

Theoretical approaches to syncretism in marked-nominative languages

Abschlussarbeit zur Erlangung
des akademischen Grades Bachelor of Arts (B.A.)

eingereicht am 12.08.2021

Sinoël Gordon Isidor Dohlen

Matrikelnummer: 3716009

sd18vidi@studserv.uni-leipzig.de



Universität Leipzig

Institut für Linguistik

Betreuerin: Ph. D. Maria Kouneli

Zweitbetreuer: Prof. Ph. D. Gregory Kobele

Contents

1	Theory and predictions	1
1.1	Case sequences	1
1.1.1	Syncretism	1
1.1.2	Containment	3
1.1.3	The hierarchy	6
1.2	Adjustments	7
1.3	Marked-nominative alignment	9
1.3.1	Possibilities & Predictions	10
2	Application	11
2.1	Datooga	12
2.2	Possessors intervene	14
2.2.1	Syncretism	15
2.2.2	Containment	15
2.2.3	Solution	16
2.2.4	West Nordic	17
2.3	Tennet	18
2.4	Haro	19
3	Conclusion	21

*I dedicate this thesis to the conlang community,
which sparked my interest in morphosyntactic alignment five years ago*

Overview

In this thesis, I review current theories of case syncretism. I then go on analyzing paradigms of three marked-nominative languages which exhibit different patterns of syncretism.

In order to account for these patterns, I decouple the case morphs from the abstract, grammatical roles they express.

This leads me to a slightly adjusted version of Zompì's (2017) hierarchy, which is able to derive all new and old data, shown in (0).

- (0) a. $S \leq \{A, P, PR\} \leq \{R, TL, PC, SC\}$
b. $[[[S] A/P/PR] R/TL/PC/SC]$

I then go on to show how this adjusted hierarchy can even account for previously underivable patterns from ergative and accusative languages.

An index of the abbreviations used in this thesis is given in the footnote on the bottom of this page.

⁰A - agent, ABL - ablative, ABS - absolutive, ACC - accusative, APPL - applicative, DAT - dative, DEF - definite, ERG - ergative, GEN - genitive, INDEF - indefinite, INST - instrumental, LOC - locative, NOM - nominative, OBL - oblique, P - patient, PC - place, PL - plural, PR - possessor, PROPER - proper name, R - recipient, S - subject, SC - source, SG - singular, SNOM - S-nominative, TL - tool, VOC - vocative, 3 - 3rd person

1 Theory and predictions

In this chapter I'll give an overview of the theoretical models used (and not used) in this thesis. I then go on explaining what marked-nominative languages are and why their examination in light of the mentioned models should interest us.

1.1 Case sequences

1.1.1 Syncretism

In recent years, the theoretical literature on case syncretism has shifted more towards using theories which model case syncretisms as sequences of cases neighbouring each other (Johnston 1996, Caha 2009, 2010 and Zompì 2017), opposed to using unordered, natural classes, the approach which previously dominated the literature going as far back as Bierwisch (1967).

Johnston proposes that for every language, a single sequence of case(s?) can be found, which is capable of deriving/describing all possible syncretisms of the given language. The basic concept is that only cases which are adjacent on the sequence can undergo syncretism with each other. I'll go over his analysis of Modern Greek in Johnston (1996:§2.2.5) in order to give an example of how it works.

	'nation' [SG]	'nation' [PL]	'friend' [SG]	'friend' [PL]	'barber' [SG]
N(OM)	éθn-os	éθn-i	fíl-os	fíl-i	kuré-as
V(OC)	éθn-os	éθn-i	fíl-e	fíl-i	kuré-a
A(CC)	éθn-os	éθn-i	fíl-o	fíl-us	kuré-a
G(EN)	éθn-us	eθn-ón	fíl-u	fíl-on	kuré-a

Table 1: Modern Greek as cited by Johnston (1996)

In table 1 we can see that the paradigm of 'nation' syncretizes N, V and A in both singular and plural, whereas the singular of 'barber' syncretizes V, A and G. The only nominal with a distinct form for each case is the singular

	‘spoon’	‘chest’
	[SG,INDEF]	[SG]
DAT	ein-em Löffel	gol-u-z
INST	mit ein-em Löffel	gol-u-z-un

Table 2: German and Budukh

of ‘friend’. Any order would be capable of deriving this one(single/specific?) paradigm. Its plural shows syncretism between N and V.

The sequence N–V–A–G can derive all of these patterns via the above mentioned method, and it is the only one which can do so besides its inverse: G–A–V–N. This bi- or unidirectionality is not a part of Caha and Zompì’s models (see 1.1.2).

I won’t go into much detail about Caha (2009,2010) since Zompì (2017:§2.1) scales back a lot of structure of his proposed hierarchy, and some of the phenomena Caha investigates are not that relevant to case syncretism as it occurs in the languages of this thesis.

For example, he treats adpositions in the same way as he does bound marking. In (Caha 2010:3–13) he points out that in German, which is usually not considered to have an instrumental¹, NPs containing *mit* ‘with’ nonetheless display a pattern almost identical to nouns marked by the instrumental in Budukh (see table 2).

The NPs *mit einem Löffel* ‘with a spoon’ and *goluzun* ‘with the chest’ both contain the respective dative forms of the nouns, the only difference being that in Budukh both are expressed by suffixes, while in German only dative is expressed by a suffix and instrumental by a preposition. For Caha this is only a matter of movement taking place in German, but not Budukh.

Most of the languages of the marked-nominative alignment lack detailed descriptions, so I’ll limit myself to analyzing bound marking².

The most relevant takeaway from Caha for this thesis is the notion of Universal Case Contiguity, which Caha describes as follows:

¹This is because there is no dedicated morpheme occurring in the NP when a given noun is serving as an instrument, like there is for accusative: *ein-en Löffel* ‘INDEF-ACC spoon’ or dative: *ein-em Löffel* ‘INDEF-DAT spoon’.

²As does Zompì for lexical nouns.

- (1) Universal (Case) Contiguity: (Caha 2009:10, emphasis mine)
- a. Non-accidental case syncretism targets contiguous regions in a sequence **invariant** across languages
 - b. The Case sequence: nominative – accusative – genitive – dative – instrumental – comitative

Unlike Johnston (1996:200–222) who rejected the possibility of a universally applicable version of the sequence, Caha argues for just that. However, like Johnston before him, he doesn't provide any analysis of ergative or tripartite languages.

I will therefore focus on Zompì's (2017) approach in this thesis, since he includes ergative and tripartite languages in his sample, thus deriving the broadest coverage of alignment patterns so far, namely prototypical nominative-accusative, ergative-absolutive and tripartite languages.

In order to do so Zompì makes some adjustments to the sequence. Instead of each slot in the sequence only being able to host exactly one case, they host natural classes made up of cases based on Marantz's (1991) dependent case theory (Zompì 2017:2).

The 'unmarked' class includes absolutive, nominative and the case employed for the S argument in tripartite languages³, 'dependent' includes ergative and accusative while 'inherent' includes all other cases besides the genitive, which is the only case Zompì isn't able to place into one of the three classes (Zompì 2017:§5.1).

1.1.2 Containment

Unlike Johnston's sequences whose directionality can be reversed without any impact on the predictions made, Zompì's model arranges the classes into an ordered hierarchy.

The argument for this hierarchical ordering is based on overt containment between dependent and oblique cases in some languages. This means that they

³Zompì calls this the S-Nominative, I will simply refer to it as the case of the S argument to avoid naming yet another pattern after the nominative.

display an ABB pattern where each letter stands in for a case marker. If two morphs are described as A and B, they are distinct, unrelated forms, while B and B share the same ‘core’.

	‘boy’ [SG]	‘boy’ [PL]
NOM	čhav-ó	čhav-é
ACC	čhav-és	čhav-én
GEN	čhav-és-koro	čhav-én-goro
DAT	čhav-és-ke	čhav-én-ge
INST	čhav-és-ar	čhav-én-car
LOC	čhav-és-te	čhav-én-de
ABL	čhav-és-tar	čhav-én-dar

Table 3: Vlakh Romani

The notion of the ban on ABA patterns (better known simply as *ABA), introduced by Bobaljik (2012) to explain the absence of certain patterns of adjectival degree morphology⁴, has since been used to explain gaps in other domains of morphology.

	‘horse’ [PL]
NOM	yuka-ñ
ACC	yuka-s
GEN	yukaśśi ⁵
INST	yuka-s-yo
LOC	yuka-s-aṃ
ABL	yuka-s-äṣ

Table 4: Tocharian A

	‘man’ [PL]
NOM	eñkw-i
ACC	eñkw-eṃ
GEN	eñkw-eṃ-ts
LOC	eñkw-eṃ-ne
ABL	eñkw-eṃ-mem

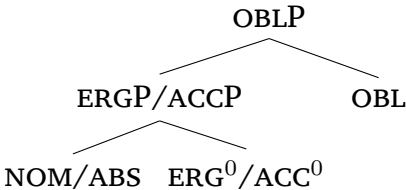
Table 5: Tocharian B

Applied to the accusative languages Tocharian A, Tocharian B and Vlakh Romani, previously discussed by Caha (2010) and repeated here in table 3–5,

⁴In a sample of over 300 languages, Bobaljik describes how adjectives systematically avoid employing a shared stem for positive and superlative, while using a different one for the comparative, like f.e. good, better, good-est (ABA). The patterns AAA, ABB and rarely ABC are attested.

we can spot four paradigms with ABB patterns. For example, the *ó* in *čhav-ó* ‘boy-NOM’ is a distinct morph and thus labeled ‘A’. Another distinct morph is *és* in *čhav-és* ‘boy-ACC’, labeled ‘B’. The morph *ésar* in *čhav-ésar* ‘boy-INST’ however overtly contains *és* and is thus also labeled ‘B’.

The oblique cases all contain the respective accusative in all four paradigms. No language is known in which this order is reversed, i.e. where the accusative would overtly contain one of the oblique cases. The fact that these overt containment relations only ever occur in one direction is used to corroborate the hierarchical ordering of inherent above dependent cases (Zompì 2017:§3.3). This leads us to the tree structure in (2a), with the corresponding bracket notation in (2b).

- (2) a.  (Zompì 2017:41)
- b. [NOM/ABS [ERG/ACC [OBL]]]

Zompì (2017:§1.4) shows that case containment can be derived within both Nanosyntax and Distributed Morphology, while also ruling out ABA patterns without explicitly banning them, just by using cumulative decomposition (see 3) together with the underlying structure in (2).

- (3) Cumulative decomposition of cases (Zompì 2017:13)
- a. NOM = [A]
 - b. ACC = [A,B]
 - c. GEN = [A,B,C]

He also claims that ergative and accusative are built upon absolutive and nominative respectively, but without the same sound morphological evidence.

The nominative forms bear no containment relation to any of the other forms in three out of the four paradigms. Only in the plural of Vlakh Romanian *čhav-é* ‘boy’, do the other forms contain the nominative. I’ll tackle this assumption in the following section.

But not only nominative-accusative languages display overt containment. In Agul (table 6) for example, the ergative serves as the base for genitive and dative as is expected by the grouping of ergative together with accusative as dependent cases.

	‘father’ [SG]	‘belt’ [SG]	‘sky’ [SG]
ABS	gaga	čal	zaw
ERG	gaga-di	čal-i	zaw-u
GEN	gaga-di-n	čal-i-n	zaw-u-n
DAT	gaga-di-s	čal-i-s	zaw-u-s

Table 6: Agul

This leads Zompì to the following generalization:

- (4) Containment generalization (Zompì 2017:65)
- a. Each inherent case contains a clause-level dependent case.
 - b. Each clause-level dependent case contains a clause-level unmarked case.

1.1.3 The hierarchy

Zompì rejects stipulating separate hierarchies for different alignments since according to him they are insufficient in explaining the side-by-side occurrence of ergative and accusative case in tripartite languages. Thus he proposes the hierarchy in (5), intended to hold for all languages regardless of alignment (Zompì 2017:35).

⁵The genitive is underlyingly likely to be *yuka-s-ši*. Gippert (1987:27) mentions assimilation of *s* to *š* in a similar context elsewhere in the language.

- (5) Zompì's universal hierarchy
{NOM,ABS,SNOM} – {ACC,ERG} – {DAT,LOC,INSTR}

As for alignment bound hierarchies, he however only considers hierarchies for ergative and accusative (6a–b), not tripartite languages (6c). This third hierarchy would solve the emergence of both ergative and accusative in tripartite languages while also obeying the dependent case ordering.

- (6) Alignment specific hierarchies
- a. {NOM} – {ACC} – {DAT, LOC, INSTR...}
 - b. {ABS} – {ERG} – {DAT, LOC, INSTR...}
 - c. {S} – {ACC, ERG} – {DAT, LOC, INSTR...}

Hierarchy (6c) is actually very elegant as a whole since, as Zompì (2017:36) notes himself, it is capable of deriving prototypical ergative and accusative as well as tripartite paradigms:

”In all of these [tripartite] 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 [...] or by SNOM-ERG syncretism [...]”

I discuss this and other improvements in the upcoming section.

1.2 Adjustments

Due to the nature of the categories nominative and absolutive, his hierarchy groups ‘overlapping’ cases into the classes. This leads to superfluous adjacencies in the hierarchy which permit syncretism and containment between logically impossible co-occurrences of cases such as nominative and ergative, absolutive and accusative as well as the ‘unmarked’ cases with each other. It also causes grammatical roles to occur in multiple positions in the hierarchy:

S is part of NOM, ABS and SNOM, A is part of NOM and ERG, P part of ABS and ACC. I highlighted these superfluous adjacencies in (7).

(7) Superfluous adjacencies in Zompì's hierarchy

$$\{\text{NOM, ABS, SNOM}\} - \{\text{ACC, ERG}\} - \{\text{DAT, LOC, INSTR}\}$$

By sticking to the restriction of only allowing a single, universal sequence (1a), as well as treating nominative and absolutive as undecomposable categories, he robs his approach of much elegance.

These superfluous adjacencies can however be circumvented by simply using (6c) for all alignments as alluded to in §1.1.3. Nominative and absolutive can be derived via contiguity of S & A and S & P respectively, thus emerging like any other form expressing multiple grammatical roles.

I will from hereon use the hierarchy in (8), reflecting (6c) with slightly adjusted labels. These labels represent grammatical roles (or arguments), which represent semantic and morphosyntactic properties of a node. Case morphs will be considered to be the realizations of these underlying roles. In table 7 you can see what a case will usually be called when covering these roles. The position of these arguments inside the hierarchy is fixed, like Caha's and Zompì's hierarchies are.

(8) Adjusted hierarchy $S \leq \{A, P\} \leq \{R, TL, PC, SC\}$

Since this disentangling is needed and compatible with Zompì's method, I will use it from hereon.

I focus on the widely used roles of S, A, P and R, as well as PR (possessor), TL (tool), PC (place) and SC (source) which I coined myself. The latter were the roles most frequently coded as bound morphs or by morphophonological rules directly on the noun among the languages examined in this thesis.

Some of the paradigms have more oblique cases than I will present here, but I will only leave out disjunct oblique cases since they aren't impacted by models of Case Contiguity anyway. The roles will be separated along my three

S	–
A	ergative
P	accusative
S & A	nominative
S & P	absolutive
R	dative
PR	genitive
TL	instrumental
PC	locative
SC	ablative

Table 7: Grammatical roles and their corresponding case labels

proposed classes by dashed lines to aid identifying where contiguity is upheld and where it isn't.

To avoid misunderstandings with morphological markedness, I will call Zompì's 'marked dependent cases' (A and P) 'intervening role/arguments' instead.

1.3 Marked-nominative alignment

In her book about the morphosyntactic typology of marked-S languages, Handschuh (2014) shows that marked-nominative languages are not simply a certain kind of nominative-accusative languages. For example, unlike Lander (2009:590) predicts, only one out of the 22 marked-nominative languages marks the possessor like the marked transitive argument, resulting in syncretism between S, A, and PR. The other 21 either use a dedicated genitive morpheme or leave possessors bare.

Another difference exists with citation forms, which are usually identical to the nominative in prototypical nominative-accusative languages even if the nominative is marked. The marked-nominative languages of her sample use the form of the P argument, as citation form without an exception. In most languages this form is equal to the bare form, but some languages overtly mark P as well.

The former are called marked-nominative languages of type 1 by König

(2006) and the latter type 2. Handschuh (2014:18) calls them formally and functionally marked respectively.

But there are functions for which marked-S languages differ from each other in the encoding of arguments. For example, emphatic S arguments are split almost evenly between being marked by the S-case or being unmarked.

Differences also exist between the number of contexts the S-case is used in (Handschuh 2014:§8.4). The Omotic languages use the S-case most frequently (66% of the studied functions), the other contexts being covered by the bare form (20%) not much more frequently than by other forms (14%); while in Oceanic the bare form appears much more frequently than the S-case, which in turn appears much more frequently than other forms (58% bare, 38% S, 4% other forms).

Thus, it is unwarranted to simply assume marked-S languages to pattern like their unmarked-S counterparts. I will therefore ‘a priori’ treat them as their own distinct alignment type. I go into detail about this in the following section.

1.3.1 Possibilities & Predictions

The languages covered by (Johnston 1996, Caha 2009, 2010 and Zompì 2017) have one thing in common: If they have an unmarked case in their paradigm, it’s the one expressing the S argument. This is by definition never the case for marked-S languages.

This also means that if there is a form in the paradigm which is overtly contained inside all other forms, it will include the S argument, since only a bare form can be overtly contained inside all other forms. In type 1 marked-nominative languages, P is always bare, which poses the question whether problems come about when an intervening role is systematically less marked than S.

Prototypical marked-S languages share the coding asymmetry of $S = A \neq P$ with prototypical accusative languages, whereas the morphological marking of A and P is like that of prototypical ergative languages. This has lead people to describe this alignment as both ‘marked nominative’ as well as ‘extended

ergative' Dixon (1994:63–67).

Investigating the syncretism patterns of marked-nominative languages might also give us more insight into their underlying syntax. If they were structurally just like prototypical accusative languages we would expect their syncretisms to follow the pattern in (9a). If they instead patterned like (9b), they would share their intervening role with prototypical ergative languages.

(9) Theoretically possible syncretisms

- a. $S = A \neq P = OBL$
- b. $S = A = OBL \neq P$
- c. $S = A \neq P \ OBL = ?$
- d. $S = A \neq P \neq OBL$

The last possibility is that none of the oblique cases in marked-nominative languages shares its form with A or P. This is possible in two ways: Either we could notice that neither A nor P ever intervene and some form other than A or P (f.e. construct state) forms the base of oblique cases (9c) or all oblique roles are expressed distinctly from any other form (9d), giving us no insight into the structure whatsoever. We might of course also find different patterns for different languages or even for different paradigms of the same language.

I will now go on exploring whether patterns of syncretism in marked nominative languages are better explained adopting an accusative (P intervenes) or ergative model (A intervenes) and also whether they are actually better thought of as the dominant 'marked nominative', the more niche 'extended ergative' or if both are adequate terms, depending on the individual languages. I will write out the S, A and P forms in the paradigms individually to illustrate the patterns the underlying grammatical roles undergo.

2 Application

Fortunately, pattern (9d) is not the only attested one.

First, I will investigate Datooga, where P serves as the intervening role, thus behaving like ‘accusative-like’ pattern. I also discuss the implications of working with \emptyset -morphs in that chapter.

I then take short detour to talk about possessors in §2.2, before continuing with Tennet in 2.3 and ending with Haro in 2.4. Both languages display syncretism of S, A and PR, but only for certain types of salient nominals. Haro also exhibits overt containment of this S = A = PR form.

2.1 Datooga

The first marked-nominative language we’ll look at is Datooga, specifically the Gisamjanga variety spoken in Northern Tanzania. Datooga belongs, along with the Omotik language, to one of the the two branches of Southern Nilotic, the other branch being comprised of the Kalenjin languages. I draw all the data about Datooga from Kießling (2007).

Datooga, like most Nilotic languages is verb-initial in pragmatically unmarked clauses. But when the subject is topicalized, it moves to the first position in the clause. Subjects in this position ‘lose’ their nominative case, a phenomenon informally known as ‘no case before the verb’ shared by many Nilotic languages of the area. Similar to many East African languages, case is marked purely by means of tone. Kießling gives a detailed description of the contexts in which the two forms in table 8 occur.

	‘cat’ [SG]	‘Datooga person’ [SG]
S	ɲáawúu-dá ⁶	dátòopáan-dá
A	ɲáawúu-dá	dátòopáan-dá
P	ɲáawùu-dà	dàtóopáan-dà
PR	ɲáawùu-dà	dàtóopáan-dà
R	ɲáawùu-dà	dàtóopáan-dà
TL	ɲáawùu-dà	dàtóopáan-dà

Table 8: Datooga (Kießling 2007:152–154)

⁶This suffix expresses singular number.

When serving as S or A arguments, nominals undergo a process which overwrites all lexical tones to the melody H(L₀)HH. This means that the first syllable and the last two syllables are assigned high tones, all other syllables receive a low tone. The lexical tone of a syllable never affects this process.

Since Datooga *náawùudà* ‘cat’ consists of only three syllables, they all surface with a high tone while for quadrisyllabic *dátòonáandá* ‘Datooga person’ the second syllable is low, first, third and fourth being high. This rule applies regardless of the semantics of the verb.

The bare form serves a much wider function. These include the the form of P, PR, R, TL, preverbal S and preverbal A arguments, prepositional complements, the vocative and the citation form (Kießling 2007:167–169).

The tone melody is never predictable in these contexts. I therefore assume that there is no rule involved which changes the form of the noun. It’s possible that a rule applies whose output is always the same as its input, essentially equivalent to the affixation of Ø-morphs. I will still refer to these forms as ‘bare’, since the data is identical in both analyses, only the theoretical objects differ.

We actually run into a problem if we don’t posit a Ø-morph rule for this paradigm. In (10a) the bare form lacks any morphological modification and is thus overtly contained by the form which underwent the tone melody process. This relationship leads to a pattern where S and A contain all other roles (10b). This order cannot be derived by Caha’s, Zompì’s or my own model. And indeed, when considering paradigms from other languages, it is evident that this is a pattern ideally ruled out to avoid overgeneration.

- (10) a. *náawùudà* < *náawùudà*-H(L₀)HH
 b. {P, PR, R, TL} < {S, A}

In (11a) both forms underwent a morphological rule. They are thus as disjoint as any rule introducing segmental morphs. Note that it would theoretically still be possible to first derive the Ø-morph form and the melody form afterwards suggesting a hierarchy along the lines of [[P/PR/R/TL] S/A].

- (11) a. $\text{ɲáawùudà-H(L}_0\text{)HH} \neq \text{ɲáawùudà-}\emptyset$
 b. $\{S, A\} \neq \{P, PR, R, TL\}$

What's crucial is that this is not the only possible order of operations, unlike for the analysis in (10). Since there is no evidence for one over the other analysis from the paradigm itself, I will go with to the \emptyset -morph analysis since it is consistent with data from other languages. We thus see that Datooga displays pattern (9a) like prototypical accusative languages.

Note however that Datooga has alternative strategies for expressing some oblique arguments. The TL argument can be expressed in two ways in Datooga. Either just in the bare form together with a derived applicative verb (12a); or also in the bare form, but preceded by *àbà* 'PREP' where the verb is underived (12b).

- (12) a. *gá-ftà* *bálláandá* *gàacéedà* *dáráwéetà*
 S3-shoot.APPL boy.NOM arrow swala.antelope
 'The boy shot the swala antelope with an arrow.'
- b. *gá-fù* *bálláandá* *dáráwéetà* *àbà* *gàacéedà*
 S3-shoot boy.NOM swala.antelope with arrow
 'The boy shot the swala antelope with an arrow.'

Since, according to Kießling, *gàacéedà* in (12a) is part of the core participants, whereas it functions as an adjunct in (12b), I consider its inclusion over *àbà* *gàacéedà* well motivated.

2.2 Possessors intervene

Before continuing with Tennet and Haro, I need to discuss the problems with syncretism and containment involving possessors first, since the PR argument plays an important part for syncretism and overt containment in these languages.

2.2.1 Syncretism

Zompì (2017:§5.1.1) gives multiple examples of systematic syncretism of PR together with S & A. This poses a major problem for his hierarchy since PR also syncretizes with R and P in other languages (see §2.2.4).

In Latin, for example, some paradigms display syncretism of the form S = A = PR. This can be seen in the paradigms of *cīvis* ‘citizen’ and *canis* ‘dog’.

	‘town’ [SG]	‘citizen’ [SG]	‘dog’ [SG]
S	urbs	cīvis	canis
A	urbs	cīvis	canis
PR	urbis	cīvis	canis
P	urbem	cīvem	canem
R	urbī	cīvī	canī
SC	urbe	cīvī/cīve	cane

Table 9: Latin

It’s not a phonological rule that introduces the *i* as can be seen with *urbs* ‘town’ where S & A are distinct from PR. Even more remarkably, *canis* must have become part of this pattern by analogy since in Classical Greek and Sanskrit the forms for ‘dog’ are distinct in these contexts (Zompì 2017:84).

2.2.2 Containment

Another counterexample Zompì (2017:86–87) gives comes from the Nakh-Daghestanian language Kryz in table 10. Kryz displays a pattern very similar to that of Agul mentioned earlier in table 6.

But instead of A serving as base for PR and R as in Agul (13a), it is PR that serves as a base for A, R and TL (13b). The latter pattern can also be seen in Budukh in table 11. All three belong to the Lezgian branch of Daghestanian.

- (13) a. A < {PR, R}
 b. PR < {A, R, TL}

	‘married man’ [SG]	‘house’ [SG]		‘ground’ [SG]	‘ground’ [PL]
S	furi	k’ul	S	нокъ	нокъ-г-и
P	furi	k’ul	P	нокъ	нокъ-г-и
PR	fura	k’ul-ci	PR	нокъ-и	нокъ-г-а
A	fura-г	k’ul-ci-г	A	нокъ-и-г	нокъ-г-а-га
R	fura-з	k’ul-ci-з	R	нокъ-и-з	нокъ-г-а-з
TL	fura-з-ина	k’ul-ci-з-ина	TL	нокъ-и-з-ун	нокъ-г-а-з-эн

Table 10: Kryz

Table 11: Budukh

These patterns seem to directly contradict each other. Not being able to offer a solution to these problems, Zompì decides to omit PR from his hierarchy.

2.2.3 Solution

This dilemma can be solved by placing PR alongside A and P in the class of intervening cases, as in (14): This position allows PR to syncretise with A as well as R through ‘sideways adjacency’ (i.e. syncretism between roles of the same class). Zompì utilizes this adjacency already for the derivation of obliques.

- (14) a. $S \leq \{A, P, PR\} \leq \{R, TL, PC, SC\}$
 b. $[[[S] A/P/PR] R/TL/PC/SC]$

Even stronger support for this adjustment comes from the Nakh language Ingush in table 12 where PR also serves as intervening role. Unlike Kryz and Budukh however, Ingush PR is realized distinctly from A and P.

The pattern, formalized in (15), is compatible with (14).

⁷Caha (2010:10) citing Blevins (2009:203).

‘hen’	
[SG]	
S	kuotam
P	kuotam
A	kuotam-uo
PR	kuotam-a
R	kuotam-a-a
TL	kuotam-a-ca

Table 12: Ingush⁷

(15) {S, P} < A ≠ PR < {R, TL}

2.2.4 West Nordic

Another problematic pattern comes from West Nordic languages as pointed out by Harðarson (2016). It constitutes a systematic counterexample to Caha’s proposed Case Contiguity, since P and R syncretize to the exclusion of PR as can be seen in the Icelandic paradigms in table 13.

	‘arm’	‘land’	‘queen’	‘tongue’
	[SG]	[SG]	[SG]	[SG]
S	arm-ur	land-∅	drottning-∅	tung-a
A	arm-ur	land-∅	drottning-∅	tung-a
PR	arm-s	land-s	drottning-ar	tung-u
P	arm-∅	land-∅	drottning-u	tung-u
R	arm-i	land-i	drottning-u	tung-u

Table 13: Modern Icelandic (Harðarson 2016:1332)

I formalized these patterns in (16). All syncretic forms are adjacent in the hierarchy in (14).

- (16) a. {S, A} ≠ PR ≠ P ≠ R
 b. {S, A, P} < PR ≠ R
 c. {S, A} < PR ≠ {P, R}
 d. {S, A} ≠ {PR, P, R}

2.3 Tennet

Tennet is a Southwest Surmic language spoken in South Sudan. Like most marked-nominative languages, it uses the bare form for P arguments, as form of citation and for copula complements. Possessors on the other hand are encoded slightly differently than in the other languages, where genitive marking or lack of marking are most common.

There is a semantically based split alignment in which human proper names take the suffix -i when acting as possessor, while common nouns take -o instead. The suffix -o is a distinct genitive morph unrelated to any other forms, but -i is also the general nominative marker of the language. The data come from Randal’s (1998) description of the language.

	‘elephant’ [SG]	‘Loham’ (name) [SG,PROPER]
S	ongol-i	Lohám-i
A	ongol-i	Lohám-i
P	ongol	Lohám ⁸
PR	ongol- <u>o</u>	Lohám-i

Table 14: Tennet (Randal 1998:225, 261, 268–270)

There is also the suffix -a which always has some kind of locational meaning, but it’s unclear whether it can be used with animate nouns. But since it attaches directly to the stem and thus doesn’t interact with any of the other case markers, it wouldn’t provide any further insight into the underlying structure.

⁸The changes in tone and vowel quality are not explained by Randal, but other proper names for example *Lokúli Lokúli-i* don’t undergo stem changes (both on page 261).

This pattern is compatible with our adjusted hierarchy in (14) and further corroborates the systemacity of S = A = PR syncretism.

- (17) a. {S, A, PR} ≠ P (proper)
 b. {S, A} ≠ PR ≠ P (common)

Again we need to use a \emptyset -morph for P in order to avoid it being contained inside all other forms. My hierarchy works for Tennesse as S, A and PR can all be brought into adjacencies in (17).

2.4 Haro

Unlike the previously discussed marked-S languages, Haro overtly marks P arguments. König (2006) distinguishes between two types of marked-nominative languages: those which mark S & A arguments overtly in the same way, P being bare (type 1) and those which mark S & A as well as P overtly, but the functionally unmarked form of a nominal is identical to that of the P argument (type 2).

The marking of S, A and P in many Indo-European languages is identical to that of Haro, but when looking at the marking of further arguments, the patterns diverge.

	‘lion(ess)’ [SG,INDEF]	‘man’ [SG,DEF]	‘boy’ [SG,DEF]	‘woman’ [SG,DEF]
S	gáárma	ʔassá-z-i	šaató-z-i*	mačča-t-i
A	gáárma	ʔassá-z-i	šaató-z-i*	mačča-t-i
PR	gáárma	ʔassá-z-i*	šaató-z-i	mačča-t-i*
P	gáárma	ʔassá-z-a*	šaató-z-a*	mačča-t-o
R	gáárma-si*	ʔassa-z-í-si	šaato-z-í-si	mačča-t-i-si
TL	gáárma-na*	ʔassa-z-í-na*	šaato-z-í-na*	mačča-t-i-na*

Table 15: Haro

Differences also exist for the functionally unmarked form. Forms of address (vocative), citation, possessors and copula complements most often take the form of the nominative (S & A) in accusative languages.

In the sample of marked-S languages these functions were the least likely to be encoded by the S-case. They were most often equal to P, except for possessors which had a dedicated genitive morph 60% of the time. Still 30% are bare, while 10% receive the form covering the S argument.

And in fact Haro's definite nominal paradigms run counter to the morph sharing patterns exhibited by prototypical accusative paradigms, but mirror some of the discussed ergative languages in using A/PR bases for oblique cases.

Interestingly, the language family Haro belongs to, Omotic, uses the form of the S-case for usages typical of prototypical accusative languages more often than any of the other families of the sample Handschuh (2014:226). Note that Haro itself wasn't in her sample though.

According to Woldemariam (2009:104), only definite nominals are always inflected for case. Indefinite nominals like *gáárma* 'lion(ess)' in table 15 are only marked for 'peripheral' cases.

Since indefinite PR arguments are left bare alongside S, A and P, she considers genitive to be one of the 'core' cases in Haro. PR being in a class alongside A and P reflects that. The fact that indefinite PR arguments are bare as well is another argument for its inclusion among the core cases.

I've marked forms I constructed based on Woldemariam's description with an asterisk, all other forms are directly taken from Woldemariam (2009:100–102, 104–106, 111).

If we were to treat type 2 marked-nominative languages such as Haro as if they were simply accusative languages, Haro's syncretism patterns would stand out as a unique and unexpected exception. Through my model we can see that marking patterns of S, A and P simply converge on the surface.

The examination of these patterns might thus be used as a diagnostic to determine whether a given language with the coding $S = A \neq P$ is underlyingly accusative (P intervenes) or type 2 marked-nominative (A intervenes), if little syntactic data is available.

3 Conclusion

I adjusted Zompì's (2017) hierarchy in order for it to also account for patterns of syncretism and overt containment in marked-S languages, without foregoing any patterns of ergative, accusative and tripartite systems the hierarchy was already able to derive.

The adjustments even enable the derivation of previously problematic patterns involving possessors without resorting to additional mechanisms like Harðarson (2016) and Starke (2017) do.

Most important was the decomposition of cases, especially nominative and absolutive, into abstract, grammatical roles. The inclusion of PR among the intervening roles A and P allows the model to generate more patterns, but these patterns are corroborated by marking patterns of many different languages. It even allows for the derivation of some previously problematic data.

We also saw again that, even when only looking at three languages, not all marked-nominative languages pattern alike. A shortcoming of my approach is that due to the scarce data, some patterns might be overgeneralized. For example the patterns in Tenneset and Haro which are limited to salient nominals.

Another open question concerns the nature of my proposed classes, as they don't correspond to Marantz's (1991) dependent cases anymore. A possibility is that the classes form based on the number of arguments in a clause.

S is always the only argument of a given clause, A and P are the two dependents of transitive clauses, while PR is the only dependent in a DP of at least two nominals. The obliques might then be in some sort of residual class. I leave this open for future research.

References

- Bierwisch, Manfred. 1967. Syntactic features in morphology: general problems of so-called pronominal inflection in German. In Cornelis H. van Schooneveld (ed.), *To honor Roman Jakobson: essays on the occasion of his 70. birthday*, vol. 31 (Janua Linguarum. Series Maior), 239–270. De Gruyter Mouton.
- Blevins, James P. 2009. Case and declensional paradigm. *The Oxford Handbook of Case*. 200–218.
- Bobaljik, Jonathan David. 2012. *Universals in comparative morphology: Suppletion, superlatives, and the structure of words*. MIT Press.
- Caha, Pavel. 2009. *The nanosyntax of case*. University dissertation.
- Caha, Pavel. 2010. The parameters of case marking and spell out driven movement. *Linguistic Variation Yearbook* 10(1). 32–77. <https://doi.org/https://doi.org/10.1075/livy.10.02cah>.
- Dixon, Robert M. W. 1994. *Ergativity*. R. M. W. Dixon (ed.). Vol. 69 (Cambridge Studies in Linguistics). London: CUP. 271.
- Gippert, Jost. 1987. Zu den sekundären Kasusaffixen des Tocharischen. *Tocharian and Indo-European Studies* 1. 22–39. <https://doi.org/http://dnb.info/1143602056/34>.
- Handschuh, Corinna. 2014. *A typology of marked-S languages*. Vol. 1 (Studies in Diversity Linguistics). Berlin: Language Science Press.
- Harðarson, Gísli Rúnar. 2016. A case for a Weak Case Contiguity hypothesis—a reply to Caha. *Nat Lang Linguist Theory* 34. 1329–1343. <https://doi.org/https://doi.org/10.1007/s11049-016-9328-x>.
- Johnston, Jason Clift. 1996. *Systematic homonymy and the structure of morphological categories: Some lessons from paradigm geometry*. University of Sydney dissertation.
- Kießling, Roland. 2007. The “marked nominative” in Datooga. *Journal of African Languages and Linguistics* 28(2). 149–191. <https://doi.org/10.1515/jall.2007.009>.
- König, Christa. 2006. Marked Nominative in Africa. *Studies in language* 30(4). 705–782.

- Lander, Yury. 2009. Varieties of genitive. In Andrej Malchukov & Andrew Spencer (eds.) (ed.), *The Oxford handbook of case*, 581–592. New York: Oxford University Press.
- Marantz, Alec. 1991. Case and Licensing, book = Proceedings of ESCOL 1991, optjournal = , optpublisher = , editor = German F. Westphal, Benjamin Ao, Hee-Rahk Chae (eds.) In, 234–253.
- Randal, Scott. 1998. A grammatical sketch of Tennet. In Gerrit J. Dimmendaal & Marco Last (eds.), *Surmic languages and cultures*, vol. 13 (Nilo-Saharan: linguistic analyses and documentation (NISA)), 219–272. Köln: Rüdiger Köppe.
- Starke, Michal. 2017. Resolving (DAT = ACC) ≠ GEN. *Glossa: a journal of general linguistics* 2(1). 104. <https://doi.org/https://doi.org/10.5334/gjgl.408>.
- Woldemariam, Hirut. 2009. Haro. In Gerrit J. Dimmendaal (ed.), *Coding participant marking: construction types in twelve african languages*, vol. 110 (Studies in Language Companion Series (SLCS)), 97–122. Köln: John Benjamins Publishing Company.
- Zompì, Stanislao. 2017. *Case decomposition meets dependent-case theories*. Università di Pisa MA thesis.

Eigenständigkeitserklärung

Hiermit erkläre ich, dass ich diese Arbeit selbstständig und nur unter Verwendung der angegebenen Quellen verfasst habe. Ich habe alle Zitate als solche gekennzeichnet.

Affidavit

I hereby declare that I have authored this thesis by myself, using only the cited literature. All quotations have been acknowledged and marked as such.

Ort, Datum

Unterschrift