



Università di Pisa

DIPARTIMENTO DI FILOLOGIA, LETTERATURA E LINGUISTICA

Corso di Laurea Magistrale in Linguistica

Case decomposition meets dependent-case theories

A study of the syntax–morphology interface

Candidato:

Stanislao ZOMPI

Relatore:

Prof. Giovanna MAROTTA

Correlatori:

Prof. Roberto PERONI

Prof. Luigi RIZZI

Contents

Abbreviations	iv
Abstract	1
1 A case containment hierarchy	3
1.1 Caha's (2009) empirical claim	3
1.2 Accidental and systematic syncretism	4
1.3 Nanosyntactic spellout	9
1.4 Deriving Universal Contiguity with and without Nanosyntax	13
2 Not overstating the case	18
2.1 (Dis)ordering the obliques	18
2.1.1 Dative and genitive: challenges from Germanic	18
2.1.2 A problem with Latin's instrumental	22
2.1.3 On locating the locative(s)	25
2.1.4 What might this mean?	29
2.2 Bringing ergatives into the picture	31
2.2.1 Ergative and accusative pattern alike	31
2.2.2 Further evidence against a rigid hierarchy of obliques	33
2.2.3 The importance of tripartite systems	34
2.3 Interim conclusions	40
3 Further evidence for a (coarser) case containment hierarchy	42
3.1 Case-based pronominal suppletion	42
3.1.1 Smith <i>et al.</i> 's (2016) generalization	43
3.1.2 A containment-based account	46
3.2 Case-based stem allomorphy in non-pronominal paradigms	48
3.3 Overt containment	52

4	Containment relations among cases, and the theory of dependent case	57
4.1	Is ergative an inherent case? Some morphological counterevidence . . .	58
4.2	The leading alternative: the theory of dependent case	62
4.3	Marantz's algorithm, structural economy, and case-distinctness	65
4.3.1	Neutral and tripartite systems: refining case-distinctness	66
4.3.2	The puzzle of ditransitives: a successive-dissimilation approach	67
4.4	On semantically based case-marking	70
4.4.1	On consistently semantic S-marking: rare or impossible?	71
4.4.2	Basque, Laz, and the heterogeneity of 'agentive' case	77
5	More case studies	81
5.1	Problems with the genitive	83
5.1.1	NOM-GEN syncretisms in Latin, Gothic, Latvian, and Bilin . . .	83
5.1.2	Converging evidence from overt containment and suppletion .	85
5.1.3	Towards an explanation: genitive as an unmarked case	87
5.2	(Pseudo-)problems with the other obliques	88
5.2.1	Some unproblematic counterexamples	88
5.2.2	More problematic counterexamples: Greek and Kashmiri	89
	Bibliography	92

Acknowledgements

My foremost thanks go to my first linguistics teacher and mentor Pier Marco Bertinetto, for tempering my fanaticism down the years, and for teaching me that language is seldom as simple as it looks. I also thank him for always encouraging me to follow my own theoretical interests and tastes, and for supporting me in that enterprise in all possible ways, not least by warmly recommending me for various scholarships abroad.

The last of those scholarships got me to the Linguistics Department at the University of Geneva, where nearly half of this thesis was actually written. The welcome I received from that Department was warmer and pleasanter than I could ever have anticipated. I thus offer my heartfelt thanks to everyone in there, and most especially to my dear friend and office mate Giuseppe Samo, as well as to the Department's director Ur Shlonsky, without whom I'd never have had any office to share in the first place.

I must also thank the Genevan Department for allowing me to present some parts of this thesis at two fairly different venues: the VariaForMea Winter School at Macolin, in December 2016, and the SynCart seminar in Geneva, in January 2017. I am most grateful to the participants in both events, and especially to Enoch Aboh, Adriana Belletti, Johan Rooryck, and, again, Ur Shlonsky; I only wish I had talked to them sooner. Many thanks, too, to Anna Marchesini, for asking the right question at the right time, and thus giving me the perfect spur to write § 4.3.2.

Finally, I must gratefully acknowledge the generous assistance of all those who helped me gather or vet the cross-linguistic data I have relied on. I am especially indebted to Hagen Blix, Omkar Koul, Ludovico Pontiggia, and Adrian Stegovec for much-needed help with Laz, Kashmiri, Latin, and Slovene, respectively. A special word of thanks is also due to Maria Polinsky for amazingly expeditious and helpful assistance with Kabardian and Khinalug. Last but not least, I am thankful to the people and organizations that helped get me the reference grammars of some of the languages I was concerned with: this includes the Interlibrary Loan Services of the Scuola Normale Superiore and of the LM2 Library of the University of Pisa, the Linguistic Electronic Archive of Ca' Foscari University, plus a couple of Russian websites that it would probably be unwise to thank *nominatim*.

Abbreviations

A	transitive subject	DAT	dative
ABL	ablative	DECL.	declension
ABS	absolutive	DEF	definite
ACC	accusative	DEP	dependent
ADV	adverbial	DM	Distributed Morphology
ALL	allative	DOM	differential object marking
APASS	antipassive	DU	dual
APPL	applicative	ECM	exceptional Case-marking
AS	absolute state	ELAT	elative (local case)
AUX	auxiliary	ERG	ergative
B&B	Baker and Bobaljik (forthcoming)	EVID	evidential
C	complementizer	F	feminine
CAL.	Calidorus (Plautus' character)	FUT	future
CAUS	causative	GEN	genitive
COM	comitative	HON	honorific
CMPR	comparative	HU	human
D	determiner	ILL	illative
		IMP	imperative

INES	inessive	PL	plural
INF	infinitive	PREP	prepositional
INSTR	instrumental	PRES	present
LOC	locative	PREV	preverb
LOCPREP	locative-prepositional	PRF	perfect
LST	lexically stored tree (=LTS)	PROX	proximal
LTS	lexical tree structure (=LST)	PSE.	Pseudolus (Plautus' character)
M	masculine	PTCP	participle
N	neuter	REFL	reflexive
NEG	negation	S	intransitive subject
NOM	nominative	SG	singular
NONF	nonfeminine	SNOM	S-nominative (see § 2.2.3)
NP	noun phrase	SSD	Surrey Syncretism Database
O	direct object	SUBJ	subjunctive
OBL	oblique	SPRL	superlative
PART	partitive	TP	Tense Phrase
PASS	passive	VP	verb phrase
PERL	perlative	VOC	vocative
PF	Phonological Form		

Abstract

In this thesis, I attempt to lay out a comprehensive survey of *ABA patterns in case morphology—that is, patterns whereby a certain realizational rule can never apply to both case *x* and case *z* without also applying to case *y*. I show, more specifically, that such *ABA patterns can be found in systematic case syncretism, in case-conditioned wholesale suppletion, and in case-conditioned stem-formative allomorphy. I also argue that, crucially, the universal *ABA patterns evidenced by all these phenomena consistently make reference to the very same case classes, as encapsulated in (i).

- (i) a. If a systematic syncretism covers both an unmarked core case (NOMINATIVE or ABSOLUTIVE) and an inherent case (DATIVE, INSTRUMENTAL, LOCATIVE, etc.), it will also cover a marked core case (ACCUSATIVE or ERGATIVE).
- b. If the same stem formative appears in an unmarked core case (NOMINATIVE or ABSOLUTIVE) and in an inherent case (DATIVE, INSTRUMENTAL, LOCATIVE, etc.), it will also appear in a marked core case (ACCUSATIVE or ERGATIVE).
- c. If the same root appears in an unmarked core case (NOMINATIVE or ABSOLUTIVE) and in an inherent case (DATIVE, INSTRUMENTAL, LOCATIVE, etc.), it will also appear in a marked core case (ACCUSATIVE or ERGATIVE).

The generalizations in (i) trace back to recent work by Baerman *et al.* (2005) and Caha (2009) on syncretism, by McFadden (2014) on stem allomorphy, and by Smith *et al.* (2016) on suppletion. Most of these works, however, albeit largely consistent with each other, invoke some or other distinctive qualification or supplementation to the general pattern. Here I argue, in contrast, that such ‘deviations’ are not adequately empirically supported, and that the match between the relevant patterns across the three phenomena at issue is indeed perfect or nearly so. This argumentation rests, among other things, on a systematic defense of (i-a) against a sample of over 100 diverse case-inflecting languages, as well as on a sustained critique of Caha’s (2009) proposals concerning several other supposed *ABA universals in syncretisms with or among inherent cases.

Next, drawing again on Caha (2009) and Smith *et al.* (2016), I interpret the patterns summarized in (i) as evidence for a containment hierarchy like (ii), according to which every inherent case contains a marked core case, and every marked core case in turn contains an unmarked one.

- (ii) [[[NOM/ABS] ACC/ERG] INHERENT]

This hierarchy, while totally unexpected under theories that treat the ergative as just another inherent case (e.g. Woolford 1997; Legate 2008), can by contrast be naturally interpreted under those theories that have the ergative assigned in much the same way as the accusative. Prominent among these approaches is Marantz's (1991) dependent-case theory, whereby accusative and ergative are both treated as dependent cases—i.e. assigned to nominals which stand in an asymmetric c-command relation to some other non-oblique nominal nearby. In this light, I propose to reformulate (ii) as (iii).

- (iii) [[[UNMARKED] DEPENDENT] INHERENT]

I then observe that the lightest cases in the containment hierarchy (nominative and absolutive) are also the same cases that are given default (or, more accurately, 'elsewhere') status in the Marantzian theory of case assignment. This naturally prompts the conjecture that there might be a structural economy principle behind this pattern—a principle whereby additional case layers such as those of dependent and inherent cases may only be deployed in case of necessity, when the lighter, unmarked cases are somehow 'not enough'. The question arises, then, why these latter are sometimes 'not enough'. The answer I will tentatively explore will be to suppose that clauses must try their best to abide by the two requirements in (iv).

- (iv) a. Selectional requirements must be satisfied.
 b. Two noun phrases that stand in an asymmetric c-command relation within the same domain must bear sufficiently distinct cases (cf. Richards 2010).

On this account, the additional case structure of ergatives and accusatives may be deployed only when needed to tell apart two otherwise-overly-similar noun phrases, while the case structure associated to obliques may be deployed only if specifically selected by some idiosyncratic lexical item.

Finally, I discuss a range of problems besetting such an account, including some of the difficulties surrounding the proper treatment of tripartite and active~inactive alignments.

Chapter 1

A case containment hierarchy

1.1 Caha's (2009) empirical claim

In a series of recent works, Pavel Caha (2009; 2010; 2013) has advanced and explored the following empirical hypothesis.

- (1) *Universal Case Contiguity* (Caha 2009: 10)
- a. Non-accidental case syncretisms can only target contiguous regions of a linear case sequence invariant across languages.
 - b. The case sequence: NOM – ACC – GEN – DAT – INSTR – COM

The core content of (1) is intuitively clear: no systematic (i.e. non-accidental) syncretism may target two members of the sequence (1b) without also targeting all other members in between. Thus, for example, there can be no language in which nominative and genitive are systematically syncretic to the exclusion of the accusative, because NOM – GEN, unlike NOM – ACC – GEN, is not a contiguous portion of (1b).

Some qualifications are in order, however, regarding a couple of questions that this formulation may easily prompt. Perhaps the most obvious one concerns those languages (certainly the majority) that do not have all the case values referenced in (1b). What are we to make of them? To answer this question, it is first of all important, according to Caha (2009), to recognize that there are at least two ways in which a language *L* may purport to lack a case *C*. On the one hand, *L* may fulfil the typical functions of *C* by means of an adposition, possibly coupled with some additional case morphology on the adposition's complement. Such was the case, for example, for Latin's missing comitative, whose functions were fulfilled by combining the preposition *cum* with a complement in ablative/instrumental case. The generalization in (1) has nothing to say about 'missing cases'

of this sort,¹ which should therefore be simply ignored (that is, momentarily excised from the case sequence) in assessing whether the relevant ‘defective’ languages obey it or not.

There is, however, also another path whereby language *L* may lack case *C*, one which involves the complete loss of any morphological distinction between *C* and some other case *C'* across all the categories and paradigms of *L*, with no adposition being called into play. This is, for instance, what we observe for the missing instrumental of Classical Greek, whose core functions were taken on by the dative with no need for accompanying adpositions. Pervasive neutralizations of this sort may, according to Caha (2009: 107), be characterized as extreme instances of systematic syncretism (‘total syncretism’, in his own terms), and therefore fall squarely within the purview of the generalization in (1).² Just as for ‘partial syncretism’, then, we are led to predict, for example, that no language may totally neutralize the opposition between genitive and instrumental without totally neutralizing the one between genitive and dative—and so forth for the rest of the case sequence (1b).

A complementary question that may arise next is what should be made of those cases that are *not* referenced in (1b), such as the vocative, the ergative, the partitive, and all the spatial cases discussed by Blake (2001: § 5.6) and Creissels (2009). The simplest answer is that these should also be outright ignored in assessing the empirical validity of (1). Obviously, however, one would like to do better than that. In particular, Caha’s cartographic approach leads us to expect that each of these cases should also be assigned its own place in the case sequence. In § 2.1.3, we shall zero in on Caha’s (2009: § 3.4) own attempt to fit the locative case(s) into the sequence, and on some of the serious difficulties encountered, but for the nonce we will have to leave the issue at that.

1.2 Accidental and systematic syncretism

One further question hovering over Caha’s alleged generalization has to do with the distinction between systematic and accidental syncretism. Caha (2009: 10) makes crucial appeal to this point, emphasizing that (1) ‘is not a *surface* generalization’ (his italics), and that its purview should accordingly be restricted to systematic syncretism alone. But how

¹ Caha’s whole theory, however, does have a good deal to say about them. Caha (2009: § 1.5) predicts, in particular, that the same sequence in (1b) should also constrain the polysemy of prepositions in natural languages, such that no preposition may, for example, fulfil the functions of both genitive and instrumental without also fulfilling those of the dative. In an antisymmetric setup modelled after Cinque (2005), Caha (2010: 36) further predicts that, ‘[i]f the expression of a particular case in [(1b)] involves a preposition, then all cases to its right do as well’ (cf. Caha 2009: 30). Regrettably, I will have nothing to say about either of these interesting extensions of his theory.

² For an overview of differing views on ‘total syncretism’, see especially Albright and Fuß (2012).

can we concretely distinguish systematic syncretisms from accidental ones?

According to Caha (2009: § 1.1), accidental syncretisms fall into two types: phonological conflation and purely accidental homonymy.³ In instances of the former, syncretism ‘is the result of phonological processes which conflate two distinct underlying representation[s]’ (Caha 2009: 11). A classic example of this, first discussed in this context by Johnston (1996: 23–25), is offered by the NOM-GEN syncretism of Classical Arabic.⁴ This syncretism, illustrated by *qāḍīn* ‘judge’ in table 1.1, targets a non-contiguous portion of (1b), thus representing a *prima facie* threat to (1).

		‘judge’ (AS,SG)	
		<i>surface</i>	<i>underlying</i>
NOM	sāriq-u-n	qāḍīn	qāḍiy-u-n
ACC	sāriq-i-n	qāḍīyan	qāḍiy-a-n
GEN	sāriq-a-n	qāḍīn	qāḍi-y-in

Table 1.1: Two Classical Arabic paradigms (Johnston 1996: 23–24)

The source of this purported counterexample, however, is fairly clearly phonological. More specifically, as Johnston (1996) suggests, *qāḍīn* ‘judge’ inflects in fact just like the unproblematic *sāriqun* ‘thief’, except for the consequences a regular phonological process of truncation.

(2) Arabic truncation: $y + V[+high] \rightarrow \emptyset / i_$ (Johnston 1996: 25)

Since this process is restricted to high vowels, the low vowel *a* of the ACC desinence can safely escape it, whereas both NOM’s *u* and GEN’s *i*, being [+high], must get truncated, thus leading to a situation of adventitious NOM-GEN homophony. Such homophony, however, need not be (and, according to Caha, should not be) reflected in morphological or syntactic analyses, being no more than a superficial by-product of phonology.

In a less imperfect world, phonological conflation would be the sole source of accidental syncretism; and furthermore, it would always be the result of quite general phonological processes, fully operative in their respective languages. Unfortunately, however, we do not live in that world. On the one hand, ‘a pattern of syncretism may be restricted to a particular phonological environment without there being any generally applicable

³ Caha (2009) actually terms the latter type ‘accidental homophony’—a choice which has the unpleasant result of making ‘accidental homophony’ a subcase of ‘accidental syncretism’, while syncretism *tout court* is normally considered a subcase of homophony *tout court*. I have tried to avoid this by substituting Caha’s ‘accidental homophony’ with ‘purely accidental homonymy’ (the adverb not being ornamental).

⁴ See Johnston (1996: 23–25) also for discussion of the other (Caha-compatible) syncretisms of Arabic.

phonological rule that would account for it’ (Baerman *et al.* 2005: 10). On the other hand, we can also find apparently anomalous syncretisms that resist any kind of phonological explanation, but still look, as Plank (1991b: 180) puts it, ‘*isolated, superficial, and in various respects local*’ (his italics).

Let us begin, once more, with a relatively clear-cut example. The Latin language has at least the following five cases: nominative, accusative, genitive, dative, and instrumental.⁵ The language offers several surface counterexamples to (1), some of which are indeed problematic for Caha’s theory. At least one of them, however, most probably is not. First, take a look at the relative pronoun’s paradigm in table 1.2.

	M.SG	F.SG	N.SG	M.PL	F.PL	N.PL
NOM	quī	quae	quod	quī	quae	quae
ACC	quem	quam	quod	quōs	quās	quae
GEN	cūius	cūius	cūius	quōrum	quārum	quōrum
DAT	cuī	cuī	cuī	quibus	quibus	quibus
INSTR	quō	quā	quō	quibus	quibus	quibus

Table 1.2: The relative pronoun in Latin

This is how the pronoun was standardly inflected in the classical period. The only case syncretisms instantiated are that between NOM and ACC and that between DAT and INSTR. Both are clearly allowed by Caha’s Universal Contiguity (1), and are, indeed, extremely widespread in the language. As is well known, nominative and accusative are systematically syncretic in all Latin neuter paradigms, both in the singular and in the plural, in continuity with a pattern reconstructed as well for Proto-Indo-European.⁶ Another well-known fact about Latin is that none of its nominal, pronominal, or adjectival paradigms can formally distinguish the dative from the instrumental in the plural. The paradigm at hand is no exception in either of these respects.

A potential problem, however, stems from the presence of an alternative form *quī* for the ablative singular of all three genders.⁷ The existence of this form leads to an anomalous NOM-INSTR syncretism in the masculine singular, one which cannot be explained away on phonological grounds, and would thus appear to run counter to Caha’s alleged

⁵ Leaving aside vocative, locative, and perhaps also partitive. Notice also that I follow Caha—and hence depart from tradition—in referring to the ancient grammarians’ sixth case as ‘instrumental’ rather than ‘ablative’, ‘because expressing an instrument (rather than source) is the most prominent meaning of this case when used in isolation’ (Caha 2009: 122).

⁶ This very systematic syncretism has been repeatedly cited as evidence that Proto-Indo-European had ergative characteristics. For a critique grounded in contemporary typological knowledge (traceable especially to the seminal work of Silverstein 1976), see Rumsey (1987).

⁷ This small problem went unnoticed to Plank (1991b), Johnston (1996), and Caha (2009), and surely would have to me, too, if it had not been for Ludovico Pontiggia’s generous help.

generalization. It is somewhat unclear, however, just how much importance should be attached to a supposed counterexample of this sort. This is, after all, the only known instance of a syncretism of this kind in the whole Latin language. And it is not simply restricted to a single inflection class—it is actually restricted to a single *word form*. In this light, it is tempting to analyze this syncretism as an instance of purely accidental homonymy, much like that between the nominal plural marker *-(e)s* and the 3SG agreement marker *-(e)s* in contemporary standard English.⁸

Several indications converge toward this conclusion. Foremost among them is the absence of any formal covariation between the two syncretic cases.⁹ Thus, as soon as we shift, for example, from the instrumental *quī* to its alternant *quō*, the syncretism immediately breaks down, there being no syncretic form *quō* alongside *quī* in the nominative. Likewise, if we turn to a different gender and focus on the neuter nominative *quod*, the syncretism breaks down again, in the absence of a form *quod* (along with *quī* and *quō*) for the instrumental.¹⁰ It turns out, furthermore, that some special stipulation concerning the idiosyncrasies of *quī* may be necessary in any event, given that the same form is also attested as a relative adverb (‘wherewith, whereby’, referring to all genders and numbers; see (3)) and even as a complementizer (‘so that, in order that’; see (4)).

(3) Latin (Plautus, *Aulularia* 3.5)

mihi quidem aequ-om 'st [...] dari [...] vehicl-a quī
 1SG.DAT really fair-NOM.N.SG is to.be.given vehicle-ACC.PL whereby
 veh-ar.
 carry-1SG.SUBJ.PASS
 ‘it really is fair that I should be given [...] vehicles whereby I may be carried.’

(4) Latin (Plautus, *Pseudolus* 1.1)

CAL. resti-m vol-o mihi emere.
 halter-ACC.SG want-1SG 1SG.DAT to.buy
 ‘I want to buy myself a halter.’

⁸ The Latin homonymy must almost certainly be regarded as an accident from the historical viewpoint, given that the two syncretic forms trace back to different stems: on the one hand, nominative *quī* comes from *quei* < **k^wo-i* (cf. QOI in early inscriptions), formed from the *s*-less nominative of the stem *k^wo-* with the addition of a deictic particle *-i*; on the other hand, ablative/instrumental *quī* is, in all likelihood, formed out of the stem *k^wi-* instantiated by the interrogative *qui-s* ‘who (NOM.SG.M/F)’. Cf. Kühner and Holzweissig (1912: 608–611), Ernout (1953: 85–87), Leumann (1977: 472–473), and Weiss (2009: 350–351).

⁹ We effectively find the archaic alternant *quei* as also used both in the nominative and in the instrumental (Kühner and Holzweissig 1912: 609, 611). However, this covariation is the result of a well-known phonological change from *ei* to *i*, operative quite generally in the language (see, *i. a.*, Niedermann 1906: 52).

¹⁰ The same goes for the (historically related) interrogative pronoun *quīs/quid*, which also admits *quī* (alongside *quō*) in the instrumental singular of all three genders. Once more, with a nominative form other than *quī*, the NOM-INSTR syncretism disappears.

PSE. *quam* *ob re-m?*
 what.ACC.F.SG for thing-ACC.SG
 ‘What for?’

CAL. *quī* *me* *faci-am* *pensil-em.*
 in.order.that 1SG.ACC make-1SG.SUBJ hanging-ACC.M.SG
 ‘To hang myself (with it).’

For all of these reasons, positing accidental homonymy in this case seems like a perfectly reasonable move; and indeed, hardly anybody would count NOM-INSTR among Latin’s systematic syncretisms.

However, much as with phonological conflation, it should be borne in mind that things are seldom so clear-cut. As a consequence, in order to deal consistently with the various less clear-cut cases, one needs some readily applicable criterion to rely on. Caha (2009: 15, 234) follows, in this regard, Johnston’s (1996: 15) proposal ‘to rely primarily on the criterion of co-extension of the homonymy under allomorphy and under different conditions of cumulation’—that is, what I have referred to above as ‘formal covariation between the syncretic cases’. Johnston’s (*ibid.*) reasoning goes as follows.

If we find that a suffix *x* in a certain context realizes^[11] properties *a* and *b*, it is entirely possible that the homonymy is accidental and of no more account than the two senses of *bank* in English. But if we find that in another context a suffix *y* also realizes properties *a* and *b*, then it becomes more likely that the homonymy is systematic.

This leads us to the following criterion, which might as well be encapsulated in the slogan: ‘accidents should not repeat’ (Caha 2009: 15).

(5) Accidental syncretisms are limited to single exponents.

This is not to say that non-accidental syncretisms *must* necessarily be replicated by different exponents. It only says that accidental syncretisms *cannot* be.¹² As a consequence, (5) may sometimes mark a syncretism as systematic, but may never mark any syncretism as undoubtedly accidental. This much suffices, however, to save us from the temptation of appealing too often to putatively accidental homonymy in the face of potential counterexamples.

¹¹Following much current work, Johnston says that stems and affixes ‘realize’, or are ‘exponents’ of, the feature values they are associated with. I shall adhere to this usage, too.

¹²Harbingers of this criterion can also be found in Plank (1991b). For discussion of some other diagnostics proposed by Zwicky (1991), see especially Johnston (1996: 13–15). See also the more recent discussion in Harbour (2008).

	‘dog’	‘horn’	‘you’ (sg)
NOM	canīs	cornū	tū
ACC	canem	cornū	tē/tēd ¹³
GEN	canīs	cornūs	tui/tis
DAT	canī	cornū	tibi
INSTR	cane	cornū	tē/tēd

Table 1.3: Three Latin singular paradigms

It is important to note, before moving on, that the criterion in (5) is met whenever a syncretism involves a single *exponent*, no matter how many *word forms* that exponent may appear in within the language. In other terms, the sheer token frequency of a certain exponent among a language’s word forms, or in utterance corpora of any kind, plays no role at all when it comes to assessing the accidentality of a syncretism, because multiple instances of the same exponent will always count as just one for such purposes. It so happens, for instance, that Latin’s NOM-GEN syncretism, instantiated by so many nouns of the so-called third declension, may eventually be characterized as accidental, since it only involves the single exponent *-īs*. On the other hand, Latin also features a sporadic syncretism between ACC and INSTR to the exclusion of GEN, which would then be ruled out by Caha’s Universal Contiguity. Albeit manifested by no more than twenty paradigms, this syncretism does qualify as systematic—distressingly for Caha’s theory—for it involves at least two different exponents. Both the NOM-GEN and the ACC-INSTR syncretisms, sketchily exemplified in table 1.3, will be taken up again, respectively, in § 5.1.1 and § 2.1.2. What matters now, however, is that they illustrate how systematicity and language-internal frequency do not necessarily correlate together.

1.3 Nanosyntactic spellout

Now that the content of Caha’s (2009) empirical claims should be clear, it is finally possible to move on to the more theoretical side of his proposal. This is couched within a broader approach to the architecture of language, which goes by the name of Nanosyntax.¹⁴ The point of departure of this approach is the idea that syntax operates on basic building blocks that are, in the vast majority of cases, *submorphemic*—smaller than individual morphemes. As a consequence, syntax cannot just build *on* morphemes: it must itself build morphemes (cf. Starke 2009: 6). Taken to its furthest consequences, this means

¹³Italics indicate forms from the archaic language.

¹⁴Apart from Caha’s (2009; 2010; 2013; 2016) own work, see especially Starke (2009) and Blix (2016: ch. 4).

that there need be no such thing as a pre-syntactic lexicon, and that, conversely, syntax may operate only on pre-lexical units—possibly maximally simple, each hosting no more than one feature.¹⁵ In such a scenario, ‘the lexicon, instead of feeding the syntax as in more traditional approaches, interprets it’ (Blix 2016: 29). Lexically stored exponents must therefore be allowed to realize not only terminal nodes of syntactic trees but also more complex structures pre-generated by the syntax. We need, in other words, to allow for non-terminal (or phrasal) spellout, alongside the more traditional terminal spellout.¹⁶

How exactly does this spellout process work? A popular Nanosyntactic answer consists of the following assumptions, on which I will briefly comment in turn.

(6) *Cyclic Override*

Spellout applies cyclically, bottom up, with lexicalization attempted at every node. Each successful spellout overrides previous spellouts.

In other words, for each tree generated by the syntax, the spellout algorithm scans its structure bottom up, and tries to find, at each node, an eligible Phonological Form (or PF). When the algorithm finds one such PF, then this PF is temporarily inserted, overwriting any material that may have been inserted before it.

What makes a particular PF eligible? The answer is primarily provided by the Superset Principle.

(7) *The Superset Principle*

- a. A lexically stored tree LST is a candidate for spellout of a syntactic structure SS if LST either matches SS or contains a constituent that matches SS.
- b. A lexical constituent LC *matches* a syntactic structure SS iff LC is identical to SS, ignoring traces and already-spelt-out constituents.

To put it less formally, a lexical tree may realize a given structure only if it ‘contains’ that structure. What happens, then, if more than one lexical tree meets this criterion? In that case, the solution is furnished by a Nanosyntactic adaptation of Pāṇini’s Elsewhere Principle. Pāṇini’s time-honored insight is essentially that, whenever two conflicting rules may apply to the same environment, precedence must be given to the more specialized rule—namely, the one that applies to fewer environments. Starke’s (2009: 4) reinterpretation of this, informally dubbed the ‘Minimize Junk’ Principle, goes roughly as follows.

¹⁵Parallels of this latter restriction may be found in Kayne (2005: 15; his (53)) and Cinque and Rizzi (2010: 54, 61).

¹⁶Precursors and parallels of Nanosyntax’s phrasal spellout include McCawley (1968), Weerman and Evers-Vermeul (2002), Neeleman and Szendrői (2007), and Radkevich (2010).

(8) *Minimize Junk (or the Nanosyntactic Elsewhere Principle)*

If several lexical items match a syntactic structure, the candidate with the fewest unused nodes wins.

Finally, there remain a few more intricate cases (especially instances of ‘overlap’ between lexical trees) that the simple system outlined so far cannot handle by itself. For such cases, Caha follows Abels and Muriungi (2008: 720) in positing a further condition ‘according to which the bottom of the spelled-out stretch must coincide with the bottom of the morpheme’.

(9) *Anchoring*

The lowest node of the structure to be spelt out must coincide with the lowest node of the lexically stored tree.

As a consequence, the LST $[c[a[b]]]$ will be an eligible spellout of $[a[b]]$, the ‘anchor’ b being correctly situated at the bottom of both structures. By contrast, the lexical tree $[a[b[c]]]$ will not be an alternative spellout, because its bottom is illicitly occupied by the unused node c , in violation of (9).

Now let us see, before focusing back on case, how this whole mechanism applies to a simple example borrowed from Starke (2009: 3) and Blix (2016: 33–35). The toy lexicon we are going to use is listed in (10).

(10) List of lexically stored trees (from Blix 2016: 33)

- | | | | |
|----|-----------------|-------------------|--|
| a. | -s | \Leftrightarrow | $[\text{PL}]$ |
| b. | <i>mouse</i> | \Leftrightarrow | $[n[\sqrt{\text{mouse}}]]$ |
| c. | <i>mice</i> | \Leftrightarrow | $[\text{PL} [n[\sqrt{\text{mouse}}]]]$ |
| d. | <i>elephant</i> | \Leftrightarrow | $[n[\sqrt{\text{elephant}}]]$ |
| e. | <i>sheep</i> | \Leftrightarrow | $[\text{PL} [n[\sqrt{\text{sheep}}]]]$ |

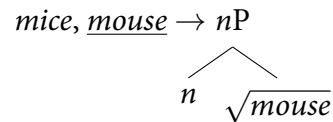
Here, then, I reproduce Blix’ (2016: 33) explanation of how the singular and plural forms of *mouse* and *elephant* may be derived.¹⁷

Let us assume, for the sake of the argument, that the structure of a noun in English is something like $[(\text{PL})[n[\sqrt{\text{root}}]]]$, with the plural head being absent in the singular. [...] The six singular and plural forms are derived as in ⟨17⟩. In ⟨17a⟩, both *mouse* and *mice* are possible candidates, but *mice* includes an unused feature PL, thus loses

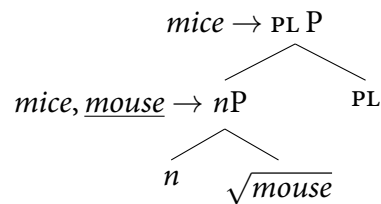
¹⁷Here I have normalized Blix’ use of quotes with respect to my usage elsewhere in this thesis, and enclosed his example numbers in angled brackets to distinguish them from my own.

to *mouse* due to EP [= the Elsewhere Principle ...] In ⟨17b⟩, on the other hand, *mouse* also wins in the first cycle, but is overwritten at the next cycle by *mice*, now the only candidate. In contrast, the plural of *elephant*, in ⟨17d⟩, results in bi-morphemicity, because at point (B) there is no LTS available that could spell out the bigger structure [...] The tree structure is replaced by phonological material, and the process repeats; now -s is the only suffix that is a possible candidate for [PL], as in ⟨17e⟩.

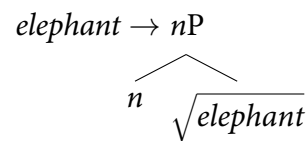
- ⟨17⟩ a. *mouse* – singular
(MINIMIZE JUNK)



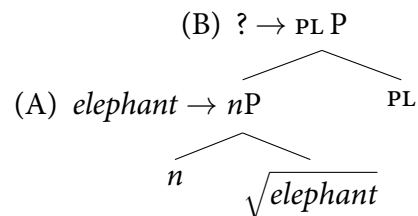
- ⟨17⟩ b. *mouse* – plural
(CYCLIC OVERRIDE)



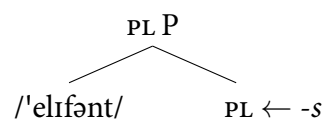
- ⟨17⟩ c. *elephant* – singular



- ⟨17⟩ d. *elephant* – plural (1)
(BIMORPHEMICITY)

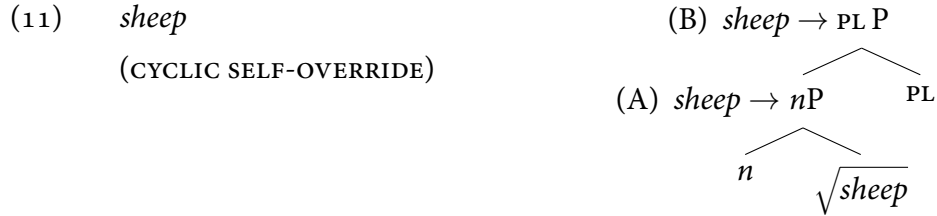


- ⟨17⟩ e. *elephant* – plural (2)
(BIMORPHEMICITY)



The most interesting case for our purposes, of course, is that of the number-syncretic form *sheep*. This is also where the Superset Principle shows its revolutionary potential most clearly. On the one hand, more traditional approaches usually treat syncretism either as an accident (in this case, for example, by positing an accidentally silent plural morpheme) or in terms of underspecification (as will be detailed below for case syncretism). In Nanosyntax, by contrast, syncretic forms are always partly *over*-specified. The LST

sheep, for example, is not somehow underspecified for number, but is, on the contrary, always lexically plural (see, e.g., (10e)). The Superset Principle, however, allows us to use it as well to realize a non-plural structure, in the absence of ‘less junky’ competitors. We may thus envisage the following spellout procedure, which Blix (2016: 35) aptly refers to as ‘cyclic self-overwriting’.



1.4 Deriving Universal Contiguity with and without Nanosyntax

The Superset-based approach to syncretism can also be applied to the domain of case. To derive the putative generalization in (1), all we need do is supplement the background laid out in § 1.3 with a convenient theory of the structure of case features. Such a theory, according to Johnston (1996: ch. 3) and Caha (2009), will have to dispose of traditional cross-classifying features¹⁸ and be couched instead in terms of ‘*sub*-classifying’ features—or, in Caha’s terminology, in terms of ‘cumulative decomposition’. The theory needs to ensure, in other words, that the featural content of each case in the case sequence be a proper subset of the content of all cases to its right. An example of case decomposition of this sort is given in (12).¹⁹

- | | |
|---|---|
| <p>(12) a. NOM = [A]</p> <p style="margin-left: 2em;">b. ACC = [A, B]</p> <p style="margin-left: 2em;">c. GEN = [A, B, C]</p> | <p style="margin-left: 2em;">d. DAT = [A, B, C, D]</p> <p style="margin-left: 2em;">e. INSTR = [A, B, C, D, E]</p> <p style="margin-left: 2em;">f. COM = [A, B, C, D, E, F]</p> |
|---|---|

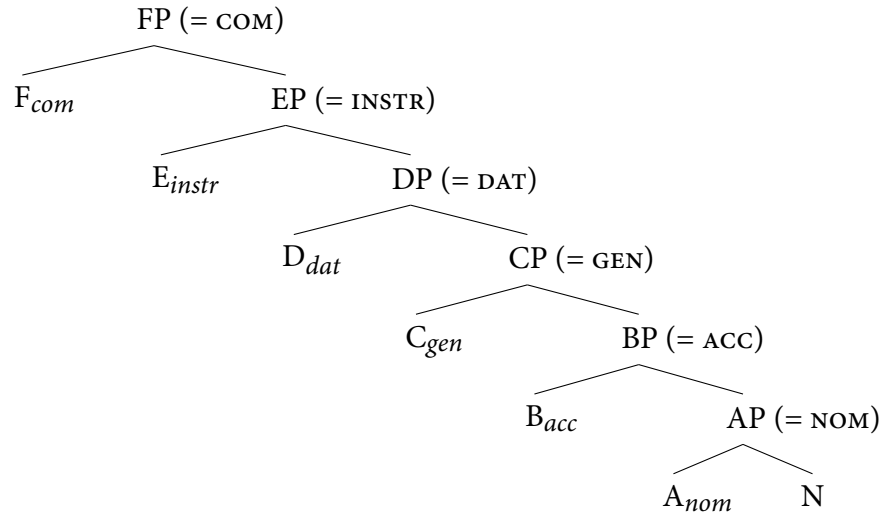
Caha’s theory, pursuant to the principle ‘One Head – One Feature’, hosts each of these atomic features in its own functional head. The functional projections corresponding to

¹⁸See Jakobson (1936) for an influential cross-classificatory proposal, and Müller (2004) and Calabrese (2008) for two recent attempts of this sort within the framework of Distributed Morphology.

¹⁹As Caha (2009: 7) correctly notes, the reverse complexity hierarchy, with comitative as the lightest case and nominative as the heaviest, would yield the very same results. His ‘decision is motivated by the traditional observation [...] that nominative is the most unmarked case, and that the oblique cases form a group differentiated from the “core” cases, nominative and the accusative’. Further supporting evidence will come from patterns of affix stacking and overt containment, which I will return to in § 3.3.

the various cases in (1b), thus, come to form a functional sequence of the sort familiar from cartographic work (Cinque and Rizzi 2010; Shlonsky 2010; Rizzi and Cinque 2016). The cartography delineated so far is represented in (13), the syntactic structure of a noun in comitative case.

(13)



To see how this derives Universal Contiguity, look at the hypothetical offending paradigm in (14).

(14)

<i>Case</i>	<i>Form</i>
NOM	α
ACC	β
GEN	α

If α is to spell out the genitive in accordance with the Superset Principle, then α 's lexical tree structure must at least improperly contain $[C_{gen} [B_{acc} [A_{nom} N]]]$. By contrast, the entry of β must be capable of realizing the accusative but not the genitive: it must therefore be $[B_{acc} [A_{nom} N]]$. In the spellout system laid out so far, both α and β are thus eligible for spellout of ACC, but α , unlike β , leaves one node unused, so that β wins out due to Minimize Junk. By the same token, however, the same competition should also arise for spellout of NOM: here, neither α nor β perfectly matches the 'realizandum' $[A_{nom} N]$, but β is again a better match, having fewer unused nodes than α (only B_{acc} compared to at least B_{acc} and C_{gen}). The nominative form should thus be either β or a still better match γ , if there is one, but it can never be α . In other words, a paradigm like (14) is, in principle, impossible.²⁰ The same reasoning can be iterated, as may be easily verified, for all the other syncretisms ruled out by (1).

²⁰An exception arises whenever our hypothetical γ , which is supposed to lexicalize $[A_{nom} N]$, is accidentally

Two further questions might arise at this point. The first one is whether there is any independent evidence for such containment relations among cases as those posited by Caha. It may effectively seem counterintuitive that the instrumental should contain the dative, or that the dative should in turn contain the genitive. As Caha (2010: 72fn15) notes, however, both of these containment relations are made transparent by the case morphology of the Lezgian language Budugh. As shown in table 1.4, the containment relation between DAT and INSTR is made overt in the singular paradigms of the third-person pronoun and of the noun *furi* ‘man’, while the relation between GEN and DAT is manifested in the latter paradigm only.²¹

	‘man’	3 SG
NOM	furi	ad
GEN	fura	anda
DAT	fura- z	anəz
INSTR	fura- z-en	anəz-n

Table 1.4: Fragments of three singular paradigms in Budugh (Caha 2010: 46)

As for the accusative, both its superset relation to NOM and its subset relation to GEN are made overt in the Vlakh Romani paradigms in table 1.5 (from Friedman 1991: 57, reported in part by Caha 2010: 41). While the former relation is manifested only by the singular of ‘girl’, the latter is made fully transparent by both paradigms in both numbers.

	‘boy’, SG	‘boy’, PL	‘girl’, SG	‘girl’, PL
NOM	čhav-ó	čhav-é	balvál	balval-á
ACC	čhav-és	čhav-én	balval-á	balval-én
GEN	čhav-és-koro... ²²	čhav-én-goro...	balval-á-koro...	balval-én-goro...
DAT	čhav-és-ke	čhav-én-ge	balval-á-ke	balval-én-ge
INSTR	čhav-és-ar	čhav-én-car	balval-á-sar	balval-én-car
LOC	čhav-és-te	čhav-én-de	balval-á-te	balval-én-de
ABL	čhav-és-tar	čhav-én-dar	balval-á-tar	balval-én-dar
VOC	čhav-éa	čhav-álen	balval-íe	balval-álen

Table 1.5: Two full Vlakh Romani paradigms (Friedman 1991: 57)

homophonous with *a*. Clearly, such syncretism does not result from the mechanics of spellout, but rather from later phonological processes or mere idiosyncrasies in the lexicon. This is why such accidents must be excluded from the purview of the theory—and from that of the ensuing generalization in (1).

²¹The genitive case I am talking about here is the one that Authier (2013) terms ‘inlocative’, namely, the one used for inalienable possession. Unusually enough, Budugh also features another genitival case (Authier’s ‘adlocative’, mostly realized by *-u*) for expressing alienable possession.

²²The genitive is followed by ellipsis to indicate the fact that there are also forms in *-i* and *-e* [e.g. *-kere* and

A few more examples of this sort will be discussed in § 3.3. The Budugh and Vlach Romani data reported above, however, should already suffice to show that the notion of a containment hierarchy of cases, counterintuitive as it may seem, is not only useful for modelling syncretism, but is also supported by more direct morphological evidence. This suggests that the notion is well worth taking seriously and exploring further.

The second question I'd like to touch on is whether we really need the whole spellout machinery of Nanosyntax in order to capture Case Contiguity, or whether such a result may also be encoded in terms of underspecification within more traditional frameworks. Nothing, as far as I can see, prevents an account of this latter sort. Let us take, for example, a theory like Distributed Morphology (Halle and Marantz 1993; Embick and Noyer 2007), in which Vocabulary Insertion is again thought to be post-syntactic, but is only allowed to target terminal nodes. Here, Starke's Superset Principle in (7) is substituted with Morris Halle's Subset Principle, whose original formulation also incorporates a convenient reinterpretation of Pāṇini's Elsewhere Principle. This is why I follow Caha (2016: § 3.1) in splitting the principle in two halves, dubbing them the 'Subset Clause' and the 'Elsewhere Clause'.

(15) *The Subset Principle* (Halle 1997: 428)

[*Subset Clause*] The phonological exponent of a Vocabulary Item is inserted into a morpheme of the terminal string if the item matches all or only a subset of the grammatical features specified in the terminal morpheme. Insertion does not take place if the Vocabulary Item contains features not present in the morpheme. [*Elsewhere Clause*] Where several Vocabulary Items meet the conditions for insertion, the item matching the greatest number of features in the terminal morpheme must apply.

Let us now see how Caha's putative generalization may also be derived in a system of this type, in combination with a featural decomposition of cases along the lines of (12), repeated here as (16).

- | | | | | |
|------|----|-----------------|----|--------------------------|
| (16) | a. | NOM = [A] | d. | DAT = [A, B, C, D] |
| | b. | ACC = [A, B] | e. | INSTR = [A, B, C, D, E] |
| | c. | GEN = [A, B, C] | f. | COM = [A, B, C, D, E, F] |

Let us focus, in particular, on a simple hypothetical paradigm like (17).

-kiri] depending on [the gender, number, and case of] the head noun' (Friedman 1991: 57, fn2; see also his p. 61).

- (17)
- | | |
|-----|----------|
| NOM | α |
| ACC | β |
| GEN | α |

In order to spell out the nominative in a DM-like system, α must bear either a proper or an improper subset of the nominative's grammatical features—that is, it must bear either only [A] or no case feature at all. On the basis of the Subset Clause in (15), α is therefore a valid competitor also for spellout of ACC, given that NOM's featural specification is itself a proper subset of ACC's. The accusative form, however, is not α but β , which means that β must be a better match than α according to Halle's (1997) Elsewhere Clause. This leaves us only three possibilities.

- (18) $\alpha = [\emptyset]$, $\beta = [B]$; $\alpha = [\emptyset]$, $\beta = [A, B]$; $\alpha = [A]$, $\beta = [A, B]$.

Under any of these possible scenarios, however, it turns out that β is a better match than α also for spellout of the genitive, for it always matches at least one more GEN feature than does α . This effectively rules out a paradigm like (17)—*quod erat demonstrandum*.

In summary, this latter argument shows that neither phrasal spellout nor the Superset Principle are strictly necessary to derive the restrictions on case syncretism that Caha proposes. What does seem minimally necessary, instead, is the combination of two key ingredients: (i) cumulative decomposition of the relevant case values and (ii) a realizational (i.e., post-syntactic) morphology in which some version of Pāṇini's Elsewhere Principle regulates competition among multiple eligible exponents.²³

²³Which is not to say, of course, that alternative accounts are impossible to construct on the basis of different premises. Luigi Rizzi (p.c.), in particular, has recently suggested to me the prolegomena to one such alternative: if we think that realizing a structure with the same exponent used for its subset is equatable to (post-syntactic) Internal Merge, whereas realizing a structure differently from any of its subsets is equatable to (post-syntactic) External Merge, then we may effectively account for the *ABA pattern of syncretism in terms of two syntactic principles already advocated on independent grounds: the (admittedly controversial) Merge-over-Move principle of Chomsky (1995: § 4.9), and Rizzi's (1990; 2001; 2004) own principle of Relativized Minimality. However, it is not very clear to me, as yet, how to give this idea an architecturally sound implementation, nor how to extend it to the parallel case of *ABA patterns in suppletion (see § 3.1). I will therefore have to leave the issue at that, waiting for further insights. For still other (more worked out) accounts, see Graf (2017) and Bobaljik and Sauerland (2017).

Chapter 2

Not overstating the case

2.1 (Dis)ordering the obliques

2.1.1 Dative and genitive: challenges from Germanic

Elegant though it is, Caha’s (2009) theory encounters several serious empirical problems. One that has recently been pointed out by Harðarson (2016) concerns the relative order of genitive and dative. As we saw in chapter 1, Caha argues that the dative always contains the genitive, on the grounds of the widespread existence of systematic ACC-GEN syncretisms to the exclusion of a non-syncretic DAT. The following cases in point are from Russian.

	‘window’, SG	‘teacher’, PL	‘100’
NOM	okn-o	učitel-ja	st-o
ACC	okn-o	učitel-ej	st-o
GEN	okn-a	učitel-ej	st-a
LOCPREP	okn-e	učitel-jax	st-a
DAT	okn-u	učitel-am	st-a
INSTR	okn-om	učitel-am-i	st-a

Table 2.1: Fragments of three Russian paradigms (Caha 2009: 12)

Further relevant examples may be found in other Slavic languages like Serbian, Slovene, Czech, and Ukrainian (Caha 2009: 237–268), in Uralic languages like Finnish, Estonian, and Northern Saami (Caha 2009: 100–102), as well as in Old English (Plank 1991b: 170–181). This last language, however, also provides some first reasons for skepticism, since it also contains several syncretisms between accusative and *dative* ‘across’ a non-syncretic genitive. Five such problematic cases are displayed in table 2.2.²⁴

²⁴The pronominal forms *me*, *ūs*, and *ēow* admit of the alternants *meč*, *ūsič*, and *ēowič*, respectively. These

	‘stone’, SG	‘friend’, SG	‘book’, SG	1SG	1PL	2PL
NOM	stān	wine	bōc	iċ	we	ġe
ACC	stān	wine	bōc	me	ūs	ēow
GEN	stān-es	wines	bōc-e, bēc	mīn	ūre	ēower
DAT	stān-e	wine	bōc, bēc	me	ūs	ēow

Table 2.2: Six Old English paradigms (Campbell 1959: 223, 243, 252, 288)

At least two of these unexpected syncretisms may likely be explained away. For ‘i-nouns’ like *wine* ‘friend’, for example, it may be argued that the *-e* of NOM/ACC and the *-e* of DAT do not share the same status: while the former is clearly part of the stem, the latter may as well be a suffix borrowed from the paradigm of ‘a-nouns’ like *stān* ‘stone’. Analogy with ‘a-nouns’ was certainly active in the genitive, whose ending was originally *-e < -i < *iz*, but was later equated to the *-es* of *stānes* (see especially Brunner 1965: 213). In this light, the dative *wine* may be plausibly analyzed as the result of contraction from **wine-e*. As for nouns like *bōc* ‘book’, the suffixes of NOM/ACC.SG and DAT.SG, albeit both null, were originally distinct in that the latter, unlike the former, triggered so-called *i*-Umlaut in the root vowel. This is documented by the more conservative alternant *bēc*, which could be used for genitive/dative only. As time went by, however, ‘many analogies intervened’ (Campbell 1959: 253), leading to conflation of NOM/ACC and DAT. Although this analogical process is certainly unexpected from Caha’s perspective, it might still be analyzed as the insurgence of a lexical idiosyncrasy concerning the root *bōc* and others like it—a diacritic that effectively exempted such roots from Umlaut effects.

By contrast, the pronominal paradigms of *iċ*, *we*, and *ġe* make for a better challenge. Their problematic syncretisms clearly cannot all be purely accidental, nor can they be explained away on phonological grounds. To cope with them, Caha (2009: § 8.4) is thus led to propose that all those offending genitives are in fact instances of the corresponding homophonous possessive adjectives—*mīn* ‘my’, *ūre* ‘our’, and *ēower* ‘your’. Caha is well aware, however, that the distinction between the pronominal genitives and the corresponding possessives was originally drawn by grammarians for a reason—namely, because the genitives of pronouns, like those of all other Old English nominals, also admit of adverbial uses. The verb *helpan* ‘to help’, for example, selects a complement in genitive case.

alternants are restricted to non-western varieties (cf. Brunner 1965: 259) and ‘are generally limited to acc.’ even though datival uses are not entirely unattested (Campbell 1959: 288). ACC-DAT covariation also concerns the quantity of *-e* in *me*, which, according to Brunner (1965: 258), may be either short or long in both the accusative and the dative form. Other alternations (*ēow-/īow-*, *ūre/ūser* etc.), are immaterial for present concerns.

- (19) God ūre help-e. (Sermo Lupi 71, quoted by Baker 2003: 181)
 God 1p.GEN help-PRES.SUBJ
 ‘God help us.’

This behavior is somewhat unexpected of run-of-the-mill possessives. Caha’s (2009: 278–282) solution, then, is to claim that these are not run-of-the-mill possessives, but rather possessives that modify a silent noun SELF, itself in genitive case. Example (19), on this account, should thus be reglossed as (20).

- (20) God ūre SELF-∅ help-e.
 God our.GEN self-GEN help-PRES.SUBJ
 ‘God help us.’

This solution works well for this particular example, given that *ūre* can also realize the genitive feminine singular of the 1PL possessive adjective. One may wonder, of course, why SELF should be feminine, or why it should be singular despite referring to a plural participant. But apart from these questions, the morphology does not pose insurmountable problems. Things are different, though, when it comes to 1SG *mīn*, which also admits of adverbial uses.

- (21) Drihten, geþenc mīn þonne...
 lord think.IMP 1s.GEN when
 ‘Lord, think of me when...’²⁵

The full paradigm of the corresponding possessive is shown in table 2.3.

	M.SG	F.SG	N.SG	M.PL	F.PL	N.PL
NOM	mīn	mīn	mīn	mīn-e	mīn-e/-a	mīn
ACC	mīn-ne	mīn-e	mīn	mīn-e	mīn-e/-a	mīn
GEN	mīn-es	mīn-re	mīn-es	mīn-ra	mīn-ra	mīn-ra
DAT	mīn-um	mīn-re	mīn-um	mīn-um	mīn-um	mīn-um

Table 2.3: The possessive *mīn* ‘my’ in Old English (cf. Campbell 1959: 262, 290)

Nowhere in this paradigm is the genitive realized as *mīn*. Rather, it may only read *mīnes*, *mīnre*, or *mīnra*. Keeping to the assumption that the silent SELF is always feminine and singular, one would have to posit an underlying genitive form *mīnre*, which should subsequently be stripped of the suffix *-re*—a process that should, unaccountably, target all and only the non-adnominal uses of this form. In fairness, this does not sound like a

²⁵From Ælfric, *Catholic Homilies* I, 37, 272; quoted by Koike (2004: 268).

satisfactory explanation.

One might then be tempted to say that there is a genuine problem here, and that the case sequence for pronouns may, for some reason, be different from the one for lexical nouns. The following Icelandic paradigms, however, advise against such a move, and suggest instead that the problem may be more general.

	NAME	‘geyser’, SG	‘old woman’, SG	1PL	2PL	2PL.HON
NOM	Hild-ur	hver	kerling	við	þið	þér
ACC	Hild-i	hver	kerling-u	okk-ur	ykk-ur	yð-ur
GEN	Hild-ar	hver-s	kerling-ar	okk-ar	ykk-ar	yð-ar
DAT	Hild-i	hver	kerling-u	okk-ur	ykk-ur	yð-ur

Table 2.4: Six Icelandic paradigms (Einarsson 1949: 33, 38, 68)

Here too, we find the ACC-DAT syncretism ‘across’ GEN which characterizes so many Germanic pronouns;²⁶ and here too, Caha’s possessive analysis faces nontrivial problems, given that no declinable possessive corresponds to the pronouns in table 2.4 (cf. Harðarson 2016: 9). More interesting, however, is the fact that, in Icelandic, this same syncretism pattern is also widespread among lexical nouns. Table 2.4 exemplifies this with three paradigms, each involving a different syncretic exponent: $-\emptyset$, $-i$, $-u$, as well as pronominal $-ur$. There is little doubt, here, that we are dealing with systematic syncretism, whatever the merits of Caha’s approach to pronominal genitives.

Harðarson (2016) argues, furthermore, that offending syncretisms of this sort are not uncommon across the Germanic family. He notes, for example, that a null affix may syncretically express ACC.SG and DAT.SG to the exclusion of GEN.SG, not only in Icelandic, but also in Old Norse and Faroese—and, we may add, also in Gothic, in contemporary German, and in Old English paradigms like that of *bōc* (see table 2.5).²⁷

In summary, even aside from the potential problems with Caha’s treatment of pronominal genitives, the evidence available from the whole Germanic family—and especially from Icelandic—shows that no universal order may be established between genitive and dative in Caha’s (2009) case sequence. Systematic syncretisms may *both* target ACC and GEN to the exclusion of DAT *and* target ACC and DAT to the exclusion of GEN: neither option appears to be universally ruled out by natural languages.

²⁶Other such cases may be found in Gothic (Braune and Heidermanns 2004: 132), Old Norse and Faroese (Harðarson 2016: 8), and modern German (Johnston 1996: 37–38).

²⁷But see Johnston (1996: § 2.2.6) and Caha (2009: § 8.4.4) for a different interpretation of the German facts. See also Harðarson (2016: § 3) for relevant discussion of the Faroese genitive and its diachronic development.

	Old Norse 'guest'	Faroese 'lip'	Gothic 'town'	German 'brother'
NOM	gest-r	vørr	baúrg-s	Bruder
ACC	gest	vørr	baúrg	Bruder
GEN	gest-s	(varr-ar)	baúrg-s	Bruder-s
DAT	gest	vørr	baúrg	Bruder

Table 2.5: Ø-syncretism in Germanic (Harðarson 2016: 4; Braune and Heidermanns 2004: 109; Johnston 1996: 35)

2.1.2 A problem with Latin's instrumental

A similar problem arises with the Latin instrumental (or 'ablative'; see fn. 5), as is partly acknowledged by Caha (2009: 287–290) himself. However, since his discussion is somewhat brief in this regard, I shall go over the relevant data in some more detail here.

Recall, first of all, that Caha's hypothesis is that the instrumental always contains both the genitive and the dative. Therefore, as a by-product of the spellout mechanism, no ACC-INSTR syncretism should ever exclude either GEN or DAT, whereas some ACC-GEN and ACC-DAT syncretisms may well exclude INSTR from their purview. Both predictions are corroborated by the fragments of Russian nominal morphology displayed in table 2.1.

The nominal morphology of Latin is also largely consistent with this pattern, its foremost syncretisms being NOM-ACC and DAT-INSTR—both of them permitted by Caha's theory. The DAT-INSTR syncretism is found in the singular of the 'second' declension and of part of the 'third' declension, as well as in all plural paradigms, across all declensions, genders, and case-inflected word classes (nouns, adjectives, participles, pronouns). The NOM-ACC syncretism, in contrast, is found in the plural of the 'fourth' and 'fifth' declensions, in the plural of part of the 'third' declension, and also in all neuter paradigms, across all numbers, declensions, and case-inflected word classes. Each of these syncretisms straddles several unrelated exponents, so that there is no doubt that they are both systematic and, indeed, deeply entrenched in the grammatical system of the language.

Somewhat less clear is the case of the syncretism between GEN and DAT, which characterizes the singular of the 'first' and 'fifth' declensions. This syncretism is never found in the plural, and although it purports to involve two different exponents (cf. *rosae* vs. *rei*), it turns out to involve only one if the relevant stem vowels are factored out. *Rosae* effectively results from *rosā* + *-ī* via diphthongization of *āī*, while *rei* results from *rē* + *-ī* via shortening of *e* before another vowel—two phonological processes operative across the board in the language (see Kühner and Holzweissig 1912: 405, 412–413; Niedermann 1906: 46).²⁸

²⁸Notice also that the shortening rule *vocalis ante vocalem corripitur* ('a vowel preceding another vowel gets

Nothing prevents us, however, from treating also this syncretism as systematic, given that GEN and DAT are adjacent in Caha's case sequence (1b).²⁹

	'rose' (1st decl., F)		'war' (2nd decl., N)	
NOM	rosa	rosae	bellum	bella
ACC	rosam	rosās	bellum	bella
GEN	rosae	rosārum	belli	bellōrum
DAT	rosae	rosīs	bellō	bellis
INSTR	rosā	rosīs	bellō	bellis

	'sea' (3rd decl., N)		'hand' (4th decl., F)		'thing' (5th decl., F)	
	SG	PL	SG	PL	SG	PL
NOM	mare	maria	manus	manūs	rēs	rēs
ACC	mare	maria	manum	manūs	rem	rēs
GEN	maris	marium	manūs	manuum	reī	rērum
DAT	marī	maribus	manuī	manibus	reī	rēbus
INSTR	marī	maribus	manū	manibus	rē	rēbus

None of the syncretisms reviewed so far runs counter to Caha's case sequence: NOM – ACC – GEN – DAT – INSTR. The problems start, however, when we bring into the picture two additional sets of paradigms—those of the neuters of the 'fourth' declension and those of the pronouns *ego* 'I', *tū* 'thou', and *sē*, a subject-controlled and occasionally long-distance reflexive used for all genders and numbers of the third person. The neuters of the 'fourth' declension exhibit, in the singular, a syncretism between NOM, ACC, DAT, and INSTR to the exclusion of GEN alone. The three pronouns, in contrast, instantiate a syncretism between ACC and INSTR to the exclusion of both GEN and DAT. Both paradigms thus contradict the predictions of Caha's case sequence, and converge to show that accusative and instrumental *can* well syncretize to the exclusion of other obliques.³⁰

In an attempt to explain away the problems posed by the neuter nominal paradigm, Caha (2009: 288, fn43) cites, if with a note of caution, a British source from the mid-eighteenth century, Key 1858 (p. 14), which 'gives a difference in length between the NOM/ACC and GEN/DAT [whereby Caha surely meant DAT/INSTR] ending, specifically *-u* for NOM/ACC and *ū* for DAT/INSTR.' However, this traditional belief, already controversial

shortened') does not apply when the vowel to be shortened is preceded, as well as followed, by another vowel (see Niedermann 1906: 46). So, for example, the GEN/DAT.SG of *diēs* 'day' is *diēi*, without any shortening.

²⁹In the following paradigms and discussion, I systematically ignore the so-called vocative case, which I have nothing new to say about.

³⁰These facts were probably at the root of Rasmus Rask's choice of the order NOM – ACC – INSTR – DAT – GEN 'in materials put together about this period [1812–1814] for a Latin grammar' (Allen and Brink 1980: 79).

	'horn' (4th decl., N)		1SG	2SG	REFL
	SG	PL			
NOM	cornū	cornua	ego	tū	—
ACC	cornū	cornua	mē	tē	sē
GEN	cornūs	cornuum	meī	tuī	suī
DAT	cornū	cornibus	mihi	tibi	sibi
INSTR	cornū	cornibus	mē	tē	sē

Table 2.6: Four Latin paradigms

among late Roman grammarians, has now been conclusively dismissed by the evidence furnished by classical poetry. Indeed, wherever the metrical texts from the classical period allow us to verify the quantity of *-u* in the relevant NOM/ACC.SG forms, it is always *-ū* that we invariably find; by contrast, ‘nowhere is a short vowel conclusively established’ (Kühner and Holzweissig 1912: 391). The ACC-INSTR syncretism of fourth-declension neuters is thus now safely established (cf. Ernout 1953: 64; Sihler 1995: 323; Weiss 2009: 252).³¹

As for the ACC-INSTR syncretism in pronominal paradigms, Johnston (1996: 49) correctly notes: ‘there seems to be little question that this homonymy is systematic. It is very old in Latin, and it co-extends under variations that appear in various periods and styles such as *mēd*, *mēmē* in the first singular, *tēd*, *tētē* in the second singular, and *sēsē* in the reflexive, each of them always functioning as both A [= ACC] and B [= INSTR]’. To these forms we may also add the archaic reflexive *sēd* (see Kühner and Holzweissig 1912: 579), as well as, of course, the standard forms *mē*, *tē*, and *sē*, for a total of three syncretic forms for each person. Even though this syncretism may well have *originated* as an accident,³²

³¹This received opinion has recently been called into question by Suárez Martínez (1996), who has systematically attempted to reanalyze all the putative evidence for *-ū* furnished by Augustan poetry. As convincingly shown by Levente (2002), however, Suárez Martínez tries to explain away some of the relevant attested examples by recourse to syntactically and/or semantically implausible analyses. At the same time, he also falls short of a comprehensive review of all the relevant attestations, failing to take account, in particular, of a crucial Pseudo-Ovidian passage (*Nux* 106) which provides key evidence in favor of *-ū*.

Caha (2009: 288, fn43) also notes that Key (1858: 14) marks the GEN.SG *cornūs* as reconstructed but unattested. This claim of Key’s is, however, totally unsupportable, witness Livy 44, 40, 8; Lucan 7, 217, and all the other occurrences of GEN.SG *cornūs* referenced by Kühner and Holzweissig (1912: 394). Cf. also Leumann (1977: 442) and Sihler (1995: 324).

³²Brugmann (1911: 415), for example, opines that the *-d* of the archaic accusatives *mēd*, *tēd*, and *sēd* does not have the same origin as the *-d* of the homophonous ablatives/instrumentals: rather, the accusative’s *d* would belong to an ancient reinforcing particle **ed* added to the original accusatives **mē*, **tē*, and **sē* (cf. *ecquid* ‘what? whatever?’ from **ed* + *quid* ‘what?’), whereas the other *-d* would be the standard suffix of the ablative/instrumental in early Latin (cf. e.g. *bellōd* ‘war.ABL/INSTR.SG’). Ernout (1953: 101), Sihler (1995: 379), and Weiss (2009: 326) do not disagree, but are more uncertain as to the identity of the particle involved in the accusative. A differing opinion is voiced by Friedrich Stolz, who, in Stolz and Schmalz (1900: 136), prefers to invoke ‘confusion’ (*Verwechslung*) between accusative and ablative/instrumental. For still

it is thus clear that, at least from some point on, it became entrenched in the grammatical system of the language. Crucial evidence to this effect comes not so much from the standard forms *mē*, *tē*, and *sē* (which may all be traced back to the respective *-d*-forms via regular phonological processes), but rather from the reduplicated emphatic alternants *mēmē*, *tētē*, and *sēsē*, which find no parallels either in the rest of their respective paradigms or in the other pronominal paradigms: these are the only three emphatic pronouns built via reduplication, and all three of them may function both as accusatives and as instrumentals. The coextension of this idiosyncrasy across the two syncretic cases is probably the strongest piece of evidence for the systematic nature of the syncretism pattern.

We can conclude, on these grounds, that Latin featured at least three phonologically unrelated exponents (the lengthening for fourth-declension neuters, the *-d* for archaic pronouns, and the combination of lengthening and reduplication for classical emphatic pronouns) which syncretically expressed both the accusative and the instrumental to the exclusion of the genitive,³³ and that, in two of those three cases, the unexpected syncretism also excluded the dative. On the basis of the criteria adopted so far, this ACC-INSTR syncretism thus fully qualifies as systematic, and hence constitutes another counterexample to Caha's (2009) universal hierarchy of obliques.³⁴

2.1.3 On locating the locative(s)

Before moving on to the broader implications of the problems reviewed above, let me discuss one last challenging case that I have left aside so far—what we may refer to as the locative prepositional, namely, a case that is used in combination with a preposition to

other hypotheses and critical discussion, see Sommer (1902: 442), Leumann (1977: 461–462), Schmidt (1978: 50–51), Szemerényi (1996: 214), and the references therein.

³³Caha (2009: 289–290) argues that Latin's pronominal genitives, much like their Old English counterparts, can better be analyzed as possessives of silent nouns, and are therefore irrelevant for the point at hand. In early Latin, however, the genitive forms at issue were totally unrelated to possessives, yet equally problematically non-syncretic: *mīs* in lieu of *meī*, *tīs* in lieu of *tui*, and presumably also **sīs* in lieu of *sui* (see Ernout 1953: 101).

³⁴Giovanna Marotta (p.c.) very rightly points out that a fourth-declension neuter like *cornū* and a singular personal pronoun like *ego* make up quite an odd couple from the viewpoint of subsequent Romance diachrony. On the one hand, nouns like *cornū* are among the first ones to 'have more or less given up inflection in the singular in Latin' (Weiss 2009: 252), while, on the other hand, the paradigms of personal pronouns are the last ones to have retained case in a great many Romance varieties, including e.g. French and Italian. The fact that the ACC-INSTR syncretism shows up only in two paradigm classes with such different degrees of stability may thus be taken to call for more caution in classifying this syncretism as systematic. I believe, though, that the variety of exponents involved and the diachronic stability of the relevant syncretism in pronominal paradigms leave little doubt that the pattern is real and worth capturing. In an attempt to provide further support for this idea, I will try in § 2.2.2 to further substantiate the more general claim that the instrumental may be syncretic with a core case to the exclusion of all other obliques, on the basis of extensive evidence from ergative and tripartite systems.

express static location. This case is absent from Caha’s first version of the case sequence, repeated here as (22), but is subsequently brought back into the picture toward the end of the third chapter of his thesis (Caha 2009: § 3.4).

(22) The case sequence: NOM – ACC – GEN – DAT – INSTR – COM (Caha 2009: 10)

Some indications as to where the locative prepositional may fit into the sequence are provided by Classical Armenian (Caha 2013; and 2009: ch. 7, § 8.6). Here, the locative prepositional of most lexical nouns is syncretic with both genitive and dative in the singular, and with the accusative in the plural.

	‘nation’, SG	‘nation’, PL
NOM	azg	azg-k’
ACC	azg	azg-s
LOCPREP	azg-i	azg-s
GEN	azg-i	azg-ac’
DAT	azg-i	azg-ac’
ABL	azg-ê	azg-ac’
INSTR	azg-aw	azg-awk’

Table 2.7: A Classical Armenian paradigm (Caha 2013: 1020)

Under Caha’s independently problematic assumption that the dative always contains the genitive, the only sequence that can capture the Armenian syncretisms in table 2.7 is one in which the locative prepositional intervenes between the accusative and the genitive:

(23) NOM – ACC – LOCPREP – GEN – DAT – ABL – INSTR

Pronominal paradigms, however, exhibit a systematic syncretism between LOCPREP and DAT to the exclusion of both ACC and GEN (see table 2.8). Caha (2009: § 8.6; 2013: 1019, fn5) attempts to explain away this problem by means of the same strategy he used for pronominal genitives in Old English. That is, he submits that Classical Armenian’s offending genitives may be possessives in disguise, in spite of morphological appearances to the contrary. The main problem with this view, however, is that this unexpected LOCPREP-DAT syncretism is not limited to personal pronouns, but also appears, for example, in the paradigms of the numeral *mi* ‘one’ and the interrogative adjective/pronoun *zi* ‘what/which’, neither of which could easily be expected to form the basis for a possessive. Indeed, the same syncretism shows up even in the paradigms of the standard possessive adjectives/pronouns, whose genitives should oddly be regarded, on Caha’s account, as posses-

	3SG	2PL
NOM	sa	dowk‘
ACC	sa	jez
LOCPREP	sma	jez
GEN	sora	jer
DAT	sma	jez
ABL	smanê	jênĵ
INSTR	sovaw	jewk‘

Table 2.8: Two Classical Armenian pronouns (Caha 2009: 291)

sives of possessives. As far as I can see, then, the shaded forms in table 2.9 cannot be taken to be anything but *bona fide* genitives.

	‘what, which’	‘my, mine,’ SG
NOM	zi(nč‘)	im
ACC	zi(nč‘)	im
LOCPREP	(h)im	imowm
GEN	êr	inoy
DAT	(h)im	imowm
ABL	imê	immê
INSTR	iw	imov

Table 2.9: Two Armenian singular paradigms (Schmitt 2007: 118, 124)

The problem disappears, however, if one abandons Caha’s assumption that the genitive must always be contained inside the dative—an assumption which the Germanic data in § 2.1.1 independently call into question. Absent this constraint, we can formulate the following exceptionless case sequence for Classical Armenian.

- (24) NOM – ACC – LOCPREP – DAT – GEN – ABL – INSTR

As it turns out, however, this sequence cannot be replicated for Czech—nor, once again, for Latin. Czech, too, has a case ‘restricted to locative (and some abstract) prepositional phrases’ (Caha 2009: 121), but unlike its Armenian counterpart, this case appears to correspond ‘to a separate layer which is higher than the possessor, but lower than the dative zone’ (*ibid.*) — see table 2.10.

In Latin, by contrast, the function of the locative prepositional is normally fulfilled by the instrumental, which Caha (2009; 2010; 2013) assumes to always follow both genitive and dative in the case sequence, contrary to what I have argued in § 2.1.2. This leads him to posit at least three distinct locative projections, resulting in the following enriched case

	‘teacher’, SG	‘good’, M.PL	‘book’, SG	‘both’
NOM	učitel	dobré	kniha	oba
ACC	učitele	dobré	knihu	oba
GEN	učitele	dobrých	knihy	obou
LOCPREP	učiteli	dobrých	knize	obou
DAT	učiteli	dobrým	knize	oběma
INSTR	učitelem	dobřími	knihou	oběma

Table 2.10: Four Czech paradigms from Caha 2009: 121

sequence (Caha 2009: 130).

(25) NOM – ACC – LOCPREP₁ – GEN – LOCPREP₂ – DAT – LOCPREP₃ – INSTR

Why so many locatives? Caha (2009: 125ff) answers this question by assigning each of them a different kind of semantics, based primarily on Czech contrasts like (26).

(26) Czech (Caha 2009: 129)

- a. u aut-a
at car-GEN
‘at/close to the car’
- b. v / na aut-ě
in on car-LOCPREP
‘in/on the car’
- c. před / pod aut-em
in.front.of under car-INSTR
‘in front of/under the car’

The key fact here is that not all Czech spatial prepositions take complements in the locative prepositional case. Rather, some take genitive complements, while still others take complements in the instrumental. Caha argues on these grounds that Czech provides evidence for all three locative projections: in terms of (25), the prepositional *autě* in (26b) would instantiate LOCPREP₂, the genitive *auta* in (26a) would instantiate LOCPREP₁, and the instrumental *autem* in (26c) would instantiate LOCPREP₃. Interestingly, however, these three cases are not distributed randomly. Rather, they each seem to be associated with a particular kind of semantics. ‘While in [(26a)] nothing is presupposed about the dimensionality of the car, in [(26b)] the car must be a dimensional object with either an interior (for IN), or a surface (for ON)’; finally, ‘the high instrumental layer occurs with so-called projective prepositions [...], which require that the object of the preposition

[...] is oriented with respect to a spatial axis, either an UP/DOWN axis, or a FRONT/BACK axis' (Caha 2009: 126, 129).

The problem is, however, that this correlation is not borne out cross-linguistically. In Classical Armenian, for example, the preposition *i/j* 'in' takes the locative prepositional, and neither the genitive nor the dative (see Schmitt 2007: 160), counter to Caha's predictions.³⁵ Likewise, a good number of Latin projective prepositions (*ante* 'in front of', *post* and *pone* 'behind', *infra* 'below') take the accusative rather than the instrumental. I take this as evidence that the locative prepositional too—just like the genitive, the dative, and the instrumental—simply resists assignment to a fixed dedicated position within a comprehensive complexity hierarchy of cases.

2.1.4 What might this mean?

In the previous subsections, I offered some empirical arguments against Caha's (2009) Universal Contiguity hypothesis on possible systematic case syncretisms. I shall now turn to the question of how that hypothesis might be weakened to cope with the problematic data reviewed so far.

Several possible ways to do so suggest themselves. One of them builds on the observation that many of the counterexamples reviewed so far involve pronominal paradigms (see tables 2.2, 2.4, 2.6, 2.8, and 2.9), and simply amounts to restricting the scope of the contiguity hypothesis to the morphology of lexical categories only. This does not strike me, however, as a very promising avenue, for it would leave us without anything to say about the several problems posed by lexical nouns (especially for locative prepositionals, and for genitives and datives across Germanic), while at the same time giving rise to difficult conceptual questions (first and foremost, why should pronouns be different, and why in this way?).

An alternative possibility—first mooted by McCreight and Chvany (1991) and by Johnston (1996), and recently revived by Harðarson (2016), apparently independently of the earlier two works—is offered by the following hypothesis.

(27) *Weak Case Contiguity*

For every natural language *L*, there exists at least one linear case sequence *CS* such that all the systematic case syncretisms of *L* target contiguous portions of *CS*. *CS* may not be the same for all languages.

Though some of the analyses I will propose in what follows will be problematically rel-

³⁵Caha (2009: 120) himself acknowledges this difficulty in his footnote 22, but leaves it unresolved.

evant to McCreight and Chvany's and Johnston's theorizing,³⁶ my primary focus in this thesis will never be on such purely geometric constraints as theirs, but rather on another plausible option for weakening Caha's theory—one which is, at least in principle, not incompatible with the one that McCreight and Chvany (1991) first suggested.

One of the main problems with (27) is that it tells us nothing about what a possible case sequence may look like, except that it must be linear. Weak Case Contiguity would thus allow, for example, also for a conceivable language-specific case sequence like (28), despite the fact that sequences of this sort are nowhere close to being attested.

(28) An impossible case sequence: *INSTR – NOM – COM – ACC – DAT

The idea that any simple linearity constraint like (27) would massively overgenerate was already foreshadowed by Plank (1991b), and was indeed one of the principal driving forces behind the inception of Caha's (2009) whole theory. So the question is, How can we preserve this fundamental insight—and correctly rule (28) out—without also incurring the whole lot of empirical problems presented so far?

The answer that I will explore in this thesis is actually very simple: cases can be divided up into larger case classes, and it is over these classes that the cross-linguistic generalizations on syncretism must be stated. What emerges from §§ 2.1.1–2.1.3, on this account, is simply the fact that genitive, dative, instrumental, and locative prepositional all belong to one and the same class, which we may simply refer to as OBLIQUE.³⁷ Given the universal order in (29), we therefore predict that none of these four cases may ever be systematically syncretic with the nominative without also being syncretic with the accusative,³⁸ but we make no prediction as to which obliques may be syncretic with one another, or which ones may be syncretic with the accusative.

(29) NOM – ACC – OBL

The alert reader, however, may have noticed some heterogeneity in the proposed hi-

³⁶This is especially the case for the discussion of Pama-Nyungan languages and other (partially or underlyingly) tripartite systems, which will occupy us in § 2.2.3.

³⁷I will remain agnostic about the possibility of further splitting this class into two subclasses of 'lighter obliques' (including at least genitive, dative, instrumental, simple and prepositional locatives, ablative, and partitive) and 'heavier obliques' (including at least the comitative and all the more complex non-prepositional spatial cases). The evidence bearing one way or another on this issue is woefully scant, because there are relatively few languages that have a synthetic comitative or complex non-prepositional spatial cases, and virtually no languages where the relevant cases participate in systematic syncretisms. Be that as it may, I will consistently ignore this hypothetical split in what follows, chiefly for ease of exposition.

³⁸As we will see in chapter 5, this prediction is borne out for all obliques except for the genitive, which seems to variably behave either as an unmarked case (on a par with nominative/absolute) or as an oblique (on a par with dative and locative) across different languages. I will offer some speculations about this in § 5.1.

erarchy in (29). There, while all oblique cases are collapsed into a single large container class, nominative and accusative are both referenced individually. This obviously raises the question of why nominative and accusative should be different in this respect. In the following section, I argue that they are in fact no different at all, but rather belong, respectively, to the classes of unmarked core cases (namely, nominative and absolutive) and of marked core cases (namely, accusative and ergative). To this end, however, I must now turn my attention to ergative case alignments.

2.2 Bringing ergatives into the picture

2.2.1 Ergative and accusative pattern alike

Like two of the main previous cross-linguistic surveys of case syncretism—to wit, Plank 1991b and Johnston 1996—Caha (2009) confines his attention almost exclusively to Indo-European languages, thus giving relatively short shrift to case-marking alignments other than nominative-accusative.³⁹ One of my intents in this thesis is, therefore, to finally incorporate ergative and tripartite case systems into the picture, drawing primarily on the fine typological work of Baerman, Brown, and Corbett (2005).⁴⁰ These authors identify three basic types of case syncretism.

(30) The typology of case syncretism of Baerman *et al.* (2005: 40):

Type 1: syncretism of the core grammatical cases;

Type 2: syncretism of a core case with a peripheral case;

Type 3: syncretism of peripheral cases.

Based on a sample of more than two hundred languages, of which just over fifty have partial case syncretism, Baerman *et al.* (2005: 40) report that non-Indo-European languages favor what they refer to as type 1, namely syncretism between nominative and accusative in nominative ~ accusative systems, and between ergative and absolutive in ergative ~ absolutive systems. Concerning their type 2, the three typologists also note: ‘Typically it is the marked core case which is affected, that is the accusative in a nominative ~ accusative system and the ergative in an ergative ~ absolutive system’ (Baerman *et*

³⁹A broader typological coverage is achieved in Caha 2010, whose primary focus is, however, not on case syncretism.

⁴⁰As is well known, an ergative case alignment is one whereby subjects of intransitive verbs (S) are formally marked in the same manner as objects of transitive verbs (O), but differently from subjects of transitive verbs (A); see, among many others, Dixon (1994) and Deal (2015). As for tripartite alignments, I shall focus back on them in much greater detail in § 2.2.3.

al. 2005: 49). These two generalizations are illustrated, for ergative systems, by the ABS-ERG syncretism of Basque and by the ERG-GEN syncretism of Burushaski, respectively.

	‘places’, DEF	‘places’, PROX	2PL
ABS	lekuak	lekuok	zuek
ERG	lekuek	lekuok	zuek
GEN	lekuen	lekuon	zuen

Table 2.11: Some partial plural paradigms from Basque (Hualde 2003: 173, 179)

	‘woman’	‘stone’
ABS	gus	dan
ERG	gús-e	dán-e
GEN	gús-mo	dán-e
DAT	gús-mur	dán-ar

Table 2.12: Fragments of Burushaski paradigms (Berger 1998: 63; Yoshioka 2012: 49)

Data like these, which are representative of widespread cross-linguistic tendencies, show how the behavior of the ergative case closely parallels that of the accusative, while the behavior of the absolutive closely parallels that of the nominative. In both kinds of systems, it is always the marked core grammatical case which shows the greater propensity to syncretize with obliques.

Most of the surface counterexamples to this generalization involve syncretism with the genitive case, which happens to pose very similar problems in nominative ~ accusative systems as well (see fn. 38). In § 5.1, I shall argue on these grounds that the genitive may not always be regarded as a run-of-the-mill oblique, but may sometimes behave like an unmarked case, more akin to nominative and absolutive than, say, to the dative or the instrumental.⁴¹ Genitives aside, however, the relevant counterexamples turn out to be vanishingly rare, and all arguably accidental.⁴²

Assuming this to be so, I shall thus formulate the following two sequences of cases—(31a) for accusative systems, (31b) for ergative systems—noting again that the genitive may have to be left out of both.

⁴¹Further evidence to this effect will come from patterns of overt containment and pronominal suppletion that I shall also discuss in § 5.1.

⁴²I know, in particular, of only two instances of syncretism between an absolutive and an oblique ‘across’ a non-syncretic ergative: the ABS-ABL syncretism of some Kashmiri nouns and the ABS-GEN-INSTR of Georgian attributive adjectives. As I shall argue in § 5.2, both may be regarded as instances of purely accidental homonymy, involving no covariation of the syncretic forms under allomorphy or under different conditions of cumulation. Similar (pseudo-)problems arise, in effect, also in nominative ~ accusative languages; cf. § 5.2.

- (31) a. NOM – ACC – OBL
 b. ABS – ERG – OBL

2.2.2 Further evidence against a rigid hierarchy of obliques

Before attempting to unify the case sequences for accusative and ergative systems, I would like to dwell a little longer on the question of the (dis)order of obliques, and particularly on the possibility of syncretisms between a core case and the instrumental. Recall from § 2.1.2 that my main argument for admitting this possibility at that point rested on a rather marginal nominal paradigm, and on the paradigms of three pronouns, in one single language—Latin. One might thus wonder whether so much should be made of so little evidence. Fortunately, the extension of the theory to ergative languages lends credence to a positive answer to this question.

While most accusative languages are admittedly reluctant to syncretize the instrumental with a core case (Latin being somewhat of an outlier in this respect), ergative languages provide extensive evidence for the availability of this option. Indeed, the propensity of ergative and instrumental to syncretize together is so remarkable as to be explicitly highlighted in several standard introductions to case and ergativity, including Blake's (2001: 24–28, 48ff) and Dixon's (1994: 57). The latter author writes:

Ergative is sometimes confined to marking A function [...] but in many languages this case form has a number of further functions – instrumental in Dyirbal and many other Australian languages, in North-east Caucasian languages such as Avar and Andi, in Chukotko-Kamchatkan, in a number of Papuan languages [...] and in both Classical and Modern Tibetan [...].

Here Dixon (1994: 57) is primarily talking of total syncretism, but, just a few lines below, he notes that something similar is occasionally found also for partial syncretism: '[i]n Alutor, from the Chukotko-Kamchatkan family, ergative coincides in form with locative for proper nouns and with instrumental for common nouns.' As a further example, Blake (2001: 24–28) cites Margany, a (now sadly extinct) Pama-Nyungan language, in which the ergative is very systematically syncretic with the instrumental in nominal paradigms, but is formally distinct from it in pronominal paradigms (see table 2.13).

Once all these languages are incorporated into the broader picture, the propensity of the Latin instrumental to syncretize with core cases begins to look a good deal less exceptional.

On the other hand, it is also worth noting explicitly that the instrumental is by no means the only oblique case that the ergative can syncretize with. Rather, ERG-GEN syn-

	‘stone’	‘boomerang’	2SG
NOM	bari	waŋal	inda
ACC	bari	waŋal	inaŋa
ERG	bari-ŋgu	waŋal-u	inda
INSTR	bari-ŋgu	waŋal-u	inundu

Table 2.13: Fragments of Margany paradigms (Breen 1981: 302ff)

cretisms are also very common cross-linguistically (cf. among others Rill 2016), as exemplified by the Burushaski data in table 2.12. Significantly rarer, but also attested, are the ERG-LOC syncretism of above-mentioned Alutor, and the ERG-DAT syncretism of Ubykh (Fell 2012: 5). Data like these clearly lend further support to a less restrictive approach to the hierarchization of obliques.

2.2.3 The importance of tripartite systems

So far, I have treated the ergative and the accusative case as though they appeared in complementary distribution across the world’s case-inflecting languages. As is well known, however, this is not quite true. In fact, there do exist case systems where the accusative and the ergative appear side by side: such systems, now commonly known as ‘tripartite’, mark the transitive subject (A), the direct object (O), and the intransitive subject (S) each with a different case. (From now on, I shall refer to these cases as ergative, accusative, and S-nominative, respectively.)

Very few languages, however, formally distinguish all three of these cases in all of their paradigms. As Blake (2001: 125) puts it:

Some languages employ both ergative and accusative case, though it is rare for a language to have both across-the-board ergative marking and across-the-board accusative marking, i.e. a full tripartite system with S, A and [O] distinguished for all classes of nominal. The only example reported in the literature is the Australian language Wangkumara (Breen 1976). The more common situation is for ergative or accusative or both to be lacking from some classes of nominal.

To wit, tripartite marking is hardly ever applied across the board, but may be more or less pervasive in different languages. It is still relatively pervasive, for example, in the North-American language Nez Perce, where it targets all nominals except for first- and second-person pronouns (Rude 1986). Virtually opposite is the situation of the Australian language Jingulu, where tripartite A-S-O marking only targets (some) pronouns, but is

not visible in any of the lexical nominals (Pensalfini 2003: 149ff, 177).⁴³

It is also interesting, however, to look at what goes on in the *non*-tripartite paradigms of these same languages. On the one hand, the non-tripartite pronominal paradigms in Nez Perce exhibit straight nominative ~ accusative case-marking alignment, which we may then describe in terms of syncretism between S-nominative and ergative.⁴⁴ On the other hand, lexical nominals in Jingulu exhibit straight ergative ~ absolutive alignment, which we may regard as the result of syncretism between S-nominative and accusative.⁴⁵

	'cat', SG/PL	1SG
SNOM	picpic	'iin
ERG	picpic-nim	'iin
ACC	picpic-ne	'iine

Table 2.14: Core cases in Nez Perce

	2PL	'child', SG
SNOM	kurrawala	wawa
ERG	kurrawala-rni	wawa-rni
ACC	kurrak-	wawa

Table 2.15: Core cases in Jingulu

These paradigms tell us at least two things. First, they suggest that the two separate hierarchies in (31), repeated here as (32), are both correct but both insufficient. We must attempt to unify (32a) and (32b) into a single comprehensive hierarchy which references both the ergative and the accusative at the same time.

- (32) a. NOM – ACC – OBL
b. ABS – ERG – OBL

Second, these paradigms show that no such unification may possibly result in a universal linear sequence of cases. As it happens, there exist both syncretisms between SNOM and ERG to the exclusion of ACC (in Nez Perce), and syncretisms between SNOM and ACC to the exclusion of ERG (in Jingulu), such that no universal order between ERG and ACC may be devised. Rather, just as we have seen for obliques, it will be necessary to restate the relevant generalizations in terms of case classes, along the lines of (33).

- (33) {NOM, ABS, SNOM} – {ACC, ERG} – {DAT, LOC, INSTR...}

Significantly, some evidence for such non-linearity can also be found within individual languages. An interesting case in point is another Pama-Nyungan language, Djapu,

⁴³I shall come back to the morphology of Jingulu pronouns in § 4.1, where they will furnish a suppletion-based argument against treating the ergative as an inherent case.

⁴⁴The paradigm of *picpic* is based on occurrences in Deal (2010: 3, 32, *et passim*). The 1SG paradigm is from Deal (2010: 34). Here I abstract away from further complications of Nez Perce, such as the systematic syncretism between ERG and GEN or the apparent contingency of tripartite case-marking on object agreement.

⁴⁵The paradigm of *wawa* is based on occurrences in Pensalfini (2003: 46, 66, 83). The 2PL paradigm is from Pensalfini (2003: 149–152). I am abstracting away from the systematic syncretism of ACC and GEN.

described by Morphy (1983). Djapu combines in its morphology all three of the case-marking alignments discussed so far — ergative~absolutive, nominative~accusative, and tripartite. More specifically, the Djapu nouns with human or higher animate referents, which Morphy refers to as [+HU], display tripartite A-S-O marking; [-HU] nouns have an ergative ~ absolutive pattern; whereas pronouns have a nominative ~ accusative pattern (cf. Morphy 1983: 33, 37, 51).⁴⁶

	‘person,’ [+HU]	‘horn,’ [-HU]	2SG
SNOM	yolŋu	dhandurruŋ	nhe
ACC	yolŋu-y	dhandurruŋ	nhuna
ERG	yolŋu-n	dhandurruŋ-dhu	nhe

Table 2.16: Core case-marking in Djapu

A very similar situation obtains in Yidiñ—as well as in Diyari, aside from minor differences in the partition of nominal classes (cf. Goddard 1982: 170–175). In all of these languages, the simplest way to account for the full range of paradigms is clearly to posit an underlying tripartite case pattern across the board. This pattern would then be most often obscured either by SNOM-ACC syncretism (as in the ‘ergative’ paradigm of ‘horn’) or by SNOM-ERG syncretism (as in the ‘accusative’ paradigm of pronouns).

Further support for such an analysis comes from case-concord phenomena, as has been noted by Melčuk (1979: 54), Goddard (1982: 172–181), Blake (2001: 25–26), and Legate (2008: 75–81). Legate (2008: 77) notes, for example, that ‘[i]n Djapu, all elements of a DP, whether continuous or discontinuous, must be marked for case and match in case.’ This straightforward generalization could not be maintained, though, if one were to analyze the demonstrative’s SNOM-ACC form *dhuwa* as an absolutive and the pronominal SNOM-ERG *nhe* as a nominative.

(34) Djapu (Morphy 1983: 84)

Dhuwa *nhe* yurru lili dha:parng rongiyi-rr.
 this.ABS you.NOM FUT hither unsuccessful return-UNMARKED
 ‘YOU will return empty-handed [but not I].’

No such problem arises, by contrast, if we analyze these forms as two S-nominatives involved in different syncretism patterns.

This same argument can be (and has been) used to advocate a similar analysis also

⁴⁶The paradigm of *yolŋu* is based on occurrences in Morphy (1983: 84, 110). The paradigm of *dhandurruŋ* is likewise based on Morphy (1983: 127, 162). The 2SG paradigm is from Morphy (1983: 51).

for languages whose S-nominative is thus a *non-autonomous*⁴⁷ case—that is, languages where the S-nominative is *always* syncretic with some other core case, but, crucially, not always with the same case. Among the other Pama-Nyungan languages, for example, ‘it is common to find the case-marking schema shown in [table 2.17]’ (Blake 2001: 24).

	NOUNS	PRONOUNS
ACC	-∅	-nya, -nha
SNOM	-∅	-∅
ERG	-lu, -ngku, -Tu ⁴⁸	-∅
INSTR	-lu, -ngku, -Tu	—

Table 2.17: Core case-marking in Pama-Nyungan (Blake 2001: 24)

Here too, we can make sense of both types of paradigms by positing an underlying tripartite pattern, systematically concealed by either SNOM-ACC or SNOM-ERG syncretism. Once again, doing otherwise, for example by saying that ‘the typical Australian language has two case systems that co-exist within it’ (Goddard 1982: 167), would prevent us from stating simple generalizations about case concord. Any non-tripartite analysis would be at a loss, for example, to explain the origin of ergative marking on the phrase *tjitji tjukutjuku-ngku* ‘small child’ in examples like the following.

(35) Pitjantjatjara (Bowe 1990: 49)

Paluru=rni [tjitji tjukutjuku] -ngku nya-ngu
 3SG.NOM=1SG.ACC child small -ERG see-PAST
 ‘He, a small child, saw me.’

In his classic textbook, Blake (2001: 26) glosses this Pitjantjatjara example as follows:

Paluru is a pronoun in A function. The form is the same as would be used for S function [...]. The bound pronoun *-rni* [...] encodes the direct object (glossed as O). The phrase *tjitji tjukutjuku-ngku* bears the ergative case marker indicating that it is to be construed with *paluru* rather than *-rni*. The simplest way to account for this ergative marking is via a general rule for predicative nominals to the effect that they agree in case with their controller, but such a rule cannot be stated if it is dependent on whether the controller bears particular case marking.

⁴⁷This term owes most of its popularity to B.J. Blake’s classic introductions to case, but actually originates in Soviet Russia. Blake (2001: 22) attributes it to Melčuk (1986: 66), who ascribes it in turn to Andrej A. Zaliznjak (1973), ‘O ponimanii termina ‘padež’ v lingvističeskix opisanijax’, in A.A. Zaliznjak, ed., *Problemy grammatičeskogo modelirovanija*, pp. 53–87 (Moskva: Nauka). My unfamiliarity with Russian has prevented me from double-checking the original source.

⁴⁸‘T stands for an alveolar stop which may assimilate to match a dental, retroflex or palatal stem-final consonant’ (Blake 2001: 24).

I therefore take it as established that there exist languages that combine SNOM-ERG and SNOM-ACC syncretisms (e.g. Djapu, Diyari, etc.), and even languages that do this so systematically as to render the S-nominative non-autonomous (e.g. Pitjantjatjara, Margany, etc.). The question now is: What do these languages tell us, in addition to what we already knew from tripartite languages that conflate either only SNOM-ERG (e.g. Nez Perce) or only SNOM-ACC (e.g. Jingulu)?

First of all, double-syncretism languages like Djapu and Pitjantjatjara highlight a hitherto unforeseen tension between Johnston's (1996) proposal and my own. Recall from § 2.1.4 that Johnston's theory of case syncretism—ultimately traceable to McCreight and Chvany (1991)—revolves around the Weak Case Contiguity hypothesis in (27), repeated here as (36).

(36) *Weak Case Contiguity*

For every natural language *L*, there exists at least one linear case sequence *CS* such that all the systematic case syncretisms of *L* target contiguous portions of *CS*. *CS* may not be the same for all languages.

As it turns out, the only way to reconcile Djapu's or Pitjantjatjara's case systems with (36) would be to posit language-specific case sequences starting with ACC. Djapu's case sequence, for example, should be something along the lines of (37).

(37) ACC – SNOM – ERG – INSTR

If construed as a containment hierarchy (along lines first sketched by Johnston himself), (37) would entail that the accusative be the 'lightest' of Djapu cases, counter to my current claim that the accusative always contains at least one other less marked case. This is the tension that I anticipated above: as it turns out, Johnston's (1996) quest for geometric constraints on case syncretisms and my own quest for substantive constraints thereon do not always lead to the same results. In this case, however, I think the two approaches are not equally well supported empirically. Rather, there seems to be fairly clear morphological evidence that the case containment relations in these Pama-Nyungan languages are actually those that my current theory predicts, rather than those predicted by Johnston: the 'lightest', least marked case is not the accusative but always the S-nominative, which is contained both in the ergative and in the accusative itself. Indeed, whenever a paradigm provides transparent indications of a containment relation between SNOM and ACC, or between SNOM and ERG, it is always SNOM that turns out to be contained in the other case, and never the other way around.

	Nez Perce 'cat'	Djapu 'person'	Mamu ⁴⁹ 1DU
SNOM	picpic	yolŋu	ŋali
ACC	picpic- ne	yolŋu- y	ŋali- na
ERG	picpic- nin	yolŋu- n	ŋali

Table 2.18: Overt containment relation between ACC and SNOM

The linearizability constraint imposed by Johnston on every language's case syncretisms thus forces us to posit case sequences that contradict whatever other evidence we have for containment relations among cases. The constraint can therefore be maintained only at the cost of depriving the sequences so obtained of any informative value concerning the decomposition of case features. In other words, Johnston's theory is either incorrect or largely irrelevant to my current concerns.

Finally, a note on one last question that I have left unresolved as yet. So far, I have not been very clear about the featural make-up of obliques across languages. Indeed, all that I have said about obliques is that they must contain (the features of) the ergative in ergative ~ absolutive systems, and (the features of) the accusative in nominative ~ accusative systems. But what happens in those systems that have both an ergative and an accusative? Does an oblique always have to contain both, or may it as well contain only one?

The case system of the Indo-Aryan language Palula (Liljegren 2016) provides an empirical argument for the latter, weaker alternative. As Baerman *et al.* (2005: 15) correctly point out, we may distinguish at least five cases in this language: S-nominative, accusative, ergative, genitive, and prepositional. 'However [...], no paradigm displays more than four distinct case forms' (*ibid.*), as can be verified in the following paradigms.

	'man', SG	'field', PL	2 PL	2 SG
SNOM	míiš	çhíitr-a	tus	tu
ACC	míiš	çhíitr-a	tus-aám	tu
ERG	míiš-a	çhíitr-am	tus-ím	thíi
PREP	míiš-a	çhíitr-am	tus-aám	tu
GEN	míiš-i	çhíitr-amii	tus-íi	thíi

Table 2.19: Three Palula paradigms (adapted from Liljegren 2016: 45, 97, 126)

The case of interest here is, of course, the prepositional. This case syncretizes with the ergative in nominal paradigms and with the accusative in plural pronominal paradigms—two syncretisms that are both consistent with the class-based containment hierarchy I am

⁴⁹Data from Mamu (a dialect of the Australian language Dyirbal) are from Dixon (1972: 50).

advocating, repeated here as (38).

$$(38) \quad \{\text{NOM, ABS, SNOM}\} - \{\text{ACC, ERG}\} - \{\text{DAT, LOC, PREP...}\}$$

A more interesting pattern, however, is the one represented by singular pronouns, each of which displays only two case forms: one for SNOM, ACC, and PREP, and another for ERG and GEN. This means that the S-nominative and an oblique (PREP) can participate in a syncretism that is ‘mediated’, so to speak, by only one of the two marked core cases (ACC), to the exclusion of the other one (ERG). The ergative is allowed, in other words, not to take part in the SNOM-PREP syncretism as long as the accusative does.

This should be impossible if all obliques had obligatorily to contain both ERG and ACC. This impossibility follows from the standard reasons given in § 1.4. In a realizational morphology regulated by the Elsewhere Principle, if $\text{PREP} \supset \text{ERG} \supset \text{SNOM}$, then no single form may qualify as the best match for Spellout of both PREP and SNOM without also being the best match for ERG—and this regardless of whether one opts for Halle’s Subset Principle or Starke’s Superset Principle. Palula’s PREP-ACC-SNOM syncretism ‘across’ a non-syncretic ergative thus leads us to reject the stronger hypothesis, and to conclude that obliques in tripartite systems need only contain either ERG or ACC, not necessarily both.

2.3 Interim conclusions

In this chapter, I have reassessed some of the evidence that motivated Caha’s (2009) proposal, and have integrated it with some new evidence both from accusative systems (especially Harðarson’s West Nordic languages) and, more importantly, from ergative and tripartite systems. On the basis of this larger body of evidence, I have concluded that most of the subgeneralizations entailed by Caha’s (2009) Universal Case Contiguity actually do not hold up.

(39) *Universal Case Contiguity* (Caha 2009: 10)

- a. Non-accidental case syncretisms can only target contiguous regions of a linear case sequence invariant across languages.
- b. The case sequence: NOM – ACC – GEN – DAT – INSTR – COM

Rather, there is only one empirically warranted generalization, which I restate in (40).

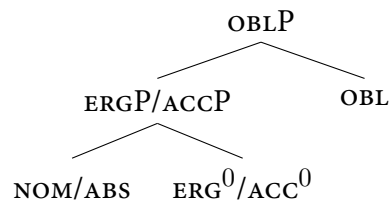
- (40) If a non-accidental syncretism covers an unmarked core case (nominative, absolutive, or S-nominative) and an oblique case (dative, locative, instrumental, prepositional), it must also cover a marked core case (ergative or accusative).

In order to capture this state of affairs in terms of containment relations, we will then have to stipulate the following hypotheses:

- (41) a. Every oblique case contains a marked core case in its representation.
b. Every marked core case contains an unmarked core case in its representation.

This in turn leads us to posit a significantly impoverished cartography by comparison to that of Caha (2009). The only case shells we may posit are one for the unmarked case,⁵⁰ one for the marked core case (which may be specified either as [ERG] or as [ACC]), and then either one oblique case shell susceptible of various featural specifications or multiple oblique case shells arranged in flexible order.⁵¹

(42)



This is the simplified cartography which I will use, at least for concreteness, throughout the following chapters.

⁵⁰If this is, indeed, anything more than the mere lack of any case (cf., among others, Bittner and Hale 1996 and Kornfilt and Preminger 2015).

⁵¹This hierarchy has already been proposed by Bobaljik (2015) and Smith *et al.* (2016) on the basis of suppletion patterns which I shall review in fair detail in § 3.1. Earlier still, the same hierarchy was also proposed by Demirok (2013: 103–104) in order to derive Bobaljik’s (2008: 303) ‘accessibility/markedness hierarchy for morphological agreement’—the idea that if, in some language, the verb agrees with anything, it agrees with some or all of the noun phrases bearing unmarked case, and if the verb agrees with anything other than unmarked noun phrases, it agrees with some or all of the noun phrases bearing dependent case (cf. also Caha 2009: § 9.1). The problem here lies not in Demirok’s explanation but rather in his *explanandum*, since there have now emerged a number of potential counterexamples to Bobaljik’s original generalizations (see especially the discussion of Semelai and Coast Tsimshian in Baker 2015: 65ff). That is why I have preferred not to include agreement in chapter 3 among the principal sources of converging evidence for (42).

[Many thanks to Hagen Blix for bringing Ö.F. Demirok’s work to my attention.]

Chapter 3

Further evidence for a (coarser) case containment hierarchy

3.1 Case-based pronominal suppletion

The term ‘suppletion’ refers to the contextually conditioned ‘wholesale replacement of a stem with a phonologically unrelated stem’ (Bobaljik 2012: 1), such as in the English alternation between nominative *she* and accusative *her*, to be contrasted with the non-suppletive pair *they* ~ *them*. As this example shows, alternations of this sort can evidently be conditioned by case alone, with number and all other ϕ -features remaining constant. This occurs, however, almost exclusively in pronominal paradigms, and never (or hardly ever) in paradigms of lexical nouns. Whatever the reason for this state of affairs,⁵² this is why the most thorough typological investigation of case suppletion to date, due to Smith *et al.* (2016), is deliberately concerned with pronouns only.⁵³

Before delving into a more detailed summary of this study, let me just outline right away what its main findings are, and how they dovetail with the other results attained so far on the basis of syncretism patterns. Just as I have argued that an unmarked core case and an oblique cannot share the same realization (i.e. be syncretic) without also sharing it with a marked core case, so Smith *et al.* find that an unmarked core case and an oblique cannot share the same *stem* without also sharing it with a marked core case. In other terms, if the relevant cases are ordered as ‘UNMARKED CORE – MARKED CORE – OBLIQUE’, the pattern that is never attested either in syncretism or in suppletion is the one we may

⁵²See Moskal (2015) for a recent proposal, and Haspelmath (2016) for skeptical discussion.

⁵³Another good reason for this is that pronominal paradigms—especially suppletive ones—are generally well detailed in descriptive grammars, such that it is relatively easy ‘to construct a large enough sample that gaps may be significant’ (Smith *et al.* 2016: 8).

characterize, following Bobaljik (2012), as the *ABA pattern, in which the accusative (or the ergative) is the only odd one out, while the nominative (or the absolutive) and the oblique share the same formal property together—be it the same suppletive stem or the same whole word form.

It is comforting to notice, furthermore, that case-driven suppletion not only conforms to the same abstract pattern we have found for syncretism, but also makes reference to the very same sequence of case classes. This sequence is repeated in (43).

(43) {NOM, ABS, SNOM} – {ACC, ERG} – {DAT, LOC, INSTR, ...}

Suppletion patterns provide support, in particular, for both the modifications to Caha's case hierarchy that I argued for in the previous chapter. Firstly, '[a]lthough consistent with a richer hierarchy, the pronominal suppletion patterns provide no compelling evidence of a richer structure than that in [(44)]' (Bobaljik 2015: 9), in accord with the idea, defended in § 2.1, that no universal hierarchy may possibly be established among the various obliques.

(44) [[[UNMARKED] ACCUSATIVE/ERGATIVE] OBLIQUE]

Secondly, suppletion patterns consistently treat the ergative on a par with the accusative—namely, as a paradigmatic intervener between the unmarked core case and the oblique cases. This finding fits neatly with the point I made in § 2.2.1, to the effect that ergative and accusative pattern alike also with respect to syncretism, jointly constituting a 'middle class' in the three-fold hierarchy (43)/(44). The match between syncretism and suppletion patterns thus appears to be perfect.

Now that this major result has been brought out, I move on to a closer look at the data that led to it and at the theoretical accounts it led to. These are the topics dealt with, respectively, in § 3.1.1 and § 3.1.2.

3.1.1 Smith *et al.*'s (2016) generalization

Smith *et al.* (2016) took as their starting point an initial sample of 160 widely diverse languages. However, 76 of these languages had no suppletion for case, while another 19 'had suppletion for case, but had only a two-way case-contrast, and hence are uninformative about the key prediction: they lack ABA patterns, but trivially so, as there is no third category to investigate' (p. 10). This leaves 65 informative languages, belonging to all three of the alignment systems that I have been discussing so far—accusative, ergative, and tripartite. In the interest of ensuring commensurability among these languages, Smith *et al.*

(2016) restrict their attention to core case-marking (absolute, (S-)nominative, ergative, accusative) and to one representative oblique case per language (typically the dative). In addition to this core data set, the authors occasionally extend their focus to additional oblique cases (or to additional languages) whenever a paradigm (or a language family) appears to deserve particular attention. ‘Throughout the study,’ however, they systematically ‘set aside the genitive case, as available sources do not consistently distinguish a genitive case (relevant to the case hierarchy) from possessive pronouns (which are not part of the hierarchy)’ (Smith *et al.* 2016: 11).⁵⁴

The key generalization that emerges from this sample revolves, to repeat, around the systematic absence of *ABA patterns, in which the nominative/absolute and an oblique would share the same stem to the exclusion of the accusative/ergative. In other words, if the accusative (or the ergative) suppletes with respect to the nominative (or the absolute), then so must all the oblique cases in the paradigm.

This typological gap appears to be all the more significant in view of the fact that, apart from *ABA, all the other possible patterns in (45) turn out to be attested, and that, except for the exceedingly rare ABC pattern, all of them are robustly attested in both accusative and ergative systems.

- (45) a. Widely attested: AAA, ABB, AAB;
 b. Very rare: ABC
 c. Unattested: *ABA

In the remainder of this subsection, I shall provide some relevant instances of each pattern, drawing on Germanic, Caucasian, and Australian material. I shall begin with the most trivial and most widespread of all patterns—the non-suppletive AAA. This can be exemplified by accusative Icelandic and ergative Lezgian, among a great many others.

Icelandic, 2SG		Lezgian, 1SG	
NOM	þú	ABS	zun
ACC	þig	ERG	za
DAT	þér	DAT	zaz

Table 3.1: AAA patterns (Einarsson 1949: 68; Haspelmath 1993: 184)

The less widespread but still widely attested ABB pattern is exemplified here by ergative

⁵⁴As a consequence of this choice, even languages with more than two cases have sometimes had to be disregarded, whenever their only oblique case happened to be the genitive. One such example, featuring only a NOM ~ ACC ~ GEN tripartition, is Modern Greek.

Georgian and, again, accusative Icelandic.⁵⁵ Here, it is the unmarked core case that acts as the odd one out, while all more marked cases share the same suppletive stem together.

Icelandic, 1SG		Georgian, 3SG	
NOM	ég	ABS	is
ACC	mig	ERG	man
DAT	mér	DAT	masa

Table 3.2: ABB patterns (Smith *et al.* 2016: 14–15)

Somewhat less widespread, but again attested across alignment types, is the pattern AAB, ‘where suppletion seems to target the “third” case in a pronoun’s paradigm, rather than suppling immediately from the second case onwards’ (Smith *et al.* 2016: 18). This is illustrated here with examples from German, Mangarayi, and Wardaman (the latter two from Australia’s Northern Territory).⁵⁶

German, 3SG.F		Mangarayi, 2SG		Wardaman, 3SG	
NOM	sie	NOM	ñangi	ABS	narnaj
ACC	sie	ACC	ñan	ERG	narnaj-(j)i
DAT	ihr	DAT	ɲanga	DAT	gunga

Table 3.3: AAB patterns (Merlan 1982: 102; Merlan 1994: 112–114)

Finally, there is one more pattern that is possibly attested—the one that involves three distinct suppletive stems, one for each case class. This ABC pattern is relatively rare but still certainly attested in other suppletion domains, such as adjectival comparative morphology, a famous example being the Latin triple *bonus ~ melior ~ optimus* for ‘good ~ better ~ best’.⁵⁷ Apparently, this pattern shows up even more rarely in the domain of case-driven suppletion. Indeed, Smith *et al.* (2016) find only one good candidate to an ABC pattern in their entire sample.⁵⁸ That is the 1SG paradigm of the East-Caucasian

⁵⁵The Icelandic case-suppletion pattern reported in table 3.2 also characterizes the 1SG paradigms of many other ancient and modern Indo-European languages, and can indeed be traced back to the reconstructed proto-language, albeit the exact preforms involved still remain extremely controversial. Cf. Katz (1998: 21ff) and references therein. It is curious, yet devoid of any historical significance, that an accidentally very similar suppletion pattern also obtains in Georgian, albeit in the 3SG paradigm rather than in 1SG.

⁵⁶Here I represent Mangarayi’s pronominal paradigms as nominative ~ accusative, but in fact the language instantiates a variant of the typical Australian alignment split, with accusative alignment in pronominal and non-neuter nominal paradigms, and ergative alignment in neuter paradigms (Merlan 1982: 52). In view of the classic arguments I reviewed in § 2.2.3, Mangarayi’s system is therefore better analyzed as a tripartite case system featuring a non-autonomous S-nominative, but I will abstract away from this here.

⁵⁷Other such examples from within Indo-European can be found in Bobaljik (2012: 29).

⁵⁸Khinalug apart, there might be further ABC patterns in the pronominal paradigms of the East-Sudanic language Murle (Arensen 1982: 89ff), but Moskal *et al.* (2015) raise the reasonable suspicion that those may actually be either AAB or even AAA.

language Khinalug.⁵⁹

	ABS	ERG	DAT
1SG	zi	yä	ás(ir)

Table 3.4: An ABC pattern from Khinalug (Kibrik *et al.* 1972: 73)

3.1.2 A containment-based account

Given how close a match there is between suppletion patterns and syncretism patterns, it comes as no surprise that they also lend themselves to very similar kinds of account. And indeed, Bobaljik (2015) and Smith *et al.* (2016) explicitly attempt to explain their findings in terms of the same containment hierarchy I have used to rule out unattested syncretisms.

(46) [[[UNMARKED] ACCUSATIVE/ERGATIVE] OBLIQUE]

Smith *et al.*'s (2016) account of the *ABA gap in case suppletion closely parallels Bobaljik's (2012) account of an analogous gap in suppletive comparative morphology (if the comparative suppletes with respect to the positive, then so must the synthetic relative superlative). Both accounts are couched within the framework of Distributed Morphology (Halle and Marantz 1993; Embick and Noyer 2007), and hinge, more specifically, on much the same three kinds of assumptions that I invoked in § 1.4, in my attempt to translate Caha's theory in more DM-compatible terms. These assumptions include:

- (i) *realizational morphology*: 'the (abstract) morpho-syntactic representation is the input to a morphological component, characterized in part by rules of exponence (vocabulary insertion) which assign phonological realizations to the terminal nodes' (Bobaljik 2012: 5);

⁵⁹Schulze (2008: 732) seems to imply that Khinalug's 1SG paradigm should in fact be regarded as an unexpected ABA pattern, as he argues that '[t]he same root [of ABS *zi*] is found in the dative *ás(ir)* < **ázir* (post-tonic devoicing), whereas the ergative/genitive stem *y-* is a Khinalug innovation.' There are good reasons, however, to think that this might not be the correct account on either the synchronic or the diachronic level. On the synchronic side, Kibrik *et al.* (1972: 72) say that '[n]ominative, ergative, and dative are expressed by suppletive forms that cannot be decomposed into stem and ending', ostensibly implying an ABC analysis. On the other hand, diachronically, the online version of Nikolayev and Starostin's (1994) *North Caucasian etymological dictionary*, s.v. *zi*, rather suggests treating the triple as an etymological AAA: 'It is most probable that the erg. form was **jāz* [...] with loss of the rare Auslaut *-z (dat. *as* should be treated as **az* + dat. marker *-s).' In sum, there does not seem to exist any *ABA pattern in either contemporary Khinalug's or reconstructed proto-Khinalug's grammar.

[My deepest thanks to Maria Polinsky for translating the relevant bits of Kibrik *et al.*'s (1972) grammar for me, and for referring me to Nikolayev and Starostin's (1994) etymological dictionary.]

- (ii) *underspecification & the Elsewhere Principle*: ‘the rules of exponence (vocabulary insertion) may be underspecified, and thus may compete to realize a given node; such competition is resolved by the Elsewhere Condition [...] ordering, in which more specific rules take precedence over more general ones’;
- (iii) *a domain-specific containment hypothesis*: something like (46) for cases, and something like [[[POSITIVE] CMPR⁰] SPRL⁰] for degrees of comparison.

How can the three ingredients (i)–(iii) contribute to deriving the impossibility of an ABA pattern like pseudo-Icelandic *þú ~ *hig ~ þér*, or pseudo-Georgian *is ~ man ~ *issa*? Bobaljik’s (2012; 2015) reasoning to this effect is very similar to that laid out in § 1.4 about impossible ABA syncretisms. In order to force the morphology to produce an ABA pattern of case suppletion, one needs to assign two phonologically unrelated stems to the two core cases. The standard way to do this within Distributed Morphology is in terms of contextual allomorphy (cf. Bonet and Harbour 2012). In (47), I take Georgian as a representative example.

- (47) a. 3SG → *ma-* / ____] ERG⁰
 b. 3SG → *is*

In order to surface only in the relatively marked ergative case, the allomorph *ma-* has to be explicitly restricted to that context in its rule of exponence (47a). The allomorph *is*, by contrast, need not be specifically associated to any context. In a framework like DM, which countenances underspecification but not overspecification, *ma-* is thus not an eligible candidate for realizing 3SG in the absolutive, since that context does not contain any ERG head and hence does not meet *ma-*’s structural description. By contrast, the underspecified allomorph *is* does qualify as a competitor for realizing 3SG both in the absolutive and in the ergative, but it is not selected in the latter context because there the Elsewhere Principle gives precedence to the more specific allomorph *ma-*.

Let us now turn to the dative. Within the current system, we can only do one of two things with it—either say no more about it, or stipulate a third exponence rule specifically dedicated to it. In the latter scenario, we end up with an (extremely marked) ABC pattern. By contrast, if we assume no exponence rules other than (47), the two allomorphs *is* and *ma-* should then compete to realize the dative 3SG, under the assumption that the dative contains ERG⁰; but then the Elsewhere Principle should let the more specific *ma-* win out again, thus resulting in an ABB pattern. The unattested *ABA pattern is thus excluded in

principle.⁶⁰

In sum, Smith *et al.*'s (2016) system allows for the derivation of all conceivable patterns (AAA, AAB, ABB, ABC) except for the illicit *ABA, the only one which is apparently unattested.⁶¹ Here are some sample accounts of each of the attested patterns.

- (48) AAA — Icelandic *þú ~ þig ~ þér*
2SG → *þ-*
- (49) AAB — Wardaman *narnaj ~ narnaj-(j)i ~ gunga*
3SG → *gunga* / *___*] ERG⁰] DAT⁰
3SG → *narnaj*
- (50) ABB — Georgian *is ~ man ~ masa*
3SG → *ma-* / *___*] ERG⁰
3SG → *is*
- (51) ABC — Khinalug *zi ~ yä ~ ás(-ir)*
1SG → *as-* / *___*] ERG⁰] DAT⁰
1SG → *yä* / *___*] ERG⁰
1SG → *zi*

3.2 Case-based stem allomorphy in non-pronominal paradigms

Much the same kind of reasoning can also be applied to case-based stem allomorphy outside of pronominal paradigms, as suggested in recent work by McFadden (2014). There, '[t]he term *stem* is used descriptively, on the one hand to contrast with root suppletion of the *go/went* type, and on the other to make it clear that the allomorphy involves a sub-part of the noun base rather than the case (or number) suffix.' We are talking, in other words, of morphological alternations like those between *maram* and *maratt-* or between *ihminen* and *ihmise-* in table 3.5.

⁶⁰Notice, however, that '[t]he notation in examples [(51)] could in principle be exploited to characterize a surface ABA pattern as an underlying ABC pattern, but one in which the [oblique] and default allomorphs are (accidentally) homophonous' (Bobaljik 2015: 5fn4).

⁶¹For an alternative account of the *ABA pattern within Nanosyntax—a framework which aspires to dispose of allomorphy rules altogether—see Starke (2014) and especially Caha (n.d.).

Tamil, ‘tree’		Finnish, ‘person’	
NOM	maram	NOM	ihmi-nen
ACC	maratt-ai	GEN	ihmi-se-n
DAT	maratt-ukku	PART	ihmi-s-tä
INSTR	maratt-aale	INES	ihmi-se-ssä

Table 3.5: Two examples of case-based stem allomorphy (McFadden 2014)

As we move away from wholesale root suppletion, it also becomes more difficult ‘to distinguish stem allomorphy from phonological readjustment of stem material’ (McFadden 2014), especially in view of the common assumption that phonological readjustment rules need not necessarily apply across the board but may as well be restricted to certain morphosyntactic environments. McFadden convincingly argues, though, that drawing such a distinction is nonetheless crucial for recognizing the relevant *ABA pattern in this domain. To this end, he uses the following leading criterion: phonological readjustment can be easily characterized in phonological terms with respect to both its triggers and its effects, while genuine allomorphy typically cannot. McFadden illustrates this with the following alternations: the first, from Tamil, is produced by *v*-epenthesis for purposes of hiatus avoidance; the second one, from Finnish, results from an essentially phonological alternation between a voiced and a voiceless variant (the former in the onset of a closed syllable, the latter in the onset of an open one).

Tamil, ‘mother’		Finnish, ‘street’	
NOM	ammaa	NOM	katu
ACC	ammaav-ai	GEN	kadu-n
DAT	ammaav-ukku	PART	katu-a
INSTR	ammaav-aale	INES	kadu-ssa

Once we factor out phonological alternations like these, the expected *ABA pattern finally emerges. More specifically, McFadden (2014) proposes the following generalization:

(52) Nominative Stem-Allomorphy Generalization⁶²

When there is stem allomorphy based on case, it distinguishes the nominative (along with any cases systematically syncretic with the nominative) from all other cases.

⁶²This is based on in-depth analysis of the following ten languages: Tamil, Northern Saami, Finnish, Estonian, Icelandic, Gothic, Latin, Sanskrit, Ancient Greek, and Russian. It is worth noting that this sample, though admittedly small, still straddles at least three branches of Indo-European, as well as two non-Indo-European families. McFadden claims, furthermore, that he ‘actually considered a far larger number of languages, but found that the vast majority have no relevant stem alternations.’

To wit, if we follow again the established order NOM-ACC-OBL, the possible patterns of case-based stem allomorphy in nominals include ABB (as in table 3.5) and AAB (to be discussed shortly), but not the theoretically excluded *ABA. These findings thus largely coincide with those of Smith *et al.* (2016) from pronominal suppletion, except for two purported differences.

One of these revolves around ABC patterns, which seem to be very occasionally attested in pronouns (recall the Khinalug paradigm in table 3.4), but are nowhere to be found in nouns: most likely, however, this asymmetry simply stems from the fact that multiple idiosyncratic alternations within a single paradigm are inherently costly, and hence functionally tolerable only in high-frequency paradigms such as pronominal ones—qualifying, furthermore, as an extremely marked option even in that domain.

The other difference between the findings of McFadden (2014) and Smith *et al.* (2016) is, by contrast, potentially more interesting, and revolves around the possibility of having an AAB pattern. Recall from the last section that Smith *et al.* found a good deal of such patterns in pronouns, and that while many of them involved full syncretism between the two As (cf. table 3.6), many others did not (cf. table 3.7).

German, 3SG.F		Krongo, 1SG		Archi, 2SG	
NOM	sie	NOM	àʔàŋ	ABS	un
ACC	sie	ACC	àʔàŋ	ERG	un
DAT	ihr	LOC	kàtí	DAT	was

Table 3.6: AAB patterns involving syncretism (Smith *et al.* 2016: 18–19)

Khinalug, 2SG.F		Wardaman, 3SG		Mangarayi, 2SG	
ABS	vî	ABS	narnaj	NOM	ñan̄gi
ERG	va	ERG	narnaj-(j)i	ACC	ñan
DAT	úx̂(ir)	DAT	gunga	DAT	ŋan̄ga

Table 3.7: AAB patterns without syncretism (Smith *et al.* 2016: 20; Merlan 1982: 102)

By contrast, all the AAB patterns in McFadden’s (2014) sample turned out to involve systematic syncretism between nominative and accusative, the clearest examples being found in classical languages.

Latin, ‘journey’		Greek, ‘dream’	
NOM	iter	NOM	ónar
ACC	iter	ACC	ónar
DAT	itiner-ī	DAT	onéirat-i

This ostensibly systematic pattern led McFadden to word his generalization the way he did in (52), and to surmise that such AAB patterns in nominal paradigms might in fact be merely epiphenomenal, insofar as the putative accusatives involved might in fact be taken to be nominatives.⁶³

The empirical robustness of the relevant pattern is, however, somewhat dubious, in that it crucially relies on a dismissal of Ancient Greek *Ablaut* as an essentially phonological alternation. McFadden seems confident that this can be done by ‘hav[ing] the exponence of certain nodes consist of or include floating, underspecified material (morae, C or V slots, etc.) that is accommodated autosegmentally’; but he cannot point to any worked-out proposal to this effect, and there are in fact good reasons to doubt that any such proposal would be feasible.⁶⁴ If we thus remain skeptical in this regard and prefer to treat Greek *Ablaut* as genuine stem allomorphy, we are then forced to conclude that AAB patterns actually *are* attested even in the absence of {A=A} syncretism.

	‘faith’	‘forearm’
NOM	písti-s	pékhy-s
ACC	písti-n	pékhy-n
DAT	píste-i	pékhe-i

We can further corroborate this conclusion if we extend our focus to adjectival paradigms. Particularly revealing are paradigms like those of Classical Greek ‘heteroclite’ adjectives, where the singular obliques take on not only quite another stem but also another declension class than the corresponding core cases, making it even clearer that no floating material may do the trick.

⁶³This may be achieved at the syntactic level via some ‘kind of Differential Object Marking, so that accusative case assignment rules would simply not apply to nouns of the relevant classes’; alternatively, one might work out a solution at the morphological level, via something like post-syntactic ‘impoverishment, so that these nouns would initially be structurally accusative, but would have the relevant head deleted or pruned before vocabulary insertion’ (McFadden 2014).

⁶⁴Particularly problematic for any such line of attack is the fact that one and the same affix might select for different kinds of stem-vocalism depending on the particular stem it is attached to. For example, the datival *-i* of the Greek third declension selects for so-called ‘medium’ vocalism (short *e*) in paradigms like that of *pístis* ‘faith’ (dat. *píste-i*), but ‘weak’ vocalism (no apophonic vowel) for themes like *patér* ‘father’ (dat. *patr-i*).

	‘big’, SG.M	‘much’, SG.M
NOM	méga-s	polý-s
ACC	méga-n	polý-n
DAT	megálō	pollō

If we accept either or both of these empirical arguments from Classical Greek, we are then led to conclude that syncretism between the core cases is not a necessary condition for an AAB pattern of allomorphy to arise. It follows, then, that the patterns of case-based allomorphy in nominal and pronominal paradigms are not qualitatively different from each other, both freely allowing for AAA, AAB, and ABB, but only marginally for ABC and not even marginally for *ABA. Since the pattern is essentially the same across both domains, no modification need therefore be made to the containment-based account sketched out in the last subsection, which will successfully apply to nominal stem-forming elements as well as to pronominal stems.

3.3 Overt containment

Though stem-allomorphy and syncretism constitute invaluable tools to probe into case structure, the evidence they furnish in favor of containment relations among cases is only necessarily indirect. It is, furthermore, inevitably non-directional, since *ABA patterns of syncretism or suppletion could be ruled out just as well by positing the reversed containment hierarchy in (53) rather than the one adopted here (cf. Caha 2009: § 1.3).

(53) [[[DAT/INSTR/LOC] ACC/ERG] NOM/ABS/SNOM]

By contrast, a more direct (and directional) probe into case structure is provided by those languages which transparently reflect the containment relations among cases in their way of stacking case affixes.

In a number of accusative languages, for instance, the nominative form is transparently contained in all the other case forms, which are built by adding one or more affixes to it. We have already seen one such example from Czech in § 2.1.3. Another case in point from outside the Indo-European family is the Dravidian language Telugu. Both paradigms are displayed in table 3.8.

	Czech 'teacher', SG	Telugu 'dog', SG
NOM	učitel	kukka
ACC	učitel-e	kukka-ni/-nu
DAT	učitel-i	kukka-ki/-ku

Table 3.8: From Caha (2009: 121); Krishnamurti and Gwynn (1985: 88)

The same formally unmarked status also characterizes the absolutive in most ergative case systems, and the S-nominative in most tripartite systems. Among the relevant examples seen so far are the ergative language Burushaski (§ 2.2.1) and the tripartite languages Djapu, Nez Perce, and Palula (§ 2.2.3).

Burushaski	'woman'	'stone'
ABS	gus	dan
ERG	gús-e	dán-e
GEN	gús-mo	dán-e
DAT	gús-mur	dán-ar

	Djapu 'person'	Nez Perce 'cat' SG/PL	Palula 2PL
SNOM	yolŋu	picpic	tus
ERG	yolŋu-n	picpic-nim	tus-ím
ACC	yolŋu-y	picpic-ne	tus-aám

In all of these paradigms, the case that marks the subject of intransitive verbs—be it the nominative, the S-nominative, or the absolutive—is transparently contained in all the other cases of the corresponding paradigm, consistent with the first part of the containment hierarchy defended here.

(54) [[[UNMARKED] ACCUSATIVE/ERGATIVE] OBLIQUE]

Caha (2010: 41–42) adduces similar evidence for a containment relation between accusative and obliques in several accusative languages. In Tocharian A and B and in Vlakh Romani, for example, the accusative case form is (in some paradigms at least) transparently contained inside most of the corresponding oblique forms (see tables 3.9 and 3.10).

⁶⁵The genitive is followed by ellipsis to indicate the fact that there are also forms in *-i* and *-e* depending on [the gender, number, and case of] the head noun' (Friedman 1991: 57, fn2).

	Tocharian A 'horse', PL	Tocharian B 'man', PL
NOM	yukañ	eñkw-i
ACC	yukas	eñkw-eṃ
GEN	yukaśši	eñkw-eṃ-ts
LOC	yukas-aṃ	eñkw-eṃ-ne
ABL	yukas-äṣ	eñkw-eṃ-mem
ALL	yukas-ac	eñkw-eṃ-ś(c)
PERL	yukas-ā	eñkw-en-tsa
COM	yukas-aśśäl	eñkw-eṃ-mpa

Table 3.9: Two partial Tocharian paradigms (Gippert 1987)

	'boy', SG	'boy', PL
NOM	čhav-ó	čhav-é
ACC	čhav-és	čhav-én
GEN	čhav-és-koro... ⁶⁵	čhav-én-goro...
DAT	čhav-és-ke	čhav-én-ge
INSTR	čhav-és-ar	čhav-én-car
LOC	čhav-és-te	čhav-én-de
ABL	čhav-és-tar	čhav-én-dar

Table 3.10: One Vlach Romani paradigm (Friedman 1991: 57)

Moskal *et al.* (2015) and Smith *et al.* (2016: 8) note, furthermore, that a similar relation also appears to hold between the *ergative* and the obliques, for example, in several Nakh-Dagestanian languages. Klimov (1994: 148), for instance, explains and exemplifies this point as follows.

A fundamental feature of the declension system in East-Caucasian languages is constituted by the so-called principle of the two stems, whereby one stem acts as the primary, formally unmarked stem and is used to build the absolutive, whereas the other one is *identical to the ergative* and *serves as the basis for the construction of all other cases* [emphasis added]. This principle is instantiated in Agul as follows.⁶⁶

⁶⁶This is my own English translation of the German translation referenced in the main text. The Russian author draws the data in table 3.11 from Aleksandr A. Magometov (1970), *Agul'skij jazyk*, p. 73 (Tbilisi: Mecniereba).

	‘father’, SG	‘belt’, SG	‘sky’, SG
ABS	gaga	çal	zaw
ERG	gaga-di	çal-i	zaw-u
GEN	gaga-di-n	çal-i-n	zaw-u-n
DAT	gaga-di-s	çal-i-s	zaw-u-s

Table 3.11: Three partial Agul paradigms

Sporadic though they are, data like these appear to corroborate also the second part of the class-based case hierarchy defended here—whereby marked core cases such as ergative and accusative are in their turn contained inside obliques.

That said, however, it should also be added that the interpretation of apparent containment relations is not always so straightforward as it looks, and that some caution should thus always be exercised with respect to data of this sort. The principal danger in this connection is constituted by the existence of phonetically null affixes, which are obviously all too easy to mistake for the lack of any affix whatsoever. One may thus be easily misled into thinking that the following Gothic paradigm might disconfirm the hierarchy in (54), the accusative/dative *baúrg* being ‘transparently’ contained in the nominative/genitive *baúrg-s*.

NOM	baúrg-s	‘town’
ACC	baúrg	
DAT	baúrg	
GEN	baúrg-s	

There are very good reasons, though, to doubt that this be the correct interpretation, and to posit instead a phonetically unrealized affix for both accusative and dative: *baúrg-∅*. On this account, nominative and accusative would not stand in any ‘transparent’ containment relation at all, since the stem would be followed by a different case suffix in each case. This move, which may otherwise seem like a merely *ad hoc* strategy for salvaging (54), is in fact supported by very the same kind of morphological evidence from within Gothic itself. See, for example, the reverse ‘containment’ pattern in the following paradigm, where it is the nominative that is seemingly contained inside the accusative.

NOM	qinô	‘wife’
ACC	qinô-n	
DAT	qinô-n	
GEN	qinô-ns	

Clearly, the paradigms of *baúrgs* and *qinô* cannot both be taken at face value at the same time—at least unless one is prepared to claim that Gothic accusatives sometimes contain and sometimes are contained in nominatives, depending on the final segment of their noun stems. It seems more sensible, in this light, to conveniently posit the existence of silent case affixes, while refraining from a hasty and superficial interpretation of such paradigms.

In conclusion, transparent containment relations provide nice evidence in support of the current theoretical proposal, but should nonetheless be always handled with some caution.

Chapter 4

Containment relations among cases, and the theory of dependent case

In the previous chapters, I attempted to demonstrate that the accusative and the ergative consistently constitute a ‘middle field’ within a universal complexity hierarchy of cases. I have argued, more specifically, that the representation of every oblique (except perhaps the genitive) contains the representation of the corresponding ergative and/or accusative, and that the representation of every ergative or accusative likewise contains that of the corresponding unmarked core case (absolutive, nominative, or S-nominative). The main piece of evidence I have adduced for this conclusion is a putative typological generalization, to the effect that no systematic syncretism may involve both an oblique and an unmarked core case without also involving either the ergative or the accusative. In chapter 3, I also presented some converging evidence from the domain of case-driven stem-allomorphy and suppletion, where the same *ABA pattern is also found. All of these phenomena point, I believe, to the conclusion that the morphology classifies cases as in (55).

(55) {SNOM, ABS, NOM}; {ERG, ACC}; {DAT, LOC, INSTR...}

Now, to the extent that one thinks these results are genuine and worth taking seriously, one may be led to expect that the syntax should also classify cases in this way. In other words, one might want to bring the morphological results laid out so far to bear on the choice among competing theories of the syntax (or morphosyntax) of case assignment, fa-

voring those theories that have the accusative and the ergative assigned in much the same fashion, and militating against those theories that do not encode any such parallelism.

4.1 Is ergative an inherent case?

Some morphological counterevidence

Prominent among the theories that deny a parallelism of ergative and accusative are those that treat the ergative as an *inherent* case—that is, a case largely exempt from structurally determined case alternations, and more typically related to the assignment of certain θ -roles (in this case, *external* or *initiator*-related θ -roles such as agent, experiencer, and instrument). Theories of this sort, nonexhaustively exemplified by Laka (2006), Legate (2008; 2012), Mahajan (2012), and Woolford (1997; 2006),⁶⁷ are currently being confronted with an increasing range of empirical challenges on the syntactic side. Řezáč *et al.* (2014: 1280) show, for example, that ergative arguments do undergo case alternations under ECM/raising-to-object in Basque, in stark contrast to Icelandic’s dative ‘quirky’ subjects, which latter are often regarded as the core instantiation of the inherent case category.⁶⁸

(56) Basque (Řezáč *et al.* 2014: 1280)⁶⁹

- a. Katu-ek sagu-ak harrapatu dituzte-la ikusi dut.
cat-d.pERG mouse-d.pAbs caught AUX.3PLA.3PLO-that seen AUX.1SGERG
‘I saw that the cats caught the mice.’
- b. Katu-ak sagu-ak harrapa-tzen ikusi ditut.
cat-d.pAbs mouse-d.pAbs catch-ing seen AUX.1SGA.3PLO
‘I saw the cats catch the mice.’

(57) Icelandic (Þráinsson 2007: 182–183)

- a. (i) Þeir hafa sofið.
3PL.NOM have.3PL slept
‘They have slept.’

⁶⁷See also the papers collected by Johns *et al.* (2006), as well as Aldridge (2008), Deal (2015), and Baker and Bobaljik (forthcoming), for further discussion and some critiques.

⁶⁸Basque, however, also appears problematic for the main alternative to the ergative-as-inherent view—namely, the theory of dependent case; cf. § 4.4.2.

⁶⁹The abbreviations ‘d.pERG’ and ‘d.pAbs’ stand for ‘definite article fused with plural ergative’ and ‘definite article fused with plural absolutive’, respectively.

- (ii) Við teljum þá hafa sofið.
1PL.NOM believe.1PL 3PL.ACC have.INF slept
'We believe them to have slept.'
- b. (i) Þeim hefur leiðst.
3PL.DAT has bored
'They have been bored.'
- (ii) Við teljum þeim hafa leiðst.
1PL.NOM believe.1PL 3PL.DAT have.INF bored
'We believe them to have been bored.'

In a similar vein, Baker and Bobaljik (forthcoming) argue that the ergative case, in several languages at least, is effectively assigned to a nominal whenever it meets the right structural conditions, in a way largely independent of its θ -role. They cite, for instance, phenomena like antipassivization or noun incorporation, which do not change the θ -role of the underlined arguments in (58), yet both impinge on their case-marking pattern.

(58) Chukchi (Baker and Bobaljik forthcoming)

- a. (i) ætlæg-e qərir-ə-rkən-en ekək.
father-ERG seek-PRES-3SGA.3SGO son.ABS
(ii) ætlæg-ən ine-lqərir-ə-rkən akka-gtə.
father-ABS APASS-seek-PRES.3SGS son-DAT
'The father is searching for the son.'
- b. (i) ætlæg-e mətqəmət kili-nin
father-ERG butter.ABS spread.on-3SGA.3SGO
(ii) ætlæg-ən mətqə-rkele-nen
father-ABS butter-spread.on-3SGA.3SGO
'The father spread the butter.'

Even more telling is example (59b), where an internal argument takes on the ergative case simply by virtue of the presence of a new absolutive argument in its vicinity, licensed by an applicative head.

(59) Shipibo (Baker and Bobaljik forthcoming)

- a. Maria-ra mawa-ke.
Maria.ABS-EVID die-PRF
'Maria died.'
- b. Nokon shino-n-ra e-a mawa-xon-ke.
1SG.GEN monkey-ERG-EVID 1SG.ABS die-APPL-PRF
'My monkey died on me.'

All of these facts suggest that ergative case is not inherent but structural in nature.

Now, what I would like to submit here is that the morphological results presented in this thesis may provide new evidence for this same conclusion, which Řezáč *et al.* (2014) and Baker and Bobaljik (forthcoming) have independently defended on syntactic grounds. I think it is possible, indeed, to construct at least three different morphological arguments against treating the ergative as an inherent case.

For one thing, it is easy to see that the class of inherent cases as defined above (θ-related cases relatively resistant to structurally determined alternations) coincides largely (perhaps perfectly) with the class I have so far referred to by the more traditional name of ‘obliques’: dative, instrumental, locative prepositional, and the like. In chapter 2, I argued at length that such cases, in all the languages where they exist, contain in their representations the core structural cases, but the containment relations internal to the obliques’ class itself are cross-linguistically variable, such that no single case containment hierarchy may constrain systematic syncretisms across all languages: the accusative, for example, is syncretic with INSTR/LOCPREP (to the exclusion of DAT) in Latin, but can be syncretic with LOCPREP alone (which can in turn be syncretic with DAT alone) in Classical Armenian.⁷⁰ Now, if the ergative were to be included in the class of oblique/inherent cases, then it too should be naturally expected to partake of the same relative ‘freedom to syncretize’ enjoyed by the other members of that class. We might expect, in particular, to find some ergative ~ absolutive language in which the ergative was syncretic only with oblique cases, and in which some other oblique was more prone to syncretize with the absolutive. However, no good instance of such a language seems to be attested: rather, whenever a syncretism involves an unmarked and an inherent case, it always turns out to (also) involve the ergative. The ergative would then qualify as somewhat exceptional among oblique/inherent cases at least in this respect.

Another generalization from chapter 2 that can be brought to bear on this issue states that no inherent case can ever syncretize with an unmarked case to the exclusion of the accusative, as was already essentially clear both to Baerman *et al.* (2005) and to Caha (2009). If the ergative were an inherent case, we would of course expect it to conform to this generalization. This is, however, evidently not the case, as is shown by those systems, commonly termed ‘tripartite’, that feature both an ergative and an accusative in the same paradigms. As we have already seen in § 2.2.3, the ergative in such systems is very often syncretic with the S-nominative to the exclusion of the accusative—a pattern flatly

⁷⁰Note that even if we were to discard the Latin accusative’s syncretism as accidental (the INSTR/LOCPREP syncretism being indisputably systematic), there still would be no possible case linearization consistent with the facts from both languages.

contradictory to the predictions of the ergative-as-inherent view, but far too widespread across diverse languages to be plausibly treated as an accident. (Relevant examples from Australia and North America are repeated here in tables 4.1–4.2.)

	‘cat(s)’	1SG
SNOM	picpic	’iin
ACC	picpic-ne	’iine
ERG	picpic-nim	’iin

Table 4.1: Core cases in Nez Perce

	‘person’	2SG
SNOM	yolŋu	nhe
ACC	yolŋu-y	nhuna
ERG	yolŋu-n	nhe

Table 4.2: Core cases in Djapu

Should we classify the ergative among inherent cases, we would then be forced, in other words, to render it the only inherent case exempt from the requirement to contain the accusative in its representation.

A similar argument from *ABA patterns is also provided by case-driven suppletion. The relevant generalization here is to the effect that no inherent case can employ the same root as the unmarked core case if the accusative employs a different, suppletive root. Once again, if the ergative were an inherent case, we would expect it to conform to this generalization—namely, that it should not be ‘more similar’ to the nominative than is the accusative. The evidence bearing on this prediction is considerably more meager than that available for syncretism. This is because, on the one hand, case-driven suppletion is restricted exclusively (or almost exclusively) to pronominal paradigms, while, on the other hand, pronominal paradigms universally exhibit a strong tendency to nominative ~ accusative alignment, as has been known since Silverstein’s (1976) seminal work. Indeed, I know of only one language showing a clear instance of a case-suppletive tripartite paradigm—that is, the Australian language Jingulu. That one language, however, is already sufficient to falsify the *ABA prediction of the ergative-as-inherent view, as its 2SG paradigm clearly displays the same root *nyama-* in the S-nominative and in the accusative, but a different root *nga(a)nk-* in the ergative. Albeit less clear-cut, the data concerning the 1SG pronoun also seem to go in the same direction.

	1SG	2SG
SNOM	ngaya	nyama
ACC	ngarr-	nga(a)nk-
ERG	ngaya-rni	nyama-rni

Table 4.3: Jingulu (Pensalfini 2003: 149–152)

To sum up, if the ergative were to be classified among inherent cases, it should then be accorded an exceptional status among them for at least three different reasons:

- (i) its unique reluctance to ever be excluded from syncretisms between the absolutive and other inherent cases;
- (ii) its ability to syncretize with an unmarked case to the exclusion of a non-syncretic accusative;
- (iii) its ability to share the same root of an unmarked case to the exclusion of a suppletive accusative.

In light of this list, it seems reasonable to conclude that the morphology simply does not treat the ergative as an inherent case, and hence that, if we are to maintain the ergative-as-inherent view, we can do so only at the cost of admitting that morphology and syntax classify cases in different ways. If we wish instead to avoid such complications and to enforce a principled parallelism between morphological and syntactic behavior, then we need a theory of case assignment that does put the accusative and the ergative in the same box. I go on to discuss one such theory in the very next section.

4.2 The leading alternative: the theory of dependent case

The theory of case assignment that best meets the demand for a parallel treatment of ergative and accusative is, no doubt, the theory of dependent case proposed by Alec Marantz (1991) and later refined, among others, by Mark Baker (2014; 2015).⁷¹ In what follows, I shall chiefly focus on Baker's implementation of the theory, which updates Marantz's original idea to the current theoretical landscape, with significant gains in empirical coverage.

Baker implements, more specifically, a Marantzian case-assigning algorithm that operates cyclically, on structural chunks roughly equatable to the Spell-Out domains of Chomsky's (2000; 2001) phase theory: some nominal projection like NP (complement of phasal D^0), some clausal projection like TP (complement of phasal C^0), and perhaps, in some languages at least, also some lower projection like VP (complement of the putative phase head v^0 , or Voice⁰). The algorithm scans each of these domains globally, in an attempt to case-mark every nominal therein according to the hierarchy in (60).

- (60) Disjunctive case hierarchy (Marantz 1991)
lexically governed case → dependent case → unmarked case.

⁷¹For precursors and related proposals, see Yip *et al.* (1987), Bittner and Hale (1996), and San Martin and Uriagereka (2000). For reinterpretations and extensions, see, among others, McFadden (2004), Bobaljik (2008), Řezáč (2011: chapter 5), and Preminger (2014: chapters 8–9).

The term *lexically governed case* refers to the case-marking that a particular lexical item (such as a quirky-case verb or an adposition) selects for one of its arguments. This notion thus corresponds rather closely to the alternative notion of *inherent case*, which I have already referred to in § 4.1. Marantz's (1991) most distinctive contributions, then, come from the other two case classes referenced in (60), which Preminger (2014: 146–147) elucidates as follows.

The term *dependent case* refers to case-marking whose appearance is contingent, roughly speaking, on the appearance of two non-lexical/oblique-marked nominals in a single locality domain. This is the category to which accusative and ergative belong. [...] *U*nmarked case refers to case marking [...] whose appearance is not contingent on the presence of any other particular lexical item or noun phrase. This is the category to which nominative and absolutive belong^[72].

How are these cases concretely assigned? I will attempt to explain it here by focusing on case assignment internal to the clause, while momentarily abstracting away from the complications introduced by the possible domainhood of VP. Let me thus go through the three main steps that the case-assigning algorithm typically performs.

- (i) First, the algorithm takes care of all the lexically selected cases in the clause: 'all noun phrases that are selected by lexical items that idiosyncratically specify a particular case marking for their arguments [...] are assigned the idiosyncratic cases in question' (Preminger 2014: 180).
- (ii) Next, the algorithm goes on to evaluate all the noun phrases that did not receive case in the previous step, and applies to them the following parameterized case-marking rule.

- (61) Let NP₁ and NP₂ be two clausemate non-oblique noun phrases.
- a. If NP₁ c-commands NP₂:
 - (i) mark NP₁ [= ERGATIVE] and/or
 - (ii) mark NP₂ [= ACCUSATIVE]
 - b. Otherwise, leave NP unmarked [= ABSOLUTIVE or (s-)NOMINATIVE]

⁷²Expanding on a distinction already explicit in Marantz (1991), Preminger (2014: 147) also adds: 'While sometimes confused with *default case* or *citation form*, this is not what the term refers to (see also Schütze 2001). In English, for example, fragment answers and other freestanding forms appear in accusative (or "objective") case [*Who's there? Me/*I.*] But the unmarked case in English is not accusative (or "objective"); it is nominative.' Cf. especially Schütze (2001)

In other words, the algorithm enters every pair of as-yet-caseless noun phrases into a sort of case competition, which results in assignment of dependent case: in an ergative language/construction, it is the higher noun phrase that receives that case, while, in an accusative language/construction, it is the lower; finally, in a tripartite-alignment language/construction, both (61a-i) and (61a-ii) are applied, so that each of the two noun phrases receives a different dependent case.

- (iii) Finally, all the noun phrases that did not receive either lexical or dependent case in the previous steps are assigned an unmarked case, which we may informally refer to as ‘nominative’, ‘S-nominative’, or ‘absolutive’, depending on which subset of (61a) applies to the language/construction under examination.

Even though the broader picture is in fact slightly more complicated than that (especially because the VP seems to constitute its own case assignment domain, at least in some languages), the Marantz–Baker theory still stands out for its elegance and for the ease with which it captures ergative and accusative systems as ‘variations of a single abstract system’ (Baker 2015: 9). The theory is particularly successful in accounting for those case alternations that keep θ -roles constant, such as those exemplified above in (58a) and (59), and repeated here as (62)–(63).

- (62) Chukchi (Baker and Bobaljik forthcoming)

- a. ətləg-e qərir-ə-rkən-en ekək.
father-ERG seek-PRES-3SGA.3SGO son.ABS
b. ətləg-ən ine-lqərir-ə-rkən akka-gtə.
father-ABS APASS-look-PRES.3SGS son-DAT
‘The father is searching for the son.’

- (63) Shipibo (Baker and Bobaljik forthcoming)

- a. Maria-ra mawa-ke.
Maria.ABS-EVID die-PRF
‘Maria died.’
b. Nokon shino-n-ra e-a mawa-xon-ke.
1SG.GEN monkey-ERG-EVID 1SG.ABS die-APPL-PRF
‘My monkey died on me.’

In both of these pairs, the underlined arguments in the (b) examples take on a different case than in the corresponding (a) examples, despite being assigned the very same θ -roles—agent in (62), patient in (63). This fact, though puzzling on ergative-as-inherent accounts, is naturally expected under the theory of dependent case, whereby ergative case-

marking is taken to depend not on the θ -role assigned but on the presence of another non-oblique noun phrase nearby. All these examples bear out this fundamental insight, featuring an ergative case form only in combination with a (primary or derived) direct object.

Apart from syntactic success stories like these, however, what is most interesting here about the theory of dependent case is how it ends up categorizing cases in essentially the same way that the morphology does: under this theory, ergative and accusative are finally put in one and the same box, and so are absolutive and nominative. It is thus easy to reformulate the containment hierarchy I have argued for in chapters 2 and 3 (cf. (64)) in terms of the same notions that the dependent case theory normally deals in (cf. (65)).

(64) $[[[\text{NOM/ABS}] \text{ACC}^0/\text{ERG}^0] \text{OBL}^0]$

- (65) a. Each *inherent* case contains a clause-level *dependent* case.
 b. Each clause-level *dependent* case contains a clause-level *unmarked* case.

4.3 Marantz's algorithm, structural economy, and case-distinctness

Now that we have seen that the theory of dependent case can provide us with the right theoretical terms to couch our containment generalizations, we can also ask ourselves whether the case containment hierarchy can, in its turn, do any explanatory work concerning the mechanics of the dependent case theory itself. In this section, I shall attempt to answer this question in the positive.

The starting point for this attempt is provided by the observation that those same cases which are given 'elsewhere'-status by the theory—namely, the (s-)NOMINATIVE and the ABSOLUTIVE—are also the lightest cases in the containment hierarchy.⁷³ This naturally prompts the conjecture that there might be a structural economy principle behind this pattern—a principle whereby additional case layers (or case features) such as those of dependent and inherent cases may only be deployed in case of necessity, when the lighter, unmarked cases are somehow 'not enough'. The question arises, then, what necessity there might be for ever using heavier cases such as the obliques, the ergative, or the accusative, rather than always using the nominative or the absolutive: why are these latter sometimes

⁷³This is reminiscent of Greenberg's (1966: 95) Universal 38: 'Where there is a case system, the only case which ever has only zero allomorphs is the one which includes among its meanings that of the subject of the intransitive verb.' What I am adding to this here is a potentially explanatory connection between this (surface-false) generalization and a specific theory of case assignment.

‘not enough’?

A promising answer would be to suppose that clauses must simultaneously abide by the two requirements in (66).

- (66) a. Selectional requirements must be satisfied.
b. Two noun phrases that stand in an asymmetric c-command relation within the same domain must bear sufficiently distinct cases.⁷⁴

On this account, the additional case structure of ergatives and accusatives may be deployed only when needed to tell apart two otherwise–overly–similar noun phrases, while the case structure associated to obliques may be deployed only if idiosyncratically selected by some specific lexical item.

Though *prima facie* attractive, this view encounters several problems. The first one has to do with the cross-linguistic validity of (66b): why is this requirement apparently inactive in the languages that exhibit so-called neutral case alignment, namely those that display no morphological case at all? Moreover, even those languages in which (66b) should in principle hold seem to permit multiple clausemate identical cases under the right circumstances. This can be illustrated, for instance, by some of the most common case-marking patterns for ditransitives, NOM-ACC-ACC and ERG-ABS-ABS.

Finally, it is also not clear that the economy principle under discussion really applies universally. If it did, one would expect that any two nondistinct noun phrases could always be dissimilated simply by dependent-case–marking either one or the other. But this clearly is not what happens in tripartite core-case alignments, which are characterized exactly by marking both noun phrases at once in such situations.

I will try to solve all of these problems in the following couple of subsections.

4.3.1 Neutral and tripartite systems: refining case-distinctness

Let us start with the problems concerning the applicability of a case-distinctness account to neutral and tripartite case systems as well as to ergative and accusative ones. The hypothesis I would like to explore here is that this problem might in fact be solved by a simple parameterization of the notion of ‘sufficiently distinct cases’ as referenced in (66b). The basic idea, in other words, is that languages are not all equally tolerant when it comes to assessing whether or not two given cases can be told apart. The baseline criterion is of course the one of accusative and ergative systems, which both regard NP₁ and NP₂

⁷⁴(66b) is clearly reminiscent of—though not immediately reducible to—Norvin Richards’ (2010: ch. 2) ‘Distinctness’ condition on linearization statements.

as case-distinct whenever there is at least one case shell that NP₁ and NP₂ do not share. On this criterion, an unmarked and an ergative case *can* be distinguished, since the DEP⁰ head of the latter is not also present on the former; the same goes, of course, for unmarked and accusative, and, *a fortiori*, for unmarked and oblique.

What I wish to suggest at this point is that tripartite case systems might simply arise as a by-product of a lesser degree of tolerance in case-distinctness evaluation. More specifically, these systems seem to require, in order to distinguish two noun phrases, that there be at least *two* case shells that those noun phrases do not share. Nominative and accusative, then, cannot count as distinct in this system, given that they share all case shells except one (the DEP⁰ case shell of the accusative); and the same goes for ergative and nominative. Only two options are therefore allowed for dyadic structures in such systems:

- (i) either S-NOMINATIVE + OBLIQUE (the latter containing *two* case shells absent in the former)
- (ii) or ERGATIVE + ACCUSATIVE (provided that the DEPENDENT case shell marked as [+ERG] counts as distinct from that marked as [+ACC]).

If we assume a case-distinctness requirement of at least two non-shared case shells, it becomes clear that there is nothing uneconomical about the double dependent-case-marking of tripartite systems. Furthermore, even in these systems, we can still witness structural economy at work in that monadic structures (where no case-distinctness requirement may apply) can only display unmarked cases, not dependent or oblique ones.

Against this background, it is also easy to envisage one more possible parameterization for so-called neutral case systems: they would be those systems capable of telling two noun phrases apart even when they share together all of their respective case shells. We end up, in other words, with the following case-distinctness typology.

- (67) What does it take for two cases to be sufficiently distinct for (66b)?
 - a. At least *two* non-shared case shells: TRIPARTITE SYSTEMS
 - b. At least *one* non-shared case shell: ERGATIVE AND ACCUSATIVE SYSTEMS
 - c. At least *zero* non-shared case shells: NEUTRAL SYSTEMS

4.3.2 The puzzle of ditransitives: a successive-dissimilation approach

Finally, let me address the puzzle posed by the apparent violation of case-distinctness in NOM-ACC-ACC, ERG-ABS-ABS, and ERG-ACC-ACC structures, such as those exemplified respectively in (68), (69), and (70).

- (68) Cuzco Quechua (adapted from Baker 2015: 231)

Yachaq-kuna-ta qilqa-y-ta yacha-chi-rqa-nku.
 learner-PL-ACC write-INF-ACC learn-CAUS-PAST-3PL
 ‘They taught the students writing.’

- (69) Burushaski (adapted from Willson 1996: 61)

Hilés-e dasín taswíir mó-oltir-imi.
 Boy-ERG girl.ABS picture.ABS 3SG.F-SHOW-3SG.M.PAST
 ‘The boy showed the girl the picture.’

- (70) Diyari (Austin 2011: 118; cf. Baker 2015: 232)

nhulu pulanha nhinha putu yingki-rna wara-yi
 3SG.NONF.ERG 3DU.ACC 3SG.NONF.ACC thing.ACC give-PTCP AUX-PRES
 ‘He gave them that thing.’

These examples are, indeed, potentially problematic for the dependent case theory. In particular, if we assume the Goal argument (‘the students’, ‘the boy’) to asymmetrically c-command the Theme (‘writing’, ‘the picture’), we should expect the presence of the latter to trigger ergative case-marking on the former, and hence to result, contrary to fact, in an ERG-ERG-ABS case pattern. Distressingly for the classical Marantzian theory, such double-ergative patterns appear to be systematically unattested (cf. Nie 2016).⁷⁵

This is clearly problematic also for my own case-distinctness-based approach, insofar as di-accusative and di-absolutive patterns each seem to use one and the same case twice. There are, as far as I can see, at least two conceivable ways to explain this problem away—and they may both be true, albeit for different languages or constructions. One viable possibility would be to surmise that, at least in some of the relevant languages, neither of a ditransitive’s internal arguments asymmetrically c-commands the other. If this were so, the two internal arguments would not enter into any case competition with each other, but would each compete with the topmost, external argument only. The relevant c-command pairs would then include only $\langle \textit{Agent}, \textit{Goal} \rangle$ and $\langle \textit{Agent}, \textit{Theme} \rangle$, but neither $\langle \textit{Goal}, \textit{Theme} \rangle$ nor $\langle \textit{Theme}, \textit{Goal} \rangle$, thus successfully deriving the observation that the two

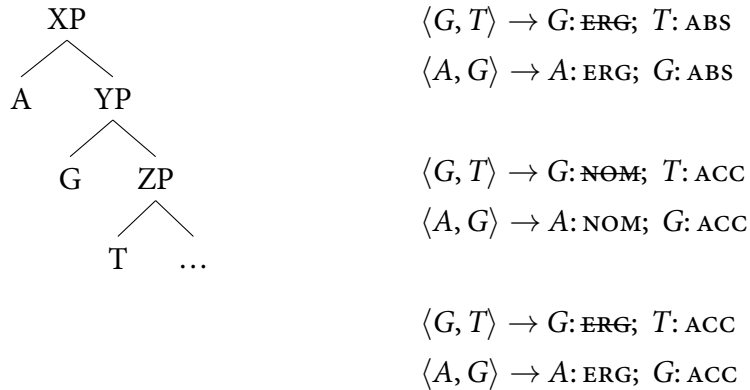
⁷⁵Dixon (2012: 257) has a couple of surface counterexamples to Nie’s (2016) generalization. One is from the North-west Caucasian language Kabardian, in which, however, what Dixon refers to as ergative is in actual fact more of an all-purpose oblique, also expressing the ‘indirect object of benefactive’, the ‘general locative of a stative verb’, the ‘general goal of a dynamic verb’, and even the ‘locus of activity’ (Colarusso 1992: 53–54). The other surface counterexample is the Brazilian isolate Trumai, in which, however, the double-ergative pattern only surfaces in causative constructions, but not with any lexical ditransitive (Guirardello 1999: 255ff, 301ff, and especially p. 311 — ‘In no other construction in Trumai a single verb can [*sic*] have two A arguments.’). One might thus argue that causatives are in fact not monoclausal but biclausal in this language (even though this view would in turn give rise to several difficulties; cf. Guirardello 1999: 310ff).

Thanks to Yining Nie, Maria Polinsky, and Barbara Stiebels for helpful discussion.

internal arguments end up with the same structural case.

Some of these ERG-ABS-ABS structures, however, might provide some evidence for a c-command asymmetry between Theme and Goal, thus invalidating the explanation suggested above. Even though I do not know of any such language, I still would like to sketch out a possible response to such hypothetical counterevidence. The option I'd like to pursue, in particular, would be to construe the dependent-case-assigning algorithm as a bottom-up scanning of all the relevant asymmetric c-command pairs. Whenever the algorithm comes across a pair of overly similar noun phrases, it attempts to dissimilate it in one of three ways: by rendering the higher noun phrase more marked than the lower, by rendering the lower noun phrase more marked than the higher, or by marking both the higher and the lower noun phrases, albeit in different ways. Crucially, I'd also like to suggest that later dissimilations may overwrite the results of the previous ones, in a way closely reminiscent of the nanosyntactic Spellout mechanism which I described in § 1.3.

Let me go through one toy example just to elucidate the logic. Let A, G, and T be an Agent, a Goal, and a Theme, respectively. Let neither of them be assigned any inherent case, and let the c-command relations among them be those schematized in the next tree.



The case-assigning algorithm proceeds from the bottom upwards. At the ZP level, there is still only one noun phrase to take care of, and no case-distinctness problem has arisen yet. At the YP level, however, one such problem arises: the c-command pair $\langle G, T \rangle$ violates case-distinctness. In an ergative ~ absolutive system, the algorithm finds a solution by marking G as ergative and by leaving T unmarked. Yet another problem, however, arises at the XP level, when two new problematic pairs are introduced: $\langle A, G \rangle$ and $\langle A, T \rangle$. The latter pair does not pose any particular problem: A is marked as ergative and T is again left unmarked. The more problematic pair here is $\langle A, G \rangle$: on the one hand, simply marking A as ergative would not suffice to turn it into a well behaved case-distinct pair (the result would still be a nondistinct pair $\langle \text{ERG}, \text{ERG} \rangle$); on the other hand, leaving A unmarked would violate the instruction to mark the higher noun phrase more than

the lower. The system thus simply un-marks G, effectively turning it into an absolutive. The same kind of procedure (schematized above next to the tree) will likewise ensure that the two internal arguments will receive the same structural case, as desired, also in nominative ~ accusative and tripartite systems.⁷⁶

To conclude, a view of dependent case assignment centered on identity avoidance is thus not only compatible with the data from ditransitive structures, but may even prove instrumental in sharpening our understanding of their case-theoretic properties.

4.4 On semantically based case-marking

There is one more problem which besets both the Marantz–Baker theory and my own reinterpretation of it, and which I would like to discuss at greater length here. The problem in question lies in the ostensible existence of semantically based case-marking, typically instantiated by the systems that Dixon (1994) refers to as ‘split-S’ and ‘fluid-S’. Both of these kinds of systems appear interesting here insofar as they do not case-mark all intransitive subjects (S) alike, but rather mark some S like direct objects (O) and some other S like transitive subjects (A), depending on the semantics of the intransitive predicate involved. In general, the S that are likelier to be marked like A are those that are closer to the prototype of an agent, initiating or intentionally participating in the event expressed, whereas the S that are likelier to be marked as O are the ones that are closer to the prototypical patient—those which passively undergo a change of state initiated by some other participant.⁷⁷ Dixon’s fluid-S and split-S systems only differ in that, in the latter, the marking of intransitive subjects is lexically determined and fixed for each verb, whereas, in fluid-S systems, the subject of one and the same intransitive verb ‘can be marked as S_A (i.e. like A) or as S_O (like O), depending on the semantics of a particular instance of use’ (Dixon 1994: 71). Either way, however, such semantically based marking clearly constitutes a problem for theories of dependent case, insofar as it purports to prove that the choice between different core cases may not (only) depend on syntactic factors like the presence or absence of other nonoblique arguments nearby, but may also depend on θ -related factors like affectedness, volitional control, and the like.

⁷⁶This view of case assignment owes to Merchant (2006), Pesetsky (2013), and Richards (2013) the key idea that subsequent case assignments may overwrite previous ones. This system also has a good deal in common with the one devised by Mark Baker (2015: 234ff) to deal with much the same data. Unlike Baker’s, however, the current system does not require positing gratuitous c-command computations inside VP even in the absence of any VP-level cases to assign. Likewise, it also disposes of the assumption that VP is a case-assignment domain (if with variable degrees of ‘hardness’) in all case-inflecting languages.

⁷⁷This deliberately informal presentation is clearly inspired by Dowty’s (1991) work on proto-roles.

In the next few pages, I shall tackle this problem in two steps. In § 4.4.1, I shall begin by reassessing the available evidence for case systems with consistently semantic S-marking, reviewing and expanding on some recent findings by Baker and Bobaljik (forthcoming). It will turn out that such systems are in reality exceedingly rare, possibly adding up to no more than two instances among the languages of the world. In § 4.4.2, I will then focus on those two putative instances, Basque and Laz, and will show that they partially differ in the behavior of their ‘ergative’ case with respect to structurally conditioned case alternations. This will lead me to suggest, *contra* Baker and Bobaljik (forthcoming), that those two languages each require a different analysis in order to be reconciled with the theory of dependent case.

4.4.1 On consistently semantic S-marking: just very rare or outright impossible?

Baker and Bobaljik (forthcoming)—henceforth ‘B&B’—begin their reassessment of the evidence for semantic S-marking by noticing a striking asymmetry between agreement and case in this connection. Their passage is worth quoting at length.

Typological sources say that the active ~ inactive/stative alignment pattern (also called a split-S or fluid-S pattern) is attested in languages of the world [...]. But our certainty that this is so is marred by the fact that these discussions generally conflate data from morphological case marking on nouns with data from agreement patterns on verbs. Mithun (1991) is a typical example: of the five active languages that she discusses at some length, four are head-marking languages (Lakhota, Guaraní, Caddo, Mohawk), and only one (Central Pomo) has overt case marking on NPs. [...]

Data from the *World Atlas of Language Structures* confirms that there is a strong interaction between Nichols’s (1986) head-marking/dependent-marking distinction and the distinction between active languages and true ergative languages. Siewierska (2005) lists 26 out of 380 languages as having an active agreement pattern, a respectable 6.8%. Indeed, in agreement-oriented languages, an active system is slightly more common than a straight ergative system (19/380, 5%). In contrast, Comrie (2005) lists only 4 out of 190 dependent marking languages as having an active case pattern (2.1%), and in this language type a straight ergative pattern is much more common than an active one (32/190, 16.8%).

In summary, while active ~ inactive agreement patterns are relatively uncommon but still robustly attested across the world’s languages, active ~ inactive case patterns turn out

to be significantly rarer.⁷⁸ Furthermore, as is also remarked by B&B, even those few languages that do exhibit such case patterns generally employ them only under tightly constrained circumstances. This stands in stark contrast to the situation of canonical morphologically ergative languages, in many of which transitivity qualifies as the only determining factor for ergative case, with no further necessary conditions to be specified.

One extra condition of this kind that is particularly common across split-S languages has to do with tense/aspect. For example, both Georgian and Svan, from the Kartvelian family, allow for so-called ergative case-marking only in combination with verbs of the aorist series (what Karvelologists often refer to as ‘Series II’), but not with verbs of either the present or the perfect series (‘Series I’ and ‘III’, respectively).⁷⁹ Similarly, in the Indo-Aryan family, Hindi and Kashmiri only allow so-called ergative or agentive case-marking in perfective clauses, and only for a very limited class of intransitive verbs (e.g. ‘cough’, ‘sneeze’, ‘shout’): in Hindi, furthermore, most of such verbs take ergative subjects only as an option, with the choice between NOM and ERG subjects being only partially determined by semantic factors.⁸⁰

Alongside similar tense/aspect-based restrictions in Drehu and Lhasa Tibetan, B&B also discuss another, rarer kind of split in Tsova-Tush, a Nakh language also known as Bats or Batsbi. In this oft-cited language, S-marking appears to be semantically based, but only on 1SG, 1PL-exclusive, 2SG, and 2PL pronouns, whereas third-person arguments conform to a canonical ergative ~ absolutive alignment (Holisky 1987). It is important to note that this split does not look like a merely morphological one like those of the underlyingly tripartite languages discussed in § 2.2.3. To wit, Tsova-Tush third-person pronouns and nouns do have a separate ergative form, but they can—and indeed must—use it only in transitive clauses.⁸¹

⁷⁸Cf. Dixon (1994: 76), also cited by B&B: ‘Split-S marking relates to the nature of the verb. It is scarcely surprising that for most languages of this type morphological marking is achieved by cross-referencing on the verb’.

⁷⁹On Georgian, cf. Harris (1981: 1–2, 46–47), among others. On Svan, cf. Tuite (1998: 18).

⁸⁰The relevant generalization, according to Mohanan (1994: 73), is that ‘when a NOM subject cooccurs with verbs that have an option between NOM and ERG subjects, the action must be nondeliberate’; if the subject is ERG, however, the action may not necessarily be deliberate. (By contrast, agentive S-marking in Kashmiri is optional only for the subject of the verb *natsun* ‘to dance’; cf. Wali and Koul 1996: 153.) The Hindi system is, in fact, even more interesting and mysterious than that, insofar as ergative marking appears to be occasionally optional (and sometimes outright ruled out) even in some perfective *transitive* clauses, in ways that appear at least partially independent from the θ -roles involved; see Mahajan (2012) for recent discussion within the ergative-as-inherent approach. It is not clear to me how all these facts might be neatly captured under the theory of dependent case.

⁸¹This split is also puzzling from a Silversteinian perspective. Indeed, first- and second-person pronouns are generally taken to designate the prototypical agents, and hence to be the ones for which agentivity should most easily be left unmarked. However, the Tsova-Tush system appears to have the asymmetry in reverse, marking agentivity more on first and second persons than on anything else. Curiously enough, though, Silverstein’s (1976) predictions are restored, as it were, in the 1PL-inclusive paradigm, which conforms to a

Subsequently, B&B also discuss what Bittner and Hale (1996: 45) refer to as ‘accusative active languages’—systems in which ‘the putative ergative case on agentive subjects is actually morphologically unmarked (\emptyset); what is overtly marked is a kind of nonagentive case found on (some) direct objects and (some) nonagentive subjects of intransitive verbs: *-al* on pronouns in Northern Pomo [Hokan, California]; *-m* in Imonda [Papuan]’ (B&B: § 5.3). B&B, however, reanalyze these languages as actually neutral languages, and convincingly argue that their putative nonagentive suffixes *-al* and *-m* should in fact be regarded as *dative* markers.

As such, their core use is on the goal arguments of ditransitive verbs [...]. These case markers can also be used on the theme arguments of monotransitive verbs, but only as so-called differential object markers (Aissen 2003; Bossong 1985). [...] This is very reminiscent of markers such as Hindi *-ko*, which serve both as differential markers for animate and/or specific direct objects, and in dative functions such as marking recipients (the most common DOM pattern, Bossong 1985 [pp. 108ff.]).

Once we think of the overt cases in these languages as datives rather than accusatives, an easy hypothesis becomes available for why they are found on the subjects of a proper subset of the unaccusative predicates: these are simply predicates that select for quirky dative case on their subjects, like those known from Icelandic [...]. This fits well with the fact that only eight known verbs have case-marked subjects in Imonda, not the whole class of unaccusatives [⁸²]. It also fits with O’Connor’s (198[7: 215]) observation that having overt case-marking on the subject in Northern Pomo sometimes expresses ‘empathy’ with the subject rather than agentivity.

Finally, the same exercise attempted by B&B can also be extended to another few languages that they omit to discuss, such as for example Late Latin, a variety frequently claimed to have conformed to a semantically based S-marking pattern (see, among others, Plank 1985; Cennamo 2011; Rovai 2014). Claims to this effect are generally part of a larger diachronic reconstruction, according to which, in the passage from Classical Latin’s nominative ~ accusative alignment to the neutral alignment of Romance languages, ‘it often was the original accusative rather than the nominative that was victorious’ (Plank 1985: 288); crucially, however, it appears that the accusative did not start this gradual takeover at random, but rather began by colonizing the realization of inactive subjects and nonverbal predicates first. As Rovai (2014: 126–127) aptly summarizes it,

Starting from the third/fourth century AD, within certain predicative constructions occurring in mid-low register texts from different areas, the argument structure is

neutral core-case alignment, using *vai* in all of A, S, and O functions (see Holisky and Gagua 1994: 42–43).

⁸²Cf. Seiler (1985: 145ff).

not identified according to an accusative alignment, but according to a semantically based one. The accusative case can in fact replace the nominative to encode non-agentive subjects of intransitive predicates such as passives [...], anticausatives [...], copular constructions [...], and other verbs denoting states [...] or uncontrolled changes of state [...]

Be that as it may, however, there still exists no evidence of any Latin variety—no matter how late or peripheral—displaying a neat one-to-one correspondence between core cases and (classes of) θ -roles, so that Rovai (2014: 139) himself bluntly acknowledges that Latin’s tendency to semantic S-marking may, ‘[t]o be sure, [...] be considered to be a minor pattern and a sub-standard feature’. More to the point, the same author (Rovai 2014: 128; emphasis mine) further specifies:

Although the phenomenon under investigation is normally referred to as “extended accusative/restricted nominative,” the restriction of the nominative has not received the same attention as the extension of the accusative. This situation may depend on the fact that, while there is direct evidence of the latter [...], the symmetrical restriction of the nominative is not easy to prove because a number of social and stylistic factors obscure the distribution of the competing alignments in the late documents. The extended accusatives are, in fact, well attested in the colloquial prose of the *Peregrinatio Aetheriae* and in a technical work on veterinary medicine like the *Mulomedicina Chironis*, but they are completely lacking in high-style writers such as Augustine and Boetius. Moreover, even in those texts where some instances are found, the vast majority of predicative constructions follow an accusative alignment, which always remains the main argument coding system in Latin sources.

Witness examples like (71), it thus appears clear that even the Latin variety of the above-mentioned *Peregrinatio* cannot provide the consistent prototype of semantic S-marking that B&B are seeking.

(71) Latin (*Peregrinatio Aetheriae* 43, 4; ~400 AD; from Rovai 2014: 129)

Ut null-us Christian-us remane-at in civitat-e.
 so.that no-NOM.M.SG Christian-NOM.M.SG be.left-3SG.PRES.SUBJ in town-INSTR.SG
 ‘So that there is no Christian [NOM] left in the city.’

Finally, I discuss one last class of languages which B&B do not discuss, but which is mentioned in Dixon (1994). Dixon (1994: 28–29) describes these languages as having ‘grammatical marking which directly describes the semantics of the conceptualisation of a particular situation without this having to be related to a prototype and filtered through

basic syntactic relations.’ He then focuses on three putative representative instances of this kind: Manipuri, Wakhi, and Foloपा. As it turns out, however, none of those three languages can make a cogent case for the existence of this even more radically semantic rationale behind case-marking. For instance, in Manipuri (Tibeto-Burman, North-east India), the suffix *-nə* may effectively mark the subject of a transitive or intransitive predicate only if it designates the controller of the action, so that there even seem to be minimal pairs suggestive of lesser or greater intentionality on the subject’s part:

(72) Manipuri (Dixon 1994: 30)

- a. ay-nə Tombə-bu theñɲi.
1SG-nə T.-DOM touched
‘I touched Tomba (intentionally).’
- b. ay Tombə-bu theñɲi.
1SG T.-DOM touched
‘I touched Tomba (unintentionally).’

However, Bhat and Ningomba (1997: 140), whose grammar Dixon (1994) could only reference as a draft, also state clearly that the ‘suffix *-nə*, denoting the actor of action sentences, can also be left unspecified provided that this non-specification does not lead to ambiguity’, as is illustrated by the following transitive and intransitive examples.

(73) Manipuri (Bhat and Ningomba 1997: 140–141)

- a. məhak-(nə) celli.
he-(nə) ran.
‘He ran.’
- b. huy-nə ma-bu ciki.
dog-nə he-DOM bit.
‘The dog bit him.’
- c. huy-du ma-bu ciki.
dog-that he-DOM bit.
‘That dog bit him.’

The system, in other words, must make provision for a massive amount of optionality, which is clearly bound to obscure the expected direct linking between event conceptualization and the actual case-marking.

Much the same goes as well for the South-eastern Iranian language Wakhi, in which ‘both transitive and intransitive subjects can be marked in either of two ways, by what have been called NOM(inative) and OBL(ique) cases’ (Dixon 1994: 34). Concerning the

choice of NOM or OBL subject marking, Elena Bashir⁸³ maintains that

an OBL subject will be more likely to appear when one or all of the following conditions obtains: (1) The action is performed as volitional and active rather than as 'passive' or stative; (2) the referential identity of the subject is to be stressed; (3) within a discourse, the subject or topic is 'new' or different from the subject of the preceding action.

But although all of these factors may certainly play some role in favoring either case-marking over the other, the fact remains that the tendencies at play are so subtle and intricate that two experienced linguists like John Payne and Boghsoh Lashkarbekov, faced with such data, ended up talking respectively of free variation and of functional equivalence between NOM and OBL subject marking (cf. again Dixon 1994: 34). This ought to suffice, I hope, to show that even this language falls short of instantiating the clean pattern that one might expect from a direct linking between case-marking and event semantics.

Finally, the last language which Dixon (1994: 32) reports to have semantically based argument marking is the Papuan language Folopa, in which both transitive and intransitive subjects may either bear the 'controller' suffix *-ne* or remain unmarked (Anderson and Wade 1988). More specifically, each verb in the language appears to have its own range of case-marking options, according to principles that are mostly semantically justifiable, such that, for example, transitive verbs incompatible with 'controller' marking include the equivalents of 'like' and 'dislike' (Anderson and Wade 1988: 8).⁸⁴ It is interesting to note, however, that, among Folopa intransitive verbs, there are some that only take \emptyset -marked subjects and some that alternate between \emptyset -marking and 'controller' *ne*-marking, but there are none that are confined to *-ne* — not even among the most prototypically 'active' ones (Anderson and Wade 1988: 6). On the other hand, it is also worth noting that *ne*-marking is possible even on subjects of rather prototypical unaccusative verbs, like *fopaarapó* 'to get mad', which describes a nonvolitional change of state (cf. Anderson and Wade, *ibid.*). Even Folopa, then, does not appear to allow for a neat one-to-one correspondence between Θ -roles/macroroles/proto-roles on the one hand and case markers on the other.

This leaves only two languages as possibly good instances of a consistently semantic S-marking pattern. I shall focus on them both in the next subsection.

⁸³E. Bashir (1986), 'Beyond split-ergativity: Subject marking in Wakhi', *CLS* 22: 14–35. Quoted (with no reference to page numbers) by Dixon (1994: 34ff.).

⁸⁴On a dependent-case approach, these exceptionally unmarked transitive subjects may be analyzed as underlying prepositional complements, unable to c-command the object and thereby to qualify for dependent-case-marking (cf. Baker 2015: § 3.2).

4.4.2 Basque, Laz, and the heterogeneity of ‘agentive’ case

The main upshot of B&B’s empirical reassessment—and of my small additions to it—is that there is virtually no known instance of a language that consistently case-marks core arguments on the basis of their ϑ -semantics only. There do exist, by contrast, a good number of languages that case-mark core arguments only on the basis of syntactic notions like Dixon’s (1994) A, S, and O: these include canonical nominative ~ accusative languages like Tamil or Classical Greek, canonical ergative ~ absolutive languages like those discussed by B&B (Tsez, Chukchi, Inuktitut, Shipibo), and overtly or underlyingly tripartite languages like Djapu or Pitjantjatjara. In all of these languages, the presence of a marked core case correlates neatly with transitivity, but there are hardly any languages where it correlates neatly with agentivity or similar semantic notions. This is rightly taken by B&B to suggest that any theory attempting to tie case directly to ϑ -semantics must be doomed from the start: indeed, the pattern that such theories would take to be the typological default turns out instead to be vanishingly rare.

This does not mean, though, that (at least descriptively) active ~ inactive case patterns are entirely nonexistent across the languages of the world. Rather, as we saw in the last subsection, such patterns do hold in some tiny corners in the grammars of a handful of languages, and there are also two languages in which the classic unergative ~ unaccusative distinction⁸⁵ is at least a very good predictor for S-marking: those are the well-known isolate Basque and the lesser-known Kartvelian language Laz, both of which mark the single argument of an unaccusative predicate just like a direct object, and the single argument of an unergative just like a transitive subject, regardless of tense-, aspect-, or person-related factors (cf. B&B: § 5.4).

The question remains, then, how the theory of dependent case may try to capture such problematic patterns. There are at least two answers to this question that have been entertained in the literature. One, passingly mentioned by Baker (2015: 54), is to say that the ‘ergative’ case of consistently semantic S-marking systems might in fact not be a dependent case at all, but rather an inherent case assigned along with agentive ϑ -roles. This would amount to saying that the ergative-as-inherent view critically discussed in § 4.1 might, after all, be correct, although only for languages like Laz and not for canonical ergative languages like Chukchi. To put it still another way, the so-called ergative case of Laz would then be fundamentally different in nature from the *bona fide* ergative of Chukchi, and rather more akin to typically ϑ -related obliques like the dative or the instrumental.

⁸⁵See, among many others, Levin and Rappaport Hovav (1995).

A different tack, ultimately taken by B&B themselves, is to suppose that some of the intransitives in these languages behave like transitives for case-theoretic purposes simply because they actually *are* transitives. B&B (§ 5.4) suggest, more specifically, ‘that they are concealed transitives, taking a second argument that is approximately meaningless and phonologically null, but nevertheless counts for triggering ergative case on its coargument by the dependent case rule in [(61a-i)].’ As B&B also remark, this kind of analysis, independently advanced by Hale and Keyser (1993), is also famously supported in Basque by the fact that most⁸⁶ of the ‘ergative’-selecting intransitives in that language also admit of an alternant syntactically realized as a formally transitive light-verb construction:

(74) Basque (Preminger 2012: 278–279)

- a. Jon-ek dantza egin d- ϕ -u- ϕ .
Jon-ERG dance do 3ABS-SG.ABS-have-3SG.ERG
- b. Jon-ek dantzatu d- ϕ -u- ϕ .
Jon-ERG dance-PTCP 3ABS-SG.ABS-have-3SG.ERG
‘Jon danced.’

Even though the ‘agentive-as-inherent’ and the ‘unergative-as-transitive’ theories may both appear stipulative to some extent, it is worth stressing that both of them do make testable predictions, and that their respective predictions are different enough so that the two theories can be comparatively appraised.

The most important point about which predictions differ has to do, more specifically, with the case-preservation properties of the so-called ergative case. Recall from § 4.1 that one of the most robust diagnostics for identifying inherent cases is their relative resistance to structurally determined case alternations. This is illustrated here by the well-known Icelandic pair (75), in which the nonoblique subject participates in a NOM/ACC alternation under raising to object, whereas the ‘quirky’ subject preserves its dative case.

(75) Icelandic (Þráinsson 2007: 182–183)

- a. (i) Þeir hafa sofið.
3PL.NOM have.3PL slept
‘They have slept.’
- (ii) Við teljum þá hafa sofið.
1PL.NOM believe.1PL 3PL.ACC have.INF slept
‘We believe them to have slept.’

⁸⁶Though not all, as pointed be Preminger (2012). See B&B’s fn. 15 also for skeptical discussion of another agreement-based argument of Preminger’s (2012) against positing an implicit object in (some) Basque unergatives.

- b. (i) Þeim hefur leiðst.
 3PL.DAT has bored
 ‘They have been bored.’
- (ii) Við teljum þeim hafa leiðst.
 1PL.NOM believe.1PL 3PL.DAT have.INF bored
 ‘We believe them to have been bored.’

Against this background, it is easy to see how the predictions of the two competing theories differ: while the agentive-as-inherent theory expects the agentive argument to preserve its case under raising just as the inherent dative does, the unergative-as-transitive theory takes the agentive to be a run-of-the-mill structural case, and thus expects it to be as prone to alternations as the structural nominative case.

Interestingly, both theories seem to be right, albeit each for a different language. On the one hand, Řezáč *et al.* (2014) have shown that the subject of an unergative does shift to the absolutive under raising to object in Basque. Significantly, the raised subject at issue may even be a chunk of an idiom, thus ruling out the hypothesis that it may also be an argument of the perception predicate in the main clause.

(76) Basque (Řezáč *et al.* 2014: 1283; idiom boldfaced)⁸⁷

- a. Kontu horretan, **ur-ak** **bide egin** du azken urteotan
 in this matter water-d.sERG way made AUX.3sERG in these last years
 ‘In this matter, things have advanced these last years.’
- b. Kontu horretan, **ur-a** **bide egiten** ikusi dugu azken urteotan
 in this matter water-d.sABS way making seen AUX.1PLA.3SGO in these last years
 ‘In this matter, we have seen things advance these last years.’

As it happens, this propensity to undergo case alternations was taken as evidence against the ‘ergative’-as-inherent analysis both by Řezáč *et al.* (2014) themselves and by Baker and Bobaljik (forthcoming).

On the other hand, Demirok (2013) argues convincingly that Laz has raising to object too, but that the ‘ergative’ subjects raised in this language do preserve their case. In this case, the suspicion to be averted is not that the case-alternating subject may be an argument of the main verb, but rather that it may never have moved from downstairs. To ensure that movement has indeed taken place, Demirok uses an argument from scope.

⁸⁷The abbreviations ‘d.sERG’ and ‘d.sABS’ stand for ‘definite article fused with singular ergative’ and ‘definite article fused with singular absolutive’, respectively.

(77) Laz (Demirok 2013: 98–99)

- a. Kat'a bere-k ma va m-dzir-u
every child-ERG 1SG.ABS NEG 1-see-3.PAST
'Not every child saw me.' [$\neg > \forall$; $*\forall > \neg$]
- b. Kat'a bere*(-k) ma va m-dzir-u do-m-a-ts'an-u
every child-*(ERG) 1SG.ABS NEG 1-see-3.PAST PREV-1-APPL-think-3.PAST
'I thought every child failed to see me.' [$\forall > \neg$; $*\neg > \forall$]

In the simple clause (77a), negation obligatorily takes scope over the universal quantifier, whereas in (77b) the reverse scope obligatorily obtains. Demirok takes this as evidence that the quantifier has actually moved to a position above negation, even though the constituent order does not allow us to see this. If we accept this argument and regard (77b) as a real instance of raising to object, then we may follow Demirok (2013) further and take the fact that 'ergative' is preserved as evidence that it is indeed an inherent case.

In summary, if we take the case-preservation diagnostic seriously enough, we are led to conclude that consistently semantic S-marking is not only an extremely rare phenomenon, but that even those few languages that do exhibit it might in fact instantiate fundamentally different phenomena under a deceptively uniform surface pattern.

Chapter 5

More case studies

In this final chapter, I will try to test the generalization defended in chapter 2, and repeated here as (78), against a sizeable sample of areally and genealogically diverse languages.

- (78)
- a. No systematic case syncretism can target both an unmarked core case (nominative or absolutive) and an oblique case (such as the dative or the instrumental) without also targeting a marked core case (accusative or ergative).
 - b. Non-systematic syncretisms result either from phonological conflation or from purely accidental homonymy. In this latter case, they can never involve multiple underlyingly distinct allomorphs.

To test this generalization, I have put together the language samples that were used in five different sources:

- (i) Baerman *et al.* 2005 — a WALS-based sample of 197 languages, mostly from outside Indo-European;
- (ii) Plank 1991b — which contributes twelve more languages, all from Indo-European except for Classical Arabic;
- (iii) Caha 2009 and Caha 2010 — which collectively contribute six more languages, all Indo-European except for Northern Saami and Bilin;
- (iv) the Surrey Syncretism Database⁸⁸ — which contributes ten more languages, all of them non-Indo-European except for Kashmiri.

This gives us a sizeable total of 225 languages, most of which are, however, not directly relevant for our purposes, either because case-marking is absent or near-absent in their

⁸⁸<http://www.smg.surrey.ac.uk/syncretism>

grammars (this is so for 123 languages from the *WALS*' sample, and for one language from the Surrey Database), or because case-marking is present but is never syncretic (such is the case for 33 languages⁸⁹ of the *WALS*' sample).⁹⁰ Once all of these irrelevant languages are factored out, we are left with a considerably downsized sample of 68 languages, of which 26 are Indo-European. These include:

- (79) a. Classical Armenian, Modern Eastern Armenian, Czech, Old and Modern English, French, Old Frisian, German, Gothic, Classical and Modern Greek, Hindi, Icelandic, Irish, Kashmiri, Latin, Latvian, Lithuanian, Old Norse, Polish, Russian, Sanskrit, Serbian, Slovene, Spanish, Ukrainian (Indo-European);
- b. Classical Arabic, Araona, Bao'an, Basque, Bilin, Boumaa Fijian, Burushaski, Chukchi, Comanche, Diyari, Finnish, Georgian, West Greenlandic, Harar Oromo, Ingush, Koryak, Krongo, Lak, Lezgian, Mangarayi, Martuthunira, Murle, Nenets, Nez Perce, Ngiyambaa, Paumari, Pitjantjatjara, Rongpo, Northern Saami, Somali, Suena, Telugu, Tsakhur, Wambaya, Warao, Yaqui, Yidiny, Yukaghir, Yup'ik, Yurok (non-Indo-European).

Fourteen of these languages contain surface counterexamples to the supposed generalization in (78).⁹¹ They are listed in (80), with italics signaling the non-Indo-European ones.

- | | | |
|------|--|-----------------------------------|
| (80) | (i) <i>Classical Arabic</i> : NOM-GEN; | (viii) <i>Ingush</i> : NOM-GEN; |
| | (ii) <i>Bilin</i> : NOM-GEN; | (ix) Kashmiri: ABS-ABL; |
| | (iii) Czech: NOM-INSTR; | (x) Latin: NOM-GEN; ⁹² |
| | (iv) <i>Georgian</i> : ABS-GEN-INSTR; | (xi) Latvian: NOM-GEN; |
| | (v) Gothic: NOM-GEN; | (xii) Lithuanian: NOM-INSTR; |
| | (vi) Classical Greek: NOM-DAT; | (xiii) Polish: NOM-GEN-DAT-LOC; |
| | (vii) Icelandic: NOM-GEN; | (xiv) Slovene: NOM-INSTR. |

⁸⁹Baerman *et al.* (2005: 227) report a figure of 34, but this results from an error of categorization. Contrary to what these authors report, Nez Perce is not exempt from case syncretism. Rather, Nez Perce first- and second-person pronouns conflate ergative and S-nominative together (cf. Deal 2010: 34).

⁹⁰Yet another, less common reason for irrelevance is the lack of a *bona fide* accusative case. Such is apparently the case for Kayardild (cf. Evans (1995: 2): 'There is no accusative case in Kayardild'). If this is so, then the NOM-LOC syncretism of one Kayardild declension (Evans 1995: 125) is in fact not a counterexample at all.

⁹¹This figure would add up to 15 if we were to include Comanche amongst the problematic languages, as Baerman *et al.* (2005) would appear to suggest. Their analysis is, however, the result of a misunderstanding of Charney's (1993: 9ff, 56ff) reference grammar. In fact, Comanche's nominative and genitive do differ, in that the latter alone triggers gemination of the initial consonant of the following word.

⁹²Here I am ignoring the extremely isolated NOM-INSTR syncretism characteristic of the masculine relative pronoun, which I have already discussed in § 1.2.

While many of these counterexamples lend themselves to fairly plain phonological explanations, or are isolated enough to be plausibly regarded as accidental homonymies, there do remain some more recalcitrant cases which arguably require that the genitive be afforded a special status among obliques. I shall start by discussing this first, and then move on to the other (I believe, more trivial) purported problems.

5.1 Problems with the genitive

5.1.1 NOM-GEN syncretisms in Latin, Gothic, Latvian, and Bilin

It may be interesting to note, first of all, that more than half of the above-mentioned counterexamples involve syncretism with the *genitive*. This asymmetry may at first be imputed to the fact that the genitive is expressed synthetically in more languages than any other oblique case (cf. Blake 2001: § 5.8). This explanation, however, cannot be sufficient, given that, with the exception of Classical Arabic, all the problematic languages at issue have other obliques alongside the genitive in their case inventories. The asymmetry is thus a first reason to suspect that there may be something else going on.

This suspicion is corroborated by the fact that several of these counterexamples are not quite so easy to explain away as are those involving the other obliques. Let us take, for example, the Latin paradigms in table 5.1.

	‘town’	‘citizen’	‘dog’
NOM	urbs	cīvis	canis
GEN	urbis	cīvis	canis
DAT	urbī	cīvī	canī
ACC	urbem	cīvem	canem
ABL	urbe	cīvī/cīve	cane

Table 5.1: Three singular paradigms from Latin

Johnston (1996: 50–51) tries to explain away this unexpected NOM-GEN syncretism via appeal to ambiguous segmentation. As he correctly points out, comparison with a less tricky paradigm like that of ‘town’ suggests that the desinences for NOM.SG and GEN.SG may not be the same: rather, the former appears to be just *-s* (cf. *urb-s*), while the latter appears to be *-is* (cf. *urb-is*). In this light, then, it seems plausible to analyze NOM.SG *cīvis* as *cīvi-* + *-s*, and GEN.SG *cīvis* as *cīvi-* + *-is* (via contraction of *i* + *i* = *ī*, the length distinction being neutralized before final *-s*). The origin of this syncretism would then be merely phonological.

A major problem with this analysis, however, is that the same NOM-GEN syncretism of *cīvīs* also developed in paradigms where the theme did not actually end with *-i*. This is the case, for example, of *canis* ‘dog’, whose theme did not originally contain any *i* (cf. Classical Greek *kýōn*, GEN *kynós*, and Sanskrit *svan*, GEN *sunah*). Rather, it appears that the *i* in its nominative was inserted between *can-* and the historical suffix *-s* at some point in Latin’s prehistory. But why should it have been inserted in the first place? A plausible explanation seems to involve precisely the analogical influence of the genitive. Thus Weiss (2009: 248), for one, explains the presence of an alternative nominative *bovis* ‘ox’, alongside *bōs* and *būs*, by saying: ‘A nominative *bovis* [...] was created on the model of the parisyllabic *i*-stems of the type NOM *cīvīs*, GEN *cīvīs*.’ Clearly, homonymies resulting from such analogical processes cannot be seriously regarded as accidental.

Much the same problem arises in the following Gothic paradigms.

	‘shepherd’, SG	‘army’, SG
NOM	haírdeis	harjis
ACC	haírdi	*hari
GEN	haírdeis	harjis
DAT	haírdja	harja

Table 5.2: Fragments of two Gothic paradigms (Kieckers 1960: 109–110)

Here, too, a phonological NOM-GEN syncretism in one paradigm provided a model for analogical formation of a *non*-phonological NOM-GEN syncretism in another paradigm. More specifically, Kieckers (1960: 110) comments: ‘The NOM *harjis* stands in place of the regular **haris*, in which a *j* was introduced from the GEN and DAT. At the same time, the fact that, in the type *haírdeis*, NOM and GEN conflated together by virtue of phonological change caused the same thing to happen in the type *harjis* as well.’⁹³ Once again, if analogy is really playing a role, any appeal to fortuitous homonymy cannot hold water.

Even more severe is the problem posed by Latvian. In this language, there are at least two different exponents, *-s* and *-ās*, which syncretically express both NOM.SG and GEN.SG. To these two we should probably also add *-us*—although it may ultimately prove feasible, in this latter case, to separate out the *-u-* as a marker of singular number, thus reducing the problem again to *-s*. Be that as it may, *-s* and *-ās* alone are already sufficient, by virtue of the criteria in (78b), to rule out any appeal to accidental homonymy in this case as well.

⁹³This is my own English translation. The German original reads: ‘Der Nom. *harjis* steht für lautgesetzliches **haris*, indem das *j* aus dem Gen., Dat. eingeführt wurde. Zugleich hat der Umstand, daß bei dem Typus *haírdeis* durch lautliche Entwicklung Nom. und Gen. zusammenfielen, bewirkt, daß das gleiche beim Typus *harjis* eintrat.’

	‘stone’, SG	‘meeting’, SG	‘market’, SG
NOM	akmens	tikšanās	tirgus
GEN	akmens	tikšanās	tirgus
ACC	akmeni	tikšanos	tirgu

Table 5.3: Fragments of three Latvian paradigms (Praulīšs 2012: 28, 30, 35)

Finally, the same problem also appears to arise in the Cushitic language Bilin, whose nominal morphology was described by Palmer (1958) and, more recently, by Appleyard (2007). Here, too, nominative and genitive can be syncretic both in the singular and in the plural, involving two different exponents, \emptyset and *-i*.

	‘dog’, SG	‘owner’, PL
NOM	gədəŋ	wanni
ACC	gədəŋ-sí	wannós
GEN	gədəŋ	wanni
DAT	gədəŋ-ád	wannəd

Table 5.4: Bilin (Palmer 1958: 382)

With this background in place, the fact that the NOM-GEN syncretisms of Icelandic and Arabic are amenable to plain phonological accounts, or that the analogous syncretism in Ingush is restricted to the sole exponent *-(i)i*, can hardly be of any interest.⁹⁴ The evidence from Latvian and Bilin (and perhaps also from Latin and Gothic) already suffices to show that syncretisms between nominative and genitive ‘across’ the accusative can sometimes be systematic, and hence that the genitive cannot always be structurally bigger than the clause-level dependent case across all languages.

5.1.2 Converging evidence from overt containment and suppletion

Before seeking a plausible explanation for this state of affairs, it may be worth pointing out that the evidence from syncretism reviewed in § 5.1.1 is by no means isolated. On the contrary, similar converging indications that the genitive does not always behave like an oblique are also provided both by case-driven suppletion and by overt containment patterns. In this latter connection, Caha (2010: 72fn13) notes, for example, that in languages like Russian and Lule Saami the plural genitive seems to be overtly contained in the corresponding accusative, as is shown in table 5.5.

⁹⁴On Icelandic and Ingush, cf. Einarsson (1949: 55) and Nichols (2011: 33, 138), respectively. On Classical Arabic, cf. Johnston (1996: 24–25), whose results I reviewed in § 1.2.

Russian	‘woman’, PL	Lule Saami	‘house’, PL
NOM	komnat-y	NOM	goad-e
ACC	komnat-y	ACC	gåd-ij-t
GEN	komnat	GEN	gåd-ij
PREP	komnat-ax	INES	gåd-ij-n
DAT	komnat-am	ILL	gåd-ij-da
INSTR	komnat-ami	ELAT	gåd-ij-s

Table 5.5: Problematic paradigms in Russian and Lule Saami (Caha 2010: 43)

Caha (2010: 72fn13) then tries to explain away these examples via appeal to a phonologically empty genitive affix (e.g. *komnat-∅*).⁹⁵ Though clearly viable for Russian, this strategy is, however, somewhat less easy to apply to Lule Saami, where the genitive affix seems to be already overtly realized as *-ij*: Caha is thus forced to reanalyze this *-ij* as a plural marker, and the nominative *-e* as a number-case portmanteau. Even more problematic is the case of the Lezgian language Kryz, where the ergative is very systematically formed out of the corresponding genitive, with very little wiggle room left for empty affixes or other alternative accounts. The situation exemplified in table 5.6 is summarized as follows by the leading specialist in the language, Gilles Authier (2009: 29):

Built out of the absolutive [...] by means of various suffixes, the genitive constitutes the ‘oblique stem’ to which the other desinences are added. Inflection is identical for all nominals except for the genitive, whose formation out of the absolutive varies depending on semantic factors^[96].

	‘tongue’, SG	‘married man’, SG	‘house’, SG
ABS	mez	furi	k’ul
GEN	miz	fura	k’ul-ci
ERG	miz-ir	fura-r	k’ul-ci-r
INSTR	miz-zina	fura-zina	k’ul-ci-zina

Table 5.6: Problematic paradigms in Kryz (Authier 2009: 30, 33, 194, 381)

Finally, the possibility that the genitive may sometimes be smaller than the clause-level dependent case is also suggested by the following suppletion pattern, characteristic

⁹⁵The same problem (along with the same possible solution) can also be found in the singular paradigm of Bilin *gədaŋ*, reported in table 5.4.

⁹⁶My own translation into English. The French original reads: ‘Formé par divers suffixes à partir de l’absolutif [...], le génitif est la “base oblique” à laquelle s’adjoignent les autres désinences. La flexion est identique pour tous les nominaux en dehors du génitif, dont la formation à partir de l’absolutif est motivée sémantiquement.’

of Icelandic's 1PL honorific pronoun, 'used in royal letters and proclamations[...] by editors of newspapers, and by some writers' (Einarsson 1949: 122).

NOM	vér
ACC	oss
GEN	vor
DAT	oss

Table 5.7: Icelandic 1PL honorific (Einarsson 1949: 68)

Here, the nominative and the genitive share the same root to the exclusion of the suppletive accusative case form, contrary to what Smith *et al.* (2016) predict should happen if the genitive were a run-of-the-mill oblique.

To sum up: all three of the containment diagnostics used so far—syncretism, allomorphy, and overt containment—consistently show that the genitive behaves sometimes like an oblique and sometimes more like an unmarked case.

5.1.3 Towards an explanation: genitive as an unmarked case

A plausible explanation for the apparent possibility of systematic syncretisms between NOM/ABS and GEN may reside in Baker's (2015: § 4.4.1) idea that the genitive might be either an unmarked case within the DP/NP domain or an inherent case, depending on the language (and perhaps variably even within one and the same language). The genitive, in other words, may be assigned, sometimes at least, not via the same mechanism that assigns locatives or instrumentals but rather by virtue of a rule like (81) (from Baker 2015: 166).

(81) If NP is not otherwise case-marked when DP/NP is spelled out, assign it *genitive*.

On this account, the genitive's variable morphological behavior would reflect a deeper kind of ambivalence in this case's syntactic nature and modes of assignment. This would lead one to expect, then, to find some degree of correlation between the syntactic and the morphological behavior of the genitive within each given language or construction; it would lead one to expect, more specifically, that those languages in which the genitive looks morphologically 'lighter' than the ergative or accusative should only assign genitive case adnominally via a rule like (81), and never use it as a θ -related case governed or selected by the verb. I am not sure, however, that this latter expectation is borne out, given that Latin, Gothic, and Icelandic all clearly admit of inherent uses of the genitive. I confess that I have no idea, at this time, as to how to make better sense of this mismatch. I

will thus have to leave this issue at that, simply acknowledging that the genitive may well have to be left out of the case hierarchy, pending further clarification of its syntactic and morphological nature(s).

5.2 (Pseudo-)problems with the other obliques

5.2.1 Some unproblematic counterexamples

Once the genitive is left to the side, it happily turns out that most of the remaining surface counterexamples can easily be shown to be either the by-product of phonological rules or totally isolated accidental homonymies. Clearly accidental, for example, is the NOM-GEN-DAT-LOC syncretism of Polish, which, as noted by Johnston (1996: 61), involves only the exponent *-i* and targets only derivative female nouns whose stem ends in *-ń*. Phonological conflation, on the other hand, is ostensibly responsible for the NOM-INSTR syncretisms of both Czech and Slovene, as has convincingly been argued by Caha (2009: 243, 249ff).⁹⁷

A phonological account is also available for Lithuanian, whose NOM.SG and INSTR.SG desinences commonly spelt as *-a* are in fact characterized by different accentual properties, the former being ‘strong attractive’, the latter ‘weak attractive’ (cf. Ambrazas 1997: 78–81). Though the contrast is most often neutralized by the interactions with the phonological properties of the stem, there are a few contexts where it does make a difference in the actual surface form—namely, when the stem is, at the same time, ‘weak’ and ‘acuted’, or when the stem’s stressed syllable and the desinence are separated by at least one intermediate syllable. In both of these contexts, main stress will ultimately fall on the desinence in the nominative, but on the stem in the instrumental.

	‘head’, SG	‘depth’, SG
NOM	galvà	gillumà
ACC	gálvą	gílumą
INSTR	gálva	gíluma

Table 5.8: Fragments of two Lithuanian paradigms (Ambrazas 1997: 80–81)

Finally, a somewhat less exciting problem arises in Georgian, where adjectives with

⁹⁷ Adrian Stegovec (p.c.) informs me that Caha’s phonological argument about Slovene, based as it was on tonal differences like that between NOM.PL *kováč-i* and INSTR.PL *kovâč-i* (‘blacksmith’), does not extend to the varieties of spoken Slovene which lack tonal distinctions, as is the case for Adrian’s own variety. However, even if the two suffixes *-i* should ultimately turn out to be phonologically identical, there would still be no other suffix involved in the syncretism, such that this counterexample might in any event be regarded as the result of an accident.

consonant stems show the kind of reduced declension exemplified by *dzveli* ‘old’ in table 5.9, with an unexpected syncretism of absolutive, genitive, and instrumental.

	old	man, SG	man, PL (archaic)
ABS	dzvel-i	k’ac-i	k’ac-n-i
ERG	dzvel-ma	k’ac-ma	k’ac-t
DAT	dzvel	k’ac-s	k’ac-t
ADV	dzvel	k’ac-ad	k’ac-t
GEN	dzvel-i	k’ac-is	k’ac-t
INSTR	dzvel-i	k’ac-it	k’ac-t

Table 5.9: Georgian paradigms (Hewitt 1995: 34, 45)

This pattern is neither as morphophonologically circumscribed as the Polish one nor as amenable to a phonological account as the Balto-Slavic ones. It is, however, restricted to the sole exponent *-i*, and is not replicated in the even more reduced declension commonly known as the ‘Old Georgian plural’, where genitive and instrumental syncretize, as expected, with the ergative, rather than with the absolutive. Pending further evidence to the contrary, I shall therefore regard this syncretism as an instance of purely accidental homonymy.

5.2.2 More problematic counterexamples:

Classical Greek and Kashmiri

Genitives apart, there are, in my whole sample, only two languages that may really look problematic for the generalization in (78): Kashmiri and Ancient Greek. Neither, however, seems to me to constitute a *bona fide* problem if all relevant details are taken into account.

Let us begin with Ancient Greek, whose syncretism between NOM.SG and DAT.SG purports to involve the two different exponents *-ā* and *-ē*.

	house, SG	honor, PL	‘tongue’, SG	‘steward’, SG
NOM	oikíā	timé	glóssa	tamíās
ACC	oikíān	timén	glóssan	tamíān
DAT	oikíā(i)	timē(i)	glóssē(i)	tamíā(i)
GEN	oikíās	timēs	glóssēs	tamíū

Table 5.10: Ancient Greek

What is unclear, though, is whether *ā* and *ē* really constitute two phonologically distinct allomorphs, rather than two allophones of one and the same allomorph. Crucial to this question is Johnston’s (1996: 41) insightful comment:

the homonyms are co-extensive under the limited amount of allomorphy involved in the distinction between *-a:* and *-e:*. But even this is arguably an automatic phonological alternation, since underlying *a:* generally surfaces as *e:* in Ionic dialects, except that in Attic *a:* generally remains after *r*, *e* and *i*.

Significantly, apart from this phonologically conditioned alternation, there is no further covariation between the two exponents at issue—not even internal to their declension class, as exemplified by the paradigms of ‘tongue’ and ‘steward’ in table 5.10. On these grounds, Johnston (1996) opts for treating this syncretism as yet another example of accidental homonymy.

Though this is certainly a plausible and welcome result, we might, however, do even better, and seriously entertain the claim that the offending NOM-DAT syncretism never actually existed at any moment in the history of spoken Greek. There is, indeed, little question that the syncretism did not exist in Classical Attic, in which the relevant first-declension endings for NOM.SG and DAT.SG were respectively *ā/ē* and *āi/ēi*, with further accentual differences between the two case forms in the paradigms of oxytone nouns: in the paradigm of *timē* ‘honor’ in table 5.10, for example, the accent is obligatorily acute in the nominative but obligatorily circumflex in the dative. Both the segmental and the accentual distinctions, however, faded away during the Hellenistic period, due to a rule that deleted the final *-i* in long-onset diphthongs, and due to a general breakdown of the classical language’s accentual system. This is, however, the very same period in which the final *-n* of the accusative ending also began to fade away, turning the unexpected *ABA pattern into a much more trivial AAA: NOM *oikía* ~ ACC *oikía* ~ DAT *oikía*. This also corresponded, interestingly enough, to a gradual erosion of the functions of the dative case—a development which would ultimately lead to the ancient dative’s total disappearance in Modern Greek as we know it today. In this connection, it is useful to quote Horrocks (2010: 116) at some length:

With the loss of the final *i*-element of the long diphthongs and the equalization of vowel length the dative singular of many classes of noun became virtually homophonous with the accusative singular (given the weakness of final *-v* [-n]). Thus 1st-declension *χώραν* [‘k^hora(n)]/χώρα [‘k^hora] ‘country’, 2nd-declension *λόγον* [‘loɣo(n)]/λόγῳ [‘loɣo] ‘word’, etc. This inevitably led to further overlapping usages even where the accusative and dative forms were clearly distinguished, as in 1st/2nd-declension plural and in the 3rd declension.

It is thus tempting to suppose that, as soon as the offending NOM-DAT homophony started to look like a systematic syncretism, the whole case system of the language somehow got reshaped to bring it back into line with the universal case containment hierarchy.

Finally, let us conclude with Kashmiri, which the Surrey Syncretism Database⁹⁸ reports to exhibit the following two patterns of ABS-ABL syncretism.

	hand, SG	hand, PL	child, SG	‘goose’, SG
ABS	ath-i	ath-i	gobur	ənz
ERG	ath-an	ath-av	gobr-an	ənz-in
ABL	ath-i	ath-av	gobr-i	ənz
DAT	ath-as	ath-an	gobr-as	ənz-is

Table 5.11: Kashmiri (Kachru 1969: 112–114; SSD report)

If these data were correct, there would indeed be good reasons to acknowledge the existence of a systematic syncretism between absolutive and ablative in this language, given that the homonymy would straddle two phonologically unrelated exponents such as *i* and \emptyset . Fortunately, however, there are fairly good reasons to doubt the accuracy of the relevant data. Matthew Baerman, the author of the SSD’s report on Kashmiri, cautiously points out that the ‘goose’-type paradigms ‘are cited by Grierson^[99] and by Zaxarin and Edel’man ^[100], but not in any other source’: in particular, there is no mention of such paradigms in the grammars of Braj B. Kachru (1969) or of Wali and Koul (1996), all native speakers of the language; furthermore, Omkar N. Koul (p.c.) has informed me that the only possible ablative form he knows of for *ənz* ‘goose’ is the totally unproblematic *ənz-i*. In light of these problems, I will therefore put aside the problematic forms reported by Grierson, as perhaps resulting from an admixture of different dialects, while obviously acknowledging that I cannot definitively dismiss them right now. I shall thus focus only on the ABS-INSTR syncretism instantiated by the suffix *-i*, whose accidental nature clearly emerges, I believe, from the comparison with the paradigms of ‘hands’ (PL) and ‘child’ (SG) in table 5.11: as is apparent from that table, the exponent *-i* can both realize the absolutive without realizing the ablative and vice versa; the appearance of syncretism only arises when the two exponence patterns accidentally intersect.

Greek and Kashmiri notwithstanding, I shall thus conclude that the generalization in (78) boasts robust empirical support, provided that we are careful enough in examining all the purported counterexamples (see especially the Kashmiri case above), and provided we allow for enough latitude regarding the cross-linguistic status of the genitive as an oblique.

⁹⁸www.smg.surrey.ac.uk/syncretism/reports/Kashmiri.pdf

⁹⁹George A. Grierson, *A manual of the Kashmiri language, comprising grammar, phrase-book, and vocabularies*. Oxford: Clarendon Press, 1911.

¹⁰⁰Boris A. Zaxarin and Džoj I. Edel’man, *Jazyk kašmīri*. Moscow: Nauka, 1971.

Bibliography

- ABELS, Klaus and Peter MURIUNGI (2008). 'The focus marker in Kîitharaka: Syntax and semantics'. *Lingua* 118: 687–731.
- AISSEN, Judith (2003). 'Differential object marking: Iconicity vs. economy'. *Natural Language & Linguistic Theory* 21: 435–483.
- ALBRIGHT, Adam and Eric FUSS (2012). 'Syncretism'. In Jochen TROMMER, ed., *The morphology and phonology of exponence*, pp. 236–288. Oxford: Oxford University Press.
- ALDRIDGE, Edith (2008). 'Generative approaches to ergativity'. *Language and Linguistics Compass* 2: 966–995.
- ALLEN, W. Sidney and Charles O. BRINK (1980). 'The old order and the new: A case history'. *Lingua* 50: 61–100.
- AMBRAZAS, Vytautas, ed. (1997). *Lithuanian grammar*. Vilnius: Baltos Lankos.
- ANDERSON, Neil and Martha WADE (1988). 'Ergativity and control in Fologa'. *Language and Linguistics in Melanesia* 19: 1–16.
- APPLEYARD, David L. (2007). 'Bilin morphology'. In Alan S. KAYE, ed., *Morphologies of Asia and Africa*. Vol. 1, pp. 481–504. Winona Lake, IN: Eisenbrauns.
- ARENSEN, Jon (1982). *Murle grammar*. Juba: University of Juba/Summer Institute of Linguistics/Institute of Regional Languages.
- AUSTIN, Peter (2011). *A grammar of Diyari, South Australia*. 2nd edn. Cambridge: Cambridge University Press. [1st edn., 1981.]
- AUTHIER, Gilles (2009). *Grammaire kryz (langue caucasique d'Azerbaïdjan, dialecte d'Alik)*. Paris: Peeters.
- (2013). 'Inalienability split in possessive NPs and the origin of the two cases with genitive function in Budugh'. *Faits de Langues* 41: 177–192.
- BAERMAN, Matthew, Dunstan BROWN, and Greville G. CORBETT (2005). *The syntax–morphology interface: A study of syncretism*. Cambridge: Cambridge University Press.
- BAKER, Mark C. (2014). 'Types of cross-linguistic variation in case assignment'. In M. Carme PICALLO, ed., *Linguistic variation in the Minimalist framework*, pp. 36–60. Oxford: Oxford University Press.
- (2015). *Case: Its principles and its parameters*. Cambridge: Cambridge University Press.

- BAKER, Mark C. and Jonathan David BOBALJIK (forthcoming). 'On inherent and dependent theories of ergative case'. In Jessica COON, Diane MASSAM, and Lisa TRAVIS, eds., *The Oxford handbook of ergativity*. Oxford: Oxford University Press.
- BAKER, Peter S. (2003). *Introduction to Old English*. Oxford: Blackwell.
- BERGER, Hermann (1998). *Die Burushaski-Sprache von Hunza und Nager. Teil I: Grammatik*. Wiesbaden: Harrassowitz.
- BHAT, D.N.S. and M.S. NINGOMBA (1997). *Manipuri grammar*. München: LINCOM.
- BITTNER, Maria and Kenneth HALE (1996). 'The structural determination of case and agreement'. *Linguistic Inquiry* 27: 1–68.
- BLAKE, Barry J. (2001). *Case*. 2nd edn. Cambridge: Cambridge University Press. [1st edn., 1994.]
- BLIX, Hagen (2016). *South-Caucasian agreement: A spanning account*. MA thesis. University of Vienna.
- BOBALJIK, Jonathan David (2008). 'Where's Phi? Agreement as a post-syntactic operation'. In Daniel HARBOUR, David ADGER, and Susana BÉJAR, eds., *Phi theory: Phi-features across interfaces and modules*, pp. 295–328. Oxford: Oxford University Press.
- (2012). *Universals in comparative morphology: Suppletion, superlatives, and the structure of words*. Cambridge, MA: MIT Press.
- (2015). 'Suppletion: Some theoretical implications'. *Annual Review of Linguistics* 1: 1–18.
- BOBALJIK, Jonathan David and Uli SAUERLAND (2017). '*ABA and the combinatorics of morphological features'. Available at: ling.auf.net/lingbuzz/003320.
- BONET, Eulàlia and Daniel HARBOUR (2012). 'Contextual allomorphy'. In Jochen TROMMER, ed., *The morphology and phonology of exponence*, pp. 195–235. Oxford: Oxford University Press.
- BOSSONG, Georg (1985). *Empirische Universalienforschung. Differenzielle Objektmarkierung in der neuiranischen Sprachen*. Tübingen: Narr.
- BOWE, Heather (1990). *Categories, constituents, and constituent order in Pitjantjatjara, an Aboriginal language of Australia*. London: Routledge.
- BRAUNE, Wilhelm and Frank HEIDERMANNS (2004). *Gotische Grammatik mit Lesestücken und Wörterverzeichnis*. 20th edn. Tübingen: Niemeyer. [1st edn., by BRAUNE alone, 1880. Subsequent editions revised by Karl HELM, Ernst A. EBBINGHAUS, and Frank HEIDERMANNS.]
- BREEN, John G. (1976). 'Ergative, locative and instrumental inflections in Wangkumara'. In R.M.W. DIXON, ed., *Grammatical categories in Australian languages*, pp. 336–339. Canberra: Australian Institute of Aboriginal Studies.
- (1981). 'Margany and Gunya'. In R.M.W. DIXON and Barry J. BLAKE, eds., *Handbook of Australian languages*. Vol. 2, pp. 274–393. Amsterdam: John Benjamins.
- BRUGMANN, Karl (1911). *Grundriss der vergleichenden Grammatik der indogermanischen Sprachen*. 2nd edn. Vol. 2: *Lehre von den Wortformen und ihrem Gebrauch*. 2nd part. Strassburg: Trübner. [1st edn., by BRUGMANN and Berthold DELBRÜCK, 1892.]
- BRUNNER, Karl (1965). *Altenglische Grammatik*. 3rd edn. Tübingen: Niemeyer. [1st edn., 1942.]

- CAHA, Pavel (2009). *The nanosyntax of case*. PhD thesis. University of Tromsø.
- (2010). 'The parameters of case marking and spell out driven movement'. *Linguistic Variation Yearbook* 10: 33–78.
- (2013). 'Explaining the structure of case paradigms by the mechanisms of Nanosyntax: The Classical Armenian nominal declension'. *Natural Language & Linguistic Theory* 31: 1015–1066.
- (2016). 'Notes on insertion in Distributed Morphology and Nanosyntax'. Available at: ling.auf.net/lingbuzz/002855.
- (n.d.). 'Suppletion and morpheme order: A unified account'. Review article on BOBALJIK (2012). Ms.
- CALABRESE, Andrea (2008). 'On absolute and contextual syncretism: Remarks on the structure of case paradigms and on how to derive them'. In Asaf BACHRACH and Andrew NEVINS, eds., *Inflectional identity*, pp. 156–205. Oxford: Oxford University Press.
- CAMPBELL, Alistair (1959). *Old English grammar*. Oxford: Clarendon Press.
- CENAMO, Michela (2011). 'Case marking of core arguments and alignment in Late Latin'. In Michèle FRUYT, Michel MAZOYER, and Denis PARDEE, eds., *Grammatical case in the languages of the Middle East and Europe. Acts of the International colloquium "Variations, concurrence et évolution des cas dans divers domaines linguistiques"* (Paris, 2–4 April 2007), pp. 297–314. Chicago: University of Chicago Press.
- CHARNEY, Jean Ormsbee (1993). *A grammar of Comanche*. Lincoln: University of Nebraska Press.
- CHOMSKY, Noam (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.
- (2000). 'Minimalist inquiries: The framework'. In Roger MARTIN, David MICHAELS, and Juan URIAGEREKA, eds., *Step by Step*, pp. 91–155. Cambridge, MA: MIT Press.
- (2001). 'Derivation by phase'. In Michael KENSTOWICZ, ed., *Ken Hale: A life in language*, pp. 1–52. Cambridge, MA: MIT Press.
- CINQUE, Guglielmo (2005). 'Deriving Greenberg's Universal 20 and its exceptions'. *Linguistic Inquiry* 36: 315–332.
- CINQUE, Guglielmo and Luigi RIZZI (2010). 'The cartography of syntactic structures'. In Bernd HEINE and Heiko NARROG, eds., *The Oxford handbook of linguistic analysis*, pp. 51–65. Oxford: Oxford University Press.
- COLARUSSO, John (1992). *A grammar of the Kabardian language*. Calgary: University of Calgary Press.
- COMRIE, Bernard (2005). 'Alignment of case marking of full noun phrases'. In Martin HASPELMATH, Matthew S. DRYER, David GIL, and Bernard COMRIE, eds., *World Atlas of Language Structure*, pp. 398–403. Oxford: Oxford University Press.
- CREISSELS, Denis (2009). 'Spatial cases'. In MALCHUKOV and SPENCER, pp. 609–625.
- DEAL, Amy Rose (2010). *Topics in the Nez Perce verb*. PhD thesis. Amherst: University of Massachusetts.

- DEAL, Amy Rose (2015). 'Ergativity'. In Tibor KISS and Artemis ALEXIADOU, eds., *Syntax – Theory and analysis: An international handbook*. Vol. 1, pp. 654–707. Berlin: Mouton de Gruyter.
- DEMIROK, Ömer Faruk (2013). *AGREE as a unidirectional operation: Evidence from Laz*. MA thesis. Istanbul: Boğaziçi University.
- DIXON, R.M.W. (1972). *The Dyirbal language of North Queensland*. Cambridge: Cambridge University Press.
- (1994). *Ergativity*. Cambridge: Cambridge University Press.
- (2012). *Basic linguistic theory*. Vol. 3: *Further grammatical topics*. Oxford: Oxford University Press.
- DOWTY, David R. (1991). 'Thematic proto-roles and argument selection'. *Language* 67: 547–619.
- EINARSSON, Stefán (1949). *Icelandic: Grammar, texts, glossary*. Baltimore: Johns Hopkins University Press.
- EMBICK, David and Rolf NOYER (2007). 'Distributed Morphology and the syntax–morphology interface'. In Gillian RAMCHAND and Charles REISS, eds., *The Oxford handbook of linguistic interfaces*, pp. 289–324. Oxford: Oxford University Press.
- ERNOUT, Alfred (1953). *Morphologie historique du latin*. 3rd edn. Paris: Klincksieck. [1st edn., 1914.]
- EVANS, Nicholas D. (1995). *A grammar of Kayardild*. Berlin: Mouton de Gruyter.
- FELL, Brian (2012). 'Case and agreement in Ubykh'. Ms.
- FRIEDMAN, Victor A. (1991). 'Romani nominal inflection: Cases or postpositions?' In Maciej GROCHOWSKI, ed., *Problemy opisu gramatycznego języków słowiańskich*, pp. 57–63. Warsaw: Polska Akademia Nauk.
- GIPPERT, Jost (1987). 'Zu den sekundären Kasusaffixen des Tocharischen'. *Tocharian and Indo-European Studies* 1: 22–39.
- GODDARD, Cliff (1982). 'Case systems and case marking in Australian languages: A new interpretation'. *Australian Journal of Linguistics* 2: 167–196.
- GRAF, Thomas (2017). 'Graph transductions and typological gaps in morphological paradigms'. Available at: ling.auf.net/lingbuzz/003373.
- GREENBERG, Joseph H. (1966). 'Some universals of grammar with particular reference to the order of meaningful elements'. In J.H. GREENBERG, ed., *Universals of language*, pp. 73–113. Cambridge, MA: MIT Press.
- GUIRARDELLO, Raquel (1999). *A reference grammar of Trumai*. PhD thesis. Rice University.
- HALE, Kenneth and Samuel Jay KEYSER (1993). 'On argument structure and the lexical expression of syntactic relations'. In Kenneth HALE and Samuel Jay KEYSER, eds., *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, pp. 53–109. Cambridge, MA: MIT Press.

- HALLE, Morris (1997). 'Distributed Morphology: Impoverishment and Fission'. In Benjamin BRUENING, Yoonjung KANG, and Martha MCGINNIS, eds., *MIT Working Papers in Linguistics 30: PF: Papers at the interface*, pp. 425–449. Cambridge, MA: MIT.
- HALLE, Morris and Alec MARANTZ (1993). 'Distributed Morphology and the pieces of inflection'. In Kenneth HALE and Samuel J. KEYSER, eds., *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, pp. 111–176. Cambridge, MA: MIT Press.
- HARBOUR, Daniel (2008). 'On homophony and methodology in morphology'. *Morphology* 18: 75–92.
- HARÐARSON, Gísli Rúnar (2016). 'A case for a Weak Case Contiguity hypothesis—a reply to Caha'. *Natural Language & Linguistic Theory* 34: 1329–1343.
- HARRIS, Alice C. (1981). *Georgian syntax: A study in relational grammar*. Cambridge: Cambridge University Press.
- HASPELMATH, Martin (1993). *A grammar of Lezgian*. Berlin: Mouton de Gruyter.
- (2016). 'Number suppletion vs. case suppletion: Does “locality” provide an explanation?' Available at: <http://dlc.hypotheses.org/902>.
- HEWITT, B. George (1995). *Georgian: A structural reference grammar*. Amsterdam: John Benjamins.
- HOLISKY, Dee Ann (1987). 'The case of the intransitive subject in Tsova-Tush (Batsbi)'. *Lingua* 71: 103–132.
- HOLISKY, Dee Ann and Rusudan GAGUA (1994). 'Tsova-Tush (Batsbi)'. In Rieks SMEETS, ed., *The indigenous languages of the Caucasus*. Vol. 4, pp. 147–212. Delmar, NY: Caravan Books.
- HORROCKS, Geoffrey (2010). *Greek: A history of the language and its speakers*. 2nd edn. Oxford: Wiley-Blackwell. [1st edn. London: Longman, 1997.]
- HUALDE, José Ignacio (2003). 'Case and number inflection of noun phrases'. In J.I. HUALDE and Jon ORTIZ DE URBINA, eds., *A grammar of Basque*, pp. 171–187. Berlin: Mouton de Gruyter.
- JAKOBSON, Roman (1936). 'Beitrag zur allgemeinen Kasuslehre'. *Travaux du Cercle Linguistique de Prague* 6: 240–288.
- JOHNS, Alana, Diane MASSAM, and Juvenal NDAYIRAGIJE, eds. (2006). *Ergativity: Emerging issues*. Dordrecht: Springer.
- JOHNSTON, Jason C. (1996). *Systematic homonymy and the structure of morphological categories: Some lessons from paradigm geometry*. PhD thesis. University of Sydney.
- KACHRU, Braj B. (1969). *A reference grammar of Kashmiri*. Urbana, Ill.: University of Illinois.
- KATZ, Joshua T. (1998). *Topics in Indo-European personal pronouns*. PhD thesis. Harvard University.
- KAYNE, Richard S. (2005). 'Some notes on comparative syntax, with special reference to English and French'. In Guglielmo CINQUE and Richard S. KAYNE, eds., *The Oxford handbook of comparative syntax*, pp. 3–69. Oxford: Oxford University Press. [Reprinted in R.S. KAYNE, *Movement and silence*, pp. 277–333. Oxford: Oxford University Press.]

- KEY, Hewitt T. (1858). *Latin grammar*. London: Bell and Daldy.
- KIBRIK, Aleksandr E., Sandro V. KODZASOV, and Irina P. OLOVJANNIKOVA (1972). *Fragmenty grammatiki xinalugskogo jazyka*. Moskva: Izdatel'stvo Moskovskogo Universiteta.
- KIECKERS, Ernst (1960). *Handbuch der vergleichenden gotischen Grammatik*. 2nd edn. München: Max Hueber.
- KLIMOV, Georgij A. (1994). *Einführung in die kaukasische Sprachwissenschaft*. Hamburg: Helmut Buske. [Translated from the Russian original *Vvedenie v kavkazskoe jazykoznanie*, Moskva, Nauka, 1986.]
- KOIKE, Takeshi (2004). *The analysis of the genitive case in Old English within a Cognitive Grammar framework, based on the data from Ælfric's Catholic Homilies, First Series*. PhD thesis. University of Edinburgh.
- KORNFILT, Jaklin and Omer PREMINGER (2015). 'Nominative as *no case at all*: An argument from raising-to-ACC in Sakha'. In Andrew JOSEPH and Esra PREDOLAC, eds., *MIT Working Papers in Linguistics 76: Proceedings of the 9th Workshop on Altaic Formal Linguistics*, pp. 109–120. Cambridge, MA: MIT Press.
- KRISHNAMURTI, Bhadriraju and J.P.L. GWYNN (1985). *A modern grammar of Telugu*. Oxford: Oxford University Press.
- KÜHNER, Raphael and Friedrich HOLZWEISSIG (1912). *Ausführliche Grammatik der lateinischen Sprache I: Elementar-, Formen- und Wortlehre*. 2nd edn. Hannover: Hahnsche Buchhandlung. [1st edn., by KÜHNER alone, 1877.]
- LAKA, Itziar (2006). 'On the nature of case in Basque: Structural or inherent?' In Hans BROEKHUIS, Norbert CORVER, Jan KOSTER, Riny HUYBREGTS, and Ursula KLEINHENZ, eds., *Organizing grammar: Linguistic studies in honor of Henk van Riemsdijk*, pp. 374–382. Berlin: Mouton de Gruyter.
- LEGATE, Julie Anne (2008). 'Morphological and abstract case'. *Linguistic Inquiry* 39: 55–101.
- (2012). 'Types of ergativity'. *Lingua* 122: 181–191.
- LEUMANN, Manu (1977). *Lateinische Laut- und Formenlehre*. 2nd edn. München: Beck. [1st edn., 1926/28].
- LEVENTE, László (2002). 'The quantity of final -u in the nominative-accusative singular of Latin u-stem neuter nouns'. *Acta Antiqua Academiae Scientiarum Hungaricae* 42: 133–140.
- LEVIN, Beth and Malka RAPPAPORT HOVAV (1995). *Unaccusativity: At the syntax-lexical semantics interface*. Cambridge, MA: MIT Press.
- LILJEGREN, Henrik (2016). *A grammar of Palula*. Berlin: Language Science Press.
- MAHAJAN, Anoop (2012). 'Ergatives, antipassives and the overt light v in Hindi'. *Lingua* 122: 204–214.
- MARANTZ, Alec (1991). 'Case and licensing'. In German F. WESTPHAL, Benjamin AO, and Hee-Rahk CHAE, eds., *Proceedings of ESCOL 1991*, pp. 234–253. Ithaca, NY: Cornell Linguistics

- Circle. [Reprinted in Eric Reuland, ed., *Arguments and case: Explaining Burzio's generalization*, pp. 11–30. Amsterdam: John Benjamins, 2000.]
- MCCAWLEY, James D. (1968). 'Lexical insertion in a transformational grammar without Deep Structure'. In Bill J. DARDEN, Charles-James N. BAILEY, and Alice DAVISON, eds., *Papers from the Fourth Regional Meeting of the Chicago Linguistic Society*, pp. 71–80. Chicago: Chicago Linguistic Society. [Reprinted in J.D. MCCAWLEY, *Grammar and meaning: Papers on syntactic and semantic topics*, pp. 155–166. New York: Academic Press, 1976.]
- MCCREIGHT, Katherine and Catherine V. CHVANY (1991). 'Geometric representation of paradigms in a modular theory of grammar'. In PLANK 1991a, pp. 91–111.
- McFADDEN, Thomas (2004). *The position of morphological case in the derivation*. PhD thesis. University of Pennsylvania, Philadelphia.
- (2014). "Why nominative is special: Stem allomorphy and case structures". Presented at the 37th Colloquium of Generative Linguistics in the Old World, Brussels.
- MEL'ČUK, Igor' A. (1979). 'The predicative construction in Dyirbal'. In I.A. MEL'ČUK, ed., *Studies in dependency syntax*, pp. 23–90. Ann Arbor.
- (1986). 'Toward a definition of case'. In Richard D. BRECHT and James S. LEVINE, eds., *Case in Slavic*, pp. 35–85. Columbus, OH: Slavica.
- MERCHANT, Jason (2006). 'Polyvalent case, geometric hierarchies, and split ergativity'. In Jackie BUNTING, Sapna DESAI, Robert PEACHEY, Chris STRAUGHN, and Zuzana TOMKOVA, eds., *Proceedings of the 42nd Annual Meeting of the Chicago Linguistic Society*, pp. 57–76. Chicago: Chicago Linguistic Society.
- 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.
- MITHUN, Marianne (1991). 'Active/agentive case marking and its motivations'. *Language* 67: 510–546.
- MOHANAN, Tara (1994). *Argument structure in Hindi*. Palo Alto, CA: CSLI.
- MORPHY, Frances (1983). 'Djapu, a Yolngu dialect'. In R.M.W. DIXON and Barry J. BLAKE, eds., *Handbook of Australian languages*. Vol. 3, pp. 1–188. Canberra: Australian National University Press.
- MOSKAL, Beata (2015). 'Limits on allomorphy: A case study in nominal suppletion'. *Linguistic Inquiry* 46: 363–376.
- MOSKAL, Beata, Peter W. SMITH, Ting XU, Jungmin KANG, and Jonathan David BOBALJIK (2015). "A number of cases of pronominal suppletion". Presented at the 38th Colloquium of Generative Linguistics in the Old World, Paris.
- MÜLLER, Gereon (2004). 'On decomposing inflection class features: Syncretism in Russian noun inflection'. In Gereon MÜLLER, Lutz GUNKEL, and Gisela ZIFONUN, eds., *Explorations in nominal inflection*, pp. 189–227. Berlin: Mouton de Gruyter.

- NEELEMAN, Ad and Kriszta SZENDRŐI (2007). 'Radical pro drop and the morphology of pronouns'. *Linguistic Inquiry* 38: 671–714.
- NICHOLS, Johanna (1986). 'Head-marking and dependent-marking grammar'. *Language* 62: 56–119.
- (2011). *Ingush grammar*. Berkeley: University of California Press.
- NIE, Yining (2016). 'Why is there NOM-NOM but no ERG-ERG?' *Toronto Working Papers in Linguistics* 36.
- NIEDERMANN, Max (1906). *Précis de phonétique historique du latin*. Paris: Klincksieck.
- NIKOLAYEV, Sergej L. and Sergej A. STAROSTIN (1994). *A North Caucasian Etymological Dictionary*. Moscow: Asterisk.
- O'CONNOR, Mary C. (1987). *Topics in Northern Pomo grammar*. PhD thesis. University of California, Berkeley.
- PALMER, Frank R. (1958). 'The noun in Bilin'. *Bulletin of the School of Oriental and African Studies, University of London* 21: 376–391.
- PENSALFINI, Rob (2003). *A grammar of Jingulu, an Aboriginal language of the Northern Territory*. Canberra: Pacific Linguistics.
- PESETSKY, David (2013). *Russian case morphology and the syntactic categories*. Cambridge, MA: MIT Press.
- PLANK, Frans (1985). 'The extended accusative/restricted nominative in perspective'. In Frans PLANK, ed., *Relational typology*, pp. 269–310. Berlin: Mouton de Gruyter.
- ed. (1991a). *Paradigms: The economy of inflection*. Berlin: Mouton de Gruyter.
- (1991b). 'Rasmus Rask's dilemma'. In PLANK 1991a, pp. 161–196.
- PRAULIŅŠ, Dace (2012). *Latvian: An essential grammar*. London: Routledge.
- PREMINGER, Omer (2012). 'The absence of an implicit object in unergatives: New and old evidence from Basque'. *Lingua* 122: 278–288.
- (2014). *Agreement and its failures*. Cambridge, MA: MIT Press.
- RADKEVICH, Nina V. (2010). *On location: The structure of case and adpositions*. PhD thesis. Storrs: University of Connecticut.
- ŘEZÁČ, Milan (2011). *Phi-features and the modular architecture of language*. Dordrecht: Springer.
- ŘEZÁČ, Milan, Pablo ALBIZU, and Ricardo ETXEPARE (2014). 'The structural ergative of Basque and the theory of Case'. *Natural Language & Linguistic Theory* 32: 1273–1330.
- RICHARDS, Norvin (2010). *Uttering trees*. Cambridge, MA: MIT Press.
- (2013). 'Lardil "case stacking" and the timing of case assignment'. *Syntax* 16: 42–76.
- RILL, Justin (2016). 'The typology of ERG=GEN'. Poster presented at the 90th Annual Meeting of the Linguistic Society of America, Washington DC.
- RIZZI, Luigi (1990). *Relativized Minimality*. Cambridge, MA: MIT Press.
- (2001). 'Relativized Minimality effects'. In Mark BALTIN and Chris COLLINS, eds., *The handbook of contemporary syntactic theory*, pp. 89–110. Oxford: Blackwell.

- RIZZI, Luigi (2004). 'Locality and left periphery'. In Adriana BELLETTI, ed., *Structures and beyond* (The Cartography of Syntactic Structures 3), pp. 223–251. Oxford: Oxford University Press.
- RIZZI, Luigi and Guglielmo CINQUE (2016). 'Functional categories and syntactic theory'. *Annual Review of Linguistics* 2: 139–163.
- ROVAI, Francesco (2014). 'Case marking in absolute constructions: further evidence for a semantically based alignment in Late Latin'. *Journal of Latin Linguistics* 13: 115–143.
- RUDE, Noel (1986). 'Topicality, transitivity, and the direct object in Nez Perce'. *International Journal of American Linguistics* 52: 124–153.
- RUMSEY, Alan (1987). 'The chimera of Proto-Indo-European ergativity: Lessons for historical syntax'. *Lingua* 71: 297–318.
- SAN MARTIN, Itziar and Juan URIAGEREKA (2000). 'Infinitival complementation in Basque'. In Xabier ARTIAGOITIA, Patxi GOENAGA, and Joseba A. LAKARRA, eds., *Erramu Boneta: A Festschrift for Rudolf de Rijk*, pp. 597–610. Vitoria: Universidad del País Vasco.
- SCHMIDT, Gernot (1978). *Stammbildung und Flexion der indogermanischen Personalpronomina*. Wiesbaden: Harrassowitz.
- SCHMITT, Rüdiger (2007). *Grammatik des Klassisch-Armenischen mit sprachvergleichenden Erläuterungen*. 2nd edn. Innsbruck: Institut für Sprachen und Literaturen der Universität Innsbruck. [1st edn., 1981.]
- SCHULZE, Wolfgang (2008). 'Towards a history of Khinalug'. In Brigitte HUBER, Marianne VOLKART, and Paul WIDMER, eds., *Chomolangma, Demavend und Kasbek. Festschrift für Roland Bielmeier zu seinem 65. Geburtstag*. Vol. 2, pp. 703–744. Halle (Saale): International Institute for Tibetan and Buddhist Studies.
- SCHÜTZE, Carson T. (2001). 'On the nature of default case'. *Syntax* 4: 205–238.
- SEILER, Walter (1985). *Imonda, a Papuan Language*. Canberra: Australian National University.
- SHLONSKY, Ur (2010). 'The cartographic enterprise in syntax'. *Language and Linguistics Compass* 4: 417–429.
- SIEWIERSKA, Anna (2005). 'Alignment of person agreement marking'. In Martin HASPELMATH, Matthew S. DRYER, David GIL, and Bernard COMRIE, eds., *World Atlas of Language Structure*, pp. 406–409. Oxford: Oxford University Press.
- SIHLER, Andrew L. (1995). *A new comparative grammar of Greek and Latin*. Oxford: Oxford University Press.
- SILVERSTEIN, Michael (1976). 'Hierarchy of features and ergativity'. In R.M.W. DIXON, ed., *Grammatical categories in Australian languages*, pp. 112–171. Canberra: Australian Institute of Aboriginal Studies. [Reprinted in Pieter MUYSKEN and Henk VAN RIEMSDIJK, ed., *Features and projections*, pp. 163–232. Dordrecht: Foris, 1986.]
- SMITH, Peter W., Beata MOSKAL, Ting XU, Jungmin KANG, and Jonathan David BOBALJIK (2016). 'Case and number suppletion in pronouns'. Available at: ling.auf.net/lingbuzz/003110.

- SOMMER, Ferdinand (1902). *Handbuch der lateinischen Laut- und Formenlehre*. Heidelberg: Carl Winter's Universitätsbuchhandlung.
- STARKE, Michal (2009). 'Nanosyntax: A short primer to a new approach to language'. *Nordlyd* 36: 1–6.
- (2014). 'Cleaning up the lexicon'. *Linguistic Analysis* 39: 245–256.
- STOLZ, Friedrich and Joseph H. SCHMALZ (1900). *Lateinische Grammatik*. 3rd edn. München: Beck'sche Verlagbuchhandlung.
- SUÁREZ MARTÍNEZ, Pedro Manuel (1996). 'La quantité de -u chez les neutres de la quatrième déclinaison latine'. In Rosén HANNAH, ed., *Aspects of Latin: Papers from the Seventh International Colloquium of Latin Linguistics, Jerusalem, April 1993*, pp. 91–98. Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft.
- SZEMERÉNYI, Oswald J.L. (1996). *Introduction to Indo-European linguistics*. Oxford: Clarendon Press. [Translated from the 4th German edn. (1990). 1st edn., *Einführung in die vergleichende Sprachwissenschaft*, Wissenschaftliche Buchgesellschaft, Darmstadt, 1970.]
- TUITE, Kevin (1998). 'A short descriptive grammar of the Svan language'. Available at: [http://www.uni-jena.de/unijenamedia/downloads/faculties/phil/kaukasiologie/svan\[slightlyrevised\].pdf](http://www.uni-jena.de/unijenamedia/downloads/faculties/phil/kaukasiologie/svan[slightlyrevised].pdf).
- WALI, Kashi and Omkar N. KOUL (1996). *Kashmiri: A cognitive-descriptive grammar*. New York: Routledge.
- WEERMAN, Fred and Jacqueline EVERS-VERMEUL (2002). 'Pronouns and case'. *Lingua* 112: 301–338.
- WEISS, Michael (2009). *Outline of the historical and comparative grammar of Latin*. Ann Arbor: Beech Stave Press.
- WILLSON, Stephen R. (1996). 'Verb agreement and case marking in Burushaski'. *Work Papers of the Summer Institute of Linguistics* 40: 1–71.
- WOOLFORD, Ellen (1997). 'Four-way case systems: Ergative, nominative, objective and accusative'. *Natural Language & Linguistic Theory* 15: 181–227.
- (2006). 'Lexical case, inherent case, and argument structure'. *Linguistic Inquiry* 37: 111–130.
- YIP, Moira, Joan MALING, and Ray JACKENDOFF (1987). 'Case in tiers'. *Language* 63: 217–250.
- YOSHIOKA, Noboru (2012). *A reference grammar of Eastern Burushaski*. PhD thesis. Tokyo University of Foreign Studies.
- ZWICKY, Arnold M. (1991). 'Systematic versus accidental phonological identity'. In PLANK 1991a, pp. 113–131.
- PRÁINSSON, Höskuldur (2007). *The syntax of Icelandic*. Cambridge: Cambridge University Press.