Notes on insertion in Distributed Morphology and Nanosyntax

Pavel Caha

Masarykova univerzita Brno

pcaha@mail.muni.cz

0. Abstract

This paper compares insertion in two realizational theories of morphology, Nanosyntax (NS, Starke 2009) and Distributed Morphology (DM, Halle and Marantz 1993). I argue that the insertion principle used in DM is incapable of governing insertion at non-terminal nodes, and should be therefore replaced by the Superset Principle used in NS.

1. Three differences between NS and DM

The theories of Nanosyntax (cf. Baunaz and Lander this volume) and Distributed Morphology (cf. Harley and Noyer 1999; Embick and Noyer 2003; Bobaljik 2015) provide each a particular theory of how to get from the basic building blocks of language to complex sentences. In practice, the research mostly focuses on linguistic units of smaller size (the traditional words or small phrases), but both theories share the idea that a general theory of syntax reaches deep into the structure of such units.

It is clear that any such theory is going to have multiple components, and the two theories diverge on a number of important questions. For instance, an important difference (**dif 1**) concerns **the nature of the basic building blocks**. Distributed Morphology projects its syntactic structures from complex objects that correspond to pre-packaged feature structures (bundles) taken directly out of the pre-syntactic lexicon. The main architectural claim of Nanosyntax is opposed to this; the idea is that the only component of grammar capable of constructing complex feature structures is syntax, which thus has to start from individual

(atomic) features. This leads to a syntax with large trees, increased number of movement steps and a quite different outlook than the tree structure representations used within DM. NS also tries to dispense with head movement, which leads to an increase in phrasal movement and remnant movement derivations. As highlighted in the introduction to this volume, Nanosyntax shares these features with Cartography and other approaches.

Another difference (**dif 2**) is the use of **phrasal spell out**, which in NS represents one of the core analytical tools, while at the same time being rarely used and often argued against in DM, as in Embick and Marantz (2008), or most recently in Embick (2014).[1] The core idea behind phrasal spell out is that pronunciation is like other operations in natural language (movement, ellipsis, etc.) in that it targets potentially non-trivial constituents.

This aspect of the NS theory is closely related to the fact that morphemes usually express multiple features, which for NS automatically means that they express multiple heads. Such morphemes will be called here portmanteau morphemes. Portmanteau morphemes must also be dealt with in DM (and other theories), and indeed DM has at least two tools to accommodate portmanteaus, including (i) the post-syntactic operation of Fusion and (ii) zero morphemes accompanied by conditioned allomorphy or Readjustment Rules.

[1] For clarity, let me mention that phrasal spell out is a type of non-terminal spell out (phrase is a type of a non-terminal node). I mention this here because there seems to be some confusion about this in the literature. For instance, Haugen and Siddiqi (2013) write that phrasal spell out „inserts entire phrases into non-terminal nodes rather than just single Vocabulary Items.“ This is incorrect, phrasal spell out inserts Vocabulary Items into phrasal nodes.

Please note that what NS calls “phrasal spell out” is often best recast in DM as the spell out of a non-terminal inside a complex head. This is because NS often uses phrasal movement derivations where DM uses head movement, and so for reasons of cross-theoretical comparison, it is best to understand “phrasal spell out” to actually mean “non-terminal spell out.”

This paper probes deeper into the particular issue of portmanteaus, because it seems that in this particular domain, a meaningful comparison can be made. I argue that neither Fusion (section 3) nor silent morphemes (section 4) represent adequate tools to deal with portmanteaus. I further demonstrate that the insertion principle used in DM (the so-called Subset Principle) is incapable of governing non-terminal lexical insertion, while the NS conception (the so-called Superset Principle) handles the relevant data with ease. The move to non-terminal spell out is also theoretically attractive; once adopted, it replaces the morphological operations proposed to deal with portmanteaus (Caha, 2009, ch. 2).

Once such morphological operations are eliminated, what emerges is the possibility of a direct mapping between syntax and pronunciation, mediated only by lexical access. This theoretical goal ultimately represents a third difference (**dif 3**) between NS and DM: is there (DM), or is there not (NS), **a separate post-syntactic morphological module/structure**, with rules and operations specific to that module? This issue – independent of the difference 2 – revolves around the questions of whether the remaining post-syntactic operations used in DM are needed to derive the surface forms of language, or whether these operations can also be dispensed with. Once again, NS is not alone here in its attempts to avoid using such operations, see, e.g., Koopman (2005) or Kayne (2010).

The specific operations that are covered under the **dif 3** (on this way of cutting the pie) are **post-syntactic operations that change constituency and/or linear order**. Specifically, Marantz (1989) as well as Embick and Noyer (2001) propose that in a number of cases, structures of the type [X [Y Z]] can be transformed onto [[Y X] Z] by **Merger**, which (according to them) takes place after syntax. According to my perception of the data, the difference to NS here is not so much about what the constituency is (I think that in most cases it is [[Y X] Z], i.e., the one which DM derives by Merger), the question is rather whether the post-syntactic Merger is the only way to derive such structures. The strategy that has been

adopted in some of the NS literature is to say that Y and X are in fact merged independently together, and only later merged with Z, yielding [[Y X] Z]; see Baunaz and Lander (this volume) for the operation Construct and also Taraldsen (this volume).[2]

Obviously, the logically independent issues covered under 2 (is there or is there not phrasal spell out) and 3 (what is the constituency of strings and how to derive it) interact. If lexicalization targets (potentially phrasal) constituents, then it matters what type of constituent structure feeds into lexicalization. To give a concrete example: it matters for spell out of X and Y whether the constituency is [X [Y Z]] (X and Y cannot spell out together) or whether it is [[X Y] Z] (X and Y can be spelled out together).

Most of the discussion of NS from DM positions has focused on the issues surrounding constituency (**dif 3**). So despite the fact that my main goal here is to look at the topic of insertion (**dif 2**), it feels wrong to ignore the concerns that researchers working in DM have about the constituency required by phrasal spell out. So I devote some remarks to this topic in section 2, before I come back to insertion in sections 3 to 5. Section 6 concludes.

2. Developing a theory of spell out in a model of syntax that keeps changing

As highlighted under **dif 1,** NS and DM have diverging opinions concerning the general outlook of syntactic structure. On the one hand, these differences are rather limited in scope, pertaining mostly to the number of heads and the amount and type of movement needed to generate surface strings. On the other hand, these differences have quite a big impact on the debate that takes place in the literature. To a certain extent, it may seem surprising that diverging opinions about movement would lead to disagreement between frameworks that are in practice concerned by the traditional morphology. But it is in fact old news that one's conception of syntax tends to be rather important for one's choice of morphological framework.

[2] In the context of DM, Bobaljik (2012, 57) also suggests the option of simply base-generating such structures.

In order to illustrate that, let me come back to the 1990s when DM itself was taking off as a framework. In one of the papers of that time, Pullum and Zwicky (1990) reflected over DM's conception of morphology from the positions of an "A-morphous" approach (based on the idea that words correspond to unstructured feature matrices and represent the atoms of structure building). The authors noted that the structured morphological representations used in DM fare quite well with Pollock's (1989) system, where complex words arise by head-movement, rather than being simply pulled out of the lexicon. To the authors, such a system involved "baroque-style derivations," where "it is not clear why affixes or stems could not be moved apart by syntactic rules to yield (e.g.) Affix Topicalization or Heavy Stem Shift constructions." They went on to claim that [Pollock's system,] "even if it worked…, it would … generate strings of the wrong sort to be input to a morphological module of a theoretically and empirically optimal sort." In effect, Pullum and Zwicky took their paper to "cast serious doubt on whether its basic operation of head-to-head movement was doing any real work." Some twenty five years later, we are where we are. Head movement has become a common stance in syntactic analyses, and DM a point of reference to compare new alternatives with.

The reason for mentioning Pullum and Zwicky's criticism here is not to agree with it; rather, the intention is to show that DM's reliance on a new syntactic tool (head movement) led to a new style of morphological analyses, and that researchers who had their doubts about the correctness of the tool were reluctant to adopt a morphological theory that seemed to require its large scale application.

But syntax has of course not stopped in the 90s, and new tools kept arriving. Consider, for instance, the development from Cinque (1994) to Cinque (2005). Both papers try to derive the order of nouns, adjectives, numerals and demonstratives. In 1994, Cinque proposed that when the noun preceded any of these, it moved from its base position at the bottom of the tree to the left of those modifiers by head-movement. In 2005, Cinque included many more

languages into his sample and noted that the original system encounters problems, because a noun may move across two modifiers without inverting their order, something that the traditional head-movement analysis could not do. So Cinque proposed that the noun actually crossed its modifiers by phrasal movement, and provided a system which restricts movements in such a way that it can generate all and only the attested orders using exclusively phrasal movement. This – among other things – meant that the complement of the noun (which often remains last in the string even when the noun moves) has to undergo obligatory extraposition (Cinque 2005, 327).[3]

With such changes in syntax taking place (cf. the literature emanating from Kayne 1994; 1998), it became possible to contemplate new models of morphology which make use of such derivations. To see where they fit in within NS, consider the form *puell-as* (discussed also in the introduction to this volume), which is an accusative plural (*-as*) of the Latin noun 'girl' (*puell*). As far as syntax is concerned, there are reasons to think that number and case occupy distinct slots in the cross-linguistically fixed sequence of projections (we will see some evidence for this in section 4). But in the morphological string, it is impossible to separate the accusative meaning from the plural meaning; *-as* is a portmanteau for both. To remove tensions such as these (a non-divisible form corresponding to two syntactic positions), NS proposes that these two projections are spelled out together as a phrase after the noun moves to the left by phrasal movement, see (1):

[3] Extraposition here refers to the movement of the noun's complement to a position in the extended NP such that it c-commands and precedes all the noun's modifiers. After that, in languages where complements follow the noun, the whole extended NP undergoes remnant movement, so that the complement ends up to the right of the whole noun phrase (again).

(1) NP **[K [Num ~~NP~~]]**

puell ----- *as* -----

Suppose now that the noun had a complement. What would happen to it? If the complement did not move, it would be carried inside the NP and end up in between the root and the affix (which does not happen). So the assumption needed to make this work is that the complement undergoes obligatory extraposition to an even higher position (minimally above the landing site of the noun), followed by a remnant movement of the noun with its affix. There is nothing new here compared to Cinque's (2005) work: complements undergo obligatory extraposition.

For many researchers, such derivations are simply too baroque to be of any interest. For instance, Embick (2014) correctly points out that "the main predictions of [phrasal spell out] derive from constituency." But keeping the 1990s ideas about syntax as a reference point, he points out that for instance, "synthetic forms should never be found with adjectives that take complements," because that would lead to the problem of the intervening complement in (1). In general terms, Embick concludes that the representations required by phrasal spell out are „in the crucial cases … incorrect." He notes, though, that „in all of these, it is possible to incorporate additional assumptions [obligatory extraposition of complements] to neutralize incorrect predictions", but for these „there is little evidence" (p. 23).

However, keeping Cinque's quite independent conclusions about obligatory extraposition in mind, one cannot help feeling that we are taking part in a similar debate that took place some 25 years ago: we simply believe in different syntaxes (one standard by now, the other still too baroque), and so we are talking cross purposes.

Embick's paper is one of the few critiques of phrasal spell out that at least acknowledges these issues. Merchant (2015) fails to mention even the possibility of extraposition when he

writes that „whenever [phrasal spell involving the root] must occur, the root should allow for no internal arguments, a prediction that is clearly false.“

Merchant's unwillingness to contemplate extraposition in these cases contrasts with his own work on ellipsis. Consider, for instance, gapping examples like: *John likes sandals and Mary ~~likes~~ stiletto heels*, concerning which van Craenenbroek and Merchant (2013:743) say that "it has become fairly standard to analyze gapping as involving movement of the remnants [*stiletto heels*] to the left followed by deletion … of the rest of the clause." Clearly, the authors are relying on extraposition to create constituents (the verb and its inflection) that are later targeted by ellipsis. The reason why Merchant denies this option for phrasal spell out (in effect subjecting phrasal spell out to a different theoretical standard than ellipsis) is a mystery to me.

Another construction where Embick (2014) as well as Haugen and Siddiqi (2013) think that phrasal spell out does not work are cases where a preposition spells out as one piece with the determiner (they mention French and Spanish respectively). They start from the idea that because of the scope relations, the preposition and the determiner cannot form a constituent to the exclusion of the noun. But since they spell out together, they conclude that phrasal spell out cannot deal with this fact.

What is slightly puzzling is that Embick (2014) ultimately does adopt the structure [[P D] N] for his own analysis, i.e., a structure which makes his point about wrong predictions of phrasal spell out irrelevant if evaluated against that structure. The reason why he denies the phrasal-spell-out model access to the correct structure is because he says that as far as syntax is concerned, the structure must be [P [D N]], and the structure which correctly describes the P-D interaction is derived by Merger. The idea is that phrasal spell out cannot make use of the post-syntactic Merger which, according to Embick, is the only way to derive the structure.

There are two observations to be made. The first one is that phrasal spell out is in principle independent of whether one does or does not have Merger, and so these things should

not be automatically lumped together. Even more so that they introduce a double standard for comparison where a terminal-spell-out model is evaluated against the correct structure, and the phrasal-spell-out model is claimed to make wrong predictions because it is evaluated against a different (and incorrect) structure. In other words, even if one adopts a theory with Merger, it is still better (as I argue in sections 3-5) to spell out the non-terminal derived by Merger than let insertion target terminals.

The second observation is related to the issue of whether syntax can or cannot build structures of the sort [[P D] N]. Here it is worth noting that apart from the French/Spanish facts, there is evidence for such a type of structures independent of phrasal spell out. For instance, Baker and Kramer (2014) discuss the fact that in Amharic, a prepositional marker which scopes over the whole extended NP is located inside (i.e., in the middle of) a constituent which is located on the left branch of that NP. This can only be captured if the preposition is included inside the left-branch modifier ([[P Modifier] N]).

Now given the converging evidence for the possibility of structures like [[P D] N], the question to ask is: can we design our theory of syntax in a way that it can accommodate the evidence without the involvement of a post-syntactic component?

Proposals addressing such issues are already out there even independently of NS (but see also Baunaz and Lander's introduction for a discussion of the operation Construct and also Taraldsen's contribution to this volume). For instance, Leu (2008) investigates structures analogous to the unexpected [[P D] N], but for a slightly different case. He looks at examples where the evidence suggests that definiteness morphemes are present in the projection of the adjective, yielding surface structures of the sort [[D A] N], against the semantically expected [D [A N]]. Leu then proposes a theory which base-generates the definiteness marker inside the projection of the adjective. The fact that definiteness scopes over the whole NP is attributed to a silent D sitting in the main projection line, so we actually have the structure [[D A] [**D** [N]]],

where the boldfaced D is silent due to a generalized Doubly Filled Comp (Koopman 1996). So once again, it seems that we independently need the structures which Embick claims to be problematic, and we also have minimally the beginning of an account. The fact that this account deviates from the structures used in the early 1990s should not distract us too much.

My ultimate point is that the reflections of NS in the DM literature resemble the welcome that DM got some 25 years ago. The reason for the resemblance is that people back then as much as now have different ideas about what syntax looks like. The syntactic structures of NS are "too baroque, and even if they worked, there is little evidence for them." I have tried to make some justice to that debate here; my main point is that the constituent structures needed for phrasal spell out (and claimed to be problematic) are needed also for reasons that have nothing to do with phrasal spell out, a point that I think should not be controversial.

The only thing that is potentially controversial is whether we should generate such structures in the syntax or not. I think that the null hypothesis is that we should minimally try to do so. Ultimately, any empirical debate about the structures needed for phrasal spell out cannot take the 1990s syntax for granted. Rather, we have to focus on questions such as these: is there any evidence for or against the claim that the constituent structure needed for phrasal spell out is accessed also by other processes (like gapping etc.).

Whatever the ultimate answer to such questions will be, in what follows, I sidestep as many of these issues as possible. My goal will be to discuss the differences that exist in the insertion procedure independently of the diverging opinions about what syntactic structures look like, and how they are constructed. The way I do this is that I assume here only structures independently adopted in DM, and show how the theory can be improved if non-terminal spell out is adopted.

3. Insertion in DM

This section illustrates some of the paradoxes that arise in classical DM, where spell out targets terminal nodes only. I also show why it is impossible to extend the insertion procedure to non-terminals, presenting what I think are unsurmountable puzzles. The solution to the puzzles will come in section 5, when we turn our attention to the Superset Principle.

3.1 The Subset Principle

Vocabulary insertion in DM is governed by the so-called Subset Principle. Its canonical formulation comes from Halle (1997), see (2). The principle has two parts, which can be called The Subset Clause and The Elsewhere Clause. The Subset Clause determines the conditions for the applicability of a Vocabulary Item. The Elsewhere Clause says what happens when several Vocabulary Items are applicable in a given environment.

(2) The Subset Principle

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

The main empirical bite of the principle is that it can nicely model syncretism. To see that on an example, consider the following fragment of the singular declension in Latin. In the table, I also indicate one possibility of attributing morpho-syntactic features to the relevant cells borrowing some of the features from Halle (1997).

(3) A fragment of the Latin declension

[+SINGULAR]	neuter [-MASC]	masculine [+MASC]
nom [+SUPERIOR]	*-um*	*-us*
acc [-SUPERIOR]	*-um*	*-um*

What we see here is that one of the exponents (*-us*) is tailor made for a particular case and gender.[4] There is no intricacy involved in saying that this marker is the pronunciation of the features [+masc, +sup]. It is much harder to say what the other marker is the pronunciation of. It can appear both with masculines and neuters, and it can appear both in the nominative and in the accusative.

The Subset Principle allows us to generate such paradigms easily. It is enough to say that the marker *–um* simply corresponds to an underspecified singular marker, as in (4a). The Subset principle then predicts that it can occur in any case and any gender, as long as the [+sg] feature (for which *–um* is specified) is present in a given feature bundle/paradigm cell. So in principle, this marker can occur in all the cells of the paradigm (3).

(4) a. *-um* = [+sg]

b. *-us* = [+sg, +sup, +masc]

The number of cells where such an 'underspecified' marker actually surfaces depends on what other markers there are. So when we add the nominative (+sup) masculine (+masc) *–us* into the set of Vocabulary Items, see (4b), this will lead to a clash between *–um* and *–us* in the nominative of the masculine paradigm (in all other cells *-us* does not qualify for insertion on

[4] I am ignoring here the potential decomposition into *–u-m* or *-u-s*, which does not change the point.

the basis of the subset clause). In this particular environment, *–us* wins, because it is more specific, and *–um* fills the rest of the cells. So the possibility to have underspecified markers coming in after syntax is very attractive; syntacticians can do their work on fully specified trees/feature structures, over which syntactic generalizations are stated. The Subset Principle takes care of the morphological detail.

3.2 The paradox

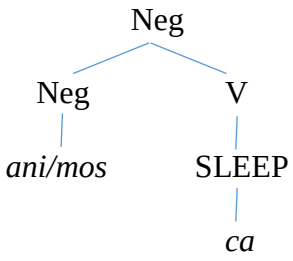
Notice now that the insertion principle (2) explicitly states that phonological exponents are "inserted into a morpheme of the terminal string." A morpheme in DM is a term for a bundle of features (a terminal), so in effect this means that insertion only applies at terminal nodes.

Are there any reasons to doubt this conclusion? Suggestions to this effect were rare in the early days of DM (I think that the first non-terminal spell out analysis in DM is Radkevich 2010, with a manuscript version going back to 2008). However, some issues did start appearing. An interesting insight into the dilemmas that the theory faced at that point (and still faces today) is provided by a careful study of negation in Korean by Chung (2007).

The starting point of Chung's discussion is the fact that sentences in Korean can be negated by attaching one of the two negative prefixes *ani* or *mos* to the verb (5a,b). Chung (2007) shows that each of the negations is a separate head in the clausal spine, and the verb combines with it by syntactic movement. The negation sits lower than T, so that the structure of the string *mos/an(i)-ca* is shown in (6).

(5)	a.	*ca-n-ta*	b.	*mos/an(i) ca-n-ta*
		sleep-PRES-DECL		NEG sleep-PRES -DECL
		'is sleeping'		'can't sleep / isn't sleeping' (Chung 2007:ex.1,2,4)

(6)



The second relevant fact is that the verb *al-* 'know,' see (7a), does not combine with any of these markers (7b), but shows a suppletive form *molu-* instead, as in (7c).[5]

(7) a. *al-n-ta*
know-PRES-DECL
'knows'

b. **mos/an(i) al-n-ta*
NEG know-PRES-DECL
'cannot / does not know'

c. *molu-n-ta*
NEG.know-PRES-DECL
'cannot / does not know'
(Chung 2007:ex.45)

What we see here in abstract terms is that where one meaning ('not sleep') has two markers (Neg and V), another meaning ('not know') has a single non-divisible marker (Neg+V). Chung takes care to show that *molu* is an actual negative form; it does not correspond to a lexical verb like 'ignore, be unaware of.' So it seems that *molu* actually expresses both Neg and V.

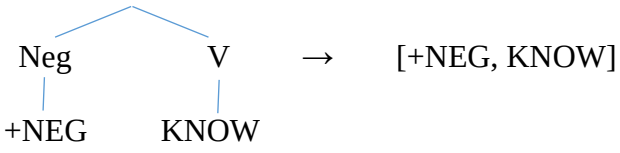
But given that insertion only targets terminal nodes, it cannot be the case that *molu* is inserted into the two terminals at the same time. So the question is whether *molu* is inserted under Neg, or under V. The choice feels forced and leads to some obvious issues. If it was

[5] The form is obviously similar to *mo(s-a)l*, and could have easily developed from the regular form. But synchronically, it is unpredictable and must be stored (to judge from Chung's paper).

inserted under V, we would expect it to combine with negation (which it doesn't). If it was inserted under Neg, we would expect it to combine with a verb (which it doesn't either).[6]

The intuition that the form *molu* in fact conveys the meaning of both 'know' *and* the negation seems hard to implement. Chung (2007) concludes that within the confines of the DM system, there is only one possible solution. What one has to say is that the structure (6) is targeted by a special operation, Fusion, which turns the layered representation into a flat node. The procedure is given in (8), taken from Chung (2007, ex.82). The lexical entry (9) is then allowed to apply, since Fusion has collapsed both features under a single terminal node.

(8)



→ [+NEG, KNOW]

(9) [+NEG, KNOW] = *molu*

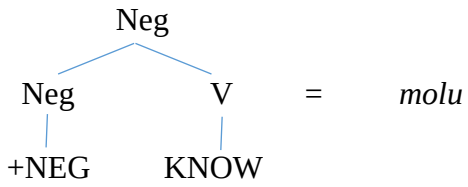
The solution in terms of Fusion makes justice to the observation that *molu* conveys both Neg and V, but it leads to a paradox (identified already in Chung 2007, ftn.22). On the one hand, Fusion must precede lexical insertion, because lexicalization targets the structures which Fusion creates. On the other hand, Fusion happens only when the lexicon contains a portmanteau for the fused heads. (Thus, for instance, Fusion cannot happen when the verb is 'sleep,' as in (5).) Thus, an operation which precedes lexicalization is conditioned by lexicalization.

[6] An anonymous reviewer suggests that *molu* may be a special allomorph for KNOW under negation, and that also negation has a special zero allomorph in the context of KNOW. I get to this type of approach in section 4.

3.3 Non-terminal spell out as a solution to the paradox

Chung then notes that a natural solution for the paradox would be to say that the lexical item *molu* spells out the whole structure in (8). The lexical entry would then look as in (10):

(10)



= *molu*

Once (10) is adopted as a possible format for a lexical item, the paradox disappears. There is no need for Fusion to apply before insertion, because the non-terminal in (8) can directly be mapped onto its pronunciation.

So the natural thing to do at this point would be to simply drop the restriction that Vocabulary Items may only be inserted at terminals. If one could do that, one would get a more general theory of insertion that applies to all syntactic nodes (without a restriction to terminals which is anyway stipulated in (2)). Moreover, the ugly paradox would disappear; we would in fact derive the fact that Fusion is driven by the existence of a particular Vocabulary Item, because insertion at a non-terminal would only be available for such items. Or wouldn't it?

3.4 The communication breakdown

Let us check. Consider for example the Vocabulary Item for SLEEP, given in (11).

(11) *ca-* = [$_{V}$ SLEEP]

Looking at the item, let us ask the question whether this lexical item can spell out the structure (6). After we have dropped the restriction to terminals, The Subset Principle now says that the

entry can spell out a node "if the item matches all or only a subset of the grammatical features" specified in that node. And since (11) corresponds to a subset of the features contained in (6), it seems that *ca-* should be able to spell out that node, in effect meaning both 'to sleep' and 'not to sleep,' a consequence that Chung notes in his ftn. 22. Generalizing this observation, Bye and Svenonius (2010, ftn12) point out that this solution would in fact lead to the expectation that every sentence is pronounced by a single morpheme. I will call this consequence 'the communication breakdown.'

Once Chung realized the grim prospects of simply extending The Subset Principle to non-terminal spell out, he made no attempts at developing the idea further. Compared to the problem of communication breakdown, the Fusion rule in (8) (which he ultimately adopts) looks quite innocent. But is there really no way to avoid this problem?

3.5 Trying to fix things by adding principles: the VIP

One possibility to avoid communication breakdown has been investigated in the work by Radkevich (2010) and Bobaljik (2012). They suggest that adopting in addition to the Subset Principle the 'Vocabulary Insertion Principle' will eliminate the problem.

(12) The Vocabulary Insertion Principle (Radkevich 2010, 8)

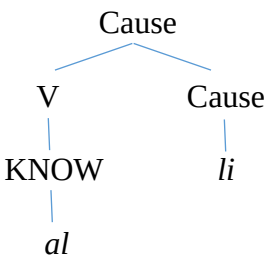
The phonological exponent of a vocabulary item is inserted at the minimal node dominating all the features for which the exponent is specified.

With the VIP in place, some of the most pressing problems of communication breakdown indeed disappear. For instance, (11) can no longer spell out the whole structure in (6), because the minimal node containing all the features of (11) is the V node. Consequently, Neg must be spelled out separately (by one of the negative markers) and we get the right result.

However, (12) ultimately fails in a slightly more complex set of cases. To see that, consider an additional fact noted by Chung (2007), namely that suppletion for negation interacts with causativization. The first piece of the relevant data is in (13a), which is a causative form of the verb 'know,' meaning 'let know, inform.' The causative component corresponds to the affix *–li*. The structure of the causative as proposed in Chung is shown in (13b). The VIP based insertion procedure correctly inserts *al-* 'know' at the V node, and *–li* spells out the causative.

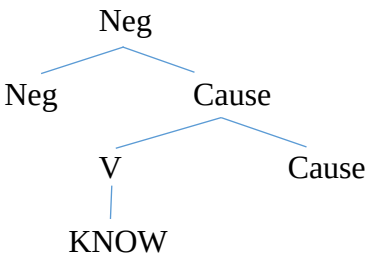
(13) a. *al-li-*
know-CAUS
'let know, inform'

b.



The question to ask is what happens when we negate the causative form, yielding the meaning of 'not to inform,' where the negation scopes over the causative, see (14), reproduced from Chung (2007:ex.86).

(14)



The prediction of The Subset Principle together with the VIP is that the lexical item (10) can be inserted on the top node of the tree (14). That is because the features of the lexical entry (Neg and V) correspond to a Subset of the features in (14), and the top node in (14) is the minimal node containing those features.

But that is a wrong result. As Chung (2007) points out, the structure (14) is actually pronounced as (15). This means that (14) cannot be spelled out by *molu* even though the Subset Principle would seem to allow this even when combined with the VIP.

(15) ani / mos al-li-ess-ta
NEG know-CAUS-PAST-DECL
'did not /could not inform' (Korean, Chung 2007, ex.58)

What is the general lesson we learn? The Subset Principle – which has been widely adopted in DM as a principle governing insertion – fails badly when used as a principle that regulates insertion at non-terminal nodes. The literature that (for good reasons) tries to extend DM by non-terminal spell out, adopted a general strategy of preventing problems by *adding* principles. Their goal is to somehow block the unbounded extension of a particular VI up to a point where this single item spells out the whole sentence. The VIP is one possible way to do this, but it fails in a particular set of cases – namely when a VI contains a feature at the bottom (V) and at the top of a tree (Neg), with one or more features intervening in between (Cause). I will refer to this as a "problem with interveners."

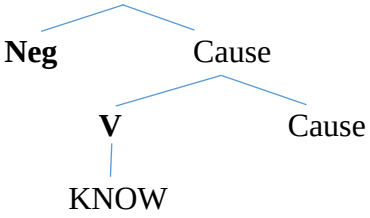
Other principles of a similar sort have been proposed, but they all ultimately fail. I discuss one of these below.

3.6 Global comparisons

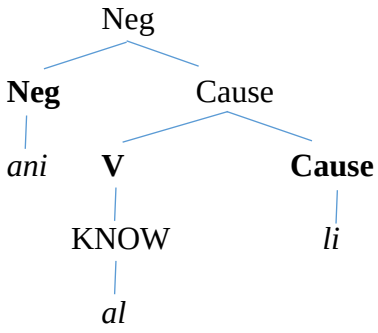
One way to avoid the problem with interveners is to say that there is actually nothing wrong with the fact that *molu* matches the structure in (14). It's just that there is a *better* way of expressing the content of (14), which expresses more features than *molu* alone.

The idea is informally depicted in (16). What we see in (16a) is that when the item *molu* spells out the whole constituent (16a), it leaves the causative feature unexpressed (expressed features are in bold). Compare to this the (correct) spell out in (16b), which – even though more complex in terms of the number of markers – expresses more features. (16c) makes explicit the logic which would lead to the result that (16b) is chosen over (16a).

(16) a. Neg = *molu*



b.



c. Compare possible lexicalizations and choose the one with fewest features unexpressed. (Cf. Siddiqui 2006: The most economical derivation will be the one that maximally realizes all the formal features of the derivation with the fewest morphemes.)

A theory roughly along these lines has been proposed in Julien (2015), and it must have been on Embick and Marantz's mind when they wrote that "when the approaches to blocking involving comparison between otherwise grammatical expressions are made explicit [in our case the comparison of *molu* and *ani-al-li*], they all necessarily involve global competitions …" Embick and Marantz (2008, 4).

We will see later that this conclusion is not completely correct, and that The Superset Principle provides a solution that avoids global competitions. Mentioning the quote here serves as an illustration that by weighing the two scenarios in (16) one against the other, we are not pursuing a useless intellectual exercise. It seems that when Embick and Marantz tried to imagine a model of spell out where *molu* blocks *mos-al* (recall (7)), they considered the "global

comparison" scenario in (16) an unavoidable consequence. This finding is important when we want to evaluate the contribution of NS to the development of a usable phrasal-spell-out model. Clearly, Embick and Marantz did not see it coming as late as 2008 (which is about the same time when the first Nanosyntax papers started appearing in press, see Caha 2007 and Fábregas 2007, who both give credit to Starke's unpublished work going back to the early 2000s).

Now even admitting that (16) is potentially computationally complex, is it going to work? There is again a set of cases where it won't. In classical instances of phrasal spell out, the expected and regular combination of A and B is blocked by an opaque form C. But there are also cases where the combination of A and B is blocked by A, i.e., a form that is identical to one of the two expected pieces. I will call this the AbAB pattern (**A b**locks **AB**).

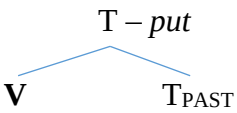
One example of such a pattern can be found in English plural formation. As Baunaz and Lander (this volume, ex.31) illustrate, English has some irregular plurals like *mice* which (apparently) block regular plurals such as **mouse-s* (cf. Siddiqi 2006, ch. 3). If we extend this treatment to *sheep* blocking **sheep-s*, we get AbAB.

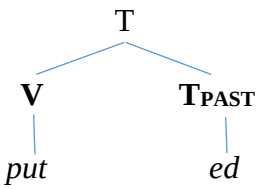
Another case of AbAB is attested for irregular past tense formation. In the classical case, **giv-ed* is blocked by *gave*. One way to encode this is to say that *gave* has the lexical entry in (17a), and it blocks the analytical spell out *giv-ed*, as shown in (17b).

(17) a. T = *gave* b. T – *gave*

Against this analysis, consider now the verb *put*. Here, the regular form **put-ed* is not blocked by a third form, say **pat*. Instead, the past tense is actually the same as the root (*put*), so we have a case of AbAB.

How can we implement this? According to the Subset Principle, *put* cannot have an entry like (17a), because it can also be inserted under V only. Therefore, it must be lexically specified as V, just like all regular verbs. But with such an entry, it is a mystery why *put* can block **put-ed*, whereas regular verbs cannot. In technical terms, when we run the global comparison model on the two lexicalization patterns in (18), we incorrectly predict that (18b) (where no features are left unexpressed) will be chosen over (18a) (where PAST is unexpressed).

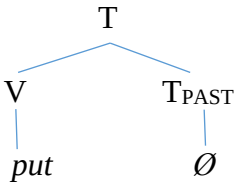
(18) a. 

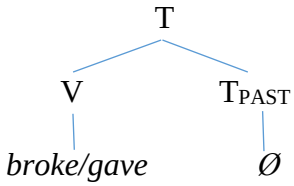
b. 

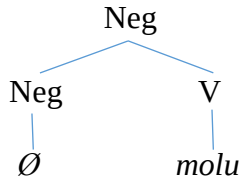
The conclusion is that even admitting global competition does not help, because the signals we are getting from the data are contradictory. For (18), it seems that it is better to use a portmanteau rather than to express all the features. But in (16), it seems that it is better to express all the features, rather than use a portmanteau. I can see no easy way out.

4. Zero morphology

Given these issues, the response within DM (and other theories) is to maintain the initial hypothesis and deny the existence of non-terminal spell out to begin with. Under this standard approach, the AbAB pattern is simply analyzed as an instance of zero morphology, see (19a).

(19) a. 

b. 

c. 

This approach can be generalized for all portmanteaus (i.e., beyond the AbAB pattern) by proposing that the overt terminal has a special allomorph (either suppletive or derived by Readjustment Rules), see (19b), and the node which conditions this special allomorph is silent. This is a classical DM solution going back to Halle and Marantz (1993:124), and it may be used as an all-purpose mechanical alternative to phrasal spell out. The analysis of *molu* 'not know' would then be as in (19c).

This solution is rather popular in a part of the DM literature. For instance, Trommer (1999; 2003) suggests that if we make use of a sufficient number of zeroes (along the lines of 19b), we may get rid of Fusion altogether. Trommer further suggests that also Impoverishment (an operation that deletes features prior to insertion) may be eliminated in a similar fashion.

While some researchers have a general feeling that the excessive use of zero morphology is to be avoided (see, e.g., Siddiqi 2006, ch 3), individual inclinations on this question are likely to vary (as Trommer's papers indicate). The natural question to ask in this context is this: is there any way to tease apart the NS approach from the zero-based proposal empirically?

In order to address this question, I turn to a particular type of data sets where the expressive power of silent morphology can be easily observed. These are the so-called *ABA patterns that were recently investigated both within DM (Bobaljik 2012, a.o.) and NS (Caha 2009, a.o.). My exposition here will be necessarily brief, but the point will hopefully get across, relying on the background to the *ABA patterns provided in the introduction to this volume.

The particular case I discuss is a fragment of the Classical Armenian declension, see (20). Sources on the language include Schmitt (1981), Halle and Vaux (1998) and Caha (2013).

(20) A fragment of a Classical Armenian paradigm (*azg* 'nation,' *hogi* 'spirit')

case	nation, sg	nation, pl	spirit, sg	unattested	decomposition
NOM	azg	azg-k'	hogi	A	[A]
ACC	azg	azg-s	hogi	B	[A, B]
LOC	azg-i	azg-s	hogi	A	[A, B, C]

The first important thing to note about the paradigms is their syncretism structure (syncretism highlighted by shading). In particular, there is a constraint on syncretism in this language, such that in the order of cases given in the table (20) (i.e., NOM – ACC – LOC), only adjacent cases can be syncretic. Specifically, we have the NOM – ACC syncretism (*azg*), the ACC – LOC syncretism (*azg-s*), and also the NOM – ACC – LOC syncretism (*hogi*). What is not found is the non-adjacent NOM – LOC syncretism (see the last column). The absence of this pattern is usually referred to as the *ABA (* for absence, ABA for the pattern that is absent).

In order to derive the *ABA, the decomposition of the cases into features must proceed as shown in the last column. If this "cumulative" decomposition is adopted, the *ABA property of the paradigm is captured regardless of whether we are adopting a DM style insertion (see Bobaljik 2012) or a NS type of theory (Caha 2009).

With the background in place, the first relevant point is that the decomposition proposed guarantees the *ABA only as along as zero morphology (or Impoverishment) are kept out of the picture. Suppose, for instance, that there was a zero marker or an Impoverishment rule applying in the locative cell, which would spell out/delete the features B and C. The effect would be that LOC is effectively turned into a NOM, as shown in the middle column of the table (21). The net effect is that the impoverished decomposition allows us to generate what on the surface looks like an ABA pattern; the *ABA is no longer derived.

(21) The interaction between *ABA and zero morphology

case	decomposition	zero applies	a surface ABA
NOM	[A]	[A]	**a**
ACC	[A, B]	[A, B]	**b**
LOC	[A, B, C]	[A, ~~B, C~~]	**a** (+ Ø)

I will not speculate here over whether there are some cases where this state of affairs is attested in some limited cases or not; the point is simply that zero morphemes do come with a price.

The second difference starts from the observation that in a paradigm which decomposes cumulatively, not only the exponents themselves exhibit the *ABA distribution, but also the morphemes that are conditioned by these features (Bobaljik 2012, McFadden 2014). This gives us an independent metric to assess the distribution of other (say number) morphemes in paradigms as either compatible or incompatible with the *ABA pattern observed. What I now want to show is that in Classical Armenian, the zero morphemes required by the terminal-only spell out yield morpheme distributions that are incompatible with the independently observed *ABA restriction.

In order to show this, I must elaborate on the Classical Armenian paradigm by adding the dative and the instrumental in the paradigms, see (22). The cases are again placed in an order where syncretism is still restricted to adjacent cells (see, e.g., *azg-i*), which leads to the decomposition shown in the last column.

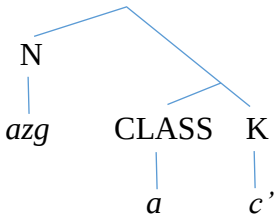
(22) *azg* ‘nation,’ *get* ‘river,’ *ban* ‘word’

case	sg	pl	sg	pl	sg	pl	features
NOM	azg	azg-k’	get	get-k’	ban	ban-k’	[A]
ACC	azg	azg-s	get	get-s	ban	ban-s	[A, B]
LOC	azg-i	azg-s	get-i	get-s	ban-i	ban-s	[A, B, C]
DAT	azg-i	azg-**a**-c’	get-o-y	get-**o**-c’	ban-i	ban-**i**-c’	[A,B,C,D]
INS	azg-**a**-w	azg-**a**-w-k’	get-**o**-w	get-**o**-w-k’	ban-**i**-w	ban-**i**-w-k’	[A,B,C,D,E]

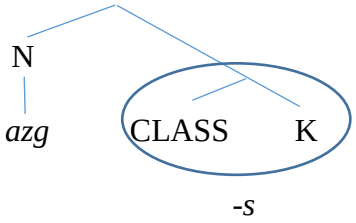
The new thing in these paradigms is that in the lines below LOC, the morphological decomposition is getting increasingly more complex, especially in the plural. Looking for instance at the dative plural (shaded), we invariably find the marker *-c’*. Interestingly, this marker does not attach to the root (whose form can be seen in the nominative singular), but instead to a vowel whose quality varies depending on the particular root. These vowels are in bold, and appear also in the instrumental (both singular and plural). They are traditionally analyzed as a separate morpheme, called the theme marker.

The question to ask is what happens to the theme markers in NOM-LOC, where we do not see them. In Caha (2013), an analysis in terms of non-terminal spell out is proposed in order to explain their disappearance. The gist of the idea is that those case markers that follow the root directly are portmanteau morphemes that spell out both the case node and the node where the theme markers reside, thus causing their absence. In (23), I illustrate the idea on the contrast between the dative and locative plural. Note that the structures in (23) are the output of several applications of movements, which I ignore here; see Caha (2013) for details.

(23) a.



b.



What is relevant now is to look at what happens when we try to recast this analysis in terms of zeroes. Such an analysis is depicted in the first two columns of the table (24) below; all the cells where we see no theme marker are analyzed as having a perfectly agglutinative structure with zero themes ($Ø_{th}$) conditioned by a particular case/number combination.

(24) The distribution of class zeroes

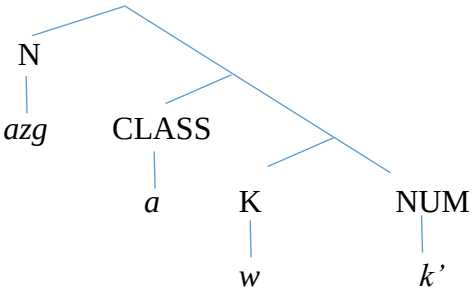
case	sg	pl	features
NOM	azg-$Ø_{th}$-ø	azg-$Ø_{th}$-k’	[A]
ACC	azg-$Ø_{th}$-ø	azg-$Ø_{th}$-s	[A, B]
LOC	azg-$Ø_{th}$-i	azg-$Ø_{th}$-s	[A, B, C]
DAT	azg-$Ø_{th}$-i	azg- a -c’	[A,B,C,D]
INS	azg- a -w	azg- a -w-k’	[A,B,C,D,E]

Within each number, the alternation between the zero theme and the overt theme is a case of a case-conditioned allomorphy. Given the decomposition of case in Classical Armenian, we expect that such an allomorphy should observe the *ABA restriction. And so it does: both zero theme markers and overt theme markers form contiguous regions in the paradigm. This is easy to encode: for instance, we can say that in the plural, the overt theme marker is conditioned by the feature D; so *-a-* will then appear in all cases that have this feature in the plural, which is the dative and the instrumental, a contiguous stretch of cases in the *ABA order.

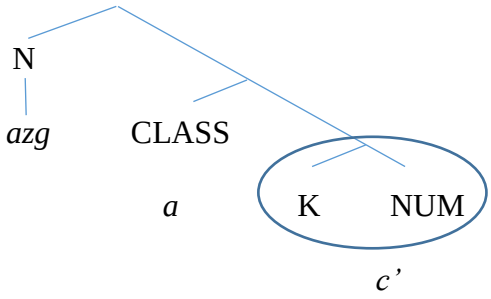
However, problems appear when we turn to the expression of number. In order to see that, consider the fact that in the instrumental plural, see the last row of (22), we obviously have an agglutinative structure, where the instrumental singular (ending in *–w*) is subject to further affixation by (plural) *–k'*, yielding a sequence where case is followed by number. This suggests that there are two separate nodes in the morphosyntactic structure, each with its own marker, and that the noun moves to their left by phrasal movement.

The question is how to analyze the cases where number and case cannot be parsed separately. The first thing to note is that this cannot be attributed to some kind of phonological coalescence. To see that, consider the fact that the syncretic DAT/LOC.SG form *azg-i* has the locative plural counterpart *azg-s*, while the dative plural is *azg-a-c'*; it simply can't be that both are derived by a phonological process starting from the string [*azg–i*]-*k'*. Therefore, both Halle and Vaux (1998) and Caha (2013) analyze *–s* and *–c'* as portmanteaus for number and case. In Caha's approach, the contrast between the portmanteau dative and the agglutinative instrumental is encoded as shown below in (25a,b).

(25) a.

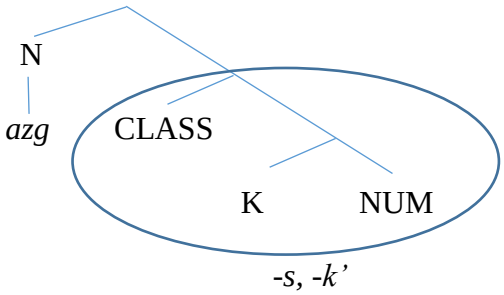


b.



For completeness, I also show in (26) how the phrasal spell out analysis handles the NOM, ACC and LOC cases (which only have a single marker). The idea is that *–s* and *–k'* spell out all the relevant categories (a combination of the proposals in (23) and (25)), see Caha (2013) for details.

(26)



The fact that *–k'* appears both as a simple number marker (in 25a) and as a portmanteau for several categories is an instance of AbAB: the expected combination of *–k'* and the case/theme markers is blocked by *–k'* itself.

Let me now turn to the question how the analysis based on zeroes handles this fact. Recall that the basic "insight" of this approach is that in cases where we only see one marker for two categories, the other category must be zero. Following this idea, there are two ways of distributing the zeroes in the plural paradigm. (i) The zero markers spell out case, as in the second column of (27). (ii) The zero markers spell out number, see the third column in (27). In both options *–k'* spells out number.

(27) The distribution of zeroes in the number and case

case	zero K	zero Num	features	portmanteau
NOM	-Ø$_{(Kase)}$ **-k'** $_{(Num)}$	-Ø$_{(Kase)}$ **-k'** $_{(Num)}$	[A]	K+NUM = *-k'*
ACC	-Ø$_{(Kase)}$ -s $_{(Num)}$	-s$_{(Kase)}$ -Ø $_{(Num)}$	[A, B]	K+NUM = *-s*
LOC	-Ø$_{(Kase)}$ -s $_{(Num)}$	-s$_{(Kase)}$ -Ø $_{(Num)}$	[A, B, C]	K+NUM = *-s*
DAT	-Ø$_{(Kase)}$ -c' $_{(Num)}$	-c'$_{(Kase)}$ -Ø $_{(Num)}$	[A,B,C,D]	K+NUM = *-c'*
INS	-w$_{(Kase)}$ **-k'** $_{(Num)}$	-w$_{(Kase)}$ **-k'** $_{(Num)}$	[A,B,C,D,E]	K = *-w* NUM = *-k'*

Whichever option we look at, we end up with a *ABA violation exhibited by the plural marker *–k'*. Its distribution is impossible to state using the logic illustrated for theme markers, because

NOM and INS do not form a natural class in the decomposition required to capture the *ABA pattern. (The same in fact holds also for Halle and Vaux's 1998 feature system, which was not designes to capture the *ABA property of Classical Armenian.)

In the specific decomposition used here, the problem of the zero approach takes the following shape: in order for the number marker *–k'* to appear in NOM, its context of insertion must be defined by the feature A (it can also be context free). Therefore, in principle, we expect *–k'* in all cases, because all cases have A. The reason why it does not appear in ACC must then be attributed to the fact that there is a competing number marker (either zero or *-s*) which is more specific; it must then be context dependent on A and B; therefore, it wins over *–k'* for insertion in the ACC. However, by this reasoning, *-k'* should not appear in INS, because this case also contains A and B, so the more specific marker should appear here as well. In effect, a paradigm that is surface compatible with *ABA is turned into one which encounters a *ABA violation.[7]

The main point stemming from the discussion is this: zeroes are a tool that DM can use to mimic the effects of phrasal spell out. In fact, this is possible to the extent that apparently any non-terminal spell out account may be mechanically translated into an account that uses conditioned allomorphy and zeroes. The main purpose of this section was to show that the type of allomorphy needed to make this work leads to patterns that violate an independently observed *ABA constraint. The reason for this is obviously not that the existence of a *ABA pattern was an illusion that is dispelled after the zeroes are placed where need be. Rather, the reason is that

[7] One can 'solve' this problem by involving an Impoverishment rule (because creating *ABA violating patterns is what Impoverishment does). One would then delete the feature B in the context of F, bleeding the context for insertion of the marker conditioned by the features A and B. It is possible to show that this leads to additional problems, but space limitations prevent me from showing this.

the zero morphemes are simply made up, and hence cannot be used as a general replacement for non-terminal spell out.

All the problems are avoided under the non-terminal spell out analysis and more generally in any analysis with portmanteau markers. On this type of approach, the Classical Armenian paradigm breaks down into two separate sub-systems. From NOM to DAT, number and case are spelled out by portmanteaus. In the last column of (27), these cells are shaded. There is no *ABA violation within this part of the paradigm. In the instrumental, there is a different sub-system, where case and number are separate. Trivially, there is no violation of *ABA in this sub-system either. The (nominative) plural *–k'* appears in both sub-systems, spelling out either just number (in the instrumental plaural), or both plural and the case feature A (an instance of AbAB).

The approach in terms of portmanteau marking is also in line with existing DM studies like Halle and Vaux's (1998) analysis; they too propose that the paradigm is not surface agglutinative, and that the number of morphological slots in the paradigm varies exactly as shown in the last column of (27). For them, the *–k'* marker also blocks a separate realization of the nominative, being inserted into a single number-case node just like *–s* and *–c'*. However, for them, the only way to ensure the variable number of slots is to rely on post-syntactic operations of Fusion and Fission that manipulate the number of insertion slots by brute force.

This point quite likely carries over to the past tense of English verbs. As Bobaljik (2012:160) as well as Lander (2015:96) note, there seems to be a *ABA constraint on the forms of irregular verbs. The relevant sequence to look at is PRESENT – PARTICIPLE – PAST (consider *come – come – came*; *shine – shone – shone*). As Bobaljik and Lander agree, this can be captured by a cumulative decomposition where the number of features monotonically grows in the relevant sequence. As a consequence, we should not be much compelled by an analysis like

(19b) that leads to parses such as *break-**ø**, broke-**n**, broke-**ø**,* because such artificial parses once again create a *ABA violation (due to the markers in bold) where none should be attested.

At this point, we have gone the full circleAdmiting the paradoxes connected to Fusion, the reviewers suggested that DM 'has no problem' because we can make use of zeroes. However, looking more closely at the distributions of such zeroes, we realize that the last hope fails as well.

In view of these issues, some of the recent work in DM (Radkevich 2010, Bobaljik 2012, Julien 2015) tries to extend it by making use of non-terminal spell out. However, we have seen that they have hard times doing so: the subset logic in (2) which works so beautifully for paradigms such as (3) breaks down when one tries to extend it to non-terminal spell out.

As mentioned, the strategy has been to avoid these problems by adding principles that somehow restrict the logic inherent in the Subset Principle; we have looked at Radkevich's VIP, or (something resembling) a global comparison model (Julien 2015). None of these fully work, and so we are still waiting for a fully functional non-terminal insertion procedure based on the Subset Principle. This may well be impossible, as Embick and Marantz seemed to think back in 2008.

Still, in a sense, all these strategies equal to giving up The Subset Principle for non-terminal spell out, because their ultimate goal is to allow underspecification only at terminal nodes, and neutralize its effects at non-terminals. This is explicitly hinted at in Svenonius and Bye (2010); in their ftn. 12, they write: "The Subset Principle […] would lead to the result that every sentence would be at most one morpheme. To avoid this empirically false result, [… w]e might for example assume that [an additional principle of] Exhaustive Lexicalization holds of projecting categories […]". Exhaustive Lexicalization Principle (Fábregas 2007) says that every feature must be lexicalized; in the passage quoted above, Bye and Svenonius note that if the principle only held for heads (each head must be lexicalized), the communication

breakdown would be avoided. Ultimately, this approach equals to admitting two insertion principles: The Subset Principle for terminals, and the "Spell Out Every Head" Principle for non-terminals. (But even that will still not explain the AbAB pattern, because here one head goes unexpressed.)

5. The Superset Principle

There is a single reason underlying all these puzzles: the insertion procedure is not governed by the Subset Principle. In order to make things work, we have to adopt the Superset Principle instead (Starke 2009). I give one possible formulation below, which is quite different from Starke's wording, but it is close to the original Subset Principle (so that we can see what the difference actually is). It is also not too remote from Starke's idea.

(28) The Superset Principle

[*The Superset Clause*:] The phonological exponent of a Vocabulary Item is inserted into a node if the item contains all (or a superset of) the grammatical features contained in the node. Insertion does not take place if the vocabulary item is not specified for all features contained in the node. [*The Elsewhere Clause*:] Where several items meet the conditions for insertion, the item containing fewer features unspecified in the node must be chosen.

Potentially, this principle may apply to terminal, non-terminal and phrasal nodes. Let me now show how adopting it solves all the puzzles we have seen.

5.1 AbAB

I start from the AbAB. With the Superset Principle in place, we can specify both *put* and *gave* in the same way, namely as irregular past tense forms, see (29a). With this lexical entry, they can spell out the non-terminal (17b), because they contain all its features. This is not the case for regular verbs, which have a lexical entry like the one in (29b). They do not contain the T node, and so they cannot be inserted at the non-terminal in (17b).

(29) a. *put, gave, sang, …* = [V T_{PAST}]

b. *kiss, locate, …* = V

c. *give, sing, …* = V

Notice now that all the verbs in (29a) can also spell out the V node only, because V is contained in them. So this explains why *put* can have the same form for the past tense and the root, while regular verbs can't.

Finally, the reason why *gave* and *sang* do not show up as the root form of the relevant verb is because of competition. The VIs *sing* and *give* also spell out the V node, and in this capacity, they contain fewer unused features, so they outcompete the past tense form in this environment.[8]

The same logic extends to the AbAB pattern exhibited by *sheep* and the Classical Armenian *–k'* (concerning the latter see Caha 2013 for details).

[8] An anonymous reviewer asks about the treatment of forms such as *told*. If they do not decompose (see Kaye 1995, 310-2 for some arguments), then their treatment is the same as for *gave* in (29a). If they do decompose, something extra needs to be said. Space limitations prevent me from going into this here, but see Caha (to appear) for the discussion of forms such as *bett-er* (which is an analogous problem of a suppletive form accompanied by what looks like regular morphology).

5.2 Korean Negation

Recall now the problem we had with *molu.* It was specified as [Neg KNOW] (recall (10)) and the question was why it can't spell out the causative form [Neg [KNOW Cause]], recall (14). The answer is now clear: it is because (10) does not contain all the features of the top node in (14); Cause is missing.

5.3 How to create a simple paradigm

Recall finally the paradigm (3), repeated below in (30) for convenience. How does the Superset Principle deal with this example?

(30) A fragment of the Latin declension

[+SINGULAR]	neuter [-MASC]	masculine [+MASC]
nom [+SUPERIOR]	*-um*	*-us*
acc [-SUPERIOR]	*-um*	*-um*

Let me first say that because of independent assumptions (recall **dif 1**), there are no complex terminals in NS, so the following discussion is artificial with respect to the actual working of the Superset Principle in that framework. Still, for completeness of the argument, let me show that the Superset Principle could also be used to handle these cases.

If we want to encode the same intuition as in (4), namely that *–um* is a marker that can appear in any cell of the paradigm, it has to contain all the potential feature combinations. Its specification is thus as shown in (31a).[9]

[9] This entry is admittedly quite inelegant, a side effect of the terminal-spell-out scenario. See Caha (2009) for a system of case decomposition which makes such lexical items look neater.

(31) a. *-um* = [+MASC, - MASC, + SUP, - SUP, + SG]

b. *-us* = [+ MASC, + SUP, + SG]

If we now specify *–us* as shown in (31b), it will only qualify for insertion in the nominative singular of the masculine gender, because it does not contain the features of any other cell. The lexical entries in (31) thus derive the paradigm (30), which shows that we haven't lost the cases that were working fine under The Subset Principle driven insertion.

6. Conclusions

By looking at the difficulties of classical DM, we come to realize that NS is not just a particular constellation of the bits and pieces that were independently available before (like Cartography and phrasal spell out). Without changing the principle of insertion, the combination of these proposals simply does not take off. In retrospect, employing the Superset Principle and thereby creating a well working theory seems trivial. At the same time, when Embick and Marantz were contemplating these issues back in 2008, the prospects of any breakthrough in this domain seemed impossible. From that perspective, the Superset Principle and NS is general have opened a whole new field of competition in grammar to look at. Unlike the half-working alternatives, it has made it possible to abandon Fusion and the paradoxes that came with it. By doing so, it creates the prospect for a simplified architecture of grammar where syntactic structures are mapped on their pronunciation by lexical access only.

6. References

Baker, Mark, and Ruth Kramer. 2014. "Rethinking Amharic prepositions as case markers inserted at PF." *Lingua 145*: pp. 141-172.

Baunaz, Lena, and Eric Lander. this volume. "Nanosyntax: The basics."

Bobaljik, Jonathan. 2012. *Universals in Comparative Morphology. Suppletion, Superlatives, and the Structure of Words.* Cambridge, MA: The MIT Press.

Bobaljik, Jonathan. 2015. "Distributed Morphology." Manuscript of an encyclopedia chapter, Uconn.

Bye, Patrick, and Peter Svenonius. 2010. "Exponence, phonology, and non-concatenative morphology." Ms. CASTL, University of Tromsø.

Caha, Pavel. 2007. "Case movement in PPs." *Nordlyd 34(2)*: pp. 239-299.

Caha, Pavel. 2009. *The nanosyntax of case*. Doctoral dissertation, CASTL, Tromso.

Caha, Pavel. 2013. "Explaining the structure of case paradigms by the mechanisms of Nanosyntax." *NLLT 31(4)*: pp. 1015-1066.

Caha, Pavel. to appear. "Suppletion and morpheme order: a unified account." *Journal of Linguistics.*

Chung, Inkie. 2007. "Suppletive negation in Korean and distributed morphology." *Lingua 117(1)*: pp. 95-148.

Cinque, Guglielmo. 1994. "On the evidence for partial N-movement in the Romance DP." In *Paths towards Universal Grammar: Studies in honor of Richard S. Kayne*, edited by Guglielmo Cinque, Jan Koster, Jean-Yves Pollock, Luigi Rizzi, and Rafaella Zanuttini, pp. 85–110. Georgetown University Press.

Cinque, Guglielmo. 2005. "Deriving Greenberg's universal 20 and its exceptions." *Linguistic Inquiry 36(3),* pp. 315-332.

Embick, David. 2007. Linearization and local dislocation: Derivational mechanics and interactions. *Linguistic Analysis 33(3-4),* pp. 2-35.

Embick, David. 2014. *On the targets of phonological realization*. Ms., UPenn.

Embick, David, and Alec Marantz. (2008). Architecture and blocking. *Linguistic Inquiry 39(1)*, pp. 1-53.

Embick, David, and Rolf Noyer. 2001. "Movement operations after syntax." *Linguistic inquiry 32(4)*, pp. 555-595.

Fábregas, Antonio. 2007. "The exhaustive lexicalisation principle." *Nordlyd 34(2)*, pp. 165-199.

Halle, Morris. 1997. "Distributed Morphology: Impoverishment and fission." In *PF: Papers at the interface*, edited by Benjamin Bruening, Yoonjung Kang, and Martha McGinnis, pp. 425–450. Cambridge, MA: MIT Press.

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 Calvert Watkins*, edited by Jay Jasanoff et al., pp. 223-240. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.

Harley, Heidi, and Rolf Noyer. 1999. "Distributed Morphology." *Glot International 4(4)*, pp. 3-9.

Haugen, Jason H., and Daniel Siddiqi. (2013). "Roots and the derivation." *Linguistic Inquiry 44*, pp. 493–517.

Julien, Marit. 2015. "Theme vowels in North Sámi: Spanning and maximal expression." *Lingua 164*, pp. 1-24.

Kaye, Jonathan. 1995. "Derivations and interfaces." In *Frontiers of phonology*, edited by Jacques Durand and Francis Katamba, pp. 289-332. London and New York: Longman.

Kayne, Richard. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Kayne, Richard. 1998. "Overt vs. Covert Movements." *Syntax 1(2)*, pp. 128-191.

Kayne, Richard. 2010. "Towards a syntactic reinterpretation of Harris and Halle (2005)." In *Romance languages and linguistic theory 2008: Selected papers from the 'Going*

Romance' Groeningen 2008, edited by Reineke Bok-Bennema, Brigitte Kampers-Manhe and Bart Hollebrandese, pp. 145-170. Amsterdam: John Benjamins.

Koopman, Hilda. 1996. "The spec head configuration." In *Syntax at Sunset: UCLA Working Papers in Syntax and Semantics*, edited by Edward Garrett and Felicia Lee, pp. 37-64. Los Angeles, CA: UCLA Dept of Linguistics.

Koopman, Hilda. 2005. "Korean (and Japanese) morphology from a syntactic perspective." *Linguistic Inquiry 36(4)*, pp. 601-633.

Lander, Eric. 2015. *The nanosyntax of the Northwest Germanic reinforced demonstrative*. Doctoral dissertation, Universiteit Ghent.

Leu, Thomas. 2008. *The internal syntax of determiners*. Doctoral dissertation, NYU.

Marantz, Alec. 1989. "Clitics and Phrase Structure." In *Alternative Conceptions of Phrase Structure*, edited by Mark Baltin and Antony Kroch, pp. 99-116. Chicago: University of Chicago Press.

McFadden, Thomas. 2014. *Noun stem alternations, the structure of case and the locality of allomorphy.* Handout of a talk at LAGB, University of Oxford, September 5th.

Merchant, Jason. 2015. "How much context is enough? Two cases of span-conditioned stem allomorphy." *Linguistic Inquiry 46*, pp. 273–303.

Pollock, Jean Yves. 1989. "Verb movement, universal grammar, and the structure of IP." *Linguistic Inquiry 20(3)*, pp. 365-424.

Pullum, Geoffrey K., and Zwicky, Arnold M. 1991. "A misconceived approach to morphology." In *Proceedings of the West Coast Conference on Formal Linguistics Vol. 10*, pp. 387-398.

Radkevich, Nina V. 2010. *On location: The structure of case and adpositions*. Doctoral dissertation, University of Connecticut.

Schmitt, Rüdiger. 1981. *Grammatik des Klassich-Armenischen.* Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.

Siddiqi, Daniel. 2006. *Minimize exponence: Economy effects on a model of the morphosyntactic component of the grammar*. Doctoral dissertation, University of Arizona.

Starke, Michal. 2009. "Nanosyntax: A short primer to a new approach to language." *Nordlyd 36(1)*, 1-6.

Taraldsen, Tarals. this volume. "Spanning vs. constituent lexicalization: The case of portmanteau prefixes."

Trommer, Jochen. 1999. *Morphology consuming syntax' resources*. Ms, Universität Potsdam.

Trommer, Jochen. 2003. Feature (non-)insertion in a minimalist approach to spellout. In *Proceedings of CLS 39*, pp. 469-480.

van Craenenbroeck, Jeroen, and Jason Merchant. 2013. Ellipsis phenomena. In *The Cambridge Handbook of Generative Syntax*, edited by Marcel den Dikken, 701–745. Cambridge: Cambridge University Press.