Unmerging Analytic Comparatives

Karen De Clercq   Guido Vanden Wyngaerd

1 Introduction

In this paper, we look at the internal structure of the analytic comparative marker *more*, arguing that it spells out nearly all the features of a gradable adjective. When this marker is merged with an adjective in the positive degree, it creates a situation of feature recursion or overlap, where *more* duplicates certain features that are also present in the adjective that it modifies. We argue that such overlap must be disallowed as a matter of principle. We present an empirical argument in favour of such a restriction, which is based on the generalisation that comparative markers which occur to the left of the adjectival root are incompatible with suppletion. This generalisation can be shown to follow from a restriction against overlapping derivations. In order to achieve such nonoverlapping derivations, an Unmerge operation may remove previously created structure.

The paper is structured as follows: in section 2 we lay out some of the prerequisites for our analysis having to do with the structure of the comparative, and the difference between analytic and synthetic comparatives in English. Section 3 presents the analysis of PRE and POST marking developed by Starke (to appear), and applies it to the case of the comparative. In section 4 we investigate the internal structure of *more*. Finally, in section 5 we present and explain the generalisation that PRE marking is incompatible with suppletion.

2 Prerequisites

2.1 The structure of the comparative

Caha (2017), De Clercq & Vanden Wyngaerd (2017) propose that the com-

parative head CMPR of Bobaljik (2012) is to be split up into two distinct heads, C1 and C2. This more fine-grained structure of the comparative receives empirical support from the regular comparative in Czech, which is formed with the suffix *-ějš-*:[1]

(1)

Pos	Cmpr	
bujar-ý	bujař-ejš-í	'merry'
červen-ý	červen-ějš-í	'red'
hloup-ý	hloup-ějš-í	'stupid'
moudr-ý	moudř-ejš-í	'wise'

Caha (2017) presents two pieces of evidence suggesting that *-ějš-* in fact consists of two parts, i.e. that it is to be segmented as *-ěj-š-*. The first is that with certain adjectives, the first exponent (*-ěj-*) disappears, as is shown in (2). The second exponent (*-š-*) systematically disappears with comparative adverbs, as shown in (3).

(2)

Pos	Cmpr	
star-ý	star-š-í	'old'
such-ý	suš-š-í	'dry'
drah-ý	draž-š-í	'expensive'
tvrd-ý	tvrd-š-í	'hard'
tich-ý	tiš-š-í	'silent'

(3)

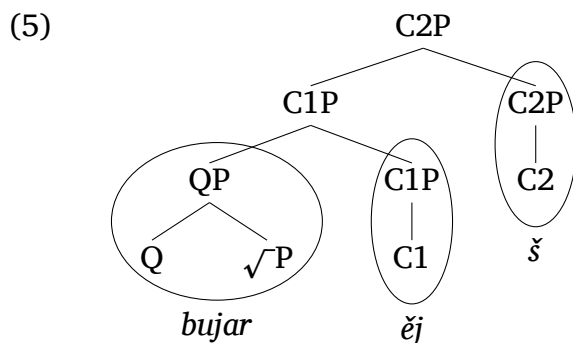| Cmpr Adj	Cmpr Adv	
červen-ěj-š-í	červen-ěj-i	'redder'
hloup-ěj-š-í	hloup-ěj-i	'sillier'
moudř-ej-š-í	moudř-ej-i	'wiser'
rychl-ej-š-í	rychl-ej-i	'faster'

This leads to the natural conclusion that these exponents spell out different heads, the C1 and C2 heads shown in (4):[2]

---

[1]The final *í/ý* is an adjectival agreement marker, which I shall henceforth ignore. The grapheme *ě* corresponds to an *e* which triggers the palatalisation of the preceding consonant, e.g. in *bujařejší* 'merrier'.

[2]The structure in (4) is adjusted to fit the surface order of Czech, which is the mirror image ordering of the underlying functional sequence (see Cinque 2005). As in Cinque's work, we analyse mirror orders as arising through a series of standard movement oper-

(4)

```
            C2
          /    \
        C1      C2
       /  \      |
      A    C1    |
      |    |     |
    bujar  ěj    š
```
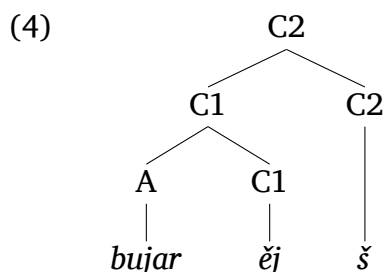
Before we proceed, we want to modify the structure in (4) somewhat. We take the head A to have internal structure itself, being composed of a root feature $\sqrt{}$ and a gradability feature Q (De Clercq & Vanden Wyngaerd to appear). The adjectival roots of the *bujar* type therefore do not merely realise the terminal node A, but a phrasal node QP, which is composed of Q and $\sqrt{}$.[3] Nanosyntax in fact assumes that there is never insertion under a terminal node, but that all insertion is insertion at the phrasal level. The full tree of *bujařejší* 'merrier', which moreover fully reflects its derivational history, therefore looks as in (5). The circles indicate which phrase is realised by which exponent.

(5)

```
                    C2P
                  /     \
               C1P       C2P
              /    \       |
            QP     C1P     C2
           /  \     |      |
          Q   √P    C1     š
          |   |     |
        bujar     ěj
```

As we shall see in section 3 below, a structure of the form $[_{C1P}$ C1] is the derivational remnant of a movement of the complement of C1 (QP in this case) to its left; in the same way, the complement of C2 (C1P) has been

---

ations, which affect phrasal nodes and move them either spec-to-spec, or, as in the case of (5) below, in a roll-up fashion. For details of how these derivations work, we refer the reader to Caha et al. (2017).

[3]We use the notation $\sqrt{}$ without any commitment at this point as to whether this is really the bottom of the hierarchy, or whether $\sqrt{}$ needs to be further decomposed into even more functional heads. In fact, we shall argue below that there is another functional head STATE between Q and $\sqrt{}$.

raised to its left.

    With this much in place, let us address the question of what explains the distinction between the adjectives in (1) (i.e. those of the *bujar* 'merry' type) and those in (2) (i.e. those of the *star* 'old' class). The distinction is a lexical one, i.e. it is not predictable from any phonological or other property. Rather than propose a lexical diacritic to distinguish the two types of adjectives, we assume (following Caha 2017) that it can be accounted for without any auxiliary assumptions by the mechanisms of nanosyntax, in particular phrasal spellout, and the assumption that lexical entries contain syntactic structures (Starke 2014). Applying these assumptions to the case of the adjectives in (2), we assume that an adjective like *star* 'old' spells out the node C1P, as shown in (6). Accordingly, its lexical entry is as in (7a). Adjectives like *bujar* 'merry' spell out a smaller node, as shown in (5); they have a lexical entry like the one in (7b).

(6)
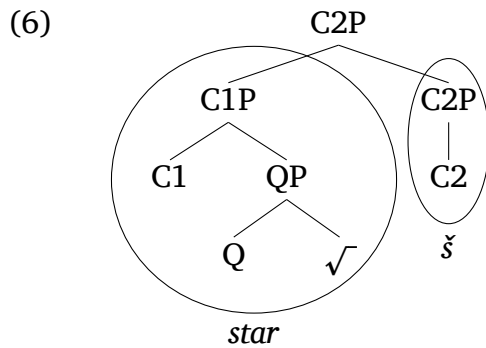
(7) $[_{C1P}$ C1 $[_{QP}$ Q $\sqrt{}$ $]]$ $\Longleftrightarrow$ /star/
            $[_{QP}$ Q $\sqrt{}$ $]$ $\Longleftrightarrow$ /bujar/

The difference between how these two classes of adjectives mark the comparative now automatically follows: since adjectives of the *bujar* 'merry' class spell out QP, they need a separate exponent for realising both C1 and C2, which explains the presence of both the *-ěj-* and the *-š-* markers in the comparative. For completeness, the lexical entries of these markers are given below:

(8)   a. $[_{C1P}$ C1 $]$ $\Longleftrightarrow$ /ěj/
       b. $[_{C2P}$ C2 $]$ $\Longleftrightarrow$ /š/

Adjectives of the *star* 'old' class, in contrast, already spell out C1 them-

selves, so that C1 cannot be realized by *-ěj-*, leaving only the C2 marker *-š-* in the comparative. Another way of representing these patterns is in terms of the table in (9), which shows the tradeoff between the size of the root and the number of suffixes realised:

(9)

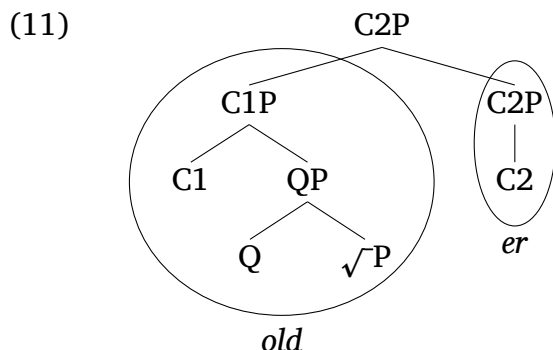| √	Q	C1	C2
bujar		ěj	š
star			š

2.2 English

Caha (2017) applies the account of Czech to the well-known distinction in English between the synthetic comparative (with *-er*), and the analytic one (with *more*). He argues that, like in Czech, the difference is a lexical one, which involves the size of the adjectival root. Adjectives with synthetic comparatives spell out C1P, those with analytic comparatives spell out QP. Accordingly, the suffix *er* spells out less structure than those with *more*. The relation between the adjectival root and the markers of the comparative can again be shown in a table akin to the one in (9):

(10)

√	Q	C1	C2
intelligent		mo	re
smart			er

What this table implies is that *-er* spells out the C2 feature (like Czech *š*), whereas *more* is internally complex and spells out both the C1 and the C2 features. The latter assumption is not entirely new, as other authors have noted that *more* is the comparative of *much,* and as such internally complex (e.g. Corver 1997, Bobaljik 2012, Embick 2007). The claim that the distinction between analytic and synthetic comparatives is lexically determined is not new either (see e.g. Bobaljik 2012: 164, Gouskova & Ahn 2016).

The tree representation of a synthetic comparative like *older* shows a striking resemblance with the one for *star* 'older' in (6) above:

(11)

```
                    C2P
           _____
          /                      \
        C1P                      C2P
       /    \                     |
     C1      QP                   C2
            /   \
           Q    √P                er
         old
```

The lexical entries for the two types of adjectives likewise are identical to those given for the two types of adjectives in Czech in (7) above (except of course for their phonologies):

(12) $[_{C1P}$ C1 $[_{QP}$ Q √]]   ⟺   /old/
         $[_{QP}$ Q √]   ⟺   /intelligent/

And the lexical entry for the English marker of the synthetic comparative is identical to that the Czech -*š*- marker (see (8b) above).

(13) $[_{C2P}$ C2] ⟺ /-er/

However, when we look at English comparatives with *more,* the parallel with Czech breaks down. While Czech marks the comparative entirely suffixally, English *more* occurs to the left of the adjective. In the following section, we investigate the reason for this difference in the light of the theory of PRE vs POST marking proposed in Starke (to appear). The examination of this question will naturally lead to a further probe into the internal structure of *more* in section 4.

3   PRE vs POST

As we saw in the previous section, the Czech comparative has exclusive POST marking: it is to the right of the stem (i.e. suffixal), and displays mirror principle ordering with respect to the underlying functional sequence. PRE marking, in contrast, involves functional material to the left of the stem, whose ordering reflects the underlying order of the functional sequence. Starke (to appear) suggestst that this difference reflects two modes

of combination, Merge-F and Merge-XP. POST marking involves exclusively Merge-F, whereas PRE marking involves both Merge-F and Merge-XP. Let us show how this works with an abstract example. Consider first a tree resulting from Merge-F:

(14)

K3
 K2
  K1
   Z
    Y   X
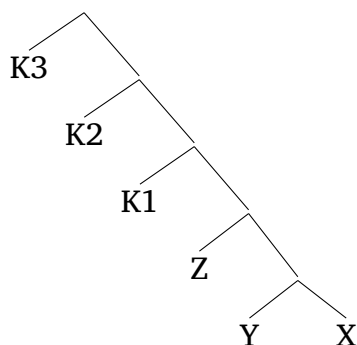
After each Merge step, the lexicon is consulted for a spellout. If no suitable lexical entry is found, spellout-driven movement will apply to create a structure that can be spelled out. This follows from the formulation of the spellout algorithm, a preliminary and incomplete version of which we provide in (15):

(15)   Merge F and
   a.   Spell out FP
   b.   If (a) fails, move the (spec of the) complement of F, and retry (a)

Suppose that in the derivation of (14), spellout according to (15a) succeeds until the derivation reaches ZP. Then, at the merger of K1, (15a) fails and (15b) applies, moving ZP (in successive steps) to the spec of K3P.[4] This will result in POST marking, as indicated in (16):

---

[4]To derive the tree in (16), the complement of K1 needs to be moved first, and then after the merger of K2 and K3 the spec of their complement. We gloss over these complications, as they are orthogonal to our concerns. For an in-depth discussion of the mechanics of the nanosyntactic spellout algorithm, we refer the reader to Caha et al. (2017).

(16)

K3P

Z

Y   X

root

K3

K2   K1P

K1

POST marker
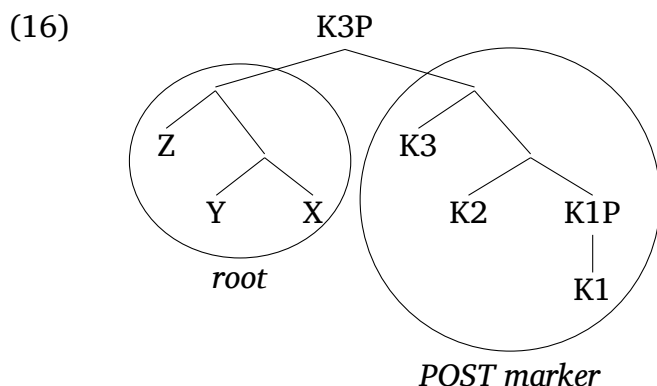
Note that it follows from the formulation of the spellout algorithm that, in the absence of movement, any lexical item must minimally spell out two features: this is because Merge applies before the lexicon is consulted, and Merge is a binary operation. A lexical item can only spell out a single feature if movement has applied (by (15b)). In either case, it also follows from the way (15) works that spellout can only apply to phrases, and never heads, as spelling out heads would require access to the post-syntactic lexicon prior to the first application of Merge.[5]

PRE markers are specifiers that are merged in a separate workspace, and that subsequently get merged into the main derivation. This process consists of two steps. The first step involves creating the PRE marker in a separate workspace, which happens through the process of Merge-F discussed above. The second step is to merge the XP so created into the main spine as a specifier by an application of Merge-XP. The product of such a merger is shown in (17):
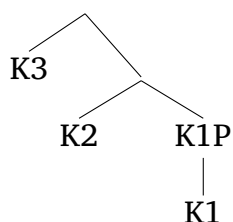
(17)

K3P

K3

K2   K1

PRE marker

Z

Y   X

root
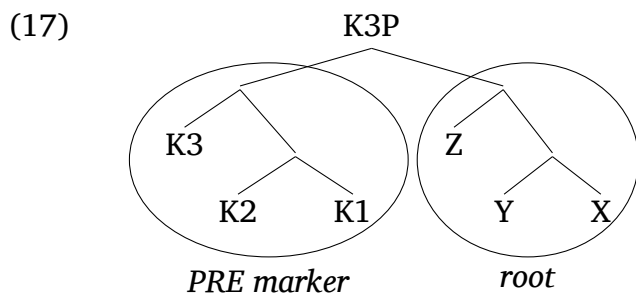
In contrast to the nonprojecting specifiers which are the result of move-

---

[5]This issue is independent of the existence of Self Merge, i.e. a First Merge step that would merge α with itself, yielding the singleton set $\{\alpha\}$ (Adger 2013).

ment, specifiers coming out of a separate workspace project in the main derivation (as discussed in Starke 2004). This is because they provide features that are required in the main derivation by the functional sequence. This can be seen by comparing (17) with the underlying functional sequence in (14) above: (17) basically provides that underlying functional sequence, except that it is 'chunked' into two different constituents, which correspond to two different pieces of phonology, each of which spells out a part of the f_{seq}.

Notice now a subtle difference between the POST marker in (16) and the PRE marker in (17): the suffix has at its bottom a K1 head that projects to K1P. This structure is the trace of the movement of the complement of K1, ZP, at an earlier stage in the derivation. In contrast, the prefix in (17) does not have a projecting K1: instead, K1 is dominated by the label of its sister, K2. This derivational difference is at the same time a lexical difference: PRE material has a binary bottom, POST material has a unary bottom:

(18)   POST: unary bottom   (19)   PRE: binary bottom
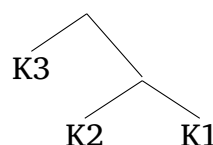
With this in mind, we can return to the derivation of English analytical comparatives, and the internal structure of *more*.

4   More

Analytic comparatives with *more* are triggered by the need to spell out the C1 feature in cases where the adjectival root does not spell out C1. Recall from the above discussion that in English C1 can either be provided by the adjectival root, or by *more*. This was shown schematically in (10), repeated here:

(10)

√	Q	C1	C2
intelligent	mo	re
smart		er

Let us now consider in more detail how this works derivationally. Suppose the syntax merges QP, consults the lexicon, and spells out QP as *intelligent*. Subsequent merger of C1 will not lead to a successful spellout, as *intelligent* does not spell out this feature, given its lexical entry (see (12) above). Applying spellout-driven movement of QP to adjoin to C1P will be to no avail, as the English lexicon has no lexical entry that spells out just C1P. This remains true after merging C2 and further movement of QP to C2P: although *-er* spells out C2, it does not spell out C1, and insertion fails. Since spellout-driven movement has reached a dead end, and in order to supply a spellout for the C1 feature, a new derivation is spawned in a separate workspace. The spellout of the C1 feature will ultimately be provided by *more*. From this logic, it follows that the C1 feature has to be present in the lexical entry for *more*. But it cannot be the case that this is the only feature that *more* spells out: in the previous section we have seen that PRE markers have a binary bottom, so they must at least spell out two features. An obvious candidate is the C2 feature, so that, as a first approximation, the entry for *more* minimally looks as follows:

(20) $[_{C2P}$ C2 C1] $\iff$ /more/

Once the separate derivation successfully spells out as *more,* it is integrated into the main projection line through Merge-XP:

(21)
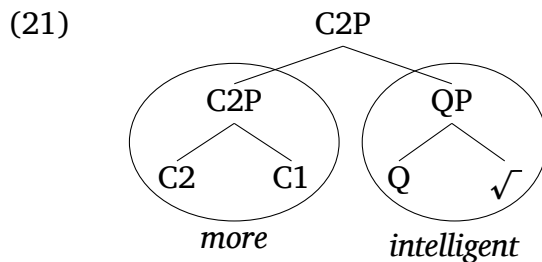
There are a few reasons, however, why the lexical entry for *more* in (20) cannot be quite right yet. In particular, there is reason to assume that *more* realises additional features beyond C1 and C2. The first reason is that *more* can not only occur as a marker of the comparative with adjectives, but also

as an adverb with verbs and as an adnominal modifier:

(22)   a.   They laughed more than I expected.
        b.   She needs to eat more vegetables.

This suggests that *more* functions as a gradable adjective (or adverb) itself, spelling out the full functional sequence of a gradable adjective in the comparative. Also, *more* does not just express the comparative of adjectives like *intelligent*, but it is also itself the (suppletive) comparative of *much*. The existence of the three steps of comparison *much–more–most* again suggests that *more* is a gradable adjective in its own right. This is further confirmed by the fact that there exist analytic 'comparatives of inferiority' with *less* (e.g. *less intelligent*). These are notoriously absent with synthetic comparatives, a property dubbed *Lesslessness* by Bobaljik (2012: 4):

(23)   Lesslessness
        No language has a synthetic comparative of inferiority.

This fact also suggests that *more* and its negative counterpart *less* have richer internal structure than the suffixal comparative marker *-er*.

    The assumption that *more* realises more features than just C1 and C2 finds additional support in an observation made in Matushansky (2013) to the effect that non-gradable adjectives (like *French, right* or *male*) do not form morphological comparatives in English:

(24)   *Frencher, *righter, *maler

However, the same adjectives can form analytic comparatives with a (coerced) gradable interpretation, i.e. as meaning 'having more properties associated with being French, right or male':

(25)   a.   *Becky's uncle is Frencher/righter/maler than Napoleon.
        b.   Becky's uncle is more French/more right/more male
              than Napoleon.

Assume that the lexical entries for these nongradable adjectives lack the gradability feature Q. Given our functional sequence assumed so far, this would imply that they spell out just $\sqrt{}$P. But we have also assumed earlier that spellouts that do not involve movement need to spell out minimally two features, given that the first consultation of the lexicon occurs after

11

the first application of Merge. We must therefore conclude that there is an additional feature between Q and $\sqrt{}$ in adjectives like *French, right, male*, etc. We shall assume that this is a feature STATE (or S for short). S can be realized by the verb *be* in a verbal environment, but also by adjectives. They become gradable adjectives through the addition of the Q feature. This means that the functional sequence for gradable adjectives must be amended as shown in (26):

(26)

```
        QP
       /  \
      Q    SP
          /  \
         S    √
```
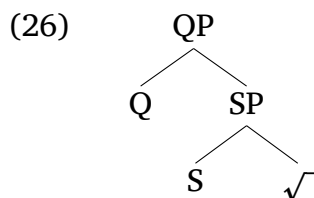
Nongradable adjectives spell out SP. They cannot appear in a morphological comparative for the same reason that *intelligenter* is ruled out: not all the features required by the functional sequence have a spellout. In the case of *intelligenter*, C1 did not get a spellout, in the case of *Frencher* it is both the C1 and the Q-feature that fail to be realised. The existence of analytic comparatives as in (25b) confirms our earlier claim that *more* also provides the gradability feature Q, so that a legitimate syntactic structure can be built where all the features of a comparative adjective are spelled out. This is depicted in (27):

(27)

C2	C1	Q	S	$\sqrt{}$
	more		French

This is almost the complete internal structure for *more*, except for one fact: *more* is the suppletive comparative of *much*. Following Starke (2014), we assume that a suppletive form contains a pointer to its nonsuppletive counterpart, i.e. the lexical entry of *more* contains a pointer to another, existing, lexical entry, namely *much*.[6] We represent this as follows:

---

[6]Extensive discussion of the nanosyntactic treatment of suppletion in degree comparison can be found in Caha (2017), De Clercq & Vanden Wyngaerd (2017), Caha et al. (2017).

(28)   C2P   ⟺   /more/
       / \
     C2   C1P
          / \
        C1   *much*
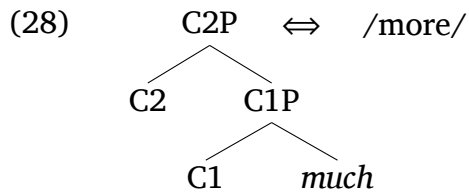
A consideration of the internal structure of *much* will therefore lead us to the full internal structure of *more*. A full investigation of this issue is beyond the scope of the present article (see De Clercq 2017 for discussion). There is, however, one important aspect of *much* that we want to consider here. This concerns the fact that a handful of English adjectives allow modification by *much* (Bresnan 1973, Corver 1997).

(29)   much alike/different/afraid/aware/reliant/dependent/offended

Since in those cases, *much* is a PRE marker, it must spell out at least two features. We take these features to be Q and S, i.e. the lexical entry of *much* is as in (30):

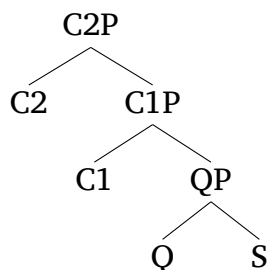(30)   QP   ⟺   /much/
       / \
      Q   S
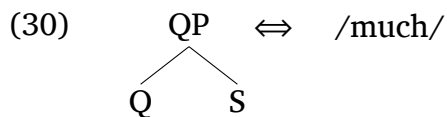
Putting this all together gives us the full functional sequence that is realized by *more* in (31). It will become important later on that *much* does not spell out a root feature, i.e. that *much* is a functional adjective, and that only lexical adjectives have the $\sqrt{}$ head.[7] For good measure, we add the tree that is realised by the negative counterpart of *more*, *less*, in (32):[8]

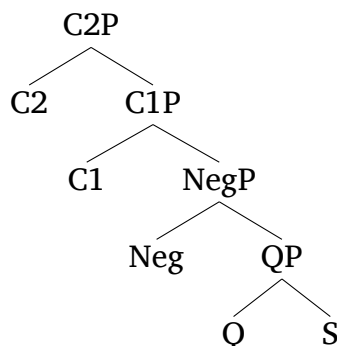[7]This difference between *much/more* and regular adjectives will potentially account for the fact that *much/more* is the only adjective that is able to modify other adjectives. Alternatively, we could treat *much/more* as adverbs, in which case an additional ADV feature would top the tree in (30).

[8]See De Clercq & Vanden Wyngaerd (to appear) for extensive discussion of the Neg feature in the adjectival f_{seq}.

(31)
```
        C2P
       /    \
     C2      C1P
            /    \
          C1      QP
                 /  \
                Q    S
```

(32)
```
        C2P
       /    \
     C2      C1P
            /    \
          C1      NegP
                 /    \
               Neg     QP
                      /  \
                     Q    S
```
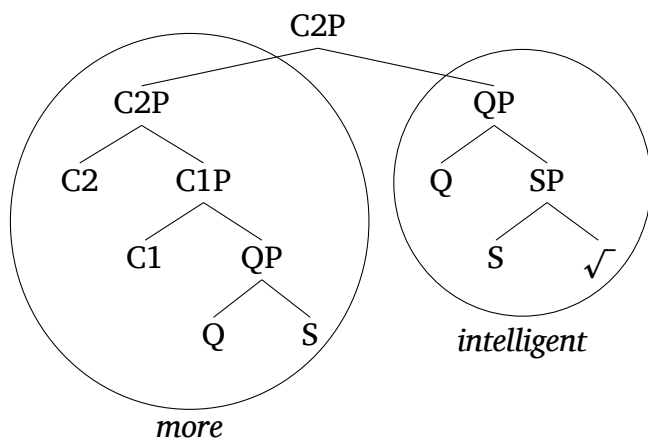
A question that arises at this point is whether PRE markers can recurse the functional sequence already spelled out in the main spine. We have seen that the PRE marker *more* spells out almost the entire functional sequence of a gradable adjective, and so when it combines with one to mark its comparative, the PRE marker replicates portions of the main spine:

(33)
```
                    C2P
         _____/      _____
        /   C2P          \    QP    \
       |   /    \          |  /  \    |
       |  C2     C1P       | Q    SP  |
       |        /    \     |     /  \ |
       |      C1      QP   |    S    √|
       |             /  \  |          |
       |            Q    S |          |
        _____/  _____/
               more          intelligent
```

In the above tree, we see a partial overlap between the features of the specifiers and those of the main spine in the features Q and S, which occur twice. In principle it is possible that the system allows feature overlap. However, in section 5 we will present an empirical argument in support of the idea that there should be no featural overlap between the specifier *more* and the main spine.

Before we provide empirical evidence for the idea that overlapping structures must be ruled out, we need to discuss how the overlap could be done away with. In principle this could be remedied in either of the two ways: remove the redundant features in the specifier, or remove them in

14

the main spine. In practice, only the latter option is available in virtue of both the Superset Principle and the way structure removal works. The correct nonoverlapping representation therefore looks like this:

(34)
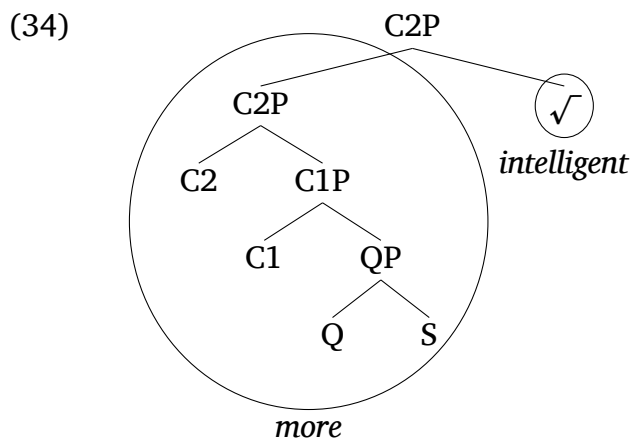
Compared to (33), we have eliminated the features Q and S from the main spine, and spelled out the one remaining $\sqrt{}$ feature as *intelligent*. The remaining structure realises all the features required to spell out a comparative degree adjective, using each feature only once. The alternative would consist in eliminating Q and S from the specifier, as shown in (35):

(35)
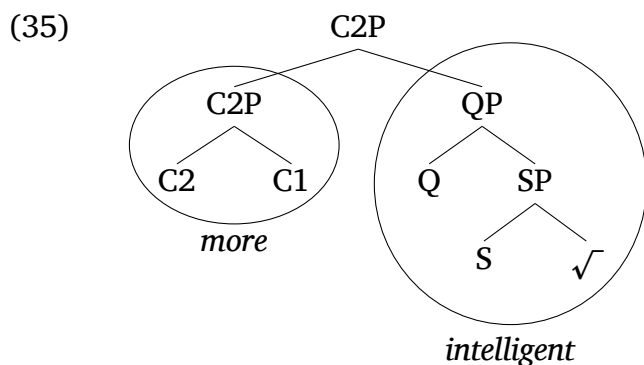
However, there are two reasons why this second type of derivation is problematic. For one thing, the spellout of *more* in the specifier violates the Superset Principle:

15

(36)   *The Superset Principle*
        A lexically stored tree can be inserted at a syntactic node iff the
        lexically stored tree contains the syntactic node.

Looking at the lexical tree for *more* in (28), we see that it does not contain the syntactic C2P node that forms the specifier in (35): the syntactic node misses the bottom part of the lexical tree, in particular the features Q and S. Spellout of the specifier therefore fails as a result of the Superset Principle. On top of that, the operation responsible for removing material merged earlier would have to start eating away features from the bottom of the structure, so as to remove the overlapping features Q and S from the specifier in (35). If it operates in the main spine, however, it removes features starting from the top of the derivation, such that the most recently merged feature is removed first, and so on. We shall call the operation in question Unmerge (see also Müller 2017, who proposes the operation Remove as a counterpart to Merge). Unmerge is the mirror image operation of Merge, and therefore can only remove features starting at the top. Unmerge (or backtracking as it is also called) is an option that can be applied if it turns out that a given derivation meets a dead end. In such a case, previous Merge steps can be undone and a different derivational strategy be attempted. So in the face of the impossibility of *intelligenter* discussed earlier, a derivation with a specifier may be attempted. This will yield *more*; this will in turn require the removal of the previously merged Q and S features in the main spine, in order to get the representation in (34) without overlap. In the next section, we shall investigate an empirical consequence of the two different derivational options in (34) and (35). It will turn out that the empirical evidence provides an additional argument for favouring (34) over (35).[9]

5   PRE marking and Suppletion

The empirical evidence that we want to discuss concerns a data pattern which relates to what Bobaljik (2012) calls the *Root Suppletion Generalisation* (RSG):

---

[9]Starke (to appear) requires overlap of a single feature with a projecting specifier. As the case of *more* shows, potential overlap may involve many more features than one.

16

(37)   *Root Suppletion Generalisation* (Bobaljik 2012: 3)
        Root suppletion is limited to synthetic (i.e., morphological) comparatives.

Data illustrating this generalisation comes from languages where the synthetic and the analytic comparative alternate, as in the examples below (Bobaljik 2012: 70-71):

(38)

	POS	Synthetic CMPR	Analytic CMPR	
French	mauvais	pire	plus mauvais	'bad'
Italian	buon-o	miglior-e	piú buon-o	'good'
Greek	kak-ós	cheiró-ter-os	pjo kak-ós	'bad'

In such cases, we see that only the synthetic comparative shows suppletion. What is striking about the data in (38), however, is that the analytic markers of the comparative are all PRE markers, and the synthetic ones are all suffixal. This is not only true for the sample of data in (38), but more generally. A detailed investigation of all the suppletive forms in Bobaljik (2012) reveals that nearly all the ones involving root suppletion have a suffixal comparative marker. For example, in Slavic only two languages (Bulgarian and Macedonian) have a prefixal comparative marker, and these are the only ones that have no comparative suppletion in the adjective for *good* (Bobaljik 2012: 45):

(39)

	POS	CMPR	SPRL
Bulgarian	dobər	po-dobər	naj-dobər
Macedonian	dobro	po-dobro	naj-dobro
Czech	dobr-ý	lep-ší	nej-lep-ší
Sorbian	dobr-y	redl-iši	
Serbian	dobar	bol-ji	naj-bol-ji
Ukranian	dobr-yj	krašč-yj	naj-krašč-yj
Ukranian	harn-yj	krašč-yj	
Russian	xoroš-ij	luč-še	(nai-luč-š-ij)

The table in (40) gives an overview of the data in Bobaljik (2012), arranged according to the type of marking. The first column gives the meaning of the adjective. The second column gives the total number of supple-

tive adjectival triples (i.e. positive, comparative, and superlative degree). The final four columns break this number down into the number of POST-marked forms, PRE-marked forms, and circumfixally marked ones, respectively. Some forms have no recognisable prefix or suffix (like *worse*), and these are listed in the final column (with PM for 'portmanteau').

(40)

	N	POST	PRE	CIRCUM	PM
GOOD	32	24	–	3	5
BIG	7	5	–	1	1
BAD	22	19	–	–	3
SMALL	9	6	–	–	3
MUCH, MANY	31	25	1	–	5
Total	101	79	1	4	17

It is striking that out of 101 suppletive triples, there is only one PRE marked form with suppletion, and four circumfixal ones, while all the others involve suffixal or portmanteau marking. The circumfixal forms are spurious, however. They come from two genetically related languages, Georgian and Svan. For Georgian, Gippert (1996) writes that the forms given by Bobaljik no longer have a comparative meaning in present-day Georgian, but only a 'superlative/elative' function, real comparatives being built analytically. And while noting that in Old Georgian comparatives 'were commonly formed with a prefixed *u-* plus a suffix', he also observes that the prefix 'is identical with the versional marker of a third person in finite verbal forms and refers to the object of the comparison'. So the prefix in all likelihood was not a marker of the comparative degree proper, but rather an agreement marker. As far as Svan is concerned, Bobaljik (2012: 108n) already expresses some doubts as to the relevance of these forms. Gippert (1996: 37) likewise states that Svan 'show[s] the same tendency as Georgian does, in that these formations are restricted to superlative/elative functions today while real comparatives are built analytically'.

The one PRE marked suppletive comparative in (40) concerns the case of Bulgarian/Macedonian. These two languages also show an unexpected ABA pattern in the adjective for *much/many*:

18

(41)

	POS	CMPR	SPRL	
Bg.	mnogo	po-veče	naj-mnogo	'much/many'
Mac.	mnogu	po-veḱe	naj-mnogu	

We shall put them aside for now, however, and take the following to be a valid generalisation:

(42)   *Generalisation on PRE-marking and Suppletion (GOPS)*
       When there is root suppletion, the marker of the comparative degree cannot occur to the left of the adjectival root.

Now before we can see what the relevance is of (42) for the question we raised at the end of the previous section, we need to discuss the analysis of suppletion. Caha (2017), De Clercq & Vanden Wyngaerd (2017) argue that suppletive roots like English *bett* are characterised by the fact that they spell out C1P, as shown in (43):

(43)
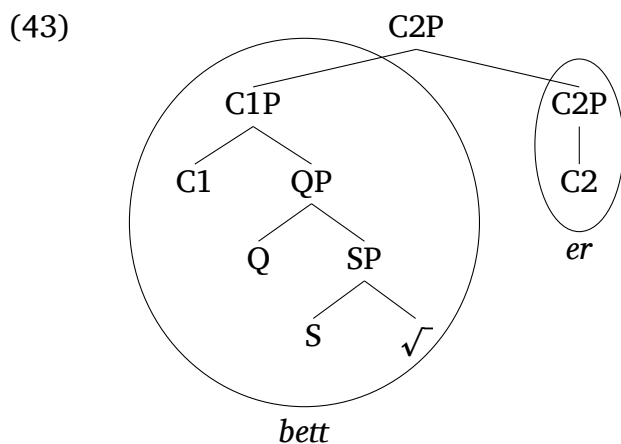
Now the generalisation in (42) can be shown to follow from this analysis, provided no feature overlap is permitted between the features of the main spine and those of a specifier. To see how this works, let us reconsider the overlapping representation of the analytical comparative in English given in (33) above, but replace the adjectival root with a suppletive one, i.e. a C1P:

(44)

C2P

C2P

C2   C1P

C1   QP

Q   S

PRE marker

C1P

C1   QP

Q   SP

S $\sqrt{}$
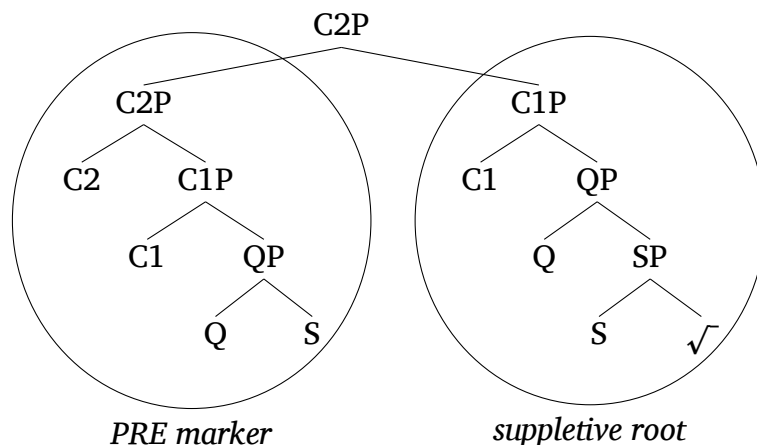
suppletive root

If overlap is allowed in principle, this type of derivation, which combines a PRE marker with a suppletive root, cannot be blocked. Even though it would not arise in English (because *better* will block the derivation of *more bett*), it could be derived in a language that lacked a lexical item like English *er* (which spells out just [$_{C2P}$ C2]). In search for spellout, a new derivation would be spawned, and as long as the language had a lexical item that provided a spellout for the missing feature C2, it could occur as a PRE marker with the suppletive root, as shown in (44). It would be highly coincidental if the theory allowed a derivation like (44), but no language with suppletion in the comparative had a PRE marker as in (44).

In contrast, if overlap is ruled out in principle, we can derive the GOPS in (42). In such a case, Unmerge (or Remove, in the terminology of Müller 2017) would need to remove anything merged higher than $\sqrt{}$ from the main spine, so as to create a representation without overlap:

20

(45)

C2P
- C2P *(PRE marker)*
  - C2
  - C1P
    - C1
    - QP
      - Q
      - S
- √ *(nonsuppletive root)*
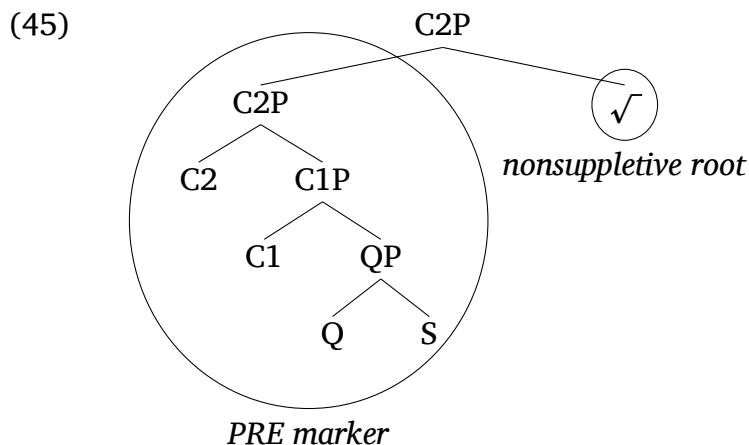
But as soon as the main derivation (on the right hand side in (45)) shrinks below C1P, it will revert to the nonsuppletive root.[10] In this way, the GOPS is derived: a PRE marker will always minimally spell out C1 and C2 (since these mark the comparative), and if overlap is not allowed, then the main spine cannot contain C1 anymore, which is the feature that characterises suppletive roots. As a result, suppletive roots are incompatible with PRE marking.

6   Conclusion

In this paper, we have considered the internal structure of the English marker of the analytic comparative, *more*. We have argued that its internal structure spells out all the heads of a gradable adjective, except for the root feature. When *more* is merged into the main spine, it triggers an Unmerge operation that removes previously created structure. In this way, we were able explain the Generalisation on PRE marking and Suppletion (GOPS) to the effect that PRE marking of the comparative markers is incompatible with suppletion. The explanation relies on the distinction proposed in Starke (to appear) between PRE and POST marking in terms

---

[10]Note that the PRE marker must not spell out all the features of a gradable adjective, because if it did and overlap is not allowed, there would be nothing left in the main spine to spell out. The ban against overlapping derivations could be taken to result from assuming a numeration that must not contain any duplicate features. This will ensure that derivations with overlap cannot arise (although Unmerge is still required if a given type of derivation reaches a dead end).

of the binary or unary nature of the feature(s) at the bottom of the tree of the marker, and on a restriction to the effect that there must not be feature overlap between PRE markers and the derivation they are merged with. In order to achieve a converging derivation without featural overlap, an Unmerge operation may remove previously generated structure in the main derivation.

References

Adger, David. 2013. *A syntax of substance*. Cambridge, MA: MIT Press.

Bobaljik, Jonathan. 2012. *Universals in comparative morphology*. Cambridge, MA: MIT Press.

Bresnan, Joan. 1973. Syntax of the comparative clause construction in English. *Linguistic Inquiry* 4. 275–343.

Caha, Pavel. 2017. Explaining Bobaljik's root suppletion generalization as an instance of the adjacency condition (and beyond). In Joseph Emonds & Markéta Janebová (eds.), *Language use and linguistic structure: Proceedings of the Olomouc linguistics colloquium 2016*. 193–208. Olomouc: Palacký university.

Caha, Pavel, Karen De Clercq & Guido Vanden Wyngaerd. 2017. The fine structure of the comparative. Ms. Masarykova Univerzita, U Ghent, KU Leuven.

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

Corver, Norbert. 1997. Much-support as a last resort. *Linguistic Inquiry* 28. 119–164.

De Clercq, Karen. 2017. The internal syntax of Q-words. *Linguistics in the Netherlands* 34. 31–46.

De Clercq, Karen & Guido Vanden Wyngaerd. 2017. *ABA revisited: evidence from Czech and Latin degree morphology. *Glossa* 2(1). 69: 1–32.

De Clercq, Karen & Guido Vanden Wyngaerd. to appear. Negation and the functional sequence. *Natural Language and Linguistic Theory* `http://ling.auf.net/lingbuzz/003794`.

Embick, David. 2007. Blocking effects and analytic/synthetic alternations. *Natural Language and Linguistic Theory* 25(1). 1–37.

Gippert, Jost. 1996. The formation of comparatives in the history of Georgian. Part I: The prehistory of the synthetic comparatives. *Enatmecnierebis Sak'itxebi* 1. 37–45.

Gouskova, Maria & Suzy Ahn. 2016. Sublexical phonotactics and English comparatives. Ms. NYU.

Matushansky, Ora. 2013. More or better: On the derivation of synthetic comparatives and superlatives in English. In Ora Matushansky & Alec Marantz (eds.), *Distributed morphology today: Morphemes for Morris Halle*. 59–78. Cambridge, MA: MIT Press.

Müller, Gereon. 2017. Structure removal: An argument for feature-driven Merge. *Glossa* 2(1). 28. 1–35.

Starke, Michal. 2004. On the inexistence of specifiers and the nature of heads. In Adriana Belletti (ed.), *Structures and beyond,* vol. 3. 252–268. Oxford: Oxford University Press.

Starke, Michal. 2014. Cleaning up the lexicon. *Linguistic Analysis* 39. 245–256.

Starke, Michal. to appear. Complex left branches, spellout, and prefixes. In Lena Baunaz, Karen De Clercq, Liliane Haegeman & Eric Lander (eds.), *Exploring nanosyntax.* Oxford: Oxford University Press.