Unmerging Analytical Comparatives

Karen De Clercq[1] Guido Vanden Wyngaerd[2]

[2]KU Leuven (Brussels)

[1]FWO/U Ghent (Ghent)

Workshop on Shrinking Trees in Morphology
Universität Leipzig, Institut für Linguistik, 27 April 2018

Introduction

(1)

	POS	CMPR
SYNTHETIC	smart	smart-er
ANALYTIC	intelligent	more intelligent

Root Suppletion Generalisation (RSG) (Bobaljik 2012)

Root suppletion is limited to synthetic (i.e., morphological) comparatives.

(2)

	Greek	POS	CMPR	
SYNTHETIC	kak-ós	cheiró-ter-os	‘bad’
ANALYTIC	kak-ós	pjo kak-ós	
ANALYTIC	kak-ós	*pjo cheir-ós	

(3)

	POS	CMPR
SYNTHETIC	good	bett-er
ANALYTIC	intelligent	more intelligent
ANALYTIC	good	*more bett

Generalisation on Suppletion and PRE-marking (GOSP)

When there is root suppletion, the marker of the comparative degree cannot occur to the left of the adjectival root.

Aims of this talk:

- refine Bobaljik's proposal on the internal complexity of CMPR by splitting up CMPR into C1 and C2
- present an account of the distribution of analytic vs synthetic comparative marking in terms of this more fine-grained structure
- show that GOSP is valid
- explain GOSP as a consequence of
  - a principled distinction in the way PRE markers differ from POST markers
  - a restriction against feature overlap
- where feature overlap arises, an Unmerge operation removes previously merged heads

The comparative: splitting up CMPR

(4)

POS	CMPR	SPRL	
bujar-ý	bujař-ejš-í	nej-bujař-ejš-í	‘merry’
červen-ý	červen-ějš-í	nej-červen-ějš-í	‘red’
hloup-ý	hloup-ějš-í	nej-hloup-ějš-í	‘stupid’
moudr-ý	moudř-ejš-í	nej-moudř-ejš-í	‘wise’

í/ý = adjectival agreement: Case, number, gender

Comparative *ějš* = *ěj+š*

2 pieces of evidence showing that -*ějš*- consists of two parts:

1. -*ěj*- disappears with certain adjectives
2. -*š*- disappears with comparative adverbs

1. *-ěj-* disappears with certain adjectives (the *star* ‘old’ class)

(5)

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’

2. *-š-* disappears with comparative adverbs

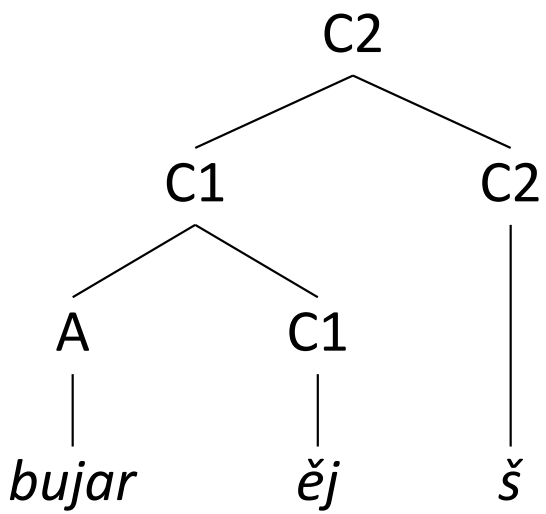
(6)

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’

- The Czech comparative suffix consists of two parts: *ěj+š*
- These two parts correspond with two syntactic heads: C1 and C2
- These two heads supersede Bobaljik's CMPR

- The Czech comparative suffix consists of two parts: *ěj+š*
- These two parts correspond with two syntactic heads: C1 and C2
- These two heads supersede Bobaljik's CMPR

(7)



Decomposing A

- the head A has internal structure, and is composed of
  - a root feature $\sqrt{}$ (some prefer …)
  - a gradability feature Q

Modelling the lexical difference between two adjective classes

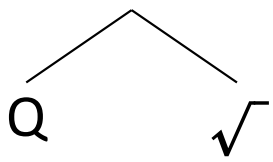
- different types of adjectival roots realise constituents of different sizes (=phrasal spellout)

(8)

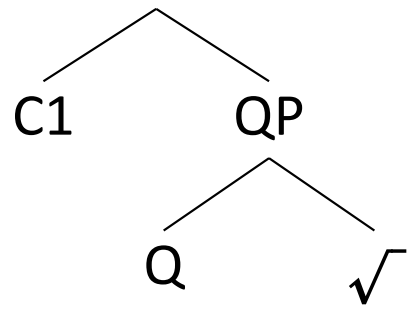
√	Q	C1	C2
bujar		ěj	š
star			š

Lexical entries for the two adjective classes

(9) QP ⟺ /bujar/



(10) C1P ⟺ /star/



Lexical entries for the comparative markers

(11) C1P ⟺ /ěj/

C1

(12) C2P ⟺ /š/

C2

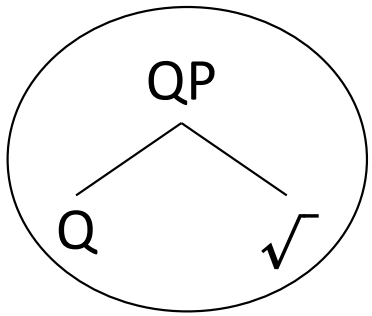
The *bujar* ‘merry’ class

- adjectival roots of the *bujar* type realise a phrasal node QP

The *bujar* ‘merry’ class

- adjectival roots of the *bujar* type realise a phrasal node QP

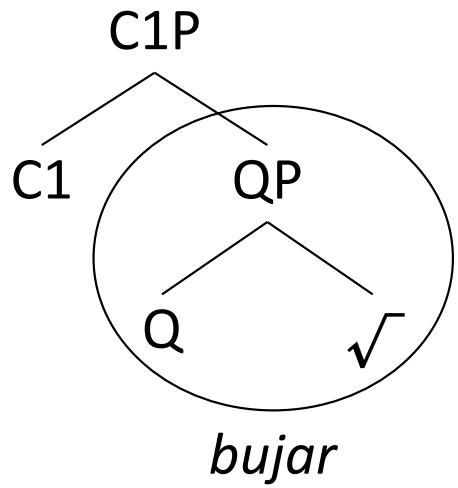
(13)



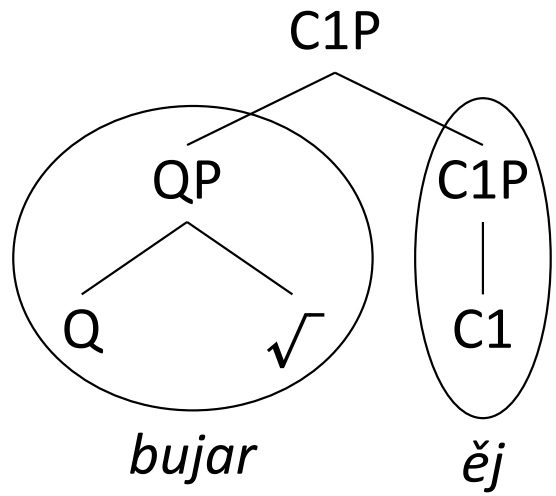
bujar

Comparative marking in the *bujar* ‘merry’ class

(14)



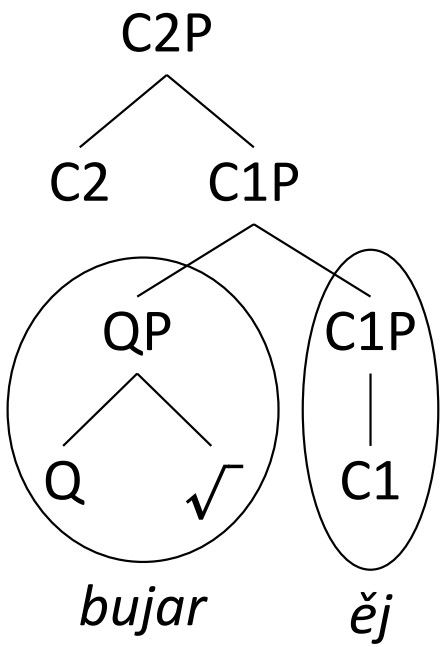
(15)



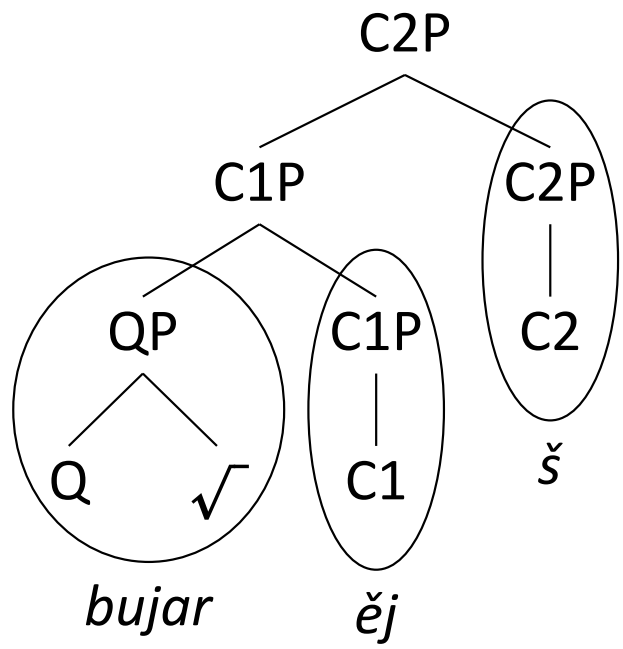
- (movement preferred over direct spellout of C1P by Faithfulness)

Comparative marking in the *bujar* ‘merry’ class

(16)



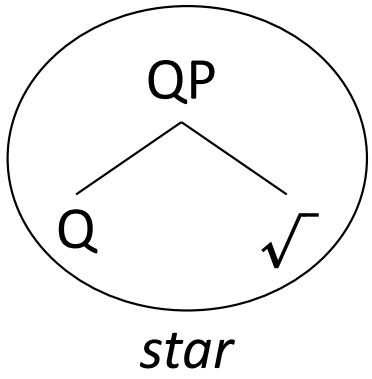
(17)



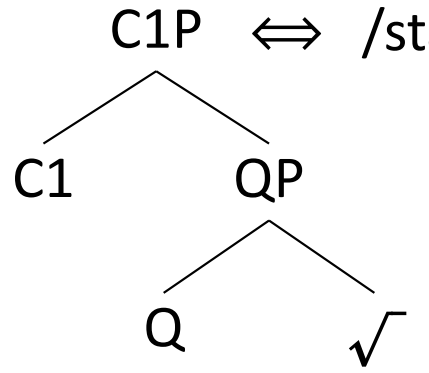
The *star* 'old' class

- adjectival roots of the *star* type realise a phrasal node QP or C1P

(18)



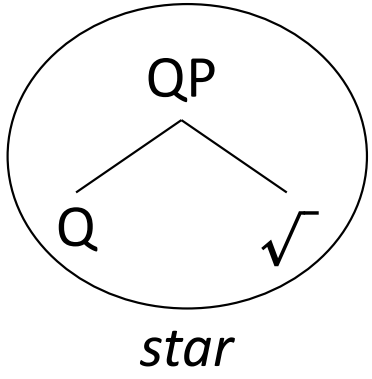
(10)



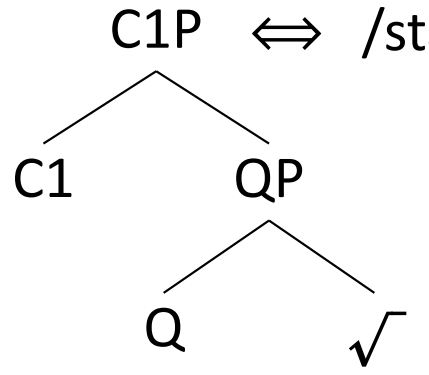
The *star* 'old' class

- adjectival roots of the *star* type realise a phrasal node QP or C1P

(18)



(10)

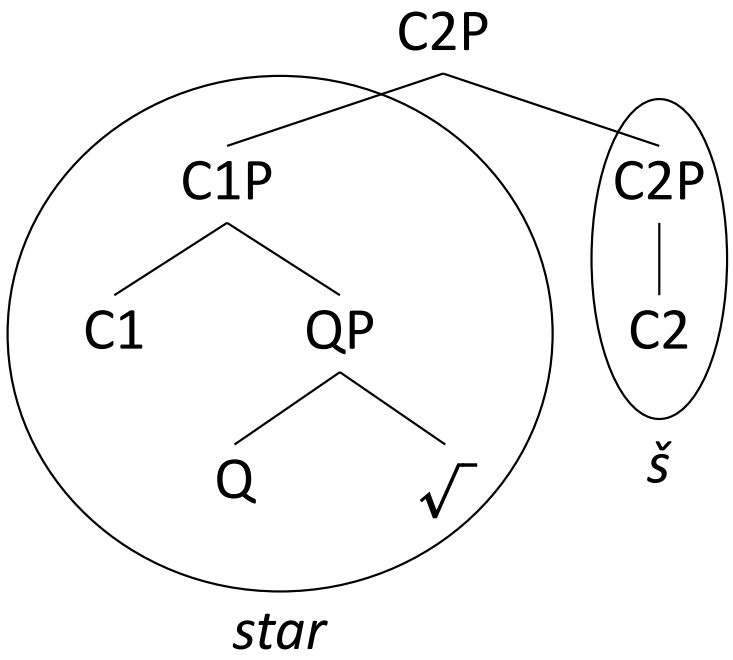


(19) *The Superset Principle*
A lexically stored tree λ can spell out a syntactic constituent σ iff λ contains σ as a subtree.

The *star* 'old' class

- in the comparative, adjectival roots of the *star* class spell out the node C1P
- this explains the absence of *ěj*

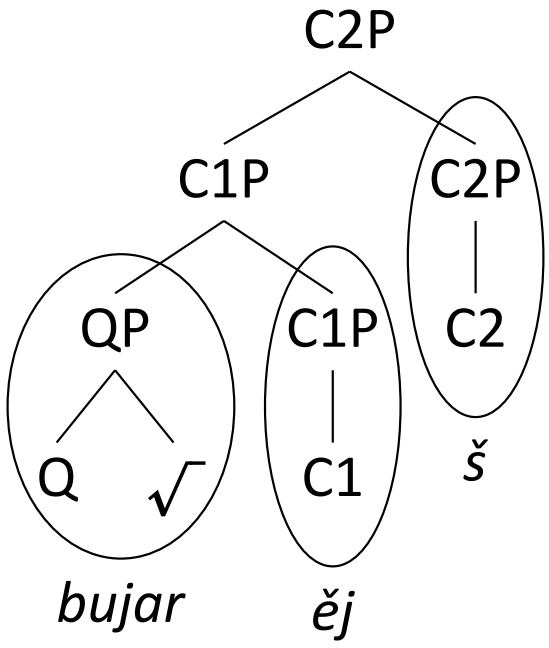
(20)



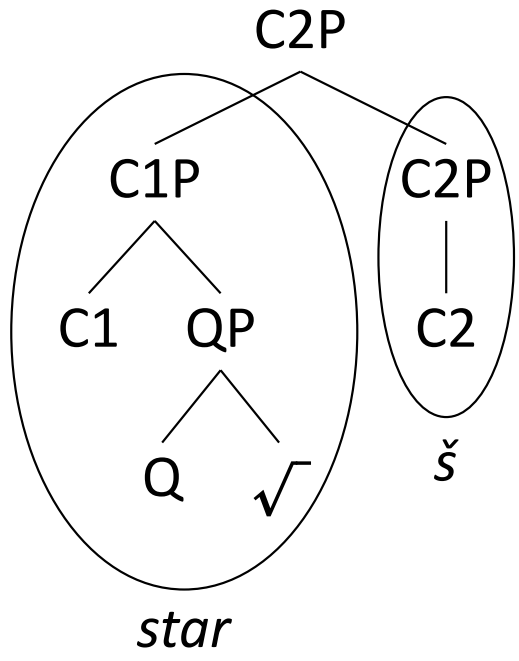
The root-affix tradeoff

- as the root grows, less suffixes get spelled out

(21)



(22)



The root-affix tradeoff

(23)

√	Q	C1	C2	
kluz		k		‘slippery’
bujar		ěj	š	‘merry’
star			š	‘old’
ostř				‘sharp’ (NE Bohemian)

(Caha et al. 2017)

English analytic-synthetic comparatives (Caha 2017*b*)

- analytic-synthetic distinction in the comparative is lexical
- it relates to the size of the root
- the comparative markers likewise also vary in size

(24)

√	Q	C1	C2
intelligent		mo	re
smart			er

English analytic-synthetic comparatives (Caha 2017*b*)

- analytic-synthetic distinction in the comparative is lexical
- it relates to the size of the root
- the comparative markers likewise also vary in size

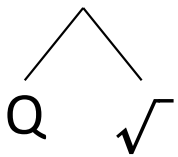
(24)

√	Q	C1	C2
intelligent		mo	re
smart			er

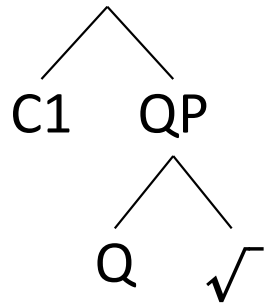
- the distribution is not determined by phonology, but by frequency (Graziano-King 1999; Bobaljik 2012)
  - *more lax, gaunt, ill, apt*
  - *older, longer, smarter, thinner*

Lexical entries for the two adjective classes

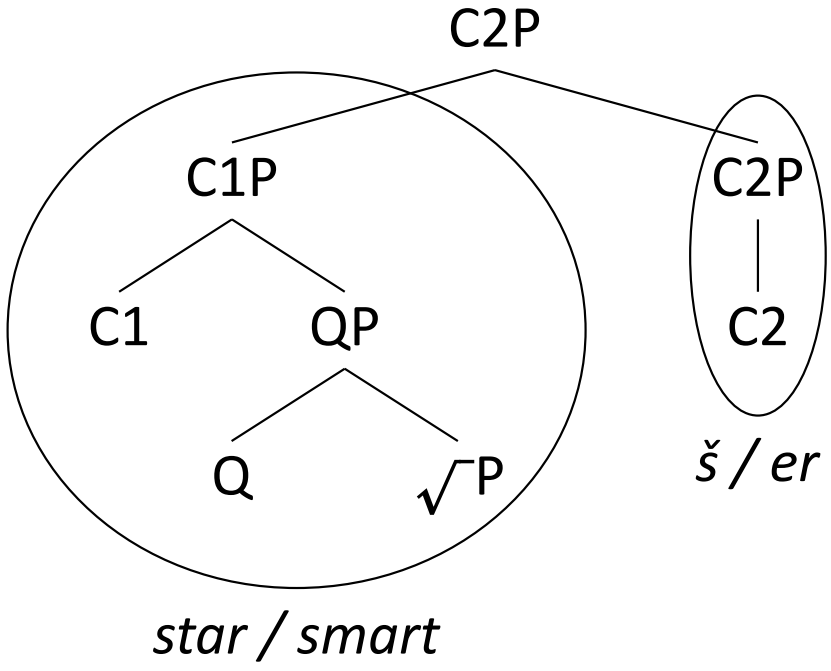
(25) QP $\Leftrightarrow$ /intelligent/



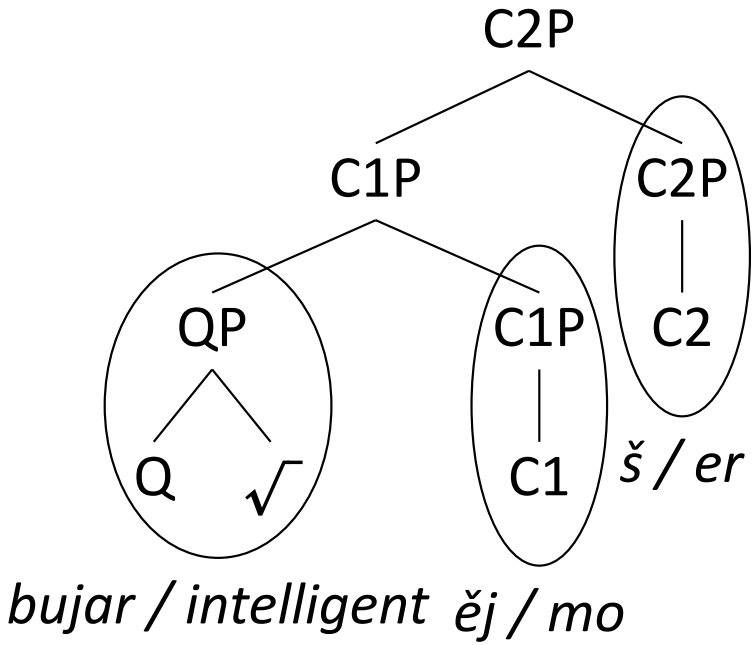
(26) C1P $\Leftrightarrow$ /smart/



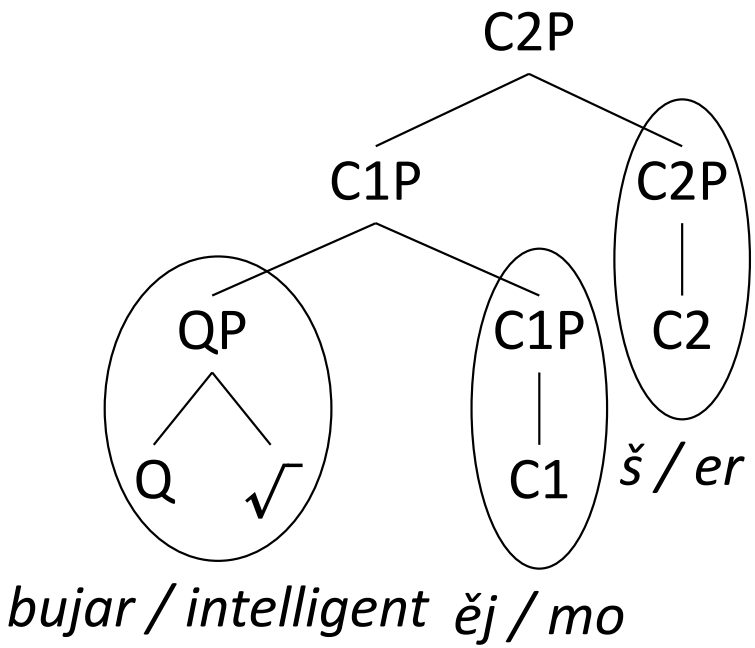
(27)



(28)



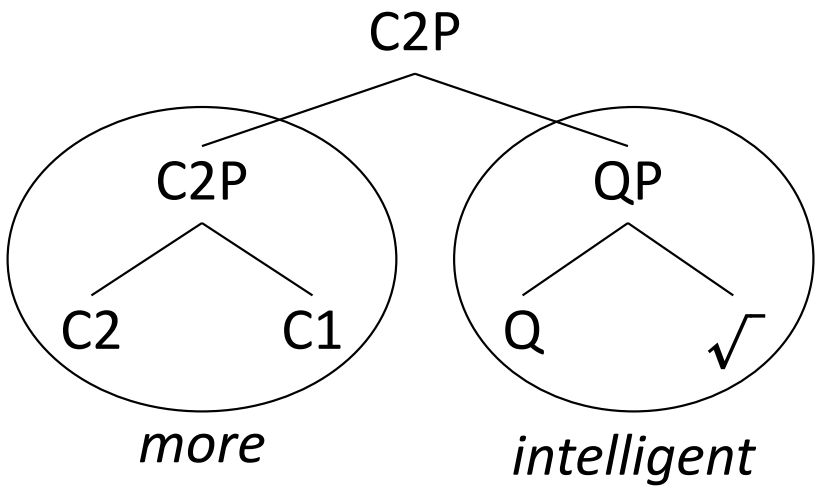
(28)



- Although we assume that *more* does spell out both C1 and C2, the above structure is plainly incorrect
- Czech comparative marking is entirely suffixal, but the English analytic marker of the comparative precedes the root

- A better structure is (29) (to be updated later):

(29)



Side note: the lexicon

The traditional lexicon

The lexicon is ‘a messy and an ugly place, full of disorder, exceptions, and cacaphony’.

The nanosyntactic Lexicon (Starke 2014)

The lexicon contains nothing but well-formed syntactic expressions.

The traditional lexicon

- The analytic-synthetic distinction: a lexical diacritic$^{[+M]}$ on certain adjectives (which triggers Merger) (Bobaljik 2012: 164)

(30) a. $[_A$ smart$^{[+M]}$ $]$
b. $[_A$ intelligent$^{[-M]}$ $]$

- The prefix-suffix distinction: lexically specified (Embick and Noyer 2007)
- The suppletive/nonsuppletive distinction: context-sensitive insertion rules
- etc.

The nanosyntactic lexicon

- The analytic-synthetic distinction: a difference in the size of lexical trees (works for both Czech and English)
- The suppletive/nonsuppletive distinction: a difference in the size of lexical trees
- The prefix-suffix distinction: a difference in the internal makeup of lexical trees

PRE vs POST

POST marking:

- suffixal
- to the right of the stem
- displays mirror principle ordering
- results from movement of the stem to the left of the POST marker

PRE marking:

- prefixal
- functional material to the left of the stem
- ordering reflects the underlying order of the functional sequence
- involves no movement, but a separately merged complex specifier

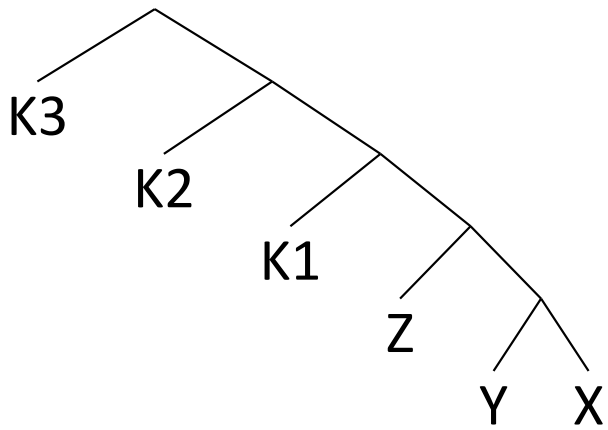
PRE vs POST

Starke (2018): two modes of combination:

- Merge-F
- Merge-XP

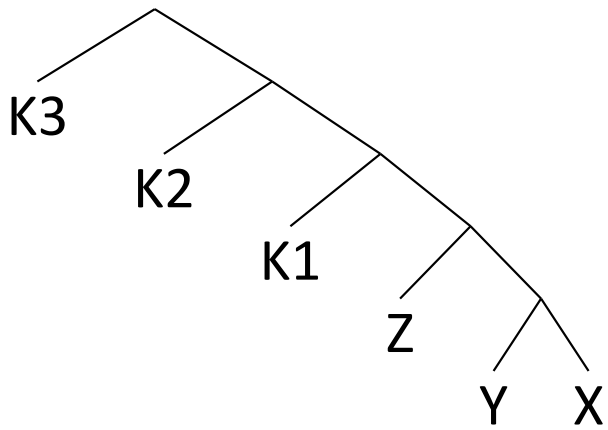
Merge-F

(31)



Merge-F

(31)



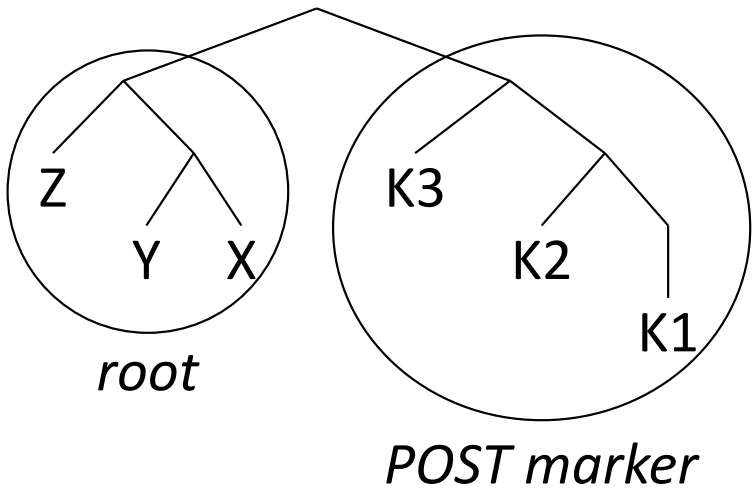
Spellout Algorithm

Merge-F and

a. Spell out FP

b. If (a) fails, move the (spec of the) complement of F, and retry (a)

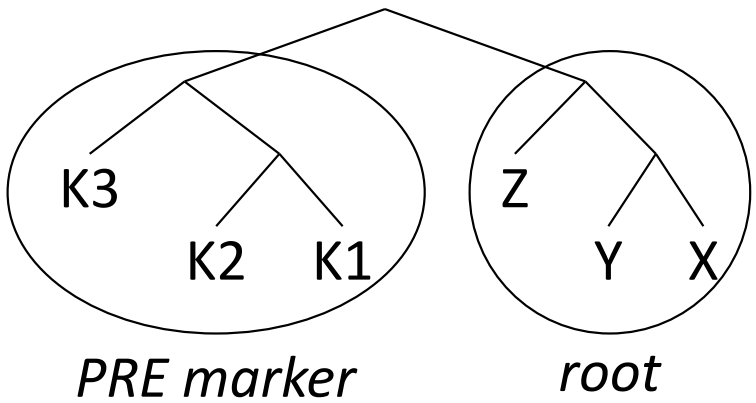
(32)



Merge-XP

- Merge-XP merges an XP as a complex specifier
- XP is merged in a separate workspace (by Merge-F), and subsequently gets merged into the main derivation

(33)



PRE vs POST

Spellout Algorithm

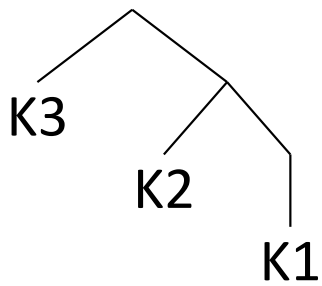
Merge-F and

a. Spell out FP
b. If (a) fails, move the (spec of the) complement of F, and retry (a)
c. If (b) also fails, spawn a new derivation providing F and merge that with the current derivation, projecting F to the top node.

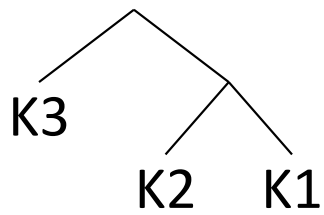
PRE vs POST

- The prefix-suffix distinction: a difference in the internal makeup of lexical trees

(34) POST: unary bottom



(35) PRE: binary bottom



PRE vs POST

- assume that Merge is always binary
- the spellout algorithm orders Merge-F before spellout
- therefore, in the absence of movement, any lexical item must minimally spell out two features

Synthetic-analytic comparatives

- C1 can either be provided by the adjectival root, or by *more*
- Analytic comparatives with *more* are triggered by the need to spell out C1, when the adjectival root does not realise C1

(36)

√	Q	C1	C2
intelligent		more	
smart			er
intelligent			er

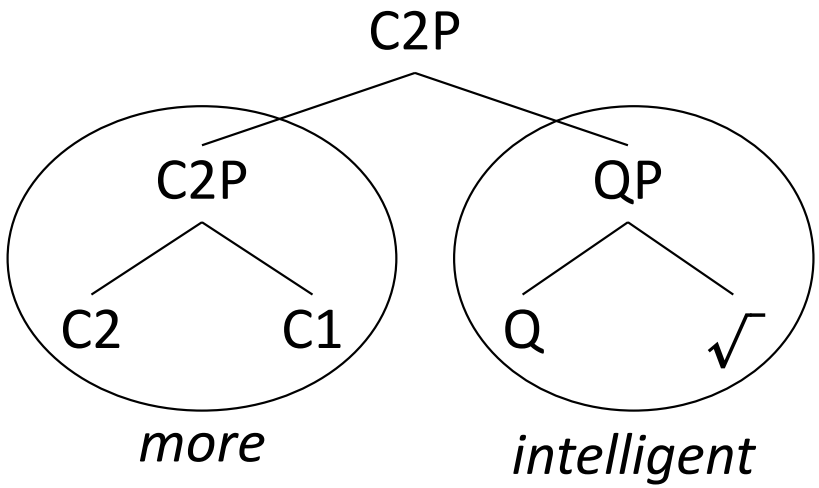
Synthetic-analytic comparatives

- C1 can either be provided by the adjectival root, or by *more*
- Analytic comparatives with *more* are triggered by the need to spell out C1, when the adjectival root does not realise C1

(36)

√	Q	C1	C2
intelligent		more	
smart			er
intelligent			er

(37)



Decomposing *more*

- *more* realises additional features beyond C1 and C2
- *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

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

Decomposing *more*

- *more* functions as a gradable adjective itself
- *more* is itself the (suppletive) comparative of *much*: *much–more–most*
- this is further confirmed by the fact that there exist analytic 'comparatives of inferiority' with *less* (e.g. *less intelligent*).
- these are absent with synthetic comparatives (*Lesslessness*; Bobaljik 2012: 4):

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

- *more* and *less* have richer internal structure than the suffixal comparative marker *-er*.

Decomposing *more*

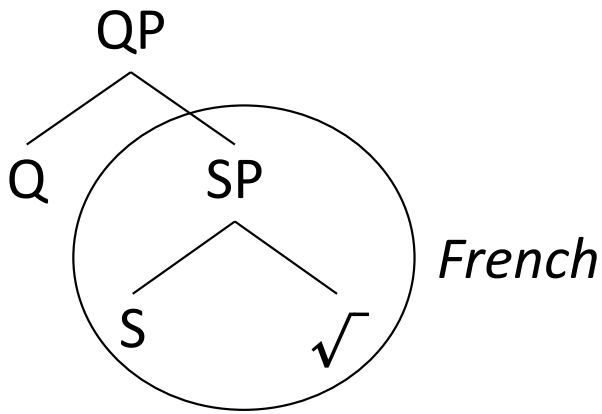
- non-gradable adjectives do not form morphological comparatives in English (40a)
- the same adjectives can form analytic comparatives with a (coerced) gradable interpretation (40b) (Matushansky 2013)

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

Decomposing *more*

- nongradable adjectives lack the gradability feature Q
- this would imply that they spell out just $\sqrt{}$
- but adjectival roots need to spell out minimally two features
- the nongradable adjectives need to spell out more than $\sqrt{}$
- there is an additional feature between Q and $\sqrt{}$: a feature STATE (or S)

(41)



Decomposing *more*

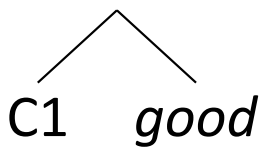
(42)

√	S	Q	C1	C2
French		more		
French				er

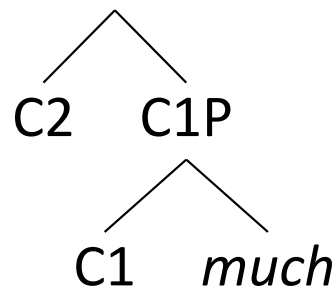
Decomposing *more*

- *more* is the suppletive comparative of *much*
- a suppletive form contains a pointer to its nonsuppletive counterpart
- the lexical entry of *more* contains a pointer to the entry for *much*

(43) C1P $\Leftrightarrow$ /bett/



(44) C2P $\Leftrightarrow$ /more/



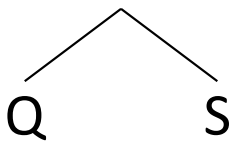
Decomposing *much*

- a handful of English adjectives allow modification by *much* (Bresnan 1973; Corver 1997)

(45) a. *much intelligent/smart/kind/ …
b. much alike/different/afraid/aware/reliant/dependent/offended

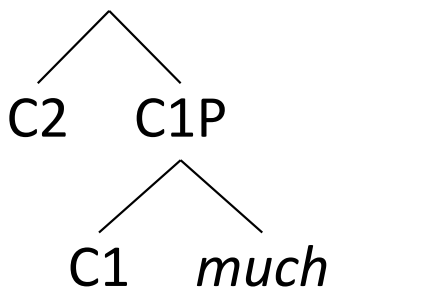
- *much* is a PRE marker, therefore it must spell out at least two features
- these features are Q and S:

(46) QP ⟺ /much/

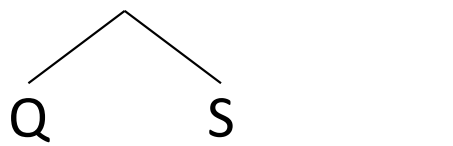


Decomposing *more*

(44)

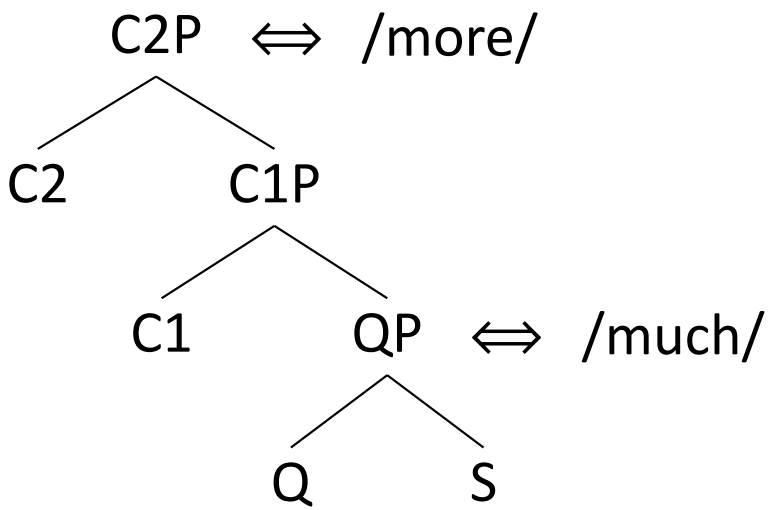


(46)



Decomposing *more*

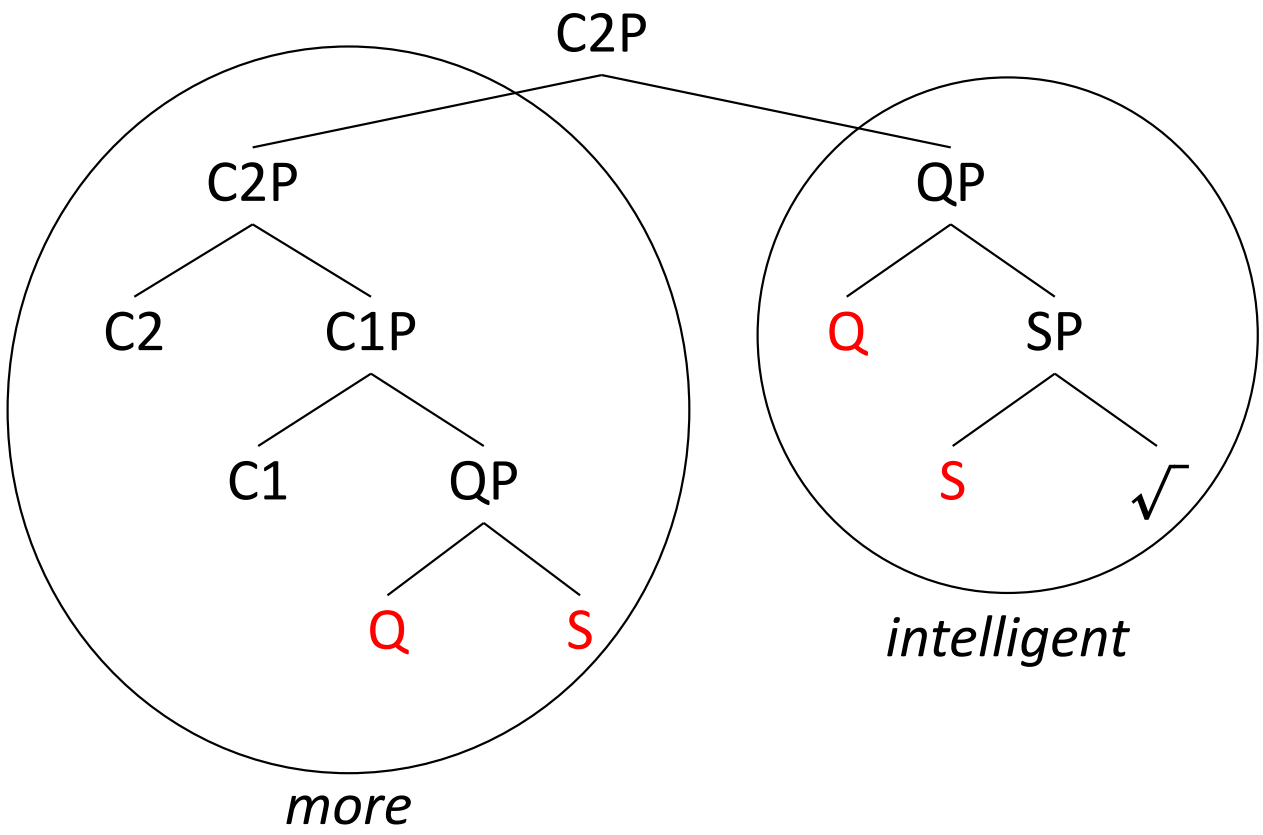
(47)



Feature overlap

- Can PRE markers recurse the functional sequence already spelled out in the main spine?

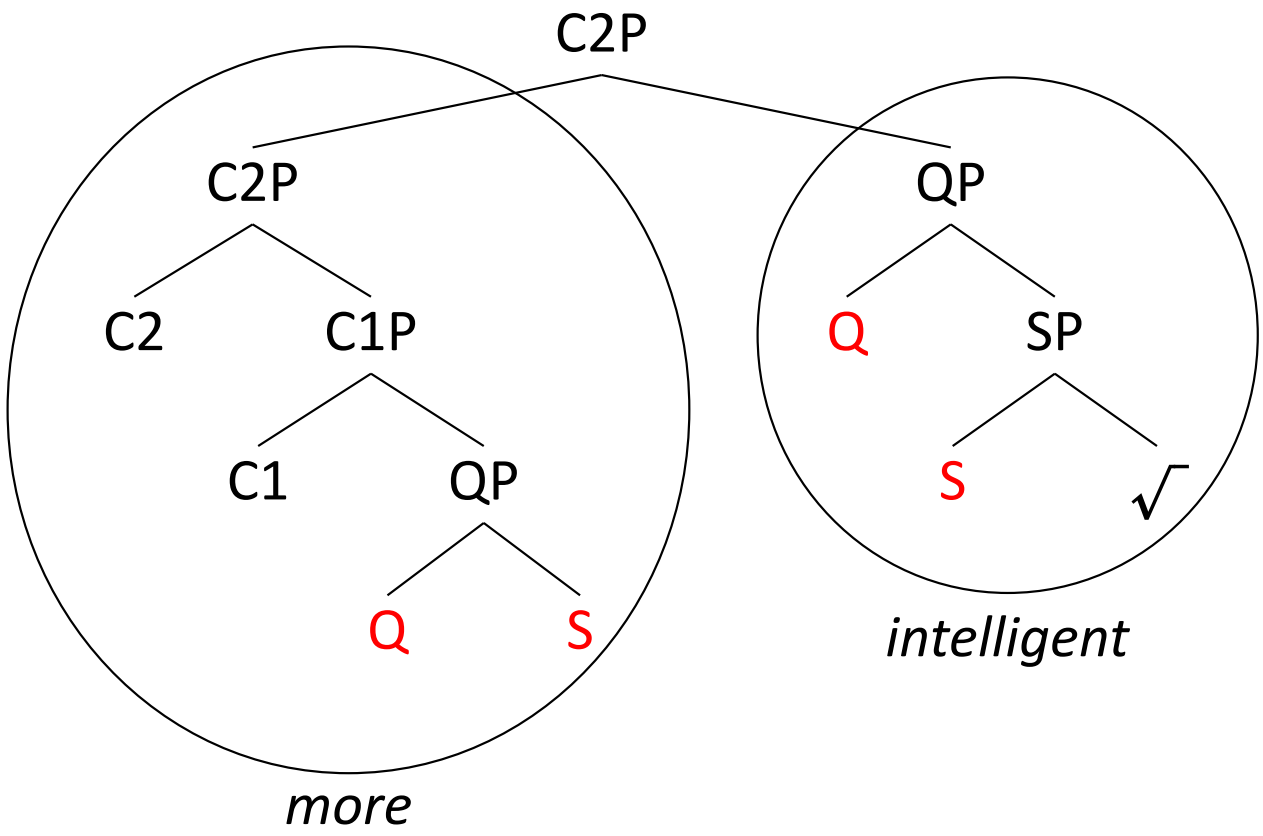
(48)



Feature overlap

- Can PRE markers recurse the functional sequence already spelled out in the main spine?

(48)

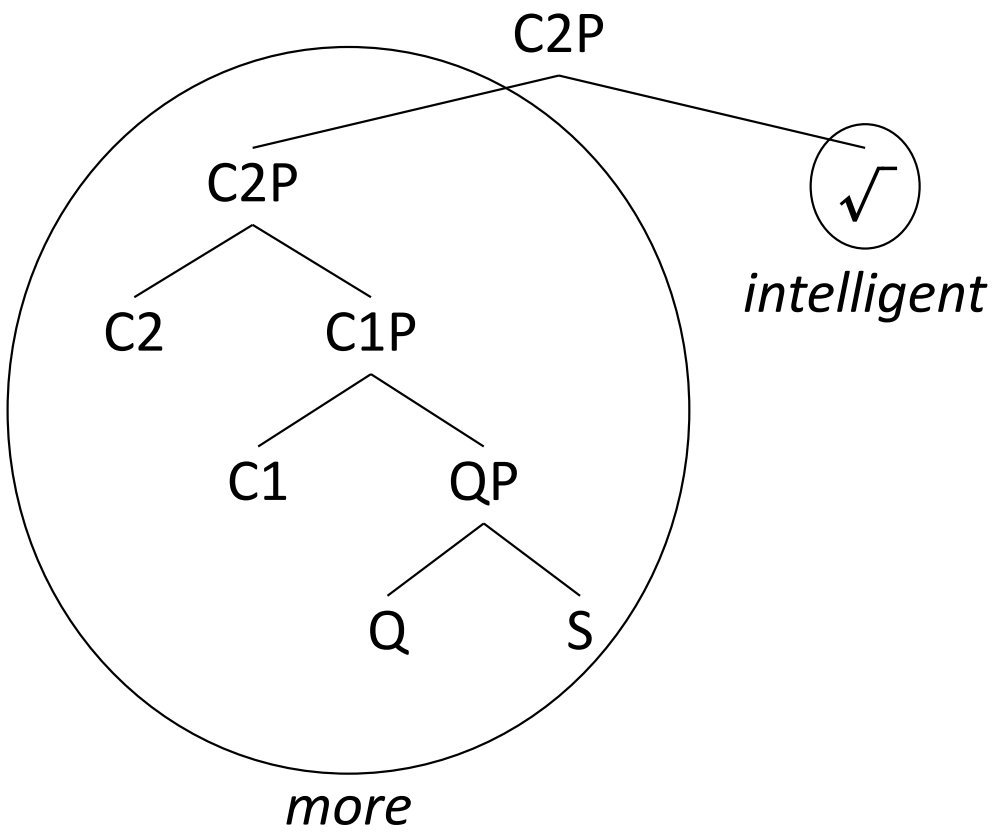


- We shall argue that such overlap is in fact disallowed

Two options for structure removal

- remove structure from the main spine

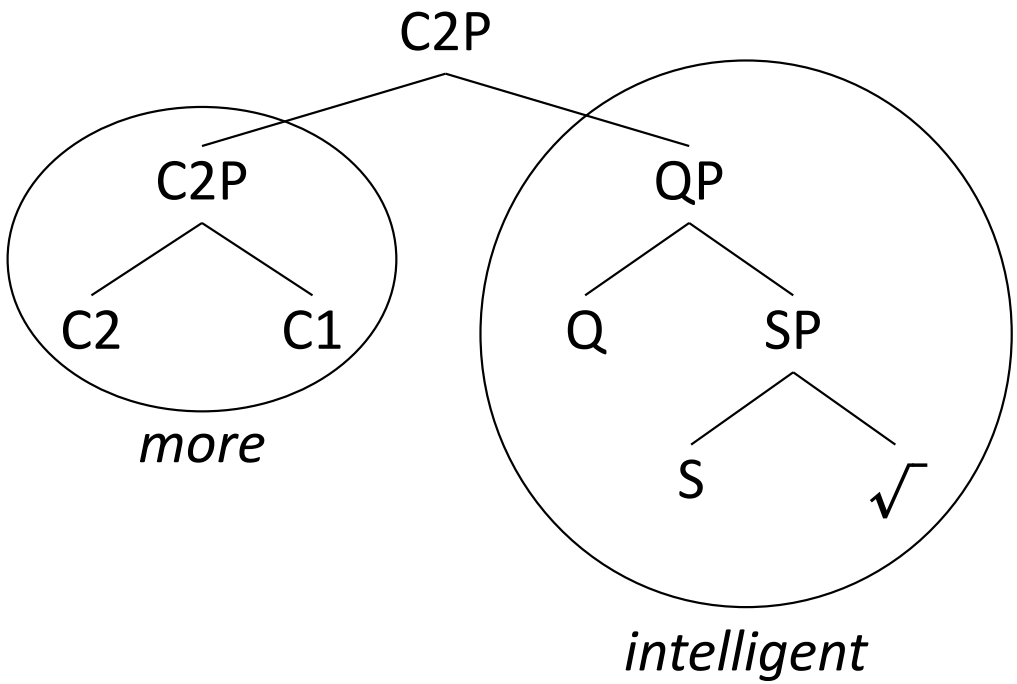
(49)



Two options for structure removal

- remove structure from the specifier:

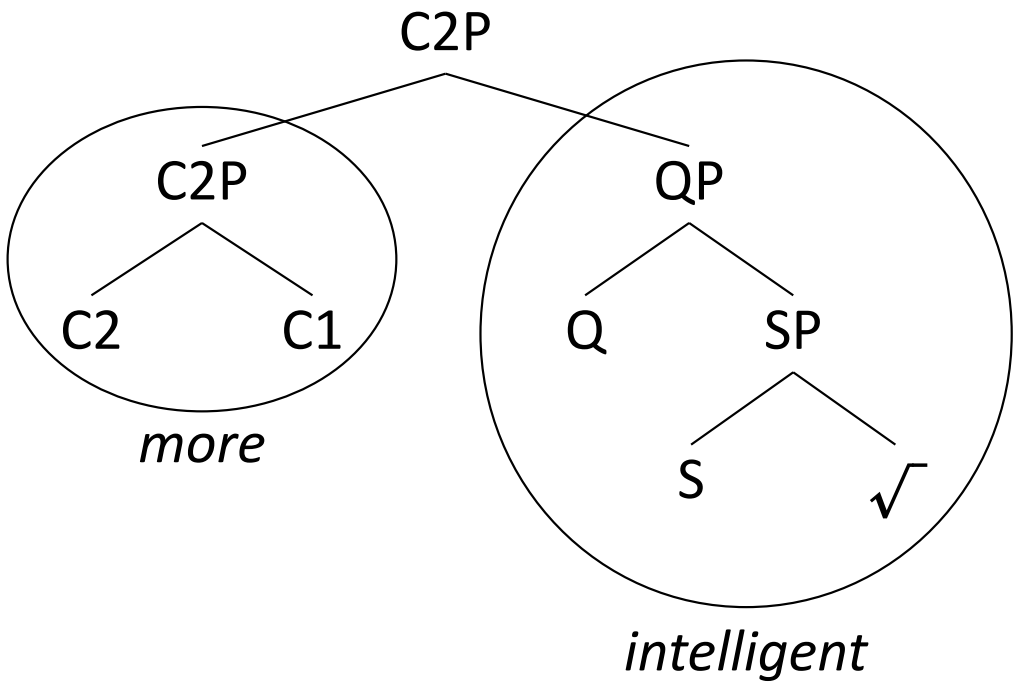
(50)



Two options for structure removal

- remove structure from the specifier:

(50)



- this second option is problematic for the Superset Principle
- structure removal would have to 'eat away' structure at the bottom of the specifier, which is arguably countercyclic

Interim summary

We have argued that

- feature overlap is banned
- structure removal takes place in the main spine

Interim summary

We have argued that

- feature overlap is banned
- structure removal takes place in the main spine

We shall now argue that

- empirical evidence from patterns of suppletion supports this

Unmerge: empirical evidence from suppletion

Generalisation on Suppletion and PRE-marking (GOSP)

When there is root suppletion, the marker of the comparative degree cannot occur to the left of the adjectival root.

Slavic

- only two languages (Bulgarian/Macedonian) have a PRE comparative marker
- no comparative suppletion in Bulgarian/Macedonian in GOOD

(51)

	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)

Bobaljik's (2012) data

(52)

MEANING	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

Bobaljik's (2012) data

(52)

MEANING	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

- out of a total of 101 suppletive triplets, only 1 is PRE-marked, while 4 are circumfixal
- the circumfixal cases are spurious

Georgian

(53)

POS	CMPR	
k'argi-i	u-mjob-es-i	'good'
	u-k'et-es-i	

- it is tempting to think of the double marking as realisations of C1 and C2 …

(54)

A	C1	C2
	es	u
	u	es
k'argi		
mjob		
k'et		

Georgian

(53)

POS	CMPR	
k'argi-i	u-mjob-es-i	'good'
	u-k'et-es-i	

- it is tempting to think of the double marking as realisations of C1 and C2 …

(54)

A	C1	C2
	es	u
	u	es
k'argi		
mjob		
k'et		

- … but probably incorrect

Gippert (1996):

- 'The Old Georgian comparatives, nowadays used with a 'superlative/elative' function only, were commonly formed with a prefixed *u-* plus a suffix that appeared either as a shorter variant, *-e* or *-o*, or as a longer, declinable one, *ēs-*
- …these formations are restricted to superlative/elative functions today while real comparatives are built analytically …
- …the prefix appearing as *u-* […] is identical with the versional marker of a third person in finite verbal forms and refers to the object of the comparison'

Old Georgian

(55)

A	C1	C2	AGR
k’argi			
mjob		es	u
k’et		es	u

Svan

(56)	POS	CMPR	
	ezär	xo-č-a	‘good’
		xo-č-ēl	

(57)	POS	CMPR	
	d͡zɣəd	xo-š-a	‘big’
		xo-š-ēl	

Bobaljik (2012: 108n):
‘Gudjedjiani and Palmaitis (1986) list four suppletive comparatives in Svan; but note also that the comparative forms in *xo-…-a* for these adjectives are used with a positive sense, and subject to further comparative formation in *xo-…-el*. It may thus be synchronically inappropriate to include these forms here.’

Gippert (1996: 37)

‘It can easily be shown that the synthetic type was inherited from Proto-Kartvelian, given that similar formations exist in the Zan languages as well as Svan; cp. Megrelian *u-magal-aš-i* ‘highest (from *magal-i* ‘high’), Laz *u-ʒgi-š-i* ‘best’, or Svan *xo-lqmaš-a* ‘strongest (from *ləqmäš* ‘strong’). Curiously enough, all sister languages show the same tendency as Georgian does, in that these formations are restricted to superlative/elative functions today while real comparatives are built analytically: Megrelian uses *umosi*, Laz, *dido*, and Svan, *gun* or *ʒġəd* as equivalents of Georgian *upro*.’

Bulgarian/Macedonian

(58)

	POS	CMPR	SPRL	
Bg.	mnogo	po-večе	naj-mnogo	‘much/many’
Mac.	mnogu	po-veḱe	naj-mnogu	

Bulgarian/Macedonian

(58)

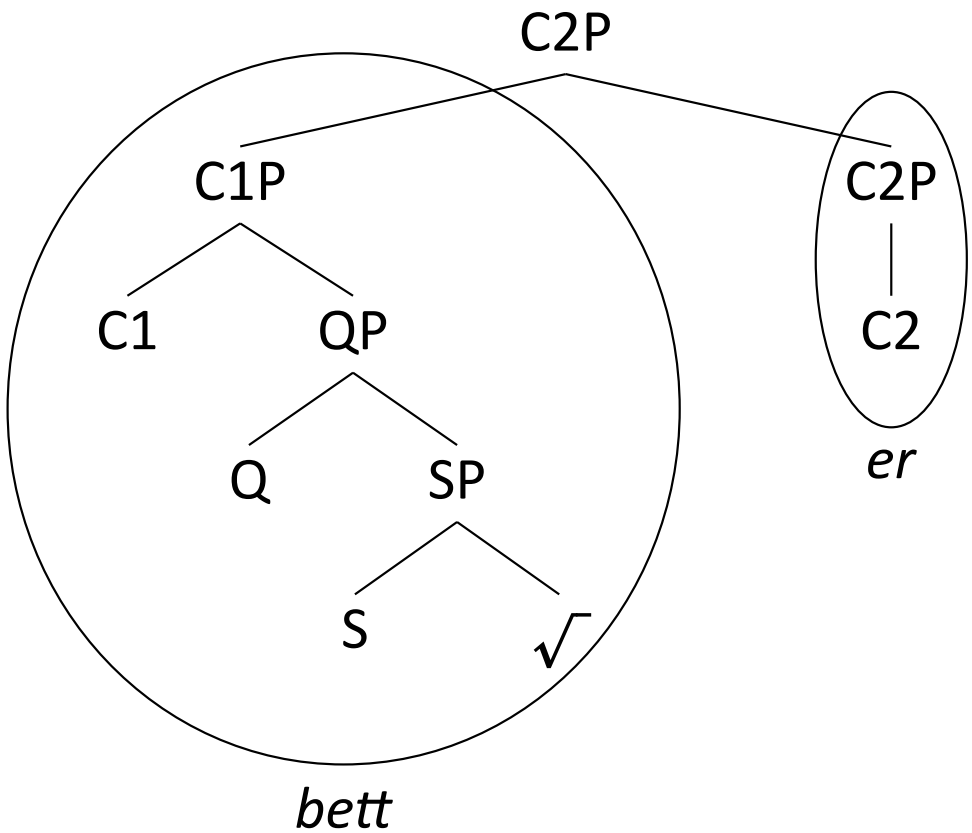
	POS	CMPR	SPRL	
Bg.	mnogo	po-večе	naj-mnogo	‘much/many’
Mac.	mnogu	po-veḱe	naj-mnogu	

- problematic for GOSP, but we set this case aside, and we take GOSP to be a valid generalisation for now

Suppletion

- suppletive roots (like English *bett*) spell out C1P (Caha 2017*a*; De Clercq and Vanden Wyngaerd 2017)

(59)



Suppletion

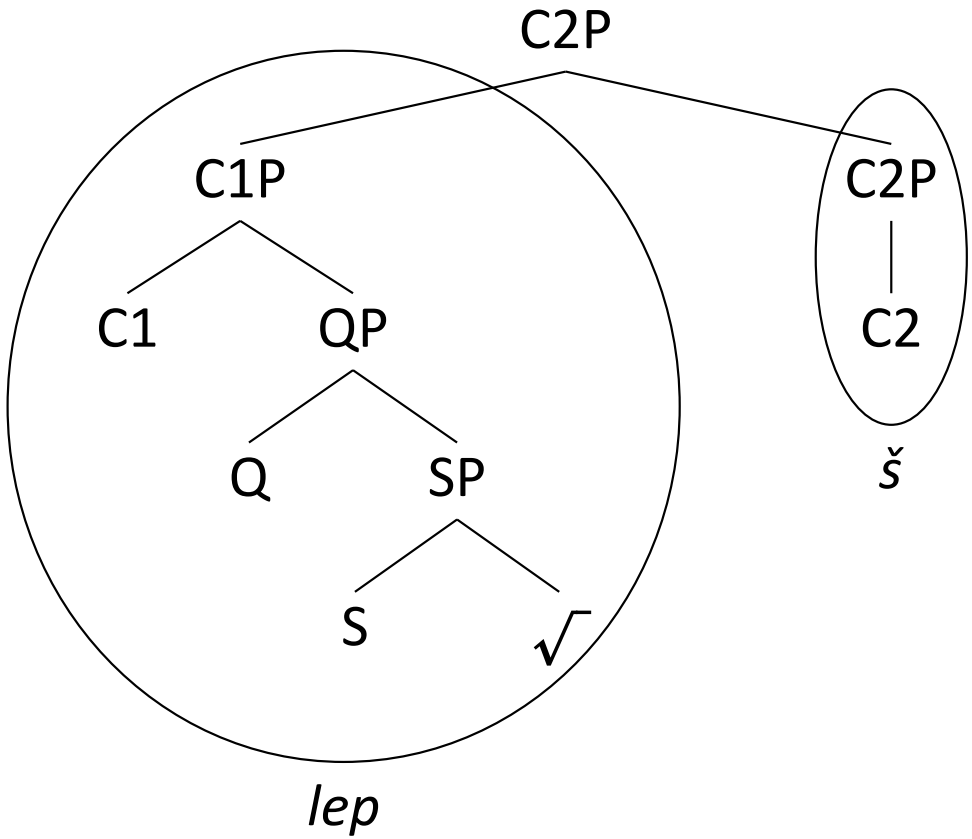
- Confirmed by the absence of *ěj* in Czech suppletive comparatives

(60)

POS	CMPR	
dobr-ý	lep-š-í	‘good’
špatn-ý	hor-š-í	‘bad’
mal-ý	men-š-í	‘little, small’
velk-ý	vět-š-í	‘big’
dlouh-ý	del-š-í	‘long’
vysok-ý	vyš-š-í	‘tall’

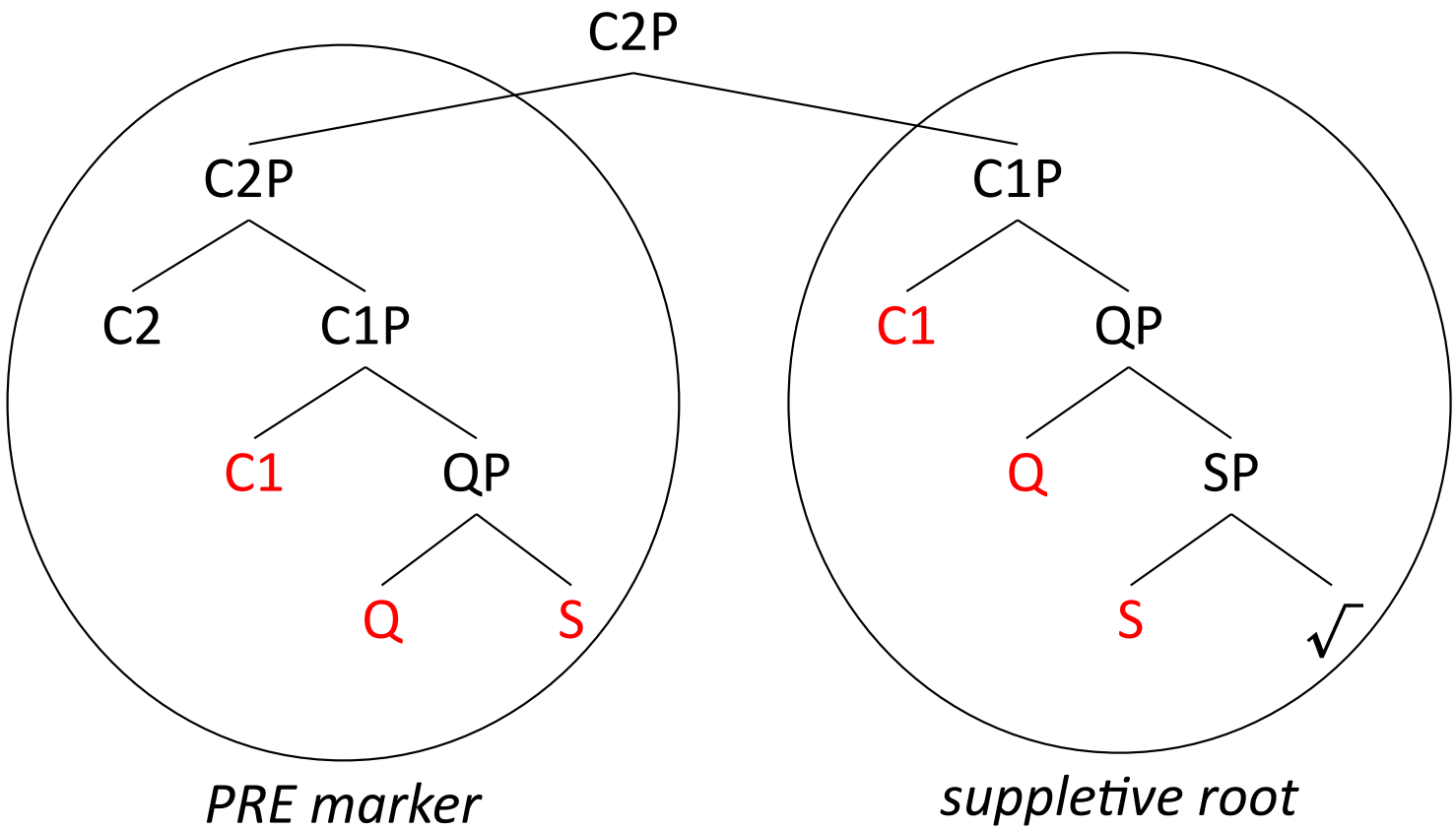
Suppletion

(61)



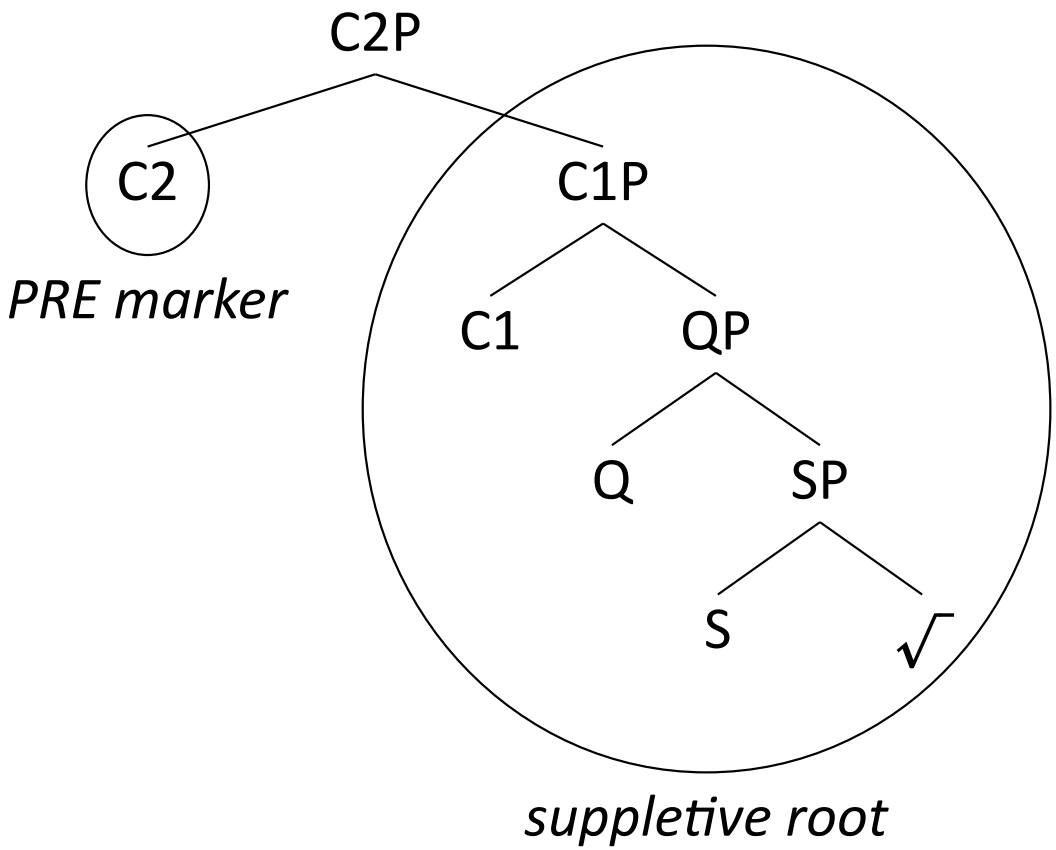
- GOSP follows from this analysis, provided feature overlap is not permitted
- consider the hypothetical situation in (62), which is not allowed:

(62)



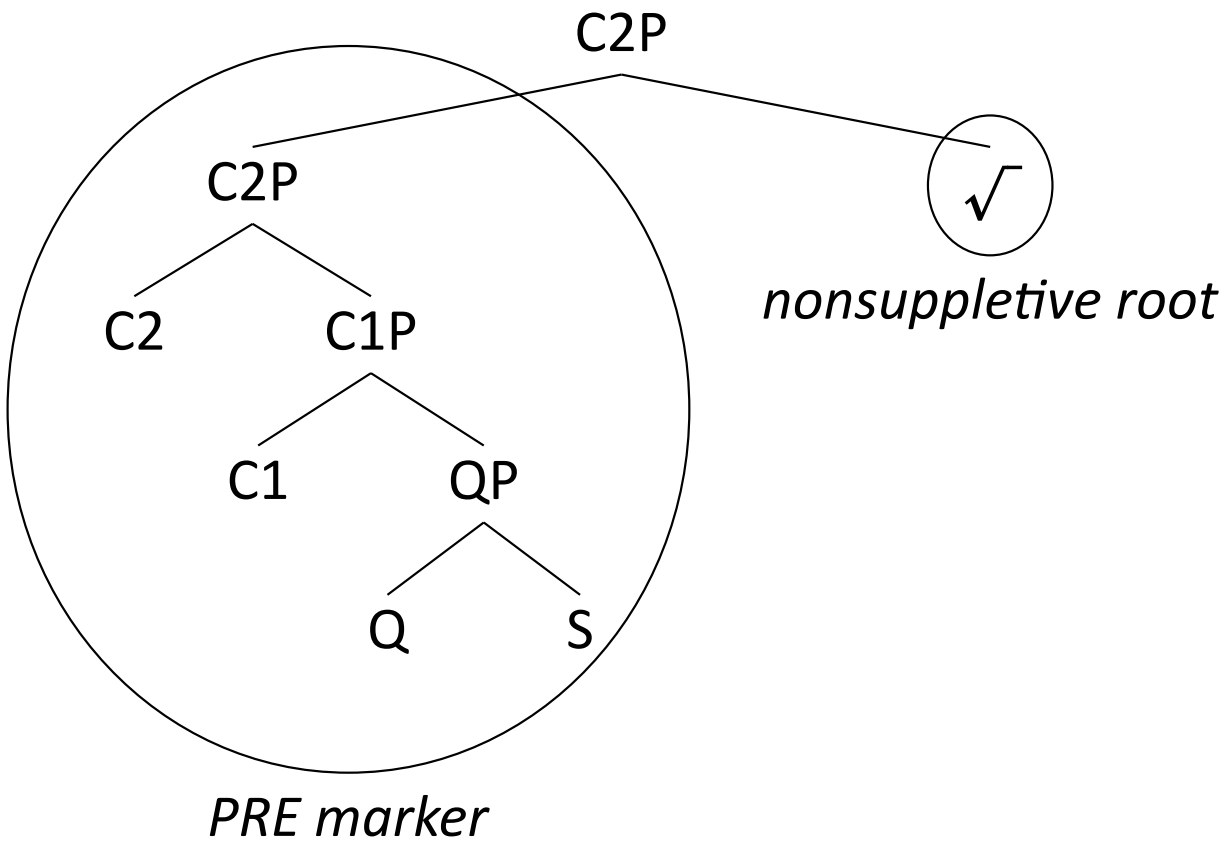
- if structure could be removed from the specifier (countercyclically), a PRE marker could combine with a suppletive root

(63)



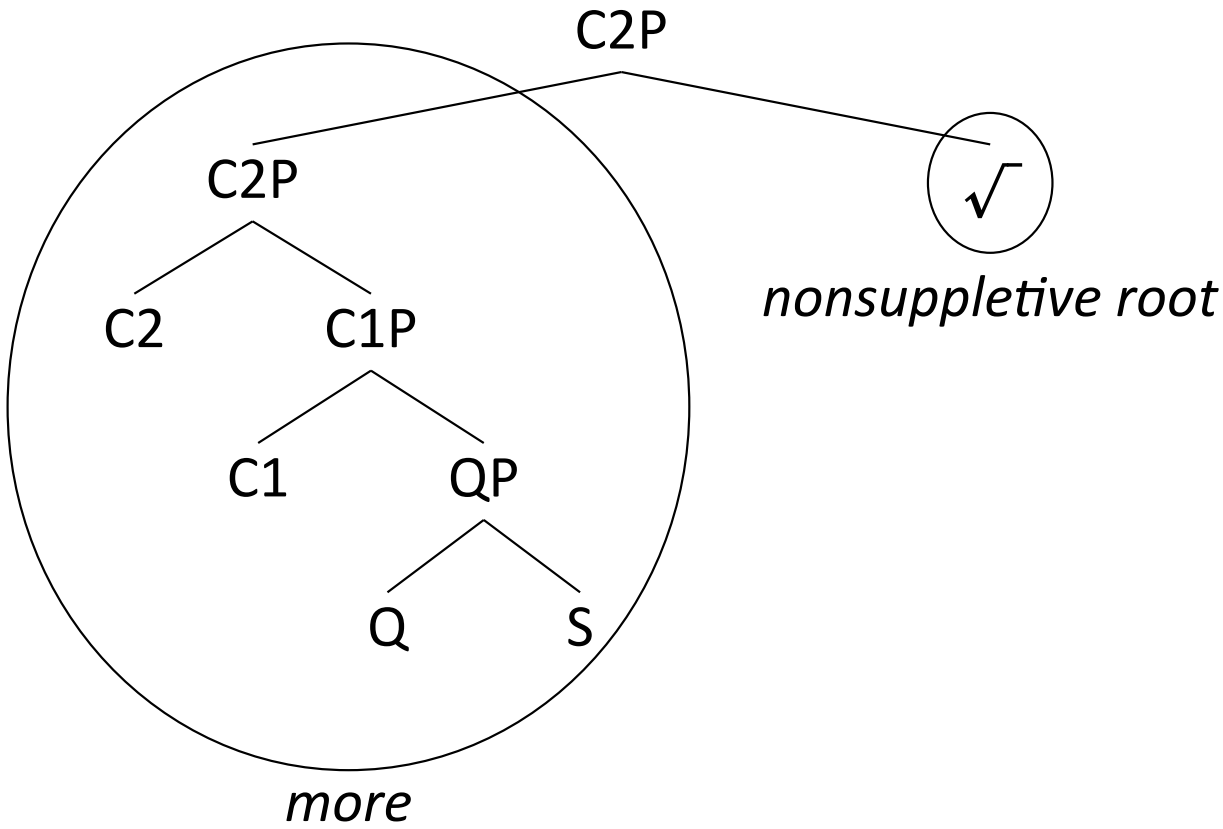
- Unmerge removes C1, Q, S from the main spine
- the remaining $\sqrt{}$ is too small to spell out a suppletive root

(64)



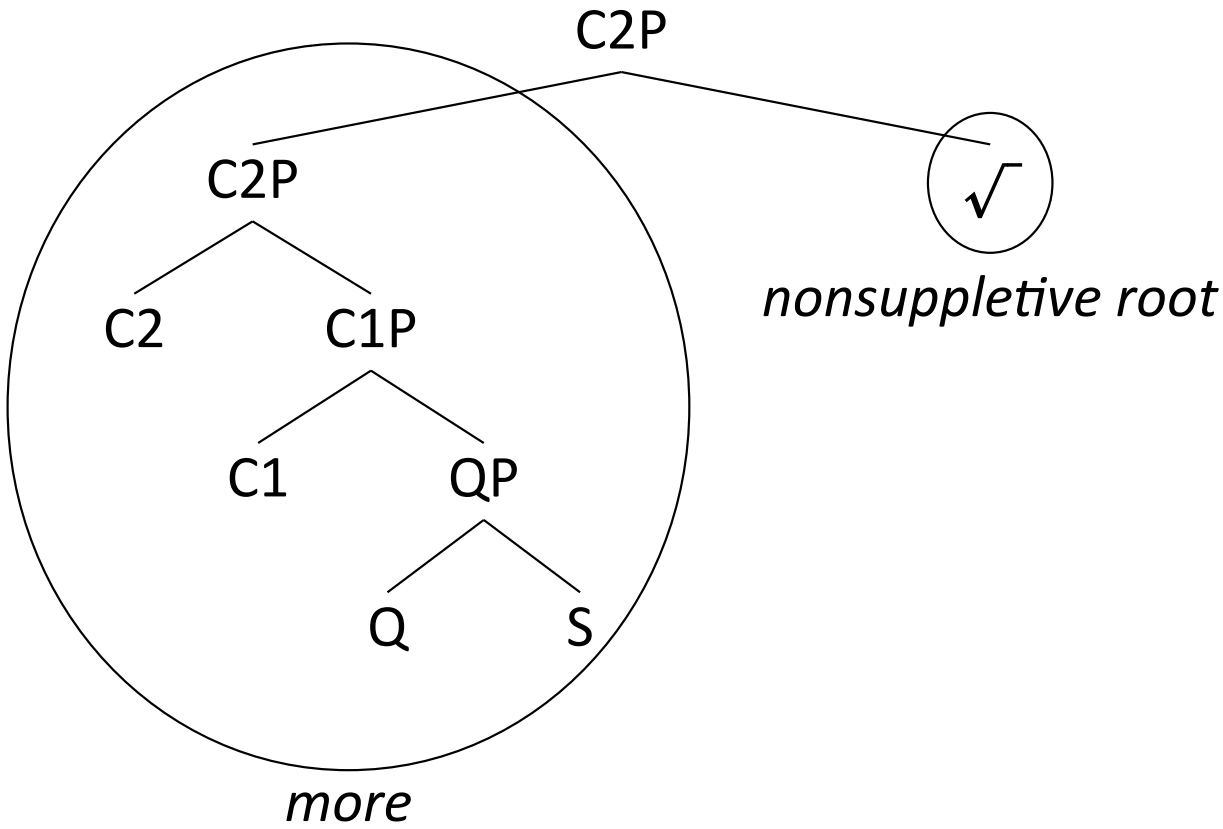
- the proposal works for English, where *more* is really big, and the adjective must consequently be small (too small to be suppletive)

(65)



- the proposal works for English, where *more* is really big, and the adjective must consequently be small (too small to be suppletive)

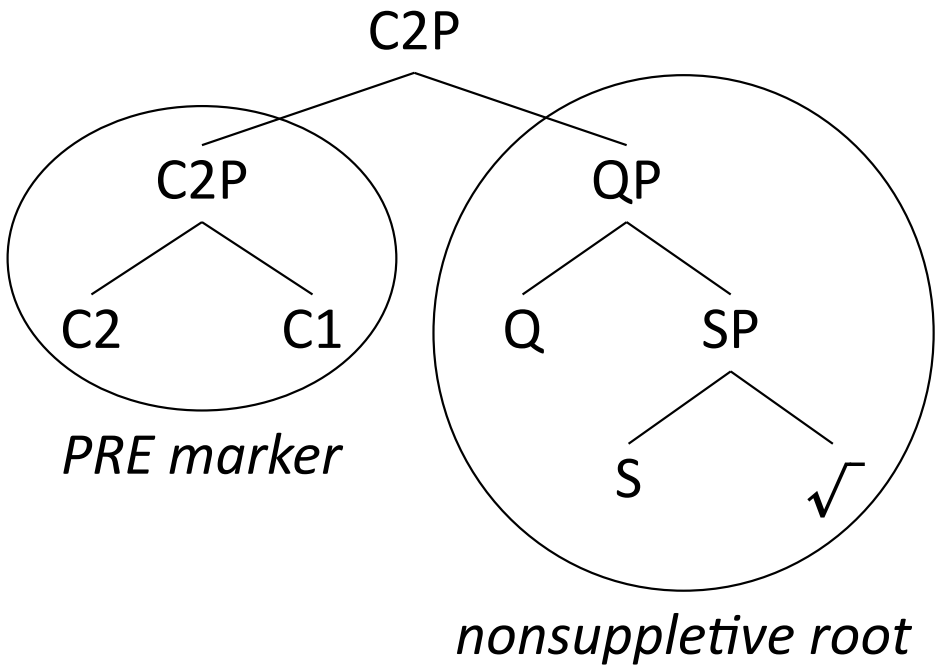
(65)



- but can this conclusion be generalised?

- languages may have PRE markers which are smaller than English *more*
- but given the Spellout Algorithm, a PRE marker needs to realise at least two features
- therefore, the adjectival root is maximally QP (=nonsuppletive)

(66)



GOSP is derived, assuming

- an fseq <C2, C1, Q, S, $\sqrt{}$>
- the Spellout Algorithm and the binary nature of Merge
- a restriction against feature overlap
- an Unmerge operation, which removes structure from the main spine

Summary

- GOSP: PRE marking is incompatible with suppletion
- GOSP follows from a ban against overlapping derivations
- an Unmerge operation may remove previously generated structure in the main derivation
- adjectival roots may vary in size
- comparative marking varies in function the size of the root: as the root grows, the marking shrinks, and vice versa
- POST marking involves lexical items with a unary bottom, PRE marking lexical items with a binary bottom

References

Bobaljik, J. (2012), *Universals In Comparative Morphology*, MIT Press, Cambridge, MA.

Bresnan, J. (1973), 'Syntax of the comparative clause construction in English', *Linguistic Inquiry* **4**, 275–343.

Caha, P. (2017*a*), Explaining Bobaljik's root suppletion generalization as an instance of the adjacency condition (and beyond), *in* J. Emonds and M. Janebová, eds, 'Language use and linguistic structure: Proceedings of the Olomouc Linguistics Colloquium 2016', Palacký university, Olomouc, pp. 193–208.

Caha, P. (2017*b*), 'Suppletion and morpheme order: Are words special?', *Journal of Linguistics* **53**(4), 865–896.

Caha, P., De Clercq, K. and Vanden Wyngaerd, G. (2017), The fine structure of the comparative. Ms. Masarykova Univerzita, U Ghent, KU Leuven.

Corver, N. (1997), 'Much-support as a last resort', *Linguistic Inquiry* **28**, 119–164.

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

Embick, D. and Noyer, R. (2007), Distributed morphology and the syntax/morphology interface, *in* G. Ramchand and C. Reiss, eds, 'The Oxford Handbook of Linguistic Interfaces', Oxford University Press, Oxford, pp. 289–324.

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

Graziano-King, J. (1999), Acquisition of comparative forms in English, PhD thesis, CUNY, New York.

Matushansky, O. (2013), More or better: On the derivation of synthetic comparatives and superlatives in English, *in* O. Matushansky and A. Marantz, eds, 'Distributed Morphology Today: Morphemes for Morris Halle', MIT Press, Cambridge, MA, pp. 59–78.

Starke, M. (2014), 'Cleaning up the lexicon', *Linguistic Analysis* **39**, 245–256.

Starke, M. (2018), Complex left branches, spellout, and prefixes, *in* L. Baunaz, K. De Clercq, L. Haegeman and E. Lander, eds, 'Exploring Nanosyntax', Oxford University Press, Oxford, pp. 239–249.