Natural Language and Linguistic Theory manuscript No.
(will be inserted by the editor)

Explaining the structure of case paradigms by the mechanisms of Nanosyntax

The Classical Armenian nominal declension

Pavel Caha

Received: date / Accepted: date

Abstract This paper looks in detail at the Classical Armenian declension. I highlight several generalizations that can be read off the surface paradigms, including restrictions on syncretism, fusional vs. agglutinative expression of categories and the emergence of unexpected thematic vowels. Subsequently, I explain these generalizations within the framework of Nanosyntax (Starke 2009, 2011). The defining features of the account are fine-grained syntactic representation (a single feature per syntactic terminal) and phrasal spell-out. I argue that these two tools allow us to replace a separate level of morphological (paradigm specific) structure by a syntactic tree.

Keywords Case, Case syncretism, Classical Armenian, paradigm, phrasal spell out, nanosyntax

1 Introduction

There seems to be a broad consensus in the literature that in order to describe the fine details of morphological paradigms, one needs something in addition to the syntactic structure and the lexicon (for some recent work in the generative tradition, see, e.g., Williams 1981, 1994, McCreight and Chvany 1991, Halle and Marantz 1993, Ackema and Neeleman 2007). The current paper focuses in detail on case declension of Classical Armenian in order to (re-)evaluate the type of evidence usually brought in favor of separate morphological (or more narrowly paradigm) structure. I argue that once syntax is made sufficiently fine-grained, with individual features corresponding to syntactic heads, paradigm structure may be dispensed with.

CASTL, University of Tromsø, 9037 Tromsø, Norway

Tel.: +47-776 45523
Fax: +47-776 44239
E-mail: pavel.caha@hum.uit.no

I start by looking at case syncretism. The term refers to a situation in which a particular noun does not morphologically distinguish two (or more) distinct cases. An example is given below. In (1), we can observe that the noun 'God' has the same form in the ablative and the genitive. In (2), we see a noun that distinguishes the two environments.[1]

(1) Syncretism of GEN-ABL

a. ordi-k' Astuc-**oy**
son-NOM.PL God-GEN
'the children of God'

b. i Astuc-**oy**
by God-ABL
'by [the benevolent] God'

(2) No syncretism of GEN-ABL

a. merjawor Daniêl-**i**
acquaintance Daniel-GEN
'an acquaintance of Daniel'

b. i Daniêl-**ê**
from Daniel-ABL
'from Daniel'

Syncretism is a pervasive feature of natural languages (Baerman et al. 2005), and the declension of Classical Armenian is a good example of this. In §2-§3, I show that in Classical Armenian, syncretism is subject to a restriction such that in a linear sequence of cases, only adjacent cases may enter this relationship. The existence of such a type of constraint has led McCreight and Chvany (1991, p.91) to argue that "syntactic features are inappropriate to the modeling of paradigms, and that geometric representations [...] can provide more insight, particularly in the description of syncretism."

This paper argues that no specific paradigm structure is needed in order to account for the data, including the observed restrictions. In §4, I demonstrate that the facts may be captured using a particular case decomposition, such that when we go left to right in the linear sequence uncovered, cases grow in complexity with each case adding a feature compared to the previous one (cf. Starke 2009, Bobaljik 2012). §5 compares the proposal to some common alternative approaches, and shows its advantages. In sum, the current theory is reductionist: a specific module dedicated to paradigm structure is not necessary in order to account for the facts (cf. Bobaljik 2002 and further work in that spirit).

At the same time, I argue in §6 that each case feature corresponds to a syntactic head (cf. Caha 2009, 2011).[2] The argument is based on the observation that case attraction in Classical Armenian is sensitive to the linear sequence that restricts syncretism. Thus, despite its reductionist nature, the theory does not deny the existence of paradigm structure. Paradigm structure is real, but not autonomous: it corresponds to the syntactic structure.

§7 is devoted to a detailed presentation of phrasal spell-out, a device that may map both terminals as well as complex phrasal nodes directly on their pronunciation (drawing on Starke 2009). With the proposal in place, §8 turns

[1] Examples from Krause and Slocum (no date).

[2] To put the idea in a context: within Cartography (Cinque and Rizzi 2010), it is generally agreed that each feature is hosted by a separate head. A related idea is presented in Kayne (2005). Starke (2009, 2011) develops these ideas in a specific framework, Nanosyntax, whose research program I follow closely.

to its consequences for the splits between agglutination and fusion (cf. Plank 1999). I use the term fusion in order to describe the fact that sometimes, case is expressed together with other inflectional categories. For example, in (3-a), it is expressed together with number. The term agglutination, on the other hand, refers to instances where such categories are expressed separately, as in (3-b). The separation of case and number is seen in comparison with the singular form, which is just like the plural minus the plural marker *-k‘*. (No such segmentation is possible in (3-a)).

(3) a. azg-i — azg-s
nation-LOC.SG nation-LOC.PL
b. azg-aw — azg-aw-k‘
nation-INS.SG nation-INS-PL

In §8 I observe that agglutination and fusion are not distributed randomly across the paradigm. Specifically, agglutination in a particular case K entails agglutination in all cases that can be found to the left of K in the linear sequence relevant for syncretism — apparently yet another paradigm structure effect. The theoretical challenge is to model the varying number of exponents in the two distinct cases shown in (3-a,b) (as well as capturing the regularity such splits are subject to): what blocks **azg-i-k‘* as the plural of *azg-i* in (3-a)?

Williams (1994, 22) says concerning a similar example that “in general, if there are two ways of filling a slot in the paradigm [*azg-i-k‘* vs. *azg-s*], only one may be used. [...] This reveals that there is a target slot to fill, which is independent of the rules for filling it, and that slot is given by the paradigm.”

Working in a Distributed Morphology model, Halle and Vaux (1998) answer the challenge differently. They propose that there is only a single terminal in syntax for both case and number (in effect blocking *azg-i-k‘* by a proposal about the syntax of Classical Armenian). However, for the instrumental plural, they are forced to propose a morphology specific rule (Fission) that produces two surface positions.[3] Thus, despite important differences in execution, both solutions entail the existence of a morphology specific level of representation (corresponding to the paradigm structure, with a particular number of slots to fill).

In §9, I offer an alternative solution in terms of phrasal spell-out. I show that under this view, the number of exponents that are needed to spell out each particular case naturally emerges from the way lexical insertion interacts with syntactic structure. There is no need for any morphological rules (like Fission) producing paradigm representations with a particular number of slots, a conclusion I strengthen in §10. Here, I focus on two consonantal declensions in the language, and show that apparently unexpected thematic vowels show up in this type of declension. I demonstrate, however, that no additional rules are needed to explain their existence; the vowels appear as a consequence of the

[3] Had there been two positions in syntax, a separate rule (presumably Fusion) would be needed to handle those cases where only a single marker appears.

proposed spell-out procedure interacting with the independently established syntactic structure.

The result is a theory of paradigm structure that eschews both syntax independent structure and syntax independent rules, leading to the claim that paradigm structure is in fact best accounted for in terms of syntactic structure, using independently needed notions such as binary branching, structural intervention and others.

2 A linear restriction on syncretism

I open the empirical discussion by introducing a strong restriction on syncretism in Classical Armenian. In establishing the generalization, I rely on the description offered in Schmitt (1981), and I also partly draw on aspects of the analysis of Classical Armenian presented in Halle and Vaux (1998).

According to standard descriptions, Classical Armenian has seven cases: nominative, accusative, locative, genitive, dative, ablative and instrumental. I will stick to this description, noting that the INS can also be used as a comitative (so potentially there can be eight cases), and that the GEN and DAT are not formally distinguished in the nominal declension (so potentially there can be six cases, with the caveat that GEN/DAT are distinct in pronouns, including demonstratives). Putting aside for now one apparent exception (to be discussed later), syncretism in case is restricted to contiguous regions in a linear sequence, which I give in (4-b).[4]

(4) *Case Contiguity (Armenian)*:
  a. Non-accidental case syncretism is restricted to target contiguous regions in the following sequence:
  b. NOM − ACC − LOC − GEN − DAT − ABL − INS

The statement (4) rules out a number of logically possible syncretisms. For instance, NOM and LOC cannot be syncretic to the exclusion of ACC, because they are not contiguous in (4-b). This restrictive prediction of (4) is borne out and the syncretism is unattested. I show that schematically in the first column of the table (5) (syncretic cells highlighted by shading).

[4] The phrase "non-accidental" in (4) is intended to exclude two sources of homophony: (i) phonological conflation, whereby two distinct underlying forms end up homophonous due to a regular phonological process, and (ii) accidental homophony. Accidental homophony plays no role in this paper; however, one instance of homophony will be analyzed as an example of phonological conflation in the sense described above (see §10.3).

(5) *The restrictions on syncretism in Classical Armenian*

	N.A.	spirit (SG.)	N.A.	word (SG.)	N.A.	Tigran (SG.)
NOM	A	hogi-ø	...	bay-ø	...	Tigran-ø
ACC	B	hogi-ø	...	bay-ø	...	Tigran-ø
LOC	A	hogi-ø	A	bay-i	A	Tigran-ay
GEN	A	hogw-oy	B	bay-i	B	Tigran-ay
DAT	B	hogw-oy	A	bay-i	B	Tigran-ay
ABL	A	hogw-oy	...	bay-ê	A	Tigran-ay
INS	...	hogw-ov	...	bay-iw	...	Tigran-aw

It is, however, not the case that NOM and LOC are never the same; see the second column of the table (5). Crucially, when this happens, then ACC is syncretic with the NOM–LOC pair, leading to a syncretism which targets a contiguous region of (4-b). Similarly, as shown in the first column, GEN and ABL are never syncretic to the exclusion of other cases. But again, such a syncretism becomes possible once it covers DAT as well, applying to a region contiguous in (4-b) (see the dark shade in the second column).

A similar prediction is that, for instance, LOC–DAT or LOC–ABL syncretisms are unattested, but LOC–GEN–DAT or LOC–GEN–DAT–ABL can be found. This is shown in the remainder of the table (5).[5]

To evaluate the restrictive power of such a linear constraint abstractly, consider some numbers. In a system with seven cases (Classical Armenian), there are 120 logically possible syncretisms. The statement (4) predicts that 99 of these are unattested, and allows only 21.

It has been argued in the literature that Classical Armenian is not unique in possessing a linear contiguity constraint, and that such a situation is quite general; see McCreight and Chvany (1991), Johnston (1996), Caha (2009) and references there.[6] If this is so, such a large scale descriptive reduction of possibilities in language after language raises a challenge for the theory of syncretism. Clearly, the theory should be able to derive this, and, preferably, offer an account that extends to other aspects of case than syncretism.

In this paper, I offer a way to understand the linear constraint (4) as a consequence of the proposal that case features are syntactic heads, ordered in a functional sequence. I also show that the same linear sequence is relevant for

---

[5] DAT–LOC syncretism across a distinct GEN is attested in the pronominal declension (e.g., *mez/jez* 'we, you, ACC/LOC/DAT,' *mer/jer* 'my, our'). The reason for this apparent exception is that the "genitive" pronoun is in fact a possessive pronoun, which formally does not belong to the paradigm, and hence, disturbs the picture.
There are several pieces of evidence to support this claim: (i) unlike regular genitives, the pronominal possessive forms take additional agreement suffixes; (ii) unlike regular genitives, the possessive forms show no syncretism with any other case; this is expected if they are not case forms, but unexplained otherwise; (iii) the "genitive" ending *-r* (*me-r/je-r* 'my, our') shares diachronic origin with morphemes routinely classified as possessive markers and not as case markers, e.g., in Latin forms such as *nos-ter* 'our.'

[6] Similar observations have been made in other domains than case, see Starke (2005/2009) Bobaljik (2007/2012), Pantcheva (2010, 2011) and Vangsnes (in press) for approaches directly related to the one pursued here.

other phenomena in the grammar, and work out a proposal how this follows from the initial proposal developed for syncretism.

2.1 Establishing the order

Let me begin by introducing a different aspect of the constraint (4), which is positively predictive. In particular, if a language does not show any syncretism, then any linear ordering of cases yields (trivially) a correct generalization. Evidence for any particular ordering is thus more convincing if supported by as many attested examples as possible.

This aspect of the linear constraint is illustrated on a representative sample of paradigms shown in the table (6). The table is organized in such a way that the cases are ordered top-down according to the sequence given in (4-b), i.e., NOM–ACC–LOC and so on. The cells in darker shade show pair-wise syncretisms of adjacent cases, and move gradually one notch down as we go in the table from left to right. Lighter shade marks cells where syncretism (irrelevantly for the main point) extends beyond a simple pair.

(6) *Attested syncretisms in Classical Armenian*

	word (SG.)	nation (PL.)	nation (SG.)	year (SG.)	river (PL.)
NOM	bay-ø	azg-k‘	azg-ø	tari-ø	get-k‘
ACC	bay-ø	azg-s	azg-ø	tari-ø	get-s
LOC	bay-i	azg-s	azg-i	tarw-oĵ	get-s
GEN	bay-i	azg-ac‘	azg-i	tarw-oy	get-oc‘
DAT	bay-i	azg-ac‘	azg-i	tarw-oy	get-oc‘
ABL	bay-ê	azg-ac‘	azg-ê	tarw-oĵê	get-oc‘
INS	bay-iw	azg-awk‘	azg-aw	tare-aw	get-owk‘

The linear order is then established as follows. In the singular of the noun ‘word,’ NOM and ACC show syncretism to the exclusion of all other cases. From the perspective of linear ordering, this means that they must be neighbors in the linear order relevant for syncretism: NOM–ACC. ACC and LOC are the same in the plural, see the shading in the plural of ‘nation.’ This leads to NOM–ACC–LOC. LOC in turn must be adjacent to GEN and DAT on the basis of the syncretism in the singular of ‘nation’: NOM–ACC–LOC–GEN/DAT.

In all the paradigms above, and in the nominal system in general, GEN and DAT are always the same. This means that their order cannot be determined internally to the nominal declension of Classical Armenian. Thus, the reason I state the order as GEN–DAT rather than DAT–GEN is not motivated by Classical Armenian alone. It is, however, justified by cross-linguistic data, which I discuss in the next section.

While the mutual order of GEN and DAT cannot be decided internally to Classical Armenian, it can be established that the ABL comes after these two cases, due to the syncretism in the plural (see ‘river’). This leads to the ordering NOM–ACC–LOC–GEN/DAT–ABL. INS shows no syncretisms in Classical

Armenian; it then comes either last or first. (It cannot come in the middle, because then it would disturb the needed adjacency between other cases.)[7]

We are thus left with four possible orderings, and hence, four possible ways to state a linear constraint on syncretism in Classical Armenian. I give them in (7):

(7) *Four possible sequences with syncretisms contiguous*

a. NOM − ACC − LOC − GEN − DAT − ABL − INS
b. NOM − ACC − LOC − DAT − GEN − ABL − INS
c. INS − NOM − ACC − LOC − GEN − DAT − ABL
d. INS − NOM − ACC − LOC − DAT − GEN − ABL

2.2 Summing up

As we have seen, the syncretisms of Classical Armenian occupy contiguous regions in the linear sequence given in (4). While correct in the sense that the constraint shows minimal violations (with one exception to be presented and explained away in §10.3), the sequence in (4) is underdetermined by the actual data. First, there is no evidence for mutual ordering of GEN and DAT; second, since INS shows no syncretisms, it can either come last or first.

3 What is the Case sequence of Armenian?

In this section, I present some considerations which favor the statement (4) over possible alternatives.

3.1 Universal Contiguity

The first argument for the ordering (7-a) comes from cross-linguistic comparison. Caha (2009) proposes a hypothesis, Universal Contiguity, which says that across languages, there is a fixed sequence of cases in which only contiguous regions show syncretism:[8]

(8) *Universal (Case) Contiguity*:

a. Non-accidental case syncretism targets contiguous regions in a sequence invariant across languages.
b. The Case sequence: NOM − ACC − GEN − DAT − INS − COM

[7] There is actually one homophony of INS with DAT, but this is due to a phonological conflation (see Halle and Vaux 1998, n.7). I come back to this later on.

[8] It has to be mentioned that Plank (1991) argues that a linear contiguity constraint is too strong for a number of languages. See Johnston (1996) and Caha (2009) for a linear analysis of some of the languages identified as problematic in Plank (1991).

In the sequence (8-b), GEN precedes DAT, and INS comes after these two. In Classical Armenian, only the order (7-a) is consistent with this cross-linguistic pattern. Hence, it is reasonable to assume that it is the correct one.

Consider a couple of examples from three different branches of Indo-European that illustrate (8). In Russian (McCreight and Chvany 1991), the ordering is unambiguously NOM–ACC–GEN–PREP–DAT–INS, see (9). Relevantly, GEN precedes DAT, and DAT precedes INS.

(9) *Syncretism in Russian (McCreight and Chvany 1991)*

	window, sg.	teacher, pl.	both, m.i.	book, sg.	100
NOM	OKN-O	učitel-ja	dv-a	knig-a	st-o
ACC	OKN-O	UČITEL-EJ	dv-a	knig-u	st-o
GEN	okn-a	UČITEL-EJ	DV-UX	knig-y	ST-A
PREP	okn-e	učitel-jax	DV-UX	KNIG-E	ST-A
DAT	okn-u	učitel-am	dv-um	KNIG-E	ST-A
INS	okn-om	učitel-ami	dv-umja	knig-oj	ST-A

The order GEN–DAT–INS appears also in Old English (Plank 1991, Caha 2009):

(10) *Syncretism in Old English (Plank 1991)*

	thing,f.	eye,n.	daughter,f.	gift,pl.
NOM	þing	dohtor	ēage	ġiefa
ACC	þing	dohtor	ēage	ġiefa
GEN	þinges	dohtor	ēagan	ġiefa
DAT	þinge	dehter	ēagan	ġiefum
INS	þinge	dehter	ēagan	ġiefum

Finally, syncretism in Sanskrit (Plank 1991, Johnston 1996) is only consistent with an order which includes GEN–DAT–INS, where the position of DAT between GEN and INS is relevant for our present concerns.

(11) *Sanskrit (paradigms from Baerman 2008)*

	god, DU.	god, DU.	fire, SG.	god, PL.	god, DU.
NOM	devāu	devāu	agnis	devās	devāu
ACC	devāu	devāu	agnim	devān	devāu
LOC	devayos	devayos	agnu	deveṣu	devayos
GEN	devayos	devayos	agnes	devānām	devayos
ABL	devābhyām	devābhyām	agnes	devebhyas	devābhyām
DAT	devābhyām	devābhyām	agnaye	devebhyas	devābhyām
INS	devābhyām	devābhyām	agninā	devāis	devābhyām

These observations allow us to understand the pattern found in Classical Armenian as a special instance of the general scenario (8). If that is so, we have to order GEN before DAT, and INS must be last rather than first.[9]

[9] Needless to say, there are a number of questions that arise once cross-linguistic perspective is taken. Two issues deserve mentioning: (i) local cases show up at places that are not

3.2 Case attraction

There is additional evidence that INS should come last (and not first) in the sequence: case attraction (Plank 1995, p.43, Blake 1994). In Classical Armenian, the complement of a noun is "normally" expressed by the GEN, as schematically depicted in (12-a). However, if the head is in ABL or INS, the dependent GEN can be "attracted." Attraction consists in replacing GEN by the case which is carried by the head noun, see (12-b,c). However, if the head noun is in another case (NOM, ACC, or LOC), attraction is unattested, see (12-d):[10]

(12) Case attraction in Classical Armenian

a. [N [N-GEN]]

b. [N-ABL [N-GEN ⇒ N-ABL]]

c. [N-INS [N-GEN ⇒ N-INS]]

d. *[N-NOM/ACC/LOC [N-GEN ⇒ N-NOM/ACC/LOC]]

A real language example of attraction is in (13).

(13) [bazmut'-eamb [zawr-**awk'**-n Hay-oc']]
crowd-INS.SG force-INS.PL-DEF Armenian-GEN.PL
'with a crowd of the Armenian forces' (Plank 1995, p.43)

In (13), the boldfaced head noun 'crowd' bears INS. The head has a complement, 'of the Armenian forces.' The head of the complement, 'forces,' would 'normally' occur in GEN. However, as a result of attraction, it bears INS (see the boldfaced affix), inherited from the head noun.

Importantly, if we choose a sequence with INS adjacent to ABL, we capture not only the restrictions on case syncretism, but also the restrictions on case attraction, as seen in (12-b-d). Specifically, in the syncretism sequence in (4), cases to the right of GEN have the power to attract it, cases to its left cannot.

3.3 Summing up

To sum up the two preceding sections: syncretism in the Classical Armenian nominal declension is restricted to contiguous regions in a linear sequence of cases. Out of four possible orderings, we have chosen the one which (i) captures additional restrictions in the language (case attraction), and (ii) is consistent with a larger cross-linguistic pattern. In the next section, I provide a way to understand the case sequence theoretically.

---

directly comparable; (ii) particular languages present challenges to the view expressed in (8) (the reviewers mention specifically Latin). I cannot possibly address these issues within the space of an article, and I can only provide references to places where these challenges are addressed. For local cases, see Caha (2009, §3.4.3-4) and Caha (to appear). For cross-linguistic issues (including the discussion of Latin), see Caha (2009, ch.3&8).

[10] It cannot be decided whether attraction does or does not occur in the dative. Since the dative is always the same as the genitive, case attraction, if active, applies vacuously.

4 Deriving Contiguity

In this section, I show that (4) follows from a particular feature decomposition, interacting with a spell-out procedure based on the Elsewhere Condition. Relevant for the issues highlighted in the introduction is going to be the conclusion that linear contiguity is expressible in terms of feature decomposition, and does not need to be encoded through paradigm-specific notational devices.

4.1 Cumulative decomposition

The first thing we need is that the number of features characteristic for each case has to grow monotonically as we move along the contiguity sequence in (4). Thus, what we need is that, for instance, NOM = [A], ACC = [A, B], LOC = [A, B, C] and so on, see (14). I will call such a decomposition "cumulative."

(14) Cumulative case decomposition
  a. NOM = [A]
  b. ACC = [A, B]
  c. LOC = [A, B, C]
  d. GEN = [A, B, C, D]
  e. DAT = [A, B, C, D, E]
  f. ABL = [A, B, C, D, E, F]
  g. INS = [A, B, C, D, E, F, G]

What is crucial to derive Case Contiguity is the formal character of the decomposition, and not the content of the features. Consequently, the proposal is compatible with a number of interpretations, not crucial to the argument (though see §6.2 for a suggestion). Let me also remark that even though (14) may suggest that case features are privative, this is not necessary. For example, NOM can be [+case], ACC can be [+case; −independent], etc. As long as we keep adding features, the resulting system can be used to derive contiguity.

4.2 Two Spell out rules

I now proceed to show how this representation contributes to understanding the contiguity requirement on syncretism. The gist of the proposal is the idea, shared among various frameworks, that lexical entries are not tailor-made for one representation only, but they can be associated to a larger number of representations. Linear Contiguity then follows if an entry may only target a set of cases that form a contiguous region on the scale in (4).

With the decomposition (14) in place, such a goal may be achieved in two ways. Either we propose that if a given lexical entry α applies to a case K, it applies to all cases that contain K. For example, if α is specified for GEN ([A,B,C,D]), it applies automatically also in DAT, ABL, INS. This strategy is known as the Subset Principle (see, e.g., Halle 1997 among many others).[11]

[11] The Subset Principle, as Halle (1997) states it, is in fact a combination of a "pure" Subset Principle and the Elsewhere Condition. Here, I ignore the Elsewhere part of the Subset Principle, and treat it as a separate rule (cf. Harley 2008).

Alternatively, we may propose that if a lexical entry α applies to a case K, it applies also to all cases contained in K. Thus, if α is specified for GEN, it applies also in LOC, ACC, NOM. This assumption is called the Superset Principle (see Starke 2009). I state the Superset Principle informally below:

(15) *The Superset Principle (preliminary version)*: A phonological exponent is inserted into a node if its lexical entry contains all features of that node.

It can be shown that if one adopts the representations in (14), both approaches derive (4) as a theorem. In what follows, I demonstrate this only for the Superset Principle (see Bobaljik 2012 for a related proposal based on the Subset Principle). This choice is made on the basis of the observation – to be justified in §9.3 – that for certain facts of Classical Armenian, the Superset Principle offers a better solution than its alternative.

The theory sketched so far constrains syncretism to contiguous regions in (4), but it is as yet incapable of dealing with syncretisms that do not include NOM. Thus, recall that the entry for a given case automatically applies to all cases contained in it, leading only to syncretisms that stretch from the particular case to NOM. This apparent problem disappears once competition among entries is taken into consideration, and the Elsewhere Condition is adopted to regulate it. The reasoning follows.

I start from the observation that there are situations where more than one entry is applicable in a given case. Suppose, for example, that there are the two following entries, α and β:

(16) a. $/\alpha/ \Leftrightarrow$ [C,B,A]
b. $/\beta/ \Leftrightarrow$ [A]

In (16-a), α is specified for the features of LOC. By the Superset Principle, it may then also be inserted in ACC and NOM. β is specified for NOM. According to the Superset Principle, this exhausts its applicability. The result is that in NOM, both rules may apply:

(17) *The range of applicable environments of α and β*

case	form
NOM	α, β
ACC	α
LOC	α

The two entries 'compete,' and the result is determined by the Elsewhere Condition. This condition says that the most specific entry wins over the others (see Kiparsky 1973, my formulation draws on Neeleman and Szendrői 2007).

(18) *The Elsewhere Condition*: In case two rules (R1 and R2) can apply in an environment E, R1 takes precedence over R2 if it applies in a proper subset of environments compared to R2.

In our example, the entry for β (R1), takes precedence over α (R2) in NOM, because it applies in a proper subset of environments. (It does not apply in ACC, LOC.) The result of the competition is that α surfaces in LOC and ACC only, a contiguous region that does not include NOM. Hence, once we introduce the Elsewhere Condition, the generative capacity of the system is increased to yield also contiguous syncretisms without NOM.

Consider now how the proposed system derives the Contiguity constraint. To see that it does, suppose that we want to encode a syncretism which would violate it, e.g., NOM and LOC are the same to the exclusion of ACC, as in the hypothetical paradigm (19). If it turns out that such a syncretism cannot be encoded by the spell-out system operating on the proposed decomposition, we prove that the system derives Contiguity, as manifested in Classical Armenian.

(19) *An offending paradigm*

case	form
NOM	α
ACC	β
LOC	α

To generate the offending paradigm, we have to come up with an entry that can appear both in the locative and the nominative. Such an entry is (20).

(20) $/\alpha/ \Leftrightarrow$ [C B A]

Now we need an entry which can spell out ACC (B, A) but not LOC (C, B, A), see (21). Such an entry provides a perfect match for ACC, and due to competition, remove it from the set of cases where the entry (20) surfaces.

(21) $/\beta/ \Leftrightarrow$ [B A]

However, the entry (21) can apply in NOM (A) as well. Hence, the entries (20) and (21) compete not only for ACC, but also for NOM, see (22). In such a situation, the rule introducing β takes precedence over α in NOM as well, because it is a better match; see (23).

(22) *The applicability of α and β*

case	form
NOM	α, β
ACC	α, β
LOC	α

(23) *The paradigm generated*

case	form
NOM	$\cancel{\alpha}$, β
ACC	$\cancel{\alpha}$, β
LOC	α

In words: whenever we get α in LOC and β in ACC, we fail to get α in NOM. This means that the system derives Contiguity, because it cannot generate offending paradigms. We thus leave this section with the conclusion that Contiguity follows from the interaction of two proposals: (i) cumulative decomposition and (ii) the Superset Principle and the Elsewhere Condition.

5 Case Contiguity and alternative approaches

This section considers two common approaches to case morphology — one based on cross-classification, and one on feature geometries. I show that these alternatives fail to capture linear contiguity — in fact, they provide virtually no constraints on possible vs. impossible syncretism.

5.1 Jakobsonian cross-classification and its descendants

In Jakobson (1962), cases are represented as collections of equipollent features (like [+/− marginal], [+/− ascriptive]), each of which *cross*-classifies the full set of cases into two groups. This system has the immediate advantage that natural classes of cases can be referred to with the help of such features, and syncretism can be restricted to these classes.[12]

From the perspective of Case Contiguity, two formal properties of such systems appear problematic (see also McCreight and Chvany 1991 and Johnston 1996). First, the expressive power of cross-classification goes beyond Contiguity. To see that, consider the cross-classification of 4 cases – NOM, ACC, LOC and GEN – by two features, X and Y (24). The particular decomposition is proposed so as to match the facts of Classical Armenian as close as possible. The natural classes definable by such a decomposition are given in (25).

(24) Cross-classification

	+Y	−Y
+X	NOM	ACC
−X	GEN	LOC

(25)
a. [+X]: {NOM, ACC}
b. [−Y]: {ACC, LOC}
c. [−X]: {LOC, GEN}
d. [+Y]: {NOM, GEN}
e. [Ø]: {NOM, ACC, LOC, GEN}

The natural classes (25-a-c) are those that we need to capture the observation that syncretism targets contiguous regions in the sequence NOM–ACC–LOC–GEN. However, the syncretism in (25-d) should be disallowed. Thus, the conclusion is that a system of cross-classification predicts a linear contiguity constraint to be false. Equivalently, the existence of a linear contiguity constraint proves such a model wrong.

The second problem is that the system also under-generates (compared to Contiguity): as things stand, there is no way to define syncretisms of three terms (contiguous or not). However, there is no known constraint to the effect

[12] Relevant recent references include (among many others) Halle and Vaux (1998), Bobaljik (2002), Müller (2003), Embick and Noyer (2007), Calabrese (2008), Harley (2008).

that syncretism can target sets of two and four cases, but not three, or any other similar restriction (see (5) for examples of three term syncretisms).

To increase the generative capacity of the system in the right direction, cross-classification is usually accompanied by a mechanism which incorporates the Elsewhere Condition (18).

With the Elsewhere Condition in place, we get a three term syncretism by the interaction of two rules, (26-a,b).[13]

(26)	a.	[+X,+Y]:	{NOM}	$\rightarrow$	/phon A/
	b.	[Ø]:	{NOM, ACC, GEN, DAT}	$\rightarrow$	/phon B/

These rules (by themselves) pick out overlapping natural classes of cases: (26-a) says that NOM is realized by /phon A/, and (26-b) requires that any member of {NOM, ACC, GEN, DAT} is pronounced as /phon B/. The two rules clash for NOM. The winner is determined by the Elsewhere Condition (18). Since /phon A/ applies in a proper subset of cases compared to /phon B/, /phon A/ takes precedence for NOM. As a result, NOM = /phon A/, ACC, GEN and DAT = /phon B/.

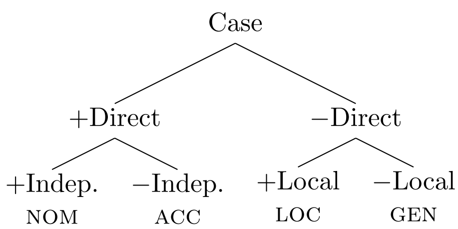
However, this is a deadly fix. The result is that any syncretism becomes possible. The reasoning is this: the entry (26-b) can, in principle, insert /phon B/ to any of NOM, ACC, GEN, DAT. Consequently, we can get any triplet of these by assuming that any one of the cases is spelled out by an entry similar to (26-a). And we can also get any pair by assuming (26-b) and that two cases of our choice are spelled out by two rules similar to (26-a).

In conclusion, if Case Contiguity is correct, Jakobsonian cross-classification is not an adequate model of case representation. Its shortcoming is purely formal, and holds regardless of the content of features, and how the pluses and minuses distribute over individual cases.

5.2 Feature geometries

A number of proposals add structure to the proposed features, and go thus in the direction of the present account.[14] For instance, adapting the proposal by Williams (1981), we could device the following feature geometry for a part of the Armenian system:

(27)



[13] The rule in (30a) reads as follows: the feature matrix [+X, +Y], corresponding to NOM, is realized by the phonology /phon A/.

[14] Relevant references include Williams (1981, 1994), Wiese (2004), McFadden (2007).

I will not comment on the content of the features, and note only a formal difference between such a geometry and Jakobsonian cross-classification (24). In particular, the feature which distinguishes among the direct cases ([independent]) is different from the feature which distinguishes among the indirect cases ([local]). This is in contrast with the cross-classification (24), where the feature which distinguishes NOM from ACC [+/−Y] is the same that distinguishes between LOC and GEN. This difference represents a step in the right direction, because it was the feature [+Y] which allowed for a non-contiguous syncretism to be singled out; see (25-d). Thus, in (27), each feature defines a set that forms a contiguous region in the syncretism sequence.

Looking into Williams' system in more detail, we can think of lexical entries as specified for non-terminal nodes of the feature tree above. A particular case is then spelled out by the entry that is closest to it in the tree hierarchy. For example, NOM-ACC syncretism is accounted for by an entry that targets the node [+direct]. Starting from either of NOM/ACC, and going upwards, this is the entry that we encounter first. In cases NOM and ACC are not syncretic, lexical items are fully specified and target corresponding terminal nodes.

Then, in order to get ACC-LOC syncretism, we must postulate an entry specified for the root node [case], applicable in all four cases. If we then further postulate tailor-made entries for NOM and GEN, ACC-LOC syncretism emerges.

With such a mechanism, however, any syncretism becomes possible once again. To show that, let me keep the entry that targets the [case] node; this entry may apply in any of the cases. With this entry in place, we get any triplet of cases by inserting a distinct marker in one terminal node of our choice. In a similar way, if we have two competing entries targeting the terminal nodes, any pair of cases can be syncretic.[15]

5.3 Summary

The previous two sub-sections have argued that two common feature systems intended to account for syncretism fail to derive linear contiguity. The reason for this are the formal properties of the feature representation employed. In particular, in these systems, each case has an 'exclusive address:' the feature combination that characterizes a given case neither contains, or is contained by any other case. In such systems, 'default' entries can be restricted in arbitrary ways by competing lexical entries targeting individual cases, with the result that any syncretism may be generated.

[15] Formally, the system makes some non-trivial predictions concerning syncretism. For example, if NOM and GEN are syncretic to the exclusion of other cases, then ACC and LOC cannot be. That is because both NOM-GEN and ACC-LOC syncretisms can be only obtained by inserting an exponent under the [case] node in the tree (27). Since only one marker can be inserted at a given node, the combination of the two syncretisms is impossible.
However, as far as I know, there is no known constraint on syncretism that has such an implicational format (if one type of syncretism is attested in a paradigm, then another type of syncretism cannot be, even though it is independently attested). If that is so, and such predictions find no empirical support, this in fact adds up to the drawbacks.

The current system differs in precisely this aspect: there is no 'exclusive address.' For example, if we have an entry apparently tailor-made for GEN (A,B,C,D), this entry also targets LOC (A,B,C), ACC (A,B), and NOM (A). Hence, the drawback of the systems discussed is not replicated.

However, the current system is not the only response Contiguity. Pursuing a similar agenda, other approaches have drifted away from features altogether (McCreight and Chvany 1991, Wiese 2003), and proposed that paradigms are (spatial) objects in their own right, corresponding to a linearly ordered set of cells. Markers are then specified for contiguous sub-spaces of such paradigms.

The main reason I do not adopt such an approach is theoretical simplicity. As we will see shortly, I argue that cumulative sub-classification can be implemented in syntax, using mechanisms that are independently needed. To the extent that such an approach turns out to be successful, it has the a priori advantage of making an independent concept of a paradigm space theoretically unnecessary. That does not mean giving up on the idea that paradigms are structured objects (as opposed to a set of cells); what it means is that paradigm structure reduces to (or derives from) syntactic structure.

6 The layered structure of case features

The general conclusion that I would like to build on in this section is that Case Contiguity, if correct, helps us establish the formal properties that a case decomposition should comply with. The crucial question is: Why should a case decomposition have such properties?

6.1 The tree structure

The answer I suggest here is based on the observation that cumulative classification is equivalent to the decomposition shown in (28), where individual features head projections in a binary branching syntactic tree:

(28) $[_{com}\ \mathrm{F}\ [_{ins}\ \mathrm{F}\ [_{dat}\ \mathrm{E}\ [_{gen}\ \mathrm{D}\ [_{loc}\ \mathrm{C}\ [_{acc}\ \mathrm{B}\ [_{nom}\ \mathrm{A\ DP}\ ]\ ]\ ]\ ]\ ]\ ]\ ]$

The tree encodes the proposal that a nominative DP is a type of syntactic constituent, in which the DP is the complement of the feature [A]. The accusative is a similar constituent, one which is built on top of the nominative ([A]) by the addition of [B], so that ACC = [A, B]. And similarly for the other cases: LOC = [A, B, C], GEN = [A, B, C, D], and so on. Thus, the proposal in essence says that the features A, B, C etc., needed for syncretism, are the primitives of the syntactic structure, and they are ordered in a universal functional sequence.

Note that the feature [B] is not "accusative." Accusative is the name of a constituent which arises as the result of merging [A] and [B] on the top of the DP in this order. To make this clear, I avoid calling the terminals "accusative" but reserve that label only for the non-terminal projections. The labels of the non-terminal nodes, such as *acc* or *gen* are chosen for clarity of presentation, and they do not imply that the label is qualitatively different from the head. I assume that the "true" label of the accusative constituent is BP, but I avoid referring to it in that way because such a label is quite opaque.

Finally, in order to terminologically distinguish the representations (14) and (28) (both are instances of cumulative decomposition), I call (14) flat cumulative decomposition, and (28) layered cumulative decomposition.

6.2 The content of the features

In the preceding sections, I have stressed that what is at stake here are the formal properties of the decomposition, rather than the actual content of the features. However, for the sake of completeness, I add a brief speculation as to what the content might be.

I will be assuming a view on feature meaning such that each feature operates on the semantics of its complement. In the prototypical case, each feature restricts the denotation of its complement to a smaller subset (even though I make use of other types of semantic contribution as well). Under this view, nominative will be the least specific case semantically, with other cases getting increasingly more specific. With this general background, I would like to distinguish four "zones" in the proposed functional sequence:

(29) a. Zone 1: NOM-ACC (structural cases)
b. Zone 2: LOC-GEN (stative cases)
c. Zone 3: DAT (goal case)
d. Zone 4: ABL-INS (source cases)

As indicated, I understand NOM-ACC as pure structural cases, compatible with a wide range of interpretations. To keep things as simple as possible, I assume that the interpretation of the feature [A] is "case." The feature [B] is then perhaps best understood as "dependent," reflecting the hypothesis that only NOM may surface as the so-called "default case."[16] A potential problem for this simple view is that ACC in Classical Armenian occurs in goal PPs. To make this fact compatible with the "structural" approach to NOM-ACC, I follow Abraham (2003) in his suggestion that the ACC in goal PPs is actually verb-governed.

The second zone comprises two "stative" cases. LOC denotes static location, and GEN a static possession (broadly understood). The reason why these two

[16] Schütze (2001) argues that English has a default accusative case, which in English is manifested only on pronouns. However, all the examples that are used to establish this claim (coordination, left dislocation, gapping, short replies) are contexts that exclude weak pronouns and require strong pronouns (see Cardinaletti and Starke 1999). This makes it more likely that the data rather suggest that *he/she* are weak nominative pronouns, whereas *him* a strong pronoun, ambiguous for NOM/ACC.
An anonymous reviewer reminds me that *he/she* are available in (some) coordination structures (whereas weak pronouns generally are not, although some counterexamples have been reported) — I leave this for future research. The reviewer also notes that *he/she* can be contrasted and stressed; but these facts are compatible with being a weak pronoun (Cardinaletti and Starke 1999, p.153-4).

should be the simplest of the semantic cases is connected to a consensus in the literature that states are basic, and that change of state meanings are derived from them (see, among many others, Kratzer 2000 and Ramchand 2008 for verbal semantics and Jackendoff 1983 for spatial expressions).

The idea for "dative" is then that the feature "E" in (28) adds a change of state meaning to the two stative cases, leading to a general "goal" reading. This subsumes the prototypical dative use as a recipient case (i.e., change of possession case); see Caha (2009, ch.5) for a defense of such an idea.[17]

Finally, ABL and INS are understood as "source" cases. The lower ABL covers not only spatial "source of motion" reading, but also a more abstract "source of event" reading, corresponding to the passive agent. INS may then be understood along similar lines (i.e., as a "source" of the event), the distinction between agent and instrument proper being both animacy and the "controled-by-agent" restriction on instruments.

The reason why the source zone should be derived by the addition of features on top of the "goal" zone is not immediately obvious, but finds an analogue in work by Pantcheva (2010, 2011). On the basis of a cross-linguistic sample, Pantcheva discovered that if a language shows morphological containment between spatial source and goal, then source always contains goal. She explains that by proposing that source paths are constructed by reversing the orientation of goal paths. If that is so, her proposal explains why the source zone is more complex than the goal zone, with the feature "F" plausibly corresponding to the type of reversative semantics she proposes.

6.3 Case attraction as evidence for a layered structure

Leaving the content of the features aside, the general interest of the proposal is twofold. Theoretically, the proposal entails that the organization of features inside morphemes is governed by the same principles as the organization of phrases inside sentences (i.e., a binary branching tree). The descriptive restriction on syncretism, namely Case Contiguity, is then just a reflex of this deeper hypothesis concerning the architecture of grammar (see Kayne 2005, Starke 2009, Cinque and Rizzi 2010).

In addition, we gain a number of empirical predictions. For instance, Caha (2011) shows that both across and within languages, movement may target various positions in between such case features. This can only be so if (28) (as opposed to (14)) is the correct representation.[18]

In this section, I provide Armenian-internal evidence from case attraction that favors the layered structure (28) over the flat representations in (14).

[17] In a number of languages, the goal zone also comprises the allative (DAT=ALL is frequent). But recall that the goal case in Classical Armenian is ACC. The suggestion is that the ACC is verb governed, and Armenian directionals thus correspond to applicatives.

[18] In languages where the noun lands in between the case features, the features are split into two groups and get realized as two separate morphemes (one preceding, one following the noun). Hence, such languages also provide evidence for the proposed containment relations; see Caha (2011) for a number of illustrations for each containment proposed.

Specifically, as noted in section §3.2, case attraction in Armenian is related to the system of syncretism in an intriguing way. I repeat the observation below:

(30) *Case Attraction (Armenian)*:

a. In the 'syncretism sequence' below, cases to the right of the GEN have the power to attract it, case to its the left cannot do so.

b. NOM – ACC – LOC – GEN – DAT – ABL – INS

In order to show how case attraction forces us to abandon flat structures, I first present an analysis of this phenomenon as the combination of (i) agreement and (ii) subsequent ellipsis. I portray this schematically in (31):

(31) Case attraction as agreement plus ellipsis
head noun-case$_i$ [[dependent noun-~~GEN~~] AGR=case$_i$]

In words: GEN marking is elided because of the presence of a structurally higher agreement marker, which bears identical case features as the head noun.[19]

The first reason for adopting (31) is that an apparently bizarre phenomenon is decomposed into two well-known and independently attested processes. To see agreement marking of possessors apply independently, consider the example in (32). It comes from Guugu Jalanji, and shows that the genitive possessor 'Dick' is marked by an AGR marker (*-du*) tracking the case of the head ('dog'):

(32) Dicki-ndamun-du kaya-ngka
Dick-GEN-ERG dog-ERG
'Dick's dog' (Guugu Jalanji, Plank 1995)

Turning to Armenian itself, it can be shown that case attraction shares two characteristics with such a process. First of all, attraction is local: it does not apply recursively to the possessor of the attracted phrase. To see that, consider (13) again (repeated below for convenience).

(33) [bazmut'-eamb [zawr-**awk'**-n Hay-oc']]
crowd-INS.SG force-INS.PL-DEF Armenian-GEN.PL
'with a crowd of the Armenian forces' (Plank 1995, p.43)

In the example, the attracted possessor (*zawr-awk'-n*) has itself a nominal dependent (*Hay-oc'*), which escapes attraction.

The same restriction has been observed for possessor agreement in Lardil by Richards (2007). The relevant data are given below, the ACC *-i* spreads from the head noun ('boomerang') to the head of its GEN dependent ('older brother'), but not on the dependent of the dependent ('boy'):

(34) [marun-ngan thabuji-kan-i] wangalk-i]]
boy-GEN older.brother-GEN-ACC boomerang-ACC
'(You broke) the boy's older brother's boomerang.'

[19] I assume that agreement involves a separate base-generation of case features on top of the agreeing constituent, GEN in our case. At LF, these features must match with the case features base-generated on the head noun.

Secondly, attraction appears to be optional. This is shown in (35), where both non-attraction (35-a) and attraction (35-b) are available:

(35) a. i knoǰ-ê t'agawor-I-n
by wife-ABL king-GEN.SG-DEF
b. i knoǰ-ê t'agawor-**Ê**-n
by wife-ABL king-ABL.SG-DEF
both: 'by the wife of the king' (Plank 1995, p.20)

This may now be seen in parallel to the fact that agreement is likewise 'optional.' In Old Georgian, for example, the possessor does not need to agree with the head noun, as in (36-a), but it may do so, as shown in (36-b):[20]

(36) a. saxl-is paṭron-ma
house-GEN mistress-ERG
'the mistress of the house' (Plank 1995, p.7)
b. gwam-isa krist-es-isa
body-GEN Christ-GEN-GEN
'of the body of Christ' (Plank 1995, p.4)

Summing up, there are at least two properties that case attraction shares with clear cases of possessor agreement (locality and optionality); this follows if case attraction involves agreement as a crucial component.

Adopting such a view brings us to the question why the genitive marker does not co-occur with the agreement marker (as it does in (34) and (36)). There are two logical options: (i) the genitive has never been there, or (ii) it has been elided on the surface. I argue here for the latter option.

Let me begin by showing that such an ellipsis process is independently needed. An example is given in (37). To set the stage, (37-a) shows that in Amharic, the possessor is preceded by the possessive marker *yä*.[21] This marker gets deleted when the whole DP is preceded by the dative marker *lä* in (37-b). (It gets deleted after other oblique markers as well.)

(37) a. yä-Gɨrma-n wändɨmm
of-Girma-ACC brother
'Girma's brother-ACC'
b. lä-~~yä~~-Gɨrma wändɨmm
to-Girma brother
'to Girma's brother' (Amharic, Baker and Kramer 2010)

[20] In Old Georgian, agreement is in fact not optional, but depends on the position of the dependent noun. Simplifying things slightly, in postnominal position, agreement applies; in pre-nominal position, it does not. I assume that in Classical Armenian, the GEN noun also has access to two distinct positions, one where agreement is obligatory and one where it is disallowed. The existence of such two distinct positions is masked by the fact that the noun moves even higher up.

[21] The example also has an accusative marker -*n*, which marks the whole DP, and surfaces following the first constituent of the DP.

Note that in Amharic, ellipsis is case sensitive: it does not apply in (37-a), i.e., in ACC (it does not apply in NOM either). The reason Armenian should be analyzed along the same lines comes from the fact that attraction is also case sensitive (30). The explanation for this emanates from an independent condition on ellipsis, which I state in (38):

(38) *Recoverability of ellipsis (case attraction):*
The feature content of the elided morpheme (genitive) has to be recoverable from the antecedent (the agreement marker).

The following table spells out the solution in detail. The first column reflects the proposal that both the genitive case and the agreement marker are underlyingly present. Combining this proposal with cumulative case decomposition, we arrive at the representations in the middle column. Ellipsis then applies to the genitive, and produces the output in the last column:

(39) An ellipsis account of the facts

	N-GEN-AGR	Ellipsis	Output
a.	N-GEN-NOM	N-~~[A,B,C,D]~~-[A]	*
b.	N-GEN-ACC	N-~~[A,B,C,D]~~-[A,B]	*
c.	N-GEN-LOC	N-~~[A,B,C,D]~~-[A,B,C]	*
d.	N-GEN-GEN	N-~~[A,B,C,D]~~-[A,B,C,D]	N-GEN
e.	N-GEN-DAT	N-~~[A,B,C,D]~~-[A,B,C,D,E]	N-DAT
f.	N-GEN-ABL	N-~~[A,B,C,D]~~-[A,B,C,D,E,F]	N-ABL
g.	N-GEN-INS	N-~~[A,B,C,D]~~-[A,B,C,D,E,F,G]	N-INS

As is evident, the condition (38) derives the observation that attraction does not happen in NOM, ACC, LOC (the cases above the double line). That is because in these cases, the proposed ellipsis cannot be recovered: the feature D finds no antecedent in any of the cases.[22]

This approach thus correctly restricts the application of case attraction in non-trivial ways, accounting for its locality, optionality, and case sensitivity. Importantly, for the account to work, we need the initial proposal that case features are cumulative: GEN must be contained in all cases to its right, and in no case to its left.

Further, the account leads to the conclusion that case features are layered. That is because of an independently established conjecture that the antecedent of an ellipsis must be a constituent (for textbook treatments see, e.g., Haegeman 1994 and Carnie 2008; see also Pesetsky 1995, Phillips 2003 and Baltin 2006 for relevant discussion). If that is so, then the features of GEN must not only be contained inside those cases which license its ellipsis, but they must form a constituent to the exclusion of other features. This in turn forces us to

[22] I assume that in NOM, ACC, LOC, Classical Armenian is thus forced to make use of the possessor position where agreement does not apply.

abandon the initial non-layered decomposition (14) as insufficient. In contrast, the facts are immediately compatible with (28).[23]

7 Phrasal spell-out

As a starting point, let me repeat that we have taken a step beyond cumulative decomposition and added the proposal that each case feature corresponds to a separate head in the tree (cf. Kayne 2005, Starke 2009, Cinque and Rizzi 2010). Theoretically, this explains the formal properties of cumulative decomposition; empirically, we find supporting evidence in phenomena such as case attraction.

But the model also raises questions concerning spell-out. In particular, since case morphemes spell out more than one feature (§4.2), and each feature is a head (§6.1), the model virtually necessitates a theory where spell-out may target more than one head. In this section, I present a phrasal spell-out algorithm with such an effect, drawing mainly on Starke (2009).[24]

The algorithm I describe takes the syntactic structure as its input, and provides the phonological representation as its output. This simplifies the architecture of grammar in the sense relevant for this paper: there is no intermediate morphological structure. In later sections, I show that such a theory is not only simpler concerning its architecture; it also captures subtle details of paradigm structure.

7.1 Three core assumptions

The proposal to be developed understands spell-out to be a cyclic bottom-up translation of syntactic structure onto phonological representation. At each node, a PF is constructed, and stored in memory until the next cycle. The intermediate PF representations do not necessarily correspond to a part of what is pronounced in the end. Intermediate representations may be scraped off, and replaced by a completely different representation in case the algorithm requires that.[25]

Coming to the actual algorithm, the first assumption I make is that it scans the structure bottom up, with lexicalization attempted at every node.

[23] Admittedly, the facts do not provide evidence for such a rich structure as (28) posits (we only have evidence for GEN being a constituent inside DAT/ABL/INS). But the point is that whenever we can find evidence for layering, we do.

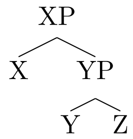
[24] Cf. McCawley (1968), Weerman and Evers-Vermeul (2002), Neeleman and Szendrői (2007), Ramchand (2008), Radkevich (2009), Taraldsen (2010), Pantcheva (2010).

[25] These stages thus do not correspond to unchangeable phases, a notion that may be close to what Chomsky (2008, 143) has in mind: "For minimal computation, as soon as the information is transferred it will be forgotten, not accessed in subsequent stages of derivation [...] Working that out, we try to formulate a *phase-impenetrability condition* PIC [...] PIC holds [...] for the mappings to the interface."
The view assumed here comes closer to the one described in Šurkalović (2011). Šurkalović proposes that "spell-out does not proceed in chunks but in concentric circles" (p.97), with the consequence that "the input to phonology at each phase is cumulative, consisting of the spell-out of the current phase together with the spell-out of the previous phases." (p.84)

(40) *Assumption 1*: Spell-out proceeds bottom up.

Thus, in (41), spell out targets Y and Z before YP, and it targets X and YP before XP.[26] It is not crucial whether Y is spelled before Z, or X before YP.

(41)



In Starke (2009), bottom-up spell-out is part and parcel of the cyclic spell-out hypothesis (cf. Starke 2010, Caha 2011, Pantcheva 2011). In particular, spell-out applies after each step of Merge. And since Merge is standardly assumed to proceed bottom up (Chomsky 1995), so does spell-out. In this paper, I simplify the model and ignore the interleaving of Merge and lexical access for ease of presentation: I use a model where syntax produces a full-fledged output representation, which is then step-wise translated onto phonological form.

Thus, for each tree derived by syntax, we go through the structure bottom up, and construct a PF at every node. The PF consists of material eligible for insertion at a given node. The principle which determines what entries are applicable is given below; it is a refined version of the Superset Principle (15).

(42) *Assumption 2*: The Superset Principle (cf. Starke 2009):
A lexically stored tree matches a syntactic node iff
   a. the lexically stored tree contains the syntactic node including the material dominated by that node OR
   b. the lexically stored tree matches all daughter nodes (ignoring traces)

Sometimes, an item matches both because of (42-a,b), but it is enough that only one condition obtains. Because of that, I save space and mention only one reason for obtaining match in such cases. When an item matches a node, it represents the whole tree dominated by that node at PF, and overrides any material that may have been inserted earlier on.

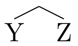
I further assume that a successful lexical search sends to PF a set of items that match the node. (If the set has more than one member, competition arises, with the winner determined by the Elsewhere Condition.)

To see the workings of the system in an example, assume that in addition to the structure (41), we have the lexical entries in (43):[27]

[26] See Bobaljik (2002) for a related proposal.

[27] In general, I assume that lexical entries are triplets <phonology, syntax, concept>. Since the conceptual contribution of case affixes is negligible (if present), I ignore it here.
Further, I assume that lexical entries are links between phonological, syntactic and conceptual representations, each representation a well-formed object in the relevant module. This answers some frequent questions about why there should be syntactic trees 'in the lexicon' (see (43-b)), and related worries about the role of syntax and the lexicon in constructing expressions. From the linking conception, it follows that the lexicon has no means to create structures of its own. Consequently, it follows that the only possible target of lexicalization

(43) a. Z ⇔ /α/
b. YP ⇔ /β/

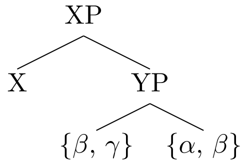


c. Y ⇔ /γ/
d. X ⇔ /δ/

How is the structure (41) spelled out using the lexical entries (43)? We start from the node Z, and determine the set of entries applicable for that node. This set comprises of (43-a,b), since each contains the node Z. Lexical search thus produces an intermediate stage of PF derivation: the set {α, β}.

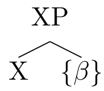
As a next step, spell-out targets Y. Y is contained in the entries (43-b,c), and hence, we obtain the set {β, γ}. At this point, we have the PF representation in (44-a) (two unordered sets). For reasons of clarity, I will be using the representation in (44-b) to represent such a translation. The advantage of (44-b) is that it shows how the representation (44-a) was derived.

(44) a. {β, γ}, {α, β}

b.


Spell-out then targets YP. Since the entry for {β} matches both daughters of YP (see (44-b)), YP is also matched by {β}, yielding the representation (45-a), which may be by convention represented as in (45-b):

(45) a. {β}

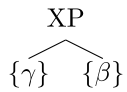
b.


This shows that when lexical search produces an item that matches a phrasal node, the whole tree is represented by this item at PF, and all material previously inserted below that node is deleted from the PF representation. Consequently, in the course of spell-out, β won as the spell-out of YP over a combination of two items, say the sequence $\gamma{>}\alpha$. Starke (2009) calls this effect "The Biggest Wins Theorem," since the entry with the biggest tree in the lexicon wins. Empirically, we get the effect that portmanteau morphemes win in competition over agglutinative forms. To mention just a couple of examples: *off* wins over **from on*; *went* over **go-ed*; in Armenian, as discussed in more detail below, the ABL.PL *c'* wins over ABL+PL *ê-k*.

Spell-out then targets X, with the only match {γ} introduced by (43-d). At this point, we have the two unordered sets in (46-a), which corresponds to the conventional representation (46-b):

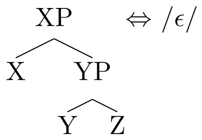
(when it comes to the syntax part of the lexical entry) is a well-formed syntactic structure. Thus, for example, [X, Y, Z] cannot be a part of a lexical entry, because syntax cannot construct ternary branching trees (cf. Kayne 1994, Chomsky 1995). Thus, despite the fact that the lexicon stores outputs of syntax, it does not duplicate it; the lexicon only links the outputs of syntax to representations in other modules.

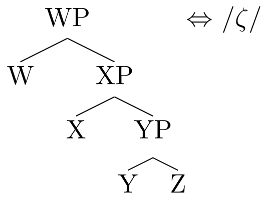
(46) a. $\{\gamma\}$, $\{\beta\}$

b. 

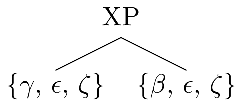
There is no match for the whole XP; hence, there is no 'overriding.' Rather, the translation procedure at XP introduces the mutual ordering of $\{\gamma\}$, $\{\beta\}$. Since the feature spelled out by $\{\gamma\}$ c-commands those spelled out by $\{\beta\}$, the phonological representation orders the sets of (46-a) as $\gamma{>}\beta$. Rewriting of previously inserted material and ordering of sets are nicely visible in the tree representations, and I am going to use such representations exclusively.

My final assumption concerns situations where the spell-out algorithm does not provide a unique candidate for spell-out. Thus, coming back to our initial example in (41), suppose that we have two additional entries:

(47) a. 

b. 

The change this brings into the picture is the following. Since each of the terminals in (41) is contained in the entries in (47), we get two more competitors for each node. Consequently, we end up with (48), instead of (46-b):

(48) 

When spell-out targets XP, we obtain the set $\{\epsilon, \zeta\}$ (each matches both daughter nodes). The unique spell out of XP is thus not determined by the principles we have introduced so far. What we need is the Elsewhere Condition:

(49) *Assumption 3: The Elsewhere Condition*: In case two rules, R_1 and R_2, can apply in an environment E, R_1 takes precedence over R_2 if it applies in a proper subset of environments compared to R_2.

In our case, the rule introducing ϵ applies in a proper subset of cases compared to the rule introducing ζ (ζ can apply wherever ϵ does, but ϵ is inapplicable for W, and WP). Hence, ϵ wins in competition for the spell-out of XP. Importantly, I assume that Elsewhere computation applies only where indispensable, i.e., in cases where the 'Biggest Wins Theorem' has not produced a unique candidate.

Recall that the Elsewhere Condition and the Superset Principle were introduced in §4.2, and they are necessary to derive Contiguity. Incorporating these assumptions replicates the result in a phrasal spell-out system.

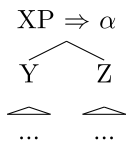
7.2 Two consequences

Let me now come to two consequences of the proposed model that will turn out to be important in accounting for various interesting phenomena in Classical Armenian.

(50) Consequences
  a. if a phrase XP is spelled out by a single morpheme in a language L, any sub-structure of XP is spelled out by a single morpheme as well (in that language)
  b. insertion may only target features that form a constituent

To see the consequence (50-a), consider the representation in (51). Suppose that in this tree, XP is spelled out by α:

(51)



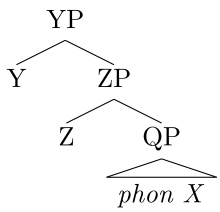
In order for that to be possible, either (52-a) or (52-b) has to hold, because of the Superset Principle, see (42):

(52) a. the lexical entry for α contains the whole tree dominated by XP, and hence, it contains also (the tree dominated by) Y and Z
  b. the lexical entry for α matches both Y and Z

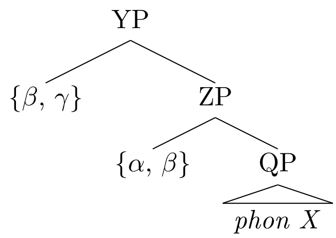
From (52), it is easy to see that for every XP, the item which matches that XP matches also its immediate sub-constituents (Y and Z), and will be used to spell out these nodes (unless there is a better match due to the Elsewhere Condition). We may now repeat this reasoning recursively for Y and Z: since α matches these nodes, it also matches their immediate sub-constituents, and so on, until we reach the terminal nodes. In conclusion, if XP may be spelled out by a single morpheme α in a particular language, then any substructure will correspond to a single morpheme, because α is always a match. This holds regardless of whether α actually surfaces (there may be a better match).

To see the corollary (50-b), suppose that we modify the example structure (41) by introducing a complement to Z, pronounced as *phon X*:

(53)



(54)



In (53), Z may be replaced by the set $\{\alpha, \beta\}$, and Y by $\{\beta, \gamma\}$, using the entries in (43). The result is shown in (54). Importantly, YP in (54) cannot be replaced by $\{\beta\}$, despite the fact that $\{\beta\}$ matches both Y and Z. That is because in YP, $\{\beta\}$ matches only one daughter (Y), but not ZP. This in turn is caused by a combination of two factors. First, the entry for β (see (43-b)) does not contain a tree identical to the ZP of (54), and second, not all daughters of ZP are matched by $\{\beta\}$ (QP is not). Hence, even though Z and Y are both contained in a single entry, they cannot be spelled out by that entry if they do not form a constituent to the exclusion of other features. I turn to this prediction in §10, and show that it explains the puzzling phenomenon of morpheme splitting. The general conclusion from such facts is that since the constituency of individual features that make up morphemes matters for spell-out, we gain evidence for positing such structures inside individual morphemes.

Finally, I note that if QP had extracted out of YP in (54), then β would match YP once again. The reasoning is this: since β matches Z, it also matches ZP where Z is the only non-trace daughter (recall that QP extracts). Further, since β now matches both ZP and Y, it also matches YP. This is once again related to the statement that Y and Z can be spelled out as one morpheme just in case they form a constituent. After QP extracts, Y and Z form a constituent to the exclusion of other features, and hence, they may be lexicalized together.

The fact that insertion ignores traces has a consequence that will become relevant. In particular, insertion at a node that dominates a trace produces the same output as insertion at its non-trace daughter. Recall the example: when ZP immediately dominates a trace, insertion at ZP produces the same results as insertion at Z. Hence, whether ZP is present in the tree has no effect at PF. I use this corollary (55) later on to simplify complex representations when focusing on their spell out.

(55) *The tree simplification convention:* nodes that dominate a trace may be skipped in spell-out representations.

In the remainder of the paper, I turn to the consequences in (50), and show their validity on empirical material.

8 The hierarchy of agglutination and fusion

I start with the prediction (50-a): if an XP is expressed by a single morpheme, any sub-structure corresponds to a single morpheme as well. By combining this prediction with cumulative decomposition, we obtain an interesting testing ground. In particular, since one case represents a substructure of another case, we predict that if a case X is expressed together with some other inflectional category, then all cases contained in X also fuse that category with case.

This prediction is nicely borne out in Armenian, as I argue below. Specifically, I focus on the plural declension and argue that the INS.PL has separate markers for class, plural and case. These categories are, however, subject to fusional expression in other cases. Specifically, NOM, ACC and LOC show only a single marker, while GEN, DAT and ABL are characterized by two markers: a class marker, and a number/case portmanteau. I summarize these claims in the table (56), where shading indicates synthetic expression of given categories:

(56) *The template for Armenian plural declension*

	STEM	CLASS	CASE	NUMBER
NOM, ACC, LOC	stem	-class&case&number		
GEN, DAT, ABL	stem	-class	-case&number	
INS	stem	-class	-case	-number

The interest of (56) lies in the observation that the same sequence which restricts syncretism and case attraction governs also the agglutinative vs. fusional expression of categories. I state this in (57):

(57) *Agglutinative vs. fusional spell-out (Armenian)*:

a. In the sequence below, if a given case is expressed together with some other category, then all cases to its left are as well.

b. NOM − ACC − LOC − GEN − DAT − ABL − INS

The current section establishes the empirical generalization (57), and the next section derives it from the three spell out assumptions.

8.1 Case and Number

I start by taking a closer look at the singular/plural distinction. In most cases, the singular and plural endings are different. This can be understood if (most) case exponents also spell out number. The exception to this is the INS.PL, which is built on top of the INS.SG by the affixation of *-k'*, a morpheme which also shows up in the NOM.PL.

(58) *The analytic nature of the instrumental plural*

	nation, SG.	nation, PL.
NOM	azg-ø	azg-**k'**
ACC	azg-ø	azg-s
LOC	azg-i	azg-s
GEN	azg-i	azg-ac'
DAT	azg-i	azg-ac'
ABL	azg-ê	azg-ac'
INS	azg-AW	azg-AW-**k'**

A reasonable hypothesis is that *-k'* marks plural. This finds confirmation in the verbal paradigm, where *-k'* differentiates the 1.SG and 1.PL agreement (59-a), and in the composition of the 2nd person pronouns (59-b):

(59) a. sirem – sirem-k'
love.1.SG – love.1-PL

b. du – du-k‘
you.SG – you-PL

Turning to NOM.PL, a question arises whether it is expressed by one or two morphemes. Under one possible hypothesis, *-k‘* is just plural, and NOM is *-ø*. I call this the agglutinative analysis. An alternative hypothesis (which I call the fusional analysis) is that *-k‘* spells out both number and case.[28] These analyses are hard to tease apart, but there is indirect evidence in favor of fusion.

First of all, the agglutinative analysis requires an additional (null) morpheme. How serious a drawback that is depends on the details of implementation; but as a general methodological guideline, we should favor an analysis that does not postulate an additional (invisible) entity, if it can be dispensed with.

Secondly, if *-k‘* is just plural, we end up with a kind of split between agglutination (NOM, INS) and fusion (other) that does not seem to be attested anywhere else (basing myself on the survey in Plank 1999 and other languages I have looked at). Specifically, in languages with such splits, a particular marking strategy always targets contiguous regions on a case hierarchy (usually grammatical vs. oblique). On the agglutinative analysis, Classical Armenian would be the only counterexample to this generalization. And since this unlikely typological feature is an artifact of an analysis, it seems that an alternative analysis (which lacks this particular consequence) is preferred.

Third, it is not only the case that a particular marking strategy (agglutination vs. fusion) occupies a contiguous region on a case hierarchy, there is a virtually unexceptional bias towards fusional expression in the NOM (see, once again, Plank 1999 for an overview).[29] As an example, consider the Mordvin definite declension, discussed in McFadden (2004).

(60) *Mordvin Definite Declension (Erźa dialect)*

Case	‘the house,’ SG.	‘the house,’ PL.
NOM	kudo-ŚŚ	kudo-TŃE
ACC/GEN	kudo-ŃT	kudo-TŃE-ń
DAT	kudo-ŃT-eń	kudo-TŃE-ń-eń
ABL	kudo-do-ŃT	kudo-TŃE-de
INE	kudo-so-ŃT	kudo-TŃE-se

What we see here is the stem *kudo-* ‘house’ inflected for number, case and definiteness. Almost all cases split case from definitness/number morphemes (these are in small caps). The exception to this is NOM.SG, where case is expressed together with these categories as the portmanteau *-ś*.[30]

[28] This analysis presupposes the Superset Principle; when *-k‘* acts as a pure plural marker, it spells out a sub-constituent of its specification.

[29] Plank (1999) finds this pattern in Brahui, Finnish, Karelian, Chukchi, Yawelmani, Tunica, Romani, Mordvin, Estonian, Tocharian A and Wakhi. I add the Prizren-Timok dialect of Serbian (Caha 2011).

[30] I have to mention that Plank (1999) identifies an additional minority pattern where – as Plank argues – there is agglutination in the grammatical cases, and fusion in the semantic

Going back to Armenian: assuming the fusional analysis of NOM.PL, case and number are expressed separately in the most marked case (INS), and fuse in other cases:

(61) *Case and Number*

	STEM	CASE	NUMBER
OTHER	stem	-case&number	
INS	stem	-case	-number

8.2 Class markers

Now compare the plural paradigm we have looked at with other plural paradigms:

(62) *The plural declension in Classical Armenian*

	nation, PL.	river, PL.	word, PL.	time, PL.
NOM	azg-k‘	get-k‘	bay-k‘	žam-k‘
ACC	azg-s	get-s	bay-s	žam-s
LOC	azg-s	get-s	bay-s	žam-s
DAT	azg-a-c‘	get-o-c‘	bay-i-c‘	žam-u-c‘
GEN	azg-a-c‘	get-o-c‘	bay-i-c‘	žam-u-c‘
ABL	azg-a-c‘	get-o-c‘	bay-i-c‘	žam-u-c‘
INS	azg-a-w-k‘	get-o-v-k‘	bay-i-w-k‘	žam-u-ø-k‘[31]

The comparison reveals a separate vocalic element between the root and the morphemes *-c‘* and *-w*. This suggests that we are looking at a morpheme, but what is this morpheme? Starting from the observation that its quality is determined by the root, whereas the quality of the plural *-k‘* is not, it is attractive to analyze this morpheme as originating locally to the root. As for its identity, Halle and Vaux (1998) take it to be a theme marker, a type of a nominal classifier. I adopt this approach here as well.

But why is the class marker absent in NOM, ACC and LOC? I suggest that in these cases, class is expressed together with number and case. In other words, *-s* and *-k‘* spell out also class. This analysis is depicted in (63) in abstract terms, and a breakdown of concrete paradigms is given in (64):

cases. This pattern represents a counterexample to the theory presented here; but there are reasons to doubt its existence.
According to Plank, this pattern arises in Chechen, Archi, Georgian and (interestingly) Classical Armenian. Including Classical Armenian on the list of languages that exhibit agglutination in the grammatical cases (specifically NOM) makes it obvious that Plank's conclusions rest on the agglutinative analysis (NOM=*-ø+k‘*). Once scrutinized, it turns out that 3 out of the 4 languages (Chechen, Archi, and Classical Armenian) are listed as agglutinating on the basis of an analysis that assumes a zero morpheme. This is a rather weak evidence.
Thus, what remains from the minority pattern under scrutiny is a single example: the so-called archaic plurals in Modern Georgian. I have to leave this example for future research.

[31] This form apparently lacks the instrumental *-w/v*. That is due to phonology: the class marker *u* and the instrumental *-w* fuse into one segment. I argue for this later in this article.

(63) *The template for Armenian declension*

	STEM	CLASS	CASE	NUMBER
NOM, ACC, LOC	stem	-class&case&number		
GEN, DAT, ABL	stem	-class	-case&number	
INS	stem	-class	-case	-number

(64) *Case, Number and Class*

	STEM	CLASS	CASE	NUMBER
NOM	stem		-k‘	
ACC, LOC	stem		-s	
GEN, DAT, ABL	stem	-a-/-o-/-i-/-u-		-c‘
INS	stem	-a-/-o-/-i-/-u-	-w-	-k

To conclude: if the morphological analysis depicted in the table above is correct, then the generalization (65) is established:

(65) *Agglutinative vs. fusional spell-out (Armenian)*:
a. In the sequence below, if a given case is expressed together with some other category, then all cases to its left are as well.
b. NOM − ACC − LOC − GEN − DAT − ABL − INS

The next section derives the generalization from the spell out algorithm introduced in §7.1 operating on the proposed functional sequence of cases.

9 Deriving the splits between agglutination and fusion

In §7.1, I introduced an algorithm that maps syntactic constituents onto their pronunciation. Thus, when number and case are expressed as one morpheme, this is encoded by spelling out a constituent containing these two categories. Knowing the constituent structure is thus essential for carrying out the analysis; and I turn to this issue now.

9.1 The order of morphemes

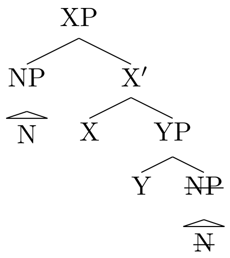
According to the view presented here, morphemes are either heads or phrases. If that is so, we have two tools to take into consideration when we derive morpheme orders: both phrasal and head movements may be involved in the derivation of the Armenian inflected nominal.[32]

One way in which phrasal movement differs from the traditional head movement (Travis 1984, Baker 1988) is that the head of the extended projection, the noun in our case, can move across two heads without inverting their order, as in (66-a), in violation of the Head Movement Constraint. A derivation which

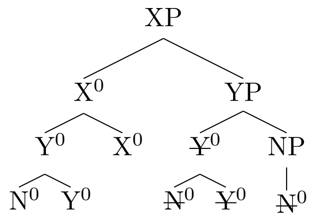
[32] In the ideal case, head-movement is treated as a special instance of phrasal movement. The analysis presented here does not allow such a reduction. The issue how to replace the type of head-movement I employ in this paper is subject to ongoing research in Nanosyntax.

obeys the Head Movement Constraint is shown in (66-b), and the movement of N across X inevitably leads to the inversion of X and Y (excorporation aside).

(66) a. Phrasal movement:



b. Head movement:



In Classical Armenian, a theory based solely on head movement leads to a wrong prediction. To see that, consider the base-generated order of the markers in question. I follow the literature and adopt the hierarchy in (67), where the NP is dominated by the projections of the Classifier (Cl), Number (Num, see Borer 2005), and finally case (K, see Bittner and Hale 1996).[33]

(67) [K [Num [Cl [N]]]]

We know that the Noun has to move higher than K, since K is a suffix; head movement then yields the sequence (68), which is empirically wrong: Num follows K in Classical Armenian (recall *-a-w-k'* CL-INS-PL).

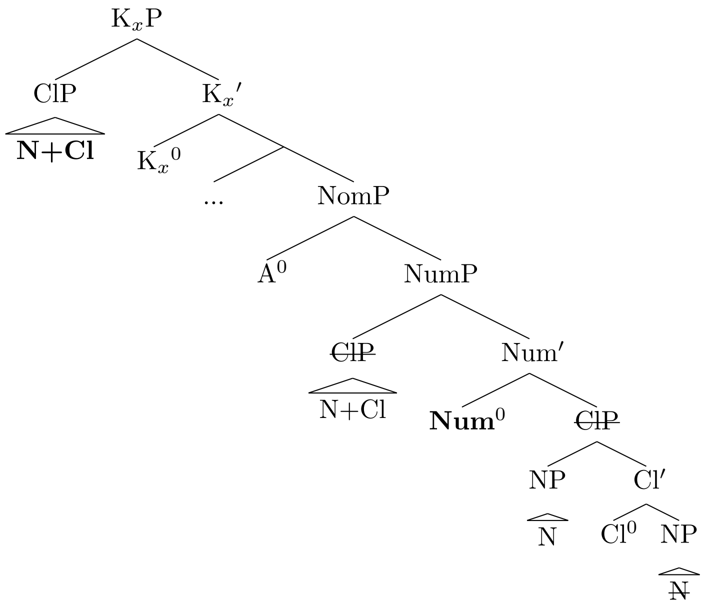
(68) *N-Cl-Num-K

What we need is a derivation involving phrasal movements (see also Koopman and Szabolcsi 2000, Julien 2007, Muriungi 2008). The simplest derivation (granted the general approach originating in Kayne 1994) is shown in (69).

---

[33] A reviewer suggests that arbitrary thematic vowels are hard to account for 'in syntax,' and require the existence of a specific morphology module — contrary to the architectural claims made here. I do not agree with this conclusion for reasons I sketch below.
In particular, the selection between a particular root and an adequate thematic vowel can be treated as idiomatic (in a technical sense). Space considerations prevent me from developing a full fledged account of idioms, and I give only its bare bones. In Nanosyntax (cf. Starke 2010), idioms like *kick the bucket* are treated as phrasal lexical entries that make reference to other entries. Thus, when a bottom-up translation derives a constituent with its constituent parts spelled out as *kick the bucket* (as opposed to *kick the pot*), that triggers the insertion of the concept DIE. Similarly, the Armenian lexicon contains a number of idioms that refer to a constituent consisting of the root and a particular thematic vowel (cf. Marantz 1995). Noting that there are idioms that contain words which do not make sense out of that idiom, roots on their own may or may not have interpretation, depending on whether any conceptual information is associated to them outside of the particular idiom. Assuming an account along these lines, the pairing of roots and thematic vowels requires a specific morphological component no more than phrasal idioms do.

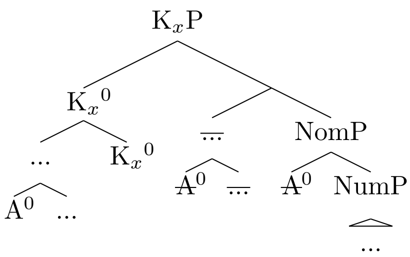
(69)



First NP and Cl invert. Then we add Num, which is crossed by its complement, leading to N-Cl-Num. Num must end up last, and hence stays in situ. I put it in bold. After we add K (specifically, the relevant number of heads), only ClP moves across K, leading to N-Cl-K-Num. This is the correct order.[34]

I assume two additional movements that do not lead to surface re-ordering, but change the constituency. First, I assume that case features undergo head-movement from A to the highest K head, see (70). This creates a constituent that contains only case features, but no number feature. This constituent is needed in the INS.PL, where *-w* targets such a constituent.

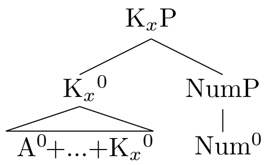
(70)



Combining the movements in (69) and (70) produces graphical clutter. To remove it, I use the tree simplification convention (55) and skip the case nodes whose head evacuates by head movement, yielding (71). Recall that such a simplification has no effect on the actual PF.

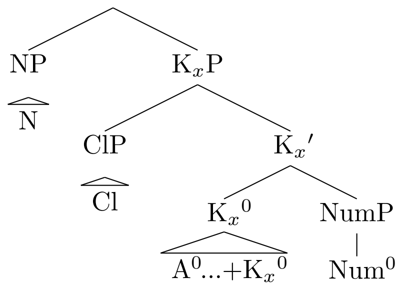
[34] I assume that the NP is head final, and may contain adjectival modifiers. Note as well that pre-nominal adjectives do not carry agreement in Armenian.

(71)



Note further that in (69), the class feature(s) do(es) not form a constituent with case and number to the exclusion of the head noun. Since we know from (64) that class may be spelled out together with case and number, we need this constituent. Thus, I assume that the head NP sub-extracts from the final landing site of the ClP. Disregarding traces, this will leave a series of non-branching projections in the Spec of K_xP. Due to the particular formulation of the Superset Principle I have offered, anything that can spell out the Class head will also spell out ClassP, and hence, I abbreviate the structure into (72). This structure will be used as a general template for insertion.

(72)



9.2 Packaging of categories by means of phrasal spell-out

Let me now turn to the splits between agglutination and fusion, as attested in Armenian. In particular, we need to derive (73) (repeated from (64)). In the table, I abstract over the quality of the theme, and abbreviate it as V.

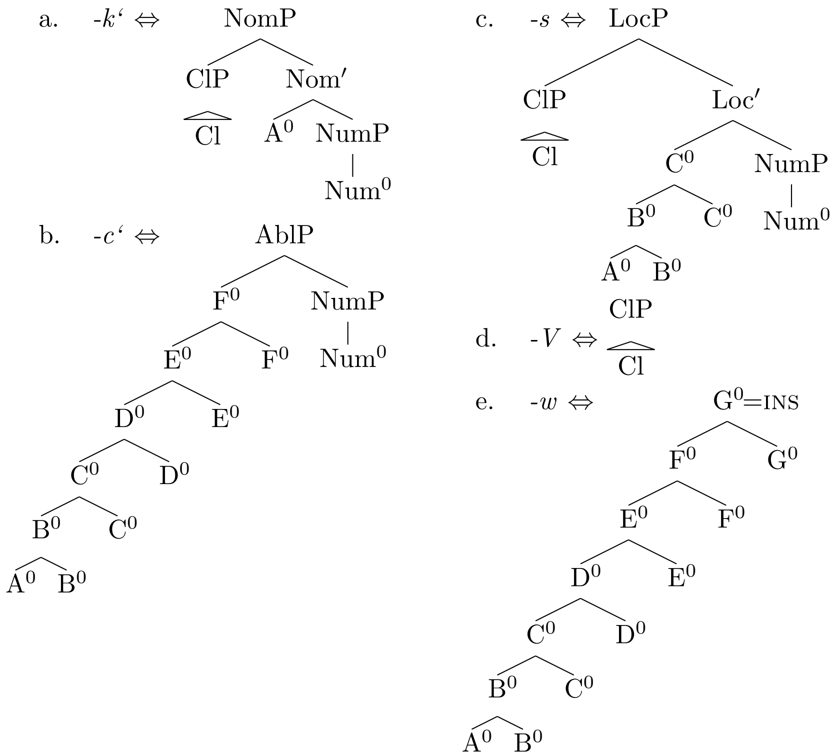
(73) *Case, Number and Class*

	stem	-class	-case	-number
NOM	stem		-k‘	
ACC, LOC	stem		-s	
GEN, DAT, ABL	stem	V		-c‘
INS	stem	V	-w-	-k

In order to show how (73) is derived, I first write down the lexical entries of the relevant markers. Then I go step by step through the structure of the individual cases. Once we see how things work, I turn to some general conclusions.

(74-a,c) give the entries for *-k‘* and *-s*. These entries are able to spell out all of class, case and number, and they differ in the number of case features they possess. Note that in the entry of *-s*, I omit the nodes that dominate traces due to head-movement of A to the highest K head, i.e., C. This is done in order to allow for easier parsing of the structure, using the convention (55).

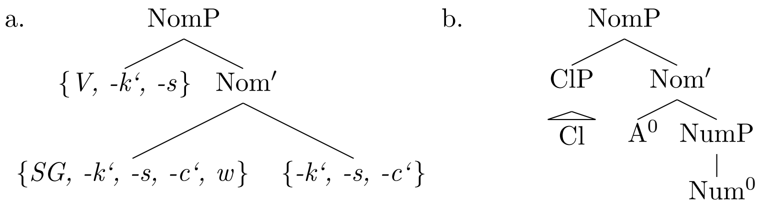
(74)



For GEN/DAT/ABL, we need two additional entries, introduced in (74-b,d). The marker -*c‘* is specified for the case features A-F (representing the ABL), but unspecified for class. In (74-d), I use the generic V for class markers, and ignore their differences (see fn. 33). Finally, -*w* is a pure case marker, specified for the complex INS head (A-G), but no class and number, see (74-e).

With the entries in place, I turn to the spell-out of the NOM.PL. Its structure is in (75-b). Starting from Num^0 (representing plural), we find three matching entries: {-*k‘*, -*s*, -*c‘*}. Since the algorithm ignores traces, the same set matches NumP, because it matches its only daughter. Thus, I replace NumP in (75-b) by the relevant set in (75-a). The order of elements in the set reflects their specificity; in case competition arises, the leftmost element of the set wins.

(75)

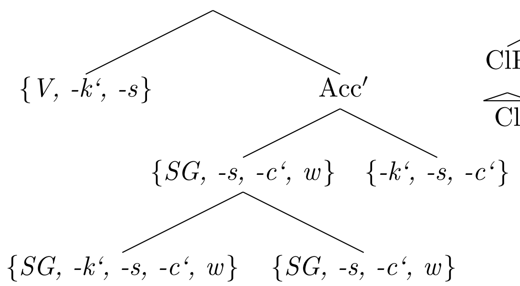


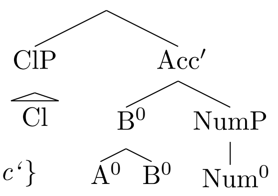
A^0 is matched by {-*k‘*, -*s*, -*c‘*, *w*}, because these entries contain A^0. I also assume that singular case markers in general may spell out A^0, and I add them into the set for completeness, even though – as we will see – they will never get inserted. Hence, I replace A^0 of (75-b) by the set {*SG, -k‘, -s, -c‘, w*} in (75-a), where SG stands for the set of singular case markers.

Turning to the complex class node, we may note that it is contained in all the class vowels, as well as in *-k‘* and *-s*. The set comprising these elements replaces the phrasal class node in (75-b) to yield the last ingredient of (75-a).

Now both daughters of Nom′ in (75-a) are matched by {*-k‘, -s, -c‘*}, hence, this set also matches Nom′. The whole NomP is then matched by {*-k‘, -s*}, since these elements match both daughters. With no more projections to spell out, competition arises with the result that *-k‘* is selected on the basis of the Elsewhere Condition (it applies in proper subset of cases compared to *-s*).

The next case I discuss is the accusative (76-b). The only difference in the spell-out procedure (compared to (75)) concerns the case node. Here, once again, the set {*SG, -k‘, -s, -c‘, w*} matches the A head (replacing it in (76-a)). Proceeding to B^0, the set {*SG, -s, -c‘, w*} applies. As a result, the complex case head can be spelled out by {*SG, -s, -c‘, w*}.

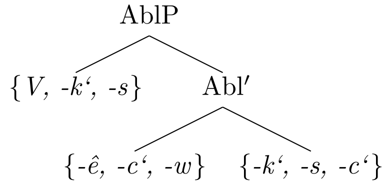
(76) a. 

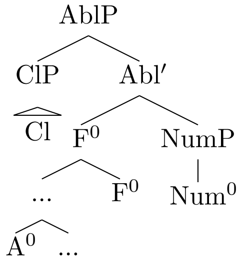
b. 

Proceeding further upwards, {*-s, -c‘*} match Acc′, since they match both daughters. {*-s*} then comes out as the only candidate for the whole AccP.

In general, every exponent that matches the AccP in (76) matches also the NomP in (75). This is the property of the system from which the generalization about agglutination and fusion follows. Since (i) in the ACC, the constituent containing class, case and number receives a single marker, and (ii) every marker that matches AccP matches also NomP, it follows that there is such a portmanteau marker available in the NOM as well. If the same marker surfaces, we get syncretism, if there is a better match, we get a different marker. But no matter which of these two situations obtains, the generalization is derived.

Let me now skip GEN/DAT and proceed directly to ABL, shown in (77-b).

(77) a. 

b. 

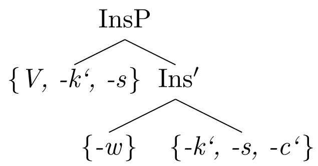
There is no difference in matching for either ClP or NumP compared to the previous scenarios. I will take a shortcut through the insertion at the complex case node, and focus on the complex head as a whole. It is clear that neither *-s* or *-k'* can spell it out (neither contains the features from D upwards). What we are thus left with is *-c'*, as well as *-w*, plus the singular ablative marker *-ê* (for which see (58)). Thus, I replace the complex case head with the set {*-ê, -c', -w*}, see (77-a). Spell-out then targets Abl′. The entry *-c'* matches both daughters, and it emerges as the sole possibility for the spell-out of this node. Note that due to the way spell-out proceeds (recall the Biggest Wins Theorem), we correctly exclude the possibility that ABL.PL gets spelled out as the combination of the ABL.SG *-ê* with *-k'*.[35]

Note as well that for any case smaller than ABL, *-c'* is in the competitor set for that node. Since *-c'* is also a competitor for the spell-out of NumP, then in any case smaller than ABL, *-c'* is going to be a candidate for a fusional spell-out of number and case. From that, it follows that fusion of case and number in the ablative forces their fusion for any smaller case.

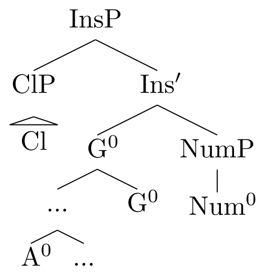
Finally, spell-out reaches the AblP node. There is no match, and hence, the two daughter sets are ordered: {*V, -k', -s*}> {*-c'*}. In the first set, a thematic vowel wins in competition. This correctly yields V>*-c'*.

Turning now to the INS.PL (78-b), there is only a single candidate for the spell-out of the complex case node: *-w*. (78-a) shows the result of matching:

(78) a.



b.



At Ins′, there is no match. The algorithm thus orders the daughter sets as {*-w*}>{*-k', -s, -c'*}. Due to competition, we arrive (correctly) at *-w-k'*. At InsP, the ordering between the class marker V (the most specific of the three eligible) and *w-k'* is introduced, yielding *V-w-k'*.

9.3 Agglutination and Fusion vs. the Subset Principle

This section turns to the following question: can we formulate an alternative analysis using the Subset Principle? The claim of this section is that this is impossible; but there are qualifications to be made along the way.

Most importantly, the analysis should deliver the splits between agglutination and fusion "for free," because this is what has been achieved in the

[35] A virtually identical situation obtains in the GEN/DAT.PL, since *-c'* is the only candidate for the spell-out of both the case and number branch of the bar node.

previous section using the Superset Principle. However, the standard way to deal with such splits in syntax-oriented frameworks such as Distributed Morphology (see, e.g., Embick and Noyer 2007) is by recourse to rules which either fuse (Fusion) or split (Fission) positions of exponence (syntactic terminals). For example, Halle and Vaux (1998) account for some of the facts discussed above by saying that (i) number and case represent a single node in syntax (that is why number and case are fused most of the time); and (ii) a special rule (Fission) creates a separate number position in the INS.PL (that is why the INS.PL is agglutinative). At a general level, it is assumed that the paradigm structure is rigid: each case has the same number of morphological slots to fill. If there are too few or too many markers, additional rules are invoked.

The approach suggested here is different. There is no definite number of slots to fill; there are only features that need spell-out. If there are comparatively few of them (as in the NOM), one marker may do for the whole lot. If there are comparatively many features (as in the INS), three markers may be needed. The borders between morphological slots are not cast in stone, they are fluid and emerge in the process of insertion. As a consequence, there is no need for special rules to fuse or split the assumed morphological positions. And, of course, once special morphological rules are gone, the need for a special morphological representation evaporates.

The conclusion that the Subset Principle is inadequate as a tool in modeling phrasal spell out presupposes that the view entertained here is what the theory should achieve. The claim is not that the Subset Principle is inadequate for the purposes it is generally used; it is inadequate if one wants to extend its use beyond what it was designed to handle.

In presenting the argument, I further take for granted that the contiguity restriction (plus case attraction and agglutination/fusion) shows that cumulative classification is on the right track. I now proceed to the reasoning.

The starting point is that under a Subset Principle approach, the marker -*k‘* is just plural, and only that. It cannot be specified for class, or case, because it shows up in contexts where case and noun class are not present — recall *sirem* ‘I love’ vs. *sirem-k‘* ‘we love’ (59). In addition, there is the fact that in the INS.PL, -*k‘* co-occurs with case and class (so it apparently does not spell them out). Thus, in order to ensure that in the NOM.PL, we do not see the class marker -*a* co-occuring with -*k‘*, -*k‘* has to be allowed to ‘extend upwards’ in the tree (75), swallowing everything on the way. This is consistent with the Subset Principle based reasoning: since Pl is a subset of the NOM.PL tree, -*k‘* may spell out the whole tree.

Why does -*k‘* not surface in the ACC.PL in (76)? Since PL is a subset of ACC.PL, -*k‘* is applicable. One is then led to propose that it is blocked by -*s*, specified as ACC.PL. (Because of such a specification, -*s* is inapplicable in NOM.PL: ACC.PL is not a subset of NOM.PL.) -*s* also spells out the locative, because ACC.PL specification represents a subset of it. Problems arise when we focus on the spell-out of GEN/DAT/ABL.PL. Taking the ablative as an example, see (77), I start by noting that -*s* must be a match for this constituent, following the general logic of the Subset Principle. Since it does not appear here, we must block it by *c'*, which is then specified as GEN.PL, so that *-c'* actually appears all the way from GEN to ABL. But what cannot be captured is the fact that the class marker *-a* emerges when *-c'* is inserted. Somehow, this has to be baked into the difference between *-s* and *-c'*. The only way to do that seems to be specifying *-s* for class, but leave *-c'* unspecified for class. (That is why the class marker appears with *-c'* but not with *-s*.) But then *-k'* must also be specified for class (since neither *-s* or *-k'* co-occur adjacent to class markers). This contradicts the initial observation that *-k'* cannot be specified for class.

At a general level, the problem is that if the unmarked case is expressed by a single marker, and that marker is allowed to spell out super-structures (the essence of the Subset Principle), then all more complex cases should have fusion as well (because the unmarked case corresponds to their subset). In other words, the generalization comes out the other way round: if we get fusion in an unmarked case, we should get it in all marked cases.

Because of this, the only way to get the facts right seems to be making INS the least marked case (the feature A), and NOM the most marked case (features A-F). But problems arise even on this view, as I argue below. Starting with the now unmarked INS.PL (now actually corresponding to the structure (75)), we specify *-k'* for plural, *-w* for the feature A (the sole feature of the INS) and the class markers spell out the class node. The fact that *-c'* does not appear as a case-number portmanteau in the INS.PL can then be encoded by specifying it as ABL.PL (now corresponding to the features A+B+Pl, i.e., a part of the tree in (76)). If that is the specification, then *-c'* is not a subset of the unmarked INS. Further, because ABL is a subset of all cases safe the INS, *-c'* will further be correctly eligible for insertion in all such cases, deriving the observation that fusion of case/number in the ABL entails their fusion from ABL to NOM. The fact that *-c'* actually appears only in DAT/GEN.PL is then another instance of blocking. In particular, the LOC/ACC.PL *-s* is going to be specified as LOC.PL (features A-E+Pl), which makes it a better competitor in case both *-c'* and *-s* apply. In addition, if class is added to the entry of *-s*, then no class marker will appear from LOC.PL to NOM.PL, i.e., the range of cases where *-s* may apply.

The place where the system breaks down is NOM.PL. In order to prevent *-s* to appear here, we have to block it by a specific NOM.PL marker (features A-F+Pl+Class). But the marker which actually appears there is *-k'*, which we already have listed in the lexicon as just [Pl]. With such a specification, however, it cannot block *-s* from occuring there. And if we change the specification of *k'*, we cannot insert it under Pl only, as required for INS.PL. We may then conclude that even under the reversal of the markedness hierarchy, we cannot give a unified account of the facts using the Subset Principle.[36]

[36] In addition, the explanation of Case Attraction in terms of ellipsis disappears under the markedness reversal.

9.4 Conclusions

This section has derived the observation that the ‘syncretism scale’ restricts fusional vs. agglutinative expression of morphological categories (each strategy forms a contiguous region with fusion at the unmarked end). This fact follows from the way cumulative classification interacts with the consequence (79-a). Importantly, the actual splits between agglutination and fusion are successfully derived only with reference to the underlying syntactic representation (and the actual lexical entries).

(79) Consequences
   a. if a phrase XP is spelled out by a single morpheme in a language L, any sub-structure of XP is spelled out by a single morpheme as well (in that language)
   b. insertion may only target features that form a constituent

10 The constituency of features

In this section, I turn to (79-b). To see its teeth in a simple example, recall the structure of the Armenian noun (72). In this structure, there is no constituent that contains number and class to the exclusion of case. Hence, there can be no class/number portmanteau to the exclusion of case. This is borne out: when number marking is separated from case (the *-k‘* in INS.SG), it is insensitive to class. When class is separated from case (various vowels in INS.SG.), it is insensitive to number (the same vowel emerges in INS.PL, see (58)).

The logic of the prediction allows us to account for more subtle effects. Suppose, for example, that two categories, α and β in (80-a), can be spelled out by a single portmanteau morpheme P. The entry of P is then as in (80-b). Now if another category, γ in (80-c), structurally intervenes between α and β, α and β cannot be spelled out by P. That is because α and β do not form a constituent in (80-c).

(80) a. $\Rightarrow$ P
   α β
   b. $/P/ \Leftrightarrow$ α β
   c. *$\Rightarrow$ P
   α γ β

What happens instead is that in (80-c), each of α and β is spelled out on its own (leaving aside an even more specific portmanteau that contains γ). I call this effect morpheme splitting: features that may correspond to a single morpheme in a given language must be spelled out by two pieces when they are separated by an intervener. I argue that morpheme splitting explains certain puzzling facts concerning consonantal declensions in Armenian.

Morpheme splitting also has consequences for the architecture of the spell-out system. In particular, if the effect is real, it demonstrates that the expression of syntactic features by phonological material is not driven by some kind of a simple economy condition, such that one merely tries to minimize the number of exponents. Instead, morphemes are sensitive to structural configurations of their features. If that is so, we gain additional justification for proposing that morphemes are structured entities.

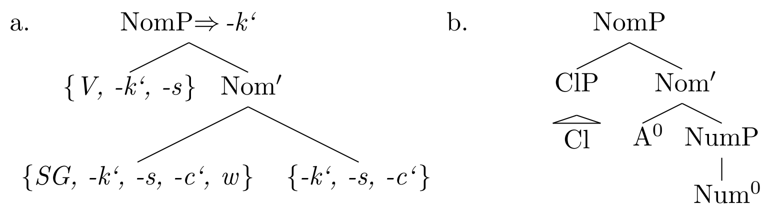
10.1 Morpheme splitting in a-stems

Let me first illustrate morpheme splitting on the familiar declension of the a-stems.

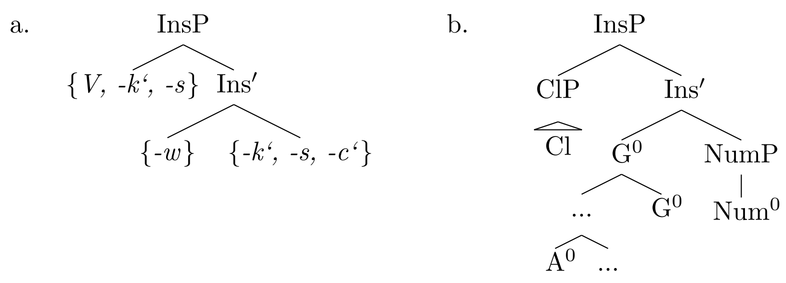
(81) *Classical Armenian, a-stem declension*

	nation, SG.	nation, PL.
NOM	azg-ø	azg-k‘
ACC	azg-ø	azg-s
LOC	azg-i	azg-s
GEN	azg-i	azg-a-c‘
DAT	azg-i	azg-a-c‘
ABL	azg-ê	azg-a-c‘
INS	azg-a-w	azg-a-w-k‘

Recall that in the INS.PL, we have identified a sequence of three morphemes: *-a-w-k‘*. I have suggested that *-a-* expresses class, which is spelled out together with other inflectional categories in the unmarked cases, i.e., in NOM, ACC and LOC. I show that on the example of the nominative, repeated from (75):

(82)


Under this analysis, the *-k‘* in NOM.PL actually spells out the features which are realized as *-a-* in the INS. Combining these statements, we realize that the reason why *-a-* and *-k‘* cannot fuse in the INS.PL is constituency: the features expressed as *-w-* intervene between the class marker and plural:

(83)


The structural intervention of *-w-* thus forces a single morpheme (*-k'*) to split in two (*-a-* and *-k'*), and reveals a hidden structure inside an apparently indivisible morpheme: the nominative plural *-k'*.

10.2 n-stems

N-stems in Classical Armenian offer another opportunity to observe the same effect. I give two examples of this declension type below. The two types differ by vowel quality in NOM/ACC/LOC.PL.

(84) *Classical Armenian n-stems*

	race, SG.	race, PL.	part, SG.	part, PL.
NOM	az-n	az-in-k'	mas-n	mas-un-k'
ACC	az-n	az-in-s	mas-n	mas-un-s
LOC	az-in	az-in-s	mas-in	mas-un-s
DAT	az-in	az-an-c'	mas-in	mas-an-c'
GEN	az-in	az-an-c'	mas-in	mas-an-c'
ABL	az-n-ê	az-an-c'	mas-n-ê	mas-an-c'
INS	az-am-b	az-am-b-k'	mas-am-b	mas-am-b-k'

The traditional approach (to be rejected) analyzes *-n-* and the preceding vowel (if any) as a single morpheme, which combines with the root to form a stem. The variation in the vocalic element of the theme marker is then seen as a variation of the stem, distinct from affixation.

However, if we compare the vowel to the rest of the system, we notice that it is identical to the vowel we observe as a case ending in a-stems. The following table shows this, with parts identical across the paradigms boldfaced. For INS, I follow Schmitt (1981, p.46) and equate *-w* and *-b* as two guises of the same morpheme, realized as *-b* after consonants, and *-w* after vowels.[37]

(85) *A-stems and n-stems decline the same (modulo n)*

	SG., A-STEM	SG., N-STEM
NOM	azg-**ø**	mas-**ø**-n
ACC	azg-**ø**	mas-**ø**-n
LOC	azg-**i**	mas-**i**-n
GEN	azg-**i**	mas-**i**-n
DAT	azg-**i**	mas-**i**-n
ABL	azg-**ê**	mas-n-**ê**
INS	azg-**a-w**	mas-**a**-m-**b**

Taking the parallel seriously, we uncover two facts. The first one is a positional asymmetry between the case marker *-i-* (precedes *-n*), and the case markers *-ê*

[37] I also assume that *m* in INS.PL is a phonological variant of the stem marker *n*.

and -*w* (follow -*n*-). The way the asymmetry cuts across the paradigm (INS + ABL vs. the rest) seems to be yet another instance of case hierarchy effects.[38]

The second observation is directly relevant for the present concerns: unlike in the singular, where we find a perfect match between the a-stem and the n-stem (modulo *n*), an unexpected vowel emerges between the root and -*n*- in NOM/ACC/LOC.PL. I put it in small caps:

(86) *A mysterious vowel in the plural of n-stems*

	PL., A-STEM	PL., N-STEM	PL., N-STEM
NOM	azg-**k‘**	mas-U-n-**k‘**	az-I-n-**k‘**
ACC	azg-**s**	mas-U-n-**s**	az-I-n-**s**
LOC	azg-**s**	mas-U-n-**s**	az-I-n-**s**
GEN	azg-**a**-**c‘**	mas-**a**-n-**c‘**	az-**a**-n-**c‘**
DAT	azg-**a**-**c‘**	mas-**a**-n-**c‘**	az-**a**-n-**c‘**
ABL	azg-**a**-**c‘**	mas-**a**-n-**c‘**	az-**a**-n-**c‘**
INS	azg-**a**-**w**-**k‘**	mas-**a**-m-**b**-**k‘**	az-**a**-m-**b**-**k‘**

The theory developed up to now allows us to understand the emergence of this unexpected vowel as a predicted instance of morpheme splitting.[39] In particular, the stem marker -*n*- structurally intervenes between the class node, and the constituent containing case/number, as indicated by its surface position in GEN-INS.PL. Because of this intervention, -*k‘* in NOM.PL cannot spell out both class and case/number features. As a consequence, the class features must be spelled out by a separate morpheme: the extra vowel. The account is fleshed out in more detail below.

First, I analyze the ‘stem marker’ -*n*- as a derivational morpheme. This proposal is based on the facts discussed in Olsen (1999). According to this study (p.113), the n-stem declension is one of the most populated ones in Classical Armenian (around 2000 items). Importantly, the vast majority of these nouns are abstract nouns derived from adjectives, nouns and verbs; example derivations include ‘old → oldness,’ ‘brother → brotherhood,’ and ‘kill → killing’ (p.547). This fact is accounted for if -*n*- attaches to a (potentially) complex base, see (87), and turns it into a noun. I further assume that the resulting complex is further dominated by the regular nominal sequence of projections, including Class, Number and Case:[40]

[38] In the present account, the ordering in LOC, GEN and DAT is due to phrasal movement. The root moves above case, without pied-piping the stem marker along.

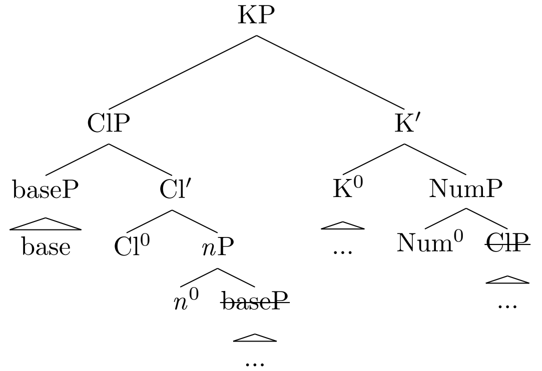
[39] Note that an analysis in terms of a purely phonological process is not viable, or at least not obviously so. To be sure, Classical Armenian does exhibit vowel-zero alternations (Schmitt 1981, p.38-39); however, they follow a different pattern than the one we are dealing with here. Thus, in the prototypical instance, we have a vowel in a closed syllable *amis* ‘month, NOM.SG’ and a zero in an open syllable *amøs-u* ‘month, GEN.SG’. This regular pattern is different from the data at hand, where both zero *mas-øn* ‘part, NOM.SG’ and a vowel *mas-in-k‘* ‘part, NOM.PL’ show up in a closed syllable. In other words, a theory based on epenthesis (as observed elsewhere in the language) would wrongly predict NOM **mas-in*. Cf. fn. 41.

[40] There is a number of nouns that are apparently underived. In fact, Olsen (1999, p.156) notes that the declension type is so productive, that any noun whose ACC.SG accidentally

(87) [K [Num [Cl [$_{noun}$ n [derivational base (A,V,N)]]]]

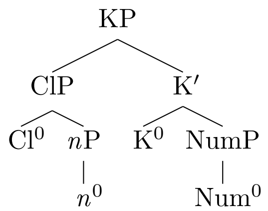
The derivation of the surface order in the plural declension starts by a series of movements of the base: first leading to [base-n], and then the base moves above the class marker [base-Cl-n]. Once it lands in the Spec of the Class marker, it pied-pipes this constituent to the left of Num: [base-Cl-n]-Num. Upon the merger of K (K [base-Cl-n]-Num), the constituent [base-Cl-n] moves across it without pied-piping Num, leading to the structure (88).

(88)

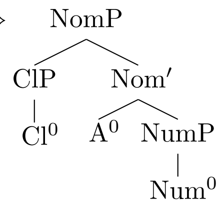


The phrasal node below the K head in (88) reflects the proposals made earlier concerning head movement of case features and tree simplification. As in the general case, the base further sub-extracts from Spec,ClP and need not concern us further. The resulting structure (with traces omitted) is shown in (89):

(89)



(90) /-*k'*/ ⇔



Now recall that -*k'* (and -*s*) spell out a constituent composed of the Class marker, K heads and number. The entry for -*k'* is repeated in (90). This entry, however, cannot spell out ClP node in (89), because of the presence of the nP node. As a result (89) gets spelled out as (91), with {-*k'*, -*s*, -*c'*} replacing Nom′, {*V*, -*k'*, -*s*} replacing Cl^0, and {n} replacing nP:

---

ended in the sequence *Cn* were re-analyzed as n-stems (i.e., *C*-n). I believe that these facts are compatible with the general view taken here, since in general, there is no obstacle in nominalizing nouns (recall ‘brotherhood’). Once again, the phrasal idiom view suggested in fn. 33 needs to be invoked.

(91)

KP
ClP {*-k‘*, *-s*, *-c‘*}
{*V*, *-k‘*, *-s*} {*-n*}

Thus, the way insertion is set up predicts that once the class marker is structurally separated from the case/number constituent by the intervening *-n-*, each must be spelled out separately. This in turn explains the appearance of the additional vocalic element between the base and the stem marker *-n-* in NOM, ACC and LOC plural. This is a non-trivial result, recall that the emergence of the vowel disturbs the otherwise perfect parallel (modulo *n*) between the a-stem and the n-stem declensions, shown below:

(92) *The mismatch between a- and n-stems in* NOM., ACC., LOC. PL.

	A-stem SG.	N-stem SG.	A-stem PL.	N-stem PL.
NOM	**-ø**	**-ø**-n	**-k‘**	-U-n-**k‘**
ACC	**-ø**	**-ø**-n	**-s**	-U-n-**s**
LOC	**-i**	**-i**-n	**-s**	-U-n-**s**
GEN	**-i**	**-i**-n	**-a-c‘**	**-a**-n-**c‘**
DAT	**-i**	**-i**-n	**-a-c‘**	**-a**-n-**c‘**
ABL	**-ê**	-n-**ê**	**-a-c‘**	**-a**-n-**c‘**
INS	**-a-w**	**-a**-m-**b**	**-a-w-k‘**	**-a**-m-**b-k‘**

As things stand, however, this picture leads also to the prediction that the vowel which appears between the noun and *-n-* is identical to the class marker which occurs also in the oblique cases. As I show in a moment, this is correct for some nouns, but it is wrong for the paradigm above. We do not get *-a-*, but *-u-* (a class marker of the u-stems).

I do not know what is the source of the difference. What is needed is a proposal of how individual class markers break down into components, for example where exactly *-n-* and *-a-* come in, and so forth. Alternatively, one can see the *-u-* as a result of ablaut, which has been suggested in Schmitt (1981, p.44). Thus, we would have an underlying *a*, subject to a morphologically triggered ablaut, yielding the surface *-u-* (see Guerssel and Lowenstamm 1996 for a relevant discussion). I leave this for future research.

On the other hand, a non-negligible number of n-stems does in fact show the predicted nominative *-a-n-k‘* and the accusative/locative *-a-n-s* (Schmitt 1981, p.103 lists about 30 such nouns). I am going to discuss one such example, namely the plural forms of the noun ‘sister in law,’ shown below:[41]

[41] This paradigm provides additional evidence against treating the vocalic marker as epenthetic. That is because the root is vowel final, and hence, the vowel does not improve the phonotactics in any way.

(93) *An alternating noun: 'sister in law'*

	sister in law, PL.
NOM	nu-(A-n)-k'
ACC	nu-(A-n)-s
LOC	nu-(A-n)-s
GEN	nu-a-(n)-c'
DAT	nu-a-(n)-c'
ABL	nu-a-(n)-c'
INS	nu-a-w-k'

In the cells stretching from NOM to ABL plural, this particular noun inflects either as a n-stem, or as an a-stem. The n-stem forms include the bracketed material, the a-stem forms exclude it. This variable behavior provides us with a minimal pair. According to the analysis, the presence of the additional *n* in NOM/ACC/LOC should lead to a structure where an additional class marker has to show up. And this is exactly what happens. Thus, compared to the a-stem forms of the same noun, the inclusion of *-n-* in NOM, ACC or LOC automatically leads to the emergence of an additional vowel. (Equivalently, the absence of *-n-* leads to the disappearance of this vowel.) The paradigm shown above thus bears out the predictions in their strongest form.

10.3 On phonological conflation

In this section, I address an apparent syncretism of the instrumental with the dative (and other cases) across the ablative in u-stems, shown in (94).

(94) *An* INS–DAT *syncretism in Classical Armenian?*

	time, SG.
NOM	žam
ACC	žam
LOC	žam-u
GEN	žam-u
DAT	žam-u
ABL	žam-ê
INS	žam-u

The discussion of this piece of data will ultimately show that the syncretism is irrelevant for the Case sequence (as given in (4)), because it is the product of a phonological conflation. Crucially from the current perspective, the discussion is also going to provide additional evidence for the analysis of morpheme splitting, and that is why I discuss this paradigm here.

The highlighted syncretism goes against the order of cases in Classical Armenian, because — as established on independent grounds — INS must be separated from DAT by ABL. Thus, we apparently get a situation in which non-adjacent layers of case show syncretism, and this cannot be accounted for by the present system. However, I am going to argue that the syncretism is the result of a phonological process which merges *u* and *w* into one segment, and thus the underlying morphology looks as depicted in (95).

(95) a. Dat: *u*
b. Ins: *u-w* ⇒ *-u*

There are two aspects of the proposal. The first aspect is that the sequence *uw* is simplified to *u* in Classical Armenian. The second aspect is the hypothesis that there is an underlying *-uw* present in the instrumental. I provide evidence for these two claims in turn.

The evidence for the existence of the relevant phonological process consists in showing that the process is attested elsewhere in the language. The following example set shows this. In (96-a), we observe an instance of a productive suffix -*wor*, which (in one of its uses) means 'bearer' (Olsen 1999, 362). When this suffix attaches to the stem *zim-u* 'armour,' we obtain the relevant sequence *uw*. The resulting form *zimu-or* then confirms that *uw* yields *u* quite generally (Olsen 1999, p.367,615, cf., Schmitt 1981, p.46).

(96) a. ałełna -wor → ałełnawor
bow -bearer
'archer'
b. zimu -wor → zimuor
armour -bearer
'soldier'

Turning to the second point: what evidence is there for the underlying presence of *-uw* in the instrumental plural of the u-stem 'time'? The evidence is provided by the declension of the noun 'day:'

(97) *The r-stem 'day' with a mysterious -b*

	time, SG.		day, SG.
NOM	žam-**ø**		aw-**ø**-r
ACC	žam-**ø**		aw-**ø**-r
LOC	žam-**u**		aw-**u**-r
GEN	žam-**u**		aw-**u**-r
DAT	žam-**u**	⇔	aw-**u**-r
ABL	žam-**ê**		aw-r-**ê**
INS	žam-**u**	⇔	aw-**u**-r-B

The declension of the r-stem 'day' shows a parallel to the n-stems in the following sense: in the cells from NOM to ABL, the noun 'day' inflects just like the u-stem 'time' with an additional *r* that occupies a structural position analogous to -*n*. What is to be noted is the presence of two markers (apart from *r*) in the INS of 'day' (*u* and *b*). This contrasts with apparently only one marker in the paradigm 'time' (*u*). The contrast sticks out when we realize that the homophonous DAT *u* does not split, but it is ordered to the left of *r*. Hence, what is initially mysterious is the emergence of the *-b* in *aw-u-r-b*; that's why I have put it in small caps in (97).

The point is that the unexpected appearance of *-b* becomes completely regular once we adopt the proposal that the INS *-u* is underlyingly *u-w*. This sequence is merged into one segment in cases where they end up adjacent, but it is preserved when they are separated by the consonantal stem marker:

(98) *The r-stem 'day' explained*

	time, SG.	day, SG.
INS	žam-**u-w** (⇒ -u)	aw-**u**-r-**b**

This provides the evidence needed for the underlying presence of *u-w* in the INS of u-stems, and supports the conclusion that the syncretism of DAT and INS across a distinct ABL is to be treated as accidental, and irrelevant for the issues under discussion.

The paradigm of 'day' is interesting also for the phenomenon of morpheme splitting. Because of the parallel behavior of *r* and *n*, we now predict that in the NOM/ACC/LOC.PL, a class marker should emerge between the root *aw* 'day' and the *r*. The prediction is borne out. As the table below shows, the class marker *-u-* (in small caps) appears where predicted, even though it has no counterpart in the (otherwise parallel) declension of the u-stem noun 'time':

(99) Morpheme splitting in r-stems

	time, PL.	day, PL.
NOM	žam-**k'**	aw-U-r-**k'**
ACC	žam-**s**	aw-U-r-**s**
LOC	žam-**s**	aw-U-r-**s**
GEN	žam-**u-c'**	aw-**u**-r-**c'**
DAT	žam-**u-c'**	aw-**u**-r-**c'**
ABL	žam-**u-c'**	aw-**u**-r-**c'**
INS	žam-**u-w-k'**	aw-**u**-r-**b-k'**

The paradigms above show that the appearance of the vocalic class marker in the relevant cells of the paradigms is not an effect of a particular class (the n-stems or the r-stems). It is a structurally governed process, which appears every time the relevant structural configuration obtains.

11 Conclusions

The main goal of this paper has been to show that if one adopts (i) fine-grained syntax, and (ii) phrasal spell-out, an interesting analysis of the Classical Armenian declension appears. Two aspects of the approach deserve final highlighting.

The first one is the fact that the observed 'case sequence' governs the behavior of at least three logically independent empirical domains (syncretism, case attraction, and agglutination vs. fusion). I believe it is a challenge for any theory to explain why these phenomena should be subject to the same regularity. The current proposal achieves this by 'cumulative decomposition:' the proposal that individual cases are characterized by a growing number of features (cf. Caha 2009). Because of the way the case sequence penetrates various aspects of Classical Armenian grammar, I think its discovery and its particular implementation may well be the main empirical contribution of this paper. If correct, the approach taken here suggests that there is only one mode of grammatical organization of smaller units into bigger chunks (i.e., binary branching tree), no matter how small (sub-morphemic features) or big (phrases) the units are.

And secondly, it is also important to pay attention to the mechanisms which derive these results on the basis of the proposed decomposition. In the current theoretical context, phrasal spell-out based on the Superset Principle (due to Starke 2009) may be the most important innovation described here. Without phrasal spell-out, syntactic terminals are the only locus for insertion, and their number determines the number of slots in the paradigm. But since the number of separate morphological slots for class and gender varies (none, 1 or 2), and the number of syntactic terminals for number and gender is constant, the mismatches have to be stated by brute force and lead to the consequence that a separate morphological structure becomes indispensable. In a phrasal spell-out system, the number of terminals may not correspond to the number of morphemes. Consequently, the number of paradigm slots may vary, and appears as an automatic consequence of the interaction between the syntactic structure and the actual entries.

References

Abraham, Werner. 2003. The myth of doubly governing prepositions in German. In *Motion, Direction and Location in Languages: In Honor of Zygmunt Frajzyngier*, edited by Erin Shay and Uwe Seibert, pp. 19–38. John Benjamind, Amsterdam.

Ackema, Peter and Ad Neeleman. 2007. Morphology ≠ syntax. In *The Oxford Handbook of Linguistic Interfaces*, edited by Gillian Ramchand and Charles Reiss, pp. 325 – 353. Oxford University Press, Oxford.

Baerman, Matthew. 2008. Case syncretism. In *The Handbook of Case*, edited by Andrej Malchukov and Andrew Spencer, pp. 219–230. Oxford University Press, Oxford.

Baerman, Matthew, Dunstan Brown, and Greville G. Corbett. 2005. *The Syntax-Morphology Interface. A Study of Syncretism*. Cambridge University Press, Cambridge.

Baker, Mark C. 1988. *Incorporation: A Theory of Grammatical Function Changing*. University of Chicago Press, Chicago.

Baker, Mark C. and Ruth Kramer. 2010. prepositions as case morphemes inserted at PF in Amharic. Handout of a talk at the BCGL5 Conference, December 2-3, at CRISSP, Brussels.

Baltin, Mark. 2006. The non-unity of VP-preposing. *Language* 82: 734–766.

Bittner, Maria and Ken Hale. 1996. The structural determination of case and agreement. *Linguistic Inquiry* 27 1: 1–68.

Blake, Barry J. 1994. *Case*. Cambridge University Press, Cambridge, second edn.

Bobaljik, Jonathan. 2002. Syncretism without paradigms: Remarks on Williams 1981,1994. In *Yearbook of Morphology 2001*, edited by Geert Booij and Jaap van Marle, pp. 53 – 86. Kluwer, Dordrecht.

Bobaljik, Jonathan. 2007. On comparative suppletion. Ms., University of Connecticut.

Bobaljik, Jonathan. 2012. *Universals in Comparative Morphology: Suppletion, superlatives, and the structure of words*. MIT Press, Cambridge, Mass.

Borer, Hagit. 2005. *Structuring Sense: An Exo-Skeletal Trilogy*. Oxford University Press, New York.

Caha, Pavel. 2009. *The Nanosyntax of Case*. Ph.D. thesis, CASTL, University of Tromsø.

Caha, Pavel. 2011. The parameters of case marking and spell out driven movement. *Linguistic variation yearbook 2010* 10: 33–78.

Caha, Pavel. to appear. Case in adpositional phrases. *Syntax and Semantics* .

Calabrese, Andrea. 2008. On absolute and contextual syncretism. In *Inflectional Identity*, edited by Asaf Bachrach and Andrew Nevins, pp. 156–205. Oxford University Press, Oxford.

Cardinaletti, Anna and Michal Starke. 1999. The typology of structural deficiency: A case study of three classes of pronouns. In *Clitics in the Languages of Europe*, edited by Henk van Riemsdijk, pp. 145–233. Mouton, Berlin.

Carnie, Andrew. 2008. *Constituent Structure*. OUP, Oxford.

Chomsky, Noam. 1995. *The Minimalist Program*. MIT Press, Cambridge, Ma.

Chomsky, Noam. 2008. On phases. In *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*, edited by Robert Freidin, Carlos P. Otero, and Maria Luisa Zubizarreta, pp. 133–166. MIT Press, Cambridge, MA.

Cinque, Guglielmo and Luigi Rizzi. 2010. The cartography of syntactic structures. In *The Oxford handbook of linguistic analysis*, edited by Bernd Heine and Heiko Narrog, pp. 51–65. Oxford University Press, Oxford.

Embick, David and Rolf Noyer. 2007. Distributed Morphology and the syntax-morphology interface. In *The Oxford handbook of linguistic interfaces*, edited by Gillian Ramchand and Charles Reiss, pp. 289–324. Oxford University Press, Oxford.

Guerssel, Mohand and Jean Lowenstamm. 1996. Ablaut in Classical Arabic measure I active verbal forms. In *Studies in Afroasiatic Grammar*, edited by Jacqueline Lecarme, Jean Lowenstamm, and Ur Shlonsky, pp. 123–134. HAG, The Hague.

Haegeman, Liliane. 1994. *Introduction to Government and Binding Theory*. Blackwell, Oxford.

Halle, Morris. 1997. Impoverishment and fission. In *Papers at the Interface*, edited by Benjamin Bruening, Y. Kang, and Martha McGinnis, no. 30 in MITWPL, pp. 425 – 449. MIT, Cambridge, MA.

Halle, Morris and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, edited by Kenneth Hale and Samuel Jay Keyser, pp. 111–176. MIT Press, Cambridge, Ma.

Halle, Morris and Bert Vaux. 1998. Theoretical aspects of Indo-European nominal morphology: The nominal declensions of Latin and Armenian. In *Mir Curad: Studies in Honor of Clavert Watkins*, edited by Jay Jasanoff, H. Craig Melchert, and Lisi Olivier, Innsbrucker Beiträge zur Sprachwissenschaft, pp. 223–240. Institut für Sprachwissenschaft der Universität Innsbruck, Innsbruck.

Harley, Heidi. 2008. When is syncretism more than a syncretism? Impoverishment, metasyncretism and underspecification. In *Phi theory: Phi-features across modules and interfaces*, edited by Daniel Harbour, David Adger, and Susana Béjar, pp. 251 – 294. Oxford University Press, Oxford.

Jackendoff, Ray. 1983. *Semantics and Cognition*. MIT Press, Cambridge, Ma.

Jakobson, Roman. 1962. Beitrag zur allgemeinen Kasuslehre: Gesamtbedeutungen der russischen Kasus. In *Selected writings, vol. 2*, pp. 23–71. Mouton, The Hague.

Johnston, Jason Clift. 1996. *Systematic Homonymy and the Structure of Morphological Categories. Some Lessons from Paradigm Geometry*. Ph.D. thesis, University of Sydney.

Julien, Marit. 2007. On the relation between morphology and syntax. In *The Oxford Handbook of Linguistic Interfaces*, edited by Gillian Ramchand and Charles Reiss, pp. 209 – 238. Oxford University Press, Oxford.

Kayne, Richard S. 1994. *The Antisymmetry of Syntax*. MIT Press, Cambridge, Ma.

Kayne, Richard S. 2005. Some notes on comparative syntax, with special reference to English and French. In *The Oxford Handbook of Comparative Syntax*, edited by Guglielmo Cinque and Richard S. Kayne, pp. 3–69. Oxford University Press, Oxford.

Kiparsky, Paul. 1973. 'Elsewhere' in phonology. In *A Festschrift for Morris Halle*, edited by Paul Kiparsky and Steven Anderson. Holt, Rinehart and Winston, New York.

Koopman, Hilda and Anna Szabolcsi. 2000. *Verbal Complexes*, vol. 34 of *Current Studies in Linguistics*. MIT Press, Cambridge, Ma.

Kratzer, Angelika. 2000. Building statives. Berkeley Linguistic Society.

Krause, Todd B. and Jonathan Slocum. no date. Classical Armenian online. Http://www.utexas.edu/cola/centers/lrc/eieol/armol-0.html.

Marantz, Alec. 1995. 'Cat' as a phrasal idiom: Consequences of late insertion in Distributed Morphology. Ms., MIT.

McCawley, James D. 1968. Lexical insertion in a transformational grammar without Deep Structure. In *Papers from the fourth regional meeting of the Chicago Linguistic Society*, edited by B. J. Darden, C.-J. N. Bailey, and A. Davidson. University of Chicago, Chicago.

McCreight, Katherine and Catherine V. Chvany. 1991. Geometric representation of paradigms in a modular theory of grammar. In *Paradigms: The Economy of Inflection*, edited by Frans Plank, pp. 91 – 112. Mouton de Gruyter, Berlin, New York.

McFadden, Thomas. 2004. *The position of morphological case in the derivation: a study on the syntax-morphology interface*. Ph.D. thesis, University of Pennsylvania.

McFadden, Thomas. 2007. Default case and the status of compound categories in Distributed Morphology. *Penn Working Papers in Linguistics (Proceedings of PLC 30)* 13.1: 225–238.

Müller, Gereon. 2003. A Distributed Morphology approach to syncretism in Russian noun inflection. In *Proceedings of FASL 12*, edited by Olga Arnaudova, Wayles Browne, Maria Luisa Rivero, and Danijela Stojanovic, pp. 353–374. Michigan Slavic Publications, Ann Arbor.

Muriungi, Peter. 2008. *Phrasal movement inside Bantu verbs: Deriving affix scope and order in Kiitharaka*. Ph.D. thesis, CASTL, Tromsø.

Neeleman, Ad and Kriszta Szendrői. 2007. Radical pro-drop and the morphology of pronouns. *Linguistic Inquiry* 38: 671–714.

Olsen, Birgit Anette. 1999. *The noun in Biblical Armenian*. Mouton de Gruyter, Berlin, New York.

Pantcheva, Marina. 2010. The syntactic structure of locations, goals, and sources. *Linguitics* 48: 1043–1081.

Pantcheva, Marina. 2011. *Decomposing Path. The nanosyntax of directional expressions*. Ph.D. thesis, CASTL, Tromsø.

Pesetsky, David. 1995. *Zero Syntax: Experiencers and Cascades*. The MIT Press, Cambridge, MA.

Phillips, Colin. 2003. Linear order and constituency. *Linguistic Inquiry* 34: 37–90.

Plank, Frans. 1991. Rasmus Rask's dilemma. In *Paradigms: The Economy of Inflection*, edited by Frans Plank, pp. 161–196. Mouton de Gruyter, Berlin.

Plank, Frans. 1995. (Re-)introducing suffixaufnahme. In *Double Case: Agreement by Suffixaufnahme*, edited by Frans Plank, pp. 3–112. OUP, Oxford.

Plank, Frans. 1999. Split morphology: How agglutination and flexion mix. *Linguistic Typology* 3: 279–340.

Radkevich, Nina. 2009. Vocabulary insertion and the geometry of local cases. Ms., UConn, downloadable at: http://ling.auf.net/lingBuzz/000958.

Ramchand, Gillian. 2008. *Verb Meaning and the Lexicon*. Cambridge University Press, Cambridge.

Richards, Norvin. 2007. Lardil "case stacking" and the structural/inherent case distinction. Ms., MIT, downloadable at: http://ling.auf.net/lingBuzz/000405.

Schmitt, Rüdiger. 1981. *Grammatik des Klassisch-Armenischen*. No. 32 in Innsbrucker Beiträge zur Sprachwissenschaft. Institut für Sprachwissenschaft der Universität Innsbruck, Innsbruck.

Schütze, Carson. 2001. On the nature of default case. *Syntax* 4(3): 205–238.

Starke, Michal. 2005. Nanosyntax class lectures. Spring 2005, University of Tromsø.

Starke, Michal. 2009. Nanosyntax. A short primer to a new approach to language. In *Nordlyd 36: Special issue on Nanosyntax*, edited by Peter Svenonius, Gillian Ramchand, Michal Starke, and Tarald Taraldsen, pp. 1–6. University of Tromsø, Tromsø. Available at www.ub.uit.no/munin/nordlyd/.

Starke, Michal. 2010. Universal grammar vs. lexically driven derivations, once more. Talk at the SmaSh workshop, University of South Bohemia, České Budějovice, June 23.

Starke, Michal. 2011. Towards elegant parameters: Language variation reduces to the size of lexically stored trees. Transcipt from a talk at Barcelona Workshop on Linguistic Variation in the Minimalist Framework. Available at http://ling.auf.net/lingBuzz/001183.

Taraldsen, Tarald. 2010. The nanosyntax of Nguni noun class prefixes and concords. *Lingua* 120 6: 1522 – 1548.

Travis, Lisa. 1984. *Parameters and Effects of Word Order Variation*. Ph.D. thesis, MIT.

Vangsnes, Øystein A. in press. Syncretism and functional expansion in Germanic wh-expressions. *Language Sciences* 34.

Šurkalović, Dragana. 2011. Modularity, linearization and phase-phase faithfulness in Kayardild. *Iberia: An International Journal of Theoretical Linguistics* 3 1: 81 – 118.

Weerman, Fred and Jacqueline Evers-Vermeul. 2002. Pronouns and case. *Lingua* 112: 301–338.

Wiese, Bernd. 2003. Zur lateinischen Nominalflexion: Die Form-Funktions-Beziehung. Ms., IDS Mannheim.

Wiese, Bernd. 2004. Categories and paradigms. On underspecification in Russian declension. In *Explorations in Nominal Inflection*, edited by Gereon Müller, Lutz Gunkel, and Gisela Zifonun, pp. 321–372. Mouton de Gruyter, Berlin.

Williams, Edwin. 1981. On the notions "lexically related" and "head of a word". *Linguistic Inquiry* 12.2: 245–274.

Williams, Edwin. 1994. Remarks on lexical knowledge. *Lingua* 92: 7 – 34.

Acknowledgements A number of people gave me comments and suggestions concerning the material presented here. I started working on the topic as a part of my dissertation-related research, and my supervisor Michal Starke has substantially influenced the initial stage of work. Jonathan Bobaljik and Hilda Koopman (in their capacity as committee members) have left their mark on this work as well, for which I am very grateful. In addition, a very early version of the paper was presented at WOTM 4 in Leipzig (2008), and I thank the audience there for feedback and interesting discussions.

Subsequently, Gillian Ramchand and Peter Svenonius have provided me with detailed comments on a draft of this material as it appeared in Nordlyd (Tromsø working papers in linguistics, 2009). Major changes have occurred in the first draft submitted to NLLT due to the comments from three anonymous reviewers, accompanied by an additional review by the handling NLLT editor, Gereon Müller. For the final round of comments, I am indebted to Marcel den Dikken. A very special thanks to Marina Pantcheva, who has read and commented on all these various versions.

I thank all these people for bringing up empirical challenges, pointing out problems of analysis, finding better ways of putting things, removing typos, correcting my English, and asking some big picture questions.