The parameters of case marking and spell out driven movement

Pavel Caha

Abstract

This paper proposes that case decomposes into a number of separate functional projections, which are ordered in a universal functional sequence. Cross-linguistic variation in position and complexity of case morphology is then accounted for in terms of differential noun phrase movement within the invariant sequence. I further investigate the possibility that variation in movement may be reduced to variation in the shape of the actual lexical entries. In order to implement this, the model of cyclic spell out by Starke (2009b) is adopted. In this model, each step of external merge is followed by lexical access. Consequently, evacuation movements may be triggered after each step of external merge in order for successful lexicalization to take place.

1   Introduction

If we understand a case marker to be a formal device that signals the grammatical role of a noun phrase (as in, e.g., Moravcsik 2009), two distinctions can be made concerning the position and the complexity of the marker.

The first distinction cuts between prefixal/prepositional and suffixal/postpositional marking.[1] Thus, for example, the Samoan example in (1-a) shows prepositional marking, while the Hungarian (1-b) shows suffixal marking:

(1)   a.   E   sogi e   le tama Ø   le ufi i   le to'i.
            PRES *cut* ERG *the boy*   ABS *the yam* INS *the knife*
            'The boy cuts the yam with the knife.'   (Moravcsik, 2009, p.242)
       b.   A   fiú   fel-vág-ja a   krumpli-t   a   kés-sel.
            The boy.NOM *up-cut-3sg the potato*–ACC *the knife*-INS

            'The boy cuts up the potato with the knife.'   (Hungarian Éva
            Dékány, p.c.)

The second division is based on the complexity of the marker, and it distinguishes between simplex (synthetic) and complex (analytic) marking. Thus, for example, the marking of "semantic cases" in Estonian (below the line) is said to be analytic, because it involves the genitive case and an additional marker:

---

[1]I remain oblivious here to the distinction between prefixes and prepositions, or suffixes and postpositions.

(2)　　An Estonian sample paradigm (adapted from Blevins 2008)

	church, SG.	church, PL.
NOM	kirik	kiriku-d
ACC	kirik-u	kiriku-d
GEN	**kirik-u**	**kiriku-te**
PART	**kirik-u**-t	kiriku-sid
ADES	**kirik-u**-l	**kiriku-te**-l
DAT/ALLA	**kirik-u**-le	**kiriku-te**-le
ABLA	**kirik-u**-lt	**kiriku-te**-lt
INES	**kirik-u**-s	**kiriku-te**-s
ILLA	**kirik-u**-sse	**kiriku-te**-sse
ELA	**kirik-u**-st	**kiriku-te**-st

In contrast, the marking in the following paradigm from Classical Armenian is synthetic; while in some cases, a distinction can be drawn between class/number/case morphemes, the case morphemes themselves appear to be undecomposable (see Caha 2009a for a detailed analysis):

(3)　　*Classical Armenian, Schmitt (1981)*

	nation, SG.	nation, PL.
NOM	azg-**ǿ**	azg-**k'**
ACC	azg-**ǿ**	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'

In this paper, I follow two related goals. The first one is to propose a theory of case that aims to capture the variation along these two dimensions, as well as the restrictions it is subject to. The theory is formulated in a framework that falls within the cartography tradition (see, e.g., Cinque and Rizzi 2008 for a recent overview; see also Cinque 2005, whose approach I follow here closely). In particular, I assume that there is a fine-grained and universal functional sequence, which is the source of underlying similarity of all human languages. The variation observed in the surface patterns is then due to distinctions in the lexicon, and/or differential application of movement processes (leaving it aside for now whether these two apparently distinct sources of variation can be unified).

Specifically, I propose that (i) each case decomposes into a number of functional projections located above the DP, which are ordered in a universal functional sequence. If that is the case, then (ii) the observed variation falls out as the result of differences in noun phrase movement within the fixed sequence. Specifically, the variation in the height of movement accounts for the variation in prefixal/suffixal marking (and the restrictions it is subject to). The second dimension of variation is accounted for by another parameter of movement, namely whether projections along its path are pied-piped or not. More specifically, when no pied-piping occurs inside the stack of functional projections responsible for case, these projections form a single constituent to the exclusion of the noun (once it evacuates). I follow Starke (2009a) and propose that such a

constituent can be subject to phrasal lexical insertion, yielding synthetic marking for case. However, when pied-piping occurs, the case features end up in (at least) two separate constituents, and hence, they cannot be spelled out by a single entry (assuming with Starke 2009a that lexical entries apply only to constituents).

Apart as an attempt to explain the facts we encounter, the enterprise is interesting from a general theoretical perspective. If correct, the proposal entails that the type of variation we find in phrasal syntax is replicated at a substantially more microscopic level. In fact, each of the terminals I will be dealing with corresponds to a unit of granularity smaller than a morpheme, perhaps to a single feature.

This part of the study thus supports the general approach to syntax suggested in Kayne (2005), Cinque and Rizzi (2008), and elaborated in a specific direction (called Nanosyntax) by Starke (2009a). Under this view, syntax does not group together pre-packaged lexical units, be they morphemes, or words. Instead, syntax is an entirely pre-lexical component which operates on single features, and groups them together into large and fine-grained syntactic trees. Morphemes – since they generally correspond to a number of features – are consequently understood as phrasal entities, constructed from atomic features by successive application of Merge (for which see Chomsky 1995).

The second goal of this paper stems from the conclusion of the first part. In particular, once variation in case marking is understood as a product of variable movement, a question arises what triggers (this type of) movement, and what is responsible for variation. To answer this question, I follow Starke (2009b), and introduce a theory where (at least certain type of) movement is driven by lexical entries. This idea is implemented in a model where spell out applies cyclically (bottom-up), after each step of external merge. If this approach turns out to be on the right track, then the two distinct sources of variation mentioned above (different lexical entries and different movements) reduce to a single source (namely different lexical entries).

2 The Nanosyntax of case

In this section, I present a model of case inflection designed to capture variation between suffixal vs. prepositional marking of case, as witnessed in many languages. The model has been developed and introduced in more detail in Caha's (2009b) *The Nanosyntax of Case*, hence the name of this section. The presentation proceeds in three steps: first I highlight an empirical generalization concerning the regularity of suffixal vs. prepositional marking, then I introduce the actual model, and finally, I illustrate its workings by means of empirical examples.

2.1 The preposition/suffix hierarchy

It is well known that not all languages have a strictly prepositional case (like Samoan) or strictly suffixal case (like Hungarian). There is a number of languages which fall somewhere in between; some cases are suffixal, and some involve a preposition in addition. Thus, for example, the instrumental in German

is expressed as a combination of both a case suffix and a preposition:[2]

(4) mit einem   Löffel
     *with a.*DAT.SG *spoon*
     'with a spoon' (German)

The division between what case is expressed as a suffix and what case is expressed by a functional prepositions is far from random. Building on previous work by Blake (1994), I have proposed in Caha (2009b) that there is in fact a linear sequence of cases in which both prefixal and suffixal marking form a contiguous region. I state this as (5):

(5) *The preposition/suffix hierarchy*
     a. If the expression of a particular case in the Case sequence (below) involves a preposition, then all cases to its right do as well.
     b. The Case sequence: NOM – ACC – GEN – DAT – INS – COM

I will now illustrate the predictions of (5) on a couple of examples, taking into account also potential counterexamples. The point in presenting the apparent counterexamples is to show that while they may be problematic for Blake's original proposal (on which I rely), they do not represent counterexamples to the formulation (5).

The differences between Blake's approach and (5) are essentially two. First, the sequence (5-b) is simplified compared to Blake's proposal by excluding the ergative, and spatial cases. Omitting these cases has various reasons which I cannot discuss in full detail for reasons of space, and their implementation into the overall picture is thus a task I leave for future research.

The second difference — important for present concerns — is that Blake views the case sequence (5-b) as a hierarchy that governs the inventory of cases in a language (where case is understood more narrowly than here, namely as a suffixal inflectional category). In essence, Blake says that if a language has a case listed on the hierarchy, it will usually have all the cases to its left. Blake's statement is thus a near equivalent of (5); however, there are certain subtle differences which have led many researchers to believe that Blake's hierarchy (and any linear statement in general) is untenable as anything stronger than a tendency. However, I do not think that this is the case; rather, the slight reformulation I have offered above takes care of a number of apparently problematic examples, and turns the hypothesis into a more accurate claim with the ambition of a language universal (subject to further testing). I illustrate some of the issues below.

Starting with a simple example, consider a two-case language such as English; here NOM and ACC show no preposition, while the genitive is introduced by *of*. On the basis of the statement (5), we now predict that all other cases require a preposition as well. This is borne out, and we find the dative *to* and the instrumental/comitative *with*.

The English system, however, does not exhaust the logical possibilities for two-case systems; for example, there could in principle be languages which show no preposition for only NOM and INS, but require a preposition for ACC, GEN, etc.

---

[2]Such languages are also illustrative of the analytical strategy; instrumental in German is expressed by a complex marker *mit*+DAT.

However, to the best of my knowledge, there are no such (or similar) languages, and this state of affairs is correctly ruled out by the statement (5): if ACC shows a preposition, then INS has to have one as well.

Moving on to three-case languages, we find a system of NOM, ACC and GEN, with prepositional marking for the remaining cases; Standard Modern Arabic can be taken as an example. Thus, Arabic shows no preposition for subjects (6-a), objects (6-a) and possessors (6-b), but requires a preposition for dative (recipient) arguments (6-a) and instruments (6-c):

(6) Standard Arabic, Islam Youssef (p.c.)
  a. aħDara-t   maryam-u   al-xiTaab-a   *(li)- muhammad-in.
      brought-AGR *Maryam*-NOM DEF-*letter*-ACC TO- *Muhammad*-GEN
      'Maryam brought the letter to Muhhamad.'
  b. umm-u   muhammad-in
      mother-NOM *Muhammad*-GEN
      'Muhammad's mother'
  c. akala   al-ħisaa'-a   *(bi-) milʕaqat-in
      ate.3RD.MASC DEF-*soop*-ACC *with-* spoon-GEN
      'He ate the soop with a spoon.'

However, there appear to be no languages where, for example, DAT has no preposition, but where GEN requires one. This is ruled out by the hypothesis: since GEN has a preposition, DAT should have one as well.

In the particular case of GEN and DAT discussed above, we can observe the advantages of the statement (5) over Blake's. Thus, a number of researchers have stressed as problematic the fact that alongside the system of Arabic (NOM – ACC – GEN), there are also three case systems consisting of NOM – ACC – DAT but no GEN; see for example the description of Elfdalian – a rather unique dialect of Swedish – in Dahl and Koptjevskaja-Tamm (2006). On the basis of the facts that the authors discuss, they conclude that "a strict hierarchy here cannot be established:" the presence of dative does not imply the presence of genitive (contrary to Blake's statement), and the presence of genitive does not imply the presence of dative (as required on the basis of the Arabic facts presented above).

However, the dialects discussed in Dahl and Koptjevskaja-Tamm (2006) do not violate the statement in (5). To see why, consider the relevant Elfdalian examples in (7), taking both examples and glosses from (Dahl and Koptjevskaja-Tamm, 2006, p.65-6):

(7) a. ǫ mainum faðer-es   garde
         *on my.*DAT *father.*DAT-POSS *farm.*DAT
         'on my father's farm'
     b. pennskrineð kullun
         *pen box.*DEF *girl.*DAT
         'the girl's pen box'

What we see here is that in order to express a possessor, Elfdalian uses a form which is identical to the dative (a cross-linguistically relatively well attested pattern). When this form is pre-nominal, as in (7-a), an additional morpheme appears; in the postnominal position (which is claimed to be the most common

one), a form identical to the dative alone is enough to stand for the genitive.[3] Thus, there is a crucial asymmetry between Arabic and Elfdalian. In particular, no preposition shows up for possessors in Elfdalian; and since there is no preposition for the genitive function, we predict nothing concerning the presence or absence of dative case suffix. In contrast, a preposition has to show up with indirect objects in Arabic, revealing the type of asymmetry between genitive and dative which is crucial for the ordering in (5), proposed by Blake.

Similar considerations make irrelevant a problem for Blake's hierarchy noted by Asbury (2008), a.o. Thus, on p.188 Asbury states: "Clearly the present position of genitive in [the Case sequence] could not be correct, since Hungarian, for example, has dative, locative, ablative/instrumental and many others, but no genitive." However, as in the case of Elfdalian, Hungarian is not problematic for the restatement (5). That is because possessors in Hungarian do not come out as PPs; they are marked either nominative, or dative (depending on their structural position). Thus, repeating the same reasoning as I presented for Eldalian: since there is no prepositional marking for GEN, we predict nothing on the basis of (5) concerning the marking for DAT and other cases higher up in the scale. What we do predict is that since Hungarian has an instrumental suffix, it has no prepositional marker for any lower case (function); and that is correct, the absence of a dedicated genitive case notwithstanding.

Thus, the conclusion I draw is that a language wide syncretism (e.g., of GEN and DAT) and quite likely other orthogonal factors influence the inventory of morphologically distinct cases in a language. However, once these effects are taken into consideration, the Case sequence in (5) seems to correctly describe what is a possible and impossible natural language.

Moving on to four-case systems, we find languages such as Icelandic or standard German with suffixes for NOM, ACC, GEN and DAT, and prepositional marking for INS and COM. These languages thus represent a predicted type, with a cut in the hierarchy between DAT and INS.

On the side of potential counterexamples, an anonymous abstract reviewer for FASL 2007 conference has brought to my attention a group of Caucassian languages, including Kabardian (Colarusso 1992), which have a dedicated instrumental case suffix, but (apparently) lack the genitive and dative. As before, however, the reason for the apparent lack of GEN and DAT case is not that the marking of GEN and DAT is prefixal; rather, the functions of GEN and DAT merge into one (suffixal) form together with ERG, called OBLIQUE by Colarusso. Thus, this state of affairs is similar to the language-wide syncretism of GEN and DAT discussed above for Elfdalian, and I have nothing new to add here.

Finally, I mention here Latin as an example of a language which has five suffixal cases (NOM, ACC, GEN, DAT and INS), but a prepositional marking for the comitative (cum).[4]

To conclude, as the examples discussed here suggest, the variation between suffixal vs. prepositional marking is subject to apparent regularity. This reg-

---

[3]See also Svenonius (to appear) for a discussion of the Elfdalian case system. Svenonius describes a different (more recent) variety of the dialect, but the relevant facts appear to be identical.

[4]Traditionally, the instrumental case is called ablative in Latin. However, when used on its own, this case expresses an instrument and not a source (leaving aside an interesting, but rather restricted class of names of towns, cities and small islands). I also ignore the vocative case (also more generally), although ultimately one would like to bring it into the picture as well.

ularity can be expressed as a linear ordering of cases (or case functions), such that a particular marking strategy (suffix vs. preposition) occupies a contiguous region on the scale. I repeat the relevant statement below, and take it to be a reasonable hypothesis with potentially universal validity, subject to further testing and refinement.

(8)   *The preposition/suffix hierarchy*
   a.   If the expression of a particular case in the Case sequence (below) involves a preposition, then all cases to its right do as well.
   b.   The Case sequence: NOM – ACC – GEN – DAT – INS – COM

2.2   Deriving the hierarchy

In this section, I turn to an implementation of (8). The essence is the idea that individual cases which make up the sequence (8-b) correspond to syntactic constituents, and they stand in structural subset/superset relations, identical to the ordering (8-b). The proposal is depicted as (9):

(9) $[_{com}$ F $[_{ins}$ E $[_{dat}$ D $[_{gen}$ C $[_{acc}$ B $[_{nom}$ 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]. Similarly, GEN = [A,B,C], etc. In the tree, the Case sequence relevant for the preposition/suffix hierarchy thus reappears as the sequence of non-terminal projections; the linear NOM–ACC–GEN–… turns into a hierarchichal [... [gen [acc [nom]]].

   I note already here that in languages where for instance, the instrumental case is expressed as a single morpheme, this morpheme will be taken to spell out a syntactic constituent containing (minimally) the features A-E (and possibly others such as number). Similar remarks apply to other case morphemes. The details of how constituents are spell out with number of features inside them is introduced later on.

   With the basic proposal in place, consider now the crucial question how a case morpheme ends up as a suffix on the noun. Assuming with Kayne (1994) that c-command maps onto linear precedence, we conclude that if nothing happens, case marking will be prefixal. Consequently, a case marker becomes a suffix only as a result of noun movement. In (10), I show what must happen if a particular case is to be expressed as a suffix: a constituent containing the noun must move to the left of all the features that a particular case is composed of.

(10) $[$ 6 $[_{com}$ F $[$ 5 $[_{ins}$ E $[$ 4 $[_{dat}$ D $[$ 3 $[_{gen}$ C $[$ 2 $[_{acc}$ B $[$ 1 $[_{nom}$ A DP ...

The numbers 1-6 indicate potential landing sites of the DP. 1 is the landing site which turns the NOM into a suffix. 2 represents the same position for the ACC and so on.

   The preposition/suffix hierarchy emerges when this picture is combined with the view that languages differ regarding the maximum height of noun movement (Cinque 2005). Thus, Cinque (2005) argues that the following examples are to be analyzed in terms of variable height of NP movement within an identical base-generated sequence, namely Dem > Num > A > NP:

7

(11) Variation in N movement: Czech, Farsi, Maasai, Kiitharaka[5]

a. ty   tři   velmi tlusté **knížky**
    those   *three*   *very* *thick* ***books***

b. un   seta   **ketabe** kheili koloft
    those   *three*   ***books*** *very* *thick*

c. kùndâ **mésai** àré   sàdan
    those ***tables*** *two*   *nice*

d. **mabuku** mara   mathatu   manene muno
    book *those*   *three*   *big*   *very*

Applying Cinque's observations and proposals to the empirical domain at hand, we expect that not all languages will move their DP to the top of the tree (10), i.e., all the way to 6. Such languages will split the features of the comitative case into two sets. Some features will be to the left of the DP, and some to its right. Features to the left will be spelled out as a preposition, features on the right as a suffix. Comitative in these languages is thus spelled out as P+DP+K. This way, the variation in the height of DP movement directly translates onto the variation in the inventory of case suffixes in a language.

Consequently, we generate a series of implications of the form: if a language expresses the genitive as a preposition (DP moves lower than 3), it also requires prepositions for all cases which are even higher up. Similarly, if a language expresses the dative as a preposition (DP moves lower than 4), it also expresses all the higher cases as prepositions. And this is exactly what the preposition/suffix hierarchy says.

2.3   Prepositions, postpositions and compound cases

In this section, I illustrate the working of the theory on empirical material. Thus, consider again the fact that in English, the expression of the genitive case involves a preposition with its complement in the accusative. In the preceding section, I have proposed to understand this in terms of case containment: the genitive structurally contains the accusative, and English shows this on its sleeve:

(12)
$$
\begin{array}{c}
\text{genitive} \\
\diagup\diagdown \\
C^0 \quad \text{accusative} \\
| \qquad \diagup\diagdown \\
\textit{of} \quad \quad \textit{him}
\end{array}
$$

Thus, in English, the DP does not cross the projection of the genitive case, as a consequence of which the genitive feature corresponds to a preposition. Now if we make sure that dative, instrumental etc. contain the genitive, as proposed in the preceding subsection, it follows that the presence of prepositional genitive will entail a prepositional dative. The reason is that if the DP cannot cross the projection of the genitive, then for any category containing the genitive, the marking will involve a preposition as well.

---

[5]Sources of data as follows. Czech: the author, Farsi: Marina Pantcheva (p.c.), Maasai: Koopman (2005, ex.3), Kiitharaka (diacritics omitted): Peter Muriungi (p.c.).

Under this view, English contrasts rather minimally with "rich case" languages where the genitive case marker includes the accusative case marker as its component part. In (13), I show a fragment of a Vlakh Romani case paradigm (the vocative is missing). Relevant is the fact that all the oblique cases are built on top of the accusative by means of an additional suffix:[6]

(13)   Vlakh Romani from Friedman (1991, p.57)

	boy, SG.	boy, PL.	girl, SG.	girl, PL.
NOM	čhav-ó	čhav-é	balvál	balval-á
ACC	čhav-és	čhav-én	balval-á	balval-én
GEN	čhav-és-koro	čhav-én-goro	balval-á-koro	balval-én-goro
DAT	čhav-és-ke	čhav-én-ge	balval-á-ke	balval-én-ge
INS	čhav-és-ar	čhav-én-car	balval-á-sar	balval-én-car
LOC	čhav-és-te	čhav-én-de	balval-á-te	balval-én-de
ABL	čhav-és-tar	čhav-én-dar	balval-á-tar	balval-én-dar

Under the approach taken here, the genitive case is derived from the accusative case by the addition of a feature, expressed as a separate morpheme in both Romani and English. This feature precedes the accusative constituent in English, but follows it in Vlakh Romani. I portray this difference as a matter of movement: the accusative DP stays below this additional feature in English (12), but it moves higher up in Vlakh Romani:[7]

---

[6]Such a decompositional paradigm structure as witnessed in Romani has led some researchers to argue that the markers which follow the ACC morpheme are in fact better classified as postpositions, but see Friedman (1991) for arguments for classifying them as case markers (in traditional terminology).Thus, Friedman notes that there are phonological interactions between the outer suffix and the accusative marker indicative of word-level phonology. Further, he notes that Romani is otherwise a prepositional language, with only a host of suffixal forms shown in the paradigm above. I take no stand on the issue; what matters is that Romani shows a pattern rather similar to English, modulo the order between the accusative base and the additional suffix.

[7]In order to broaden the database, I note that Vlakh Romai is not the only language that shows GEN-ACC containment; the accusative is contained inside the genitive in Kazakh (Pakendorf 2007, p.94), or the West Tocharian plural declension (Gippert 1987). Among the more familiar languages, a similar state of affairs seems to obtain in the plural of some Latin declensions, as illustrated below.
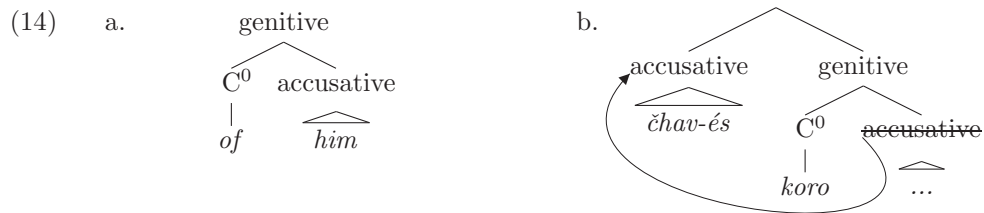
(i)   A subset of Latin plural declensions

|	star, PL.	slave, PL.	day, PL.
NOM	stell-ae	serv-ī	di-ēs
ACC	stell-**ās**	serv-**ōs**	di-**ēs**
GEN	stell-**ār**-um	serv-**ōr**-um	di-**ēr**-um
DAT	stell-īs	serv-īs	di-ēbus
INS	stell-īs	serv-īs	di-ēbus

The presentation presupposes the historical change of s to r in intervocalic positions, an independently attested process in Latin (e.g., in infinitives).

Potentially, there are problematic languages where the GEN-ACC containment comes out the other way round, but I believe that alternative analyses are available for the problematic cases. To the best of my knowledge, the counterexamples are exhausted by: certain Russian plurals (Timberlake 2004) and their analogues across Slavic, Lule Saami (data from Bruce Morén-Duolljá, p.c.), and Malayalam (Jayaseelan 2009). These data are interesting since they present a potential support for an alternative decomposition of case proposed in Pesetsky (2007), according to whom the accusative contains the genitive. See the paradigms below:

(ii)   Problematic paradigms in Russian and Lule Saami (partial)

(14)   a.   genitive
       C⁰ accusative
       |
       of *him*

  b.   accusative genitive
       čhav-és C⁰ ~~accusative~~
            |
            koro …

Proceeding to a new type of example, recall that in Arabic, NOM, ACC and GEN are expressed by suffixes; on the other hand, DAT includes the preposition *li* added on top of the genitive. Following the same logic as we did for English, the structure (15) emerges:

(15)   dative
        D⁰ genitive
        |
        li *muhammad-in*
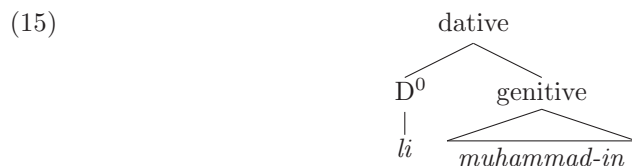
Repeating the crucial reasoning: as shown by the emergence of the preposition, the noun in Arabic does not cross the projection of dative. And since instrumental and comitative include the dative, they will likewise include a preposition.

Again, under this view, Arabic contrasts only minimally with languages where the expression of the dative case includes the genitive marker as its proper part. In the table below, I give as an example the singular declension of Ingush (Caucasian, Blevins 2009, p.203, attributed to Nichols 1994):

(16) Singular noun paradigm from Ingush

Russian	woman, PL.	Lule Saami	house, PL.
NOM	**komnat**-y	NOM	goad-e
ACC	**komnat**-y	ACC	**gåd-ij**-t
GEN	**komnat**	GEN	**gåd-ij**
PREP	**komnat**-ax	INE	**gåd-ij**-n
DAT	**komnat**-am	ILL	**gåd-ij**-da
INS	**komnat**-ami	ELA	**gåd-ij**-s

However, the Russian example potentially involves a phonologically empty affix. If that is so, it cannot be decided whether this affix is preserved in the ACC or not. This contrast with the evidence I drew from Latin, Romani and the other languages mentioned, where there is a dedicated accusative suffix, subject to further affixation. For Lule Saami, a potential analysis would follow the track of den Dikken (2006) and suggest that the genitive *gådij* is in fact a possessor of a noun (which is either silent, or expressed by *-t*), a situation independently attested in Ugro-Finnic.

As far as the situation in Malayalam goes, the claim seems to be based on a particular analysis that can be disputed. Thus, Rosmin Mathew (p.c.) tells me that the assumed genitive form is unable to surface as a possessor without additional affixation, and is in fact best described as a type of a stem. The stem surfaces without additional affixes only in compounds, hence the potential confusion, since Russian formations similar to compounds serve as one of Pesetsky's arguments for the claim that the accusative embeds the genitive. Similarly, the Malayalam grammar I have consulted (Asher and Kumari 1997) does not lend support to the view that the genitive form serves as the basis of the accusative.

	'hen'
NOM	kuotam
ERG	kuotam-uo
GEN	kuotam-a
DAT	kuotam-a-a
ALL	kuotam-a-ga
INS	kuotam-a-ca
LAT	kuotam-a-ğ
COMPAR	kuotam-a-l

Thus, if this proposal is on the right track, Arabic and Ingush datives differ only concerning the order of the genitive constituent and the feature D:[8]

(17) a.   dative

D⁰   genitive

li   *muhammad-in*

b.   genitive   dative

kuotam-a   D⁰   ~~genitive~~

a   ...
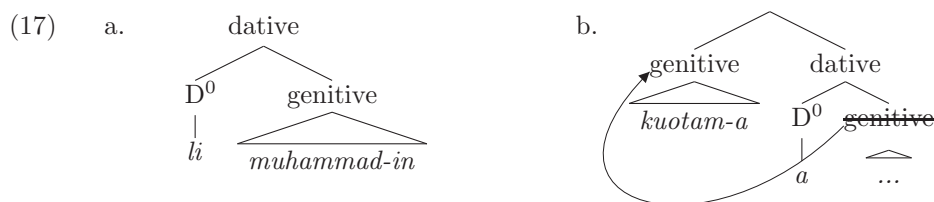
The same logic extends to the instrumental. Thus, languages such as Icelandic or German need a preposition to express an instrument, and this preposition attaches on top of the dative:

(18)   mit einem   Löffel
        with a.DAT.SG *spoon*
        'with a spoon' (German)

The structure for German is in (19-a); the DP does not cross the instrumental feature, and thus, the feature is expressed to the left of the DP as a preposition. Crucially for the preposition/suffix hierarchy, if the instrumental is contained

---

[8]Examples of other languages similar to Ingush include Djabugay (Australia, Embick 2008, p.96-7), Erźa Mordvin (Ugro-Finnic, McFadden 2004, p.241) or Estonian (Blevins 2008) (note that in the last two languages, the dative coincides with the illative and allative respectively, with the case form traditionally labeled as one of the latter).

Potential (and instructive) counterexamples (dative inside genitive) occur in some Australian languages (Aljawara, Andegerebinha, Arabana, Awabagal, Garadjari, Walbiri, see the appendix of Blake 1977). However, there are reasons to believe that these surface counterexamples are irrelevant, because two independent processes occurring more generally across Australian languages conspire to yield a surface form which is deceptive. The first thing to know is that in about 50 percent of Australian languages, the dative and genitive are not distinguished (Blake 1977, p.38). On the basis of this common scenario, I hypothesize that in the problematic languages, it is also in fact the case that GEN=DAT.

If that is so, the extra morpheme which occurs with the genitives is then to be analyzed as orthogonal to case marking, a line of analysis which is independently plausible in view of another process common in Australia, namely Suffixaufnahme (see Plank 1995, Austin 1995). This type of approach is independently supported in some of the problematic languages; thus, for example, in Aljawara, Andegerebinha and Arabana, the extra piece that derives the "genitive" from dative is a morpheme identical to the accusative. In Awabagal, the extra morpheme (*-pa*) is identical to a morpheme that occurs in the locative case (*kin-pa*), with locatives likewise known to get further agreement affixes in Australian. Thus, I think it is reasonable to regard these surface counterexamples as instances of GEN/DAT syncretism, with an additional (agreement like) morpheme attaching to the common GEN/DAT form inside NPs. This is in fact the same analysis as in the case of the Elfdalian (7-a).

inside the comitative, then a preposition will necessarily occur in the comitative if it appears in the instrumental.

(19) a. instrumental

    b.
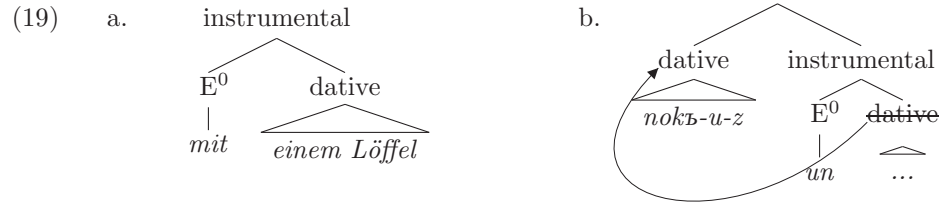
There are languages with suffixes rather than prepositions which replicate the German pattern. One such language is Budukh, with five sample paradigms in (20). Budukh thus contrasts with German only minimally, regarding the presence/absence of movement of the dative constituent, see (19-b).

(20) Sample paradigms from Budukh, spatial cases omitted (Dešeriev 1967)

	ground, sg.	ground, pl.	man, sg.	chest, sg.	3.sg.pron.
NOM	nokъ	nokъ-r-i	fur-i	gol-u	ad
GEN	nokъ-u	nokъ-r-a	fur-a	gol-a	an-da
ERG	nokъ-u-r	nokъ-r-a-ra	fur-ra	gol-ə-r	an-ə-r
DAT	**nokъ-u-z**	**nokъ-r-a-z**	**fur-a-z**	**gol-u-z**	**an-ə-z**
INS	**nokъ-u-z**-un	**nokъ-r-a-z**-ən	**fur-a-z**-en	**gol-u-z**-un	**an-ə-z**-ən

Finally, Latin has no comitative, and uses the preposition *cum* that attaches on top of an instrumental. A similar case is found in Georgian, where the comitative is formed by the postposition *gan* which follows the instrumental form (Stolz et al. 2009, p.607). I do not show the structures here to save space, since the general logic should be familiar by now.

2.4 Summary

In this section, I have first presented reasons to believe that the statement (21) correctly describes the attested and unattested patterns that pertain to the expression of case by prepositions and suffixes.

(21) *The preposition/suffix hierarchy*
    a. If the expression of a particular case in the Case sequence (below) involves a preposition, then all cases to its right do as well.
    b. The Case sequence: NOM − ACC − GEN − DAT − INS − COM

Then I have proposed that the hierarchy follows if individual cases are understood as syntactic constituents, related by containment relation dictated by the Case sequence (21-b). As a result, we arrive at the structure (22):

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

If this sequence is underlyingly present in all languages, then variation as well as the restrictions it obeys arise as a result of differences in the height of (extended) noun (phrase) movement within this sequence (drawing on proposals made in Cinque 2005):

(23) $[\,6\,[_{com}$ F $[\,5\,[_{ins}$ E $[\,4\,[_{dat}$ D $[\,3\,[_{gen}$ C $[\,2\,[_{acc}$ B $[\,1\,[_{nom}$ A DP ...

Thus, the variation in preposition/suffix inventory reduces to an independently needed syntactic operation, namely movement, which is known to vary in precisely the way needed for the present empirical domain.

In addition, the proposal neatly ties together the class of languages which express particular cases as P+case combinations, and languages where a similar pattern occurs in the postnominal position.

3   The analytic/synthetic division

Once the preposition/suffix division is understood in terms of movement *height*, a question immediately arises whether there is also variation in movement *type* (presence vs. absence of pied-piping), an option which is now predicted. The answer I suggest in this section is that indeed, there is such a variation, and it corresponds to the second parameter under investigation, namely the division between analytic (complex) and synthetic (simplex) expression of case.

To see what variation in movement type looks like, let me come back to the case discussed by Cinque (2005). Here, we find alongside the Kiitharaka example (24-a) also the Gungbe order (24-b):

(24)   a.   mabuku **mara** *mathatu* MANENE MUNO
            book   *those three*   *big*   *very*
       b.   àgásá ɖÀXÓ *àtɔ̀n*   **éhè** lɔ́ lɛ́
            crabs big   NUMERAL DEM SPF NUMBER
            'These specific three big crabs'   (Gungbe, Aboh 2004, p.92)

Kiitharaka and Gungbe do not contrast in the phrase-initial position of the noun. However, they differ in the order of the demonstrative (boldfaced), numeral (in italics) and adjective (in small caps). Whereas in Kiitharaka, the order is the base-generated Dem > Num > A, in Gungbe, it is the exact opposite. In Cinque's (2005) theory, this difference is analyzed in terms of differential pied-piping. Thus, in Kiitharaka, the noun phrase moves to the initial position alone, and triggers no pied-piping along the way. In Gungbe, pied-piping is triggered at every step. Thus, when the noun crosses the adjective, yielding [N A], it pied-pipes the adjective to the next higher landing site above the numeral, leading to the intermediate order [[N A] Num]. Then, the demonstrative is merged, and when the noun phrase moves to its left, it again pied-pipes the whole constituent along (this is also the dertivation proposed in Aboh 2004, p.92).

Now I proceed to show that precisely these two distinct movement types are the factor responsible for the analytic/synthetic distinction. Let me ilustrate this by means of an empirical example.

Thus, as we have seen above, the difference between Arabic on the one hand and Ingush on the other can be expressed as a difference in noun movement. Specifically, the noun stays below the dative in Arabic, but it moves above the dative in Ingush (pied-piping the genitive along):

(25)   a.

b.
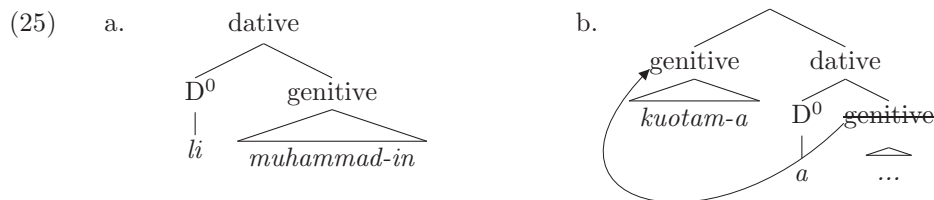
It is, however, very often the case that the dative case is not expressed analytically at all (namely as GEN plus an extra morpheme), but rather as an undecomposable suffix, see the Classical Armenian "synthetic" paradigms below:

(26)   *Classical Armenian dative is synthetic*

	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'

The derivation of such synthetic examples relies on the existence of phrasal spell out (see in particular Starke 2009a and related work by Ramchand 2008, Caha 2009b, Taraldsen to appear or Pantcheva to appear; see also McCawley 1968, Weerman and Evers-Vermeul 2002, Neeleman and Szendrői 2007 or Radkevich 2009). The essence of the proposal is that in Armenian – like in Ingush – the extended NP moves higher than the projection of dative; however, unlike in Ingush, it does not pied-pipe the genitive constituent along, see (27-a) for Armenian, and (27-b) for Ingush.

(27)   a.
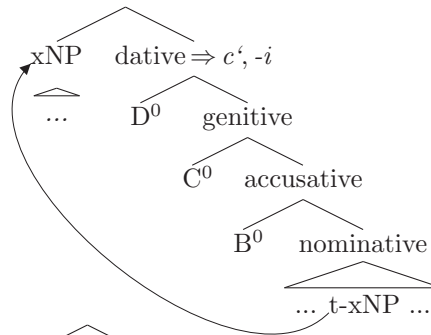
b.
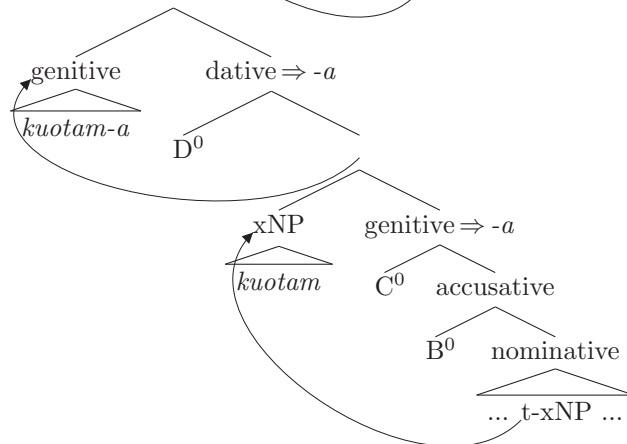
In the structure (27-a), the Armenian dative markers (-*i*, -*c'*) are then inserted to replace a whole phrasal constituent, containing the dative as well as the

genitive. This is indicated in (27-a) by the arrow which indicates phrasal lexical insertion.

In Starke (2009a), phrasal insertion is restricted to constituents. If that is so, there can be no synthetic marker of the Armenian type in Ingush, simply because there is no constituent that contains both D^0 and the genitive (to the exclusion of the extended NP). Consequently, derivations with pied-piping (such as (27-b)) lead necessarily to analytic morphology; derivations without pied-piping (such as (27-a)) give rise to synthetic morphology.

Thus, the difference in analytic/synthetic case morphology between Armenian and Ingush once again reduces to a difference in movement: pied-piping at the genitive in Ingush, and its absence in Armenian.

At this point, the first part of the paper finishes, and a more tentative enterprise is on its way. Before we embark on this new territory, let me sum up the discussion up to now.

At the beginning of the paper, I have focussed on two dimensions along which case morphology differs: (i) position, and (ii) complexity of the markers. At present, these dimensions are understood as (i) variation in the absolute height of noun movement within a universal sequence of case projections, and (ii) type of movement involved (pied-piping or not).

As already said before, this result has an obvious theoretical advantage over possible alternatives: it explains and constrains the observed variation in terms of a device (movement) which is independently known to vary precisely along the lines needed to deliver the facts. The additional interest of this particular example is the microscopic scale at which the variation is observed. Namely, it concerns the interaction of phrasal movement with units of syntactic structure smaller than morphemes. If on the right track, such analysis supports the emerging view on syntax according to which syntax is a completely pre-lexical devise, operating on units much smaller than lexical entries. To use a quote from Starke (2009a): "Syntax does not build on morphemes. Syntax builds morphemes."

4 Spell-out driven movement: motivations

A number of questions remain, as always. The one which is perhaps most pressing is what triggers the movements which are assumed to take place; why is it the case that the noun moves higher in Arabic than in English; why is the genitive pied-piped in Ingush but not in Czech; why do we get pied-piping at the genitive in Ingush but not at the accusative, etc.

In order to provide an answer to these issues, I will follow a recent proposal by Starke (2009b) to the effect that these parameters are controlled by material stored in the lexicon. In a nutshell, the idea is that since the Armenian lexicon contains an entry capable of spelling out all the dative features, the derivation proceeds without pied-piping so that the right type of constituent is derived. In Ingush, there is no such entry; consequently, pied-piping is triggered at the genitive, as a result of which the dative (and higher cases) can be lexicalized as a separate constituent by an additional morpheme.

These ideas are technically implemented in a model where each step of external merge is followed by lexical access, and lexical access then determines how the derivation proceeds. In particular, a successful insertion of a morpheme

at a phrasal node requires that the material to be lexicalized by a single morpheme forms a constituent. If there is a lexical entry such that spell out applies successfully at the root node, no movement takes place. However, if spell out fails, various types of evacuation movements may be required for insertion to take place. After these are performed, the structure is input to further external merge, and so on. The model is introduced in detail in the next section.

The purpose of this section is to provide some empirical justification for going down this road. The particular class of cases I will be interested in show variation in the height and type of movement within a single dialect, where the ultimate landing site as well as the type of movement depends on the class of the moving element. For example, as I discuss below in more detail, there is a dialect of Serbian where singulars move higher than plurals; this is not straightforwardly captured if (this type of) movement is triggered by some property of the landing site (a strong nominal feature, or a similar device), since both singulars and plurals are quite likely capable to satisfy a strong nominal feature. However, such variation is expected if (at least some) movement is dependent on lexicalization resources: since these turn out to be different for distinct classes of items (a situation known as allomorphy), a difference in the height and type of movement may arise as a consequence.

4.1 Variation in the height of movement within a single dialect

As the first example, consider the Bulgarian paradigm (28).

(28)   a.   Tazi duma m-i   e nepoznata.
             that word I-DAT *is unfamiliar*
             'That word is unfamiliar to me.'
     b.   Tazi duma e nepoznata {*Ø /na} sina mi.
             that word is unfamiliar   *to son my*
             'That word is unfamiliar to my son.'
     c.   Tazi duma e nepoznata {*Ø /na} Stojan.
             that word is unfamiliar   *to Stojan*
             'That word is unfamiliar to Stojan.'

In (28-a), we see an adjective whose argument is expressed by the dative clitic *mu*. (The clitic leaves the AP and appears adjacent to the verb, the clitic position in Bulgarian.) In (28-b) or (28-c), the argument is headed by a lexical noun, in which case a preposition must appear. (I show this both for a common noun (28-b) and a proper name (28-c) for reasons that will become clear shortly.) In the present framework, that means that clitics must move higher up than all non-clitics; an asymmetry that needs to be accounted for.

Similar contrasts between clitics and non-clitics can be reproduced in French and other Romance languages; however, identical asymmetries extend beyond the simple clitic/non-clitic distinction. For example, Hendriks (1976, p.106) points out that a dialect closely related to Bulgarian – the Radoža-Vevčani dialect of Macedonian – not only replicates the contrast between clitics (the dative *mu* in (29-a)), and common nouns, see (29-b), but in addition, a restricted class of masculine personal nouns in the singular (essentially proper names)

patterns with clitics, and has a dedicated dative suffix, see (29-a) again:[9]

(29) a. dai m-u   stöjan-u
          give him-DAT *stojan*-DAT
          Give (it) to Stojan!
      b. go   dáde na sín-a   mi
          it.ACC *gave to son*-ACC *my*
          'He gave it to my son.'

In the present framework, this again means that there is an asymmetry in terms of movement: clitics and personal proper names move higher up than the rest of nominal items.

   As another example, consider the case marking asymmetry from one of the Prizren-Timok dialect group of Serbian, see (30), reproduced from Sobolev (2009, p.723).

(30) One of the Prizren-Timok dialects

'woman'	SG	PL
NOM	žen-a	žen-e
ACC	žen-u	žen-e
GEN/DAT	žen-i	na žen-e

In this dialect, we again see a split that concerns suffixal/prepositional marking of dative arguments, with the difference that the split now runs between the singular and plural: singular nouns move above dative, plurals have to stay low.

   Moving on to perhaps more controversial instances of what I believe to be the same phenomenon, consider the curious behavior of certain Serbian proper names (noted first, to my knowledge, in Bošković 2006). In (31), I show the behavior of the indeclinable noun *Meri* (understood as the name of a horse), which shows up as such in NOM, ACC, GEN and DAT environments, see (31-a-d) respectively.

(31) Serbian (Bošković 2006, Monika Bader, p.c.)
      a. Meri   ide
          *Mary*NOM *walks*
          'Mary is walking'
      b. Uzgajać konja   je kupio MeriACC.
          breeder horses.GEN AUX *bought Meri*
          'The horse breeder bought Meri.'
      c. Bojao   sam se   Meri
          I.was.afraid AUX REFL *Meri*.GEN
          'I was afraid of Meri.'
      d. Petar je doneo Meri   slamu
          Peter AUX *brought Meri*.DAT *hay*.ACC
          'Peter brought hay to/for Meri.'

However, it is impossible to use this noun as an object of instrumental selecting verbs, such as *ovladati* 'govern.' This is shown in (32); (32-a) illustrates the fact

---

[9]The proper names that appear with the dative suffix may also appear as a PP, i.e., *na stöjana* 'to Stojan.ACC' would be also fine in (29-a). The point here is that other nouns have to include the preposition.

17

that the verb requires an instrumental object and (32-b) shows that the name *Meri* cannot be used here:

(32) a. Oni su   ovladali   Andorom
          they AUX *conquered Andorra*INS
          'They conquered Andorra.'
      b. *Džokej je   ovladao   Meri.
          jockey AUX *conquered Meri*
          'The jockey conquered Meri.'

The only way to make the structure licit with just the bare noun is to insert a preposition:

(33) Džokej je   ovladao   s(a) Meri.
      jockey AUX *conquered with Meri*
      'The jockey conquered Meri.'

A possible way to understand these facts is by proposing that nouns like *Meri* move only as high as dative, but no further. As a consequence, their paradigm is "defective:" a preposition must be inserted to spell out the additional feature.

(34) The defective paradigm of *Meri*

	Meri	Andora
NOM	Meri	Andor-a
ACC	Meri	Andor-u
GEN	Meri	Andor-e
DAT	Meri	Andor-i
INS	s(a) Meri	Andor-om

If this interpretation is correct, then once again, a movement asymmetry cuts across two distinct classes of Serbian feminine names.

4.2   Variation in the type of movement within a single dialect
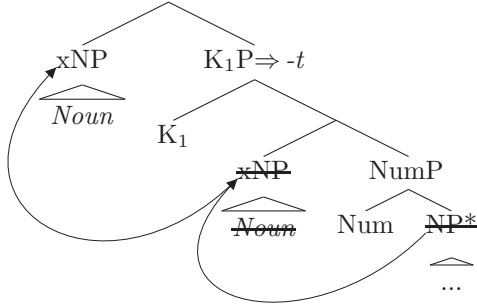
Similar variation also obtains for the second parameter of variation, namely pied-piping. This type of inter-language variation is best attested for the analytic/synthetic expression of case and some other inflectional category. Consider, for example, the following Finnish paradigms.
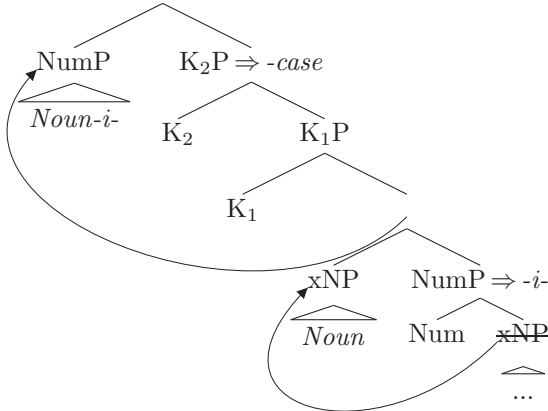
(35)   Finnish (from Kiparsky 2001)

	bear, sg.	bear, pl.
NOM	karhu	karhu-**t**
ACC	karhu-n	karhu-**t**
GEN	karhu-n	karhu-j-en
PART	karhu-a	karhu-j-a
ESS	karhu-na	karhu-i-na
TRANS	karhu-ksi	karhu-i-ksi
INE	karhu-ssa	karhu-i-ssa
ELA	karhu-sta	karhu-i-sta
ILL	karhu-un	karhu-i-hun
ADE	karhu-lla	karhu-i-lla
ABL	karhu-lta	karhu-i-lta
ALL	karhu-lle	karhu-i-lle
ABE	karhu-tta	karhu-i-tta

The table shows that number and case are synthetic in NOM and ACC, and analytic in the rest of the cases. In the present framework, where synthetic morphology corresponds to a cyclic movement without any pied-piping, that means that number is not pied-piped across case in NOM and ACC. However, it must be pied-piped in the oblique cases:

(36)   Structural case in Finnish:

xNP $\Leftarrow$ *Noun* $K_1P \Rightarrow$ *-t*
  K_1
    xNP $\Leftarrow$ *Noun*   NumP
      Num   NP*
        ...

(37)   Obliques in Finnish:

NumP $\Leftarrow$ *Noun-i-* $K_2P \Rightarrow$ *-case*
  K_2 K_1P
    K_1
      xNP $\Leftarrow$ *Noun*   NumP $\Rightarrow$ *-i-*
        Num   xNP
          ...

Examples where such variation in pied-piping must be acknowledged for projections within the case stack are somewhat harder to come by. However, there appear to be cases like that in my native Czech, see the tables below.

(38)   The singular/plural of three Czech nouns

	mouse, F	castl, M.	sea, N
GEN.SG.	myš-i	hrad-u	moř-e
DAT.SG.	myš-i	hrad-u	moř-i
GEN.PL.	myš-í	hrad-ů	moř-í
DAT.PL.	myš-í-m	hrad-ů-m	moř-í-m

The table shows paradigms where in the singular, both genitive and dative are undecomposable; in the plural, however, the dative is plausibly built on top of the genitive by means of an extra *m*. Within the framework proposed, that means that singular shows a cyclic movement of the noun (without any pied-piping), while in the plural, the genitive is pied-piped when the noun moves above the dative (as in Ingush).

4.3   Conclusions

If the general approach pursued here is correct, then – as the examples discussed indicate – the absolute height of movement as well as its type are not parameters that are fixed for the language as a whole. For example, clitics commonly contrast with full nouns, but finer distinctions exist. To give an example again: singulars and plurals may differ both in height of movement (as they do in one of the Prizren Timok dialects) and in the type of movement (Czech).
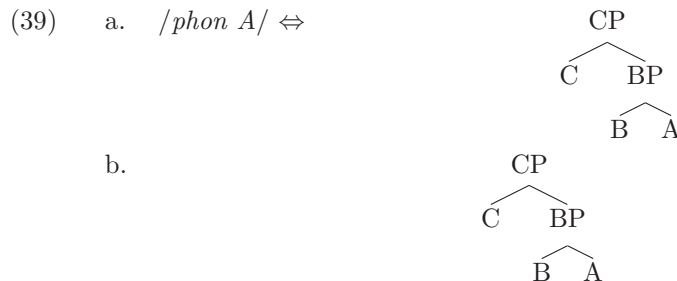
If that is the case, we need a theory of movement which allows for such a variation. Yet, we would also prefer if the theory does not achieve this goal at the cost of allowing too much freedom. The following section presents a theory with these two properties (due to Starke 2009b). In this approach, the way movement proceeds is dependent on lexical resources.

5   Spell-out driven movement: an implementation

In the model to be introduced, movement is triggered by phrasal spell out. I thus start by introducing the basics of such a spell out procedure, relying on Starke (2009a).

5.1   Phrasal spell out

Suppose that (39-a) is a lexical entry (a pairing of sound and syntax, with conceptual information ignored), and (39-b,c) are a syntactic structures.
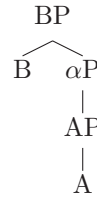
(39)   a.   */phon A/* ⇔

CP
├── C
└── BP
     ├── B
     └── A

   b.

CP
├── C
└── BP
     ├── B
     └── A

c.
$$\text{BP} \overbrace{} \quad \text{B} \quad \text{A}$$

The right part of the entry (39-a), called "Lexically Stored Tree" (LST) is identical to the structure (39-b). If syntax constructs (39-b), the entry (39-a) applies at the root node, as a result of which this structure is pronounced as *phon A*.

As Starke (2009a) proposes, the same can happen in (39-c). Even though the syntactic structure is not identical to the LST, the LST in fact contains a part that is identical to it. Such a matching condition is called the Superset Principle (in analogy to the Subset Principle of Distributed Morphology), because the lexical entry applies to any node it is a superset of.
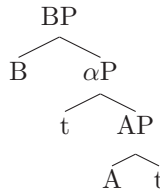
(40)   *The Superset principle (Starke 2009a)*: A lexically stored tree matches a syntactic node iff the lexically stored tree contains the syntactic node.

The insertion procedure ignores traces. Thus, assume that (41-a) is a lexical entry, and (41-b,c) are structures where (i) the complement of A has moved out, and (ii) a phrase has extracted out of the position immediately dominated by αP. I use the notation α, since I wish to remain neutral on certain issues that will emerge later on. In any case, the entry (41-a) matches both (41-b) and (41-c) if traces are ignored by matching.

(41)   a. $/phon\ B/ \Leftrightarrow$
                                                  BP
                                         B αP
                                              AP
                                              A

      b.   BP / B αP / t AP / A t

      c.   AP / A t

Note, however, that (41-a) does not match any of (40-b,c), since it does not contain those trees (it has the non-branching AP node in addition). Similarly, (40-a) does not match any of (41-b,c).

5.2   Cyclic spell out

With these entries in place, suppose that the derivation merges together the features A nd B, forming a BP:

(42)   BP / B A

In the model proposed by Starke (2009b), each step of merge is followed by lexical access. Lexical access proceeds bottom up, and insertion is thus first attempted at each of the terminals B and A. Each of the terminals is contained in the LST for (39-a), and hence, each terminal is spelled out by this entry.

Crucially, spell out then proceeds to the phrasal node BP. This node is contained in the LST in (39-a), and thus, /phonA/ is inserted at BP. Insertion at a non-terminal node causes the non-pronunciation of any material contained in the node; in this case, the material previously inserted at A or B is "overridden" by the spell out of BP. The consequence is that spell out of a non-terminal, if successful, blocks any potential analytic expression; famoulsy, *went* blocks *go-ed*.

The derivation then continues by merging C, forming a CP:

(43)
$$
\begin{array}{c}
\text{CP} \\
\overbrace{\text{C}\quad\text{BP}} \\
\qquad\overbrace{\text{B}\quad\text{A}}
\end{array}
$$

Since spell out proceeds bottom up, we first have to spell out C (the spell out of BP has already been discussed). C is a trivial sub-constituent of the LST in (39-a), making it possible for insertion to take place. Proceeding to CP, we again may use (39-a), with the result that any material previously used is deleted. (43) is thus pronounced as /phon A/.

5.3 Evacuation movements

The derivation then continues in a similar fashion, deriving eventually an XP. Suppose now that XP merges with the feature A, yielding an AP:

(44)
$$
\begin{array}{c}
\text{AP} \\
\overbrace{\text{A}\quad\text{XP}} \\
\qquad\overbrace{\ldots}
\end{array}
$$

Spell out first targets lower nodes, as described above, arriving eventually at AP. The root node AP finds no match in the lexicon. However, there is a potential candidate in the entry (41-a), repeated below, which contains AP.

(45)   /*phon B*/ ⇔
$$
\begin{array}{c}
\text{BP} \\
\overbrace{\text{B}\quad\alpha\text{P}} \\
\qquad\mid \\
\qquad\text{AP} \\
\qquad\mid \\
\qquad\text{A}
\end{array}
$$

The AP in this entry, however, contains only one of the daughters, the feature A. This entry, as Starke (2009b) proposes, triggers an evacuation movement of XP. I state it below as (46):

(46)   *Spell out driven movement:* In case a phrasal node can be spelled out after the evacuation of a sub-constituent, then evacuation takes place.

As a consequence of (46), XP extracts, yielding (47).

(47)

αP
XP   AP
... A   t

Here, the notation αP comes in play, since I wish to remain neutral concerning the labeling of the root node in (47); it may be an AP, or XP, depending on considerations that are not crucial for the discussion.

After the evacuation movement, AP is spelled out as (45). Note that the material used to spell out XP is not over-ridden, since XP is not contained inside AP.

5.4   Movement without pied-piping

After the evacuation movement, the structure is input to further external merge. Thus, suppose that B merges with (47), yielding (48):

(48)
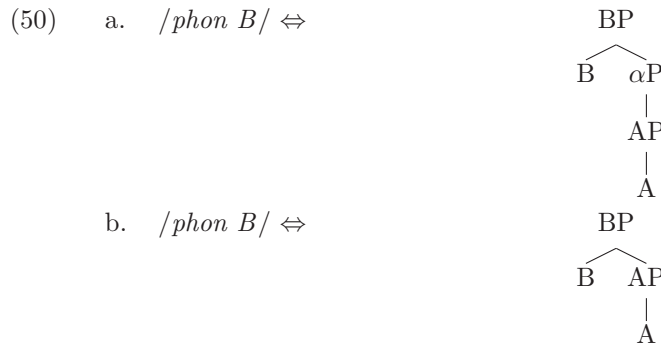
BP
B αP
   XP   AP
   ... A   t

Again, BP finds no match in the lexicon, due to presence of XP. However, if XP extracts, BP could be spelled out by (45). And thus again, an evacuation movement is triggered:

(49)

βP
XP   BP
...   B αP
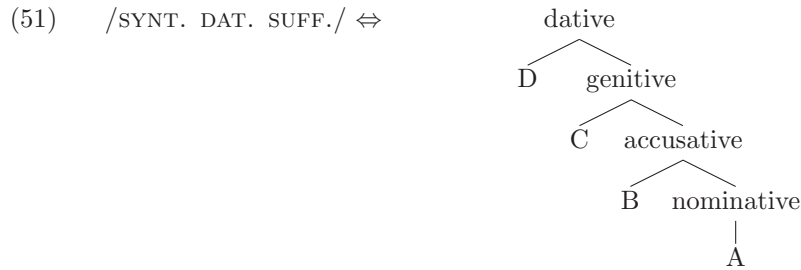   t   AP
            A   t
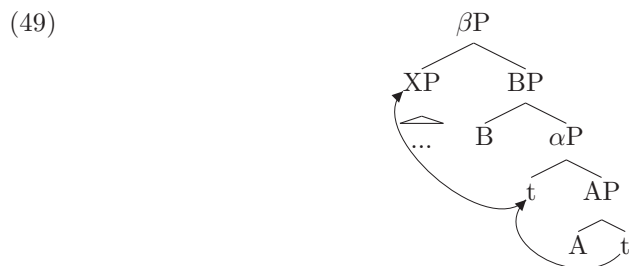
As is now clear from inspection, the derivation (49) is largely equivalent to a structure of successive-cyclic XP movement which triggers no pied-piping. Note that cyclic movement here takes place within a stretch of a functional sequence which is to be spelled out by a single entry, (45).

As a matter of convenience, from now on, I will be omitting in lexical entries the nodes labeled by Greek letters. Thus, (50-a) becomes (50-b). The purpose of the abbreviated entry (50-b) is to avoid clutter, the important point being that the entry (50-b) forces a successive cyclic evacuation of the original sister of A.
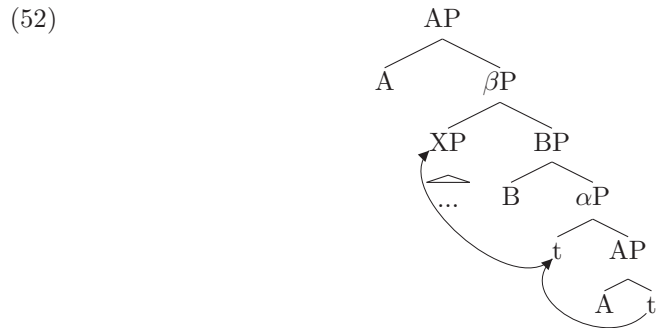
(50) a. /*phon B*/ ⇔

```
              BP
             /  \
            B    αP
                 |
                 AP
                 |
                 A
```

  b. /*phon B*/ ⇔

```
              BP
             / \
            B   AP
                |
                A
```

Returning to empirical facts, the result is that the entry of a synthetic dative suffix (of the Armenian type) looks as follows:

(51) /SYNT. DAT. SUFF./ ⇔

```
              dative
             /    \
            D    genitive
                 /    \
                C    accusative
                     /    \
                    B   nominative
                        |
                        A
```

This entry will force successive cyclic movement of the extended NP out of the nominative, leading eventually to its complete extraction out of the dative. This is the required result, since the dative is suffixed in Armenian.
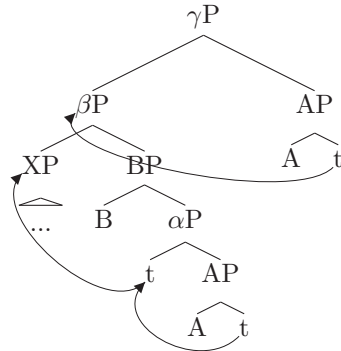
5.5 Roll up

Once we move beyond a single entry, we get a roll up type of movement. Thus, suppose now that our βP (derived in (49)) is embedded under the feature A:

(52)

```
              AP
             /  \
            A    βP
                /  \
              XP    BP
             / \   /  \
            ...  B    αP
                     /  \
                    t    AP
                        / \
                       A   t
```
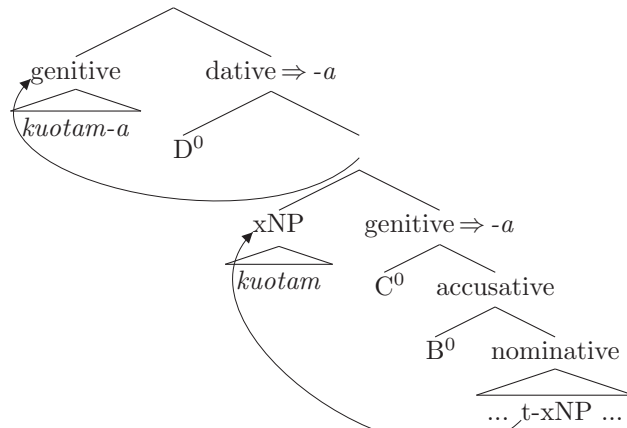
In (52), the root node AP can be spelled out by (50-b) in case βP extracts. Given the possibility to spell out the root node if an evacuation movement takes place, the movement is triggered:
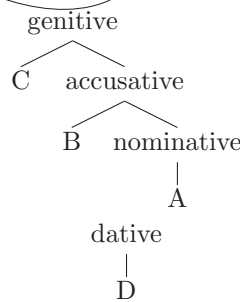
(53)

γP

βP   AP

XP   BP   A   t

...   B αP

t   AP

A   t

What we get is a cyclic movement of XP inside βP, and a roll up (XP pied-pipes βP) when we reach boundaries between lexical entries. This means that the derivation in Ingush, repeated below, is triggered by the combination of the entries in (55).

(54)

genitive   dative $\Rightarrow$ -*a*

kuotam-a   D^0

xNP   genitive $\Rightarrow$ -*a*

kuotam   C^0 accusative

B^0 nominative

... t-xNP ...

genitive

C   accusative

B   nominative

A

dative

D

(55)   a.   /SYNT. GEN. SUFF., -*a*/ $\Leftrightarrow$

     b.   /ANAL. DAT. SUFF., -*a*/ $\Leftrightarrow$

First, the extended NP moves out of the genitive without pied-piping, since this stretch of the functional sequence is covered by a single entry. The whole genitive then needs to be pied-piped out of the dative, so that (55-b) can successfully apply at the dative phrasal node.

     Note that on the general level, the system developed up to now leads to the consequence that evacuation movements (both cyclic and roll up) always include the head of the extended projection, the noun in the particular case at hand. That is because in order to create the right type of constituent for lexicalization, the evacuation movement has to remove the complement of the phrasal morpheme in full, including crucially the head of the projection.

This consequence is a welcome result. In particular, the theory of Cinque (2005) (as well as its alternative explored in Abels and Neeleman 2009) requires this as one of the constraints on the relevant class of movements. If the theory presented here turns out to be tenable in a broader perspective, its additional merit is that it derives this restriction.[10]
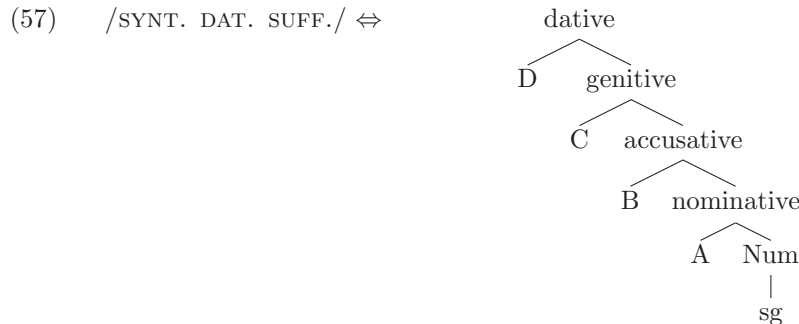
5.6 A mixed language: Czech

At this point, we are ready to tackle languages where one class of nouns has a synthetic marking (cyclic movement), and another class shows analytic marking (roll up). I have mentioned Czech as an example, with the data repeated below.

(56) The singular/plural of three Czech nouns

	mouse, F	castl, M.	sea, N
GEN.SG.	myš-i	hrad-u	moř-e
DAT.SG.	myš-i	hrad-u	moř-i
GEN.PL.	myš-í	hrad-ů	moř-í
DAT.PL.	myš-í-m	hrad-ů-m	moř-í-m

For the singular, we need a synthetic suffix, which I give in (57), ignoring the distinctions between the three classes in the table above. The fact that the application of the entry is limited to the singular is expressed by proposing that it spells out the (set of) number projection(s) responsible for the singular interpretation:

(57)   /SYNT. DAT. SUFF./ ⇔

$$
\begin{array}{l}
\text{dative} \\
\diagup\diagdown \\
\text{D} \quad \text{genitive} \\
\qquad \diagup\diagdown \\
\qquad \text{C} \quad \text{accusative} \\
\qquad\qquad \diagup\diagdown \\
\qquad\qquad \text{B} \quad \text{nominative} \\
\qquad\qquad\qquad \diagup\diagdown \\
\qquad\qquad\qquad \text{A} \quad \text{Num} \\
\qquad\qquad\qquad\qquad\quad | \\
\qquad\qquad\qquad\qquad\quad \text{sg}
\end{array}
$$

This affix will trigger cyclic evacuation of the extended NP, just in case the NP is constructed with the singular number. Note that as a result of the singular specification in (57), this affix is inapplicable in the plural, and hence, we need additional entries.

The two entries which I assume (still ignoring the class differences) are below:

---

[10]I add, however, that there is an additional issue to be dealt with, namely examples of straight morpheme orders in postnominal position (e.g., the order N-K-Num of Classical Armenian, see Caha (2009a) for a detailed discussion of the empirical facts). Such orders are quite rare, and they have been in focus because they violate Baker's (1988) Mirror Principle. The technology for such cases will have to somewhat weaken the initial setup, and allow for derivations where the complement of (phrasal) K does not evacuate in full, but only the actual NP moves, leaving Num in situ. I cannot deal with this case for reasons of space, as its treatment requires the introduction of an additional device, the so-called pointer (Starke 2009b).

(58)   a.   /SYNT. GEN. SUFF., *-í, -ů̊*/ ⇔

genitive
C   accusative
B   nominative
A   Num
│
pl

   b.   /ANAL. DAT. SUFF., *-m*/ ⇔

dative
│
D

These entries will force cyclic movement to the left of the genitive, and then a step of a roll up movement across the dative, as required by the data (see also the discussion of Ingush in the preceding subsection).
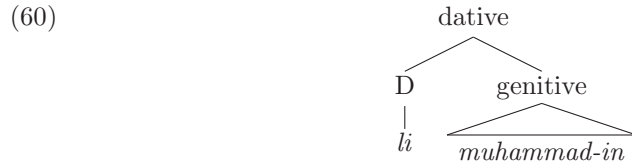
The only thing that needs to be taken care of is that we do not get a roll up derivation in the singular. Thus, suppose that the derivation in the singular has reached the stage where the constituent corresponding to the genitive singular has been spelled out (having triggered an evacuation movement). Then, the feature D is merged, yielding the dative singular:

(59)

dative
D^0 αP
xNP   genitive ⇒ *-phon*
⌒
...   C   accusative
... t-xNP ...

When spell out targets the dative, two options arise. (i) A cyclic movement of xNP is triggered, and (57) is inserted at the root node of (59). (ii) Pied-piping of αP is triggered, and (58-b) is inserted at the root node. The latter option is empirically incorrect, and must be ruled out. In order to achieve such and similar effects, the theory presented by Starke (2009b) contains a ranking of movement types, to the effect that no pied-piping is preferred over pied-piping. As a consequence, (i) is chosen over (ii), and the right result (a synthetic dative) emerges.
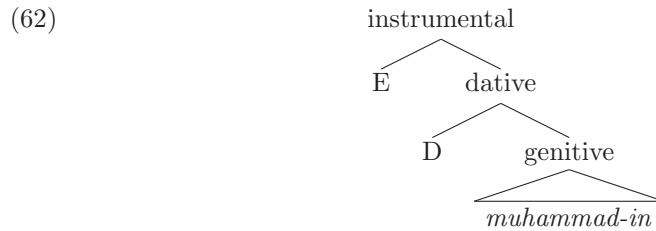
5.7   Prepositions

So far, we have seen the derivation of various types of suffixal marking, but it is not yet clear how movement may fail to obtain. The simplest example would be one where there is no lexical entry whatsoever containing the root node of the derivation. The following could be an example. In Arabic, there is a genitive suffix which forces the displacement of the NP above the genitive. Suppose that this much has been accomplished in the derivation. When we add the feature D, forming a dative, the preposition *li* is first inserted under the terminal D:

(60)

$$\text{dative}$$

```
              dative
              /    \
            D      genitive
            |      /      \
            li    muhammad-in
```

(61) $/li/ \Leftrightarrow D$

If there is no entry containing the "dative" node, there will be no reason for any displacement, and (60) would be successfully derived. While the final picture will have to be more complicated than this, the general insight – which I will preserve – is that if there is no entry containing a given phrasal node, this will lead to the lack of movement.

The complication we now have to deal with is the fact that in the instrumental, the structure of which is in (62), the instrumental preposition *bi* attaches to the genitive, and not on top of the dative PP. This means that *bi* cannot be inserted under E alone, but it must in fact spell out a constituent with both D and E.

(62)

```
           instrumental
            /        \
          E         dative
                    /      \
                   D      genitive
                          /      \
                       muhammad-in
```

The problem is then simply the fact that the most obvious way to derive this constituent – namely extracting the genitive out of the instrumental – will lead to a postposition, rather than a preposition.

The way I will deal with this situation here is by proposing that the entry of *bi* is as given in (63):

(63) $/bi/ \Leftrightarrow$

```
            E
           / \
          D   E
```
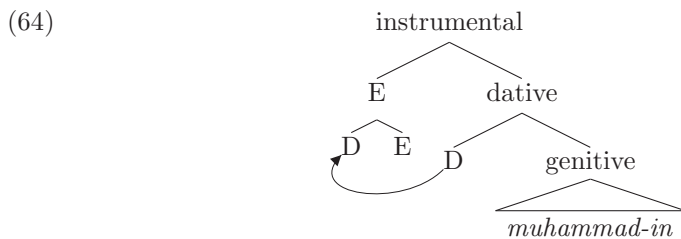
This entry has the crucial property that it lacks the projection of the feature D, namely the dative constituent. As a result of that, it will not force any movement in (60). Similarly, there is no point evacuating the genitive in (62), simply because (63) would still not be a match for the instrumental node in (62) (see the discussion following (41)).

Instead, what the entry (63) leads to is a structure that is traditionally understood to arise by head-movement. In other words, when the derivation reaches the stage (62), head-movement is triggered by the entry (63), yielding (64) as an output.[11, 12]

---

[11]M. Starke (p.c.) suggests that the constituent corresponding to [D E] is in fact base-generated, and moves cyclically as a chunk already from the lowest position. I leave the details of the proposal aside for reasons of space.

[12]A possible concern about this proposal is that head-movement reverses the base-generated order of the features D and E without an accompanying NP movement. If D and E were each lexicalized by a separate morpheme, this option would lead to an unattested type of ordering

(64)

                    instrumental
                   /   \
                  E   dative
                 / \   /   \
                D   E   D   genitive
                                   /   \
                              muhammad-in

Repeating what is crucial: the lack of a lexical entry for a given non-terminal causes the switch from suffixal to prepositional marking. This is the last parameter needed.
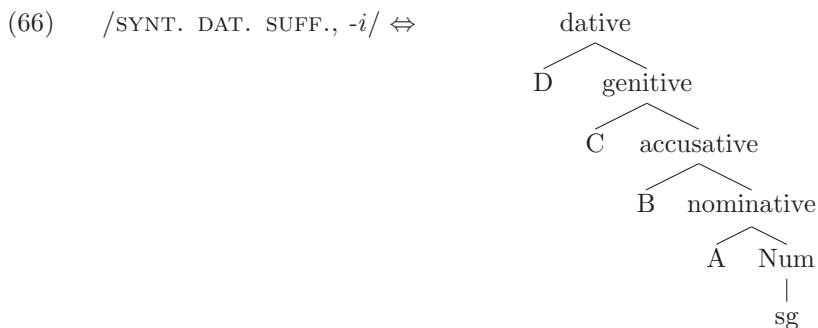
5.8 A mixed dialect of the Prizren-Timok group

To show how the proposal developed up to now handles data from a language with variable height of movement, recall the paradigm (65):

(65)   One of the Prizren-Timok dialects

'woman'	SG	PL
NOM	žen-a	žen-e
ACC	žen-u	žen-e
GEN/DAT	žen-i	na žen-e

To correctly model this dialect in the system proposed, we end up with the following entry for the dative singular:

(66)   /SYNT. DAT. SUFF., *-i*/ ⇔

                              dative
                             /   \
                            D   genitive
                                  /   \
                                 C   accusative
                                      /   \
                                     B   nominative
                                          /   \
                                         A   Num
                                                   |
                                                   sg

This entry will force a series of evacuation movements of the noun phrase, with the result that the phonological exponent will end up suffixed.

In the plural, we have the accusative marker *-e*, see (67-a), and the preposition *na* with the entry in (67-b):

(see Cinque 2005 for a general discussion and Caha 2009b, §2.6.2 for a discussion of ordering in the domain of case morphemes). However, in the case at hand, such a movement is only allowed in case D and E are spelled out by a single morpheme, which renders their ordering opaque. Importantly, no such movement will be triggered when D and E are spelled out separately, and thus, admitting head movement in such a constrained fashion does not, in fact, lead to orderings incompatible with the generalizations observed in the literature.

(67) a. /SYNT. ACC. SUFF., *-e*/ ⇔

accusative
├── B
└── nominative
    ├── A
    └── Num
        └── Pl

  b. /DAT. PREP., *na*/

D
├── D
└── C

In sum, there is no phrasal entry that can apply at the genitive node in the plural (note that (66) does not qualify), as a result of which the noun stops moving below this node, and a preposition appears.

5.9 Summing up

To sum up the mechanics in one paragraph: each external merge is followed by lexical access. If the root node created by external merge is directly contained in a lexical entry, no movement takes place. Similarly, if there is no entry containing the root node, movement does not happen either. However, if the root node can be matched by a phrasal lexical entry after a step of movement, then such a movement is triggered. This may be either a roll up movement, a cyclic movement without any pied-piping, or a head-movement type of operation, depending on the particular entry we are dealing with.

6   Conclusions

If on the right track, the proposal leaves us with the general conclusion that variation in movement height and type (within an extended projection) can be reduced to the expected source of variation, namely the lexicon.

In addition, there are two potential advantages over an alternative where this type of movement is driven by syntax-internal devices (features). The first advantage is empirical: spell out driven movement allows us to deal with interesting cases where variation in movement obtains within a single dialect.

The second advantage is theoretical, and admittedly tentative. In particular, the proposal holds the promise of deriving the requirement that the relevant type of movement always has to move the head of the extended projection. The leading idea is this: since the purpose of evacuation is to create a separate constituent, the complement must be removed in full, including the head noun. As highlighted above, whether this result can be maintained in view of mirror violating orders, remains a question for future work.

References

Abels, Klaus and Ad Neeleman. 2009. Universal 20 without the LCA. In *Merging Features: Computation, Interpretation and Acquisition*, edited by José M. Brucart, Anna Gavarró, and Jaume Solà. Oxford University Press, Oxford.

Aboh, Enoch Oladé. 2004. *The morphosyntax of complement-head sequences. Clause structure and word order patterns in Kwa*. Oxford University Press, Oxford.

Asbury, Anna. 2008. *The morphosyntax of case and adpositions*. Ph.D. thesis, Universiteit Utrecht.

Asher, R. E. and T. C. Kumari. 1997. *Malayalam*. Descriptive Grammars. Routledge, London.

Austin, Peter. 1995. Double case marking in Kanyara and Mantharta languages, Western Australia. In *Double case: Agreement by Suffixaufnahme*, edited by Frans Plank, pp. 363 – 379. Oxford University Press, New York.

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

Blake, Barry J. 1977. *Case marking in Australian languages*. Australian institute of Aboriginal studies, Canberra.

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

Blevins, James P. 2008. The structural determination of case and agreement. *Linguistica Uralica* 44 4: 241–267.

Blevins, James P. 2009. Case and declensional paradigm. In *The Oxford Handbook of Case*, edited by Andrej Malchukov and Andrew Spencer, pp. 200 – 218. OUP, Oxford.

Bošković, Željko. 2006. Case checking versus case assignment and the case of adverbial nps. *Linguistic Inquiry* 37: 522–533.

Caha, Pavel. 2009a. Classical armenian declension. In *Nordlyd 36: Special issue on Nanosyntax*, edited by Peter Svenonius, Gillian Ramchand, Michal Starke, and Tarald Taraldsen, pp. 77–112. University of Tromsø, Tromsø. Available at www.ub.uit.no/munin/nordlyd/.

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

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

Cinque, Guglielmo. 2005. Deriving Greenberg's universal 20 and its exceptions. *Linguistic Inquiry* 36: 315 – 332.

Cinque, Guglielmo and Luigi Rizzi. 2008. The cartography of syntactic structures. *STiL – Studies in Linguistics, CISCL Working Papers* 2: 42 – 58.

Colarusso, John. 1992. *A grammar of the Kabardian language*. The University of Calgary Press, Calgary.

Dahl, Östen and Maria Koptjevskaja-Tamm. 2006. The resilient dative and other remarkable cases in Scandinavian vernaculars. *Sprachtypologie und Universalienforschung (STUF)* 59: 56–75.

Dešeriev, Ju. D. 1967. The Budukh language [in russian]. In *The Languages of the Peoples of the USSR*, edited by V.V. Vinogradov, vol. 4, pp. 643–659. Nauka, Moscow.

den Dikken, Marcel. 2006. When Hungarians agree (to disagree) – The fine art of 'Phi' and 'Art'. Ms. CUNY.

Embick, David. 2008. Localism versus globalism in morphology and phonology. Ms., University of Pennsylvania.

Friedman, Victor A. 1991. Romani nominal inflection: Cases or postpositions? In *Problemy opisu gramatycznego języków słowiańskych, (Studia gramatyczne, Vol. 11)*, pp. 57–63. Polish Academy of Sciences, Warszaw.

Gippert, Jost. 1987. Zu den sekundären Kasusaffixen des Tocharischen. *Tocharian and Indo-European Studies* 1: 22–39.

Hendriks, P. 1976. *The Radoža-Vevčani dialect of Macedonian*. The Peter de Ridder Press, Lisse.

Jayaseelan, K. A. 2009. Case movement and theories of case. Paper read at The Third Workshop of the International Research Project on Comparative Syntax and Language Acquisition, Nanzan University, March 12.

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. 2001. Structural case in Finnish. *Lingua* 111: 315–376.

Koopman, Hilda. 2005. On the parallelism of DPs and clauses: Evidence from Kisongo Maasai. In *Verb First*, edited by Andrew Carnie, Sheila Dooley, and Heidi Harley, p. 281301. John Benjamins, Philadelphia.

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.

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

Moravcsik, Edith A. 2009. The distribution of case. In *The Oxford Handbook of Case*, edited by Andrej Malchukov and Andrew Spencer, pp. 231–245. OUP, Oxford.

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

Nichols, Johanna. 1994. Ingush. In *The Indigenous Languages of the Caucasus, vol. 4: Northeast Caucasian Languages*, edited by Rieks Smeets, pp. 79–145. Caravan Books, Delmar, NY.

Pakendorf, Brigitte. 2007. *Contact in the prehistory of the Sakha (Yakuts): Linguistic and genetic perspectives*. Ph.D. thesis, Universiteit Leiden.

Pantcheva, Marina. to appear. The syntactic structure of locations, goals, and sources. *Linguitics* .

Pesetsky, David. 2007. Undermerge...and the secret genitive inside every Russian noun. Handout of a talk at FASL 16.

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

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.

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

Sobolev, Andrej. 2009. Case and declensional paradigm. In *The Oxford Handbook of Case*, edited by Andrej Malchukov and Andrew Spencer, pp. 716–729. OUP, Oxford.

Starke, Michal. 2009a. 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. 2009b. Nanosyntax class lectures. Fall 2009, University of Tromsø.

Stolz, Thomas, Cornelia Stroh, and Aina Urdze. 2009. Varieties of comitative. In *The Oxford Handbook of Case*, edited by Andrej Malchukov and Andrew Spencer, pp. 601–608. OUP, Oxford.

Svenonius, Peter. to appear. An Övdalian case system. In *Studies in Övdalian Syntax*, edited by Kristine Bentzen and Henrik Rosenkvist. John Benjamins, Amsterdam.

Taraldsen, Tarald. to appear. The nanosyntax of nguni noun class prefixes and concords. *Lingua* .

Timberlake, Alan. 2004. *A reference grammar of Russian*. Cambridge University Press, Cambridge.

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