (variously referred to as the Elsewhere Principle, specificity, etc.) ensures that only the most specific of those allomorphs (roughly, the one that's applicable to the fewest contexts) will actually apply.

On this approach, we can maintain that the context specification of the dark-grey root allomorph *an(d)-* is indeed $\{\pi_{\text{PART}}, \#_{\text{PL}}\}$ as in (6), and then say that the light-grey root allomorph *va-* is simply the elsewhere case, i.e. has an empty context specification.

(7) $\sqrt{\text{GO}} \rightsquigarrow \begin{cases} \textit{an(d)-} \,/\, \_\_\ \pi_{\text{PART}}, \#_{\text{PL}} \\ \textit{va-} \end{cases}$

As a result, the *va-* is now eligible to expone $\sqrt{\text{GO}}$ in all contexts, but only gets to actually do so when it is not outcompeted by the more specific *an(d)-*, i.e. anywhere except in participant-plural contexts.

In this system, the set of contexts in which an allomorph is eligible is still a natural class, but the set of contexts in which it actually applies no longer need be — it might just be the result of subtracting one or more smaller natural classes out of a larger natural class. Even so, the system remains fairly restrictive: if an allomorphic alternation partitions a paradigm into two sets of contexts, then at least one of those two (the one in which the more specific allomorph applies) still has to be a natural class. This means, among other things, that the analyses we hone in on based on the allomorph distribution in one paradigm fragment make predictions about the allomorph distribution to be found in the rest of the paradigm.

For example, given the paradigm fragments in (4)–(5), we concluded in (7) that the root allomorph *an(d)-* must have at least π_{PART} and $\#_{\text{PL}}$ in its context specification, which in turn now predicts that that same allomorph should never apply in singular or non-participant contexts in any other part of the paradigm either. Expanding the paradigm fragment to encompass not just the present indicative, but also the present subjunctive, bears out this prediction in Italian, where the allomorph *and-* does remain confined to participant-plural contexts across both moods.[2]

(8) Italian 'go', PRES.IND and PRES.SJNV

	INDICATIVE		SUBJUNCTIVE	
	SG	PL	SG	PL
1	vado	andiamo	vada	andiamo
2	vai	andate	vada	andiate
3	va	vanno	vada	vadano

However, things don't go as smoothly in Languedocien, where the *an-* allomorph turns out to apply in all of the present subjunctive — not just in participant-plural contexts, but in singular and non-participant contexts too. In the Languedocien present, the allomorph *va-* therefore applies in the singular and 3PL of the indicative, while the allomorph *an-* applies in the participant-plural contexts of the indicative as well as in all the subjunctive. This means, distressingly enough, that neither allomorph picks out a natural class.

(9) Languedocien 'go', PRES.IND and PRES.SJNV (Maiden 2018: 195)

	INDICATIVE		SUBJUNCTIVE	
	SG	PL	SG	PL
1	vau	anam	ane	anem
2	vas/vai	anatz	anes	anetz
3	va	van	ane	anen

This is what Aronoff (1994) refers to as a *morphome* — an allomorph distribution partitioning a paradigm into equally "unnatural" sets of contexts.

1.3 Proliferating morphomes?

Faced with morphomes, we might be tempted to give up on any hope of formulating a restrictive theory of allomorph distribution, and conclude that an allomorphic alternation can divide up a paradigm into any arbitrary partition (or maybe even cover) thereof.

[2] In fact, things break down even in Italian as soon as we venture beyond the present of the finite verb. Once again, I'm setting these complications aside, as the current examples are just for illustrative purposes.

Indeed, there are authors who do entertain this conclusion. At the beginning of his book, for instance, Maiden (2018: 11–12) illustrates the notion of a morphome by way of the fictitious example in (10), which he characterizes as "deliberately quite extreme," but which his general approach does not categorically rule out.

(10) An invented morphome (Maiden 2018: 12)

	INDICATIVE			SUBJUNCTIVE		
	PAST	PRES	FUT	PAST	PRES	FUT
1SG	karep-i	karep-j	perak-l	karep-ɲ	karep-ɔ	karep-m
2SG	karep-k	karep-p	karep-l	karep-θ	perak	karep-l
3SG	perak-ɲ	karep-s	karep-m	karep-ɲ	karep-a	karep-h
1PL	karep-i	perak-θ	karep-t	karep-v	karep-u	karep-a
2PL	karep-d	karep-v	karep	karep-v	karep-f	perak-a
3PL	karep-ɲ	karep-y	karep-t	perak-v	karep-u	karep-i

There are reasons, however, not to adopt quite such a drastic reaction. As a methodological principle, weakening our theory to the point of virtual empirical vacuity would of course be warranted only if we had clear signs that just about anything goes. However, it turns out that's not the case. For one thing, no attested morphome picks out such a set of contexts as the one highlighted in (10)—a set so unnatural that not even any of its non-singleton subsets appear to be natural classes. In fact, we don't even need to push so hard in order to find systematic cross-linguistic gaps in attested allomorph distributions. As we will see in more detail in §2.1.1, several such gaps have recently been uncovered even in paradigms of as few as three cells—see (11) for one of many generalizations illustrating this point, all of them evidently unexpected on an "anything goes" type of approach.

(11) *Smith* et al.*'s (2019) generalization*
No stem allomorph can appear in both a nominative and the corresponding dative to the exclusion of the corresponding accusative.

An alternative reaction in the literature has essentially set morphomes aside and focused instead on cross-linguistic gaps like (11), which a restrictive approach does have a shot at deriving. The problem, however, is that, while the first type of reaction has arguably led to weakening theories too much, this second type has led to further strengthening of a theory that already seemed too strong to begin with. In particular, in order to derive generalizations like (11), part of the literature has combined the already restrictive theory of natural classes and Pāṇinian ordering together with increasingly restrictive assumptions about the distribution of morphosyntactic features, i.e. about what natural classes are there to begin with. Unsurprisingly, rather than chipping away at the morphome problem, this has made it bigger: now that new assumptions have made certain sets of contexts no longer characterizable as natural classes, even paradigms that Pāṇinian ordering used to easily capture via appeal to those classes are suddenly becoming, *de facto*, morphomic.

We will discuss this issue at great length in §2.2, but to get a first taste of it, take a look, for instance, at the Albanian paradigm in (12).

(12) Albanian distal demonstrative (Newmark, Hubbard & Prifti 1982: 122)

	MASCULINE		FEMININE	
	SG	PL	SG	PL
NOM	a-**i**	a-**t**-a	a-**j**-o	a-**t**-o
ACC	a-**t**-ë	a-**t**-a	a-**t**-ë	a-**t**-o
DAT	a-**t**-ji	a-**t**-yre	a-**sa**-j	a-**t**-yre

At first glance, this paradigm would look like a poster child for the elegance and power of Pāṇinian ordering: rather than listing all the contexts where the stem allomorph *-t-* applies, we can just appropriately specify *-sa-* for DAT.SG.FEM contexts and *-i/j-* for NOM.SG contexts, so as to leave *-t-* as an elsewhere to apply whenever there is no better match.

(13) $$\text{DIST} \rightsquigarrow \begin{cases} \textit{-sa-} \;/\; \_\_\; \gamma_{\text{FEM}},\ \#_{\text{SG}},\ \varkappa_{\text{DAT}} \\ \textit{-i/j-} \;/\; \_\_\; \#_{\text{SG}},\ \varkappa_{\text{NOM}} \\ \textit{-t-} \end{cases}$$

However, as we'll see in §2.1.2, an attractive account of the generalization in (11) involves the proposal that the nominative has no proprietary case features, i.e. that there is no such thing as the $\varkappa_{\text{NOM}}$ referenced in (13).[3] On this approach, so-called nominative contexts are therefore no longer a natural class, nor, as a consequence, are nominative-singulars: there is no set of features that all and only such contexts share. The problem with this assumption, as we will discuss at length in §2.2, is that it makes the analysis in (13) unstatable: if nominative-singular contexts do not form a natural class, it becomes impossible to specify *-i/j-* for them, and hence to correctly capture the distribution of *-i/j-* and *-t-*. As a result, even a paradigm that seemed very much not morphomic has, in the end, turned out to be.

[3] Or, equivalently, that $\varkappa_{\text{NOM}}$ should effectively be understood as just $\varkappa$ — a nondescript case feature that the nominative shares with all other cases.

1.4 The challenge we are faced with

This preliminary outline, albeit brief and cursory, should hopefully have given the reader a sense of the kind of challenge we're faced with. On the one hand, we need to weaken our theory of allomorph distribution so that we can finally capture attested morphomes such as (9) in the Languedocien verb and, under certain assumptions, (12) in the Albanian demonstrative. On the other hand, we also don't want to weaken our theory so much as to allow for Maiden's (2018) imaginary monstrous morphome in (10), nor as to make it impossible to derive (with the right ancillary assumptions about the distribution of morphosyntactic features) cross-linguistic generalizations such as Smith *et al.*'s (2019) in (11). The rest of this dissertation takes on this challenge.

Chapter 2

*ABA, an account of it, and its empirical problems

We closed the previous chapter by establishing the need for a middle ground between, on the one hand, an overly permissive "free-morphomes" approach that would allow for any possible allomorph distribution whatsoever and, on the other hand, an overly restrictive "morphome-free" approach that would make each allomorph eligible only in a natural class of contexts, with Pāṇinian ordering resolving some eligibility overlaps. In this chapter, we will nail down just how much more restrictive than the first approach, and how much more permissive than the second, our new approach is going to have to be. To this end, we will zoom in on a narrow empirical domain where the excesses of both previous approaches will be particularly evident — the domain of stem allomorphy in (pro)nominal paradigms crossing case and number.

2.1 A generalization and a popular account of it

2.1.1 *ABA in case-conditioned stem allomorphy

The approach that more immediately runs into a problem in the domain at hand is the "free-morphomes" approach, whose pervasive overgeneration issues have recently been brought out by a flurry of empirical generalizations commonly referred to as **ABA* (following Bobaljik's 2012 terminology). These are cross-linguistic universals of the following form: "No allomorph in any language applies in both context C_1 and context

C_3 without also applying in context C_2."[4] In particular, (pro)nominal stem allomorphy[5], and not just obeys generalizations of this sort, too — a discovery due to Smith *et al.* (2019).

(14) *Smith* et al.*'s (2019) generalization (I)*
In nominative-accusative languages, no stem allomorph can appear in both a nominative and the corresponding dative to the exclusion of the corresponding accusative.

Smith *et al.* (2019) propose this generalization based on a wide cross-linguistic survey of case-conditioned stem allomorphy in pronouns. They discover that, if one adopts the order NOM ≺ ACC ≺ DAT, a search through the nominative-accusative languages in their sample turns up various instances of AAA (no case-conditioned stem allomorphy at all, as in Turkish in (15)), of ABB (ACC and DAT sharing the same stem allomorph to the exclusion of NOM, as in Brahui), and of AAB (NOM and ACC sharing the same stem allomorph to the exclusion of DAT, as in Old Saxon) — as well as very rare instances of ABC (a different stem allomorph for each of the three cases, as in Albanian). However, there are no instances of ABA.

(15)

	Turkish 1PL	Brahui 1SG.F	Old Saxon 3SG.F	Albanian DEM.SG.F	*
NOM	**biz**	**ī**	**si**-u	-**j**-o	
ACC	**biz**-i	**kan**-e	**si**-a	-**t**-ë	
DAT	**biz**-e	**kan**-ki	**i**-ru	-**sa**-j	

Outside the domain of nominative-accusative languages, Smith *et al.* (2019) found a twin generalization in force across absolutive-ergative languages.

(16) *Smith* et al.*'s (2019) generalization (II)*
In absolutive-ergative languages, no stem allomorph can appear in both an absolutive and the corresponding dative to the exclusion of the corresponding ergative.

The presence of AAA, ABB, AAB, and ABC patterns with respect to the order ABS ≺ ERG ≺ DAT (a presence to be contrasted once again with the absence of ABA) is illustrated in (17).

[4] Although here I'm stating it as just a constraint on contextual allomorphy, "*ABA" has also been used to describe similar constraints on affixal syncretism of the kind discovered (but not yet referred to by that name) by Caha (2009).

[5] It's also worth clarifying that, here and elsewhere, what we mean by "(contextual) allomorphy" is, in effect, suppletive allomorphy. That is, we're excluding surface-phonological allomorphy from the purview of our generalizations.

(17)

	Basque 2SG	Georgian 3SG	Wardaman 3SG	Khinalug 1SG	*
ABS	**hi**	**is**	**narnaj**	**zɨ**	
ERG	**hi**-k	**ma**-n	**narnaj**-(j)i	**yä**	
DAT	**hi**-ri	**ma**-s	**gunga**	**as**(-ɨr)	

Here I'll also follow Smith *et al.* in unifying the generalizations in (14) and (16) by grouping NOM and ABS together under the label of "unmarked case," and ACC and ERG together under the label of "dependent case" — two notions borrowed from Marantz's (1991) configurational theory of structural case assignment (cf. also Yip, Maling & Jackendoff 1987).

(18) *Smith* et al.*'s (2019) generalization (unified version)*
No stem allomorph can appear in both an unmarked case and the corresponding dative to the exclusion of the corresponding dependent case.

I will use these terms here because they permit more concise and readable prose, but the correctness of the dependent-case view of ACC and ERG (which I offered some additional arguments for in Zompì 2017, 2019) will be irrelevant to the proposals I'll make here.

2.1.2 A popular account

An "anything goes"–type of approach obviously has no way to derive the generalization in (18). By contrast, the more restrictive approach holds some more initial promise. In particular, building on work by Bobaljik (2012) on a similar *ABA pattern found with the triple positive–comparative–superlative in the domain of adjectival morphology (cf. Caha 2009, 2013 on syncretism among case markers), Smith *et al.* (2019) propose to explain the generalization in (18) as follows.

As a first ingredient of their account, they assume a version of the restrictive approach to allomorph distribution that we have been discussing so far, which I articulate more explicitly in (19)–(20). (Here I adopt the following terminology: in a rule of exponence "M ⇝ *e* / __ C," *e* is the *exponent*, M is its *target specification*, and C its *context specification*.)

(19) *Underspecification*
Exponent *e* is eligible to expone morpheme M iff
a. the target specification of *e* is a subset of M
b. and the context specification of *e* is part of the context of M.

(20) *Pāṇinian ordering*
If exponents e_1 and e_2 are both eligible to expone morpheme M, and e_1 is more specific than e_2 (i.e. the set of morphemes that e_1 is eligible to expone is a proper subset of the set of morphemes that e_2 is eligible to expone), e_2 doesn't expone M.

The second ingredient of Smith *et al.*'s (2019) account is a substantive assumption about the decomposition of traditional cases into morphosyntactic features. Specifically, they assume what Caha (2013) calls a *cumulative decomposition*: the case features of the unmarked case constitute a proper subset of those of the dependent case, which in turn constitute a proper subset of the case features of the dative, as in (21)–(22).[6]

(21) *Cumulative decomposition in nominative-accusative languages*

a. Case features of NOM: ∅
b. Case features of ACC: $\{\varkappa_{\text{ACC}}\}$
c. Case features of DAT: $\{\varkappa_{\text{ACC}}, \varkappa_{\text{DAT}}\}$

(22) *Cumulative decomposition in absolutive-ergative languages*

a. Case features of ABS: ∅
b. Case features of ERG: $\{\varkappa_{\text{ERG}}\}$
c. Case features of DAT: $\{\varkappa_{\text{ERG}}, \varkappa_{\text{DAT}}\}$

To see how these ingredients jointly derive *ABA, consider, for example, the paradigm {NOM_{X}, ACC_{X}, DAT_{X}}, where the subscript X appearing throughout is meant to indicate that the three contexts differ only in case features, with all their other features held constant.

(23)
a. Suppose exponent *b* expones morpheme M in context ACC_{X}, and exponent *a* expones M in context DAT_{X}.
b. Since *b* applies in ACC_{X}, and ACC_{X} is part of DAT_{X}, *b* will also be eligible in DAT_{X}.
c. To block *b* from applying in DAT_{X}, *a* must therefore be more specific than *b*.
d. But then, if *a* were eligible to apply in ACC_{X}, *a* would block *b* there, too.
e. Since that is by hypothesis not the case, *a* must be eligible in DAT_{X} but ineligible in ACC_{X} — and this can only be if *a* is specified for a context containing $\varkappa_{\text{DAT}}$.
f. However, this in turn entails that *a* is ineligible in the $\varkappa_{\text{DAT}}$-free context NOM_{X} — hence deriving the *ABA generalization in (14), as desired.

With ABS_{X}, ERG_{X} and $\varkappa_{\text{ERG}}$, respectively substituted for NOM_{X}, ACC_{X}, and $\varkappa_{\text{ACC}}$, the account is replicated for ergative systems.

[6] So long as the relevant subsethood relations hold, any number of additional features can be added without affecting the validity of the account.

2.1.3 Converging evidence for a cumulative decomposition

Underspecification and Pāṇinian ordering are not the only ingredients of Smith *et al.*'s (2019) account that had been independently argued for. In fact, cumulative decompositions of cases also have a history of independent empirical backing. In this subsection, I'll focus on two strands of evidence commonly adduced in their favor.

Case competition in free relatives

The first type of evidence comes from case-mismatch resolution in free relative constructions. In such constructions, the matrix clause and the relative clause may impose conflicting case requirements on the shared constituent. While some languages resolve such conflicts based solely on which clause each case is assigned in (by always giving priority to the case assigned inside the matrix, or always to the case assigned inside the relative) and while other languages disallow such conflicts altogether, there are famously also some languages that resolve such conflicts on the basis of a "strength hierarchy" of cases (Harbert 1978, 1982, 1992; Pittner 1995; Vogel 2001; Daskalaki 2009; Bergsma 2019). In all the languages of this type, the strength hierarchy at play is: DAT ≫ DEP ≫ UNM.[7]

(24) Gothic (Harbert 1992: 109, 111)

a. bi [þamm -ei anafulhun þai sinistans]
by(+DAT) which.DAT COMP commended(+ACC) the.NOM elders
'according to what the elders commended' DAT ≫ ACC

b. [Þan -ei frijos] siuks ist
who.ACC COMP you.love sick is
'The one you love is sick.' ACC ≫ NOM

c. [Þamm -ei leitil fraletada] leitil frijod
who.DAT COMP little is.forgiven little loves
'The one to whom little is forgiven loves little.' DAT ≫ NOM

The cumulative decomposition of case can help us make sense of this pattern, together with the assumption that languages like Gothic resolve the mismatch by forming a union of the two mismatching feature sets (much in the spirit of unification in non-transformational syntactic frameworks): if DAT ⊂ DEP ⊂ UNM, then DAT ∪ DEP = DAT, DEP ∪ UNM = DEP, DAT ∪ UNM = DAT. See Bergsma (2019) for a grafting-based implementation of this idea within a nanosyntactic framework.

[7] In fact, all such languages are nominative-accusative — something I take to be an accident.

Accessibility hierarchies in case-discriminating agreement

The second strand of evidence for cumulative case decompositions comes from the following universal hierarchy of accessibility to predicate–argument φ-agreement, which Bobaljik (2008) has established by drawing on Moravcsik (1978).

(25) *The Moravcsik–Bobaljik accessibility hierarchy*
Unmarked case ≫ Dependent case ≫ Lexical/Inherent case

The hierarchy should be read as follows: If, in some language, predicates φ-agree with anything, they φ-agree with some or all of the noun phrases bearing unmarked case; if, in some language, predicates φ-agree with anything other than unmarked noun phrases, they φ-agree at least with some or all of the noun phrases bearing dependent case.

An example of a language in which only nominative NPs are accessible to φ-agreement is Hindi, illustrated in (26).

(26) Hindi (from Bobaljik 2008: 309)

a. raam-ne roṭii khaayii thii
Ram(MASC)-ERG bread(**FEM**).NOM eat.PF.**FEM** be.PF.**FEM**
'Ram had eaten bread.' ERG NOM

b. siita-ko laṛke pasand the
Sita(FEM)-DAT boys(**MASC.PL**).NOM like be.PST.**MASC.PL**
'Sita likes the boys.' DAT NOM

We can contrast this with Nepali in (27), where both nominative and ergative NPs are accessible to φ-agreement. As Bobaljik argues, however, no language instantiates the "mirror-image" of Hindi or Nepali, i.e. a system in which only ergative, only dative, or only ergative-or-dative subjects get φ-agreed with.

(27) Nepali (Bobaljik 2008: 310–311)

a. maile yas pasal-mā patrikā kin-ē
1SG.ERG DEM.OBL store-LOC newspaper.NOM buy-PST.**1SG**
'I bought the newspaper in this store.' ERG NOM

b. malāī timī man par-ch-au
1SG.DAT **2MASC.HON**.NOM liking occur-NPST-**2MASC.HON**
'I like you.' DAT NOM

Caha (2009: 296) proposes to derive the Moravcsik–Bobaljik generalization from the cumulative decomposition of case. He reasons: "the agreement hierarchy could follow from the interaction of agreement computation and decomposed case under the assumption that case layers intervene between the Probe and the Goal. […] Specifically, if a particular layer of case counts as a blocker of the Agree relation, all cases which contain that layer will likewise block the relation." So, if only the dative acts as a blocker in language *L*, dative noun phrases will not be agreed with in *L*, but both dependent- and unmarked-case noun phrases will be — the situation we've seen in Nepali in (27). By contrast, if in language *L′* it's the dependent case that acts as a blocker, then agreement in *L′* will access neither dependent-case nor dative noun phrases, because, given the cumulative decomposition, both will include the dependent-case layer. *L′* will thus only allow agreement with noun phrases in unmarked case — as we've seen is the case for Hindi in (26). All of Bobaljik's (2008) other implicational universals follow in a similar way.

2.2 Some multidimensional puzzles

So far, this chapter has painted a rather rosy picture: Smith *et al.*'s (2019) account appears to derive a nontrivial generalization from the combination of just two assumptions, which have both been independently argued for. Indeed, it would be hard to improve on this if the account really predicted just the desired generalization. However, closer inspection reveals that the account predicts something stronger than that — in fact, something far too strong.

The problem has already been noticed in passing by McFadden (2018: 25) and Smith *et al.* (2019: 1043fn21) themselves, and then discussed more extensively (although with an exclusive focus on nominative-accusative languages) by Christopoulos & Zompì (2022) and Caha (2023). It is an instance of something we already remarked on in §1.2: due to the very restrictiveness of the system we've been assuming for allomorph selection, the analyses we hone in on based on one fragment of a paradigm entail strong predictions about the rest of that paradigm as well. Specifically, we'll see that, while Smith *et al.*'s (2019) account makes the right predictions about unidimensional paradigm fragments (ones whose members only differ from each other in case features), it makes fundamentally incorrect predictions when it comes to the multidimensional paradigms that those fragments are a part of (larger paradigms whose members differ both in case and along other inflectional dimensions, such as grammatical number or gender).

2.2.1 Multidimensional puzzle #1: Non-elsewhere unmarked-case stems

Let's consider, for example, a unidimensional ABB or ABC pattern, and let's try to replicate the same reasoning we outlined in §2.1.2. Since exponent *a* applies in context UNM_X, and UNM_X is part of DEP_X, *a* will also be eligible in DEP_X. To block *a* from applying in DEP_X, exponent *b* must therefore be more specific than *a*; but then, if *b* were eligible in UNM_X, *b* would block *a* there too, contrary to fact. Hence *b* must be eligible in DEP_X but ineligible in UNM_X, i.e. must be specified for a context containing a dependent-case–specific feature ($\varkappa_{\text{ACC}}$ or $\varkappa_{\text{ERG}}$). Crucially, this predicts that *b* should also not be eligible to apply in any $\varkappa_{\text{ACC}}$-less or $\varkappa_{\text{ERG}}$-less context (i.e. any nominative or absolutive context) anywhere else in the rest of the full multidimensional paradigm.

Sometimes, things do work that way. For example, restricting our focus to paradigm bipartitions for the moment, we may see the same ABB pattern cutting across SG and PL, as in the Rutul paradigm in (28), or an ABB pattern in one dimension whose exponent *a* shows up in an AAA pattern in another dimension, as in the Icelandic paradigm in (29).

(28) Rutul ‘who’ (Ibragimov 1978: 86)

$$\text{WHO} \rightsquigarrow \begin{cases} \textit{hal} \;/\; \_\_\; \varkappa_{\text{ERG}} \\ \textit{vuš} \end{cases}$$

	SG	PL
ABS	**vuš**	**vuš**-er
ERG	**hal**-a	**hal**-dəbiš-æ
DAT	**hal**-də	**had**-dəbiš-də

(29) Icelandic (Einarsson 1949: 68)[8]

$$2 \rightsquigarrow \begin{cases} \textit{ykk-} \;/\; \_\_\; \#_{\text{PL}},\ \varkappa_{\text{ACC}} \\ \textit{þ-} \end{cases}$$

	2SG	2PL
NOM	**þ**-ú	**þ**-ið
ACC	**þ**-ig	**ykk**-ur
DAT	**þ**-ér	**ykk**-ur

But this prediction is also quite often falsified. For instance, Christopoulos & Zompì (2022) offer various examples of the pattern in (30), with an ABB in the singular whose exponent *b* shows up throughout the plural. Specifically, they find such patterns in “the 1st-person pronoun in Modern Greek (Holton *et al.* 2012: 113) and Latvian (Prauliņš 2012: 54), the demonstrative pronouns in Eastern Armenian (Dum-Tragut 2009: 130–1), Gothic (Braune & Heidermanns 2004: 134–5) […], Old English (Hogg & Fulk 1992: 192–5), Old Norse (Barnes 2008: 63–64), Sanskrit (Mayrhofer 1978: 58–59) and Tocharian (Krause & Slocum 2007–10: [§31]), the Gothic relative pronoun (Braune & Heidermanns 2004: 136), the Latvian ‘emphatic’ pronoun *pats* (Prauliņš 2012: 465) and the [feminine] 3rd-person pronou[n] of Afrikaans (Donaldson 1993: 123)” (Christopoulos & Zompì 2022: 10).

8 Also the honorific second person: *þér* ~ *iður* ~ *iður* (Einarsson 1949: 68). A similar second-person paradigm is also found in Thuringian (Spangenberg 1990: 281).

(30) Icelandic 'the/this/that' (MASC) (Einarsson 1949: 70)

	SG	PL
NOM	**s**-á	**þ**-eir
ACC	**þ**-ann	**þ**-á
DAT	**þ**-eim	**þ**-eim

Christopoulos & Zompì refer to stems like *s*- in the paradigm in (30) as "non-elsewhere nominative stems" (abbreviated as "NENSs"), but the pattern is also instantiated in several absolutive-ergative languages: apart from the Kidero Tsez dialect in (31), it's most clearly found in the paradigm for 'girl' in Tabasaran (Babaliyeva 2013: 23–24, 26; cf. also Alekseev & Šikhaleva 2003: 44, 53, 109), and plausibly also in the reflexive of Archi (Bond & Chumakina 2016: 69) and in the paradigm for 'day' in Lezgian (Haspelmath 1993: 80). This is why here I'll switch to the alignment-neutral label of "non-elsewhere unmarked-case stems."

(31) Tsez dialect of Kidero addressee-proximal (class II) (Imnajšvili 1963: 110)

	SG	PL
ABS	**how**-žo	**hem**-žedi
ERG	**hem**-e-l'ā	**hem**-žedā
DAT	**hem**-e-l'ár	**hem**-žedár

Non-elsewhere unmarked-case stems are also found in two additional, rarer patterns involving unidimensional ABBs — the one in (32) and the one in (33), both only attested in nominative-accusative languages so far.[9]

(32) Eastern Frisian (Matras & Reershemius 2003: 22)

	3M.SG	3PL
NOM	**h**-äi	**z**-äi
ACC	**h**-um	**h**-ör
DAT	**h**-um	**h**-ör

(33) Yiddish (Jacobs 2005: 185)

	1SG	1PL
NOM	**ix**	**m**-ir
ACC	**m**-ix	**undz**
DAT	**m**-ir	**undz**

[9] The pattern in (32) is also found in the unrelated second-person paradigm of Pfälzisch (Palatinate) dialects (Green 1990: 254). As for the pattern in (33), I'll come back to it (and especially to the question of just how widespread it really is) in §4.3.1.

Finally, non-elsewhere unmarked-case stems are also attested in both of the only two unidimensional ABC patterns that have been documented so far: these are the Albanian pattern in (34) from the nominative-accusative camp and the Khinalug pattern in (35) from the absolutive-ergative camp.

(34) Albanian, feminine distal demonstrative (Newmark, Hubbard & Prifti 1982: 122)

	SG	PL
NOM	a-**j**-o	a-**t**-o
ACC	a-**t**-ë	a-**t**-o
DAT	a-**sa**-j	a-**t**-yre

(35) Khinalug (Khvtisiashvili 2013: 125)

	1SG	1PL.EXCL
ABS	**zɨ**	**y**-ir
ERG	**y**-ä	**y**-ir
DAT	**as**(-ir)	**ʃır**-ì

2.2.2 Multidimensional puzzle #2: Non-elsewhere core-case stems

The problem carries over to AAB patterns, too.

By the same logic we deployed in §2.1.2 and §2.2.1, given a unidimensional AAB, we should predict that exponent *b* should be specified for contexts containing $\varkappa_{\text{DAT}}$, and hence ineligible to apply in any $\varkappa_{\text{DAT}}$-less (i.e. NOM/ABS or ACC/ERG) contexts. Again, sometimes things do work out that way, as in the Old High German paradigm in (36), where the same AAB pattern cuts across singular and plural.

(36) Old High German (Braune 2004: 243)

$$3\text{F} \rightsquigarrow \begin{cases} i\text{-} \ / \ \_\_ \ \varkappa_{\text{DAT}} \\ si\text{-} \end{cases}$$

	3F.SG	3F.PL
NOM	**sī**, **si**(-u)	**si**-o
ACC	**si**-a, **si**-e	**si**-o
DAT	**i**-ru, **i**-ro	**i**-m, **i**-n

But again, things often work differently. For example, certain neuter nouns in Latin appear to instantiate the pattern in (37), with an AAB in the singular whose exponent *b* shows up in both NOM and ACC in the plural (cf. McFadden 2018: 6–7; Christopoulos & Zompì 2022: §5.1; Caha 2023: §6 for similar examples from Classical Greek and Slovenian, and for further discussion).[10] Following the model of "non-elsewhere nominative stems"

[10] I'll come back to the significance of this pattern for Christopoulos & Zompì's (2022) theory in §2.3.

and "non-elsewhere unmarked-case stems", we might refer to stems like *femur* in (37) as "non-elsewhere core-case stems."

(37) Latin 'thigh' (Weiss 2009: 240)

	SG	PL
NOM	**femur**	**femin**-a
ACC	**femur**	**femin**-a
DAT	**femin**-ī	**femin**-ibus

Non-elsewhere core-case stems, too, are found on both sides of the accusative/ergative divide. In fact, an even clearer example is furnished by split-ergative Megrelian in (38).[11]

(38) Megrelian
(Rostovtsev-Popiel 2020: 544)

	1SG	1PL
NOM	**ma**	**čki**(-n)
ERG	**ma**	**čki**(-n)
INSTR	**čki**-m-it	**čki**-n-it

Finally, two additional patterns involving non-elsewhere core-case stems are instantiated by stages of the history of German — see the paradigms in (39) and (40).

(39) Old High German
(Braune 2004: 243)

	3N.SG	3N.PL
NOM	**i**-ʒ	**si**-u
ACC	**i**-ʒ	**si**-u
DAT	**i**-mu/**i**-mo	**i**-m/**i**-n

(40) Modern German

	3M.SG	3PL
NOM	**er**	**sie**
ACC	**ih**-n	**sie**
DAT	**ih**-m	**ih**-nen

2.2.3 Multidimensional puzzle #3: Non-elsewhere singular stems

Finally, yet another potential instance of the same problem would arise if we were to adopt a cumulative decomposition, not just for case, but also for number. This is an issue

[11] I'm including the instrumental instead of the dative in (38) because the dative is metasyncretic with both absolutive and ergative in pronouns.

I've purposely skirted so far; in particular, I haven't explicitly ruled out the possibility of an exponent only being eligible in singular contexts without being eligible in plural ones. Admitting this possibility, however, is not an uncontroversial move.

Corbett (2000), for example, observes that, whenever there is a language-wide formal-markedness asymmetry between singular and plural, it's always the singular that has zero-marking. Furthermore, in the domain of agreement, Béjar & Rezac (2009) note that there seem to be probes selectively searching for plural (i.e. probes able to search past a closer, singular goal to agree with a farther, plural one) but no probes selectively searching for singular (i.e. no probes able to search past a closer, plural goal to agree with a farther, singular one). In a similar spirit, Nevins (2011) further notices the absence of Number Case Constraint effects. Béjar & Rezac (2009), Nevins (2011), and Preminger (2017) thus take all of this to point to a cumulative decomposition for number too, such that the number features of the singular constitute a proper subset of those of the plural.[12]

(41) a. Number features of SG: $\varnothing$
 b. Number features of PL: $\{\#_{\text{PL}}\}$

If we were to assume something along the lines of (41), then we would run up against a similar problem to the ones we discussed in §§2.2.1–2.2.2. For example, in the paradigms in (42)–(45), we find an exponent *a* in ACC/ERG/DAT.SG but a different exponent *b* in ACC/ERG/DAT.PL. Given (41), this pattern should only be derivable by specifying *b* for contexts containing $\#_{\text{PL}}$ — but this would predict *b* to never be found in singular contexts, contrary to fact.[13] Stems like *enn-* in Malayalam (42), *hon-* in Basque (43), *i-* in Yiddish (44), and *d-* in Hunzib (45) might therefore be referred to as "non-elsewhere singular stems."

[12] But see especially Harbour (2014) for a sustained defense of a non-privative view. Smith *et al.* (2019) follow Calabrese (2005), Nevins (2010), and Moskal (2018) in pursuing a markedness-based reconciliation between singular-plural asymmetries and Harbour's (2014) views. Whether the singular cannot be selectively referenced because it is the unmarked value of a bivalent feature (+ in ±ATOMIC) or because it simply does not exist, the problems with (42)–(45) arise either way.

[13] While the pattern in (44)–(45) is comparatively rare (Yiddish and Hunzib being the only examples I've found so far among accusative and ergative languages, respectively), the pattern in (42)–(43) is cross-linguistically common: instances abound in Turkic, Dravidian, and Nakh-Dagestanian (see Smith *et al.* 2019 for further examples).

(42) Malayalam (Nair 2012: 35)

	1SG	1PL.EX
NOM	**ñān**	**ñaŋ**-ŋaḷ
ACC	**enn**-e	**ñaŋ**-ŋaḷ-e
INSTR	**enn**-āl	**ñaŋ**-ŋaḷ-āl

(43) Basque proximal (Hualde 2003: 178)

	SG	PL
ABS	**hau**	**hau**-ek
ERG	**hon**-ek	**hau**-ek
DAT	**hon**-i	**hau**-ei

(44) Yiddish (Jacobs 2005: 185)

	3F.SG	3PL
NOM	**z**-i	**z**-ej
ACC	**z**-i	**z**-ej
DAT	**i**-r	**z**-ej

(45) Hunzib (van den Berg 1995: 60)

	2SG	2PL
ABS	**m**-ə	**m**-iž-e
ERG	**m**-ə	**m**-iž-e
DAT	**d**-ibi	**m**-iž-uu

Taken in isolation, the existence of non-elsewhere singular stems could just be regarded as evidence against the cumulative number decomposition in (41) and in favor of non-cumulative alternatives on the market (such as those mentioned in fn. 12). However, the fact that the exact same problem also arises with case, for which the evidence for cumulative decomposition is considerably stronger from both a morphology-internal (§2.1.2) and a morphosyntactic (§2.1.3) standpoint, suggests that the number decomposition in (41) might actually still be correct, and that non-elsewhere singular stems might just be another instance of a more pervasive problem inherent to the exponent-selection system we've discussed so far.

2.3 Previous attempts at a solution

Although the present discussion is to my knowledge the first place where all of these patterns have been systematically brought together, the issue they highlight is not new to the literature. In particular, among the works that have already identified and tackled it, there are two that have pursued a more conservative approach to it than I will pursue here — McFadden (2018) and Christopoulos & Zompì (2022). Both of these works attempt to retain the exponent-selection system based on underspecification and Pāṇinian ordering, and argue that the tension between *ABA on the one hand and non-elsewhere unmarked- and core-case stems on the other should rather be solved by amending the inventory of natural classes that exponents can be made eligible to apply in. In this subsection, I briefly review these proposals, and argue that they both fail to generalize to the full range of attested non-elsewhere unmarked- and core-case stems.

2.3.1 McFadden's (2018) markedness-based solution

The first previous attempt at a solution I'll discuss here is due to McFadden (2018), who bases his proposal on a narrower class of data than the one I've been considering here. Specifically, here I've tried to take account of all known paradigms that instantiate case-conditioned allomorphy — and, among these, I've let pronominal paradigms feature prominently, as they are the ones that exhibit such allomorphy most often (cf. Moskal 2015) and with the richest variety of patterns. By contrast, McFadden (2018) focuses on lexical nouns' paradigms, and more specifically on the allomorphic exponence of what he analyzes as instances of the categorizing head *n* — i.e., in more naive terms, stem formatives like *-ō*/*-in-* in Latin (46), intervening between the root proper and number/case morphemes.

(46) Latin 'human being' (Ernout 1953: 45–46)

	SG	PL
NOM	hom-**ō**	hom-**in**-es
ACC	hom-**in**-em	hom-**in**-es
DAT	hom-**in**-ī	hom-**in**-ibus

Within this set of data, the only instances of non-elsewhere unmarked- and core-case stems that McFadden finds are patterns like that of Latin (46), with a NOM.SG contrasting with the rest of the paradigm (cf. Icelandic (30) and Tsez (31) in §2.2.1), or like that of Latin (47), with NOM.SG and ACC.SG patterning alike in contrast to the rest of the paradigm (cf. (37) and Megrelian (38) in §2.2.2).

(47) Latin 'thigh' (Weiss 2009: 240)

	SG	PL
NOM	**femur**	**femin**-a
ACC	**femur**	**femin**-a
DAT	**femin**-ī	**femin**-ibus

Having identified the empirical problem in these terms, McFadden proceeds to tackle it as follows. As a first step, he observes that, for all the non-elsewhere core-case stems like *femur* in his sample, NOM.SG and ACC.SG (in fact, each NOM and its corresponding ACC) are systematically fully syncretic. He takes this fact to suggest that non-elsewhere core-case stems are really nothing more than non-elsewhere unmarked-case stems in disguise: "One way to think about this […] is that with these nouns the accusative form really *is* just the nominative form" (McFadden 2018: 8; emphasis his), as a result of either immediately post-syntactic deletion of κ_{ACC} (an operation commonly referred to as *Impoverishment*) or "a kind of Differential Object Marking, such that accusative assignment would simply not apply to the relevant nouns" (*ibid.*). Under either implementation of this view, the only multidimensional puzzle that would then remain for us to solve would be the one posed by the non-elsewhere unmarked-case stem in (46).

As for (46), McFadden (2018: 8) suggests that, contrary to initial appearances, the dark-grey set of contexts where the allomorph *-in-* applies might in fact be a natural class: on the assumption that nominative is the only unmarked case and that singular is the unmarked number, the dark-grey set would be the class of all contexts involving some — any — marked feature. McFadden himself remains noncommittal about the concrete implementation of this idea, but one possibility would be, for example, to posit a markedness metafeature ^{m} shared by both κ_{ACC} and $\#_{\text{PL}}$ (to be rewritten as "κ^{m}_{ACC}" and "$\#^{m}_{\text{PL}}$", respectively), and to include such a metafeature in the context specification of an exponent like dark-grey *-in-*.

(48) $n^0 \rightsquigarrow \begin{cases} \text{-}in\text{- } / \_\_\ ^{m} \\ \text{-}\bar{o} \end{cases}$

This solution, however, runs into several problems. First of all, it works well for the pattern in (46) and its apparent variant in (47), but it does not readily extend to other attested non-elsewhere unmarked- and core-case stems such as those we saw in (32) and (39), reproduced here as (49) and (50), respectively.

(49) Eastern Frisian (Matras & Reershemius 2003: 22)

	3M.SG	3PL
NOM	**h**-äi	**z**-äi
ACC	**h**-um	**h**-ör
DAT	**h**-um	**h**-ör

(50) Old High German (Braune 2004: 243)

	3N.SG	3N.PL
NOM	**i**-ʒ	**si**-u
ACC	**i**-ʒ	**si**-u
DAT	**i**-mu/**i**-mo	**i**-m/**i**-n

The problem with these patterns is that, rather than setting the nominative (and possibly accusative) singular apart from everything else, they set apart the nominative (and possibly accusative) plural — so that, if we wanted to replicate McFadden's account, we would have to say that in these languages it's actually the singular that is marked and the plural that is unmarked. This move, however, would lack any independent motivation, and would run into especially thorny problems in some varieties of Old Saxon where paradigms like (46) and ones like (50) appear to happily coexist side by side — which would force us, under McFadden's approach, to stipulate that plural is the marked number in some paradigms but not in others, even internal to one and the same language.[14]

(51) Old Saxon (Holthausen 1921: 115, 117 [Anm. 1])

	3M		DEM.N	
	SG	PL	SG	PL
NOM	**se**	**the**-a	**i**-t	**si**-u
ACC	**the**-na	**the**-a	**i**-t	**si**-u
DAT	**the**-mu	**the**-m	**i**-m	**i**-m

Furthermore, even if we did restrict our focus only to patterns like (46)–(47), McFadden's markedness-based account still wouldn't generalize correctly to more complex multidimensional paradigms — ones that cross not just case and number but also other inflectional dimensions such as gender or deixis type. In such paradigms, McFadden's account (at least in its simplest form) would predict that the markedness-conditioned allomorph should block the elsewhere allomorph, not just in marked-number and marked-case contexts, but also in marked-gender/marked-deixis-type contexts. On the usual assumption that any given opposition has only one unmarked member (if any), this would in turn entail that the markedness-conditioned allomorph should consistently block the elsewhere allomorph in all but one gender/deixis type. However, as pointed out by Christopoulos & Zompì (2022: §3.3), this prediction is falsified by paradigms like Latvian (52), where the supposed markedness-conditioned allomorph *paš-* fails to block the supposed elsewhere *pat-* in the NOM.SG of either gender, or like Tocharian A (53), where supposedly markedness-conditioned *c-* fails to block *s-* in the NOM.SG of either deixis type.[15]

[14] 3S.M.NOM *se* (preserving a Proto-Germanic allomorphy pattern) coexists with the analogical-leveling form *th(i)e* (Holthausen 1921: 117).

[15] Christopoulos & Zompì (2022: §3.3) also show that both gender and deixis type can condition stem allomorphy (including in Tocharian itself), so the problem can't be defused by conjecturing that such features might just be inert for allomorphy-conditioning purposes.

(52) Latvian emphatic pronoun ‘self’ (Prauliņš 2012: 465)

	MASC		FEM	
	SG	PL	SG	PL
NOM	**pat**-s	**paš**-i	**pat**-i	**paš**-as
ACC	**paš**-u	**paš**-us	**paš**-u	**paš**-as
DAT	**paš**-am	**paš**-iem	**paš**-ai	**paš**-ām

(53) Tocharian A masculine demonstrative (Krause & Slocum 2007–10: §31)

	DISTAL		PROXIMAL	
	SG	PL	SG	PL
NOM	**s**-am	**c**-em	**s**-äs	**c**-es
ACC=DAT	**c**-am	**c**-es-äm	**c**-as	**c**-es-äs

In view of these two problems, the notion of “markedness” that is required to go along with McFadden’s (2018) proposal ends up looking more like an *ad-hoc* diacritic than like any independently justified notion of markedness currently out there. As a result, the proposal appears not to provide a genuine solution to the problem posed by non-elsewhere unmarked-case and core-case stems.

2.3.2 Christopoulos & Zompì’s (2022) NOM- and SG-specific features

Christopoulos & Zompì (2022) — henceforth, “C&Z” — offer an alternative solution that improves on some of the issues with McFadden’s (2018) proposal but still falls short of full generality.

C&Z observe that, given a Pāṇinian allomorph-selection system like the one we discussed in §1.2, three contexts C_1, C_2, and C_3 don’t necessarily have to be involved in a fully cumulative decomposition ($C_1 \subset C_2 \subset C_3$) in order to give rise to a *ABA pattern. Rather, the only conditions they must meet to give rise to *ABA in such a system are the two in (54).

(54) *Necessary conditions for C_1, C_2, C_3 to give rise to *ABA in a Pāṇinian system*

a. $C_2 \subset C_3$ so that, if allomorph *b* applies in C_2 and allomorph *a* applies in C_3, it must be because *a* has in its context specification some of the features in $C_3 - C_2$;

b. $(C_3 - C_2) \cap C_1 = \varnothing$ so that, by having in its own context specification some feature from $C_3 - C_2$, allomorph *a* will necessarily be ineligible to apply in C_1, thereby deriving *ABA.

With this backdrop in place, C&Z argue that a partially cumulative decomposition such as (55)[16] might "allo[w] us to have the best of both worlds" (p. 17), i.e. to straightforwardly capture non-elsewhere unmarked-case stems — as well as, maybe, some apparent non-elsewhere core-case stems — while crucially preserving an account of *ABA.

(55) *Partially cumulative decomposition in nominative-accusative systems*

a. Case features of NOM: $\{\varkappa_{\text{NOM}}\}$

b. Case features of ACC: $\{\varkappa_{\text{ACC}}\}$

c. Case features of DAT: $\{\varkappa_{\text{ACC}}, \varkappa_{\text{DAT}}\}$

On the one hand, the decomposition in (55) meets both of the conditions in (54), and so still derives *ABA within a Pāṇinian exponent-selection system. On the other hand, this case decomposition also crucially makes nominatives a natural class — and therefore, if coupled with a non-cumulative number decomposition whereby singular and plural are also natural classes (e.g. (56)), it directly captures the non-elsewhere unmarked-case stems in (49), (52), and (53) without running into any of the markedness-theoretic issues that beset McFadden's proposal (see (57)–(59)).

(56) *Non-cumulative decomposition of number*

a. Number features of SG: $\{\#_{\text{SG}}\}$

b. Number features of PL: $\{\#_{\text{PL}}\}$

(57) Eastern Frisian (Matras & Reershemius 2003: 22)

$$3\text{M} \rightsquigarrow \begin{cases} z\text{-} \,/\, \_\_ \, \#_{\text{PL}}, \varkappa_{\text{NOM}} \\ h\text{-} \end{cases}$$

	3M.SG	3PL
NOM	**h**-äi	**z**-äi
ACC	**h**-um	**h**-ör
DAT	**h**-um	**h**-ör

[16] Like McFadden, C&Z focus only on accusative systems, though their proposal may be replicated for ergative systems by replacing $\varkappa_{\text{NOM}}$ with $\varkappa_{\text{ABS}}$ and $\varkappa_{\text{ACC}}$ with $\varkappa_{\text{ERG}}$.

(58) Latvian emphatic pronoun 'self' (Prauliņš 2012: 465)

$$\text{SELF} \rightsquigarrow \begin{cases} \textit{pat-} \ / \ \_\_ \ \#_{\text{SG}}, \varkappa_{\text{NOM}} \\ \textit{paš-} \end{cases}$$

	MASC		FEM	
	SG	PL	SG	PL
NOM	**pat**-s	**paš**-i	**pat**-i	**paš**-as
ACC	**paš**-u	**paš**-us	**paš**-u	**paš**-as
DAT	**paš**-am	**paš**-iem	**paš**-ai	**paš**-ām

(59) Tocharian A masculine demonstrative (Krause & Slocum 2007–10: §31)

$$\text{DEM.M} \rightsquigarrow \begin{cases} \textit{s-} \ / \ \_\_ \ \#_{\text{SG}}, \varkappa_{\text{NOM}} \\ \textit{c-} \end{cases}$$

	DISTAL		PROXIMAL	
	SG	PL	SG	PL
NOM	**s**-am	**c**-em	**s**-äs	**c**-es
ACC=DAT	**c**-am	**c**-es-äm	**c**-as	**c**-es-äs

Alongside these positive results, however, C&Z's proposal also faces several challenges from both a morphosyntactic and a morphology-internal front. For one thing, while C&Z do demonstrate that *ABA is by itself consistent with both a fully cumulative and a partially cumulative decomposition, they don't address any of the morphosyntactic evidence which, as we reviewed in §2.1.3, appears to unambiguously point to the fully cumulative alternative. As Caha (2019) pointed out in response to an earlier version of C&Z's work, the fact that the nominative displays a special propensity to lose in case competition (or, we might add, to let φ-agreement go through undisturbed) specifically follows if we assume that the nominative's case features form a proper subset of every other case's, but it remains a mystery if we assume otherwise — a mystery that C&Z leave unresolved.

What's maybe even more distressing, however, is a problem that arises from the morphology-internal front: as it stands, C&Z's proposal easily captures non-elsewhere unmarked-case stems such as those we just saw again in (57)–(59), but it still has no direct way to account for non-elsewhere core-case stems such as those repeated in (60)–(61). This is because, under C&Z's partially cumulative decomposition in (55), NOM and ACC still do not constitute a natural class together to the exclusion of DAT.

(60) Latin 'thigh' (Weiss 2009: 240)

	SG	PL
NOM	**femur**	**femin**-a
ACC	**femur**	**femin**-a
DAT	**femin**-ī	**femin**-ibus

(61) Old High German (Braune 2004: 243)

	3N.SG	3N.PL
NOM	**i**-ʒ	**si**-u
ACC	**i**-ʒ	**si**-u
DAT	**i**-mu/**i**-mo	**i**-m/**i**-n

There are two main ways one might try to reconcile C&Z's account with these patterns. The first one would be to capitalize on the fact, first observed by McFadden (2018), that such patterns tend to involve systematic full-word syncretism between the nominative and the accusative in which the non-elsewhere core-case stem applies (see e.g. *femur* being both NOM.SG and ACC.SG in (60), or *si-u* being both NOM.PL and ACC.PL in (61)). One could then propose on these grounds, again essentially following McFadden, that all these patterns involve some feature-manipulation process that, between syntax proper and exponent selection, turns the relevant accusative context into the corresponding nominative — non-elsewhere core-case stems thus being nothing but unmarked-case stems in disguise.

The alternative approach, which is the one C&Z pursue, would instead be to try and explain away all existing non-elsewhere core-case stems in terms of (possibly lexically restricted) phonological alternations, rather than morphosyntactically conditioned allomorphies. In the Latin paradigm in (60), since NOM.SG and ACC.SG are the only two suffixless contexts, one might posit a phonological process turning *femin-* into *femur* whose conditioning factor would not be structural proximity to any case or number features, but rather linear adjacency to the right edge of the phonological word (p. 20). Similarly, in the Old High German paradigm in (61), one might observe that *i-* only applies before consonant-initial suffixes while *si-* only does so before vowel-initial ones, and then attempt to devise a phonological account accordingly (C&Z, p. 23fn33).

Both of these strategies, however, lay themselves open to slippery-slope criticisms. Within the feature-manipulation strategy, for example, turning an accusative into the corresponding nominative would require not just deleting $\varkappa_{\text{ACC}}$ but replacing it with $\varkappa_{\text{NOM}}$ — but then, if this kind of feature manipulation is possible, what isn't? In particular, why couldn't one replace $\varkappa_{\text{ACC}}$ with $\varkappa_{\text{NOM}}$ exclusively in the context of $\varkappa_{\text{DAT}}$, so as to sneakily generate an ABA pattern as in (62)?

(62) **Something we do not want:**

a. **Pre-feature-manipulation decomposition:**

i. Case features of NOM: $\{\varkappa_{\text{NOM}}\}$

ii. Case features of ACC: $\{\varkappa_{\text{ACC}}\}$

iii. Case features of DAT: $\{\varkappa_{\text{ACC}}, \varkappa_{\text{DAT}}\}$

b. **Feature-manipulation rule**

$$\varkappa_{\text{ACC}} \rightarrow \varkappa_{\text{NOM}} \;/\; \_\_ \;\varkappa_{\text{DAT}}$$

c. **Post-feature-manipulation decomposition:**

i. Case features of NOM: $\{\varkappa_{\text{NOM}}\}$

ii. Case features of ACC: $\{\varkappa_{\text{ACC}}\}$

iii. Case features of DAT: $\{\varkappa_{\text{NOM}}, \varkappa_{\text{DAT}}\}$

d. **Exponence**

$$\sqrt{\ldots} \rightsquigarrow \begin{cases} a \;/\; \_\_ \;\varkappa_{\text{NOM}} \\ b \end{cases}$$

Likewise, within the morphophonological strategy, if the phonology is allowed to effect such idiosyncratic and powerful processes as turning *-in-* into *-ur* at the end of a word, what exactly will it not be allowed to do? In particular, how could we be sure that it won't use its power to once again let ABA in through the backdoor?

Finally, besides being potentially too powerful, the two strategies we've just outlined may also not be powerful enough, since C&Z themselves (p. 22fn31) mention some non-elsewhere core-case stems that prove especially refractory to them both. These are the patterns of so-called *heteroclite* adjectives in Classical Greek — adjectives whose stem is context-dependently exponed not just by different allomorphs, but specifically by allomorphs belonging to different declension classes. In the paradigm in (63), for example, the stem allomorph *méga-* applies in NOM.SG.NEUT, ACC.SG.NEUT, NOM.SG.MASC, and ACC.SG.MASC, and combines with (sometimes null) affixes of the athematic declension, whereas the stem allomorph *megál-* applies in all the rest of the paradigm and combines with affixes of the thematic declension (characterized by the presence of theme vowels like *-o-* or *-ē-*).

(63) Classical Greek ‘great’ (Schwyzer 1939: 584)

	NEUTER		MASCULINE		FEMININE	
	SG	PL	SG	PL	SG	PL
NOM	**méga**	**megál**-a	**méga**-s	**megál**-oi	**megál**-ē	**megál**-ai
ACC	**méga**	**megál**-a	**méga**-n	**megál**-ūs	**megál**-ēn	**megál**-ās
DAT	**megál**-ōi	**megál**-ois	**megál**-ōi	**megál**-ois	**megál**-ēi	**megál**-ais

The non-elsewhere core-case stem in (63) is doubly difficult for C&Z to explain away. On the one hand, it can’t possibly be reanalyzed as a non-elsewhere unmarked-case stem in disguise, as it does not involve full-word syncretism between each relevant nominative and the corresponding accusative — witness the difference between NOM.SG.MASC *méga-**s*** and ACC.SG.MASC *méga-**n***. On the other hand, it also can’t be easily reanalyzed as a phonological alternation conditioned by the phonological properties of the affix, given that the exponence of the affix itself (whether it will be picked out from the athematic- or thematic-declension series) appears to be conditioned by the exponence of the stem (whether it will be realized as *méga-* or *megál-*, respectively); this is, after all, what the phenomenon of heteroclisis is all about.[17] Neither of the available strategies for defusing non-elsewhere core-case stems thus appears to hold much promise in this case. I will therefore take this as a strong indication that, despite hopes to the contrary, such stems do exist and are in need of a real morphosyntactic account.

[17] McFadden (2018: 28) points out that the significance of this case study is diminished by the fact that we don’t quite understand how declension classes are represented in the first place — a sentiment echoed by C&Z (p. 22fn31). However, McFadden also points out that even these adjectives all comply with *ABA, which suggests it’d be a mistake to leave them out of the scope of our account.

2.A Subsets and substrings: a cautionary note

Before moving on to the next chapter, I'd like to devote a short aside to some critical discussion of an additional argument that C&Z submit in favor of their partially cumulative decomposition in (55). In their §6.2, they write:

> Bobaljik (2012) argues, based on a *ABA generalization […], that the positive degree of an adjective is featurally properly contained in its comparative, and that the comparative is in turn featurally properly contained in its superlative. Concomitantly, Bobaljik also shows that there are several languages where the positive form is regularly a substring of the corresponding comparative form […], as well as several languages where the comparative form is a substring of the corresponding superlative form […]. There are also languages where no surface containment is observed between positives and comparatives […], or between comparatives and superlatives […, but] there seem to be no languages (in Bobaljik's sample) where the superlative form is a substring of the corresponding comparative, or where the comparative is a substring of the corresponding positive.

Tentatively generalizing this correlation beyond the domain of adjectival degree morphology, C&Z propose the conjecture in (64).

(64) *Christopoulos & Zompì's (2022) conjecture*
Given two feature sets F and F' such that $F \subset F'$ […], the string of exponents realizing F' may not be a substring of the string of exponents realizing F.

C&Z then proceed to use the conjecture in (64) to compare alternative decomposition of case with respect to how well they each fare, when coupled with (64) itself, in predicting the attested substring relations. C&Z observe, for example, that a fully cumulative case decomposition, in combination with (64), produces the following predictions:

(65) *Predictions of (64) together with a <u>fully</u> cumulative case decomposition*
  a. Since NOM's case features are assumed to be a proper subset of ACC's, no ACC form should be a substring of the corresponding NOM form.
  b. Since NOM's case features are assumed to be a proper subset of DAT's, no DAT form should be a substring of the corresponding NOM form.
  c. Since ACC's case features are assumed to be a proper subset of DAT's, no DAT form should be a substring of the corresponding ACC form.

By contrast, the partially cumulative decomposition in (55), in combination with (64), will only replicate the prediction in (65c) (= (66)), but, by positing no subset relation between NOM's case features and the other cases', it won't replicate either of the predictions in (65a)–(65b).

(66) *Predictions of (64) together with a <u>partially</u> cumulative case decomposition*
Since ACC's case features are assumed to be a proper subset of DAT's, no DAT form should be a substring of the corresponding ACC form. (= (65c))

Which decomposition's predictions are better empirically supported?

C&Z claim that the distinctive predictions of the fully cumulative decomposition (the ones in (65a)–(65b)) are both faced with several counterexamples such as the ones in (67), whereas the only prediction replicated by the partially cumulative decomposition (the one in (66)) is in fact empirically borne out. The partially cumulative decomposition, C&Z suggest, appears to get it exactly right.

(67) ACC and DAT as substrings of the corresponding NOM (C&Z, p. 27)

	Icelandic 'valley'		Icelandic 'horse'		Gothic 'town'
	SG	PL	SG	PL	SG
NOM	<u>dal-ur</u>	<u>dal-i-r</u>	<u>hest-ur</u>	<u>hest-a-r</u>	<u>baúrg-s</u>
ACC	<u>dal</u>	<u>dal-i</u>	<u>hest</u>	<u>hest-a</u>	<u>baúrg</u>
DAT	<u>dal</u>	döl-um	hest-i	hest-um	<u>baúrg</u>

There are two problems with this argument, though. The first one is conceptual: it is unclear why the conjecture should be expected to be true in the first place. As C&Z (p. 26fn35) themselves note, "[i]n a theory that draws a distinction between null exponents and the absence of an exponent, it is not obvious what would be enforcing [(64)] — in particular, what would block the possibility that, e.g., F may be mapped onto an overt exponent while F' may be mapped onto a null exponent." The second problem is empirical: contrary to what C&Z suggest, it appears that there do exist cases in which the DAT form is a substring of the corresponding ACC form, as exemplified by the Mangarayi and Latvian paradigms in (68).

(68) a. Mangarayi unit-augmented participant pronouns (Merlan 1982: 102)

	1EXCL	1INCL	2
NOM	ŋi-r	ŋa-r	ṇu-r
ACC	<u>ŋi-r-a-ŋan</u>	<u>ŋa-r-a-ŋan</u>	<u>ṇu-r-a-ŋan</u>
DAT	<u>ŋi-r-a</u>	<u>ŋa-r-a</u>	<u>ṇu-r-a</u>

b. Latvian singular participant pronouns (Prauliņš 2012: 54)

	1	2
NOM	es	t-u
ACC	man-i	t-ev-i
DAT	man	t-ev

C&Z's claim that the featural relation between ACC and DAT has a different empirical signature than the featural relations between NOM and ACC or between NOM and DAT, therefore, turns out to be incorrect. In fact, each pair of cases from {NOM, ACC, DAT} allows for substring relations going both ways: ACC and DAT forms are substrings of NOM forms in Icelandic and Gothic, while NOM forms are substrings of both ACC and DAT forms in Mangarayi; ACC forms are substrings of DAT forms in Icelandic (see the paradigm of 'horse, SG' in (67)), while DAT forms are substrings of ACC forms in Mangarayi and Latvian.

If the conjecture in (64) were true, this state of affairs would entail that there could be no cumulativity at all in the decomposition of case, i.e. that no two cases in {NOM, ACC, DAT} could stand in a featural subset–superset relation. However, I think a quite different conclusion should be drawn: given that we do have convincing evidence for a fully cumulative decomposition of case (recall the arguments from §2.1.3) and that such a decomposition further promises to give us a handle on *ABA effects, I believe we should just conclude that the conjecture in (64) is incorrect, and that substring relations such as those highlighted in (67)–(68), based on isolated paradigms in their respective languages, are not necessarily informative as to featural subset–superset relations or the lack thereof.

Chapter 3

*ABA across dimensions: A new generalization

In Chapter 2, I focused on a positive result from unidimensional paradigms and on a non-result from multidimensional ones. The positive result was a *ABA generalization (due to Smith *et al.* 2019) on allomorph distributions over unidimensional paradigms of three contexts that would differ from each other solely in case — a given unmarked-case context, the corresponding dependent-case context, and the corresponding dative context. As for the non-result, we saw that, while the foregoing generalization itself imposed no restrictions on how different (*ABA-compliant) unidimensional paradigms may combine into a multidimensional one, certain attempts to derive that generalization did entail some such restrictions, and that they turned out to be systematically incorrect.

Faced with this interim summary, we may be tempted to take our end goal to be an account that still derives unidimensional *ABA but no longer makes any additional predictions about multidimensional paradigms at all. Before setting out after this relatively unambitious goal, however, we should ask ourselves if there really are no generalizations to be made specifically concerning multidimensional paradigms. After all, if we've only got a unidimensional generalization so far, it may be just because unidimensional paradigms are all we've really been looking at. Now that we've finally extended our focus to multidimensional paradigms too, is there really nothing more we can learn about them?

In this chapter, I'll embark on a search for a novel, specifically multidimensional generalization by first focusing, once again, on stem allomorphy in simple case-number paradigms. I will then test the generalization emerging from this initial case study against increasingly complex (pro)nominal and adjectival paradigms, and finally against suppletive verbal paradigms as well.

3.1 Back to the case study

3.1.1 What is attested

Let's start by taking a look at the various ways in which stem allomorphy can partition[18] our familiar paradigm {UNM.SG, DEP.SG, DAT.SG, UNM.PL, DEP.PL, DAT.PL}. The attested partitions of this paradigm add up to a total of 24. Besides the obvious monopartition in (69), these include the number-based bipartition in (70), as well as the nine bipartitions we already came across in Chapter 2, repeated in (71)–(79).[19]

(69) Old English (Hogg & Fulk 1992: 197)

	3M.SG	3M.PL
NOM	**h**-ē	**h**-ī
ACC	**h**-ine	**h**-ī
DAT	**h**-im	**h**-im

(70) Latvian (Prauliņš 2012: 54)

	2SG	2PL
NOM	**t**-u	**jū**-s
ACC	**t**-evi	**jū**-s
DAT	**t**-ev	**ju**-ms

(71) North West Lovari (Wagner 2012: 65)

	3M.SG	3PL
NOM	**vou**	**vou**-n
ACC	**le**-s	**le**-n
DAT	**le**-s-ke	**le**-n-ge

(72) Icelandic (Einarsson 1949: 68)

	2SG	2PL
NOM	**þ**-ú	**þ**-ið
ACC	**þ**-ig	**ykk**-ur
DAT	**þ**-ér	**ykk**-ur

(73) Icelandic 'the/this/that' (Einarsson 1949: 70)

	SG	PL
NOM	**s**-á	**þ**-eir
ACC	**þ**-ann	**þ**-á
DAT	**þ**-eim	**þ**-eim

(74) Eastern Frisian (Matras & Reershemius 2003: 22)

	3M.SG	3PL
NOM	**h**-äi	**z**-äi
ACC	**h**-um	**h**-ör
DAT	**h**-um	**h**-ör

[18] I'm once again restricting my focus to every *partition* of paradigm Π (every set of nonempty, disjoint subsets of Π whose big union is Π) rather than every *cover* of Π (every set of nonempty, possibly overlapping subsets of Π whose big union is Π)—i.e. I'm assuming that no context admits more than one stem allomorph. This is mostly true, but sometimes apparently not (e.g. doublets such *people* vs *persons* etc.). I'll set such issues aside.

[19] We've already seen in Chapter 2 that (71), (73), (76), (78), and (79) all have parallels in ergative languages—cf. (28), (31), (38), (43), and (45) respectively. (70) is also attested in such languages (e.g. Burushaski; cf. Noboru 2012: 80).

(75) Old High German (Braune 2004: 243)

	3F.SG	3F.PL
NOM	**sī**, **si**(-u)	**si**-o
ACC	**si**-a, **si**-e	**si**-o
DAT	**i**-ru, **i**-ro	**i**-m, **i**-n

(76) Classical Greek 'great' (Schwyzer 1939: 584)

	M.SG	M.PL
NOM	**méga**-s	**megál**-oi
ACC	**méga**-n	**megál**-ūs
DAT	**megál**-ōi	**megál**-ois

(77) Old High German (Braune 2004: 243)

	3N.SG	3N.PL
NOM	**i**-ʒ	**si**-u
ACC	**i**-ʒ	**si**-u
DAT	**i**-mu/**i**-mo	**i**-m/**i**-n

(78) Malayalam (Nair 2012: 35)

	1SG	1PL.EX
NOM	**ñān**	**ñaŋ**-ŋaḷ
ACC	**enn**-e	**ñaŋ**-ŋaḷ-e
INSTR	**enn**-āl	**ñaŋ**-ŋaḷ-āl

(79) Yiddish (Jacobs 2005: 185)

	3F.SG	3PL
NOM	**z**-i	**z**-ej
ACC	**z**-i	**z**-ej
DAT	**i**-r	**z**-ej

I'll also include here, for the sake of completeness, a potential eleventh attested bipartition, although I should emphasize that its existence is considerably more dubious. This is the pattern reproduced in (80), which I've only found attested in North Central Westphalian dialects, and which is reported to alternate, even in those dialects, with the pattern in (75) — *se* being an available option alongside *iär* for ACC.F.SG (Durrell 1990: 80).[20]

(80) North Central Westphalian (Durrell 1990: 80)

	3F.SG	3PL
NOM	**se**	**se**
ACC	**iä**-r	**se**
DAT	**iä**-r	**iä**-r

[20] I'll come back to this pattern in §4.3.2.

There are also eleven securely attested tripartitions in total — the eight in (81)–(88), all exemplified here for the first time, as well as the three we've already seen in Chapter 2, repeated in (89)–(91).

(81) Modern Eastern Armenian (Dum-Tragut 2009: 124)[21]

	2SG	2PL
NOM	**du**	**du**-k‘
ACC	**k‘**-ez	**j**-ez
DAT	**k‘**-ez(a)nic	**j**-ez(a)nic

(82) Wardaman (Merlan 1994: 112–114)[22]

	3SG	3PL
ABS	**narnaj**	**narnaj**-bulu
ERG	**narnaj**-(j)i	**narnaj**-buluyi
DAT	**gunga**	**wurrugu**

(83) Latin (Ernout 1953: 100)[23]

	1SG	1PL
NOM	**egō**	**nō**-s
ACC	**m**-ē	**nō**-s
DAT	**m**-ihi	**nō**-bis

(84) Yiddish (Jacobs 2005: 185)[24]

	3N.SG	3PL
NOM	**e**-s	**zej**
ACC	**e**-s	**zej**
DAT	**i**-m	**zej**

(85) Tsakhur class-III reflexive (Schulze 1997: 40)

	1SG	1PL
ABS	**wudʒ**	**jidʒ**-bə
ERG	**tʃɪ**-n	**tʃɪ**-n
DAT	**tʃɪ**-s	**tʃɪ**-s

(86) Modern German

	3N.SG	3PL
NOM	**e**-s	**sie**
ACC	**e**-s	**sie**
DAT	**ih**-r	**ih**-nen

(87) Modern German

	2SG	2PL
NOM	**d**-u	**ihr**
ACC	**d**-ich	**eu**-ch
DAT	**d**-ir	**eu**-ch

(88) North Central Westphalian (Durrell 1990: 80)

	3N.SG	3PL
NOM	**et**	**se**
ACC	**et**	**se**
DAT	**et**	**iä**-r

[21] There is no reason to think that the *-k‘* in *duk‘* is the same exponent as the *k‘-* in *k‘ez*: suffixal *-k‘* is the regular exponent of NOM.PL in the language. Also, the same pattern is also found in the reflexive paradigm of absolutive-ergative Tanti Dargwa (Sumbatova 2020: 157).
[22] Same pattern in the 3FEM paradigm of nominative-accusative Swabian (Russ 1990: 353).
[23] Same pattern in the second-person paradigm of absolutive-ergative Avar (Forker 2020: 251).
[24] Same pattern in the first-person paradigm of absolutive-ergative Lak (Friedman 2020: 214).

(89) Yiddish (Jacobs 2005: 185)

	1SG	1PL
NOM	**ix**	**m**-ir
ACC	**m**-ix	**undz**
DAT	**m**-ir	**undz**

(90) Albanian 'that' (Newmark, Hubbard & Prifti 1982: 122)

	F.SG	F.PL
NOM	a-**j**-o	a-**t**-o
ACC	a-**t**-ë	a-**t**-o
DAT	a-**sa**-j	a-**t**-yre

(91) Modern German

	3M.SG	3PL
NOM	**er**	**sie**
ACC	**ih**-n	**sie**
DAT	**ih**-m	**ih**-nen

Finally, there are two attested quadripartitions — the one in (92) from German, which I'm introducing here for the first time, and the one from Khinalug, which I already presented in Chapter 2, and which I repeat here in (93),

(92) Modern German

	1SG	1PL
NOM	**ich**	**wir**
ACC	**m**-ich	**uns**
DAT	**m**-ir	**uns**

(93) Khinalug (Khvtisiashvili 2013: 125)

	1SG	1PL.EXCL
ABS	**zɨ**	**y**-ir
ERG	**y**-ä	**y**-ir
DAT	**as**(-ir)	**ʃir**-ì

This exhausts the typology of attested partitions of the paradigm to the best of my knowledge.

3.1.2 Towards the generalization

In order to get to the generalization we want, it may be useful to introduce, at this point, a couple of notions that will help us better frame some of the problems we've been grappling with. These are the notion of a *Russian-doll triple*, and specifically of *unidimensional* and *multidimensional Russian-doll triples*, as defined in (94), (95a), and (95b), respectively.

(94) A triple of feature sets $\langle F_1, F_2, F_3 \rangle$ is a *Russian-doll triple* iff $F_1 \subset F_2 \subset F_3$.

(95) A Russian-doll triple $\langle F_1, F_2, F_3 \rangle$ is

a. *unidimensional* iff all the features in $F_3 - F_1$ belong to one and the same inflectional dimension (e.g. they're all case features);

b. *multidimensional* otherwise.

Assuming cumulative decompositions for both case and number, examples of unidimensional Russian-doll triples include the ones in (96), while examples of multidimensional ones include those in (97).

(96) a. ⟨NOM.SG, ACC.SG, DAT.SG⟩ $(= \langle \varnothing, \{\kappa_{\text{ACC}}\}, \{\kappa_{\text{ACC}}, \kappa_{\text{DAT}}\} \rangle)$

b. ⟨NOM.PL, ACC.PL, DAT.PL⟩ $(= \langle \{\#_{\text{PL}}\}, \{\#_{\text{PL}}, \kappa_{\text{ACC}}\}, \{\#_{\text{PL}}, \kappa_{\text{ACC}}, \kappa_{\text{DAT}}\} \rangle)$

(97) a. ⟨NOM.SG, NOM.PL, ACC.PL⟩ $(= \langle \varnothing, \{\#_{\text{PL}}\}, \{\#_{\text{PL}}, \kappa_{\text{ACC}}\} \rangle)$

b. ⟨NOM.SG, ACC.SG, ACC.PL⟩ $(= \langle \varnothing, \{\kappa_{\text{ACC}}\}, \{\#_{\text{PL}}, \kappa_{\text{ACC}}\} \rangle)$

With these notions in place, we are now in a better position to diagnose one major problem with the Pāṇinian approaches we explored in Chapter 2, namely the fact that such approaches incorrectly grant the same status to all Russian-doll triples, by predicting them all — unidimensional and multidimensional ones alike — to give rise to *ABA patterns. The problem with this, as we've seen, is that a difference does exist in the facts: while both of the unidimensional examples in (96) systematically bear out the *ABA prediction, both of the multidimensional ones in (97) disprove it, as we find NOM.SG and ACC.PL patterning alike to the exclusion of NOM.PL in some paradigms (e.g. (74), partially repeated in (98)) and to the exclusion of ACC.SG in others (e.g. (78), partially repeated in (99)).

(98) Eastern Frisian

	3M.SG	3PL
NOM	**h**-äi	**z**-äi
ACC	**h**-um	**h**-ör

(99) Malayalam

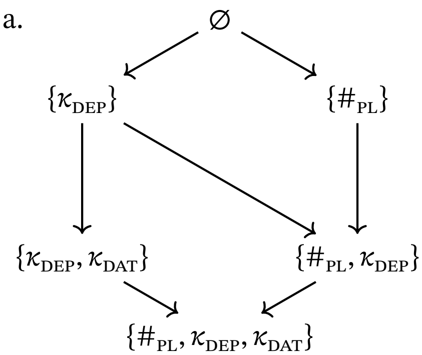
	1SG	1PL.EX
NOM	**ñān**	**ñaŋ**-ŋaḷ
ACC	**enn**-e	**ñaŋ**-ŋaḷ-e

To understand where this difference between unidimensional and multidimensional Russian-doll triples may come from, it's useful to visualize those triples more transparently. To this end, I'm going to start using the Hasse diagram of our six-context paradigm as ordered by the proper-inclusion relation ⊂.[25] (In lieu of the more explicit featural representation in (100a), I'll often use the more compact format in (100b), where *N* stands

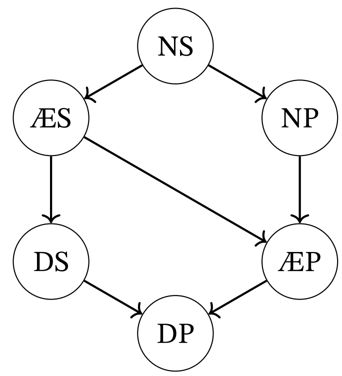
[25] In such diagrams, the *infimum* is usually placed at the bottom and the *supremum* at the top, but this would force the morphologically minded reader to some mental gymnastics, given that in traditional paradigms it's rather the nominative/absolutive (the smallest of cases) that usually sits at the top. I'll therefore deviate from common practice and flip the diagram upside down, while using directed edges pointing from each subset to its supersets as a (hopefully discreet enough) reminder of the deviation.

for ‘nominative/absolutive’, *Æ* for ‘accusative/ergative’, *D* for ‘dative’, *S* for ‘singular’, and *P* for ‘plural’.)

(100) a.



b.



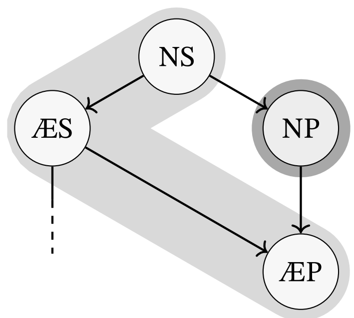
If we squint at this diagram for long enough, we start to notice that there is something special about a unidimensional Russian-doll triple like, for example, ⟨NOM.SG, ACC.SG, DAT.SG⟩: if we are to go down a directed path from its first member (NOM.SG) to its third member (DAT.SG) within the diagram, we have no other choice but to pass through its second member (ACC.SG)—or, to put it differently, there are no other Russian-doll triples in the paradigm that have both the same first and same third member as this one. The same “mandatory-route” property, of course, also holds of the other unidimensional triple ⟨NOM.PL, ACC.PL, DAT.PL⟩. By contrast, this property doesn’t hold of either of the multidimensional Russian-doll triples in (97): in order to go down a directed path from NOM.SG to ACC.PL, we may go through either NOM.PL or ACC.SG, and so neither ⟨NOM.SG, NOM.PL, ACC.PL⟩ nor ⟨NOM.SG, ACC.SG, ACC.PL⟩ can lay exclusive claim to the ⟨NOM.SG, ..., ACC.PL⟩ format.

The observation about the coexistence of multiple Russian-doll triples starting with NOM.SG and ending with ACC.PL becomes all the more relevant when we look back in this light at the paradigms in (98)–(99), repeated once more in (101)–(102). What we can notice now is that paradigms which instantiate an ABA pattern for the Russian-doll triple ⟨NOM.SG, NOM.PL, ACC.PL⟩, like (101), always simultaneously instantiate AAA for the “alternate-route” triple ⟨NOM.SG, ACC.SG, ACC.PL⟩—and, conversely, that paradigms which instantiate ABA for ⟨NOM.SG, ACC.SG, ACC.PL⟩, like (102),[26] always instantiate AAA for the “alternate-route” ⟨NOM.SG, NOM.PL, ACC.PL⟩.

[26] Or, granting its existence, like the less securely attested Westphalian pattern in (80).

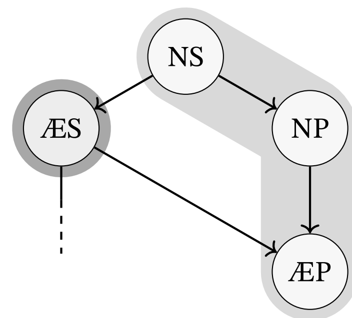
(101) Eastern Frisian

	3M.SG	3PL
NOM	**h**-äi	**z**-äi
ACC	**h**-um	**h**-ör



(102) Malayalam

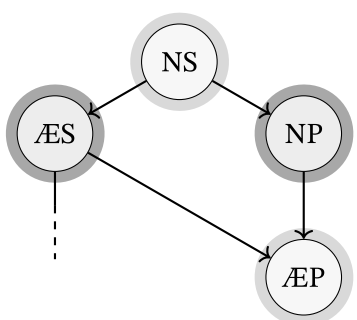
	1SG	1PL.EX
NOM	**ñān**	**ñaŋ**-ŋaḷ
ACC	**enn**-e	**ñaŋ**-ŋaḷ-e



What we never find, instead, is a paradigm where NOM.SG and ACC.PL pattern alike to the simultaneous exclusion of both NOM.PL and ACC.SG, as illustrated in the hypothetical paradigm templates in (103)–(104).

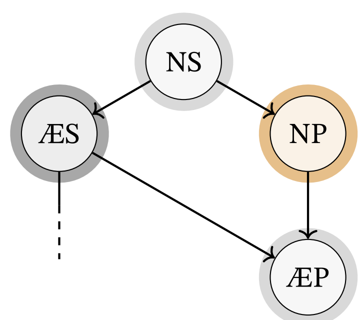
(103)

*	SG	PL
NOM		
ACC		



(104)

*	SG	PL
NOM		
ACC		



The exact same finding is replicated with an analogous pair of multidimensional Russian-doll triples — ⟨ACC.SG, ACC.PL, DAT.PL⟩ and ⟨ACC.SG, DAT.SG, DAT.PL⟩. Here, too, each triple is individually challenged by ABA patterns — the former, by the attested partition in (77), partially repeated in (105), as well as by the one in (91);[27] the latter, by the attested partition in (79), partially repeated in (106), as well as by the one in (90).

[27] Also by the less securely attested Westphalian one in (80).

However, what's consistent across all these paradigms is the fact that, whenever either of the two triples has an ABA pattern, the other one has AAA.

(105) Old High German

	3N.SG	3N.PL
NOM	**i**-ʒ	**si**-u
ACC	**i**-ʒ	**si**-u
DAT	**i**-mu/**i**-mo	**i**-m/**i**-n

(106) Yiddish

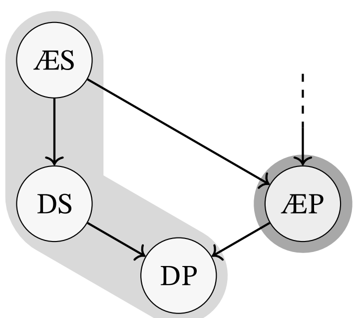
	3F.SG	3PL
NOM	**z**-i	**z**-ej
ACC	**z**-i	**z**-ej
DAT	**i**-r	**z**-ej

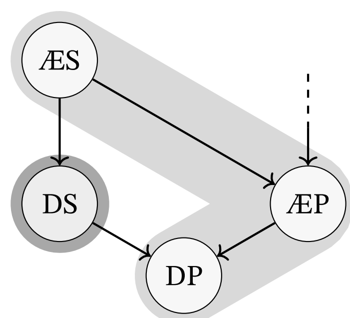
(107) Modern German

	3M.SG	3PL
NOM	**er**	**sie**
ACC	**ih**-n	**sie**
DAT	**ih**-m	**ih**-nen

(108) Albanian 'that'

	F.SG	F.PL
NOM	a-**j**-o	a-**t**-o
ACC	a-**t**-ë	a-**t**-o
DAT	a-**sa**-j	a-**t**-yre





The correct generalization, which I'll refer to as *minimal-compliance *ABA*, thus finally emerges.

(109) *Minimal-compliance *ABA*
If Russian-doll triple ⟨X, Y, Z⟩ in paradigm Π instantiates an ABA pattern, then there must be some other Russian-doll triple ⟨X, W, Z⟩ in Π that instantiates an AAA pattern.

Switching back to the graph-theoretic terminology we've occasionally been using, we may equivalently rephrase the generalization as follows. Let's suppose there is some directed path from context C_1 to context C_2 in our Hasse diagram—i.e. that $C_1 \subset C_2$—and let's suppose furthermore that the same allomorph a applies in both C_1 and C_2. Next, let's define a notion of *covering* as in (110).

(110) *Covering*
Allomorph *a* covers path *P* iff *a* applies in every context in *P*.

Now, while the broadly Pāṇinian approaches from Chapter 2 incorrectly predicted that *a* would cover every directed path from C_1 to C_2, we are weakening that universal quantification into an existential one: as long as *a* covers at least one "alternate route" from C_1 to C_2, we can let it fail to cover any number of other paths.

(111) *Minimal-compliance *ABA (equivalent formulation)*
Let H be the Hasse diagram of a paradigm as ordered by the $\subset$-relation. If there is some directed path in H from context C_1 to context C_2, and allomorph *a* applies in both C_1 and C_2, then *a* must cover some directed path in H from C_1 to C_2.

It's worth emphasizing that this proposal, with its existential quantification over directed paths, weakens the predictions of its universal-quantification predecessors only where they need to be weakened — where the domain of quantification, i.e. the set of directed paths between the two contexts in question, is large enough to let us have ABA patterns on some of those paths by virtue of having AAA on some others. By contrast, when the domain of quantification is a singleton, i.e. when between the two contexts in question there's only one directed path, the predictions of the existential-quantification proposal remain every bit as strong as those of its universal-quantification predecessors: that one path cannot have ABA. That's how the current proposal derives the classical, unidimensional *ABA generalization as a corollary: what's special about unidimensional triples like ⟨UNM.SG, DEP.SG, DAT.SG⟩ and ⟨UNM.PL, DEP.PL, DAT.PL⟩, from this perspective, is not their unidimensionality *per se*, but rather what I've previously referred to as their "mandatory-route" property — the fact that they each pick their starting point and end point, so to speak, in such a way as to reduce the number of viable paths to 1. Unidimensional *ABA thus follows as a special subcase of the minimal-compliance *ABA generalization in (109)/(111).

3.1.3 A few numbers

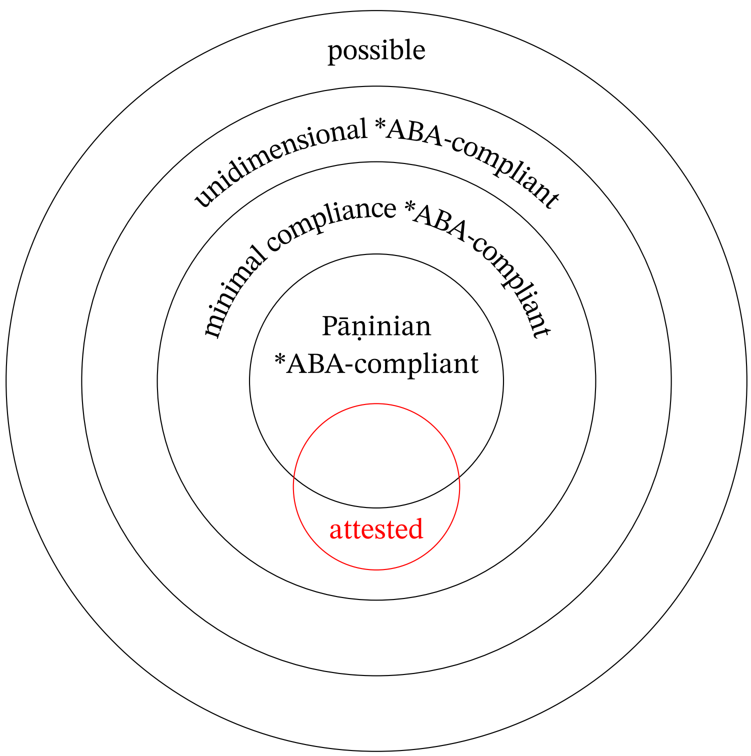
With this new multidimensional *ABA generalization in place, we can now also see in detail how its restrictiveness compares both to that of the unidimensional *ABA generalization in (112) and to that of the Pāṇinian, "maximal-compliance" multidimensional *ABA in (113).

(112) *Unidimensional *ABA*
No unidimensional Russian-doll triple can instantiate an ABA pattern.

(113) *“Pāṇinian” multidimensional *ABA*
No Russian-doll triple (be it uni- or multidimensional) can instantiate an ABA pattern.

To make the comparison concrete, we may ask, for each of the *ABA constraints under discussion, how many partitions of our six-context paradigm {UNM.SG, DEP.SG, DAT.SG, UNM.PL, DEP.PL, DAT.PL} are permitted by it. There are four sets of such partitions to be considered here — each set a superset of the next. The largest one is the set of all 203 possible partitions of our paradigm (203 being the *Bell number* B_6); next is the set of all those partitions that comply with the unidimensional *ABA constraint in (112), for a subtotal of 135; then we have the set of partitions that obey the minimal-compliance *ABA generalization in (109)/(111), for a subtotal of 93; and, finally, there is the set of partitions that comply with the “Pāṇinian” multidimensional *ABA constraint in (113), for a subtotal of 70. I represent all of this in the Euler-Venn diagram in (114).

(114)



Given our initial observation that the “Pāṇinian” *ABA constraint in (113) was demonstrably too restrictive, and given our concern that unidimenional (112) by itself might not be restrictive enough, the fact that our new minimal-compliance *ABA generalization falls roughly halfway between the two in this regard appears to be a welcome result. Be that as it may, if one were to compare the number of partitions we currently predict to be possible and the number of partitions that we find attested — 93 *vs* 25, respectively — one could still legitimately worry that even our new generalization might not be restrictive enough quite yet.

However, this worry can be substantially mitigated, I think, by taking into account an important independent factor that contributes to winnow the range of attested partitions. In particular, it should be kept in mind that only two of the attested partitions (the German one in (92) and the Khinalug one in (93)) have as many as four cells, and that no attested partition has more than four — a skew that in all likelihood doesn't have anything to do with constraints on allomorph distribution *per se*, but rather with the inherent difficulty of learning a mapping between a single abstract morpheme on the one hand and so many different allomorphs on the other.

If we try to control for this skew by restricting our focus to bipartitions, the results immediately look a lot more encouraging: out of the 31 possible bipartitions of our six-context paradigm,[28] the unidimensional constraint in (112) permits 17, our minimal-compliance *ABA in (109) permits 14, the "Pāṇinian" constraint in (113) permits only 8 — and 11 appear to be attested.

(115) Comparing the restrictiveness of various conceivable *ABA constraints

	How many partitions of {UNM.SG, DEP.SG, DAT.SG, UNM.PL, DEP.PL, DAT.PL} into *n* cells...				
	are possible	obey unidimensional *ABA	obey minimal compliance *ABA	obey Pāṇinian multidimensional *ABA	are attested
$n \in N$	203	135	93	70	25
$n = 2$	31	17	14	8	11

We can thus not only see the undergeneration problem with the "Pāṇinian" constraint (113) at its clearest; we can also see almost all of the space of possibilities predicted by our own (109) being filled up by attested patterns. I take this to be a strong indication that (109) is on the right track.

Before setting out to devise an account that derives this new generalization on principled grounds, I will first devote the rest of this chapter to further strengthening the case for it on the basis of two new empirical case studies. In §3.2, I'll test the new generalization against more complex pronominal paradigms involving grammatical gender alongside case and number, with a focus on the well-known morphomic patterns in

[28] 31 is the *Stirling number of the second kind* $\left\{ {6 \atop 2} \right\}$.

the pronominal inflection of Germanic languages. In § 3.3, I'll bring the generalization to bear on an altogether different domain, namely φ-agreement–conditioned stem suppletion in verbal paradigms.

3.2 Putting the generalization to the test (I): Morphomic interactions with gender

3.2.1 Why this case study

As a first step, I'm going to test the new minimal-compliance *ABA generalization against a minimal extension of the paradigm featured in our case study so far. More specifically, I'm going to enrich that paradigm with a third inflectional dimension — that of grammatical gender — in addition to those of case and number, and check to what extent the resulting complex allomorphy patterns align with our novel predictions.

Although small and seemingly harmless, this extension is in fact the potential source of a number of difficulties. In a nutshell, the problem is that, in order for our generalization to produce any new predictions to test in this domain (beyond those we already spelled out and tested for bidimensional case-number paradigms), we need to make some substantive assumptions about the decomposition of gender, but this turns out to be no easy task at all. As it happens, gender is a contender for being, in Corbett's (1991: 1) words, "the most puzzling of the grammatical categories," as natural languages appear to vary wildly with respect to the number of genders they involve, the criteria (semantic, formal, completely arbitrary, or some combination of the above) whereby they assign genders to nominals, and the markedness asymmetries they exhibit — so much so that even the notion itself of a universal gender decomposition might be feared to just be misconceived.

Rather than taking a stand in this debate (see especially Kramer 2015 for a defense of the viability and fruitfulness of a uniformitarian position), here I'd like to sidestep it entirely by restricting my focus to a relatively homogeneous subset of a particular language family — i.e. to West Germanic languages such as Yiddish, standard German, and (long-extinct but well documented) Old High German. While this choice might seem arbitrary and limiting at first, I believe it has several points going for it.

First of all, as we will see in more detail momentarily, we happen to have an unusually wide array of relevant data from at least one of these languages (standard German) to justify a particular gender decomposition, and a strictly cumulative one at that — that is, a decomposition that enables our minimal-compliance *ABA generalization to produce an especially large set of predictions. In addition, while we don't have quite as rich an array of evidence for the same decomposition when it comes to all the rest of the West

Germanic subfamily, I believe that the languages I'll focus on show enough similarity in the basics of their gender systems to still let us at least tentatively make uniform assumptions without having to worry too much about the anti-uniformitarian objections mentioned above. And finally (but perhaps most importantly), while the subfamily exhibits enough uniformity in its gender systems to permit meaningful comparisons, it also displays enough internal variation in allomorph-distribution patterns to still present us with a nontrivial empirical challenge to meet. In particular, West Germanic pronominal paradigms turn out to be specifically rich in morphomic patterns— ones that prove difficult to capture on a broadly Pāṇinian approach to exponent selection, even if non-cumulative decompositions for case and gender were to be adopted. Should even the most puzzling of these morphomes prove consistent with our generalization, this would therefore be encouraging evidence in its favor.

3.2.2 Assumptions about gender decomposition

Old High German, modern standard German, and Yiddish all have a system of three grammatical genders: neuter, masculine, and feminine.

The gender-assignment system can be described, as a first approximation, in terms of a fundamental distinction between human-denoting nouns, whose grammatical gender overwhelmingly tends to correlate with their semantic gender, and non-human-denoting nouns, whose grammatical gender is mostly arbitrarily specified in the lexicon (although several more or less robust regularities on a semantic, morphological, and phonological basis do exist; see e.g. Köpcke & Zubin 1984 on German). This split has led Kramer (2015: 130) to posit a distinction between *interpretable* masculine and feminine features for nouns of the first group and *uninterpretable* masculine and feminine features for nouns of the second group— a distinction that I will set aside in what follows, given that, while real and important elsewhere, it ends up being systematically neutralized in observable exponence.

Moving on to the concrete representation of the three genders, Kramer (2015: *ibid.*) treats the German neuter as the absence of gender features. I will follow her in this analysis, which is consistent, among other things, with the observation that expletives are systematically neuter in all three of the languages in question. Concerning the status of masculine and feminine, however, Kramer (2015) and I will slightly differ: while she makes use of the binary feature ±FEM (coming in an interpretable and uninterpretable version, as per above), I will rather attempt, in keeping with the spirit of the rest of this dissertation, to recast her treatment in terms of privative features, with a specific view to reflecting markedness asymmetries as featural subset-superset relations.

In particular, German provides several converging strands of evidence to the effect that its feminine gender is more marked than its masculine. These include both morphological asymmetries (the feminine being generally derived from the masculine by the addition of a suffix — never the other way around) and other asymmetries including, for example, for the purposes of ellipsis parallelism (Bobaljik & Zocca 2011).

(116) German (Bobaljik & Zocca 2011: 144–145)

a. *Masculine antecedents license feminine ellipses*

Mein Onkel ist (ein) Österreicher, und meine Tante auch.
my uncle is (an) Austrian.MASC and my aunt too
'My uncle is (an) Austrian and my aunt is too.'

b. *Feminine antecedents don't license masculine ellipses*

Meine Tante ist (eine) Österreicher-in, und mein Onkel auch.
My aunt is (an) Austrian-FEM and my uncle too
Intended: 'My aunt is (an) Austrian and my uncle is too.'

In order to represent this markedness asymmetry in structural terms, we are therefore led to posit the decomposition in (117) — which I will tentatively extend, as foreshadowed in §3.2.1, from modern German alone to the other two cognate languages as well.

(117) a. Gender features of NEU: ∅

b. Gender features of MASC: $\{\gamma_{\text{ANIM}}\}$

c. Gender features of FEM: $\{\gamma_{\text{ANIM}}, \gamma_{\text{FEM}}\}$

As a result of its strict cumulativity, (117) will produce an especially rich set of testable predictions in combination with the decompositions we've already adopted for case and number, as the relevant Hasse diagram will soon help us appreciate in the next subsection.[29] Even before drawing that diagram out, however, we can already immediately see one prediction that (117) makes with regard to unidimensional Russian-doll triples: a neuter context and the corresponding feminine should never share a stem allomorph to the exclusion of the corresponding masculine. This *ABA generalization is, to the best of my knowledge, borne out without exception across West Germanic.[30]

3.2.3 Three morphomic patterns

We can start, without further ado, with the first of our full pronominal paradigms — the modern German one in (118).

[29] In particular, notice that, should it turn out that masculine and feminine are in fact in no inclusion relation to each other, the theory would be *a fortiori* consistent with the available data. In general, any shift from a more strictly cumulative to a "flatter" decomposition will always weaken, and never strengthen, the predictions of the account.

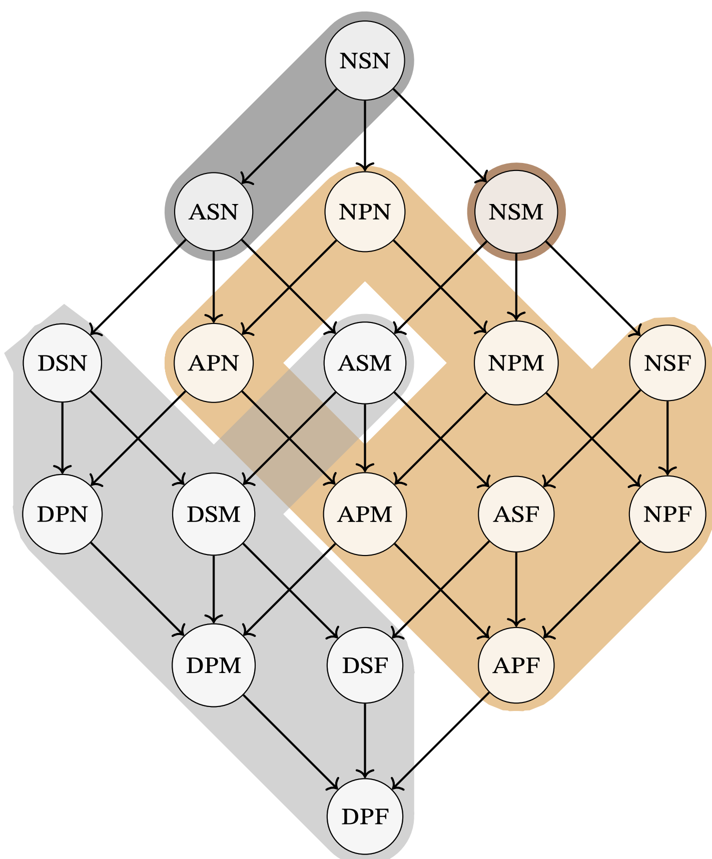
[30] If we were to adopt an alternative cumulative decomposition NEU ⊂ FEM ⊂ MASC, by contrast, we would be faced with unidimensional ABA counterexamples — see e.g. ACC.SG.NEU ***i**-ʒ* ∼ ACC.SG.FEM ***s**-ie* ∼ ACC.SG.MASC ***i**-n(an)* in Old High German (124) below.

(118) Modern German third-person pronoun

	NEUTER		MASCULINE		FEMININE	
	SG	PL	SG	PL	SG	PL
NOM	**[ɛ]**-s	**s**-ie	**[eː]**-r	**s**-ie	**s**-ie	**s**-ie
ACC	**[ɛ]**-s	**s**-ie	**ih**-n	**s**-ie	**s**-ie	**s**-ie
DAT	**ih**-m	**ih**-nen	**ih**-m	**ih**-nen	**ih**-r	**ih**-nen

A quick look at the paradigm suffices to reveal its morphomic nature: its allomorph distribution is impossible to capture via a simple allomorph-selection system in terms of natural classes and Pāṇinian ordering — regardless of whether or not we adopt strictly cumulative decompositions for case, gender, or number. In particular, even if neuter, masculine, singular, nominative, and non-dative (i.e. nominative/accusative) all constituted natural classes, we would still be left with at least one unnatural set of contexts: *s*- applies in the non-dative of feminines and plurals, whereas *ih*- applies in all datives and in ACC.SG.MASC. Even such a puzzling distribution, however, still obeys our minimal-compliance *ABA generalization, as can be seen in (119).

(119)



In particular, the diagram in (119) highlights the presence of several Russian-doll triples instantiating ABA, with each of them being "salvaged," however, by the existence of an AAA "alternate route" with the same starting point and endpoint, as detailed in (120). The generalization is unproblematically obeyed.

(120) ABA in Russian-doll triples in (118):

ASM APM DPM (but ASM DSM DPM !)
ASM ASF DSF (but ASM DSM DSF !)
ASM APM DPF
ASM ASF DPF } (but ASM DSM DPM DPF and ASM DSM DSF DPF !)
ASM APF DPF

An interesting variation on the same morphomic theme can be found in our second West Germanic paradigm — the Yiddish one in (121).

(121) Standard Yiddish third-person pronoun (Jacobs 2005)

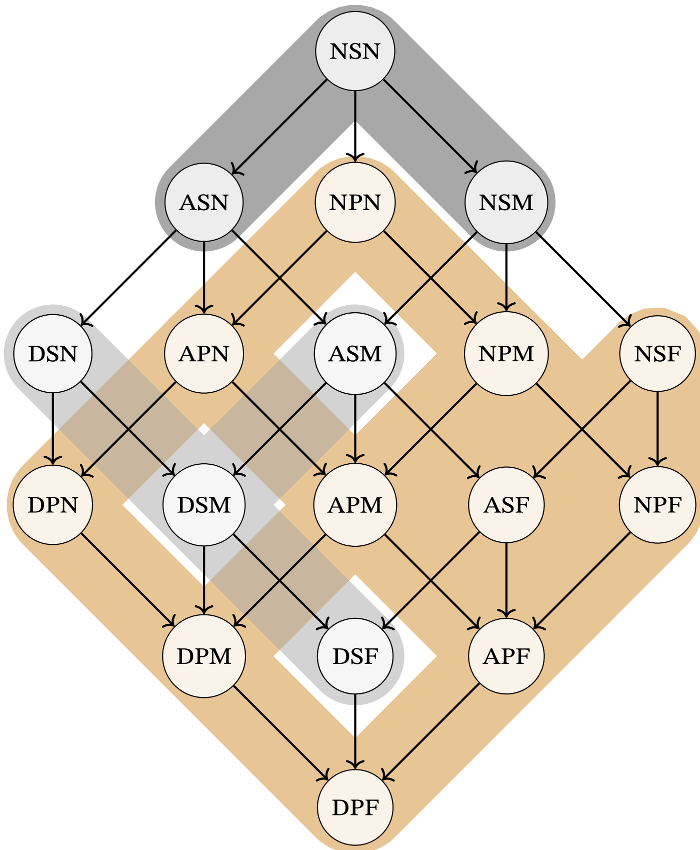
	NEUTER		MASCULINE		FEMININE	
	SG	PL	SG	PL	SG	PL
NOM	**[ɛ]**-s	**z**-ej	**[ɛ]**-r	**z**-ej	**z**-i	**z**-ej
ACC	**[ɛ]**-s	**z**-ej	**i**-n	**z**-ej	**z**-i	**z**-ej
DAT	**i**-m	**z**-ej	**i**-m	**z**-ej	**i**-r	**z**-ej

Two differences with respect to German (118) are worth noting. First, while standard German contrasts a short/lax *ɛ* for NOM/ACC.SG.NEUT and a long/tense *eː* for NOM.SG.MASC, Yiddish (at least in its non-western varieties — cf. Jacobs 2005: 30[31]) lacks the relevant phonological contrast entirely, and so appears to use the same stem allomorph in both sets of contexts.[32] Second, in Yiddish the sibilant allomorph spreads over all of the plural part of the paradigm, applying in DAT.PL as well as in NOM/ACC.PL — although it still remains confined to just NOM/ACC in the feminine singular. The resulting pattern is even more unnaturally morphomic: *ɛ*- applies in NOM.SG.MASC and NOM/ACC.SG.NEUT, *z*- applies throughout the plural and in NOM/ACC.SG.FEM, and *i*- applies in all DAT.SG's as well as in ACC.SG.MASC. As can be checked in (122)–(123), however, this morphome too turns out to be straightforwardly compatible with our minimal-compliance *ABA generalization.

[31] Cf. also [ɛs] and [ɛɾ] in Kleine's (2003: 264) text in IPA transcription.

[32] In §4.3.1, however, we will get to question whether there aren't reasons to nonetheless distinguish the NOM/ACC.SG.NEUT and the NOM.SG.MASC stem allomorphs even in Yiddish.

(122)



(123) ABA in Russian-doll triples in (121):

ASM ASF DSF (but ASM DSM DPF !)
ASF DSF DPF (but ASF APF DPF !)
NSF DSF DPF (but NSF ASF APF DPF and NSF NPF APF DPF !)

Finally, let us conclude with the Old High German paradigm in (124).

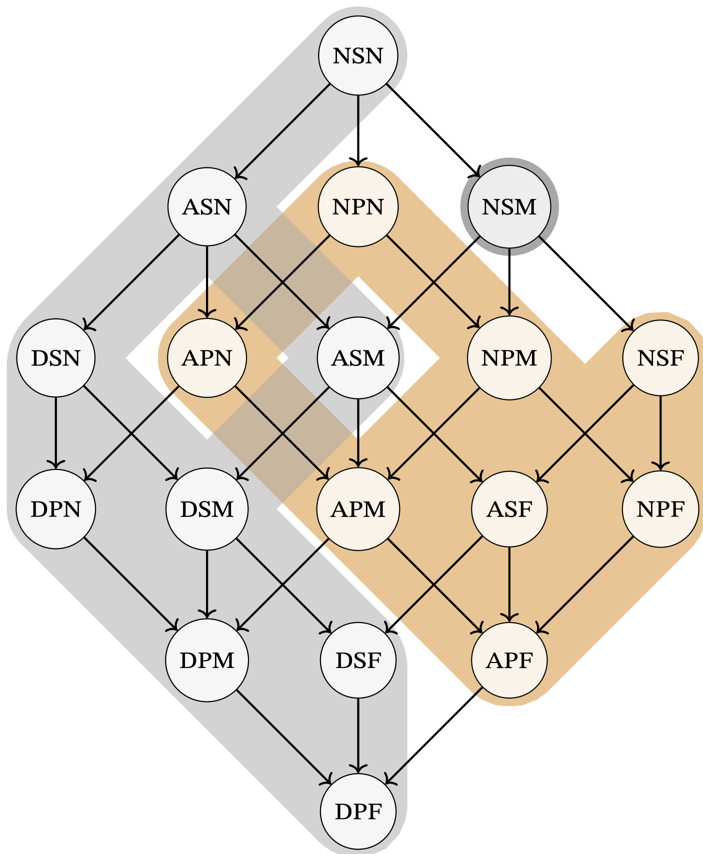
(124) Old High German third-person pronoun (Braune 2004: 243)

	NEUTER		MASCULINE		FEMININE	
	SG	PL	SG	PL	SG	PL
NOM	**i**-ʒ	**s**-iu	**ër**	**s**-ie	**s**-ī, **s**(-iu)	**s**-io
ACC	**i**-ʒ	**s**-iu	**i**-n(an)	**s**-ie	**s**-ia, **s**-ie	**s**-io
DAT	**i**-mu, **i**-mo	**i**-m, **i**-n	**i**-mu, **i**-mo	**i**-m, **i**-n	**i**-ru, **i**-ro	**i**-m, **i**-n

The allomorph-distribution in this paradigm is virtually identical to the one in modern German (118), except for NOM/ACC.SG.NEUT being subsumed under the *i*-allomorph's territory. Once again, this means that, even if we were to grant that NOM.SG.MASC was a natural class for which *ër* could be positively specified, we would still be left with at least one unnatural class: *s-* appears to apply (as in modern German) in the NOM/ACC of the feminine and of the plural, while *i-* applies in all datives, in the neuter singular, and in ACC.SG.MASC.

As a result, the Russian-doll triples instantiating ABA patterns in the diagram in (125) are almost too many to count (and certainly too many to list all out here). Each of them, however, is "salvaged" by at least one "alternate route" instantiating AAA.

(125)



The unproblematic compliance of these notorious morphomes with our novel generalization seems to me to provide encouraging evidence in its support. With this result in hand, I'll move on in the next section to a more decidedly different empirical domain.

3.3 Putting the generalization to the test (II): Agreement-conditioned verbal suppletion

When it comes to extending the minimal-compliance *ABA generalization to a new empirical domain, an obvious direction to look is toward the one φ-feature (or rather, in the terminology employed here so far, the one φ-inflectional-dimension) that's been conspicuously missing up until now — that is, person. In this section, I'm going to do this specifically through an extensive study of verbal stem allomorphy as conditioned by person and number agreement — a phenomenon which, while not particularly typologically common (cf. Bybee 1985: 93), displays a rich enough variety of attested patterns, I believe, to still lend credence to the significance of certain robust cross-linguistic gaps.

In the following subsections, I will begin by introducing my proposals about the decomposition of first, second, and third person;[33] then, I'll proceed to illustrate all the attested partitions of {1SG, 2SG, 3SG, 1PL, 2PL, 3P} to the best of my knowledge; and I will finally sketch out an extension of the predictions of the generalization to further interactions involving both φ-agreement and mood.

3.3.1 The feature structure

The usual markedness diagnostics we've appealed to so far (features of expletives, agreement failures/defaults, transparency to omnivorous agreement, etc.[34]) converge in pointing to third person as the least marked person — and, just as I did with singular and neuter, I'll take this as sufficient reason for identifying it with the absence of person features (cf. Harley & Ritter 2002; Béjar & Rezac 2009, etc., and *contra* Nevins 2007; Harbour 2016; Ackema & Neeleman 2018, as well as all references in fn. 34). When it comes to first and second person, however, it turns out to be impossible to enforce a strictly cumulative decomposition either way, as both $1 \subset 2 \subset 3$ and $2 \subset 1 \subset 3$ would give rise to unidimensional Russian-doll triples that would immediately be challenged by ABA counterexamples.[35]

[33] Since virtually all of my data points will come from the Romance family, with the remaining few from Zapotecan and Slavic, I will have no data relevant to clusivity at all.

[34] Several apparent instances of omnivorous third-person agreement have been reported from Algonquian (see especially Grishin 2023). I will set them aside here.

[35] Operstein (2015: 55; 2017: 765) explicitly claims that the Zaniza Zapotec verb 'come' instantiates the pattern in (127), but she doesn't provide full paradigms.

(126) Catanzaro 'give' preterit (one variant) (Maiden 2018: 290)

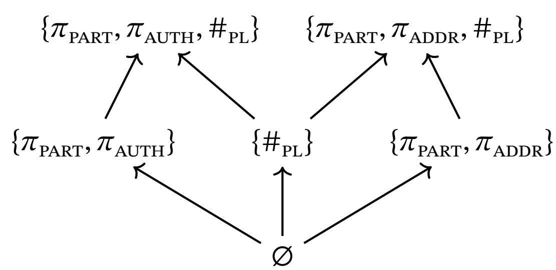
	SG	PL
1	dɛtsi	dɛ́tsimu
2	dunasti	dunástivu
3	dɛtsa	dɛ́tsaru

(127) Zaniza Zapotec (Operstein 2015)

	SG	PL
1		
2		
3		

Things do work out nicely, however, if we assume the feature structure in (128) (cf. again Harley & Ritter 2002; Béjar & Rezac 2009).[36,37]

(128) The feature structure of φ-agreement (to be slightly revised in (208))



Notice that in (128) I'm also including a participant feature π_{PART}, shared by first and second person to the exclusion of third person, even though its presence does not impinge in any way on the relevant inclusion relations. The reason why I'm representing π_{PART} nonetheless is twofold. Firstly, I expect many readers to be familiar with the overwhelming evidence (from omnivorous agreement, Person-Case Constraint effects, etc.) that the literature has adduced for such a feature — evidence that I don't want to appear to be either unaware or skeptical of. Secondly, I'm also glancing ahead, with malice aforethought, to the fact that, even though minimal-compliance *ABA itself only cares about which feature sets are subsets of which others, and not about which overlap with which, the account that I'll develop in Chapter 4 to derive minimal-compliance *ABA will actually care about both things, so that reference to a feature like π_{PART} will in fact prove crucial for allowing that account to derive several of the attested patterns.[38]

[36] Given that traditional paradigms usually have third person in their bottommost row, I can now finally revert to the common practice of putting the *supremum* at the top and the *infimum* at the bottom of Hasse diagrams.

[37] The partial order 1|2 > 3 is also one of those adopted by Moradi (2021), though she adopts other orders alongside of it too. She does so in the interest of preserving Graf's (2019) monotonicity condition, which predicts the "Pāṇinian," maximal-compliance multidimensional *ABA constraint in (113).

[38] The alert reader might then have a follow-up question — that is, why am I not representing $\{\pi_{\text{PART}}\}$ or $\{\pi_{\text{PART}}, \#_{\text{PL}}\}$ as their own nodes in the Hasse diagram? The reason is that that diagram is meant to represent

Finally, one more thing bears emphasis before we delve into the data — the fact that, in this paradigm, we can most clearly see the power of our minimal-compliance *ABA generalization in action. Although the decomposition in (128) doesn't give rise to any unidimensional Russian-doll triple, it does give rise to four multidimensional ones, which the generalization makes nontrivial testable predictions about. These are:

(129) a. 3SG cannot share an exponent with 1PL to the simultaneous exclusion of both 3PL and 1SG.

b. 3SG cannot share an exponent with 2PL to the simultaneous exclusion of both 3PL and 2SG.

At the same time, however, the generalization does allow for minimally compliant ABA patterns like 3SG, 3PL, and 1PL all patterning alike to the exclusion of 1SG, or like 3SG, 1SG, and 1PL all patterning alike to the exclusion of 3PL. As we will see momentarily, the taxonomy of attested partitions of the paradigm {1SG, 2SG, 3SG, 1PL, 2PL, 3P} does contain several such patterns, but does not pose any serious threats to the either of the generalizations in (129).

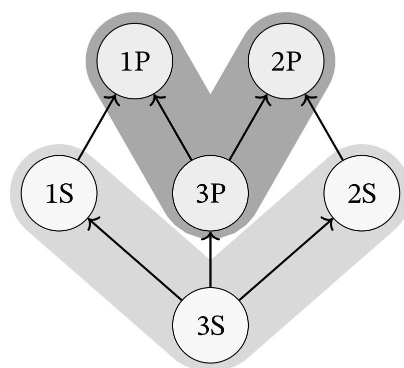
3.3.2 Cross-linguistic (but mostly Romance-internal) variation

In this subsection, I lay out what I believe is the most complete survey to date of φ-agreement–conditioned verb-stem suppletion, drawing mostly on work on Romance verbal morphology (especially Maiden 2004, 2018, Maiden *et al.* 2010, Kühlert 2023, and the references they cite).

Let's start with the easiest bipartitions to capture: PL *vs* elsewhere, 1 *vs* elsewhere, 2 *vs* elsewhere, PART *vs* elsewhere, 1PL *vs* elsewhere, 2PL *vs* elsewhere, PART.PL *vs* elsewhere.

(130) Mussomeli 'give' (Maiden *et al.* 2010)

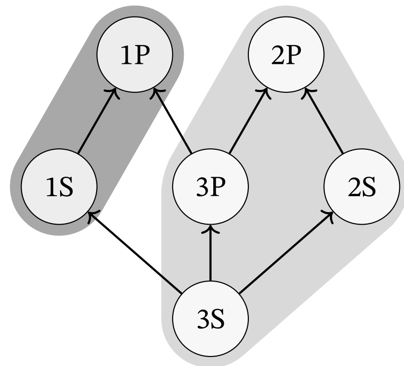
	SG	PL
1	dʊ́ɲɲ-ʊ	d-ámʊ
2	dʊ́n-ɪ	d-átɪ
3	dʊ́n-a	d-ánnʊ



inclusion relations between the various morphosyntactic contexts that a given abstract morpheme might find itself in at the end of the syntactic computation, and I'm assuming that the syntax may never generate a π_{PART} feature without concomitantly generating either a π_{ADDR} or a π_{AUTH}.

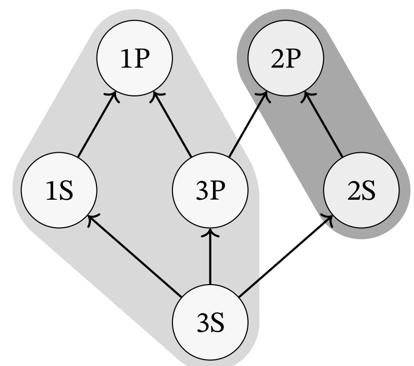
(131) Zaniza Zapotec (Operstein 2015: 55) (cf. fn. 35)

	SG	PL
1		
2		
3		



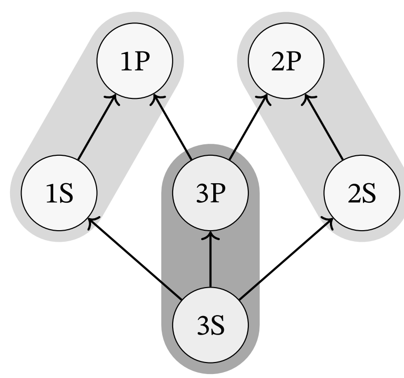
(132) Catanzaro ‘give’ preterit (one variant) (Maiden 2018: 290)

	SG	PL
1	dɛtsi	dɛ́tsimu
2	dunasti	dunástivu
3	dɛtsa	dɛ́tsaru



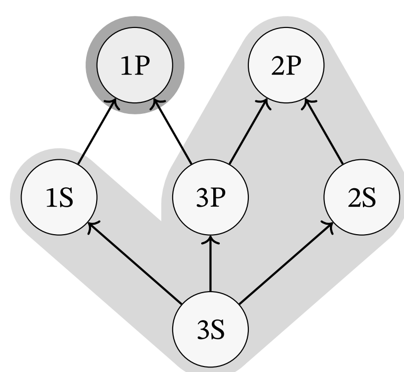
(133) ‘Be’ in some Tuscan varieties

	SG	PL
1	sono	semo
2	sei	siete
3	è	enno



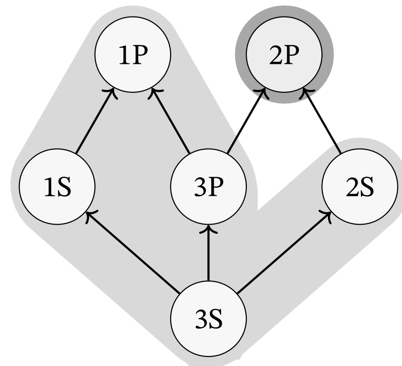
(134) Cairo Montenotte ‘go’ (Maiden *et al.* 2010)

	SG	PL
1	vag	induma
2	vai	vei
3	va	van



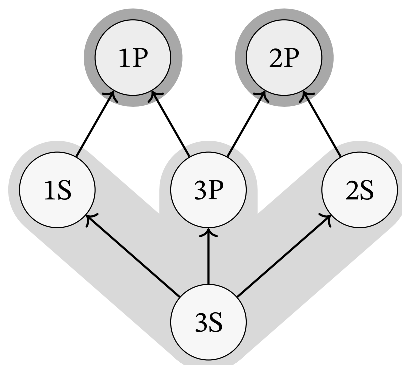
(135) Serrastretta ‘give’ (Maiden 2004: 230)[39]

	SG	PL
1	ðuɲu	ðunamu
2	ðuni	ðati
3	ðuna	ðúnanu



(136) Italian ‘go’, a.m.o.

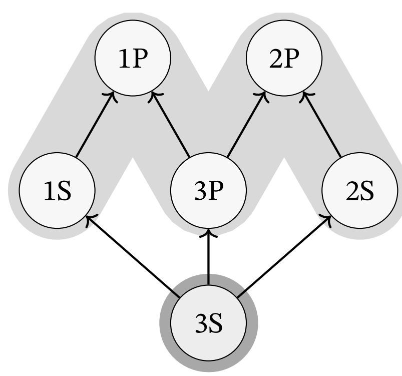
	SG	PL
1	vado	andiamo
2	vai	andate
3	va	vanno



Next are some patterns that, although harder to capture, still have no Russian-doll triple instantiating ABA.[40]

(137) Matera ‘be’ (Festa 1917: 160)

	SG	PL
1	so	symə
2	sy	sytə
3	jøtə	sondə

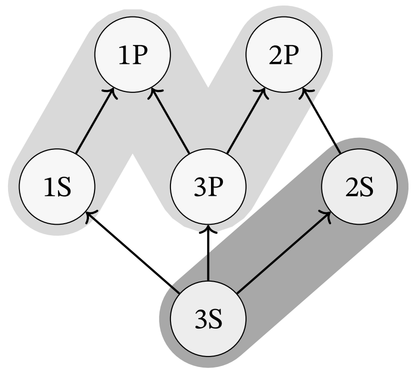


[39] Also ‘go’ in Pont-de-Vaux (Maiden *et al.* 2010), European Portuguese (*ibid.*) and Monza (Benincà, Parry & Pescarini 2016: 194).

[40] Many other Romance copular paradigms might fit the profiles of (137) and (138), but these are, I think, the cleanest examples.

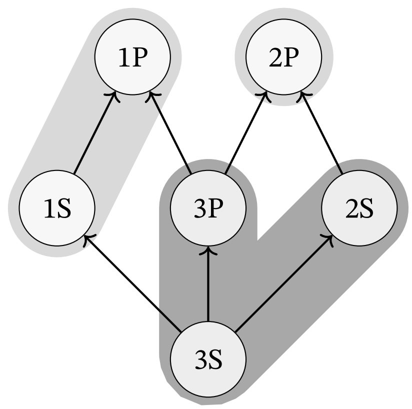
(138) Romanian ‘be’

	SG	PL
1	sunt	suntem
2	ești	sunteți
3	este	sunt



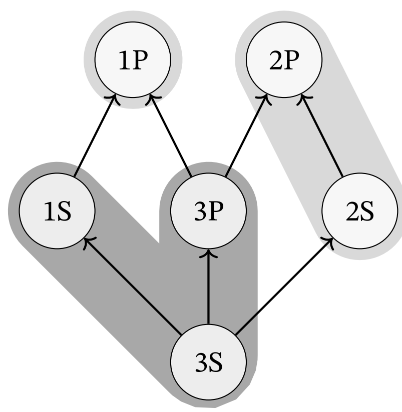
(139) Platta ‘go’ (Maiden 2018: 106)[41]

	SG	PL
1	mɔn	məin
2	vas	məis
3	va	van



(140) Catanzaro ‘give’ preterit (other variant) (Maiden 2018: 290)[42]

	SG	PL
1	dɛtsi	dunammi
2	dunasti	dunástivu
3	dɛtsa	dɛ́tsaru

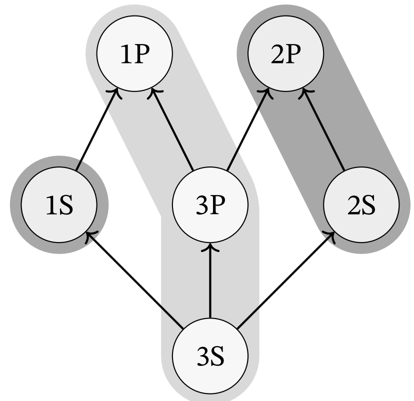


[41] Also, Nones ‘be’ (Enrico Flor, p.c.) and one variant of Acri ‘give’ (Maiden 2004: 230).

[42] Also all irregular preterites in Standard Italian, if they count as allomorphy rather than readjustment (e.g. 1sg *ébb-i*/*ébb-i*, 3sg *ébb-e*/*ébb-e*, 3pl *ébb-ero*/*ébb-ero* vs 2sg *av-ésti*, 1pl *av-émmo*, 2pl *av-éste*).

(141) Bergamasque ‘go’ (Zanetti 2005: 172)[43]

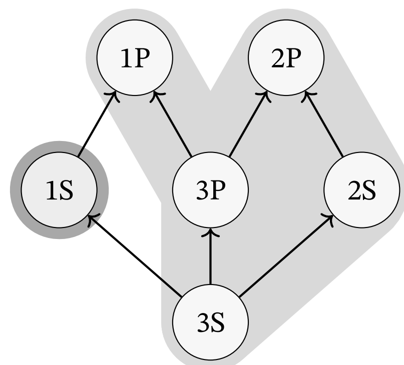
	SG	PL
1	ndo	va
2	nde	andí
3	va	va



And finally a few with ABA in Russian-doll triples — but still in keeping with (129)!

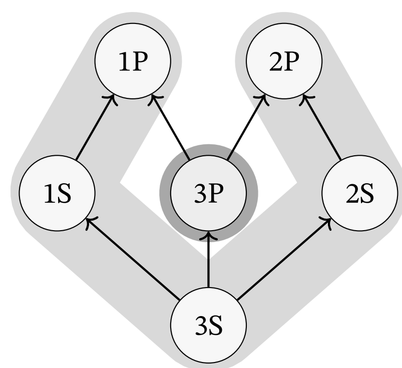
(142) Baucina ‘give’ (Maiden 2004: 230)[44]
3SG 1SG 1PL (but 3SG 3PL 1PL!)

	SG	PL
1	duɲu	damu
2	dai̯	dati
3	da	dannu



(143) Polish ‘be’
3SG 3PL 1PL (but 3SG 1SG 1PL!)
3SG 3PL 2PL (but 3SG 2SG 2PL!)

	SG	PL
1	jestem	jesteśmy
2	jesteś	jesteście
3	jest	są



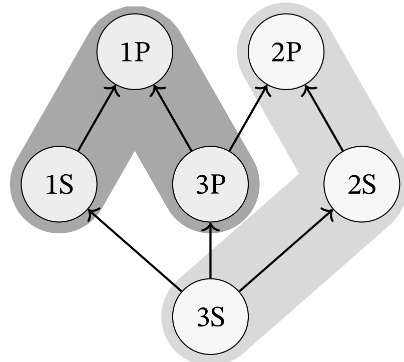
[43] The alternation between *and-* and *nd-* is purely phonological — conditioned by whether the subject clitic (not reported in (141)) ends with a vowel or a consonant (*nd* vs *and*, respectively).

[44] Also the so-called L-pattern (Maiden 2018: ch. 5, among others) — e.g. Spanish 1SG *kep-o* vs 2SG/3SG/etc. *kaβ-es/e/emos/éis/en* — if that qualifies as *bona-fide* suppletion.

(144) Latin 'can'

3SG 3PL 2PL (but 3SG 2SG 2PL!)

	SG	PL
1	possum	possumus
2	potes	potestis
3	potest	possunt



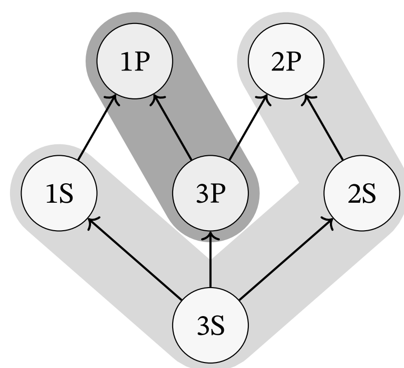
This concludes, to the best of my knowledge, the attested typology of ways in which φ-agreement-conditioned suppletion partitions our paradigm into two cells. Out of all 31 possible bipartitions, 24 are permitted by our minimal-compliance *ABA generalization in combination with the decomposition in (128); 15 of these 24 bipartitions are securely attested.[45]

Finally, there are a couple more attested bipartitions whose status is somewhat more uncertain between *bona-fide* suppletion and morphophonological alternation. While the one in (145) is no cause for worry as it wouldn't challenge our predictions either way, the one in (146) (which Martin Maiden has christened the "U-pattern") deserves some closer inspection as it potentially constitutes a counterexample to our prediction in (129a), repeated here as (147).

(145) Bavarian 'go' (Plank 2013)

3SG 3PL 2PL (but 3SG 2SG 2PL!)

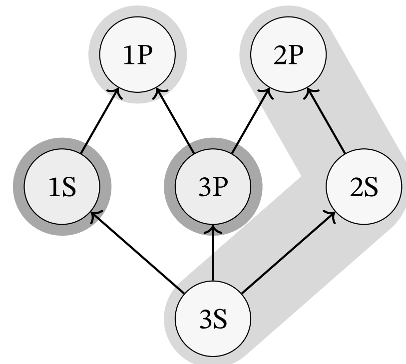
	SG	PL
1	ge:	gɛŋ-ɐn
2	ge:-sd	ge-ts
3	ge:-d	gɛŋ-ɐn



[45] The number of possible bipartitions of this six-member paradigm is of course the same as for the other six-member paradigm involving case and number (the Stirling number of the second type $\left\{\begin{matrix}6\\2\end{matrix}\right\}$, i.e. 31), but the number of generalization-compliant partitions is now higher (24 here *vs* 14 for case and number) because of the non-strictly cumulative decomposition of person. It's therefore hopefully not an accident that the number of attested patterns is also too (15 here *vs* 11 for case and number).

(146) Italian 'be worth' (the "U-pattern")
3SG 1SG 1PL and 3SG 3PL 1PL !

	SG	PL
1	válg-o	val-jámo
2	vál-i	val-éte
3	vál-e	válg-ono



(147) 3SG cannot share an exponent with 1PL to the simultaneous exclusion of both 3PL and 1SG.

I think it's important to distinguish at least two broad types of instances of the potentially problematic "U-pattern" in (146). On the one hand, we have cases like contemporary standard Italian *valere* 'be worth' (148): here the alternation -g- ~ ∅, while most likely not a part of Italian regular phonology, certainly looks like something that a sensible phonological rule could effect; and furthermore, as confirmed by a quick look at a bigger portion of the paradigm, it is exceptionlessly conditioned by a phonological trigger—whether the stem is followed by a front vocoid or not.[46] I therefore take the alternation to be the result of a lexically restricted phonological process, and thus to fall outside of the scope of an empirical typology of morphosyntactically conditioned allomorphy.

(148) Contemporary standard Italian 'be worth'

	INDICATIVE		SUBJUNCTIVE	
	SG	PL	SG	PL
1	válg-o	val-jámo	válg-a	val-jámo
2	vál-i	val-éte	válg-a	val-játe
3	vál-e	válg-ono	válg-a	válg-ano

Contrast this with the case of Old Tuscan *potere* 'can, be able to' in (149).[47] (Ignore the regular phonological alternations between *wɔ*, *ɔ*, and *o*, which are just conditioned by stress and syllable structure.)

[46] Other such alternations turn out to not only be conditioned by phonological properties of the environment, but to even be fully subsumable under regular phonology. See Steddy (2015) and especially Flor (2023) for such an approach to Italian velar palatalization, and Steriade (2022: §4.2.2) for broader discussion.
[47] The indicative present of *potere* in current Italian has reverted to the original Latin pattern in (144), which is unproblematically consistent with the prediction in (129a).

(149) Old Tuscan ‘can, be able to’

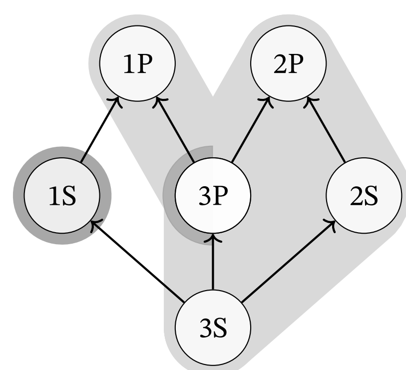
	INDICATIVE		SUBJUNCTIVE	
	SG	PL	SG	PL
1	pɔ́ss-o	pot-émo	pɔ́ss-a	poss-jámo
2	pwɔ́t-i	pot-éte	pɔ́ss-i/-e	poss-játe
3	pwɔ́t-e	pɔ́ss-ono	pɔ́ss-a	pɔ́ss-ano

Here, if we had restricted our focus to the present indicative alone, we might still have come away with the impression that the alternation was phonologically conditioned — once again, by whether or not the stem was followed by a front vocoid. However, in contrast to what we saw with (148), expanding our focus to the subjunctive here no longer confirms that hypothesis, but actively disproves it, as the allomorph *pɔss-*/*poss-* turns out to apply throughout the subjunctive regardless of the frontness or backness of what follows it. The resulting distribution in (149) proves actually quite difficult to make sense of in phonological terms, and therefore, as an instance of *bona-fide* non-phonologically conditioned allomorphy, it threatens to pose a real challenge to our prediction in (129a).

It is therefore crucial to note, in this connection, that (149) does not in fact include all the possible forms for ‘can’ in Old Tuscan — and specifically that Rohlfs (1968: 282–283) reports a form *pɔ́nno* as commonly alternating with *pɔ́ssono* in 3PL (Dante, for one, uses both). If 3PL *pɔ́nno* is analyzed as underlyingly /pɔt+no/, with exceptional lack of a theme vowel and subsequent assimilation to avoid an illicit cluster *-tn-*, then there does in fact exist a directed path from 3SG to 1PL all covered by one and the same allomorph /pɔt-/.

(150) Old Tuscan ‘can, be able to’
3SG 1SG 1PL but 3SG 3PL 1PL
(alongside 3SG 3PL 1PL)

	SG	PL
1	pɔ́ss-o	pot-émo
2	pwɔ́t-i	pot-éte
3	pwɔ́t-e	pɔ́n-no/ pɔ́ss-ono



I therefore conclude that the U-pattern is, despite initial appearances, actually not a real threat to our predictions in (129).

3.3.3 Adding mood to the picture

Finally, before moving on to the next chapter, I'd like to at least briefly bring mood into the picture, as I think it might be particularly helpful in bringing out how the puzzles we grappled with in the (pro)nominal domain are really parallel to some of the classical "morphomic" puzzles posed by verbs.

In the (pro)nominal domain, we saw that a broadly Pāṇinian approach to allomorph selection together with strictly cumulative case- and number-decompositions led us to expect that, given an ABB pattern in the singular, we might find the same ABB pattern in the plural, but not a BBB or an AAA pattern. We also saw, however, that AAA and BBB patterns in the plural are both actually attested (cf. (151)). An analogous problem arises for AAB patterns in the singular too, of course (cf. (152)).

(151)

	N.W. Lovari		Icelandic DEM		Malayalam	
	3M.SG	3PL	M.SG	M.PL	1SG	1PL.EX
NOM	**vou**	**vou**-n	**s**-á	**þ**-eir	**ñān**	**ñaŋ**-ŋaḷ
ACC	**le**-s	**le**-n	**þ**-ann	**þ**-á	**enn**-e	**ñaŋ**-ŋaḷ-e
OBL	**le**-ske	**le**-nge	**þ**-eim	**þ**-eim	**enn**-āl	**ñaŋ**-ŋaḷ-āl

(152)

	Old H. German		Ancient Greek 'big'		Yiddish	
	3F.SG	3F.PL	M.SG	M.PL	3F.SG	3PL
NOM	**sī**, **si**(-u)	**si**-o	**méga**-s	**megál**-oi	**z**-i	**z**-ej
ACC	**si**-a/-e	**si**-o	**méga**-n	**megál**-ūs	**z**-i	**z**-ej
DAT	**i**-ru/-ro	**i**-m/-n	**megál**-ōi	**megál**-ois	**i**-r	**z**-ej

The exact same problem arises in verbal paradigms. Assuming a Pāṇinian approach to allomorph selection, we would expect a PART.PL-*vs*-everything-else pattern in the indicative to potentially coexist with the same pattern in the subjunctive, but we wouldn't expect the PART.PL allomorph from the indicative to spread all over the subjunctive— nor, if the indicative were the absence of mood features, would we expect the indicative's other ("elsewhere") allomorph to spread into the subjunctive (as this would require the PART.PL allomorph to be specified for indicative mood). Once again, however, all three possibilities are actually attested (Maiden 2018: ch. 6; Kühlert 2023: 272–273).

(153) Italian 'go'

	INDICATIVE		SUBJUNCTIVE	
	SG	PL	SG	PL
1	vado	andiamo	vada	andiamo
2	vai	andate	vada	andiate
3	va	vanno	vada	vadano

(154) Languedocien 'go' (Maiden 2018: 195)

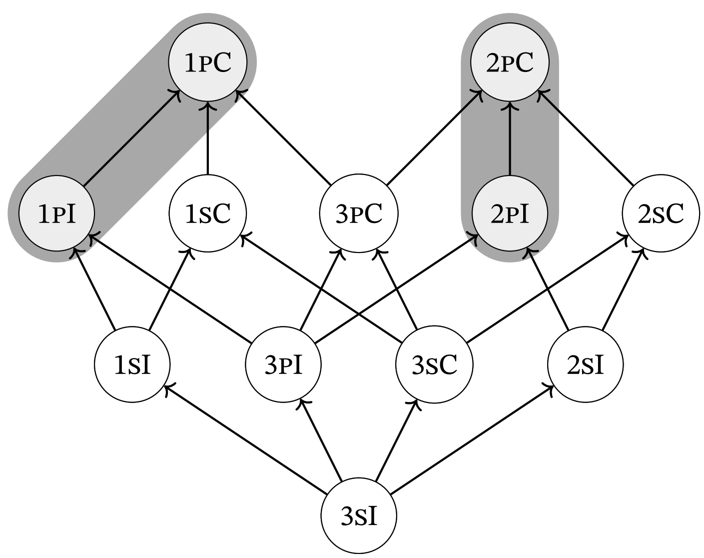
	INDICATIVE		SUBJUNCTIVE	
	SG	PL	SG	PL
1	vau	anam	ane	anem
2	vas/vai	anatz	anes	anetz
3	va	van	ane	anen

(155) Cascinagrossa 'go' (Maiden *et al.* 2010)

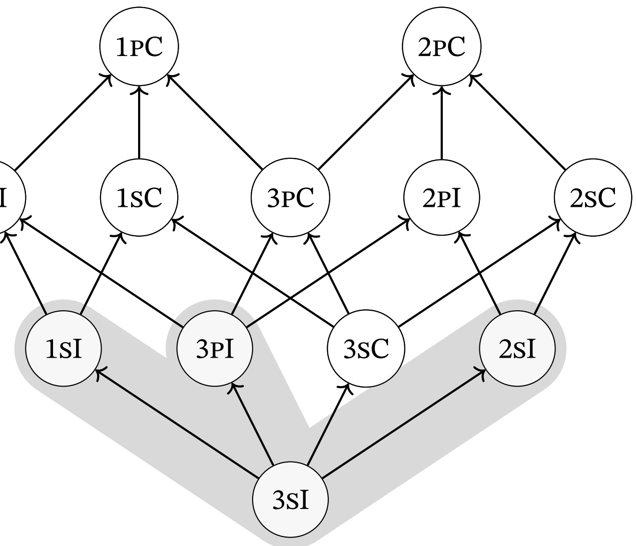
	INDICATIVE		SUBJUNCTIVE	
	SG	PL	SG	PL
1	vag	andómma	vaga	vagen
2	vɛ	andɛ́	vag	vagi
3	va	vaŋn	vaga	vagen

Just as we saw for case and number paradigms, however, none of the three possibilities violate the minimal-compliance *ABA generalization.

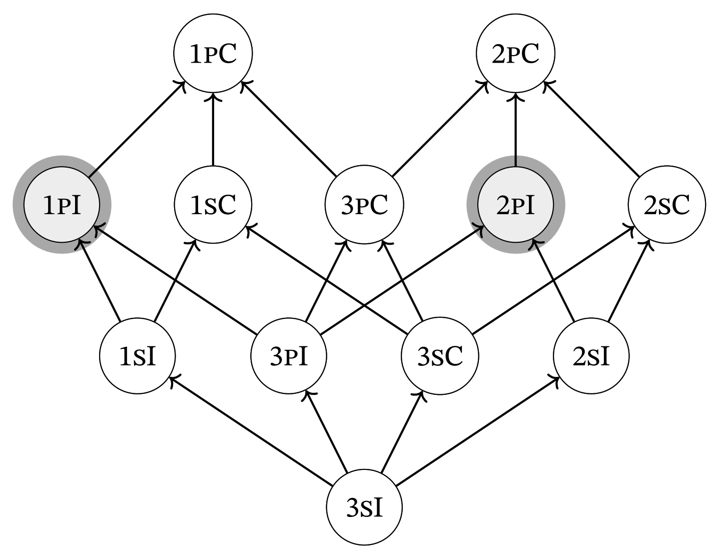
(156) Italian



(157) Languedocien



(158) Cascinagrossa



As I foreshadowed in Chapter 1, the (pro)nominal and verbal problems thus clearly deserve a unified solution — one that derives the minimal-compliance *ABA generalization on principled grounds.

Chapter 4

An Optimality-Theoretic proposal

In this chapter, I'm going to develop a proposal that will derive, among other things, the minimal-compliance *ABA generalization from Chapter 3.

4.1 The proposal

4.1.1 Background

To make our desideratum more concrete, let's focus on how that generalization plays out in a portion of the simple case-number paradigm that served as our main case study in Chapters 2–3. In particular, let's look at the two Russian-doll triples that start with NOM.SG and end with ACC.PL — ⟨NOM.SG, ACC.SG, ACC.PL⟩ and ⟨NOM.SG, NOM.PL, ACC.PL⟩. Once again, what we want is a theory whereby either of these triples can instantiate ABA as long as the other one instantiates AAA.

We've already witnessed one theory fall short of this desideratum. That theory, based on Underspecification and Pāṇinian ordering, maps each abstract morpheme M onto the most specified available exponent that is not overspecified with respect to M — i.e. that is not specified for any target feature that M does not contain, nor for any context feature that the context of M does not contain. Faced with our two Russian-doll triples, the theory would incorrectly predict that neither could ever instantiate ABA.

Something I haven't mentioned so far, however, is that Underspecification is not a unanimously adopted assumption. In particular, there is a family of approaches, commonly grouped under the rubric of Nanosyntax (Caha 2009; Starke 2009; Baunaz & Lander 2018; Starke 2018), which replace it with its very opposite, namely Overspecification. According to these alternatives, an exponent can be specified for features that are extraneous to its *exponendum*;[48] what it cannot do, however, is be underspecified,

[48] When talking about Nanosyntax, I'll use the term *exponendum* instead of *morpheme*, because, despite

i.e. fail to be specified for any of the features that its exponendum does involve. If we then combine this theory of eligibility with our familiar principle of Pāṇinian ordering (repeated in (159), only with *exponendum* substituted for *morpheme*), we'll see our criterion for exponent selection get even more thoroughly turned on its head: rather than picking out the most specified non-overspecified exponent, we're going to pick out the least specified non-underspecified exponent.

(159) *Pāṇinian ordering*
If exponents e_1 and e_2 are both eligible to expone exponendum E, and the set of exponenda that e_1 is eligible to expone is a proper subset of the set of exponenda that e_2 is eligible to expone, then e_2 does not expone E.

If we try to just substitute this "reverse" exponence logic for our previous one,[49] however, it's not clear that we're going to be any better off. In fact, it seems as though we're bound to run into the exact same problem as before, by once again wrongly predicting no Russian-doll triples to ever instantiate ABA. To see why that's the case, consider just one of our two Russian-doll triples — say, ⟨NOM.SG, NOM.PL, ACC.PL⟩, i.e. $\langle\varnothing, \{\#_{\text{PL}}\}, \{\#_{\text{PL}}, \varkappa_{\text{ACC}}\}\rangle$ — and assume that it does instantiate ABA.

(160)
a. By hypothesis, exponent *b* applies in NOM.PL, which means, by Overspecification, that *b* must at least be specified for $\#_{\text{PL}}$.
b. Again by Overspecification, *b* is thus also eligible to apply in NOM.SG. However, by hypothesis, it's not *b* that applies in NOM.SG; it's exponent *a*.
c. By Pāṇinian ordering, this means that *a* must be less specified than *b*.
d. But then, if *a* was eligible to also apply in NOM.PL, it would win there, too.
e. The fact that that's not the case means that *a* must be eligible to apply in NOM.SG but ineligible to apply in NOM.PL.
f. This is only possible if *a* is not specified for $\#_{\text{PL}}$ — the only feature present in NOM.PL but not in NOM.SG.
g. But then, by Overspecification, *a* isn't eligible in ACC.PL either — hence *ABA.

It's easy to see that the proof can be replicated for the alternative-route triple ⟨NOM.SG, ACC.SG, ACC.PL⟩ (i.e. $\langle\varnothing, \{\varkappa_{\text{ACC}}\}, \{\#_{\text{PL}}, \varkappa_{\text{ACC}}\}\rangle$) by simply substituting ACC.SG for NOM.PL

being both realizational and piece-based, Nanosyntax actually has no place for the classical concept of the morpheme (see, among others, Starke 2009 and Caha 2018).

[49] This is, of course, not a fair characterization of the Nanosyntactic approach to the problem. Nanosyntax is, in fact, a new approach to the very architecture of the mapping between syntax and exponent, and in most of its current instantiations rejects the notion of contextual allomorphy altogether to subsume it under portmanteau morphology (see e.g. Caha, De Clercq & Vanden Wyngaerd 2019). The exercise in the main text is only intended to approximate a "minimal pair" between the logics of underspecification and overspecification for illustrative purposes.

throughout and κ_{ACC} for $\#_{PL}$ in (160f). Once again, therefore, we're stuck with too strong a *ABA generalization — maximal- rather than minimal-compliance.

What crucially remains to be explored, however, is what would happen if the possibilities of Underspecification and Overspecification were to be combined together. It might at first be unclear what such a system would even look like, but fortunately a few attempts in this direction have already been made in the literature, especially in the context of efforts to rethink the whole mapping from morphosyntax to surface phonology in terms of ranked violable constraints. Specifically, the violable constraints at issue are the ones defined in (161) (cf. Ackema & Neeleman 2005, Müller 2020, and especially Wolf 2008),[50] which are simply the morphosyntactic-featural counterparts of the more familiar phonological-featural faithfulness constraints MAX and DEP from McCarthy & Prince (1995).[51]

(161) a. MAX: Assign one star to the mapping between abstract morpheme M and exponent *e* for every morphosyntactic feature that is contained in the context of M but not in the context specification of *e*.

b. DEP: Assign one star to the mapping between abstract morpheme M and exponent *e* for every morphosyntactic feature that is contained in the context specification of *e* but not in the context of M.

From this perspective, Wolf (2008: 68ff) remarks that the effects of classical Pāṇinian Underspecification may be mimicked by means of an inviolable DEP and a violable MAX (the competition being restricted to the available exponents that obey DEP, among which only the one with the fewest MAX violations emerges as the winner), while the effects of Overspecification may be mimicked by an inviolable MAX and a violable DEP (the competition being restricted to the available exponents that obey MAX, among which only the one with the fewest DEP violations emerges as the winner).[52]

As Wolf (2008) also notes, however, once we avail ourselves of a grammatical theory that lets us have "highly conflicting" constraints thanks to "a general means of resolving their conflicts" (Prince & Smolensky 1993: §1.1), it becomes unproblematic to overcome this dichotomy and to countenance both some underspecification and some overspecification at once, by having both MAX and DEP as active — but crucially violable —

[50] The MAX-side of the story has a longer history, including Noyer (1993), Bonet (1994), Trommer (2001), Wunderlich (2001), Kiparsky (2005), and Strigin (2007).

[51] Previously introduced under the names of PARSE and FILL, respectively, by Prince & Smolensky (1993). I will adopt the more familiar names MAX and DEP here, despite the risk of confusion arising from the fact that DEP is also the beginning of the word *dependent* (as in *dependent case*).

[52] To be more accurate, Wolf focuses on alternative formulations of Pāṇinian ordering that hinge on counting features rather than on subsethood relations between sets of eligibility contexts (Halle's (1997) and Starke's (2009), rather than Kiparsky's (1982)), and whose effects are tehrefore not just mimicked but actually exactly matched by MAX and DEP respectively.

constraints. In §4.1.3, I will go back to our two Russian-doll triples to show that this will finally derive minimal-compliance *ABA effects, but let me first set up the system in full detail in §4.1.2.

4.1.2 The system in full detail

In this subsection, I will explicitly articulate every aspect of my proposal — how it situates itself within the broader architecture of grammar, what constraints it involves, how those constraints interact to evaluate possible mappings from abstract morphemes to exponents, and what exactly the inputs and outputs of those mappings can(not) be.

Architectural implications

Like the more traditional Pāṇinian approaches we explored in Chapters 2–3, the current proposal too is couched within a post-syntactic and morpheme-based (or, in Stump's 2001 terms, *realizational* and *lexical*) approach to morphology — that is, I merely need to still assume that, sometime after embedding a given abstract morpheme within a given morphosyntactic context, the grammar is going to be tasked with mapping that abstract morpheme onto an exponent. While the approaches we discussed in Chapters 2–3 contend that the competition among exponents that's at the core of this mapping is adjudicated by Underspecification and Pāṇinian ordering, I'm arguing instead that the competition in question is adjudicated by a set of violable and sometimes-conflicting morphosyntactic faithfulness constraints (more on which immediately below). Crucially, no revisions of the general architecture of the grammar — and, in particular, no repeal of the modular separation between exponent selection and phonology — are in any way required.

Again in the interest of keeping my readership as broad as possible, I'm also remaining noncommittal regarding the existence or nonexistence of morphosyntactic-feature manipulations such as Impoverishment, taking place between the syntax proper and exponent selection. In what follows, however, I will articulate my proposal as if assuming such operations to be unavailable — both in order to provide a proof of principle to the effect that a lean, restrictive system can accomplish several nontrivial results, and in order to preempt the worry that such operations, if left unconstrained, might threaten our account of *ABA-related generalizations (e.g. by impoverishing the dative — but not the accusative! — into the corresponding nominative; cf. Caha 2018; Zompì 2019).

Constraints

The only constraints that I take to be operative in adjudicating the competition among available exponents are just instances of Max and Dep relativized to single inflectional dimensions. For the simple paradigm of our main case study, the relevant inflectional dimensions are going to be just case and number, yielding a total of four relevant constraints (162). For more complex paradigms, we're just going to add to our constraint set both a dedicated MAX and a dedicated DEP for every inflectional dimension that we're going to add.

(162) a. MAX(#): Assign a star for every number feature that is contained in the context of the abstract morpheme but not in the context specification of the exponent.
b. MAX(K): Assign a star for every case feature that is contained in the context of the abstract morpheme but not in the context specification of the exponent.
c. DEP(#): Assign a star for every number feature that is contained in the context specification of the exponent but not in the context of the abstract morpheme.
d. DEP(K): Assign a star for every case feature that is contained in the context specification of the exponent but not in the context of the abstract morpheme.

This means that, just as in previously discussed approaches, we're going to evaluate competing exponents only in terms of their morphosyntactic-featural faithfulness to the relevant abstract morpheme and its context. No other considerations — in particular, no markedness constraints of any sort — play a role in the evaluation.

Rankings and EVAL

As in Prince & Smolensky's (1993) Optimality Theory (OT), the competition among available exponents is resolved based on how the operative constraints are *ranked* relative to each other. More specifically, I'm going to conjecture that all and only the linear orders over the constraint set are possible rankings, and that, for each language L, there's going to be at least one ranking that all of L's data are consistent with.[53]

Given an abstract morpheme in its morphosyntactic context, a set of available exponents for it (*candidates*), and a constraint ranking $C_1 \gg \cdots \gg C_n$, the competition for the mapping from morpheme to exponent is adjudicated by the standard EVAL component of OT. Each constraint, starting from the top-ranked C_1 downward, narrows down the set of candidates in play: first, we filter out every candidate e such that there is some other candidate that violates C_1 strictly fewer times than e does; then we take the set of candidates that survived this first cut and we narrow it down further by applying the same procedure for C_2; and so on and so forth, until either we narrow the set of candidates in play down to a singleton or we reach the bottom-ranked constraint C_n — whichever happens earlier.

[53] In §4.3.3, we will see one challenge that Icelandic poses for this latter conjecture.

The constraint ranking therefore has the property of so-called *strict domination* ("Each constraint has absolute priority over all the constraints lower in the hierarchy"— Prince & Smolensky 1993: §1.1) and each evaluation is performed in parallel fashion ("the Input → Output map has no internal structure: all possible variants are [...] evaluated in parallel"— Prince & Smolensky 1993: §1.2).

Contexts and candidates

Finally, what are the possible inputs and outputs of the mapping? That is, what are the possible contexts to be faithful to, and what are the possible candidates from which to choose the most faithful one?

Here I'm going to continue to adopt the cumulative decompositions of case and number that we've been adopting so far— which means that, for the simple paradigm of our main case study, the possible contexts are all those in (163).

(163) $\begin{bmatrix}\#: & \\ \text{K}: & \end{bmatrix}, \begin{bmatrix}\#: & \\ \text{K}: & \varkappa_{\text{DEP}}\end{bmatrix}, \begin{bmatrix}\#: & \\ \text{K}: & \varkappa_{\text{DEP}}, \varkappa_{\text{DAT}}\end{bmatrix}, \begin{bmatrix}\#: & \#_{\text{PL}} \\ \text{K}: & \end{bmatrix}, \begin{bmatrix}\#: & \#_{\text{PL}} \\ \text{K}: & \varkappa_{\text{DEP}}\end{bmatrix}, \begin{bmatrix}\#: & \#_{\text{PL}} \\ \text{K}: & \varkappa_{\text{DEP}}, \varkappa_{\text{DAT}}\end{bmatrix}$

As for the candidates, given an abstract morpheme to expone in a context, these are simply the contextual allomorphs that happen to be available for that morpheme in the vocabulary of the language in question. In particular, I assume that such allomorphs may bear any of the possible case/number contexts in (163) as their case/number context specifications. Each language will arbitrarily determine how many allomorphs it's going to have for each morpheme, and which of these possible context specifications each of those allomorphs will bear.

Notice, in particular, that I'm assuming that no exponent may have $\varkappa_{\text{DAT}}$ in its context specification without also having $\varkappa_{\text{DEP}}$. In fact, I'd like to more generally assume that the entailment relations that hold between features in the syntax also hold between them in exponents' possible context specifications— an inviolable constraint that exponents have to comply with in order to even just enter the competition. This assumption, which will play a crucial role in allowing the current theory to derive unidimensional *ABA (see §4.1.4), strikes me as obeying a certain intuitive rationale: after all, precisely because $\varkappa_{\text{DAT}}$'s presence entails $\varkappa_{\text{DEP}}$'s in the syntax, the morphosyntactic contexts including $\{\varkappa_{\text{DAT}}\}$ coincide with the ones including $\{\varkappa_{\text{DEP}}, \varkappa_{\text{DAT}}\}$, so that specifying an exponent to appear in the former set of contexts should be equivalent to specifying it to appear in the latter.

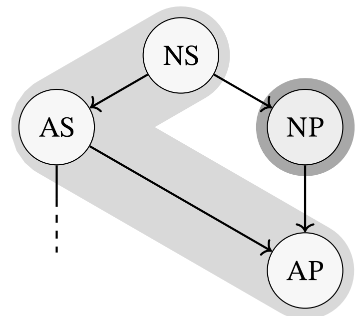
4.1.3 The key appeal

Now that we have the whole system in place, we can finally go back to our two Russian-doll triples to see how it delivers the generalization we wanted.

Let's first try to derive an ABA pattern on the triple ⟨NOM.SG, NOM.PL, ACC.PL⟩ while having AAA on ⟨NOM.SG, ACC.SG, ACC.PL⟩, as in (164).

(164) Eastern Frisian

	3M.SG	3PL
NOM	**h**-äi	**z**-äi
ACC	**h**-um	**h**-ör



Let's specify *z*- for $\{\#_{\text{PL}}\}$ and *h*- for $\{\kappa_{\text{ACC}}\}$. If we do so, then NOM.PL and ACC.SG will be trivial cases: in the former, *z*- will be fully faithful to the context, and therefore win regardless of the ranking; in the latter, *h*- will.

(165) Exponents for East Frisian 3M: $z\text{-} :: \begin{bmatrix} \#: & \#_{\text{PL}} \\ \text{K}: & \end{bmatrix}$ $h\text{-} :: \begin{bmatrix} \#: & \\ \text{K}: & \kappa_{\text{ACC}} \end{bmatrix}$

a. Nominative plural → *z*- is the fully faithful candidate

b. Accusative singular → *h*- is the fully faithful candidate

The interesting cases are the other two. In NOM.SG, absent a candidate with an empty context specification, we're bound to incur one DEP violation either way. By ranking DEP(#) above DEP(K), we're going to make sure that bearing a spurious $\#_{\text{PL}}$ will be more harshly penalized than bearing a spurious κ_{ACC}, thereby handing victory to *h*- (see (166a)). Conversely, in ACC.PL, we're bound to have one MAX violation either way. By ranking MAX(K) above MAX(#), we're going to make sure that lacking an expected κ_{ACC} will be more penalized than lacking an expected $\#_{\text{PL}}$, thereby handing victory to *h*- again (see (166b)). The pattern is thus unproblematically generated.

(166) a. Nominative singular

Input :: $\begin{bmatrix} \#: & \\ \text{K}: & \end{bmatrix}$	DEP(#)	DEP(K)
z- :: $\begin{bmatrix} \#: & \#_{\text{PL}} \\ \text{K}: & \end{bmatrix}$	*	
☞ *h*- :: $\begin{bmatrix} \#: & \\ \text{K}: & \kappa_{\text{ACC}} \end{bmatrix}$		*

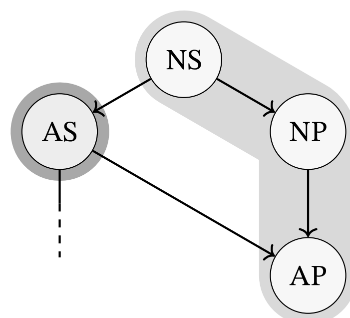
b. Accusative plural

Input :: $\begin{bmatrix} \#: & \#_{\text{PL}} \\ \text{K}: & \kappa_{\text{ACC}} \end{bmatrix}$	MAX(K)	MAX(#)
z- :: $\begin{bmatrix} \#: & \#_{\text{PL}} \\ \text{K}: & \end{bmatrix}$	*	
☞ *h-* :: $\begin{bmatrix} \#: & \\ \text{K}: & \kappa_{\text{ACC}} \end{bmatrix}$		*

Things are not so different when it comes to the pattern in (167), featuring an ABA pattern on ⟨NOM.SG, ACC.SG, ACC.PL⟩ and an AAA pattern on ⟨NOM.SG, NOM.PL, ACC.PL⟩.

(167) Malayalam

	1SG	1PL.EX
NOM	**ñān**	**ñaŋ**-ŋaḷ
ACC	**enn**-e	**ñaŋ**-ŋaḷ-e



In terms of the candidates' context specifications, the strategy is going to remain the same as for the previous pattern — namely, specifying the two candidates for $\{\#_{\text{PL}}\}$ and $\{\kappa_{\text{ACC}}\}$, respectively. Therefore, NOM.PL and ACC.SG will once again be trivial cases, each with a fully faithful candidate that will win irrespective of the ranking.

(168) Exponents for Malayalam 1EX: *ñan-* :: $\begin{bmatrix} \#: & \#_{\text{PL}} \\ \text{K}: & \end{bmatrix}$ *enn-* :: $\begin{bmatrix} \#: & \\ \text{K}: & \kappa_{\text{ACC}} \end{bmatrix}$

a. Nominative plural → *ñan-* is the fully faithful candidate

b. Accusative singular → *enn-* is the fully faithful candidate

As for the other two contexts, we are once again going to have a DEP violation in NOM.SG either way, as well as a MAX violation in ACC.PL either way. In order to make sure that in each of these two contexts the outcome turns out different from what we saw in (166), all we need to do is swap around the relative ranking of both DEP constraints (DEP(K) ≫ DEP(#) this time) and MAX constraints (MAX(#) ≫ MAX(K) this time). The desired pattern emerges unproblematically once again.

(169) a. Nominative singular

Input :: [#: ; K:]	DEP(K)	DEP(#)
☞ *ñan-* :: [#: $\#_{PL}$; K:]		*
enn- :: [#: ; K: κ_{ACC}]	*	

b. Accusative plural

Input :: [#: $\#_{PL}$; K: κ_{ACC}]	MAX(#)	MAX(K)
☞ *ñan-* :: [#: $\#_{PL}$; K:]		*
enn- :: [#: ; K: κ_{ACC}]	*	

Now that we've shown that the system can generate both of the relevant minimally compliant patterns, we also need to show that it can<u>not</u> generate the non-compliant pattern — the one whereby NOM.SG and ACC.PL pattern alike to the simultaneous exclusion of both ACC.SG and NOM.PL. In other words, we need to prove that, if exponent *a* wins in both NOM.SG and ACC.PL, then *a* will <u>have to</u> win in ACC.SG or NOM.PL or both.

So let's suppose that exponent *a* does win in both NOM.SG and ACC.PL. Now, for each input context, let's draw out an abbreviated tableau, where, below each constraint *C*, we just write:

- "W" if *a* scores strictly fewer violations of *C* than does any other candidate (i.e. if *C* would pick *a* as the unique <u>winner</u>);
- "T" if *a* scores no more violations of *C* than any other candidate, but *a* scores exactly as many violations of *C* as some other candidate *b* (i.e. if *C* would pick *a* as a winner in a <u>tie</u> with *b*);
- "L" if *a* scores strictly more violations of *C* than some other candidate *b* (i.e. if *C* would declare *a* as a <u>loser</u>).

In each of the contexts where *a* ultimately wins, the top portion of the tableau will thus be a (possibly empty) sequence of T's immediately followed by at least one W. (If there were nothing but T's, *a* would ultimately end up in a tie, and if there were any L's to the left of the leftmost W, *a* would lose.)

Now there are three possibilities:

(170) a. Leftmost W in NOM.SG and leftmost W in ACC.PL are in the same column:

	...	C_k	...
NOM.SG	(T ... T)	W	...
ACC.PL	(T ... T)	W	...

b. The leftmost W in NOM.SG is to the left of the leftmost W in ACC.PL:

	...	C_k	...	C_m	...
NOM.SG	(T ... T)	W	...	...	...
ACC.PL	(T ... T)	T	(T ... T)	W	...

c. The leftmost W in NOM.SG is to the right of the leftmost W in ACC.PL:

	...	C_k	...	C_m	...
NOM.SG	(T ... T)	T	(T ... T)	W	...
ACC.PL	(T ... T)	W	...	...	...

Given a ranking $C_1 \gg \ldots \gg C_n$, there is thus a top portion of that ranking $C_1 \ggg \ldots \ggg C_k$ such that,

- for $1 \leq i < k$, C_i yields T in both NOM.SG and ACC.PL
- and C_k, $\begin{cases} \text{in (170a), yields W in both NOM.SG and ACC.PL} \\ \text{in (170b), yields W in NOM.SG and T in ACC.PL} \\ \text{in (170c), yields T in NOM.SG and W in ACC.PL} \end{cases}$

Now let's first tackle the constraints to the left of C_k, which yield T in both NOM.SG and ACC.PL. It is crucial to remember, at this point, that each of our constraints only cares about a single inflectional dimension, and therefore will never favor distinct candidates for two different input contexts unless those two contexts differ with respect to that inflectional dimension. This means that a number-relativized constraint (be it MAX(#) or DEP(#)), having yielded T in NOM.SG, will always also yield T in other singular contexts such as ACC.SG, and having yielded T in ACC.PL, it will always also yield T in other plural contexts such as NOM.PL. Likewise, a case-relativized constraint that has yielded T in NOM.SG will always also yield T in NOM.PL, and one that has yielded T in ACC.PL will also yield T in ACC.SG. This guarantees that, if any such constraint (one that's blind to all but one inflectional dimension) yields T in both NOM.SG and ACC.PL, it will also yield T in both NOM.PL and ACC.SG. The exact same reasoning also applies to C_k in (170a): if C_k yields W in both NOM.SG and ACC.PL, it is guaranteed to also yield W in both NOM.PL and ACC.SG.

(171) The leftmost W in NOM.SG is to the left of the leftmost W in ACC.PL:

	…	C_k	…
NOM.SG	(T … T)	W	…
ACC.SG	(T … T)	W	…
NOM.PL	(T … T)	W	…
ACC.PL	(T … T)	W	…

What about C_k in the other two scenarios in (170b) and (170c)? Here we can't know for sure what's going to happen in ACC.SG and what in NOM.PL — but we do know what's going to happen in at least one of them. In (170b), for example, we know that C_k yields W in NOM.SG. Althought we can't tell whether C_k is relativized to case or number, we can tell that, if it's relativized to case, it's also going to yield W in the other nominative context NOM.PL, whereas, if it's relativized to number, then it's also going to yield W in the other singular context ACC.SG. Likewise, in (170c), we know that C_k yields W in ACC.PL; so, if it's relativized to case, it will also yield W in the other accusative context ACC.SG, while, if it's relativized to number, it will also yield W in the other plural context NOM.PL. In each of these scenarios, C_k will surely yield a W in either ACC.SG or NOM.PL.

(172) a. The leftmost W in NOM.SG is to the left of the leftmost W in ACC.PL:

	…	C_k	…
NOM.SG	(T … T)	W	…
ACC.SG	(T … T)	W	…
NOM.PL	(T … T)	T	…
ACC.PL	(T … T)	T	… W

or

	…	C_k	…
NOM.SG	(T … T)	W	…
ACC.SG	(T … T)	T	…
NOM.PL	(T … T)	W	…
ACC.PL	(T … T)	T	… W

b. The leftmost W in NOM.SG is to the right of the leftmost W in ACC.PL:

	…	C_k	…
NOM.SG	(T … T)	T	… W
ACC.SG	(T … T)	W	…
NOM.PL	(T … T)	T	…
ACC.PL	(T … T)	W	…

or

	…	C_k	…
NOM.SG	(T … T)	T	… W
ACC.SG	(T … T)	T	…
NOM.PL	(T … T)	W	…
ACC.PL	(T … T)	W	…

This provides us with the last piece we were missing to finally rule out the non-compliant pattern. Now we know that all the constraints to the left of C_k (if any) are going to yield T's in both ACC.SG and NOM.PL, and that C_k will finally deliver the decisive W for exponent a in ACC.SG or NOM.PL or both: no "double ABA" is possible.

4.1.4 Deriving unidimensional *ABA

The alert reader might've noticed that the proof I sketched out in the last subsection crucially relied on the fact that each of the four contexts in play (NOM.SG, ACC.SG, NOM.PL, and ACC.PL) had the same case as one other one of them (e.g. NOM.SG has the same case as NOM.PL) and the same number as yet another one of them (e.g. NOM.SG has the same number as ACC.SG). However, this condition clearly doesn't hold of the contexts in a unidimensional Russian-doll triple: none of the members of ⟨NOM.SG, ACC.SG, DAT.SG⟩, for example, has the same case as either of the other two. Our minimal-compliance–*ABA result, therefore, cannot be straightforwardly replicated for unidimensional Russian-doll triples; we'll need to sketch out a different proof to show that it does.

The initial setup of this proof, however, will be essentially the same as before. Let's assume that a candidate *a* wins in both a nominative and the corresponding dative, and once again, for each input context, let's write an abbreviated '*a*-centric' tableau, where below each constraint *C* we write "W" if *a* is the unique winner for *C*, "L" if *a* is a loser for *C*, and "T" if *a* is in a tie as one of the winners for *C*. Once again, in each of the contexts where *a* ultimately wins, the top portion of the tableau will thus be a (possibly empty) sequence of T's immediately followed by at least one W — specifically as in one of the three possibilities in (173).

(173) a. The leftmost W in NOM is to the left of the leftmost W in DAT:

	…	C_k	…	C_m	…
NOM	(T … T)	W	…	…	…
DAT	(T … T)	T	(T … T)	W	…

b. The leftmost W in NOM is to the right of the leftmost W in DAT:

	…	C_k	…	C_m	…
NOM	(T … T)	T	(T … T)	W	…
DAT	(T … T)	W	…	…	…

c. Leftmost W in NOM and leftmost W in DAT are in the same column:

	…	C_k	…
NOM	(T … T)	W	…
DAT	(T … T)	W	…

Let's start by looking at (173a). Given that C_k is yielding different outcomes in two input contexts that by assumption only differ in case features, C_k can't be relativized to any inflectional dimension other than case; rather, it must be either MAX(K) or DEP(K).

Suppose C_k is MAX(K). Then a tie in the DAT context is only possible if our candidate a has the exact same case-featural context specification as some other candidate b; however, if such were the case, then a couldn't possibly be the sole winner in the NOM context — so C_k must be DEP(K). This leaves us with three possible scenarios:

(174) a. $a :: \varnothing \quad \sim b :: \{\kappa_{\text{ACC}}\}$

b. $a :: \varnothing \quad \sim b :: \{\kappa_{\text{ACC}}, \kappa_{\text{DAT}}\}$

c. $a :: \{\kappa_{\text{ACC}}\} \quad \sim b :: \{\kappa_{\text{ACC}}, \kappa_{\text{DAT}}\}$

In each of these scenarios, a would get an L from MAX(K) in the DAT context (by losing to b), and given that, by assumption, none of the constraints $C_1, \ldots, C_m$ are to yield an L in DAT, this means that MAX(K) must be lower-ranked than C_m. This means, in turn, that all constraints $C_1, \ldots, C_m$ with the sole exception of C_k/DEP(K) must be relativized to inflectional dimensions other than case, and therefore each yield the same outcome across all three rows.

(175) a. Scenario (174a)

	…	$C_k =$ DEP(K)	…	C_m	(…)	MAX(K)
NOM	(T … T)	W	(T … T)	W	(…)	T
ACC	(T … T)	T	(T … T)	W	(…)	L
DAT	(T … T)	T	(T … T)	W	(…)	L

b. Scenario (174b)

	…	$C_k =$ DEP(K)	…	C_m	(…)	MAX(K)
NOM	(T … T)	W	(T … T)	W	(…)	T
ACC	(T … T)	W	(T … T)	W	(…)	L
DAT	(T … T)	T	(T … T)	W	(…)	L

c. Scenario (174c)

	…	$C_k =$ DEP(K)	…	C_m	(…)	MAX(K)
NOM	(T … T)	W	(T … T)	W	(…)	T
ACC	(T … T)	W	(T … T)	W	(…)	L
DAT	(T … T)	T	(T … T)	W	(…)	L

In each of these scenarios, a is therefore the ultimate winner in the ACC context too.

Now let's move on to (173b), repeated here as (176).

(176)

	…	C_k	…	C_m	…
NOM	(T … T)	T	(T … T)	W	…
DAT	(T … T)	W	…	…	…

By the same logic we applied above, C_m must be either DEP(K) or MAX(K) here too. Suppose it's DEP(K). Then a tie in the NOM context would only be possible if a had the exact same case-featural context specification as som other candidate b; however, if such were the case, then a couldn't possibly be the sole winner in the DAT context — so C_k must be MAX(K). This once again leaves us with three possible scenarios.

(177)
a. $a :: \{\kappa_{\text{ACC}}\}$ $\sim b :: \varnothing$
b. $a :: \{\kappa_{\text{ACC}}, \kappa_{\text{DAT}}\}$ $\sim b :: \varnothing$
c. $a :: \{\kappa_{\text{ACC}}, \kappa_{\text{DAT}}\}$ $\sim b :: \{\kappa_{\text{ACC}}\}$

In each of these scenarios, a gets an L from DEP(K) in the NOM context. By the same logic we applied above, then this means that DEP(K) must be lower-ranked than C_m — which in turn means that all constraints $C_1, \dots, C_m$ with the sole exception of C_k/MAX(K) must be relativized to inflectional dimensions other than case, and hence each yield the same outcome across all three rows.

(178) a. Scenario (177a)

	…	$C_k = $ MAX(K)	…	C_m	(…)	DEP(K)
NOM	(T … T)	T	(T … T)	W	(…)	L
ACC	(T … T)	W	(T … T)	W	(…)	L
DAT	(T … T)	W	(T … T)	W	(…)	T

b. Scenario (177b)

	…	$C_k = $ MAX(K)	…	C_m	(…)	DEP(K)
NOM	(T … T)	T	(T … T)	W	(…)	L
ACC	(T … T)	W	(T … T)	W	(…)	L
DAT	(T … T)	W	(T … T)	W	(…)	T

c. Scenario (177c)

	…	$C_k = $ MAX(K)	…	C_m	(…)	DEP(K)
NOM	(T … T)	T	(T … T)	W	(…)	L
ACC	(T … T)	T	(T … T)	W	(…)	L
DAT	(T … T)	W	(T … T)	W	(…)	T

Once again, *a* is thus the ultimate winner in the ACC context too.

Finally, let's get to scenario (173c), repeated here as (179).

(179)

	…	C_k	…
NOM	(T … T)	W	…
DAT	(T … T)	W	…

Here, too, we have three possible scenarios to consider:

(180) a. C_k is MAX(K);

b. C_k is DEP(K);

c. C_k is relativized to an inflectional dimension other than case.

If C_k is MAX(K) (scenario (180a)) and *a* is the unique winner in the NOM context, then *a* must be the sole candidate available in the vocabulary (or else any existing alternative candidate would of course have no violations of MAX(K) in NOM either, and would therefore join *a* in a tie there); *a* must thus be the ultimate undisputed winner in the ACC context, too. Similarly, if C_k is DEP(K) (scenario (180b)) and *a* is the unique winner in DAT, *a* must again be the sole candidate in play—and hence must again be the ultimate winner in ACC, too. Finally, if C_k is not relativized to case (scenario (180c)), it will of course yield the same victorious outcome in ACC as it does in both NOM and DAT—but then the question arises: Could a case-relativized constraint C_j outrank C_k and make *a* the ultimate loser in ACC only, as hypothetically outlined in (181)?

(181) One last scenario to rule out

	…	C_j	…	C_k	…
NOM	(T … T)	T	(T … T)	W	…
ACC	(T … T)	L	(T … T)	W	…
DAT	(T … T)	T	(T … T)	W	…

Fortunately, however, it's easy to show that the scenario in (181) can never take place. If C_j were MAX(K), then, in order to get an L from it in ACC, *a* would have to specified for no case features, as opposed to some other available candidate *b* specified for at least $\varkappa_{\text{ACC}}$. If such were the case, however, *a* could not get a T in the DAT context—it would lose to *b* in DAT, too. Similarly, if C_j were DEP(K), then, in order to get an L from it in ACC, *a* would have to be specified for $\{\varkappa_{\text{ACC}}, \varkappa_{\text{DAT}}\}$, as opposed to some other candidate *b* specified for a smaller set of case features. But then again, *a* could not get a T in the NOM context—it would be bound to lose to *b* in NOM too. This finally concludes our case-by-case proof of the impossibility of generating unidimensional *ABA.

While the two proofs I've sketched out in the current and previous subsection are rudimentary and limited in scope, the core results demonstrably generalize. In particular, an explicit implementation in Python,[54] kindly provided by Enrico Flor, makes it possible to show by brute-force exhaustion that no combination between any of the possible candidate sets and any of the possible constraint rankings is able to generate a violation of minimal-compliance *ABA in any of the abstract paradigms considered here. For the time being, I will therefore content myself with this encouraging result, while leaving a fully general deductive proof for future work.

4.2 Capturing a pronominal morphome

While being restrictive enough to derive the desired generalization, the system is also powerful enough to straightforwardly account for certain morphomic patterns. In this section, in particular, I'm going to apply it to an already familiar example, namely the morphomic Old High German paradigm from §3.2.3, repeated here in (182).

(182) Old High German third-person pronoun (Braune 2004: 243)

	NEUTER		MASCULINE		FEMININE	
	SG	PL	SG	PL	SG	PL
NOM	**i**-ʒ	**si**-u	**ër**	**si**-e	**sī**, **si**(-u)	**si**-o
ACC	**i**-ʒ	**si**-u	**i**-n(an)	**si**-e	**si**-a, **si**-e	**si**-o
DAT	**i**-mu, **i**-mo	**i**-m, **i**-n	**i**-mu, **i**-mo	**i**-m, **i**-n	**i**-ru, **i**-ro	**i**-m, **i**-n

To get this off the ground, let's extend to gender the assumption that candidates are inviolably subject to entailment relations between features—in this case, no γ_{FEM} without γ—and let's obviously add MAX(Γ) and DEP(Γ) to the constraint set.

With these natural adjustments in place, the system can now unproblematically capture (182), as illustrated in (183). (For reasons of space, I'm leaving out the bottom-ranked constraints DEP(K) and DEP(#) since they are never decisive, and I'm entirely omitting the tableaux for dative contexts, where it should be easy to see that top-ranked MAX(K) will always ensure the victory of *i*-.)

(183) *i*- :: $\begin{bmatrix} \Gamma: & \\ \#: & \\ \text{K}: & \kappa_{\text{ACC}}, \kappa_{\text{DAT}} \end{bmatrix}$ *ër* :: $\begin{bmatrix} \Gamma: & \gamma_{\text{ANIM}} \\ \#: & \\ \text{K}: & \end{bmatrix}$ *si*- :: $\begin{bmatrix} \Gamma: & \gamma_{\text{ANIM}}, \gamma_{\text{FEM}} \\ \#: & \#_{\text{PL}} \\ \text{K}: & \kappa_{\text{ACC}} \end{bmatrix}$

[54] Accessible at stanislaozompi.net/aba.

a. Nominative singular neuter

UR :: $\begin{bmatrix} \Gamma: & \\ \#: & \\ K: & \end{bmatrix}$	Mx(K)	Mx(#)	Dp(Γ)	Mx(Γ)
☞ *i-* :: $\begin{bmatrix} K: & \kappa_{ACC}, \kappa_{DAT} \end{bmatrix}$				
ër :: $\begin{bmatrix} \Gamma: & \gamma_{ANIM} \end{bmatrix}$			*	
si- :: $\begin{bmatrix} \Gamma: & \gamma_{ANIM}, \gamma_{FEM} \\ \#: & \#_{PL} \\ K: & \kappa_{ACC} \end{bmatrix}$			**	

b. Accusative singular neuter

UR :: $\begin{bmatrix} \Gamma: & \\ \#: & \\ K: & \kappa_{ACC} \end{bmatrix}$	Mx(K)	Mx(#)	Dp(Γ)	Mx(Γ)
☞ *i-* :: $\begin{bmatrix} K: & \kappa_{ACC}, \kappa_{DAT} \end{bmatrix}$				
ër :: $\begin{bmatrix} \Gamma: & \gamma_{ANIM} \end{bmatrix}$			*	
si- :: $\begin{bmatrix} \Gamma: & \gamma_{ANIM}, \gamma_{FEM} \\ \#: & \#_{PL} \\ K: & \kappa_{ACC} \end{bmatrix}$			**	

c. Nominative plural neuter

UR :: $\begin{bmatrix} \Gamma: & \\ \#: & \#_{PL} \\ K: & \end{bmatrix}$	Mx(K)	Mx(#)	Dp(Γ)	Mx(Γ)
i- :: $\begin{bmatrix} K: & \kappa_{ACC}, \kappa_{DAT} \end{bmatrix}$		*		
ër :: $\begin{bmatrix} \Gamma: & \gamma_{ANIM} \end{bmatrix}$		*	*	
☞ *si-* :: $\begin{bmatrix} \Gamma: & \gamma_{ANIM}, \gamma_{FEM} \\ \#: & \#_{PL} \\ K: & \kappa_{ACC} \end{bmatrix}$			**	

d. Accusative plural neuter

UR :: [Γ: #: $\#_{\text{PL}}$ K: κ_{ACC}]	Mx(K)	Mx(#)	Dp(Γ)	Mx(Γ)
i- :: [K: $\kappa_{\text{ACC}}, \kappa_{\text{DAT}}$]		*		
ër :: [Γ: γ_{ANIM}]	*	*	*	
☞ *si-* :: [Γ: $\gamma_{\text{ANIM}}, \gamma_{\text{FEM}}$ #: $\#_{\text{PL}}$ K: κ_{ACC}]			**	

e. Nominative singular masculine (*ër* is the fully faithful candidate)

f. Accusative singular masculine

UR :: [Γ: γ_{ANIM} #: K: κ_{ACC}]	Mx(K)	Mx(#)	Dp(Γ)	Mx(Γ)
☞ *i-* :: [K: $\kappa_{\text{ACC}}, \kappa_{\text{DAT}}$]				*
ër :: [Γ: γ_{ANIM}]	*			
si- :: [Γ: $\gamma_{\text{ANIM}}, \gamma_{\text{FEM}}$ #: $\#_{\text{PL}}$ K: κ_{ACC}]			*	

g. Nominative plural masculine

UR :: [Γ: γ_{ANIM} #: $\#_{\text{PL}}$ K:]	Mx(K)	Mx(#)	Dp(Γ)	Mx(Γ)
i- :: [K: $\kappa_{\text{ACC}}, \kappa_{\text{DAT}}$]		*		*
ër :: [Γ: γ_{ANIM}]		*		
☞ *si-* :: [Γ: $\gamma_{\text{ANIM}}, \gamma_{\text{FEM}}$ #: $\#_{\text{PL}}$ K: κ_{ACC}]			*	

h. Accusative plural masculine

UR :: $\left[\begin{array}{ll} \Gamma: & \gamma_{\text{ANIM}} \\ \#: & \#_{\text{PL}} \\ \text{K}: & \kappa_{\text{ACC}} \end{array}\right]$	Mx(K)	Mx(#)	Dp(Γ)	Mx(Γ)
i- :: $\left[\text{K}: \ \kappa_{\text{ACC}}, \kappa_{\text{DAT}}\right]$		*		*
ër :: $\left[\Gamma: \ \gamma_{\text{ANIM}}\right]$	*	*		
☞ *si-* :: $\left[\begin{array}{ll} \Gamma: & \gamma_{\text{ANIM}}, \gamma_{\text{FEM}} \\ \#: & \#_{\text{PL}} \\ \text{K}: & \kappa_{\text{ACC}} \end{array}\right]$			*	

i. Nominative singular feminine

UR :: $\left[\begin{array}{ll} \Gamma: & \gamma_{\text{ANIM}}, \gamma_{\text{FEM}} \\ \#: & \\ \text{K}: & \end{array}\right]$	Mx(K)	Mx(#)	Dp(Γ)	Mx(Γ)
i- :: $\left[\text{K}: \ \kappa_{\text{ACC}}, \kappa_{\text{DAT}}\right]$				**
ër :: $\left[\Gamma: \ \gamma_{\text{ANIM}}\right]$				*
☞ *si-* :: $\left[\begin{array}{ll} \Gamma: & \gamma_{\text{ANIM}}, \gamma_{\text{FEM}} \\ \#: & \#_{\text{PL}} \\ \text{K}: & \kappa_{\text{ACC}} \end{array}\right]$				

j. Accusative singular feminine

UR :: $\left[\begin{array}{ll} \Gamma: & \gamma_{\text{ANIM}}, \gamma_{\text{FEM}} \\ \#: & \\ \text{K}: & \kappa_{\text{ACC}} \end{array}\right]$	Mx(K)	Mx(#)	Dp(Γ)	Mx(Γ)
i- :: $\left[\text{K}: \ \kappa_{\text{ACC}}, \kappa_{\text{DAT}}\right]$				**
ër :: $\left[\Gamma: \ \gamma_{\text{ANIM}}\right]$	*			*
☞ *si-* :: $\left[\begin{array}{ll} \Gamma: & \gamma_{\text{ANIM}}, \gamma_{\text{FEM}} \\ \#: & \#_{\text{PL}} \\ \text{K}: & \kappa_{\text{ACC}} \end{array}\right]$				

k. Nominative plural feminine

UR :: $\begin{bmatrix} \Gamma: & \gamma_{\text{ANIM}}, \gamma_{\text{FEM}} \\ \#: & \#_{\text{PL}} \\ \text{K}: & \end{bmatrix}$	MX(K)	MX(#)	DP(Γ)	MX(Γ)
i- :: $[\text{K}: \ \kappa_{\text{ACC}}, \kappa_{\text{DAT}}]$		*		**
ër :: $[\Gamma: \ \gamma_{\text{ANIM}}]$		*		*
☞ *si*- :: $\begin{bmatrix} \Gamma: & \gamma_{\text{ANIM}}, \gamma_{\text{FEM}} \\ \#: & \#_{\text{PL}} \\ \text{K}: & \kappa_{\text{ACC}} \end{bmatrix}$				

l. Accusative plural feminine (*s*- is the fully faithful candidate)

The case study in (183), I think, is an especially clear showcase for the Optimality-Theoretic logic of constraint interaction. In particular, we can see that top-ranked MAX(K) will invariably ensure the victory of exponent *i*- (the only candidate specified for $\{\kappa_{\text{DAT}}\}$) in all dative contexts; then, the immediately lower-ranked MAX(#) will ensure the victory of *s*- throughout the plural, except in precisely those dative contexts; and finally the gender-relativized constraints will restrict *s*- to the feminine whenever its plural feature isn't required. The logic of constraint interaction makes this otherwise puzzling paradigm appear completely unproblematic.

4.3 Further predictions

So far, I've been presenting the account as if it were able to generate all and only the patterns that obey the minimal-compliance *ABA generalization. However, it turns out that that's not the case: the account does derive the generalization insofar as it is indeed unable to generate any of the patterns that violate it, but it also derives something slightly stronger than that, insofar as it is also unable to generate certain patterns that the generalization would by itself permit. To roughly quantify this difference in restrictiveness, we can look once again at our simple case-number paradigm. You may remember that this six-context paradigm admits of 203 possible partitions, 110 of which violate minimal-compliance *ABA. While the current account is unable to generate any of those 110, it turns out to also be unable to generate 24 out of the remaining 93 partitions that do obey minimal-compliance *ABA.[55]

[55] These results are once again within my reach only thanks to Enrico Flor's Python implementation (accessible at stanislaozompi.net/aba).

In the following two subsections, I'm going to discuss whether and to what extent this additional restrictiveness should be regarded as a positive result. In particular, I'm going to focus on two types of patterns that turn out to be systematically impossible for the account to generate despite being consistent with minimal-compliance *ABA — what I'll refer to as the "diagonal" pattern (§4.3.1) and the "stairstep" pattern (§4.3.2).

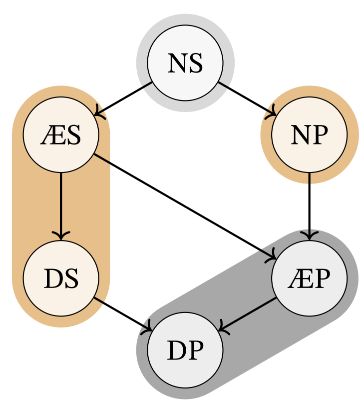
4.3.1 Further predictions (I): Impossible diagonals and the *mir*-problem

The first type of pattern that the current account is structurally unable to generate is the one I'll call the "diagonal" pattern — any partition that groups into the same cell $\text{CASE}_a.\text{NUMBER}_b$ and $\text{CASE}_x.\text{NUMBER}_y$ to the exclusion of both $\text{CASE}_a.\text{NUMBER}_y$ and $\text{CASE}_x.\text{NUMBER}_b$.

While some of these patterns are ruled out by minimal-compliance *ABA (specifically, those where the case features of CASE_a form a subset of those of CASE_x and the number features of NUMBER_b form a subset of those of NUMBER_y — e.g. NOM.SG and ACC.PL), some other such patterns are not (specifically, those where the case features of CASE_a form a subset of those of CASE_x but it's now the number features of NUMBER_y that form a subset of those of NUMBER_b — e.g. NOM.PL and ACC.SG). Among these latter diagonal patterns (those that do obey minimal-compliance *ABA), none are generable by the account, but one appears to be attested.

(184) Yiddish (Jacobs 2005: 185)

	1SG	1PL
NOM	**ix**	**m**-ir
ACC	**m**-ix	**undz**
DAT	**m**-ir	**undz**



This pattern seems to be fairly rare — possibly restricted to Yiddish and very close cognates such as Bavarian dialects. While one might identify other potential instances across Slavic (with the exception of Bulgarian), this identification appears somewhat questionable, since these languages' relevant non-nominative forms in the singular always contain an *-n-* which the NOM.PL lacks (and which the other plural forms, by contrast, do contain). See, for example, the paradigm in (185).

(185) Old Church Slavonic (Leskien 1990: 109)

	1SG	1PL
NOM	azŭ	my
ACC	mene	ny
DAT	mĭně	namŭ

Furthermore, even if Slavic were to be counted as a case in point, it would still only be the same cognate first-person paradigm (the allomorphs in play respectively going back to the same three Proto-Indo-European roots *$eǵoh_2$, **me*, and **ns*/**nes* — cf. Fortson 2010: §7.4), with no other examples in any language of the group. Were we to count cognate-form series (as in Bobaljik's (2012) seminal study), the total would still add up to just 1.

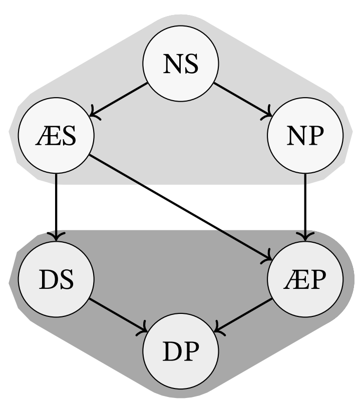
In view of the rarity of the pattern, I will take it for the time being to be the result of accidental homophony, with the *m-* in ACC/DAT.SG and the one in NOM.PL being represented as distinct exponents in the vocabulary — and I will also note in passing that, if we adopt a constraint on such homophonies along the lines of Bobaljik's (2012: 35) *Antihomophony* ("Learners avoid positing a contextual allomorph of a morpheme μ that is homophonous with the default exponent of μ"), we would still not be violating it here, as the two homophones would be specified for $\{\kappa_{\text{ACC}}\}$ and $\{\#_{\text{PL}}\}$, respectively.

4.3.2 Impossible stairsteps and the role of DOM

The second type of pattern that our current account turns out to be inherently ill-suited to generate is the one I'll call the "stairstep" pattern — any bipartition of the form $\{\{\text{NOM}_x, \text{ACC}_x, \text{NOM}_y\}, \{\text{DAT}_x, \text{ACC}_y, \text{DAT}_y\}\}$, with x and y subscripts standing for two values of some other inflectional dimension (such as number or gender) orthogonal to case.

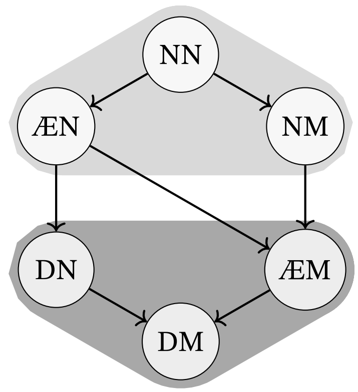
(186) "Stairstep" pattern in case–number paradigms

	SG	PL
NOM		
ACC		
DAT		



(187) "Stairstep" pattern in case–gender paradigms

	NEU	MASC/ANIM
NOM		
ACC		
DAT		



While the hostility toward such patterns might be a good thing for interactions between case and number (where stairsteps are indeed virtually unattested), it initially looks like a bad thing when it comes to interactions between case and gender. In the rest of this subsection, however, I'm going to argue that the apparent instances of stairstep patterns might not pose an insurmountable problem after all.

Let me begin with an example which seems to me particularly easy to explain away. This is found in Modern Eastern Armenian.

(188) Modern Eastern Armenian (Dum-Tragut 2009: 130–131)

	INANIMATE	ANIMATE
NOM	s-**a**	s-**a**
ACC	s-**a**	s-**ra**-n
DAT	s-**ra**-n	s-**ra**-n

At first glance, this would look exactly like an instance of the pattern in (187). What bears emphasis here, however, is the fact that Modern Eastern Armenian lacks grammatical gender altogether, but displays a language-wide pattern of Differential Object Marking (DOM) that is directly conditioned by animacy: inanimate direct objects are always exponed like nominatives, and animate direct objects always like datives. This makes it reasonable to think, then, that the accusative in this language is just never assigned to inanimate objects in the first place, and that the stairstep paradigm in (188) should be recast as the stairstep-free one in (189).

(189)

	INANIMATE	ANIMATE
NOM	s-**a**	s-**a**
ACC	—	s-**ra**-n
DAT	s-**ra**-n	s-**ra**-n

A superficially similar example is furnished by Yiddish below.

(190) Yiddish

	3N.SG	3M.SG
NOM	**ɛ**-s	**ɛ**-r
ACC	**ɛ**-s	**i**-n
DAT	**i**-m	**i**-m

However, despite the superficial similarity, independent evidence pointing to DOM is significantly scarcer in the Yiddish case: in particular, in order to mimic that analysis here, we'd have to assume that the DOM in question applies along the lines of arbitrary grammatical gender (specifically by preventing neuters alone from getting assigned accusative) — a factor rarely (if ever) reported among the ones that tend to condition DOM cross-linguistically (see e.g. Bossong 1985, among many others).

A potentially more attractive alternative, then, partly inspired by the fact that in closely related German the vowels of NOM/ACC.SG.NEUT and NOM.SG.MASC are phonologically distinct, would be to recast the distribution in (190) as in (191) — noting, for instance, that while the [ɛ] of NOM.SG.MASC survives in the closely related declension of the article (the Yiddish NOM.M.SG article is *dɛr*), the [ɛ] of NOM/ACC.SG.NEUT does not (the NOM.N.SG article being *dɔs* rather than *dɛs*!).

(191)

	3N.SG	3M.SG
NOM	**ɛ**-s	∅-ɛr
ACC	**ɛ**-s	**i**-n
DAT	**i**-m	**i**-m

Finally, the possibility of blaming DOM-like phenomena seems even less convincing for the only case to my knowledge of a potential stairstep pattern involving case and number — the North Central Westphalian paradigm in (192). In particular, in order to do the trick, DOM here would have to be active in the singular but not in the plural — an even more unusual pattern within the familiar cross-linguistic landscape of DOM.

(192) North Central Westphalian (Durrell 1990: 80)

	3F.SG	3PL
NOM	**se**	**se**
ACC	**iä**-r	**se**
DAT	**iä**-r	**iä**-r

It might then be more encouraging to note, from this perspective, that the very existence of (192) is in doubt: not only is it only reported for a handful of dialects in North Central Westphalia—but even there it's said to alternate with a DAT-vs-everything-else pattern where *se* also serves as ACC.SG.FEM (Durrell 1990: 80).

I shall therefore leave it as an open question whether stairstep patterns really exist or not, and consequently whether the theory should be enriched or amended in specific ways to make room for them.

4.3.3 Further predictions (III): Impossible coexistences of patterns

Finally, if coupled with the conjecture that we're never going to need more than one constraint ranking per language, the account also produces predictions as to which patterns may or may not coexist with which other ones within a single language. Almost all of these predictions appear to be accurate. For example, the account correctly rules out the possibility of coexistence of (193) and (194), as there doesn't exist a single constraint ranking capable of generating them both. For analogous reasons, it also rules out the coexistence of (194) and (195).

(193)

	SG	PL
NOM		
ACC		
DAT		

(194)

	SG	PL
NOM		
ACC		
DAT		

(195)

	SG	PL
NOM		
ACC		
DAT		

At the same time, the account also predicts that some nontrivial cooccurrence patterns should be possible, such as the following one from Lezgian, generable via a ranking where MAX(#) ≫ DEP(K), MAX(K) ≫ DEP(#).

(196) Lezgian (Haspelmath 1993: 80)

'water'	SG	PL
NOM	jad	jat-ar
ERG	c-i	jat-ar-i
DAT	c-i-z	jat-ar-i-z

(197) Lezgian (Haspelmath 1993: 80)

'year'	SG	PL
NOM	jis	s-ar
ERG	s-a	s-ar-i
DAT	s-a-z	s-ar-a-z

(I'm omitting the tableaux for the contexts that are perfectly matched by one of the candidates — obviously the winner. Also, since none of the VIs are specified for κ_{DAT}, every dative will pattern with the corresponding ergative.)

(198) Exponents for ‘water’: *jad-* :: $\begin{bmatrix}\#: & \#_{\text{PL}} \\ \text{K}: & \end{bmatrix}$ *c-* :: $\begin{bmatrix}\#: & \\ \text{K}: & \kappa_{\text{ERG}}\end{bmatrix}$

a.

UR :: $\begin{bmatrix}\#: & \\ \text{K}: & \end{bmatrix}$	MAX(#)	DEP(K)	MAX(K)	DEP(#)
☞ *jad-* :: $\begin{bmatrix}\#: & \#_{\text{PL}} \\ \text{K}: & \end{bmatrix}$				*
c- :: $\begin{bmatrix}\#: & \\ \text{K}: & \kappa_{\text{ERG}}\end{bmatrix}$		*		

b.

UR :: $\begin{bmatrix}\#: & \#_{\text{PL}} \\ \text{K}: & \kappa_{\text{ERG}}\end{bmatrix}$	MAX(#)	DEP(K)	MAX(K)	DEP(#)
☞ *jad-* :: $\begin{bmatrix}\#: & \#_{\text{PL}} \\ \text{K}: & \end{bmatrix}$			*	
c- :: $\begin{bmatrix}\#: & \\ \text{K}: & \kappa_{\text{ERG}}\end{bmatrix}$	*			

(199) Exponents for ‘year’: *jis* :: $\begin{bmatrix}\#: & \\ \text{K}: & \end{bmatrix}$ *s-* :: $\begin{bmatrix}\#: & \#_{\text{PL}} \\ \text{K}: & \kappa_{\text{ERG}}\end{bmatrix}$

a.

UR :: $\begin{bmatrix}\#: & \\ \text{K}: & \kappa_{\text{ERG}}\end{bmatrix}$	MAX(#)	DEP(K)	MAX(K)	DEP(#)
jis :: $\begin{bmatrix}\#: & \\ \text{K}: & \end{bmatrix}$			*	
☞ *s-* :: $\begin{bmatrix}\#: & \#_{\text{PL}} \\ \text{K}: & \kappa_{\text{ERG}}\end{bmatrix}$				*

b.

UR :: $\begin{bmatrix}\#: & \#_{\text{PL}} \\ \text{K}: & \end{bmatrix}$	MAX(#)	DEP(K)	MAX(K)	DEP(#)
jis :: $\begin{bmatrix}\#: & \\ \text{K}: & \end{bmatrix}$	*			
☞ *s-* :: $\begin{bmatrix}\#: & \#_{\text{PL}} \\ \text{K}: & \kappa_{\text{ERG}}\end{bmatrix}$		*		

There is one wrinkle, though. The two patterns in (200) and (201) are also predicted to be unable to cooccur (the former requires high Dep, while the latter requires high Max), and yet they do cooccur in Icelandic.[56]

(200) Icelandic (Einarsson 1949: 68)

	2sg	2pl
nom	**þ**-ú	**þ**-ið
acc	**þ**-ig	**ykk**-ur
dat	**þ**-ér	**ykk**-ur

(201) Icelandic 'the/this/that' (Einarsson 1949: 70)

	sg	pl
nom	**s**-á	**þ**-eir
acc	**þ**-ann	**þ**-á
dat	**þ**-eim	**þ**-eim

This might either mean that the restrictive single-ranking assumption will have to be abandoned to make provisions for different rankings in different paradigms (what we may refer to as *co-morphologies*, after Orgun's 1996 "co-phonologies") or that maybe the plural of participant pronouns — sometimes called an *associative* plural (cf. Corbett 2000: 83, Ackema & Neeleman 2018: 5ff and references therein) — is different enough from the plural of nonparticipants to deserve its own dedicated set of Max and Dep constraints, ranked independently from Max(#) and Dep(#).[57] Here I'm going to tentatively lean toward this latter alternative, which has the merit of generating a novel testable prediction: such discrepancies as the one between (200) and (201) should only be able to arise between a participant paradigm and a non-participant one.

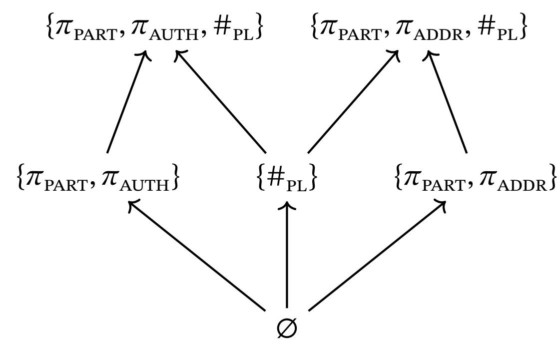
4.4 Applying the proposal to verbal paradigms

Given the feature structure in (202), we can easily extend the previous section's account to agreement-conditioned stem allomorphy in verbs.

[56] Note that the only other varieties with the pattern in (200), namely Thuringian dialects (Spangenberg 1990: 281), lack any example of the pattern in (201).

[57] The idea of a special participant-specific associative plural will come in handy again shortly in §4.4.1, for independent reasons.

(202) The feature structure of person–number agreement (to be revised in (208))



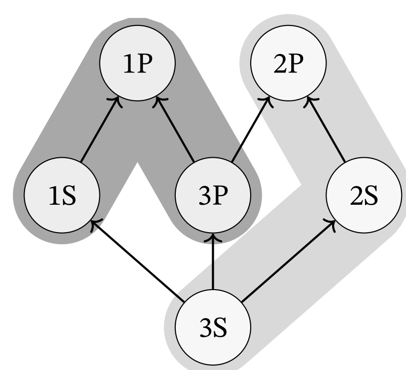
In this case, the relevant violable constraints will of course include the familiar MAX(#) and DEP(#) as well as the newly introduced MAX(Π) and DEP(Π). Furthermore, as per above, I'm going to maintain that available candidates are subject to inviolable entailment relations between features (in this case, no π_{AUTH} or π_{ADDR} in the absence of π_{PART}). More specifically, for the sake of simplicity and restrictiveness, I'm going to assume that the relevant context specifications that exponents can bear are all and only the relevant contexts that the syntax can generate and provide as an input — i.e. just the six feature sets represented in the nodes of the diagram in (202).

The system set up in this way can once again be proven by exhaustion to be incapable of generating any pattern that would violate minimal-compliance *ABA.[58] At the same time, it can also be shown to capture traditionally "morphomic" patterns that obey minimal-compliance *ABA but violate the more restrictive Pāṇinian *ABA — such as, for example, the Latin pattern in (144), repeated here in (203).

(203) Latin 'can'

3SG 3PL 2PL (but 3SG 2SG 2PL!)

	SG	PL
1	possum	possumus
2	potes	potestis
3	potest	possunt



This pattern can straightforwardly be captured by means of the vocabulary in (204) — for instance under the partial ranking I exemplify in (205).

[58] This result, oto, is made possible by the Python implementation that Enrico Flor has kindly provided me with (accessible at stanislaozompi.net/aba).

(204) $\textit{pot-} :: \begin{bmatrix}\#: & \#_{\text{PL}} \\ \Pi: & \pi_{\text{PART}}, \pi_{\text{AUTH}}\end{bmatrix}$ $\quad$ $\textit{poss-} :: \begin{bmatrix}\#: & \\ \Pi: & \pi_{\text{PART}}, \pi_{\text{ADDR}}\end{bmatrix}$

(205) a. First singular

UR :: $\begin{bmatrix}\#: & \\ \Pi: & \pi_{\text{PART}}, \pi_{\text{AUTH}}\end{bmatrix}$	Mx(Π)	Dp(Π)	Mx(#)	Dp(#)
pot- :: $\begin{bmatrix}\#: & \\ \Pi: & \pi_{\text{PART}}, \pi_{\text{ADDR}}\end{bmatrix}$	*	*		
☞ *poss-* :: $\begin{bmatrix}\#: & \#_{\text{PL}} \\ \Pi: & \pi_{\text{PART}}, \pi_{\text{AUTH}}\end{bmatrix}$				*

b. Second singular (*pot-* is the fully faithful candidate)

c. Third singular

UR :: $\begin{bmatrix}\#: & \\ \Pi: & \end{bmatrix}$	Mx(Π)	Dp(Π)	Mx(#)	Dp(#)
☞ *pot-* :: $\begin{bmatrix}\#: & \\ \Pi: & \pi_{\text{PART}}, \pi_{\text{ADDR}}\end{bmatrix}$		**		
poss- :: $\begin{bmatrix}\#: & \#_{\text{PL}} \\ \Pi: & \pi_{\text{PART}}, \pi_{\text{AUTH}}\end{bmatrix}$		**		*

d. First plural (*poss-* is the fully faithful candidate)

e. Second plural

UR :: $\begin{bmatrix}\#: & \#_{\text{PL}} \\ \Pi: & \pi_{\text{PART}}, \pi_{\text{ADDR}}\end{bmatrix}$	Mx(Π)	Dp(Π)	Mx(#)	Dp(#)
☞ *pot-* :: $\begin{bmatrix}\#: & \\ \Pi: & \pi_{\text{PART}}, \pi_{\text{ADDR}}\end{bmatrix}$			*	
poss- :: $\begin{bmatrix}\#: & \#_{\text{PL}} \\ \Pi: & \pi_{\text{PART}}, \pi_{\text{AUTH}}\end{bmatrix}$	*	*		

f. Third plural

UR :: [#: $\#_{\text{PL}}$; Π:]	Mx(Π)	Dp(Π)	Mx(#)	Dp(#)
pot- :: [#: ; Π: $\pi_{\text{PART}}, \pi_{\text{ADDR}}$]		**	*	
☞ *poss-* :: [#: $\#_{\text{PL}}$; Π: $\pi_{\text{PART}}, \pi_{\text{AUTH}}$]		**		

(Another suitable partial ranking would be DEP(Π ≫ MAX(Π), MAX(#), DEP(#), as the reader can easily verify.)

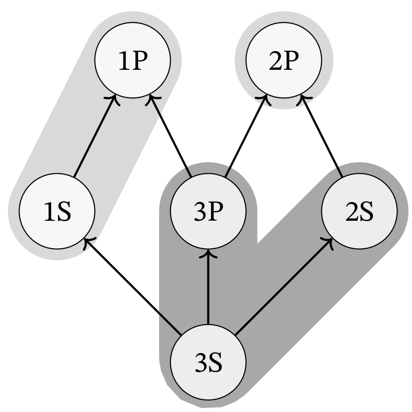
The only *prima-facie* undergeneration problems that the system faces come from the paradigms from Platta (139), Catanzaro (140), and Bergamo(141) — but I believe there are convincing fixes for each of them.[59] I turn to them in the next two subsections.

4.4.1 The need for an associative plural

Let me begin with the two patterns from Platta and Catanzaro, which are essentially mirror images of each other.

(206) Platta 'go' (Maiden 2018: 106)

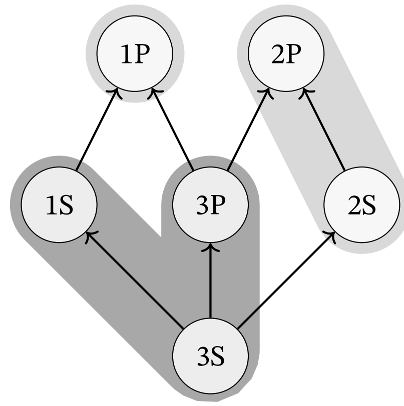
	SG	PL
1	mɔn	məin
2	vas	məis
3	va	van



[59] The reader may suspect that these undergeneration issues might be due to the current restrictive assumptions about possible exponent specifications — but that is demonstrably not the case. In particular, even if we were to allow exponents to be specified for π_{PART} alone without being further specified for either π_{AUTH} or π_{ADDR}, or if we were to let them be overspecified for both π_{AUTH} and π_{ADDR} at once, the three patterns at issue would still remain impossible to generate.

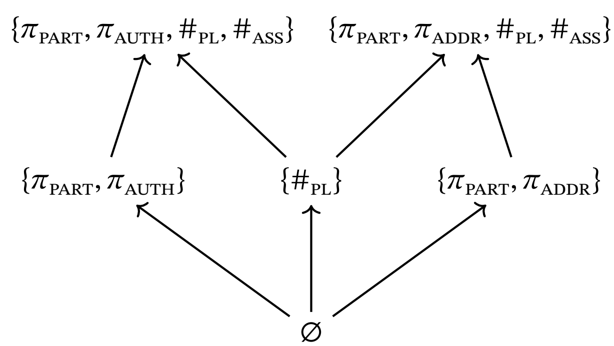
(207) Catanzaro 'give' preterit (other variant) (Maiden 2018: 290)

	SG	PL
1	dɛtsi	dunammi
2	dunasti	dunástivu
3	dɛtsa	dɛ́tsaru



This undergeneration problem can most simply be solved by following a large portion of the literature in considering the plural of 1PL and 2PL to be of an essentially different kind than the one of 3PL — specifically, an *associative* plural (see Corbett 2000: 83, Ackema & Neeleman 2018, and references therein). If we revise the feature structure in (202) along the lines of (208) — and we assume once again that candidates are inviolably filtered by entailments between features, so that no candidate may have $\#_{\text{ASS}}$ without $\#_{\text{PL}}$ — then we can still derive all the patterns from §3.3.2 and none of the minimal-compliance–*ABA violations, but we can now also finally derive the patterns in (206)–(207).

(208) The feature structure of person–number agreement (revised version)



For example, we may derive the pattern in (206) as follows:

(209) a. First singular

UR :: $\begin{bmatrix} \#: & \\ \Pi: & \pi_{\text{PART}}, \pi_{\text{AUTH}} \end{bmatrix}$	Mx(#)	Dp(Π)	Mx(Π)	Dp(#)
v- :: $\begin{bmatrix} \#: & \#_{\text{PL}} \\ \Pi: & \end{bmatrix}$			**	*
☞ *m*- :: $\begin{bmatrix} \#: & \#_{\text{PL}}, \#_{\text{ASS}} \\ \Pi: & \pi_{\text{PART}}, \pi_{\text{AUTH}} \end{bmatrix}$				**

b. Second singular

UR :: $\begin{bmatrix} \#: & \\ \Pi: & \pi_{\text{PART}}, \pi_{\text{ADDR}} \end{bmatrix}$	Mx(#)	Dp(Π)	Mx(Π)	Dp(#)
☞ *v*- :: $\begin{bmatrix} \#: & \#_{\text{PL}} \\ \Pi: & \end{bmatrix}$			**	*
m- :: $\begin{bmatrix} \#: & \#_{\text{PL}}, \#_{\text{ASS}} \\ \Pi: & \pi_{\text{PART}}, \pi_{\text{AUTH}} \end{bmatrix}$		*	*	**

c. Third singular

UR :: $\begin{bmatrix} \#: & \\ \Pi: & \end{bmatrix}$	Mx(#)	Dp(Π)	Mx(Π)	Dp(#)
☞ *v*- :: $\begin{bmatrix} \#: & \#_{\text{PL}} \\ \Pi: & \end{bmatrix}$				*
m- :: $\begin{bmatrix} \#: & \#_{\text{PL}}, \#_{\text{ASS}} \\ \Pi: & \pi_{\text{PART}}, \pi_{\text{AUTH}} \end{bmatrix}$		**		**

d. First plural (*m*- is the fully faithful candidate)

e. Second plural

UR :: $\begin{bmatrix} \#: & \#_{\text{PL}}, \#_{\text{ASS}} \\ \Pi: & \pi_{\text{PART}}, \pi_{\text{ADDR}} \end{bmatrix}$	Mx(#)	Dp(Π)	Mx(Π)	Dp(#)
v- :: $\begin{bmatrix} \#: & \#_{\text{PL}} \\ \Pi: & \end{bmatrix}$	*		**	
☞ *m*- :: $\begin{bmatrix} \#: & \#_{\text{PL}}, \#_{\text{ASS}} \\ \Pi: & \pi_{\text{PART}}, \pi_{\text{AUTH}} \end{bmatrix}$		*	*	

f. Third plural (*v*- is the fully faithful candidate)

The pattern in (207) will be obtained in exactly the same way, except for π_{ADDR} replacing π_{AUTH} in the light-grey allomorph's context specification. (The symmetry in status between π_{ADDR} and π_{AUTH} should make this fairly easy to see even without tableaux.)

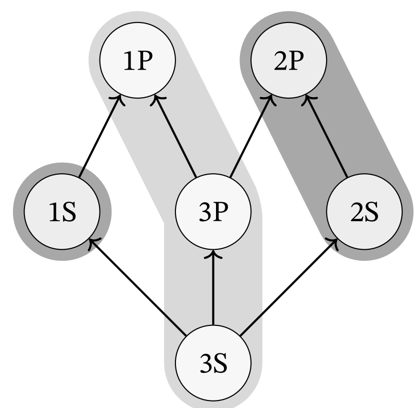
(210) *dɛt-* :: $\begin{bmatrix} \#: & \#_{\text{PL}} \\ \Pi: & \end{bmatrix}$ *duna-* :: $\begin{bmatrix} \#: & \#_{\text{PL}}, \#_{\text{ASS}} \\ \Pi: & \pi_{\text{PART}}, \pi_{\text{ADDR}} \end{bmatrix}$

4.4.2 "Deep" syncretisms again

Finally the last apparent undergeneration problem is the one posed by the Bergamasque paradigm in (211).

(211) Bergamasque 'go' (Zanetti 2005: 172)

	SG	PL
1	ndo	va
2	nde	andí
3	va	va



I believe this is, in a way, the verbal counterpart of the Modern Eastern Armenian DOM pattern in (188)–(189)— both problems being the result of an incorrect interpretation of the categories involved in the paradigm. Just as what we called "accusative" for Armenian inanimates turned out, on closer inspection, to most likely just be a nominative, so the first plural of Bergamasque turns out, on closer inspection, to most likely be an impersonal with default 3SG agreement. The metasyncretism between 1PL and 3SG is exceptionless,[60] and even reflexive clitics lack a special 1PL form. I therefore conclude that the paradigm should be recast as the unproblematic (212).

(212)

	SG	PL
1	ndo	—
2	nde	andí
3	va	va

[60] The only apparent exception involves the copula (3SG *ɛ* ∼ 1PL *s-ɛ* in the present; 3SG *ía/éra* ∼ 1PL *s-ía/s-éra* in the imperfect; Zanetti 2005: 164), but this too might be analyzed as syncretism in the verb form (*ɛ/ía/éra*) with lexically triggered allomorphy on the subject clitic.

4.4.3 Morphomes dissolved?

Before closing this chapter, I'd like to go back to an observation I made in §3.3.3, to the effect that the variety of patterns we see in the domain of verbal φ-agreement and mood is in fact parallel to what we observe in the domain of case and number, and that therefore we're owed a unified solution to both sets of issues. In that spirit, I'd like to emphasize that the current MAX/DEP-based account straightforwardly captures all three of the patterns in (153)–(155) (repeated here as (213)–(215)) if we assume the feature structure in (216) (with the indicative's mood features forming a subset of the subjunctive's) and of course the mood-specific constraints MAX(M) and DEP(M).

(213) Italian 'go'

	INDICATIVE		SUBJUNCTIVE	
	SG	PL	SG	PL
1	vado	andiamo	vada	andiamo
2	vai	andate	vada	andiate
3	va	vanno	vada	vadano

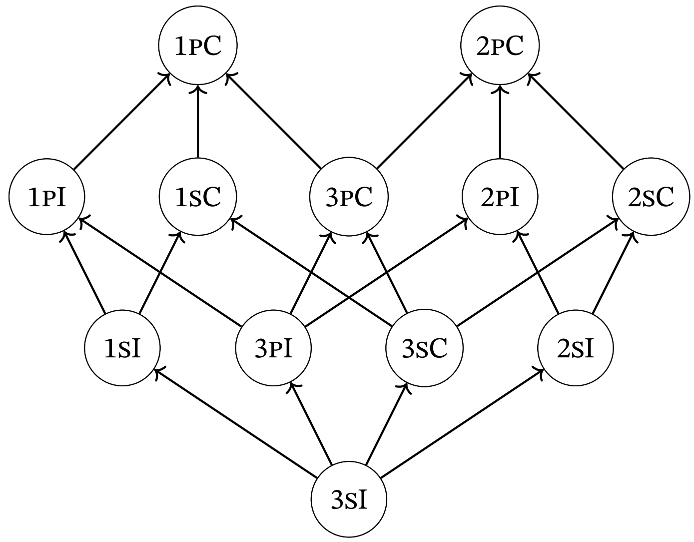
(214) Cascinagrossa 'go' (LINK)

	INDICATIVE		SUBJUNCTIVE	
	SG	PL	SG	PL
1	vag	andómma	vaga	vagen
2	vɛ	andɛ́	vag	vagi
3	va	vaŋn	vaga	vagen

(215) Languedocien 'go' (Maiden 2018: 195)

	INDICATIVE		SUBJUNCTIVE	
	SG	PL	SG	PL
1	vau	anam	ane	anem
2	vas/vai	anatz	anes	anetz
3	va	van	ane	anen

(216) Feature structure for verbs (including mood)



In Italian, *and-* is specified for person and number features, whereas *va-* isn't specified for anything. High-ranked DEP(Π) and DEP(#) restrict *and-* to PART.PL contexts—and, since neither allomorph is specified for mood, mood never makes a difference. In Languedocien, *an-* is specified for person, number, and mood, whereas *v-* is again specified for nothing. A top-ranked MAX(M) ensures that *an-* will win in all the subjunctive, but, in the indicative, next-ranked DEP(Π) and DEP(#) will restrict it to PART.PL contexts alone. Finally, the dialect from Cascinagrossa has the same ranking as Languedocien (MAX(M) ≫ DEP(Π), DEP(#) ≫ …), except that *φ*- and mood-specifications are divided up between allomorphs: *and-* is specified for person and number, but it's *v-* that is specified for mood.

At the same time, the account also predicts the impossibility of mix-and-match between different non-leveling patterns across indicative and subjunctive. Roughly put, this is because a given exponent's *φ*-specifications will either play out the same way across different moods or be completely overruled in a certain mood due to higher-ranked mood-relativized constraints, with no room left for richer interactions between the two.[61] We therefore predict the impossibility of patterns like, for example, PART.PL *vs* everything else in the indicative but 1PL *vs* everything else in the subjunctive, or 2PL *vs* everything else in the indicative but PL *vs* SG in the subjunctive—a prediction that converges with the results recently reported from Romance by Kühlert (2023: Ch. 4).[62]

[61] Such interactions would constitute the verbal equivalent of the stairstep pattern we discussed in §4.3.2.

[62] Kühlert (2023: Ch. 4) also shows that there's often φ-sensitive allomorphy in the indicative accompanied by leveling throughout the subjunctive, but never φ-sensitive allomorphy in the subjunctive accompanied by leveling throughout the indicative—a gap that does not follow from my current account and which I'll have to leave outside of its purview.

(217) An invented morphome (Maiden 2018: 12)

	INDICATIVE			SUBJUNCTIVE		
	PAST	PRES	FUT	PAST	PRES	FUT
1SG	karep-i	karep-j	perak-l	karep-ɲ	karep-ɔ	karep-m
2SG	karep-k	karep-p	karep-l	karep-θ	perak	karep-l
3SG	perak-ɲ	karep-s	karep-m	karep-ɲ	karep-a	karep-h
1PL	karep-i	perak-θ	karep-t	karep-v	karep-u	karep-a
2PL	karep-d	karep-v	karep	karep-v	karep-f	perak-a
3PL	karep-ɲ	karep-y	karep-t	perak-v	karep-u	karep-i

A fortiori, I can also finally deliver on the promise I initially made in Chapter 1: while the attested "morphomes" such as (215) are straightforwardly derived, Maiden's (2018) monstrous morphome, repeated in (217), remains well beyond our generative reach.

Chapter 5

Conclusions

In this thesis, I've argued that an approach to exponent selection in terms of underspecification and Pāṇinian ordering systematically runs into problems when it comes to capturing the distribution of contextual allomorphs in multidimensional paradigms. I've then proposed to solve those empirical problems by replacing underspecification with a well defined mix of under- and overspecification — a system in which allomorphs can both be underspecified and be overspecified with respect to the contexts they're inserted into, with each of these departures from the perfect match being penalized but not necessarily fatal. More concretely, I've implemented this idea Optimality-Theoretically in terms of a strict-domination ranking of violable Max and Dep constraints, each evaluating the exponents' faithfulness to the given morphosyntactic context with respect to a single inflectional dimension (grammatical case, number, gender, etc.).

Here, before closing, I'd like to briefly flag certain architectural assumptions that underpin the account, some potential complications I've abstracted away from in setting it up, and certain avenues for future research that these reflections may open up.

First, I'd like to insist once more on the architectural conservatism of the proposal. Just as in many previous, non-Optimality-Theoretic approaches, I too am assuming a mapping between exponence loci on the one hand and exponents with context specifications on the other; just as in most previous approaches, I'm also assuming that this mapping is adjudicated via a competition; and, just as in those approaches, I'm assuming that the competing exponents and their context specifications have to be learned as part of the language's vocabulary. The only respect in which I'm minimally differing has to do with the metric that adjudicates the competition. As discussed especially by Wolf (2008: 68ff), rather than applying Pāṇinian ordering only "from below" by regarding overspecification as a disqualifying factor (Distributed Morphology) or than applying Pāṇinian ordering only "from above" by regarding underspecification as a disqualifying factor (Nanosyntax), the current system unifies (violable-constraint versions of) both principles into a single coherent metric. However, in contrast to Wolf (2008), Rolle (2018), and others, it bears emphasizing that the OT metric I'm proposing here crucially involves only such morphosyntactic faithfulness constraints, and nothing else — particularly no markedness, either morphological, phonological, or otherwise. The tableaux I've been drawing out thus simply represent attempts to find the most featurally faithful exponent for a context, given the relevant language's available vocabulary resources — with the constraint-interaction logic of OT serving the sole purpose of mediating between potentially conflicting faithfulness demands.

Next, I'd also like to call attention to the extreme simplicity of the proposed system — a trait I view both as a virtue and, in some senses, as a limitation. On the one hand, the very low number of moving parts I've allowed for has made it easy to derive results and test out specific complex empirical cases, and it has also allowed me to offer a proof of principle that even such a lean and restrictive approach can cover quite a bit of empirical ground. On the other hand, this has also led me to systematically abstract away from potential theoretical complications that are quite possibly directly relevant — most especially the possibility of pre-exponence feature manipulations (such as Impoverishment or Rules of Referral) implicated in metasyncretism and potentially elsewhere. Looking into the future of this project, I see the investigation of these additional moving parts as one of the most natural next steps: are certain unexpected patterns (such as the sporadical "stairsteps" discussed in §4.3.2) systematically showing up only together with metasyncretisms? To the extent that this can be adduced as evidence for pre-exponence feature manipulations, what are the constraints that we'll need to posit on such manipulations in order to prevent them from reintroducing ABA through the backdoor (cf. Caha 2018; Zompì 2019)? Could their application, too, be constrained in an Optimality-Theoretic fashion?

Finally, the other most natural extensions of this project are simply empirical. In particular, in this thesis I've restricted myself to the phenomenon of contextual allomorphy, and more specifically on how it plays out in two categorial domains (case/number/gender-conditioned stem suppletion in pronouns and nouns, and φ-agreement–conditioned stem suppletion in verbs), but case studies could easily be multiplied in both directions. On the one hand, in Zompì (2019) I argued that case-affixal syncretism and case-conditioned stem suppletion systematically obey the same unidimensional *ABA patterns. To the extent that the two phenomena's patterning alike is consistent and principled, affixal syncretism should thus serve as an ideal follow-up to this study. On the other hand, recent research has uncovered evidence for cumulative decompositions in an ever growing number of categorial domains. The moment any two such categories interact in jointly conditioning allomorphy or in undergoing fusional affixal exponence together, they become potential case studies on which to test the ideas proposed in this dissertation.

Bibliography

Ackema, Peter & Ad Neeleman (2005). “Word-formation in Optimality Theory.” In Pavol Štekauer & Rochelle Lieber (eds.), *Handbook of word-formation*, pp. 285–313. Dordrecht: Springer.

——— (2018). *Features of person: From the inventory of persons to their morphological realization*. Cambridge, MA: MIT Press.

Alekseev, Mikhail E. & Sabrina Kh. Šikhaleva (2003). *Табасаранский язык*. Moscow: Academia.

Aronoff, Mark (1994). *Morphology by itself: Stems and inflectional classes*. Cambridge, MA: MIT Press.

Babaliyeva, Ayten (2013). *Études sur la morphosyntaxe du tabasaran littéraire*. Doctoral dissertation. Paris: École Pratique des Hautes Études.

Barnes, Michael (2008). *A new introduction to Old Norse – Part I: Grammar*. London: Viking Society for Northern Research.

Baunaz, Lena & Eric Lander (2018). “Nanosyntax; The basics.” In Lena Baunaz, Karen De Clercq, Liliane Haegeman & Eric Lander (eds.), *Exploring Nanosyntax*, pp. 3–56. Oxford: Oxford University Press.

Béjar, Susana & Milan Rezac (2009). “Cyclic Agree.” *Linguistic Inquiry* 40: 35–73.

Benincà, Paola, Mair Parry & Diego Pescarini (2016). “The dialects of northern Italy.” In Adam Ledgeway & Martin Maiden (eds.), *The Oxford guide to the Romance languages*, pp. 185–205. Oxford: Oxford University Press.

van den Berg, Helma (1995). *A grammar of Hunzib (with texts and lexicon)*. Munich: LINCOM.

Bergsma, Fenna (2019). “Mismatches in free relatives – grafting nanosyntactic trees.” *Glossa: A Journal of General Linguistics* 4: 119.

Bobaljik, Jonathan David (2008). “Where’s Phi? Agreement as a postsyntactic operation.” In Daniel Harbour, David Adger & Susana Béjar (eds.), *Phi theory: Phi-features across modules and interfaces*, pp. 295–328. Oxford: Oxford University Press.

——— (2012). *Universals in comparative morphology: Suppletion, superlatives, and the structure of words*. Cambridge, MA: MIT Press.

Bobaljik, Jonathan David & Cynthia Levart Zocca (2011). "Gender markedness: The anatomy of a counter-example." *Morphology* 21: 141–166.
Bond, Oliver & Marina Chumakina (2016). "Agreement domains and targets." In Oliver Bond, Greville G. Corbett, Marina Chumakina & Dunstan Brown (eds.), *Archi: Complexities of agreement in cross-theoretical perspective*, pp. 43–76. Oxford: Oxford University Press.
Bonet, Eulàlia (1994). "The Person-Case Constraint: A morphological approach." In Heidi Harley & Colin Phillips (eds.), *MIT Working Papers in Linguistics 22: The morphology-syntax connection*, pp. 33–52. Cambridge, MA: MITWPL.
Bossong, Georg (1985). *Empirische Universalienforschung. Differenzielle Objektmarkierung in der neuiranischen Sprachen*. Tübingen: Narr.
Braune, Wilhelm (2004). *Althochdeutsche Grammatik I: Laut- und Formenlehre*. 15th edn., revised by Ingo Reiffenstein. Tübingen: Niemeyer.
Braune, Wilhelm & Frank Heidermanns (2004). *Gotische Grammatik mit Lesestücken und Wörterverzeichnis*. 20th edn. Tübingen: Niemeyer.
Bybee, Joan (1985). *Morphology: A study in the relation between meaning and form*. Amsterdam: John Benjamins.
Caha, Pavel (2009). *The nanosyntax of case*. Doctoral dissertation. Tromsø: University of Tromsø.
——— (2013). "Explaining the structure of case paradigms by the mechanisms of Nanosyntax: The Classical Armenian nominal declension." *Natural Language and Linguistic Theory* 31: 1015–1066.
——— (2018). "Notes on insertion in Distributed Morphology and Nanosyntax." In Lena Baunaz, Karen De Clercq, Liliane Haegeman & Eric Lander (eds.), *Exploring Nanosyntax*, pp. 57–87. Oxford: Oxford University Press.
——— (2019). "Contextual allomorphy vs. portmanteaus: The case of suppletive nominatives." Talk given at BCGL 12. Brussels. URL: https://www.crissp.be/wp-content/uploads/2019/10/BCGL_12_paper_4.pdf.
——— (2023). "Root and stem allomorphy without multiple exponence: The case of special nominatives." Ms. Masaryk University, Brno.
Caha, Pavel, Karen De Clercq & Guido Vanden Wyngaerd (2019). "The fine structure of the comparative." *Studia Linguistica* 73: 470–521.
Calabrese, Andrea (2005). *Markedness and economy in a derivational model of phonology*. Berlin: Mouton de Gruyter.
Christopoulos, Christos & Stanislao Zompì (2022). "Taking the nominative (back) out of the accusative: Case features and the distribution of stems in Indo-European paradigms." *Natural Language and Linguistic Theory*.
Corbett, Greville G. (1991). *Gender*. Cambridge: Cambridge University Press.

Corbett, Greville G. (2000). *Number*. Cambridge: Cambridge University Press.

Daskalaki, Evangelia (2009). *(Mis)matching patterns in Greek free relatives*. Doctoral dissertation. Cambridge: University of Cambridge.

Dimmendaal, Gerrit J. (1983). *The Turkana language*. Dordrecht: Foris.

Donaldson, Bruce C. (1993). *A grammar of Afrikaans*. Berlin: Mouton de Gruyter.

Dum-Tragut, Jasmine (2009). *Armenian: Modern Eastern Armenian*. Amsterdam: John Benjamins.

Durrell, Martin (1990). "Westphalian and Eastphalian." In Charles V.J. Russ (ed.), *The dialects of modern German: A linguistic survey*, pp. 59–90. London: Routledge.

Einarsson, Stefán (1949). *Icelandic: Grammar, texts, glossary*. Baltimore, MD: Johns Hopkins University Press.

Ernout, Alfred (1953). *Morphologie historique du latin*. 3rd edn. Paris: Klincksieck.

Festa, Giovanni Battista (1917). "Il dialetto di Matera." *Zeitschrift für romanische Philologie* 38: 129–162.

Flor, Enrico (2023). "Base-derivative faithfulness in Italian palatalization." Ms. MIT.

Forker, Diana (2020). "Avar." In Maria Polinsky (ed.), *The Oxford handbook of languages of the Caucasus*, pp. 243–279. Oxford: Oxford University Press.

Fortson, Benjamin W., IV (2010). *Indo-European language and culture: An introduction*. 2nd edn. Oxford: Blackwell.

Friedman, Victor A. (2020). "Lak." In Maria Polinsky (ed.), *The Oxford handbook of languages of the Caucasus*, pp. 201–241. Oxford: Oxford University Press.

Graf, Thomas (2019). "Monotonicity as an effective theory of morphosyntactic variation." *Journal of Language Modelling* 7: 3–47.

Green, W.A.I. (1990). "The dialects of the Palatinate (*Das Pfälzische*)." In Charles V.J. Russ (ed.), *The dialects of modern German: A linguistic survey*, pp. 241–264. London: Routledge.

Grishin, Peter (2023). "Omnivorous third person agreement in Algonquian." *Glossa: A Journal of General Linguistics* 8.

Halle, Morris (1997). "Distributed Morphology: Impoverishment and Fission." In Benjamin Bruening, Yoonjung Kang & Martha McGinnis (eds.), *MIT Working Papers in Linguistics 30: PF: Papers at the interface*, pp. 425–449. Cambridge, MA: MIT.

Harbert, Wayne (1978). *Gothic syntax: A relational grammar*. Doctoral dissertation. Urbana-Champaign, IL: University of Illinois.

——— (1982). "On the nature of the matching parameter." *The Linguistic Review* 2: 237–284.

——— (1992). "Gothic relative clauses and syntactic theory." In Irmengard Rauch, Gerald F. Carr & Robert L. Kyes (eds.), *On Germanic linguistics: Issues and methods*, pp. 109–146. Berlin: Mouton de Gruyter.

Harbour, Daniel (2014). “Paucity, abundance, and the theory of number.” *Language* 90: 185–229.

——— (2016). *Impossible persons*. Cambridge, MA: MIT Press.

Harley, Heidi & Elizabeth Ritter (2002). “Person and number in pronouns: A feature-geometric analysis.” *Language* 78: 482–526.

Haspelmath, Martin (1993). *A grammar of Lezgian*. Berlin: Mouton de Gruyter.

Hogg, Richard M. & Robert D. Fulk (1992). *A grammar of Old English*. Vol. 2: *Morphology*. Oxford: Blackwell.

Holthausen, Ferdinand (1921). *Altsächsisches Elementarbuch*. Heidelberg: Winter.

Holton, David, Peter Mackridge, Irene Philippaki-Warburton & Vassilios Spyropoulos (2012). *Greek: A comprehensive grammar of the modern language*. London: Routledge.

Hualde, José Ignacio (2003). “Case and number inflection of noun phrases.” In José Ignacio Hualde & Jon Ortiz de Urbina (eds.), *A grammar of Basque*, pp. 171–187. Berlin: Mouton de Gruyter.

Ibragimov, Garun Kh. (1978). *Рутульский язык*. Moscow: Nauka.

Imnajšvili, David S. (1963). *Дидойский язык в сравнении с гинухским и хваршийским языками*. Tbilisi: Izdatel’stvo Akademii nauk Gruzinskoj SSR.

Jacobs, Neil G. (2005). *Yiddish: A linguistic introduction*. Cambridge: Cambridge University Press.

Khvtisiashvili, Tamrika (2013). *Principal aspects of Xinaliq phonology and morphosyntax*. Doctoral dissertation. Salt Lake City: University of Utah.

Kiparsky, Paul (1982). “Lexical morphology and phonology.” In In-Seok Yang (ed.), *Linguistics in the morning calm*, pp. 3–91. Seoul: Hanshin Publishing Company.

——— (2005). “Blocking and periphrasis in inflectional paradigms.” In Geert Booij & Jaap van Marle (eds.), *Yearbook of Morphology 2004*, pp. 113–136. Dordrecht: Springer.

Kleine, Ane (2003). “Standard Yiddish.” *Journal of the International Phonetic Association* 33: 261–265.

Köpcke, Klaus-Michael & David A. Zubin (1984). “Sechs Prinzipien für die Genuszuweisung im Deutschen: Ein Beitrag zur natürlichen Klassifikation.” *Linguistische Berichte* 93: 26–50.

Kramer, Ruth (2015). *The morphosyntax of gender*. Oxford: Oxford University Press.

Krause, Todd B. & Jonathan Slocum (2007–10). “Tocharian online.” URL: https://lrc.la.utexas.edu/eieol/tokol.

Kühlert, Niels Torben (2023). *A morphological typology of non-root alternations: Invasion, suppletion, and allomorphy*. Doctoral dissertation. Cambridge, MA: Harvard University.

Leskien, August (1990). *Handbuch der altbulgarischen (altkirchenslavischen) Sprache: Grammatik, Texte, Glossar*. 10th edn., revised by Johannes Schröpfer. Heidelberg: Winter.

Maiden, Martin (2004). "When lexemes become allomorphs: On the genesis of suppletion." *Folia Linguistica* 38: 227–256.

——— (2018). *The Romance verb: Morphomic structure and diachrony*. Oxford: Oxford University Press.

Maiden, Martin, John Charles Smith, Silvio Cruschina, Marc-Olivier Hinzelin & Maria Goldbach (2010). "Oxford online database of Romance verb morphology." http://romverbmorph.clp.ox.ac.uk/.

Marantz, Alec (1991). "Case and licensing." In German F. Westphal, Benjamin Ao & Hee-Rahk Chae (eds.), *Proceedings of ESCOL 1991*, pp. 234–253. Ithaca, NY: Cornell Linguistics Circle.

Matras, Yaron & Gertrud Reershemius (2003). *Low German (East Frisian dialect)*. Munich: LINCOM.

Mayrhofer, Manfred (1978). *Sanskrit-Grammatik mit sprachvergleichenden Erläuterungen*. Berlin: Mouton de Gruyter.

McCarthy, John & Alan Prince (1995). "Faithfulness and reduplicative identity." In Jill Beckman, Suzanne Urbanczyk & Laura Walsh Dickey (eds.), *University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory*, pp. 249–384. Amherst, MA: University of Massachusetts, Graduate Linguistic Student Association.

McFadden, Thomas (2018). "*ABA in stem-allomorphy and the emptiness of the nominative." *Glossa: A Journal of General Linguistics* 3 (8).

Merlan, Francesca C. (1982). *Mangarayi*. Amsterdam: North-Holland.

——— (1994). *A grammar of Wardaman: A language of the Northern Territory of Australia*. Berlin: Mouton de Gruyter.

Moradi, Sedigheh (2021). *Monotonicity in morphosyntax*. Doctoral dissertation. Stony Brook, NY: Stony Brook University.

Moravcsik, Edith A. (1978). "Agreement." In Joseph H. Greenberg (ed.), *Universals of human language*. Vol. 4: *Syntax*, pp. 331–374. Stanford, CA: Stanford University Press.

Moskal, Beata (2015). "Limits on allomorphy: A case study in nominal suppletion." *Linguistic Inquiry* 46: 363–376.

——— (2018). "Excluding exclusively the exclusive: Suppletion patterns in clusivity." *Glossa: A Journal of General Linguistics* 3 (130).

Müller, Gereon (2020). *Inflectional morphology in Harmonic Serialism*. Sheffield: Equinox.

Nair, Ravi Sankar S. (2012). "A grammar of Malayalam." *Language in India* 12: 1–135.

Nevins, Andrew (2007). "The representation of third person and its consequences for person-case effects." *Natural Language and Linguistic Theory* 25: 273–313.

——— (2010). *Locality in vowel harmony*. Cambridge, MA: MIT Press.

——— (2011). "Multiple Agree with clitics: Person complementarity vs. omnivorous number." *Natural Language and Linguistic Theory* 29: 939–971.

Newmark, Leonard, Philip Hubbard & Peter R. Prifti (1982). *Standard Albanian: A reference grammar for students*. Stanford, CA: Stanford University Press.

Noboru, Yoshioka (2012). *A reference grammar of Eastern Burushaski*. Doctoral dissertation. Tokyo: Tokyo University of Foreign Studies.

Noyer, Rolf (1993). "Optimal words: Towards a declarative theory of word-formation." Paper presented at ROW-1, New Brunswick, NJ.

Operstein, Natalie (2015). *Zaniza Zapotec*. Munich: LINCOM.

——— (2017). "Suppletion in Zapotec." *Linguistics* 55: 739–782.

Orgun, Cemil Orhan (1996). *Sign-based morphology and phonology with special attention to Optimality Theory*. Doctoral dissertation. Berkeley, CA: University of California.

Pittner, Karin (1995). "The case of German relatives." *The Linguistic Review* 12: 197–231.

Plank, Frans (2013). "The redistribution of suppletive stems: GEH- 'go' and STEH- 'stand' in Upper German dialects." Ms.

Prauliņš, Dace (2012). *Latvian: An essential grammar*. London: Routledge.

Preminger, Omer (2017). "Assessing the morphosemantic program for φ-features." Talk given at the 10th Brussels Conference on Generative Linguistics. URL: https://omer.lingsite.org/files/BCGL10.pdf.

Prince, Alan & Paul Smolensky (1993). "Optimality Theory: Constraint interaction in generative grammar." Report no. RuCCS-TR-2. New Brunswick, NJ: Rutgers University Center for Cognitive Science.

Rohlfs, Gerhard (1968). *Grammatica storica della lingua italiana e dei suoi dialetti*. Vol. 2: *Morfologia*. Torino: Einaudi.

Rolle, Nicholas R. (2018). *Grammatical tone: Typology and theory*. Doctoral dissertation. Berkeley, CA: University of California.

Rostovtsev-Popiel, Alexander (2020). "Megrelian." In Maria Polinsky (ed.), *The Oxford handbook of languages of the Caucasus*, pp. 529–569. Oxford: Oxford University Press.

Russ, Charles V.J. (1990). "Swabian." In Charles V.J. Russ (ed.), *The dialects of modern German: A linguistic survey*, pp. 337–363. London: Routledge.

Schulze, Wolfgang (1997). *Tsakhur*. Munich: LINCOM.

Schwyzer, Eduard (1939). *Griechische Grammatik*. München: Beck.

Smith, Peter W., Beata Moskal, Ting Xu, Jungmin Kang & Jonathan David Bobaljik (2019). "Case and number suppletion in pronouns." *Natural Language and Linguistic Theory* 37: 1029–1101.

Spangenberg, Karl (1990). "Thuringian." In Charles V.J. Russ (ed.), *The dialects of modern German: A linguistic survey*, pp. 265–289. London: Routledge.
Starke, Michal (2009). "Nanosyntax: A short primer to a new approach to language." *Nordlyd* 36: 1–6.
——— (2018). "Complex left branches, spellout, and prefixes." In Lena Baunaz, Karen De Clercq, Liliane Haegeman & Eric Lander (eds.), *Exploring Nanosyntax*, pp. 239–249. Oxford: Oxford University Press.
Steddy, Sam (2015). "Palatalisation across the Italian lexicon." *Rutgers Optimality Archive* 1229.
Steriade, Donca (2022). "Priscianic word formation: Morphomes, referrals and alternatives." Ms. MIT.
Strigin, Anatoli (2007). "Paradigmatic scale building in OT: The case of strong verbs in German." Ms. Zentrum für Allgemeine Sprachwissenschaft, Berlin.
Stump, Gregory T. (2001). *Inflectional morphology: A theory of paradigm structure*. Cambridge: Cambridge University Press.
Sumbatova, Nina (2020). "Dargwa." In Maria Polinsky (ed.), *The Oxford handbook of languages of the Caucasus*, pp. 147–200. Oxford: Oxford University Press.
Trommer, Jochen (2001). *Distributed Optimality*. Doctoral dissertation. Potsdam: University of Potsdam.
Vogel, Ralf (2001). "Case conflict in German free relative constructions: An optimality theoretic treatment." In Gereon Müller & Wolfgang Sternefeld (eds.), *Competition in syntax*, pp. 341–375. Berlin: Mouton de Gruyter.
Wagner, Peter (2012). *A grammar of North West Lovari Romani*. Doctoral dissertation. Prague: Charles University.
Weiss, Michael (2009). *Outline of the historical and comparative grammar of Latin*. Ann Arbor, MI: Beech Stave Press.
Wolf, Matthew (2008). *Optimal Interleaving: Serial phonology–morphology interaction in a constraint-based model*. Doctoral dissertation. Amherst, MA: University of Massachusetts.
Wunderlich, Dieter (2001). "How gaps and substitutions can become optimal: The pronominal affix paradigms of Yimas." *Transactions of the Philological Society* 99: 315–366.
Yip, Moira, Joan Maling & Ray Jackendoff (1987). "Case in tiers." *Language* 63: 217–250.
Zanetti, Umberto (2005). "Il bergamasco." In Andrea Rognoni (ed.), *Grammatica dei dialetti della Lombardia*, pp. 139–189. Milan: Mondadori.
Zompì, Stanislao (2017). *Case decomposition meets dependent-case theories*. MA thesis. University of Pisa.

Zompì, Stanislao (2019). “Ergative is not inherent: Evidence from *ABA in suppletion and syncretism.” *Glossa: A Journal of General Linguistics* 4 (73).