

Knut Tarald Taraldsen*

An introduction to Nanosyntax

<https://doi.org/10.1515/lingvan-2018-0045>

Received July 23, 2018; accepted February 5, 2019

Abstract: This article provides an introduction to Nanosyntax – an approach to the syntax/lexicon interaction originating from work by Michal Starke. I present some of the motivation for allowing lexical items to replace whole phrases rather than only terminal elements, present the basic principles taken to govern the lexicalization procedure and show how syncretism patterns are handled in Nanosyntax.

Keywords: the syntax/lexicon relation; phrasal lexicalization; nanosyntax

1 Overview

This short article is intended to provide an introduction to Nanosyntax (NS) – an approach to the syntax/lexicon interaction that is perhaps best known for assuming that a single morpheme (lexical item) lexicalizes, i.e. replaces, a syntactic phrase as opposed to just a single terminal (syntactic head).

The next six sections aim at presenting the basic motivation for NS and at familiarizing the reader with some central assumptions about the nature of lexical insertion.¹ For space reasons, the important topic of spell-out driven movement (Pantcheva 2011; Starke 2018) is only alluded to in passing.

The Envoi offers a puzzle for the interested reader to try her teeth on.

2 The smallest building blocks of syntactic structures

Chances are a linguistics student will at some point be told that lexical items (words or morphemes) are the minimal building blocks of syntactic structures. This means that lexical items are uniquely associated with the terminal nodes (X^0 -constituents or “heads”) of the tree representation of a syntactic structure.

This view is the mainstream view and is adopted whether one thinks that inflected forms, e.g. *boys*, are formed in a morphological component before they are used to build syntactic structures (the lexicalist position) or formed in the syntax by associating *boy* and *s* with two adjacent syntactic terminals, e.g. N^0 and $Number^0$. It is also adopted in theories such as Distributed Morphology (DM) that take the sound bits of morphemes (i.e. the “vocabulary items” of DM) to be introduced to spell out syntactic terminals by replacing a set of syntactic features only after all syntactic operations have applied (“post-syntactic/late lexical insertion”).

However, there are cases where a single morpheme seems to spell out more than just a single terminal. Consider, for example, the fact that the plural form of *man* is *men*. On the basis of *boy/boys* and similar pairs, many linguists (in particular, adherents of DM) would claim that there are two separate syntactic terminals, e.g. N^0 and Num^0 , one spelled out by a noun and the other by *s* (if associated with the feature [plural]), i.e. $N^0 \rightarrow boy$, $Num^0 \rightarrow s$. But then, what about *men*, which looks like a single non-decomposable morpheme?

This is a problem for any theory holding that morphemes stand in a one-to-one correspondence to syntactic terminals. But of course there are proposals as to how one may get around it. Here, I will only consider the most explicit recent proposals, i.e. the accounts generally developed within the DM framework, and argue that these proposals are highly problematic.

¹ Concise introductions to Nanosyntax are also to be found in Baunaz and Lander (2018) and Starke (2009). Caha (2018) is an in depth discussion of the issues touched upon in Section 2.

*Corresponding author: Knut Tarald Taraldsen, Department of Linguistics, The Arctic University of Norway, Tromsø, Norway; and Department of Czech Language, Masaryk University, Brno, Czechia, E-mail: knut.taraldsen@uit.no

Proponents of DM offer two ways of dealing with cases like *men*. One way is to posit a morphological process, Fusion, that turns two adjacent terminals into one prior to lexical insertion (Halle and Marantz 1993: 136, 1994: 277). On this view, *men* would be inserted into the single terminal created by fusing N^0 and Num^0 . The obvious problem with this view is that Fusion cannot apply freely. It must apply obligatorily in some cases so that **mans* (as the plural of *man*) is excluded, and must be inapplicable in most cases to allow for regular plurals like *boys*. The generalization seems to be that N^0 and Num^0 must fuse just in case there happens to be a lexical item capable of spelling out the terminal created by Fusion, e.g. *men*, but this generalization isn't easy to capture within DM where lexical insertion is governed by the Subset Principle:

(1) The Subset Principle

A morpheme A with the lexical entry $A \leftrightarrow F$ (where F is a set of syntactic features) can replace a terminal T associated with the feature set F if and only if F is a subset of G .

Presumably, the new terminal created by Fusion should be associated with the union of the feature sets associated with N^0 and Num^0 . But then, (1) entails that any lexical item that can lexicalize N^0 or Num^0 , can also lexicalize the outcome of fusing N^0 and Num^0 , i.e. both just *boy* or just *s* should be able to mean 'boys':²

- (2a) the feature sets of N^0 and Num^0 : $N^0 = A = \{x_1, x_2, \dots, x_n\}$, $Num^0 = B = \{y_1, y_2, \dots, y_m\}$
 b the feature set of fused $N^0 + Num^0$: the union of A and $B = \{x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_m\}$
 c lexical entries: *boy* $\leftrightarrow A'$ (a subset of A), *s* $\leftrightarrow B'$ (a subset of B)

So, it seems necessary to stipulate that the application of Fusion is controlled by the lexicon in the sense that Fusion can apply if and only if the lexicon contains a morpheme associated with a feature set with members from both feature sets associated with the fused terminals. For example, if *man* has the entry $man \leftrightarrow F$, *s* has the entry $s \leftrightarrow G$ and *men* has the entry $men \leftrightarrow H$ where H is the union of F and G , $N^0 = F$ and $Num^0 = G$ can fuse, and the new terminal can be lexicalized by *men*, but not by *man* or *s* given the Elsewhere Principle:

(3) The Elsewhere Principle (EP)

Given a terminal $T = F$ and lexical entries $A \leftrightarrow G$ and $B \leftrightarrow H$ with G and H both subsets of F , A wins if and only if G is a larger subset of F than H .

In this scenario, Fusion not only can apply, but must, since otherwise **mans* would be generated (with *man* lexicalizing just N^0 and *s* lexicalizing Num^0).

Although making the application of Fusion contingent upon the existence of certain lexical entries gives the right results, a conceptual issue remains as emphasized in Caha (2018): In DM, Fusion is part of set of "morphological rules" that are taken to apply before lexical insertion, i.e. at a point of the derivation when the lexicon has not yet been consulted.

The other way out of the problem with *men* would be to posit null morphemes, e.g. *men* might be parsed as *men-∅* with \emptyset lexicalizing Num^0 adjacent to N^0 lexicalized by *men* (abstracting away from the question how N^0 gets to be lexicalized by *men* rather than *man* in this case), adopting an idea by Halle and Marantz (1993: 124). But, as the detailed discussion in Caha (2018: Section 4) shows, this approach comes with unsurmountable problems of its own.

In DM, the terminal nodes of a syntactic tree are represented as sets of syntactic features, so-called feature bundles. In view of the widely accepted hypothesis that syntax is really driven by syntactic features, one might say that the ultimate building blocks of syntax are these features. But in DM, the syntactic features are buried inside unstructured sets associated with terminal nodes rather than being terminal nodes themselves. A competing view, typically held by advocates of the cartographic approach to syntax, holds that there is a one-to-one correspondence between syntactic features and syntactic terminals (see Cinque and Rizzi 2008).

² The feature set associated with N^0 is taken to include all (non-phonological) features needed to differentiate the nominal root *boy* from other roots, e.g. *man*, but a different view is developed in Caha et al. (2018).

On this view, the issues arising in connection with *man/men* will arise much more widely simply because the number of morphemes spelling out a sentence is often smaller than the number of syntactic features used to build the corresponding syntactic structure.

3 Phrasal lexicalization

We have seen that DM doesn't provide a coherent account of plural forms like *men* and have also just suggested that the problem is actually much more general since the ultimate building blocks of syntactic structures are likely to be syntactic features. But the problems dissolve once we adopt proposals by Michal Starke (lectures at the University of Tromsø and elsewhere, Starke 2009, 2011) and give up the traditional assumption that lexical items can only be inserted at terminal nodes.³

We can then simply say that *men* all by itself lexicalizes the entire phrase $[N^0 [PL^0]]$ when N^0 has the features of the root *man* (see footnote 2). In NS, this requires that the lexicon contains the entry in (4), where MAN stands for whatever feature set characterizes the root *man*:⁴

$$(4) \quad \textit{men} \leftrightarrow [{}_{\text{PIP}} [{}_{\text{NP}} \textit{MAN}] [PL]]$$

The operative principle is the Superset Principle (henceforth: SuperP):

(5) The Superset Principle

A morpheme A with the entire $A \leftrightarrow T$ can lexicalize a syntactic structure S if and only if S is identical to a subtree of T.

Since there is no entry like (6), the plural form of *boy* will still be *boys*, given the two entries (7)–(8):

$$(6) \quad X \leftrightarrow [{}_{\text{PIP}} [{}_{\text{NP}} \textit{BOY}] [PL]]$$

$$(7) \quad \textit{boy} \leftrightarrow [{}_{\text{NP}} \textit{BOY}]$$

$$(8) \quad s \leftrightarrow PL$$

Lexical insertion is taken to apply to any node dominating a tree matching the tree associated with some morpheme M in the lexicon. This excludes **mans*. Assuming bottom-up lexicalization, the entry in (4) and $\textit{man} \leftrightarrow [{}_{\text{NP}} \textit{MAN}]$ will initially associate the NP with *man* and PL will be associated with *s* in $[{}_{\text{PIP}} [{}_{\text{NP}} \textit{MAN}] [PL]]$, but these choices will be overridden by *men* at the top node given (5) and the entry in (4). That is, the fact that the existence of portmanteau forms like *men* is controlled by lexical entries, which is a problem for Fusion, is directly accounted for in NS.

Since what I have just said about *men* generalizes to all portmanteau morphemes, there is no need for Fusion or null morphemes. Nor is there any need to posit other morphological processes such as Fission splitting one terminal into two to the extent that Fission is typically invoked when a feature set which is generally associated with a single morpheme, sometimes is split over two distinct morphemes. From a NS perspective, we have at least two different terminals in both cases, and these terminals are mostly, but not always, lexicalized by a single morpheme. In fact, NS, unlike DM, paves the way towards a theory without a special morphological component mediating between syntax and lexicalization.⁵

We now turn to another general principle controlling lexical insertion in NS in conjunction with SuperP.

³ NS actually assumes that lexical insertion never applies to terminals, but only to phrasal nodes.

⁴ Or $\textit{men} \leftrightarrow [{}_{\text{PIP}} PL [{}_{\text{NP}} N]]$ if the NP moves across the functional head only when the initial structure $[{}_{\text{PIP}} PL [{}_{\text{NP}} N]]$ cannot be lexicalized as is the case for $[{}_{\text{PIP}} PL [{}_{\text{NP}} \textit{BOY}]]$ if lexicalization only applies to phrasal nodes: Then, *s* cannot lexicalize PL in $[{}_{\text{PIP}} PL [{}_{\text{NP}} \textit{BOY}]]$ where there is no phrase containing just PL, but can lexicalize PL in $[{}_{\text{PIP}} [{}_{\text{NP}} \textit{BOY}] [{}_{\text{PIP}} PL]]$ derived by movement not leaving a copy of the NP (visible to the lexicalization procedure). For ease of exposition, I take NP to move across PL in all cases in this section.

⁵ Unless the spell-out driven movement operations alluded to in footnote 4 and section 7 are viewed as morphological processes.

4 The Elsewhere Principle

To see why the statement in (5) is called the Superset Principle, consider what it predicts for the distribution of a morpheme A with the lexical entry in (9):

$$(9) \quad A \leftrightarrow [_{XP} X [_{YP} Y [_{ZP} Z]]]$$

Suppose that the syntax builds all the trees in (10), because bottom-up syntactic structure building can stop at any of the nodes XP, YP and ZP.⁶ As formulated in (5), the SuperP allows A to lexicalize not only the structure in (10)a, but also the smaller structures in (10)b-c:

$$(10)a \quad [_{XP} X [_{YP} Y [_{ZP} Z]]] = A$$

$$b \quad [_{YP} Y [_{ZP} Z]] = A$$

$$c \quad [_{ZP} Z] = A$$

The node YP in (10)b is identical to a subtree of the tree in (9). Hence, A can also lexicalize (10)b, and similar reasoning obviously also applies to (10)c. If we look at syntactic structures as structured sets of syntactic features (each of which being a syntactic terminal node), (5) allows A to lexicalize any set of features of which the features associated with A in the lexicon is a superset. This fact will play a crucial role in the account of syncretism that we will see shortly.

In passing, I want to point out a difference between NS and DM with respect to the morphological representation of syntactic features. According to the SuperP, every syntactic feature in a tree lexicalized by a morpheme A is spelled out in the sense that it is also associated with A in its lexical entry. In DM, however, lexical insertion goes by the Subset Principle:

- (1) The Subset Principle
 A morpheme A with the lexical entry $A \leftrightarrow F$ (where F is a set of syntactic features)
 can replace a terminal T associated with the feature set G if and only if F is a subset of G.

Accordingly, the morpheme A lexicalizing a feature bundle G need not be associated with every feature in G in its lexical entry. This may be significant in view of Fábregas's (2007) arguments for an "exhaustive lexicalization principle".

At this point, however, one also notes that a question arises in cases where the three structures in (10) are not in fact lexicalized by the same morpheme A. Suppose, for example, we have the following relation between structures and exponents:

$$(11)a \quad [_{XP} X [_{YP} Y [_{ZP} Z]]] = A$$

$$b \quad [_{YP} Y [_{ZP} Z]] = B$$

$$c \quad [_{ZP} Z] = B$$

Since A occurs in (11)a, we must have the entry in (12):

$$(12) \quad A \leftrightarrow [_{XP} X [_{YP} Y [_{ZP} Z]]]$$

Then, the SuperP also allows A to occur in (11)b-c. So why does B, with the entry in (13), appear instead?

$$(13) \quad B \leftrightarrow [_{YP} Y [_{ZP} Z]]$$

The answer comes from the Elsewhere Principle (the EP):

- (14) The Elsewhere Principle

⁶ It is also assumed that the unfolding of an extended projection follows a functional sequence in the sense that if the functional sequence has $[\dots [X [Y [Z \dots]]]]$, the syntax cannot build $[X [Z \dots]]$ skipping Y.

If two morphemes A and B are both able to lexicalize a syntactic structure S, but the distribution of B predicted by the SuperP is properly included in the distribution of A, B wins.

The principle in (14) is an “elsewhere principle” in the sense that when A and B compete for insertion at the same syntactic node, the more specific one of the two is preferred, taking “more specific” to mean “having a more limited distribution”.⁷ To see how (14) works, consider how it will prevent A from occurring in (11)b-c given the entries (13) and (12):

Clearly, the distribution of A as predicted by the SuperP properly includes the distribution of B, since the SuperP allows A to lexicalize all three structures in (11), while B can only lexicalize (11)b-c. By the EP, then, B blocks A for (11)b-c.

The EP too plays an important part in accounting for syncretism and the shape of paradigms.⁸

5 Syncretism and the shape of paradigms

While the SuperP requires that every feature (terminal node) in the input tree T has a match in the tree associated with the morpheme M that lexicalizes it, M may have additional features not found in T. As noted in the preceding section, the morpheme A in (9) may lexicalize all the structures in (10) because (10)b-c are also subtrees of the tree linked to A in (9). Caha (2009) is a detailed study of how this fact can be exploited to account for syncretism in case paradigms. His account is grounded in arguments to the effect that there is a case-hierarchy ... Dative > Genitive > Accusative > Nominative and that the members of this hierarchy should be decomposed into structured sets of privative features as in (15), where each feature is a single syntactic terminal:

- (15) Dative = $[_{WP} W [_{XP} X [_{YP} Y [_{ZP} Z]]]]$
 Genitive = $[_{XP} X [_{YP} Y [_{ZP} Z]]]$
 Accusative = $[_{YP} Y [_{ZP} Z]]$
 Nominative = $[_{ZP} Z]$

In conjunction with the SuperP, this predicts the existence of the syncretism patterns AAAA, AAAB, AABB, ABBB, AABC, ABBC and ABCC. AAAA emerges if the lexicon contains the entry in (16)a, but no entry like (16)b with S a proper subtree of $[_{WP} W [_{XP} X [_{YP} Y [_{ZP} Z]]]]$ whereas the other patterns come out if there is one or more additional entries like (16)b:

- (16)a $A \leftrightarrow [_{WP} W [_{XP} X [_{YP} Y [_{ZP} Z]]]]$
 b $B \leftrightarrow S$

Adding $B \leftrightarrow [_{YP} Y [_{ZP} Z]]$ to (16)a, for example, leads to AABB, given the EP, since as far as the SuperP goes, the distribution of this B is properly included in the distribution of the A in (16)a. Similarly, adding $B \leftrightarrow [_{XP} X [_{YP} Y [_{ZP} Z]]]$ to (16)a leads to ABBB, and adding $B \leftrightarrow [_{ZP} Z]$ instead yields AAAB.

If the lexicon also contains $C \leftrightarrow [_{ZP} Z]$ in addition to (16)a and $B \leftrightarrow [_{YP} Y [_{ZP} Z]]$ or $B \leftrightarrow [_{XP} X [_{YP} Y [_{ZP} Z]]]$, we get AABC or ABBC, and with $C \leftrightarrow [_{YP} Y [_{ZP} Z]]$, we obtain ABCC. For example:

- (17)a lexical entries:
 $A \leftrightarrow [_{WP} W [_{XP} X [_{YP} Y [_{ZP} Z]]]]$
 $B \leftrightarrow [_{YP} Y [_{ZP} Z]]$
 $C \leftrightarrow [_{ZP} Z]$

⁷ In DM, which abides by the Subset Principle, the candidate with the smallest feature set has a wider potential distribution than the candidate with the bigger feature set, which is therefore declared the winner by the Elsewhere Principle.

⁸ But an alternative to the EP is proposed in Caha et al. (2018).

b derivation:

Dative =	$[_{WP} W [_{XP} X [_{YP} Y [_{ZP} Z]]]]$	= A (by the SuperP)
Genitive =	$[_{XP} X [_{YP} Y [_{ZP} Z]]]]$	= A (by the SuperP)
Accusative =	$[_{YP} Y [_{ZP} Z]]$	= B (by the SuperP and the EP)
Nominative =	$[_{ZP} Z]]$	= C (by the SuperP and the EP)

Importantly, all the patterns allowed are actually documented in natural languages. There is, however, one pattern that cannot be generated in accordance with the SuperP and the EP, namely the ABA pattern, which apparently is not found in any paradigm (when the cells of the paradigm are ordered according to independently justified criteria).⁹ This is the empirical generalization currently known as the *ABA Generalization.

To see why the ABA pattern cannot be generated, consider a paradigm where the cells are ordered by decreasing feature structures as in (15). To generate an ABA pattern, the system must allow A to appear in some cell followed by one or more morphemes distinct from A in lower cells before A reappears in a yet lower cell. To get the higher A in place, we need a lexical entry for A that associates it with one of the two structures at the top of (15). The entry in (16)a, for example, will give us A as the top element of the paradigm emerging from (15):

(18)	Dative =	$[_{WP} W [_{XP} X [_{YP} Y [_{ZP} Z]]]]$	= A
	Genitive =	$[_{XP} X [_{YP} Y [_{ZP} Z]]]]$	
	Accusative =	$[_{YP} Y [_{ZP} Z]]$	
	Nominative =	$[_{ZP} Z]]$	

Then, we need an entry for B which associates it with a proper substructure of $[_{WP} W [_{XP} X [_{YP} Y [_{ZP} Z]]]]$, e.g. (19):

(19)	B ↔	$[_{XP} X [_{YP} Y [_{ZP} Z]]]]$
------	-----	-----------------------------------

Now, B will appear in the cell right under A in the paradigm, blocking A by the EP:

(20)	Dative =	$[_{WP} W [_{XP} X [_{YP} Y [_{ZP} Z]]]]$	= A
	Genitive =	$[_{XP} X [_{YP} Y [_{ZP} Z]]]]$	= B
	Accusative =	$[_{YP} Y [_{ZP} Z]]$	
	Nominative =	$[_{ZP} Z]]$	

But notice that now A cannot reappear in any of the cells remaining to be filled, since the SuperP allows B to lexicalize these too, and the EP continues to favor B over A.

6 Cumulative decomposition

At this point, a comment on NS vs. DM is in order. The treatment of syncretism described above depends on the different members of a paradigm not being characterized in syntactic terms by a single feature each (binary or not), but rather by sets of primitive features related by set inclusion as in (15). This falls out naturally from phrasal lexicalization in a cartographic approach to syntactic representations where each syntactic terminal is a single feature and syntactic structures are composed by merging the features/terminals one by one starting from the bottom of a “functional sequence” and stopping at different points.

However, one might of course also assume cumulative decomposition within a theory that rejects phrasal lexicalization and takes syntactic terminals to be associated with feature bundles, as in DM. (21) is like (15) except that features are not represented as the terminals of a syntactic tree, but rather as members of unstructured sets each of which is associated with a single syntactic terminal:

⁹ For accounts of *ABA outside NS, see Bobaljik (2012), Bobaljik and Sauerland (2018).

- (21) Dative = { W, X, Y, Z }
 Genitive = { X, Y, Z }
 Accusative = { Y, Z }
 Nominative = { Z }

Lexical Insertion controlled by The Subset Principle and DM's version of The Elsewhere Principle can then produce the licit syncretism patterns, but not any ABA patterns. The dative will be syncretic with the genitive, for example, if the lexicon contains the entry in (22)a, but not also the one in (22)b:

- (22)a $A \leftrightarrow \{ X, Y, Z \}$
 b $B \leftrightarrow \{ W, X, Y, Z \}$

An ABA pattern cannot arise, because an A filling the higher cells in the paradigm will stop extending downwards only when it reaches a cell associated with a feature set smaller than the set associated with A in the lexicon. At this point, some B will appear. But A will never come back, since each cell down from the one now filled by B is also associated with a feature set smaller than that of A.

Thus, the choice between DM and NS cannot be based on their ability to exclude ABA-patterns.¹⁰

7 Only a constituent can be lexicalized by a single morpheme

It follows from the SuperP that a single lexical item can only replace a set of terminals in a syntactic tree when those terminals form a constituent:

- (5) The Superset Principle
 A morpheme A with the entry $A \leftrightarrow T$ can lexicalize a syntactic structure S if and only if S is identical to a subtree of T.

For example, morpheme A with the lexical entry in (23) cannot lexicalize X and Y in a tree like (24):

- (23) $A \leftrightarrow [\alpha X [\beta Y]]$
 (24) $[\alpha X [\beta Y [\gamma Z]]]$

This is because X and Y in (24) do not form a constituent, hence not a subtree of (23).

The restriction of lexicalization to constituents is no impediment to our analysis of portmanteau suffixes. The ones in (25), for example, are easily accounted for if we assume the syntactic structures in (26) (where A is the adjectival root, m = masculine and f = feminine):

- (25) masculine feminine (Italian)
 singular *ross-o* *ross-a* 'red'
 plural *ross-i* *ross-e*

- (26) masculine feminine
 singular $[\alpha A [\beta \text{Sg} [\gamma m]]]$ $[\alpha A [\beta \text{Sg} [\gamma f]]]$
 plural $[\alpha A [\beta \text{Pl} [\gamma m]]]$ $[\alpha A [\beta \text{Pl} [\gamma f]]]$

The lexical entries are:

- (27)a $o \leftrightarrow [\beta \text{Sg} [\gamma m]]$ c $a \leftrightarrow [\beta \text{Sg} [\gamma f]]$
 b $i \leftrightarrow [\beta \text{Pl} [\gamma m]]$ d $e \leftrightarrow [\beta \text{Pl} [\gamma f]]$

¹⁰ Neither NS nor DM excludes ABA-patterns arising from accidental homonymy. The *ABA Generalization rests on the assumption that instances of accidental homonymy can be weeded out by statistical methods, e.g. as in Bobaljik and Sauerland (2013).

The lexicalization in (28) is legitimate with respect to the SuperP, because the two lower heads in (26) form a constituent:

	masculine	feminine
(28) singular	[A [Sg [m]]] <i>o</i>	[A [Sg [f]]] <i>a</i>
plural	[A [Pl [m]]] <i>i</i>	[A [Pl [f]]] <i>e</i>

It is likely that the suffixes in (25) spell out higher heads that the adjective has moved across, e.g. as an instance of spell-out driven movement of the sort discussed in Pantcheva (2011), Starke (2018) and Caha et al. (2018).¹¹ We then have the structures in (29) instead of those in (26) and must assume that traces are invisible to the lexicalization procedure which therefore treats (29) as (26) or that spell-out driven movement simply doesn't leave traces:

	masculine	feminine
(29) singular	[_β A [_β Sg [_γ m [_δ A]]]]]]]	[_β A [_β Sg [_γ f [_δ A]]]]]]
plural	[_β A [_β Pl [_γ m [_δ A]]]]]]]	[_β A [_β Pl [_γ f [_δ A]]]]]]]

But the restriction to constituents seems to rule out portmanteau prefixes. That is, there should be no language in which portmanteau morphemes like those in (25) are prefixes:

	masculine	feminine	(inverse Italian)
(30) singular	<i>o-ross</i>	<i>a-ross</i>	
plural	<i>i-ross</i>	<i>e-ross</i>	

In particular, the lexical entries in (27) cannot produce the portmanteau prefixes in (30) from the structures in (31) where there is no constituent [Sg/Pl [m/f]]:

	masculine	feminine
(31) singular	[_β Sg [_γ m [_α A]]]]]]	[_β Sg [_γ f [_α A]]]]]]
plural	[_β Pl [_γ m [_α A]]]]]]]	[_β Pl [_γ f [_α A]]]]]]]

Yet, there are in fact portmanteau prefixes, e.g. the nominal class prefixes in Bantu languages. As suggested in (32) (a small sample of the Xhosa paradigm), the shape of a class prefix seems to be determined by both gender and number:¹²

	gender 1	gender 4	
(32) singular	<i>u-m-ntu</i> 'person'	<i>i-si-kolo</i> 'school'	(classes 1 and 7)
plural	<i>a-ba-ntu</i>	<i>i-zi-kolo</i>	(classes 2 and 8)

This seems like prima facie evidence that restricting lexical insertion to constituent is not viable, and in fact, there is a theory of lexicalization that like NS allows a single morpheme to lexicalize more than a single terminal, but does not require those terminals to form a constituent. Instead, they must be part of a "span", i.e. a sequence of terminals (x_1, \dots, x_n) in a syntactic structure where each x_i c-commands x_{i+1} and there is no terminal y such that x_i c-commands y and y c-commands x_{i+1} , e.g. as in Ramchand (2008), Svenonius, (2012, 2014). On this approach, morphemes with the entries in (33) would effortlessly lexicalize the two topmost terminals in (34) (where g_1 and g_4 are gender features/heads):

¹¹ The authors cited argue that spell-out driven movement is a last resort option applying only when lexicalization of some phrase would otherwise fail; see footnote 4.

¹² The text discussion leaves aside the initial vowel (the "augment") in (32). This is a separate morpheme that falls away in certain syntactic contexts.

- (33)a $m \leftrightarrow (\text{Sg}, g1)$ c $si \leftrightarrow (\text{Sg}, g4)$
 b $ba \leftrightarrow (\text{Pl}, g1)$ d $zi \leftrightarrow (\text{Pl}, g4)$

- (34)a $[\alpha \text{ Sg } [\beta g1 [\gamma ntu]]]$ c $[\alpha \text{ Sg } [\beta g4 [\gamma kolo]]]$
 b $[\alpha \text{ Pl } [\beta g1 [\gamma ntu]]]$ d $[\alpha \text{ Pl } [\beta g4 [\gamma kolo]]]$

However, it is not obvious that the nouns in (32) actually have the right-branching structures in (34). In particular, the correct structures might be taken to be the ones in (35), where the terminals Sg/Pl and g1/4 form a constituent, i.e. a specifier on the main projection line:

- (35)a $[\delta [\alpha \text{ Sg } [\beta g1]] [\gamma ntu]]$ c $[\delta [\alpha \text{ Sg } [\beta g4]] [\gamma kolo]]$
 b $[\delta [\alpha \text{ Pl } [\beta g1]] [\gamma ntu]]$ d $[\delta [\alpha \text{ Pl } [\beta g4]] [\gamma kolo]]$

There is a problem with these structures, though. Nominal functional heads like Sg/Pl and g1/4 are generally assumed to merge on top of a noun forming a strictly right-branching extended projection. The structures in (35) do not conform to this assumption. But those in (36) do:

- (36)a $[\delta [\alpha \text{ Sg } [\beta g1 [\gamma N_1]]] [\epsilon ntu]]$ c $[\delta [\alpha \text{ Sg } [\beta g4 [\gamma N_4]]] [\epsilon kolo]]$
 b $[\delta [\alpha \text{ Pl } [\beta g1 [\gamma N_1]]] [\epsilon ntu]]$ d $[\delta [\alpha \text{ Pl } [\beta g4 [\gamma N_4]]] [\epsilon kolo]]$

Of course, assuming that those in (36) are the right structures, makes the portmanteau prefixes in (32) compatible with constituent lexicalization only if the class prefixes also lexicalize the extra $N_{1/4}$ in addition to Sg/Pl and g1/4:

- (37)a $m \leftrightarrow [\alpha \text{ Sg } [\beta g1 [\gamma N_1]]]$ c $si \leftrightarrow [\alpha \text{ Sg } [\beta g4 [\gamma N_4]]]$
 b $ba \leftrightarrow [\alpha \text{ Pl } [\beta g1 [\gamma N_1]]]$ d $zi \leftrightarrow [\alpha \text{ Pl } [\beta g4 [\gamma N_4]]]$

In other words, an analysis based on (36) will claim that the nominal class prefixes in Bantu languages incorporate a noun which according to Taraldsen et al. (2018) should be viewed as a classifier.

Evidence that this analysis is correct comes from (among other things) the fact that the gender feature that controls the choice of prefix cannot in general originate from the noun hosting it:

- (38)a $u\text{-}m\text{-}ntu$ ‘person’ vs. $u\text{-}bu\text{-}ntu$ ‘humanity’ (class 1 vs. class 14)
 b $u\text{-}m\text{-}thi$ ‘tree’ vs. $u\text{-}lu\text{-}thi$ ‘stick’ (class 3 vs. class 11)

This is consistent with the analysis in (36)–(37) which take the gender feature associated with the prefix to come from the classifier noun.¹³ Taraldsen et al. (2018) and Taraldsen (2018) argue that similar evidence can also be found in some languages with nominal class prefixes outside the Bantu family.

Notice, however, that even if nominal roots must combine with classifier nouns in Bantu, it doesn’t follow that the classifier and its associated number and gender heads must form a specifier rather than merge one by one to form strictly right-branching structures like (39):

- (39) $[\alpha \text{ Sg } [\beta g1 [\gamma N_1 [\epsilon ntu]]]]]$

Ultimately, this must follow from Starke’s (2018) proposal that specifiers like those in (37) are formed by one of a set of spell-out driven operations applying just in case the strictly right-branching structures cannot otherwise be lexicalized (see Starke’s article for the details). In particular, if, as mentioned in footnote 3, lexical insertion only applies to phrasal nodes, never to terminals, Sg, g1 and N_1 in a structure like (39) cannot be lexicalized at all unless there is some morpheme that can lexicalize the whole structure (in which case no prefix will surface).

¹³ Actually, gender ultimately turns out to be irrelevant to the analysis of the class prefixes in Bantu, if Taraldsen et al. (2018) are correct.

8 Envoi

My purpose in writing this article has been to stimulate curiosity about Nanosyntax. The works cited in the References provide a good starting point for further exploration. But the best way to proceed is perhaps to try resolving some puzzles. So, to conclude I offer a problem.

In Czech, most adjectives form comparatives as in (39) (from Caha et al. 2018):

(39) *mil-ý* ‘nice/kind’ – *mil-ěj-š-í* ‘nicer/kinder’

(Note that *ý* and *í* are agreement markers.)

Some adjectives have shorter comparative forms:

(40) *star-ý* ‘old’ – *star-š-í* ‘older’

Also, suppletive roots occur in some comparatives:

(41) *dobr-ý* ‘good’ – *lep-š-í* ‘better’

Caha et al. (2018: Section 6.4) note that suppletion only occurs in short comparatives, i.e. with adjectives that follow the pattern in (40). In other words, there are no cases like (42) (modulo the special cases discussed in Caha et al. (2018: Section 6.5)):

(42) **A-ý* – *B-ěj-š-í*

Why should this be so? (A plausible answer is found in Caha et al. (2018 Section 6.4).)

References

- Baunaz, Lena & Eric Lander. 2018. Nanosyntax: The basics. In Lena Baunaz, Liliane Haegeman, Karen De Clercq & Eric Lander (eds.), *Exploring nanosyntax*, 3–56. Oxford: Oxford University Press.
- Bobaljik, Jonathan. 2012. *Universals in comparative morphology*. Suppletion, superlatives, and the structure of word. Cambridge, MA: The MIT Press.
- Bobaljik, Jonathan & Uli Sauerland. 2013. Syncretism distribution modeling: Accidental homonymy as a random event. In Nobu Goto, Koichi Otaki, Atsushi Sato & Kensuke Takita (eds.), *Proceedings of GLOW in Asia IX*, 31–53. Mie, Japan: Mie University.
- Bobaljik, Jonathan & Uli Sauerland. 2018. ABA and the combinatorics of morphological features. *Glossa: A Journal of General Linguistics* 3(1). 15.
- Caha, Pavel. 2009. The nanosyntax of case. PhD dissertation. University of Tromsø.
- Caha, Pavel. 2018. Notes on insertion in Distributed Morphology and nanosyntax. In Lena Baunaz, Liliane Haegeman, Karen De Clercq & Eric Lander (eds.), *Exploring nanosyntax*, 57–87. Oxford: Oxford University Press.
- Caha, Pavel, Karen De Clercq & Guido Vanden Wyngaerd. 2018. The fine structure of the comparative. Available at <<http://ling.auf.net/lingbuzz/003790>>.
- Cinque, Guglielmo & Luigi Rizzi. 2008. The cartography of syntactic structures. *CISCL Working Papers on Language and Cognition 2*, University of Siena (distributed by MIT Working Papers in Linguistics), 43–59.
- Fábregas, Antonio. 2007. The exhaustive lexicalization principle. *Nordlyd* 34(2). 165–199.
- Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In K. Hale & S. J. Keyser (eds.), *The view from building 20; Essays in honor of Sylvain Bromberger*, 111–176. Cambridge, MA: The MIT Press.
- Halle, Morris & Alec Marantz. 1994. Some key features of distributed morphology. In Andrew Carnie, Heidi Harley & Tony Bures (eds.), *Papers in Phonology and Morphology. MIT Working Papers in Linguistics* 21. 275–288.
- Pantcheva, Marina. 2011. Decomposing path: The nanosyntax of directional expressions. PhD dissertation. University of Tromsø.
- Ramchand, Gillian. 2008. *Verb meaning and the lexicon: A first-phase syntax*. Cambridge: Cambridge University Press.
- Starke, Michal. 2009. Nanosyntax: A short primer to a new approach to language. *Nordlyd* 36(1). 1–6.

- Starke, Michal. 2011. Towards elegant parameters: Language variation reduces to the size of lexically stored trees. Available at <<http://ling.auf.net/lingbuzz/001183>>.
- Starke, Michal. 2018. Complex left branches, spell-out and prefixes. In Lena Baunaz, Liliane Haegeman, Karen De Clercq & Eric Lander (eds.), *Exploring nanosyntax*, 239–249. Oxford: Oxford University Press.
- Svenonius, Peter. 2012. Spanning. Available at <<http://ling.auf.net/lingbuzz/001351>>.
- Svenonius, Peter. 2014. Spans and words. In Daniel Siddiqi and Heidi Harley (eds.), *Morphological metatheory*, 201–222. Amsterdam: John Benjamins.
- Taraldsen, Tarald. 2018. Spanning vs. constituent lexicalization: The case of portmanteau prefixes. In Lena Baunaz, Liliane Haegeman, Karen De Clercq & Eric Lander (eds.), *Exploring nanosyntax*, 88–110. Oxford: Oxford University Press.
- Taraldsen, Knut Tarald, Lucie Taraldsen Medová & David Langa. 2018. Class prefixes as specifiers in Southern Bantu. *Natural Language & Linguistic Theory* 36(4). 1339–1394.