Allomorphy without Context Specification: Case study of Czech Adjectival Stems

Pavel Caha Karen De Clercq
Guido Vanden Wyngaerd

Abstract

This paper investigates stem-marker allomorphy in Czech adjectives. It shows that an analysis based on the frequently used context-sensitive rules comes at the expense of having to postulate widespread accidental homophony or disjunctive rules. The paper further demonstrates that the allomorphy can be accounted for within an approach based on portmanteau realisation of features, specifically the version of Nanosyntax proposed in Starke (2018), although alternative implementations are conceivable. Along the way, we explore a fine-grained decomposition of adjectival meaning and we also discuss the implications of these observations for the general issues surrounding context-sensitive rules compared to other systems of dealing with allomorphy.

1 Introduction

In Czech, there is a sizeable class of adjectives, which, morphologically speaking, correspond to a root directly followed by agreement. An example is in (1).

(1) mlad- ý
young AGR
'young'

Another large class of adjectives has a special marker n, which occurs between the root and the agreement marker, as in (2). We sometimes refer to this morpheme as an 'augment.'

(2) snad- **n**- ý
easy- N AGR
'easy'

The reasons for analysing n in (2) and elsewhere as an independent morpheme are the following. First of all, some (though not all) roots that require n exist independently of the stem marker, see, for instance, (3).

(3) a. jas
light.NOM.SG
'bright light'
b. jas-**n**-ý
light-N-AGR
'bright, clear'

Secondly, the stem marker n is sometimes absent in the comparative, as in (4):

(4) a. snad-**n**-ý
easy-N-AGR
'easy'
b. snaz-š-í
easy-CMPR-AGR
'easier'

Another indication of the morphemic status of n is the sheer number of adjectives whose stem ends in n. Table 1 provides a representative (even if not exhaustive) list of adjectives in the two classes. The size of the table indicates that both classes are relatively well represented in Czech. The zero class appears to be slightly less numerous than the n class: the list of adjectives in the Ø class is near exhaustive, while more adjectives could easily be added to the list of the n adjectives.[1]

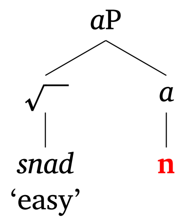
The morphological complexity of adjectives with the stem marker n suggests a structure for the positive degree with at least two un-

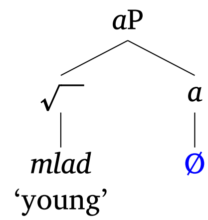
[1] The Ø in Table 1 only indicates the absence of any marking between the root and the agreement marker. In the analysis that we shall develop, we do not assume this zero morpheme.

Table 1: The two classes of adjectives (Ø vs. *n*)

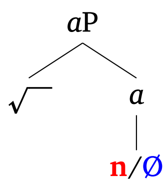
POS	GLOSS	POS	GLOSS
bíl-Ø-ý	‘white’	cen-**n**-ý	‘valuable’
blb-Ø-ý	‘stupid’	čer-**n**-ý	‘black’
bos-Ø-ý	‘barefoot’	čest-**n**-ý	‘honest’
čast-Ø-ý	‘frequent’	děs-**n**-ý	‘horrible’
čil-Ø-ý	‘lively’	div-**n**-ý	‘strange’
čir-Ø-ý	‘pure’	drob-**n**-ý	‘tiny’
čist-Ø-ý	‘clean’	drs-**n**-ý	‘rough’
dlouh-Ø-ý	‘long’	hluč-**n**-ý	‘noisy’
dobr-Ø-ý	‘good’	hnus-**n**-ý	‘disgusting’
drah-Ø-ý	‘expensive’	hod-**n**-ý	‘kind’
drz-Ø-ý	‘cheeky’	hroz-**n**-ý	‘horrible’
hloup-Ø-ý	‘stupid’	jas-**n**-ý	‘clear’
hněd-Ø-ý	‘brown’	jem-**n**-ý	‘smooth’
hol-Ø-ý	‘naked’	klid-**n**-ý	‘calm’
hust-Ø-ý	‘dense’	krás-**n**-ý	‘beautiful’
chab-Ø-ý	‘weak’	lev-**n**-ý	‘cheap’
chud-Ø-ý	‘poor’	mast-**n**-ý	‘fatty’
chor-Ø-ý	‘sick’	mír-**n**-ý	‘peaceful’
jist-Ø-ý	‘secure’	moc-**n**-ý	‘powerful’
krut-Ø-ý	‘cruel’	mož-**n**-ý	‘possible’
mal-Ø-ý	‘small’	nemoc-**n**-ý	‘sick’
mil-Ø-ý	‘lovely’	něž-**n**-ý	‘tender’
mlad-Ø-ý	‘young’	pěk-**n**-ý	‘pretty’
modr-Ø-ý	‘blue’	pev-**n**-ý	‘firm’
nah-Ø-ý	‘naked’	pl-**n**-ý	‘full’
nov-Ø-ý	‘new’	prázd-**n**-ý	‘empty’
plach-Ø-ý	‘timid’	rov-**n**-ý	‘straight’
ploch-Ø-ý	‘flat’	sil-**n**-ý	‘strong’
pust-Ø-ý	‘barren’	sla-**n**-ý	‘salty’
rychl-Ø-ý	‘fast’	slav-**n**-ý	‘famous’
slab-Ø-ý	‘weak’	sluš-**n**-ý	‘kind’
slep-Ø-ý	‘blind’	skrom-**n**-ý	‘modest’
star-Ø-ý	‘old’	smut-**n**-ý	‘sad’
tepl-Ø-ý	‘warm’	snad-**n**-ý	‘easy’
such-Ø-ý	‘dry’	straš-**n**-ý	‘terrible’
tich-Ø-ý	‘quiet’	špat-**n**-ý	‘bad’
tup-Ø-ý	‘blunt’	šťast-**n**-ý	‘happy’
tvrd-Ø-ý	‘hard’	tuč-**n**-ý	‘fat’
zdrav-Ø-ý	‘heathy’	vol-**n**-ý	‘free’
zl-Ø-ý	‘evil’	vliv-**n**-ý	‘influential’
zlat-Ø-ý	‘golden’	vtip-**n**-ý	‘funny’
živ-Ø-ý	‘living’	zálud-**n**-ý	‘wicked’

derlying positions, as in (5a). In the tree, there is a root position lexicalized by *snad* 'easy' and a position labelled little *a* (an adjectivizing head) lexicalized by the n morpheme.

(5) a.


b.


If the structure (5a) is adopted, then, for the adjectives that do not take an augment, one could argue that *a* is lexicalized by a zero morpheme, as in (5b). The proposals in (5) ultimately boil down to the idea that we are dealing with two allomorphs of little *a*, as in (6).

(6)


This article compares two approaches to the allomorphy in (6). One makes use of context-sensitive rules, while the other one is based on portmanteau lexicalisation. Let us sketch the two alternatives below.

Under the context-sensitive approach (see amongst others Siegel 1977, Halle & Marantz 1993, Moskal & Smith 2016, Choi & Harley 2019), one would specify the stem marker n as a context-free realisation of little *a* (based on the fact that adjectives with n are more numerous), while Ø would appear in the context of a restricted set of roots, namely those in the left-hand column in Table 1.

(7) Context-sensitive rules

a. $[a] \Leftrightarrow$ n

b. $[a] \Leftrightarrow$ Ø / $\sqrt{\text{YOUNG, WHITE, \ldots}}$ __

An alternative way of depicting this proposal is in (8): the idea is that depending on which specific root lexicalises the $\sqrt{}$ position, a particular adjectival marker appears (Ø or n). The two markers are not distinguished in terms of the features that they realise, but purely by context.

(8)

	$\sqrt{}$	*a*
EASY	*snad*	*n*
YOUNG	*mlad*	*Ø*

The second possible analysis is phrased in terms of portmanteau lexicalisation (see McCawley 1968, Halle & Marantz 1993, Williams 2003, Chung 2007, Neeleman & Szendroi 2007, Radkevich 2010). In this analysis, the root of the adjectives of the Ø class lexicalises two heads at the same time, that of the root and the little *a*, as shown in (9a).

(9) Portmanteau realisation

a. $\sqrt{\text{YOUNG}} + a \Leftrightarrow$ *mlad*

b. $[a] \Leftrightarrow$ **n**

In this approach, there is no zero morpheme Ø: the absence of *n* arises simply as a result of the fact that there is no need to lexicalise little *a* with roots that are capable of doing this on their own, as shown on the second line of (10).

(10)

	$\sqrt{}$	*a*
EAS-Y	*snad*	*n*
YOUNG	*mlad*	

On this approach, the stem marker n is the only marker lexicalising little *a*, and it appears with all roots that fail to lexicalise this head, as on the first row in (10).

This article shows that these are not equivalent analyses, as also discussed in Embick & Marantz (2008), Bobaljik (2012: 146-152), Banerjee (2021a,b) and Caha (To appear). This paper contributes to this debate by pointing out that the portmanteau analysis in (10) is capable of describing patterns of stem distribution in an elegant and straightforward manner, while a theory based on context-sensitive rules must postulate a proliferation of accidentally homophonous lexical entries, or disjunctive specifications.

The specific pattern demonstrating the advantage of the portmanteau approach concerns the distribution of the stem markers in the positive and the comparative. Specifically, not all adjectives that have the stem marker n in the positive also have it in the comparative, and *vice versa*. In total, four different root classes can be distinguished based on whether they have n or zero in the positive and the comparative, as depicted in Table 2.[2]

Table 2: Four different root classes

	POS	CMPR
I	n	n
II	Ø	n
III	Ø	Ø
IV	n	Ø

As we can see, Class I roots have n both in the positive and in the comparative. Class II roots have n only in the comparative. Class IV roots have n only in the positive degree. Finally, Class III roots do not have n in any of the two forms.

The main goal in this paper is to establish the empirical pattern and show that if context-sensitive rules are used, the distribution of stem markers as in Table 2 can only be described if one invokes multiple accidentally homophonous lexical entries. In contrast, the same pattern can be straightforwardly modelled within the portmanteau approach, relying on the Nanosyntax model of lexicalisation as proposed in Starke (2018).

The paper is organised as follows. In Section 2, we describe the distribution of the stem markers n and Ø in the positive. Section 3 turns to the comparative, and Section 4 discusses the shortcomings of context-sensitive rules. In Section 5, we establish our assumptions regarding the structure of the positive and comparative degrees. In Section 6, we present the portmanteau analysis of classes I-III, and Section 7 shows how the distribution of stems in Class IV can be captured. Section 8 describes the Nanosyntactic derivations in detail, and Section 9 concludes.

[2] The ordering of the classes from I to IV is not random, but reflects (for the classes I-II-III) increasing root size, as will be explained in section 6.

2 The positive

This section discusses the possible reasons why certain roots appear with the stem marker n while other roots appear with no overt stem marker (Ø). We show that the presence/absence of the stem marker cannot be determined by inspecting the phonology or the meaning of the root. The quality of the stem marker also cannot be predicted from the morphological category of the base. We therefore conclude that the selection between the root and the stem marker is governed by an arbitrary lexical class of the root.

We begin by showing that the choice between n and Ø is not governed by phonology. The reason is that many adjectives with phonologically similar roots belong in different classes. We have used one such example in (1) and (2) (with the roots *mlad* and *snad*), and many similar cases can be found in Table 1. The strongest case can be provided by a pair of homophonous adjectival roots meaning 'left' and 'cheap.' These different morphemes happen to have the same phonology, namely *lev-*. Yet, one of them (the one meaning 'left') lacks the augment (11a), while the other root with the meaning 'cheap' requires it (11b).

(11) a. lev- Ø- á noha
left AGR leg
'(the) left leg'

b. lev- n- á noha
cheap N AGR leg
'(the) cheap leg'

The point is that it cannot be determined by looking at the phonology of the root (i.e., *lev-*) whether the root will be followed by n or Ø.[3]

An analogous point can be made about the meaning of the root, namely, it cannot be determined by considering the meaning alone whether the root is going to combine with n or -Ø. Thus, there are near synonymous roots like *hrub-* and *drs-*, both meaning 'rough,' with one of them combining with Ø, see (12a), and the other with n, see (12b).

[3]The fact that root of the adjective 'cheap' is indeed *lev-* rather than an unsegmented string *levn-* can be determined on the basis of verbs such as *s-lev-i-t* 'make cheaper,' or *s-lev-a* 'discount,' where the root appears without the adjectival stem marker.

(12) a. hrub- Ø- á pokožka
rough AGR skin
'a rough skin'
b. drs- n- á pokožka
rough N AGR skin
'a rough skin'

The difficulty to find some consistent common meaning across all the roots in either of the classes can be further verified by inspecting Table 1.

Finally, the quality of the stem marker cannot be predicted from the morphological category of the base. Thus, one cannot say that n is found with nominal roots, while Ø is found with adjectival roots. We now demonstrate the perils of such an approach on a couple of examples.

To begin with, it does seem to be the case that a number of adjectives with n may be considered denominal. For example, the adjective *čest-n-ý* 'honest' has a related noun *čest* 'honour,' and may thus be considered a denominal adjective, see (13) (also compare the case of *jas* 'light' in (3) above).

(13) a. čest- n- ý
honest N AGR
'honest'
b. čest
honour

However, not all adjectives with n are denominal. For example, the root of the adjective *skrom-n-ý* 'modest' cannot be used without the augment at all, as illustrated in (15).[4]

(14) a. skrom- n- ý
modest N AGR
'modest'
b. *skrom
Int: 'modesty'

Thus, it cannot be concluded that the morphological category of the root (nominal vs. not) allows one to uniquely determine what kind

[4]The adjectives with n are thus similar to English adjectives in *y*, which often have nominal bases (*cheeky, tricky, fatty*) but not all of them do (*happy, pretty, sloppy*).

of stem marker will be found in the adjectival use of a root.

The same can be demonstrated by the following pair of examples. The example (15) shows that the noun *stříbr-o* 'silver' has a corresponding color/material adjective in n:

(15) a. stříbr- o
silver NOM.NEUT.SG
'silver (metal)'
b. stříbr- n- ý
silver N AGR
'silver (color/material)'

However, the same type of adjective derived from the noun *zlat-o* 'gold' lacks n, despite the fact that the noun appears rather similar to *stříbr-o* in its semantics and morphological structure (both are neuter nouns belonging in the same declension). This is shown in (16).[5]

(16) a. zlat- o
gold NOM.NEUT.SG
'gold (metal)'
b. zlat- Ø- ý
gold AGR
'golden (color/material)'

We conclude from this that neither phonology, nor semantics, or the morphological category of the root allows speakers to predict what kind of stem marker will be used in the adjectival form of a root. Therefore, we shall treat the quality of the stem marker as an arbitrary (lexical) property of the root. This arbitrary nature of the pairing between the root and the stem marker fits well with both of the possible accounts sketched in Section 1. On both accounts, which stem a root has (n vs. Ø) is an arbitrary property of a lexical entry. On the contextual approach, this is due to an arbitrary lexical specification of the context of insertion. On the portmanteau approach, the arbitrariness of selection translates into the arbitrariness of lexical storage as well: some roots are stored as capable of lexicalising *a*, while other roots are unable to do so.

[5] In other Slavic languages, such as Bulgarian, the stem marker n does actually show up in both cases, so Bulgarian has *zlat-o* ∼ *zlat-en* 'gold ∼ golden' on analogy to *srebr-o* ∼ *srebər-en* 'silver (N) – silver (A)'. This further illustrates the idiosyncrasy of these forms.

3 The comparative

This section discusses the distributional pattern of the stem-marker n in the comparative, which was already mentioned in Section 2, and which is repeated here in Table 3.

Table 3: Four different root classes

	POS	CMPR
I	n	n
II	Ø	n
III	Ø	Ø
IV	n	Ø

The main observation is that each of the two stem classes we discussed for positive degree adjectives may have or lack the stem-marker n in the comparative. As a result, each class is further subdivided into two subclasses depending on the quality of the stem marker in the comparative. Classes I and II have the stem-marker n in the comparative, but differ in that Class I has n in the positive, while Class II has Ø. Classes III and IV have Ø in the comparative, but differ in that Class IV has n in the positive, while Class III has Ø.

The table shows that for a given adjective, the presence or absence of n is not only determined by the left context, i.e., by a particular root, but also by the right context, i.e., whether the stem is followed by a comparative marker or not. In some classes, the presence of the comparative makes n disappear (in Class IV), other times, n only appears in the comparative (in Class II). In Classes I and III, the stem marker (whether n or Ø) is unaffected by the degree.

Let us now illustrate these patterns with examples. To present them clearly, let us first mention that in Czech, the comparative itself has two allomorphs: *š* and *ějš* (Křivan 2012). Their distribution is once again arbitrarily determined by the root. We illustrate the arbitrary nature in Table 4, where two roots that are similar in their phonology and meaning each combine with a different allomorph.

In our analysis of these patterns in Section 5, we shall adopt the analysis by Caha et al. (2019) and analyse *ějš* as a sequence of two morphemes, namely *ěj* and *š*. The reason for proposing the decomposition is that both allomorphs share an identical piece, namely *š*,

Table 4: Two allomorphs for CMPR

POS	CMPR
slab -ý	slab -*š* -í
chab -ý	chab -*ějš* -í
weak -AGR	weak -CMPR -AGR

and that these two pieces lead an independent life. In Table 4, it has been already noted that in the comparative of certain adjectives, only *š* appears while *ěj* is absent. In addition, comparative adverbs of adjectives in *ěj-š*, like *chab* 'weak' (17a), have the form *ěj-i*, with *ěj* present and *š* absent, (17b).

(17) a. chab -ěj -š -í
weak CMPR CMPR AGR
'weaker'

b. chab -ěj -i
weak CMPR ADV
'more weakly'

In our description, we will not dwell on this decomposition, but we come back to it later in Section 5.

Since the allomorphy between *š* and *ějš* is determined arbitrarily by the base to which they attach, we could expect that each of the Classes I-IV in Table 3 could have two subclasses, depending on whether the comparative has *š* or *ějš*. In reality, the situation is simpler in that when the comparative adjective has the stem marker n, then the comparative is always *ějš*. This is a result of the fact that n (in a local fashion) controls the allomorphy of the comparative, and so only one comparative allomorph is allowed in Classes I and II. However, when n is absent in the comparative, we find both allomorphs, with their distribution determined by the root. As a result, Classes III and IV as given in Table 3 divide even further into two additional subclasses, depending on the allomorph of the comparative.

With the background in place, consider Table 5, which depicts all the relevant classes as described up to now.[6]

[6]The forms given in the table conform to native speaker intuitions, and they can also be found in the corpus study of the Czech comparative by Křivan (2012), see p. 42, 40,

Table 5: The distribution of n and Ø in the positive and the comparative

		POSITIVE	COMPARATIVE	GLOSS
Class I		pěk-n-ý	pěk-n-ěj-š-í	'pretty'
Class II		žádouc-Ø-í	žádouc-n-ěj-š-í	'desirable'
Class III	a.	bujar-Ø-ý	bujař-Ø-ej-š-í	'merry'
	b.	star-Ø-ý	star-Ø -š-í	'old'
Class IV	a.	pozd-n-í	pozd-Ø-ěj-š-í	'late'
	b.	snad-n-ý	snaz-Ø -š-í	'easy'

We can see that in Classes III and IV, where the comparative morpheme directly follows the root, we find two different sub-classes as a function of the *š* ∼ *ěj-š* allomorphy. In all comparatives, the final vowel *í* is the NOM.SG agreement marker.

The most important thing for our theoretical claims about allomorphy is the existence of Classes I-IV, and the fact that these are arbitrary morphological classes. We have discussed the arbitrary nature of the stem difference (*n* vs. Ø) for the positive. Now we discuss the same issues for the comparative.

The first thing to know is that the absolute majority of adjectives belong in Classes I and III. Thus, for most adjectives, the issue of whether or not an adjective has *n* or Ø in the comparative reduces to the issue of what stem marker there is in the positive. Since this on its own is a matter of arbitrary class membership, we must conclude that the difference between *n* vs. Ø in the comparative is also a matter of arbitrary class membership.

Against this background, let us discuss the minority Classes II and IV, which change the class marker between the positive and the comparative. Let us first discuss Class IV, which loses the stem marker *n* in the comparative. We will be looking at this class side by side with Class I, which has the stem marker *n* both in the positive and in the comparative. The question is whether it can be somehow predicted by looking at the positive alone which adjective is going to keep *n* in

37, 39 for Classes I-IV respectively. Note further that the table somewhat simplifies the facts, in so far as certain adjectives may belong into different classes, as we shall see shortly.

the comparative, and which is not. Leaving aside the fact that there are only few adjectives that lack *n* in the comparative, the answer seems to be that the loss of *n* is unpredictable.

In Table 6 we provide two examples of adjectives with a comparative in *š*. Note that when the comparative *š* attaches to the root without an intervening *n*, there is a consonant mutation from *d* to *z* (a process attested elsewhere in the language).

Table 6: Adjectives with CMPR in *š* that may lose n in CMPR

POS	CMPR		GLOSS
snad-n-ý	snaz-Ø-š-í	snad-n-ějš-í	'easy/easier'
easy-N-AGR	easy-ø-CMPR-AGR	easy-N-CMPR-AGR	
zad-n-í	zaz-Ø-š-í	zad-n-ějš-í	'one which is (more) in the back'
back-N-AGR	back-Ø-CMPR-AGR	back-CMPR-AGR	

As the data show, leaving out *n* with these roots is optional, rather than obligatory, i.e. they can belong either to Class I or to Class IV.[7] Still, the possibility to omit -*n* contrasts with adjectives that are phonologically and/or semantically analogous, and where dropping the *n* is absolutely impossible, see Table 7. These adjectives unambiguously belong in Class I.

Table 7: Adjectives that keep n in CMPR

POS	CMPR		GLOSS
chlad-n-ý	*chlaz-Ø-(ěj)š-í	chlad-n-ějš-í	'cold(er)'
cold-N-AGR	cold-ø-CMPR-AGR	cold-N-CMPR-AGR	
před-n-í	*přez-Ø-(ěj)š-í	před-n-ějš-í	'one which is (more) in the front'
front-N-AGR	front-Ø-CMPR-AGR	front-N-CMPR-AGR	

In sum, the majority pattern is one where adjectives keep the stem marker *n* as we go from the positive to the comparative (Class I). However, there are a few adjectives that allow for n to be missing

[7] The corpus data presented in Křivan (2012: 39) indicate that for the adjective *snadný* 'easy,' dropping *n* is the preferred option (the ratio is approximately 2:1).

(Class IV). As to which adjectives these are is not predictable, and must be memorised on a case by case basis.

The situation is similar with the adjectives that take the *ějš* allomorph of the comparative. Here some adjectives also optionally drop *n* and, when this happens, they exhibit Class IV pattern. An example of such an adjective is provided in Table 8. As said, preserving the *n* seems to be an option, although dropping it is strongly preferred. The corpus data from Křivan (2012) indicate that the ratio is approximately 270:1 in favour of Class IV behaviour, i.e. in favour of dropping the stem marker *n*. To a native speaker, the comparative that preserves the *n* also sounds decidedly worse.

Table 8: optional n drop with *ějš* CMPR

POS	CMPR		GLOSS
pozd-n-í late-N-AGR	pozd-Ø-ějš-í late-CMPR-AGR	??pozd-n-ějš-í late-N-CMPR-AGR	'late(r)'

The adjectives in Table 9 display only Class I behaviour. They show that the reason for dropping the *n* in Table 8 is not phonological in some obvious sense, since phonologically similar roots maintain it. We therefore proceed under the assumption that the presence vs. absence of *n* in various positive and comparative forms is to be treated as allomorphy triggered by arbitrary morphological classes.

Table 9: n retention with *ějš* CMPR

POS	CMPR		GLOSS
prázd-n-ý empty-N-AGR	*prázd-Ø-ějš-í empty-N-CMPR-AGR	prázd-n-ějš-í empty-CMPR-AGR	'empty/emptier'
před-n-í front-N-AGR	*před-Ø-ějš-í front-CMPR-AGR	před-n-ějš-í front-N-CMPR-AGR	'(more) important'

Let us now turn to the other minority class, Class II, which we consider alongside Class III. Both classes lack the stem marker *n* in the positive. The overwhelming majority of such adjectives also lacks *n* in the comparative, forming what we call Class III. However, there are a few adjectives that require *n* in the comparative. All of these adjectives have something in common. Specifically, Křivan (2012: 40) points out that all Class II adjectives are homophonous with present participles in *c* (roughly analogous to English *ing*). While this restriction is interesting, our main point is going to be that the specific adjectives that belong in this class are still unpredictable. (Put simply, not all adjectival participles pattern like this.)

Let us begin the discussion of Class II by pointing out that the restriction to present participles is indeed morphological, rather than phonological: it does not concern the sound corresponding to the grapheme *c*, i.e. [ts], but the morpheme *c* in its capacity as a participial marker. Thus, when an adjectival base ends in the sound *c* that is not a participial marker, then *n* is found both in the positive and in the comparative. This is illustrated in Table 10 for the root *prac* 'work,' and for the bound root *vzác* 'rare.'

Table 10: Adjectives with roots ending in *c* that take n

POS	CMPR	GLOSS
prac-n-ý work-N-AGR	prac-n-ějš-í work-N-CMPR-AGR	'(more) demanding'
vzác-n-ý rare-N-AGR	vzác-n-ějš-í rare-N-CMPR-AGR	'(more) rare'

The point of presenting the adjectives in Table 10 is to show that there is nothing phonologically wrong about having the stem marker *n* follow a *c*-final base in the positive. Yet this is impossible for Class II adjectives like 'desirable,' which must have Ø in the positive, and *n* in the comparative, as is shown in Table 11.

Table 11: Class II adjective in *c*

POS	CMPR	GLOSS
žádouc-Ø-í desirable-Ø-AGR	žádouc-n-ějš-í desirable-N-CMPR-AGR	'(more) desirable'

Having determined that we are dealing with a morphological class rather than a phonological class, let us zoom in on the adjectival present participles. Our main point is that even in this small niche, 'predictability' of behaviour is severely restricted. In other words, while it is the case that only adjectival present participles exhibit this pattern, it is not the case that all adjectival present participles exhibit the relevant pattern.

For instance, the present participle *vroucí*, derived from the verb *vřít* 'to boil,' has an idiomatic reading 'heartfelt', and can be graded in this reading, see Table 12. When the comparative is formed, the stem marker *n* appears.

Table 12: The adjectival present participle of *vřít* 'boil'

POS	CMPR	
vrouc-Ø-í boiling-Ø-AGR	vrouc-n-ějš-í boiling-N-CMPR-AGR	'(more) heartfelt'

However, most participles do not accept any form of suffixal comparative marking. We illustrate this in Table 13 on the participle derived from the verb *šokovat* 'to shock,' see Table 13. This participle has an adjectival reading because it can be preceded by 'very' (Wasow 1977); see the second column in Table 13. However, despite being adjectival, the participle 'shocking' does not accept any comparative suffixes, whether with *n* or without it.

Table 13: The adjectival present participle of *šokovat* 'shock'

POS	DEG modification	CMPR	GLOSS
šokujíc-í shocking-AGR	velmi šokujíc-í very shocking-AGR	*šokujíc(-n)-ějš-í shocking-N-CMPR-AGR	'(very) shocking'

We thus conclude that membership in Class II is subject to lexical idiosyncrasy, rather than predictable on the basis of information that is independently available when the positive degree is inspected. Therefore, what we need is a theory of how each adjectival base combines with the stem marker (*n* vs. Ø) in a given degree construction. These selectional requirements must be sensitive to the identity of a specific root/base, rather than valid for all adjectives across the board.

Before we look into the details of how this works, we again note that context-sensitive rules and portmanteau lexical entries are (in general terms) appropriate devices for the task, because they indeed involve stipulations over the content of specific lexical items (in the portmanteau approach) or stipulations over the sensitivity to other lexical items (contextual allomorphy). However, as we now proceed to show, context-sensitive rules cannot capture the distribution of *n* and Ø without postulating multiple instances of accidental homophony, while the portmanteau approach can.

4 Context-sensitive rules

This section shows that in an approach based on context-sensitive rules, the distribution of stems in Table 5 (repeated here as Table 14 for convenience) necessarily leads to a proliferation of accidental homophony or disjunctive context specification.

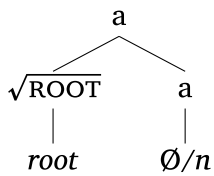
Table 14: The distribution of n and Ø in the positive and the comparative

		POSITIVE	COMPARATIVE	GLOSS
Class I		pěk-n-ý	pěk-n-ěj-š-í	'pretty'
Class II		žádouc-Ø-í	žádouc-n-ěj-š-í	'desirable'
Class III	a.	bujar-Ø-ý	bujař-Ø-ej-š-í	'merry'
	b.	star-Ø-ý	star-Ø -š-í	'old'
Class IV	a.	pozd-n-í	pozd-Ø-ěj-š-í	'late'
	b.	snad-n-ý	snaz-Ø -š-í	'easy'

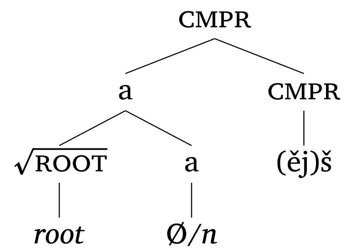
The first step on this path is to adopt a specific structure for comparatives. In the seminal work on comparatives by Bobaljik (2012), it has been proposed that the structure of the comparative contains that of the positive (see also Smith et al. 2019). This is depicted in (18), where the comparative structure in (18b) adds an additional CMPR head on top of the positive, which is in (18a).[8]

[8]When we develop the portmanteau-based account in Section 6, we shall enrich this

(18) a. POSITIVE



b. COMPARATIVE



With the structures in place, let us come back to the Vocabulary Items (VIs) needed to get the correct realisation of the little *a* head. Recall from (7) that the initial strategy was to say that n is the default realisation of little *a*, see (19a).

(19) a. *a* ⇔ n
b. *a* ⇔ Ø / Class III __

The effect of this context-free rule is that it inserts the stem marker n after every root, unless counteracted by more specific VIs. The VI that we originally used to restrict its application was the one in (19b). Due to its context specification, this is a more specific VI and it therefore wins over the more general one, realising little *a* as zero in the context of Class III roots. These two VIs, taken together, correctly model the pattern found in Classes I and III in Table 14.

Let us now consider how we can extend or modify this set of VIs to account for the newly discovered Classes II and IV. The issue is that as the VIs currently stand in (19), we expect Ø realisation of little *a* only in Class III, and n in all other contexts. For Classes II and IV, this is correct only partially. For example, in Class IV, we see n only in the positive, but not in the comparative.

To prevent n from appearing in the comparative of Class IV, we need to have a VI that applies in this environment and realises little *a* as Ø, thereby blocking the insertion of n. The VI in question is given in (20).

(20) *a* ⇔ Ø / Class IV __] CMPR

This VI competes with the two VIs in (19), and they jointly deliver a pattern as observed in Class IV, with the default realisation of little *a* in the positive, and a Ø realisation in the comparative.

structure by decomposing the comparative into three separate heads, but we will maintain the containment of the positive inside the comparative.

However, the cost to pay is that we have introduced accidental homophony into the system: we now have two different VIs inserting the same exponent Ø, namely (19b) and (20). It is meaningless to simply unify the two VIs introducing Ø by saying that Ø is inserted either in the context of Class III roots or in the comparative of Class IV roots, since the VI thus created would have to be specified for a disjunction of environments. As Christopoulos & Zompí (2023: 11) put it in their recent work, specifying a single rule for a disjunctive context has exactly the same effect as accidental homophony. In doing so, '…admitting [disjunctive] rules […] (without anything else restricting possible disjunctions) opens the door to describing any type of exponent distribution under any theory of features by simply listing the contexts in which each exponent appears.'

To summarise, the distribution of stem markers in Class IV is problematic for context-sensitive rules, since these rules cannot capture it without invoking accidental homophony, or VIs with disjunctive context specifications (which basically amounts to the same thing).

The distribution of stem markers in Class II is problematic in the same way. Based on the rules in (19), we again expect to see n both in the positive and in the comparative, since there is nothing blocking the application of the elsewhere rule (19a). However, empirically, we only find n in the comparative of Class II.

In order to prevent the context-free n from realising little *a* in the positive degree of Class II adjectives, we need to postulate a VI that only applies in the positive degree of this class. A VI like this is given in (21).

(21) *a* ⇔ Ø / Class II __] in the absence of CMPR

This VI introduces yet another instance of accidental homophony, since, once again, the exponent is Ø. However, the VI is problematic in yet another respect, namely the fact that it is triggered by a negative condition. Allowing negative specifications of contexts is again subject to the criticism voiced by Christopoulos & Zompí (2023) that it allows any kind of distribution of exponents. An anonymous reviewer further points out that it is doubtful whether negative conditions may serve as a trigger for a rule, as they are more reminiscent of a filter on the output of a grammar.

Before we leave this section, it is worth pointing out that the very same issues arise if we consider Ø to be the default exponent, as in (22a). In this setting, the Ø will appear in any environment where it is not blocked by a more specific VI. Once again, the problem is that the set of environments where Ø does not appear is a disjunctive set of environments, which means we must have several different VIs for what looks like a single marker, namely n. As in the scenario where n is the default marker, we end up with three accidentally homophonous VIs, as shown in (22). For good measure, we also repeat the VIs needed for the first scenario, where n is the default, in (23).

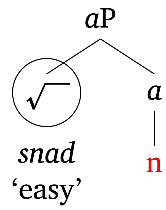
(22)
a. $a \Leftrightarrow$ Ø
b. $a \Leftrightarrow$ n / Class I ___
c. $a \Leftrightarrow$ n / Class II ___] CMPR
d. $a \Leftrightarrow$ n / Class IV ___] in the absence of CMPR

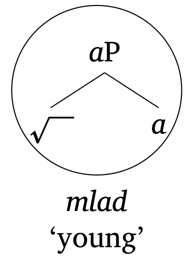
(23)
a. $a \Leftrightarrow$ n
b. $a \Leftrightarrow$ Ø / Class III ___
c. $a \Leftrightarrow$ Ø / Class IV ___] CMPR
d. $a \Leftrightarrow$ Ø / Class II ___] in the absence of CMPR

Summarising, the pattern of distribution observed for the stem markers *n* and Ø cannot be captured by context-sensitive rules without a proliferation of accidentally homophonous morphemes. It also requires negative specifications, i.e., the conditioning of rule application by the absence of specific features. In the remainder of this article, we show that if we adopt the Nanosyntax model of phrasal lexicalisation, the distribution of the stem markers can be captured without invoking either accidental homophony or negative specifications.

5 Setting up the features

The main idea behind the portmanteau account is that roots, stored in the post-syntactic lexicon, may be lexically associated to constituents of different sizes. For example, considering the simple structures in (24), we can have a set of roots that is lexically specified as lexicalising just the root node (24a), while other roots are associated to a larger constituent, and may lexicalise all the features of the positive, as in (24b).

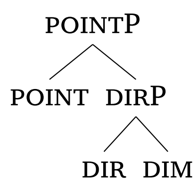
(24) a. 

b. 

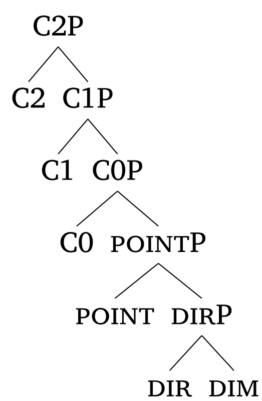
The interest of the analysis (24) is that the root controls the allomorphy between Ø and n without the need for any contextual specification. When the root lexicalises little *a*, there simply is no need to realise it for the second time by n. Another interesting consequence is that this analysis requires no dedicated Ø morpheme since the absence of n is due to phrasal lexicalisation. As a result, the conundrum of how to set up the distribution of this morpheme using context-sensitive rules simply never arises.

In developing the analysis, we shall adopt a higher resolution decomposition of the positive and the comparative degrees than the one we have assumed so far. In (25a), we depict the structure of the positive. In this proposal, the positive decomposes into three features, DIM, DIR and POINT (see Vanden Wyngaerd et al. 2020).

(25) a. POSITIVE



b. COMPARATIVE

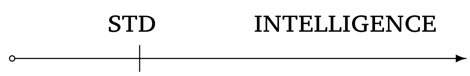


The features inside DIRP (DIR and DIM) represent within syntax the scale underlying the semantics of gradable adjectives, sometimes referred to also as the measure function (see, e.g., Kennedy 1999). The features DIM and DIR represent two independent components of each adjectival scale, namely its DIMension and DIRection. The dimension corresponds to some domain of measurement, like VELOCITY or SIZE. The DIRection encodes the ordering of the values of a particular dimension. The DIRection can be positive or negative, distinguishing antonymous adjectives that share the same dimension, e.g., *fast* vs. *slow*, or *tall* vs. *short*.[9]

The feature POINT represents a POINT on the scale that the argument of the adjective must exceed. In the positive degree, the placement of this point is contextually given, and represents the standard according to which the adjective is evaluated. The picture in (26) shows the step-wise composition: DIM and DIR give us the scale in (26a), the projection POINT introduces a point (the standard) that the argument of the adjective must exceed (26b).

(26) a.

b.



As mentioned, in the absence of any specific instructions where the POINT is to be placed, its position is determined by context. This accounts for the fact that in the positive degree, the standard is context-dependent: who or what counts as *fast* or *smart* depends on the comparison class. With certain classes of adjectives (the so-called minumum-standard adjectives), the POINT coincides with the zero point on the scale (Kennedy & McNally 2005). With such adjectives, it is enough that the argument of the adjective exceeds this zero standard for the positive degree to hold. For example, for a door to count as *open*, it is enough for it to attain a non-zero degree of openness. Similarly for the adjective *wet*: for a chair to count as wet, a non-zero degree of wetness is enough.

Let us now move to the structure of the comparative in (25b). As is obvious, our analysis of the comparative inherits from Bobaljik (2012) the idea that the structure of the positive is contained inside the comparative. At the same time, our structure differs from Bobaljik's in that it has three comparative heads, rather than just one.[10]

The most immediate reason for us to propose the existence of

[9]De Clercq & Vanden Wyngaerd (2019) provide empirical evidence that the distinction between positive and negative adjectives is syntactically relevant.

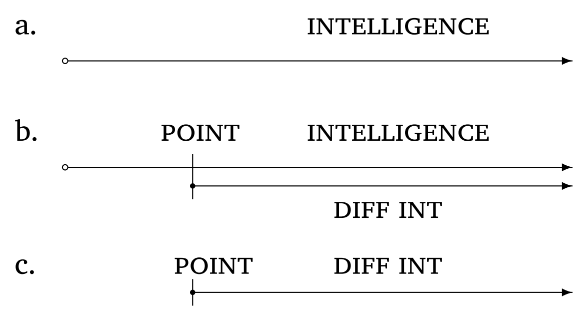
[10]See Caha (2017), De Clercq & Vanden Wyngaerd (2017) and Caha et al. (2019) who propose two comparative heads for Czech.

three comparative heads are the Class II adjectives, where the comparative has three extra morphemes compared to the positive (they are placed inside the rectangle):

(27) žádouc-í ∼ žádouc-n-ěj-š-í
desirable-AGR desirable-N-CMPR-CMPR-AGR
'desirable ∼ more desirable'

Thus, in order to accommodate the three extra morphemes, we propose three extra heads. As far as their semantics is concerned, we think that the best way to make sense of them is to adopt the idea by Kennedy & Levin (2008), who propose that comparatives involve the construction of a new scale, which is derived from the one of the positive degree. If this is so, then the heads C0-C1-C2 essentially construct a new scale with the three relevant parameters of dimension, direction and point.[11]

This idea is informally depicted in (28), with (28a) showing the scale of intelligence (DIR + DIM). Then a POINT is placed on the scale, see (28b). Unlike in the positive, the position of the point in the comparative is not determined by the context, but by the *than*-phrase. For instance, in an example like *smarter than John*, the POINT in (28b) represents the degree to which John is smart.

(28)


Once the position of the point is determined, a new scale is constructed along the same dimension, see the lower scale in (28b), but with its starting point shifted to the value of the *than*-phrase (see Kennedy & Levin 2008 for details). This derived scale is called a 'difference measure function.' The abbreviation DIFF below the lower scale in (28b) stands for the difference measure function, while INT

[11]We are indebted to Michal Starke (seminar, Brno, 2019) for this idea that we further develop in the remaining paragraphs of this section.

marks that this is a scale of intelligence.

Our hypothesis is that the construction of the difference measure function (on the basis of the positive) is the job of the heads C0 and C1. We already established that constructing a scale requires the projections of DIR and DIM; C0 can then be understood as representing the dimension of the new scale. The dimension introduced by C0 is always anaphoric to the dimension of the scale of the positive degree (INTELLIGENCE in our case). This has the consequence that, whatever dimension we find in the positive (HEIGHT, VELOCITY, INTELLIGENCE, etc.), the comparative will always have the same dimension. C1 then encodes, within syntax, the direction of the new scale (which can again be positive, as in *more intelligent than John*, or negative, as in *less intelligent than John*).

Finally, as in the positive, a POINT must be introduced on the derived scale in (28c), and the argument of the adjective must exceed this degree. Kennedy & Levin (2008) propose that the difference measure function corresponds to a type of a scale that has a zero standard, so as soon as the argument has a non-zero value on the derived scale (i.e., as long as it exceeds the degree of the *than*-phrase), it satisfies the condition. The introduction of the point on the derived scale is the role of C2.

In sum, we are suggesting that there are up to three different heads that derive the comparative meaning from the positive. This move is not only motivated by the complexity of the comparative morphology in Czech, but we have also argued that it is compatible with the semantic composition of the comparative as proposed in Kennedy & Levin (2008). Specifically, the three heads of the comparative are similar to the heads DIR, DIM and POINT found in the positive degree, and construct a derived measure function on the basis of the scale belonging to the positive-degree adjective.

In what follows, we shall assume these structures for the positive and the comparative, and we shall use them to explain the four classes of adjectives with regard to the patterning of the stem-marker n.

6 Accounting for Classes I-III

In this section, we use the proposed representations to provide an account of Classes I-III. We start by summarising the facts we discussed above in Table 2, but now including the two subclasses of Classes III and IV, yielding a total of six classes. This is shown in Table 15. Classes I-III are separated from Class IV by a line. The goal in this section is to demonstrate that Classes I-III can be easily accounted for by specifying different root sizes for the adjectives of these Classes, assuming the highly articulated adjectival projection proposed in the previous section.

Table 15: Six classes of adjectives

	POS	CMPR
I	n	n-ěj-š
II	Ø	n-ěj-š
IIIa	Ø	Ø-ěj-š
IIIb	Ø	Ø-Ø-š
IVa	n	Ø-ěj-š
IVb	n	Ø-Ø-š

In our account, we assume that syntax constructs abstract layered representations such as those provided in (25a) and (25b). These structures are mapped on the phonological representation after syntax by (late) lexical insertion. The specific model of lexicalisation we assume is the Nanosyntax framework as described in Starke (2018) and much related work (Baunaz & Lander 2018, Caha et al. 2019, Wiland 2019, Janků 2022, Cortiula 2023, Caha et al. To appear, etc.).

In this section, we do not dwell on the technical algorithmic aspects of the lexicalisation procedure, to which we return in Section 8. For now, we depict the outcome of the procedure in the form of lexicalisation tables. Lexicalisation tables keep track of which morpheme lexicalises which feature in the complex tree representation, abstracting away from how exactly this lexicalisation was achieved, allowing readers unfamiliar with the framework to understand the essence of the proposal.

An example of a lexicalisation table is given in (29) below for Class I.

(29) Class I

	DIM	DIR	POINT	C0	C1	C2
POS	root I		n			
CMPR	root I		n		ěj	š

The column headings represent the features in the functional hierarchy that we have argued for in the previous section, and the different rows represent the forms of the positive and the comparative of Class I adjectives, respectively. A form that spans different columns is a portmanteau morpheme, i.e. one that lexicalises more than one feature. The black shading below C0-C2 indicates that these projections are absent in the positive degree.

Our account of Class I is based on the idea that roots in this class only lexicalise DIM and DIR. Since such roots fail to lexicalise all the features of the positive, they need the stem marker n to lexicalise the POINT projection. This is shown on the first line of the lexicalisation table (30).

The analysis of the comparative is depicted on the bottom row in (29). It assumes that the stem marker n can lexicalise not only POINT, but also the lowest comparative projection C0. The markers *ěj* and *š* then lexicalise C1 and C2 respectively.

The lexical items needed for this analysis are provided in (30).

(30) a. /Class I root/ ⇔ [DIR DIM]
 b. /n/ ⇔ [POINT C0]
 c. /ěj/ ⇔ [C1]
 d. /š/ ⇔ [C2]

In order for the analysis to work smoothly, we need to make explicit an assumption, which is that a lexical item may lexicalise all the features it is specified for, or a subset of them. We need this to allow the stem marker n to lexicalise just the POINT feature in the top row in (29). Since POINT is contained in the entry (30b), n may be used to lexicalise just this feature.

We will refer to this as the Superset effect, and we will take it for granted in the following discussion. In the theory of Nanosyntax, this effect follows from the matching condition, which defines when an entry matches the structure, and it is sometimes referred to as The Superset Principle (Starke 2009). We come to the exact wording of the matching condition in Section 7.

Let us now turn to Class II. Our analysis of this class assumes that, since the roots of this class have no overt stem marker in the positive, they can lexicalise all the features of the positive, as shown in (31). All the other lexical items keep their specification as in (30).

(31) /Class II root/ ⟺ [DIR DIM POINT]

With this lexical entry, the root is able to lexicalise all the features of the positive without the need for any extra stem marker (see the top row in (32)).

(32) Class II

	DIM	DIR	POINT	C0	C1	C2
POS	root II					
CMPR	root II		n		ěj	š

In the comparative, the feature C0 cannot be lexicalised by the root, since the root does not contain this feature. Therefore, the stem marker -*n* appears as the lexicalisation of this feature. Since *n* is also specified for POINT, recall (30b), it also lexicaclises this feature. The comparative markers *ěj* and *š* lexicalise C1 and C2, respectively, which is what they also do in Class I.[12]

Let us now turn to the Class IIIa. This class can be accounted for under the assumption that its roots lexicalise all the features from DIM up to and including C0, as in (33).

(33) /Class IIIa root/ ⟺ [DIR DIM POINT C0]

With this specification, the root can lexicalise all the features in the positive, as on the first row of (34). Recall that due to the Superset effect, the root need not lexicalise all the features it is specified for, which makes it a good match in the positive degree, despite the absence of C0 in the structure of the positive.

(34) Class IIIa

[12]In principle, the feature POINT could also be lexicalised by the root in the comparative, and *n* would then only lexicalise C0. However, as we shall show, this option is not compatible with the algorithmic lexicalisation procedure running over constituent structure, as described in Starke (2018). Since both alternatives lead to the same result (i.e., a root followed by *n-ěj-š*), we do not need to decide the issue here.

	DIM	DIR	POINT	C0	C1	C2
POS	root IIIa					
CMPR	root IIIa				ěj	š

In the comparative, the roots of Class IIIa lexicalise all the features up to C0, explaining the absence of n, as well as the presence of *ěj-š*, as depicted on the bottom row in (34).

Finally, roots of Class IIIb can be accounted for under the assumption that they lexicalise all the features from DIM to C1, see (35).

(35) /Class IIIb root/ ⟺ [DIR DIM POINT C0 C1]

In the positive, the root will (again) lexicalise all the features (not using its C0 and C1 specification due to the Superset effect). We depict this on the top row in (36).

(36) Class IIIb

	DIM	DIR	POINT	C0	C1	C2
POS	root IIIb					
CMPR	root IIIb					š

In the comparative, the root lexicalises all the features up to C1, leaving space only for *š*.

Summarising, we have now provided a rather simple account for Classes I-IIIb, without the need to use any context-sensitive rules. Most notably, we have provided an explanation for why in Class II, the marker n is only needed in the comparative. Recall that this was unclear previously, where this class required a zero marker that blocked n (a default lexicalisation of little *a*) in the positive. This was a problematic entry, because of the accidental homophony and the negative specification of the context (in the absence of CMPR). None of these problems arises in the current approach, which dispenses with zero markers and replaces their effects by portmanteau lexicalisation.

Our analysis also explains the fact that the alternation between the two allomorphs of the comparative marker (*ěj-š* vs. *š*) only arises if the comparative morphemes are directly adjacent to the root, and not if **n** intervenes. This is because the size of **n** is fixed (it cannot go higher than C0), so that it can only trigger one allomorph. Roots, in contrast, can have varying sizes, and as a result of that they can give rise to allomorph alternation. In the following section, we turn to Class IV.

7 The proposal for Class IV

This section deals with some interesting analytical challenges offered by the Class IV roots. Let us first remind ourselves of the facts, repeated for convenience in Table 16.

Table 16: Six classes of adjectives

	POS	CMPR
I	n	n-ěj-š
II	Ø	n-ěj-š
IIIa	Ø	Ø-ěj-š
IIIb	Ø	Ø-Ø-š
IVa	n	Ø-ěj-š
IVb	n	Ø-Ø-š

The table shows how, from Class I to Class IIIb, root size increases monotonically, as shown by the increasing number of 'zero' morphemes. Furthermore, the size of the root correlates inversely with the number of overt morphemes that follow the root. Class IV does not fit into this picture, in that the presence of n in the positive suggests a small root (like those of Class I), but then the absence of n in the comparative is puzzling. Looking at the problem from a more general perspective, Class IV reveals a tension between morphological and structural containment. Recall from Section 5 that we have adopted here a version of the proposal by Bobaljik (2012), according to which the comparative degree structurally contains the positive degree. In terms of morphology, however, the Class IV comparative is not built on top of the positive: the positive base is truncated, since the stem marker n has to be deleted before the comparative marking is added.

This type of marking is quite rare, contradicting a candidate universal proposed by Grano & Davis (2018: 133).

(37) *Candidate Universal*
Universally, the comparative form of a gradable adjective is derived from or identical to its positive form.

This universal is violated by the type of marking observed in Class IV, since the comparative here is neither derived from nor identical to the positive degree.

In technical terms, the challenge stems from the fact that when we look at the comparative of Class IV, we see that it lacks n and has only *ěj-š* (Class IVa) or just *š* (Class IVb). This suggests that such roots must be specified for all the features up to C0 or C1 respectively, as on the comparative rows in (38) and (39).

(38) Class IVa

	DIM	DIR	POINT	C0	C1	C2
POS	root IVa		n			
CMPR	root IVa				ěj	š

(39) Class IVb

	DIM	DIR	POINT	C0	C1	C2
POS	root IVb		n			
CMPR	root IVb					š

Given the capacity of the roots to lexicalise all the features from DIM all the way up to C0 or C1 in the comparative (including the stretch DIM-DIR-POINT), the question arises why these adjectives need n in the positive (as the realisation of POINT). The first row of the relevant tables (38) and (39) presents lexicalisations with the n included, but this is exactly the puzzling thing: in the comparative, the exact same features (DIM-DIR-POINT) are lexicalised by the root, so why is this not an option in the positive? Why does n have to appear here?

The question is relevant also in the light of the fact that in Class III, we argued that the ability of the root to lexicalise all the features up to C0 automatically entailed that these roots can also lexicalise all the features of the positive (via the Superset Effect). This was shown in the lexicalisation Table (34) above, repeated for convenience in (40).

(40) Class IIIa

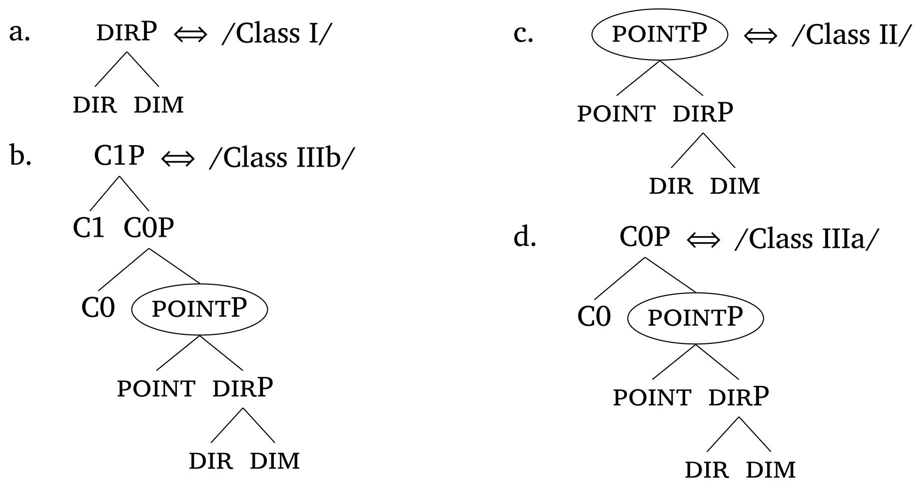
	DIM	DIR	POINT	C0	C1	C2
POS	root IIIa					
CMPR	root IIIa				ěj	š

What the Class III roots display is a property that we can informally refer to as shrinkability: a bigger root (size C0P or C1P) can shrink to lexicalise a smaller syntactic structure that is contained in it (POINTP). This shrinkability of roots is, in fact, a case of the Superset effect. Put in those terms, the puzzle presented by the Class IV roots is why they cannot shrink to size POINTP, given that the Class III roots can do exactly that.

However, it turns out that it is possible to resolve this tension in a way that maintains the assumption that the positive is structurally contained in the comparative. We shall do so by adapting a proposal by Blix (2022), who addresses a similar challenge in the domain of number marking. To see how Blix' idea works, we must consider additional details of the lexicalisation procedure. Once the details are clarified, it turns out that the lexicalisations as given in (38) and (39) are in fact possible, while still preserving the idea that the comparative is built on the positive.

The main feature of the solution stems from the fact that in Nanosyntax, lexical items do not link phonology to a random collection of features. Rather, they link phonology to well-formed syntactic representations (Starke 2014). Given this idea, (41a-d) provide the updated lexical items of the roots belonging to Classes I-IIIb. The entries still contain the same features as before, but since lexical entries link syntactic representations to phonological representations, the features are structured. And it is precisely this aspect of their format that is going to provide a solution to the conundrum presented by Class IV roots. (The meaning of the circled nodes will be clarified shortly.)

(41)



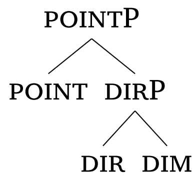
With the updated format of lexical items in place, let us now also present the Matching Condition on lexicalisation (Starke 2009), which is often referred to also as the Superset Principle.

(42) *Matching Condition*
A lexically stored tree matches a syntactic node iff the lexically stored tree contains the syntactic node.

The Matching Condition makes it clear that lexicalisation does not only care about the number of features, but it imposes a stronger requirement, namely constituent identity: a lexical entry only matches a syntactic tree iff it contains an identical tree. Therefore, lexical items can lexicalise the three features of the positive (DIM-DIR-POINT) only if these features form a constituent inside the lexical entry.

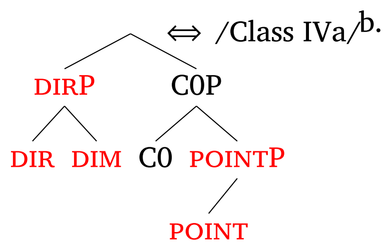
It turns out that roots of classes II, IIIa and IIIb do contain the constituent corresponding to the positive degree, as highlighted by the circles in (41). The fact that a tree identical to the syntactic tree of the positive is literally contained in the lexical trees of (41) is what allows these roots to lexicalise the structure of the positive without any additional morphology. For convenience, (43) presents the structure of the positive as originally presented in (25a).

(43) POSITIVE

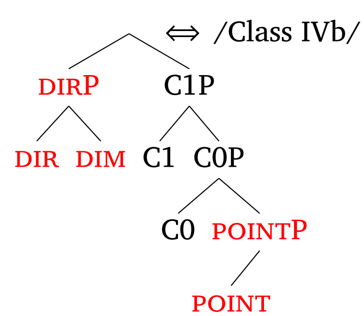


With this background in place, consider the lexical entries that we are proposing for Class IVa and IVb in (44a,b) respectively (taking inspiration from the proposal put forth in Blix 2022):

(44) a.



b.



While the constituency of the lexical entries in (44) may appear unusual, they still adhere to the general idea that lexical items are nothing but links between well-formed syntactic representations and phonology. In this approach, such lexical items are in fact to be expected: it is a logical possibility that a particular constituent moves in the syntax (e.g., the constituent [DIR DIM]), and once it lands higher up in the structure, such a structure (containing a moved constituent in a higher position) is linked to phonology. And this is precisely what the lexical entries in (44) are: they link phonology to a structure where the DIRP has moved from within POINTP to the left.

For us, the most relevant point about these lexical items is that they do contain a constituent that has all the features from DIM up to C0 or C1: this is the top node of the entries. This constituent is relevant for the lexicalisation of the comparative, where all the relevant features are lexicalised by the root. However, there is no constituent inside the entries that contain the three features of the positive. To make this clear, the features of the positive are highlighted, and it is obvious that they do not form a constituent (let alone one identical to (43)). As a result, these entries cannot lexicalise the positive: they do not contain a constituent identical to it.

To understand how the positive is lexicalised, it is relevant to note that the features DIM and DIR still form a constituent in both (44a,b), in the form of the complex left branch. Therefore, the roots can still lexicalise these two features. This is because the derivation of the positive proceeds bottom-up, first merging DIM and DIR. This will create a syntactic constituent that is identical to a constituent contained in (44), so that the root may lexicalise this syntactic constituent. This is exactly what happens in the positive, as depicted in the lexicalisation tables in (38) and (39), repeated below for convenience. However, the feature POINT cannot be lexicalised along with these two features, and that is why n appears in the positive (in a manner that we shall make precise in the next section).

(45) Class IVa

	DIM	DIR	POINT	C0	C1	C2
root-n (POS)	root IVa		n			
root-ěj-š (CMPR)	root IVa				ěj	š

(46) Class IVb

	DIM	DIR	POINT	C0	C1	C2
root-n (POS)	root IVb		n			
root-š (CMPR)	root IVb					š

In the next section, we introduce the lexicalisation system of Nanosyntax in detail, and show how the lexical items of roots interact with the syntactic derivation so that the observed patterns are derived. In particular, we shall show how the syntactic derivation will generate constituents that are identical to those in (44), allowing the roots of Class IV to lexicalise all the features up to C0P (Class IVa) or C1P (ClassIVb), as shown in (45) and (46).

8 The derivations

This section provides the basic principles of lexicalisation in Nanosyntax, and shows how they derive the lexicalisation tables of Classes I-IV, as described above in Sections 6 and 7. This is thus the last step in demonstrating that context-sensitive rules and portmanteau-based lexicalisation are not equivalent systems, and that portmanteau-based lexicalisation allows one to model the facts without the need to introduce multiple cases of accidental homophony or negative conditions on insertion.

Like much of the work in standard minimalism, Nanosyntax (Starke 2018) assumes that syntactic structures are built step by step by successive application of Merge. In addition, Nanosyntax adopts the idea of cyclic lexicalisation. This means that whenever a new feature is merged, the structure is sent to the interface, where it is determined whether the structure can be lexicalised or not. A structure can be lexicalised if it finds a matching item in the lexicon, where matching lexical items are selected based on constituent identity (recall the Matching Condition in (42)).[13]

When the structure is sent to interface and a matching lexical item is found, the derivation may continue by adding additional features. However, when no matching item is found, the structure is rejected at the interface and sent back to syntax for adjustments. These adjustments correspond to various types of evacuation movements, whose type and order is defined by the so-called lexicalisation algorithm.

(47) Lexicalisation algorithm
  a. Merge F and lexicalise FP
  b. If (a) fails, move the Spec of the complement of F, and lexicalise FP
  c. If (b) fails, move the complement of F, and lexicalise FP
  d. If (c) fails, go back to the previous cycle and try the next option for that cycle

The clause (47a) of the algorithm implements the idea of cyclic lexicalisation: always when a feature is merged, a matching item must be found for the newly created FP. (47b-d) specify various repair options, namely Spec movement, complement movement and backtracking. We will introduce these derivational options as we go.

These operations apply algorithmically, i.e. in the order specified, and blindly, i.e., without any knowledge of whether they succeed or not. The success of each operation is always evaluated against the lexicon, which is why each operation is followed by 'lexicalise FP'.

Finally, let us mention that the evacuation movements are triggered

[13] In minimalist terms, the interface must check that the structure is 'legible' at the interface, i.e., whether it can be mapped onto PF using the set of lexical items available for a given language.

by the need to lexicalise the structure, and not by the need to check features. This means that they are different from standard feature-driven movements. In Nanosyntax, they are not only different in what triggers them, but also in their implementation. For instance, it is assumed that they do not leave a trace, since they do not give rise to two interpretive positions (unlike feature-driven movements). Therefore, whenever a phrase will be moved due to evacuation movements, the base position will be simply left empty.

With the background in place, let us now turn to the step-by-step derivation of the individual six classes that we have distinguished above. Each class is discussed in a separate subsection.

8.1 Class I

The Class I roots require n both in the positive and in the comparative. In the comparative, n is followed by *ěj-š*. The pattern is depicted in the lexicalisation table in (48).

(48) Class I

	DIM	DIR	POINT	C0	C1	C2
POS	root I		n			
CMPR	root I		n		ěj	š

In (49), we give the lexical items relevant for this class.

(49)

a. DIRP ⇔ /Class I root/
 [DIRP DIR DIM]

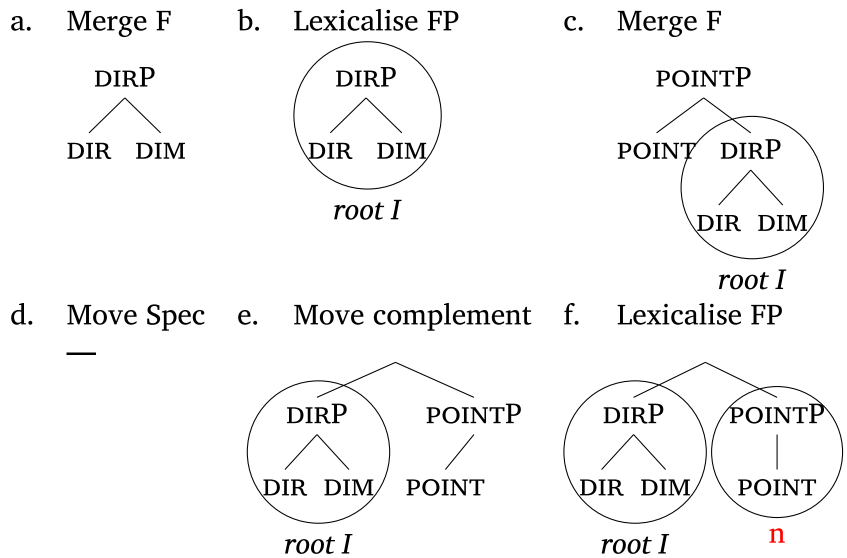
b. C0P ⇔ /n/
 [C0P C0 [POINTP POINT]]

c. C1P ⇔ /ěj/
 [C1P C1]

d. C2P ⇔ /š/
 [C2P C2]

In (50), we present the full stepwise derivation of the positive degree. The derivation starts by merging DIM and DIR together, producing DIRP (50a). After the features have been merged, the resulting phrase must be lexicalised. The DIRP in (50a) can be lexicalised by the root, because the entry (49a) is identical to (50a). Successful lexicalisation is depicted in (50b) by circling the DIRP. After successful lexicalisation, Merge F continues by adding another feature, namely POINT, producing (50c).

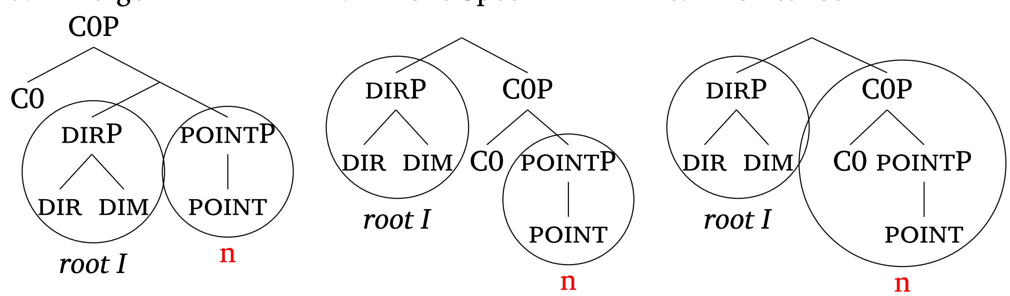
(50)



After POINT is merged, POINTP must be lexicalised. However, the root in (49a) does not contain a constituent identical to (50c), therefore, lexicalisation fails. This means that evacuation movements are triggered following the lexicalisation algorithm (47). The first evacuation movement is the movement of the Spec of the complement of the newly added feature. However, the complement in this case has no Spec, so this step does not lead to any change, and lexicalisation fails again (50d). The next step is complement movement. We start from (50c) and evacuate the complement of POINT out of POINTP. The structure we get is in (50e), where the base position of the extracted phrase is simply left empty, in line with our assumption that evacuation movements do not leave traces. Since a POINTP with a single daughter is contained in the lexical item for n (49b), lexicalisation succeeds, yielding (50f), which represents the correct lexicalisation of the positive degree of the Class I roots.

If we want to derive the comparative form, the derivation must add additional features, starting with C0. The merger of C0 is shown in (51a).

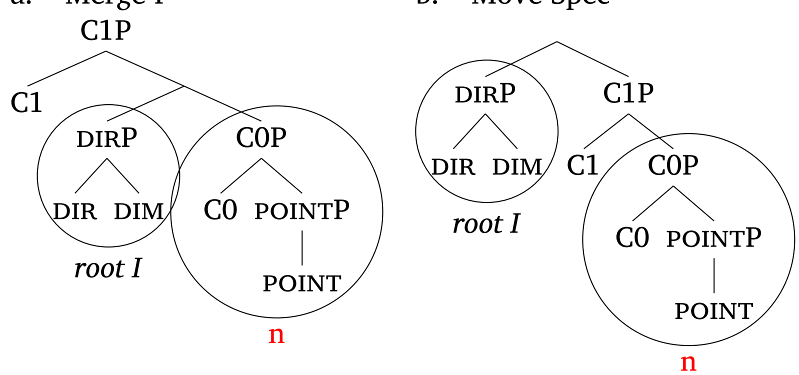
(51) a. Merge F b. Move Spec c. Lexicalise FP



The structure (51a) does not find a direct match in the lexicon, and so evacuation movements have to apply. Moving the Spec of the complement yields the structure (51b). As we can see in (51c), the remnant C0P finds a match in the lexicon, so C0 is lexicalised along with POINT by n, still in perfect compliance with the lexicalisation Table (48).

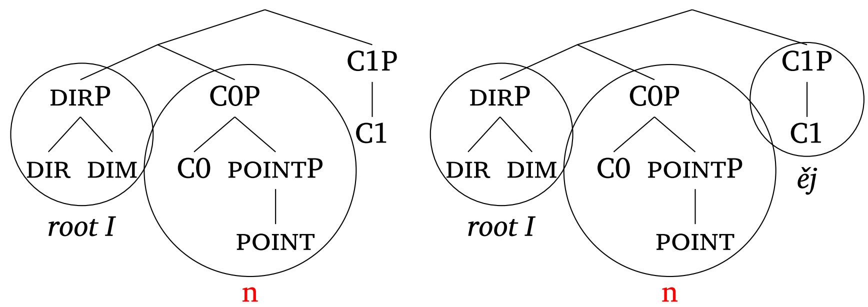
The derivation now continues by merging C1 (52a). As there is no match for C1P in the lexicon, Spec movement is tried, yielding (52b).

(52) a. Merge F b. Move Spec



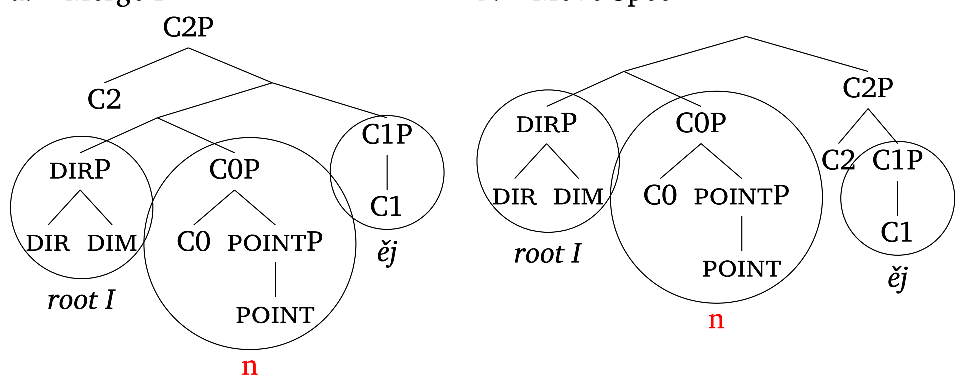
However, there is still no match for C1P, so the derivation goes back to (52a), and moves the whole complement, yielding (53a). In this structure, lexicalisation of the remnant C1P succeeds, and the whole structure is lexicalised as *root-**n**-ěj*, still in accordance with the lexicalisation Table (48).

(53) a. Move complement b. Lexicalise FP



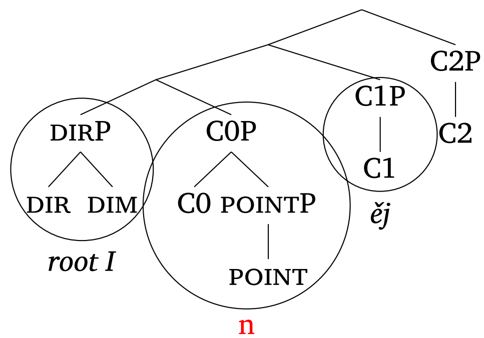
The derivation continues by merging the final feature of the comparative, C2, yielding (54a). This structure cannot be lexicalised, since there is no match for the C2P in (54a). Therefore, Spec movement is tried, yielding (54b).

(54) a. Merge F b. Move Spec

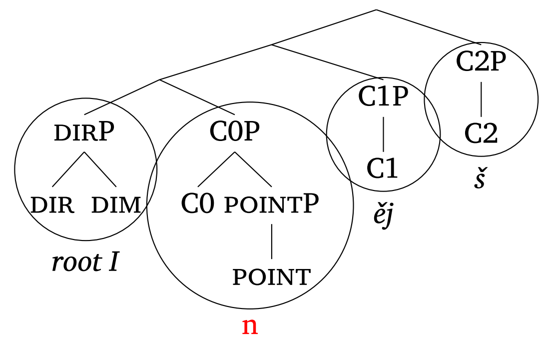


However, even in (54b), there is no match for C2P, containing now only two features, C2 and C1. Therefore, the derivation goes back to (54a) and moves the whole complement of C2, yielding (55a). In this structure, the remanant C2P contains only a single daughter, the feature C2. This C2P can be lexicalised by the lexical item *š* (55b). As a result, we have derived the correct comparative of Class I, namely the root followed by ***n***-*ěj*-*š*, exactly in line with the lexicalisation Table (48).

(55) a. Move complement



b. Lexicalise FP



Concluding, this section has shown how the lexicalisation table in Class I arises from the interaction of the proposed lexical entries and the standard Nanosyntactic lexicalisation algorithm (47).

8.2 Class II

The Class II roots do not require any extra morpheme in the positive, while in the comparative, they required the stem marker n, followed by *ěj-š*. The lexicalisation table is depicted in (56).

(56) Class II

	DIM	DIR	POINT	C0	C1	C2
POS	root II					
CMPR	root II		n		ěj	š

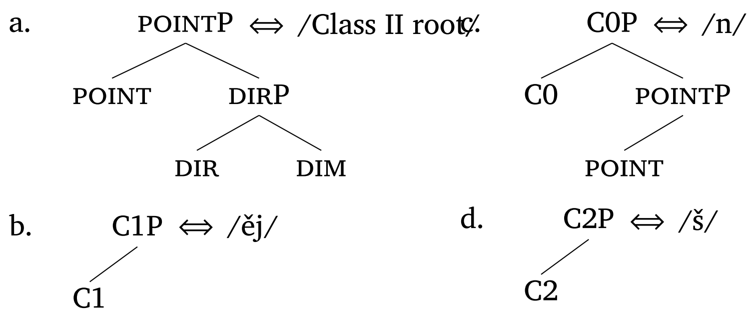
It will become relevant in the course of the discussion that the lexicalisation of the comparative in Class II is the exact same lexicalisation as in the comparative of Class I, compare the bottom row of Table (56) with the Class I table in (57), repeated from (48).

(57) Class I

	DIM	DIR	POINT	C0	C1	C2
POS	root I		n			
CMPR	root I		n		ěj	š

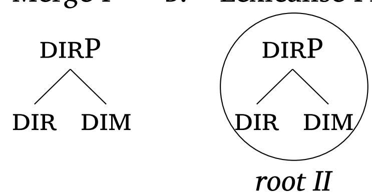
In (58), we give the lexical items relevant for Class II.

(58)



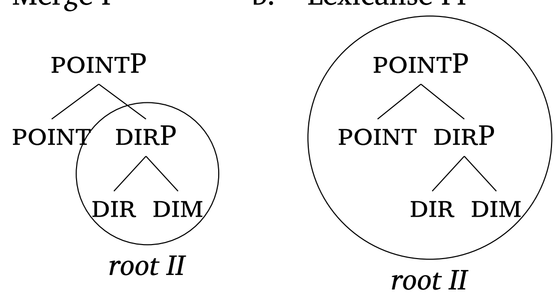
The derivation of the positive degree is quite simple. We merge DIM and DIR (59a), producing a phrase that can be directly lexicalised by the root (59b).

(59) a. Merge F b. Lexicalise FP



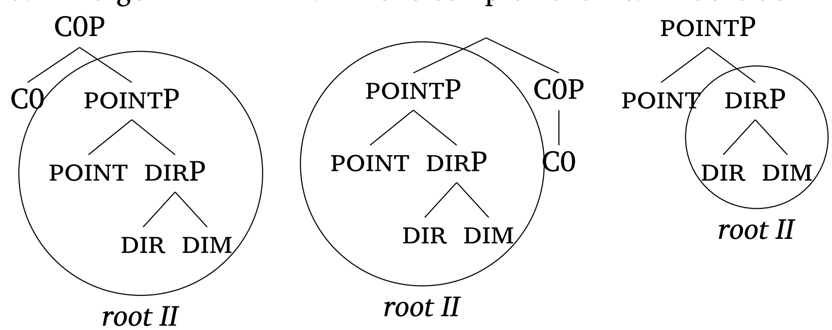
When POINT is added, as in (60a), a direct lexicalisation is possible again (60b). This final tree represents the correct structure for the positive degree in this class, featuring the root and only the root.

(60) a. Merge F b. Lexicalise FP



A slightly more complex derivation arises for the comparative. First, C0 is merged on top of the structure of the positive, yielding (61a). This structure cannot be lexicalised by the root, and evacuation movements are therefore triggered. Since the complement has no Spec, this rescue operation fails. When complement movement is tried, we get the structure in (61b). Here we have a C0P with a single daughter, namely C0. Amongst our entries in (58), C0 is only contained in the entry for n in (59c). However, the entry does not contain a constituent identical to the C0P as it is found in (61b), as the C0P of (59c) itself contains a POINTP, which the C0P of (61b) lacks. As a result, we must conclude that lexicalisation fails in (61b). This is the reason why in the lexicalisation Table (56), the middle row (where n lexicalises just C0) is not a viable option.

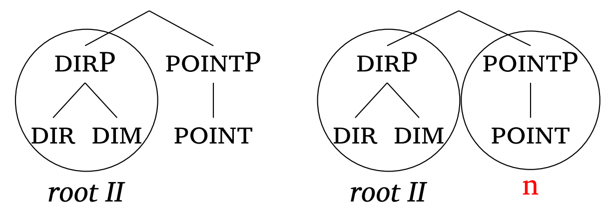
(61) a. Merge F b. Move complement c. Backtrack



Now, in the lexicalisation algorithm (47), complement movement is the last type of evacuation movement. When this movement fails, there are no more movements to be tried. As a last resort, the so-called backtracking step of the algorithm is activated. The idea behind backtracking is this: in our attempts to derive the comparative, we have added the feature C0 on top of a particular structure, and we could not lexicalise it no matter what we tried. Backtracking makes the derivation go back to the previous cycle and lexicalise it differently, changing the original structure. The C0 is added again, but since it is added to a different structure, lexicalisation of C0 has a chance to succeed.

The backtracking step therefore tells us to return to the previous cycle (i.e., back to when POINT was added) and try a different option for that cycle. The stage of the derivation when we added POINT is shown in (60a). Originally, i.e., in (60b), we used lexicalisation without movement. However, since that ultimately failed when we added C0, the algorithm tells us to try some other option. The next option of the algorithm is to move the Spec of the complement, but this option is undefined (62a). Therefore, we go back to (61c) and perform complement movement, yielding (62b). In this structure POINTP can be lexicalised as n (62c).

(62) a. Move Spec — b. Move complement c. Lexicalise FP



Notice that, as we are trying to produce the comparative, the derivation led us back to the positive, and made us reconsider our previous lexicalisation of POINTP: in the positive, it was lexicalised just with the root. Now, due to backtracking, POINTP is lexicalised by the root and n, exactly as on the bottom row of Table (56).

Another point to note is that now POINTP is lexicalised in exactly the same way as with Class I roots, whose derivation was given in (50). Therefore, should we now attempt the derivation of the comparative starting from (62c), we would follow *exactly* the same steps as in Class I, starting from (51) and proceeding all the way to (55). Therefore, the lexicalisation of the comparative in Class II is exactly the same as the lexicalisation of the comparative in Class I (i.e., *root-n-ěj-š*). This is the correct result, as already pointed out at the start of this section.

Concluding, this section has shown how the lexicalisation table in Class II arises from the interaction of the proposed lexical entries and the standard lexicalisation algorithm.

8.3 Class III

The Class III roots are characterised by the fact that they have no n, neither in the positive, nor in the comparative, where the root is directly followed by *ěj-š* (Class IIIa) or by *š*. The lexicalisation tables are depicted in (63) and (64).

(63) Class IIIa

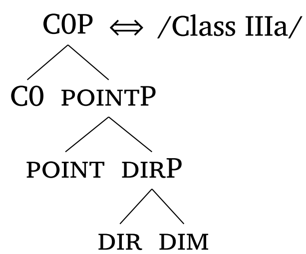
	DIM	DIR	POINT	C0	C1	C2
POS	root IIIa					
CMPR	root IIIa				ěj	š

(64) Class IIIb

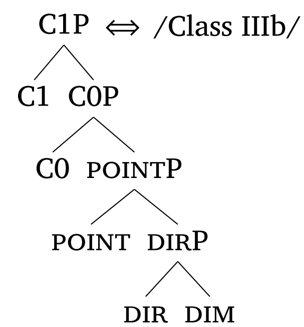
	DIM	DIR	POINT	C0	C1	C2
root (POS)	root IIIb					
root-š (CMPR)	root IIIb					š

The entries for the roots are provided in (65); we do not repeat the functional morphemes (n, *ěj*, *š*) as they remain the same throughout.

(65) a.

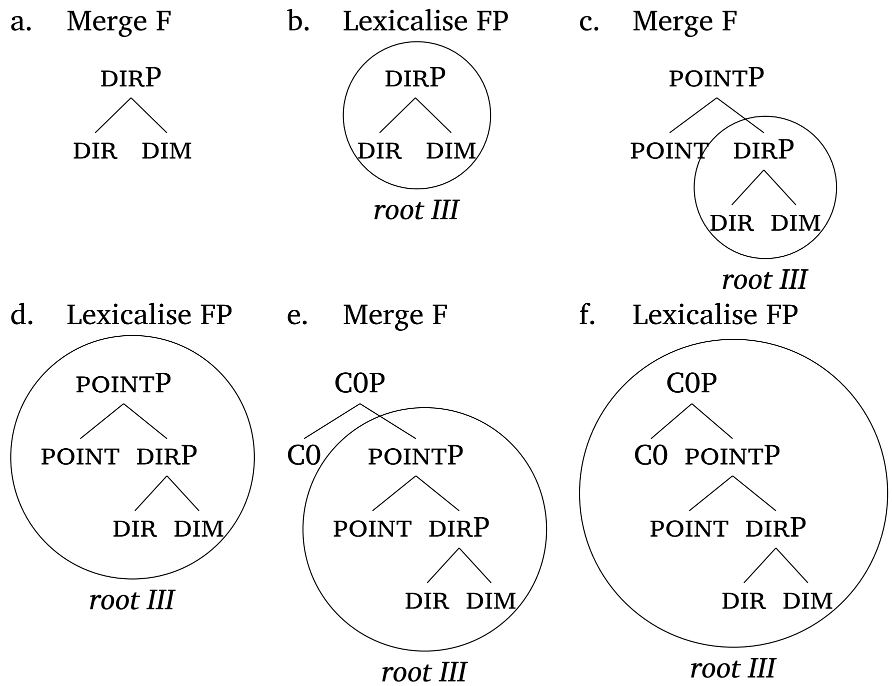


b.



The start of the derivation is the same for both classes of roots. In the derivation of the positive, we only merge features and directly lexicalise, ultimately deriving the positive form in (66d), which is correctly derived as lexicalised by the root only.

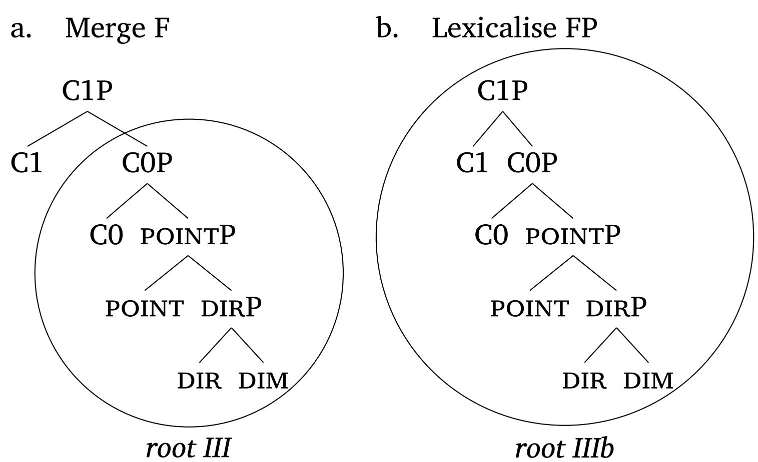
(66)



Deriving the comparative also starts the same for both subclasses of roots: when C0 is added, as in (66e), both types of roots can lexicalise this constituent without any need for movement (66f).

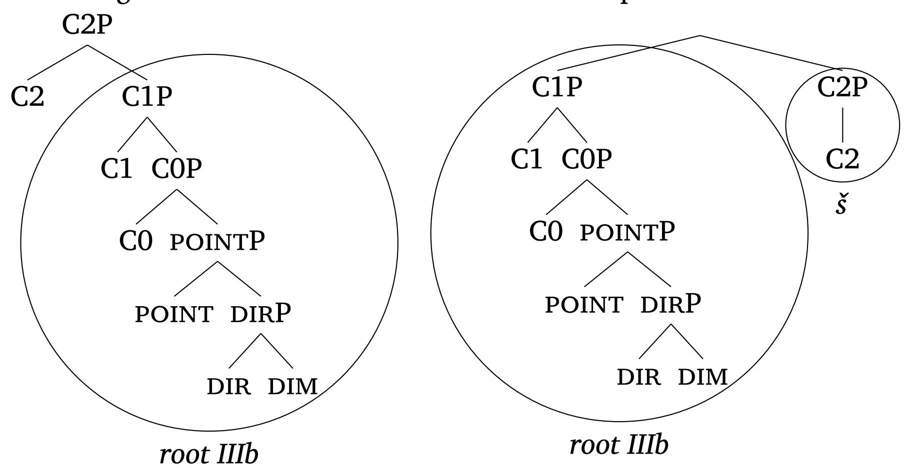
From there, the derivation diverges. When C1 is introduced in the derivation, as in (67a), only Class IIIb roots can lexicalise this constituent directly (67b).

(67)



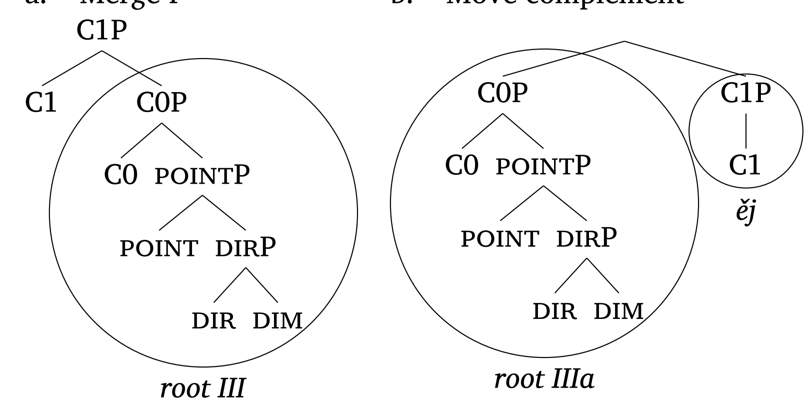
This line of derivation then derives comparatives with just *š*: when C2 is introduced on top of (67b), as in (68a), direct lexicalisation fails. Since the complement has no Spec, we show directly the result of complement movement in (68b), where the complement of C2 (i.e., C1P) evacuates out of C2P. The remnant C2P can then be lexicalised by *š* (68b).

(68) a. Merge F b. Move complement



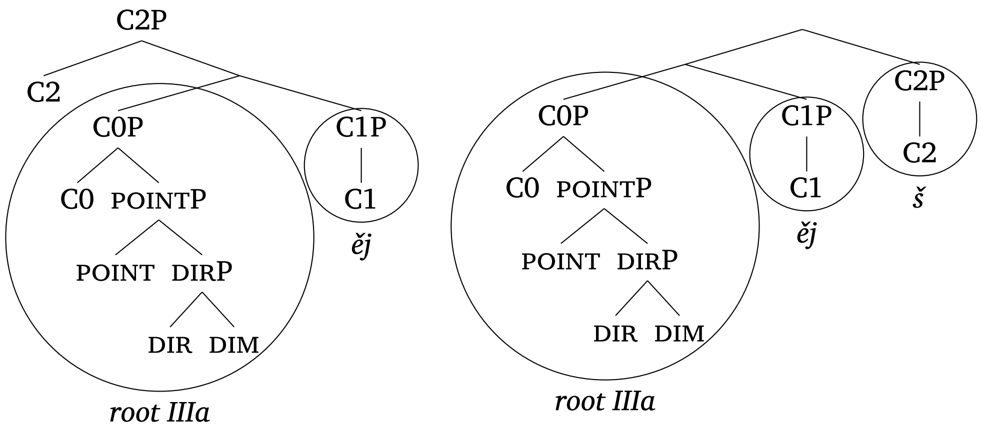
The derivation of the Class IIIa roots is different. With these roots, when C1 is introduced, as in (69a), lexicalisation fails. Therefore, evacuation movements are tried. Since the complement of F2 has no Spec, we show directly the result of complement movement in (69b), along with the successful lexicalisation.

(69) a. Merge F b. Move complement



The derivation then continues by merging C2 (70a). Since direct lexicalisation fails, evacuation movements are tried. Spec movement yields a structure that cannot be lexicalised (not shown), so that complement movement is applied. The structure ensuing from this step of movement is shown in (70b), along with the successful lexicalisation.

(70) a. Merge F b. Move complement



In sum, this section has shown how the lexicalisation tables in Classes IIIa and IIIb arise from the interaction of the proposed lexical entries and the standard lexicalisation algorithm.

8.4 Class IV

The Class IV roots have the stem marker n in the positive, but lose it in the comparative, where the root is directly followed by *ěj-š* (Class IVa) or by *š* (Class IVb). The lexicalisation tables are given in (71) and (72).

(71) Class IVa

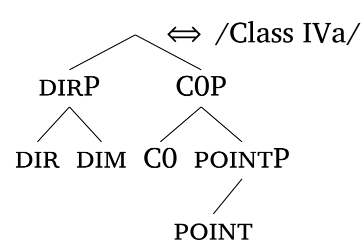
	DIM	DIR	POINT	C0	C1	C2
POS	root IVa		n			
CMPR	root IVa				ěj	š

(72) Class IVb

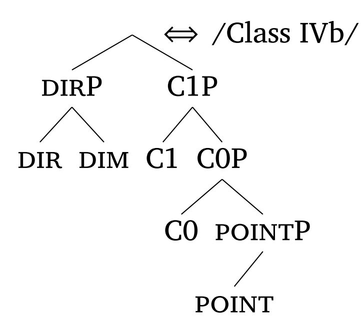
	DIM	DIR	POINT	C0	C1	C2
POS	root IVb		n			
CMPR	root IVb					š

In (44) above, we have already proposed the lexical entries needed to derive the behaviour of this class. We repeat them in (73) for convenience.

(73) a.

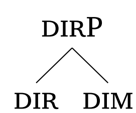


b.

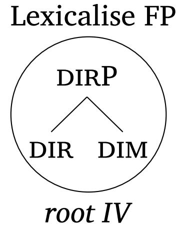


In the remainder of this section, we describe the derivations step by step. The start of the derivation is the same as in all the classes we have discussed so far, with the merger of DIR and DIM (73a). At this point, both types of Class IV roots can lexicalise this constituent (73b).

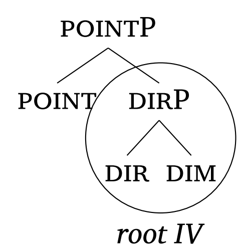
(74) a. Merge F



b. Lexicalise FP



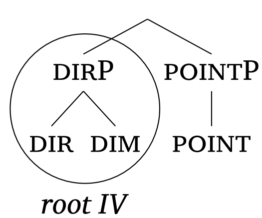
c. Merge F



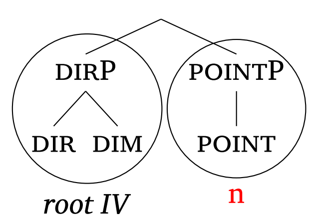
d. Move Spec

—

e. Move complement



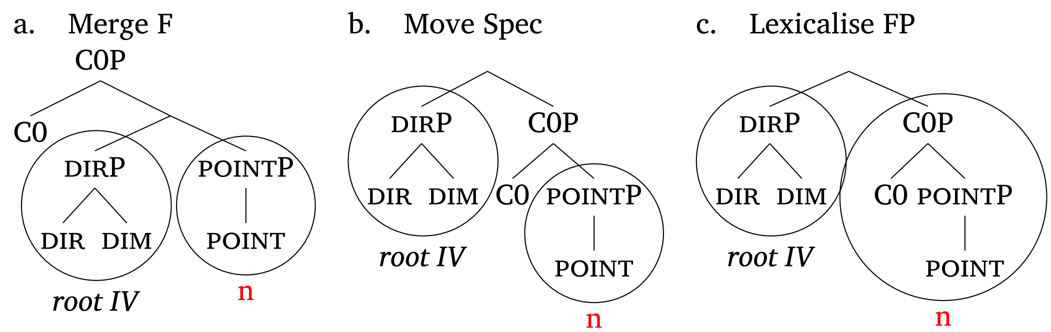
f. Lexicalise FP



When POINT is merged (74c), neither of the lexical entries in (73) contains the relevant constituent, which is why evacuation movements are tried. Spec movement leads to no result, since the complement of POINT in (74c) has no Spec. Complement movement is therefore tried next, yielding (74e). The remnant POINTP is lexicalised by n (74f), and this is the correct form of the positive.

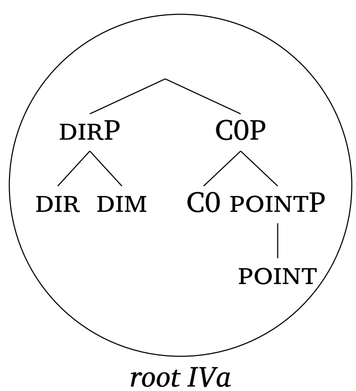
The derivation of the comparative proceeds as follows. First, C0 is merged to (74f), yielding (75a). Since there is no match, Spec movement takes place (75b). As in Class I, the remnant C0P can be lexicalised by n (75c).

(75)



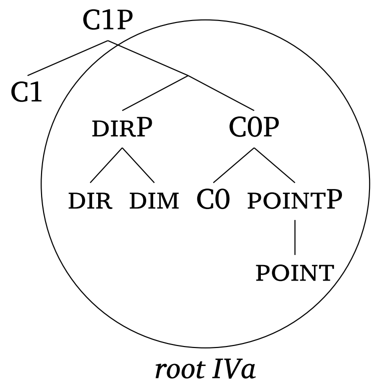
Note at this point that for roots of Class IVa, the structure (75c) is identical to the lexical item corresponding to the root, recall (73a). Thus, for Class IVa roots, the structure (75c) is realised by the root alone, as in (76).

(76)

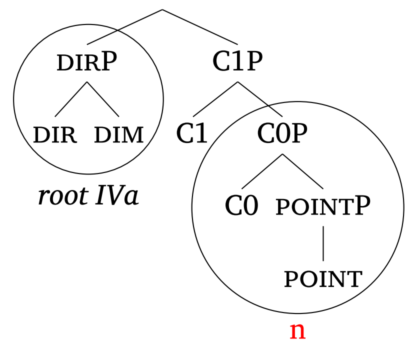


The derivation continues from (76) by merging C1 on top of it, yielding (77a). Since lexicalisation fails, Spec moves, producing (77b).

(77) a. Merge F

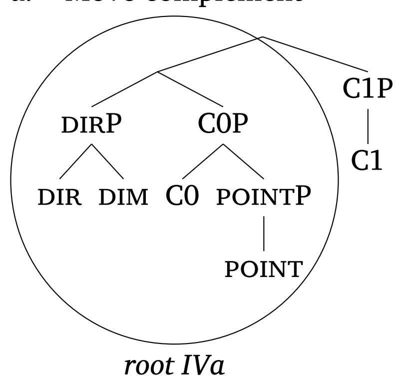


b. Move Spec

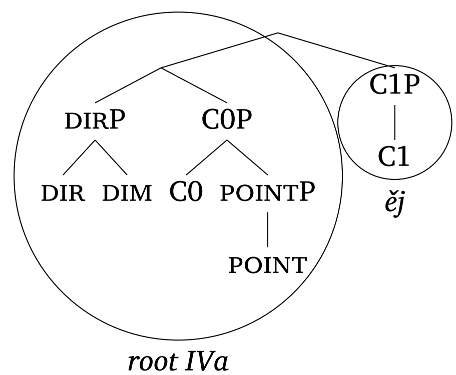


There is no lexical item matching C1P, the derivation therefore goes back to (77a) and moves the complement. We obtain (78a), where the remnant C1P can be lexicalised by *ěj*, see (78b).

(78) a. Move complement

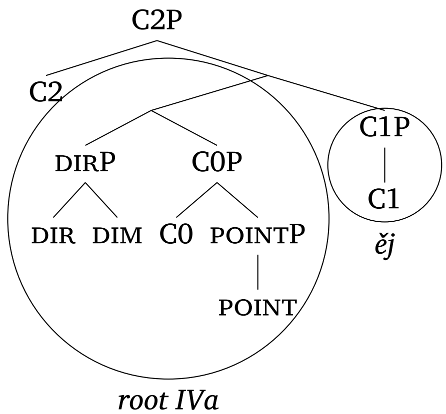


b. Lexicalise FP

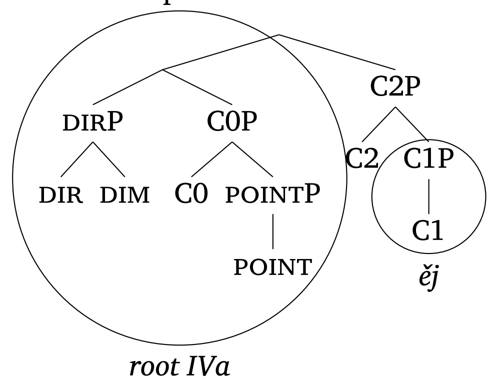


The derivation continues by merging the final feature of the comparative, namely C2, yielding (79a). This structure cannot be lexicalised (there is no match for the C2P in (79a)). Therefore, Spec movement is tried, yielding (79b). However, even in (79b), there is no match for C2P, containing now only two features, C2 and C1.

(79) a. Merge F

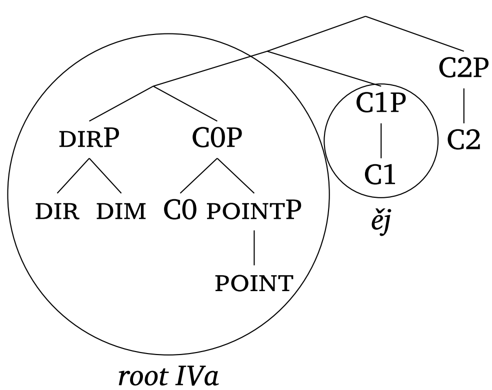


b. Move Spec

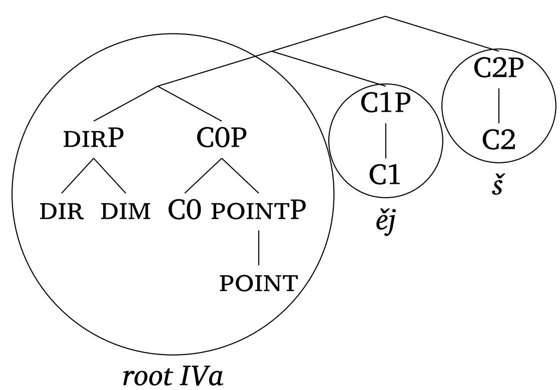


Therefore, the derivation goes back to (79a) and moves the whole complement of C2, yielding (80a). In this structure, the remnant C2P contains only a single daughter, the feature C2. This C2P can be lexicalised by the lexical item inserting *š* (80b).

(80) a. Move complement



b. Lexicalise FP



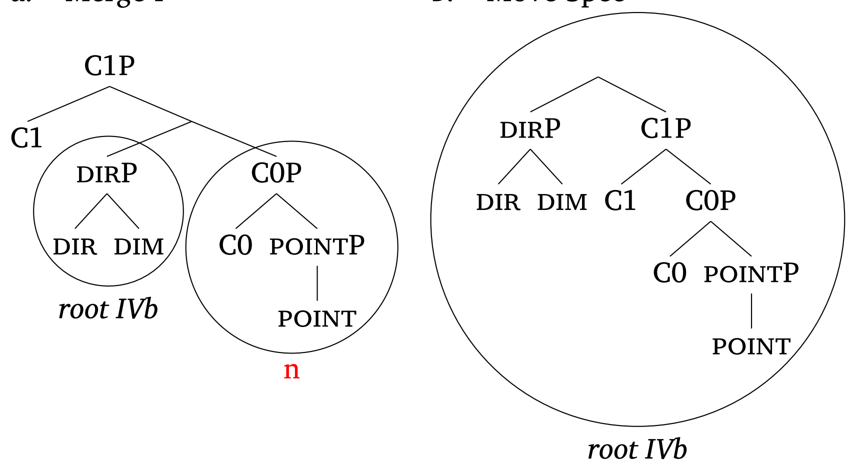
The derivations for the positive and the comparative correspond exactly to what we have in the lexicalisation table (71) above, repeated in (81) below, but with a minor change. To mark the fact that the IVa roots cannot shrink to POINTP size, we have added an asterisk in the POINT column. Even though the roots are of size C0P, they cannot shrink to POINTP because of the shape of their lexical entry. Observe that these roots can shrink to size DIRP or smaller, since constituents of this size are contained in their lexical entry.

(81) Class IVa

	DIM	DIR	POINT	C0	C1	C2
POS	root IVa		n			
CMPR	root IVa		*		ěj	š

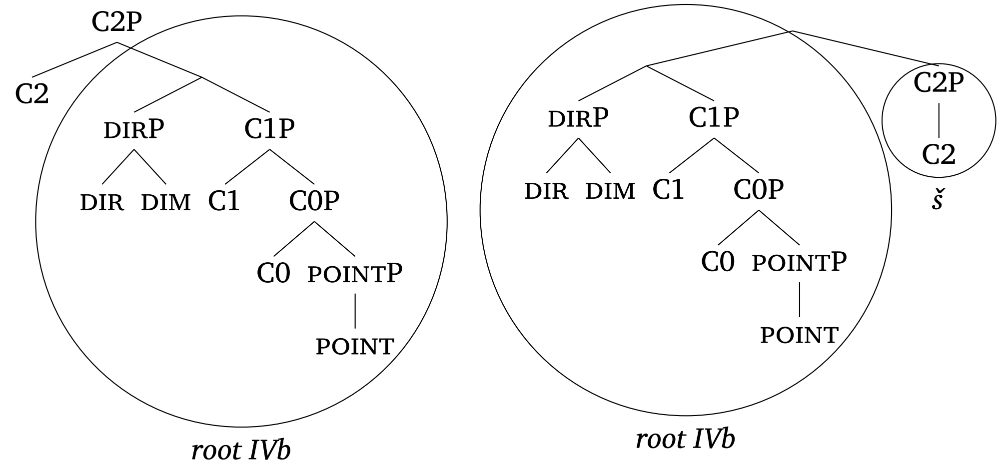
The only thing that remains is to finish the derivation of Class IVb. In order to do so, we go back to the stage of the derivation where Classes IVa and IVb diverge, namely a thte point where C1 is merged. In the case of the Class IVa roots, this required (unsuccessful) spec movement, followed by complement movement, and lexicalisation of the remnant C1P as *ěj*. In the case of the Class IVb roots, however, starting from (82a), Spec movement is going to be successful, since it creates a structure that exactly matches the lexical specification of the Class IVb roots (73b). As a result, the whole tree in (82b) is lexicalised by the root.

(82) a. Merge F b. Move Spec



When C2 is subsequently added (83a), lexicalisation without movement fails. Evacuation movements therefore apply, first Spec movement (not shown, since it fails) and ultimately complement movement, which yields (83b). This structure corresponds to the correct comparative structure of Class IVb, where the root in the comparative is followed by *š* only.

(83) a. Merge F b. Move Complement



For completeness, we repeat the lexicalisation table for this class, again with asterisks added to those cells that represent sizes to which the lexical tree cannot shrink.

(84) Class IVb

	DIM	DIR	POINT	C0	C1	C2
POS	root IVa		n			
CMPR	root IVa		*	*		š

At this point, we have completed the demonstration of the claim that a portmanteau-based approach to allomorphy is capable of encoding the distribution of the stem markers in Czech without the need to introduce any accidental homophony, disjunctive specifications, or negative conditions on insertion.

As a final point in this section, we want to reflect on the general properties of the analysis of Class IV. Recall that this is a special class where the comparative is not based on the positive, violating a proposed candidate universal by Grano & Davis (2018). As has become clear, a crucial aspect of our account is that the lexical entries are not defined purely in terms of how many features they contain, but also how these features are structured. For instance, Class IIIa roots have the same features as Class IVa roots, but they differ in whether they can lexicalise the positive on their own. The difference is in the structure of the features, not in their number.

This presupposes that lexical entries must contain structured constituents, rather than unstructured collections of features, a possibility entertained in Vanden Wyngaerd (2018), as well as in the spanning literature (e.g. Ramchand 2008, Svenonius 2012, 2016, Merchant 2015). However, our results indicate that to deliver all the necessary contrasts (especially between Classes III and IV), structured lexical entries are required.

9 Conclusion

The goal of this paper was to compare two approaches to allomorphy: one based on context-sensitive rules, and the other based on portmanteau lexicalisation. Our general point is that the theories are not equivalent and that the types of allomorph distributions that can be described by the two systems diverge. In particular, we argued that in order to capture the distribution of the stem marker *n* in Czech adjectives, an account based on context-sensitive rules requires postulating multiple homophonous markers and invokes negative conditions on their context. In contrast, the portmanteau-based approach handles these facts without any of these unwanted stipulations. We further provided a detailed demonstration of how this is achieved in the portmanteau-based approach using a standard version of Nanosyntax (Starke 2018).

Along the way, we noted that some of these patterns are interesting in and of themselves, in particular Class IV, which violates a candidate universal proposed in Grano & Davis (2018). The candidate universal says that the comparative degree should always be based on the form of the positive, which is not the case in Class IV, as the stem marker n disappears in the comparative. The system we have proposed allows this pattern to be captured without giving up on the idea, argued convincingly for in Bobaljik (2012), that the structure of the positive is contained in the comparative.

References

Banerjee, Neil. 2021a. *On the interaction of portmanteaux and ellipsis*. Boston: MIT dissertation.

Banerjee, Neil. 2021b. Two ways to form a portmanteau: Evidence from ellipsis. In Patrick Farell (ed.), *Proceedings of the Linguistic Society of America*, vol. 6 1. 39–52.

Baunaz, Lena & Eric Lander. 2018. Nanosyntax: the basics. In Lena Baunaz, Karen De Clercq, Liliane Haegeman & Eric Lander (eds.), *Exploring nanosyntax*. 3–56. Oxford: Oxford University Press.

Blix, Hagen. 2022. Interface legibility and nominal classification: A nanosyntactic account of Kipsigis singulatives. *Glossa* 7(1). 1–46.

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

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

Caha, Pavel. To appear. Nanosyntax: some key features. In Artemis Alexiadou, Ruth Kramer, Alexc Marantz & Isabel Oltra-Massuet (eds.), *The Cambridge handbook of distributed morphology*. Cambridge: Cambridge University Press.

Caha, Pavel, Karen De Clercq, Michal Starke & Guido Vanden Wyngaerd. To appear. Nanosyntax: state of the art and recent developments. In Pavel Caha, Karen De Clercq & Guido Vanden Wyngaerd (eds.), *Nanosyntax and the lexicalisation algorithm*. Oxford: Oxford University Press.

Caha, Pavel, Karen De Clercq & Guido Vanden Wyngaerd. 2019. The fine structure of the comparative. *Studia Linguistica* 73(3). 470–521. doi:https://doi.org/10.1111/stul.12107.

Choi, Jaehoon & Heidi Harley. 2019. Locality domains and morphological rules. Phases, heads, node-sprouting and suppletion in Korean honorification. *Natural Language & Linguistic Theory* 37. 1319–1365.

Christopoulos, Christos & Stanislao Zompí. 2023. Taking the nominative (back) out of the accusative. *Natural Language & Linguistic Theory* 41. 879–909.

Chung, Inkie. 2007. Suppletive negation in Korean and distributed morphology. *Lingua* 117. 95–148.

Cortiula, Maria. 2023. *The nanosyntax of Friulian verbs. An analysis of the present and past in Tualis Friulian*. Brno: Masaryk University dissertation.

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. 2019. Negation and the

functional sequence. *Natural Language & Linguistic Theory* 37(2). 425–460.
Embick, David & Alec Marantz. 2008. Architecture and blocking. *Linguistic Inquiry* 39(1). 1–53.
Grano, Thomas & Stuart Davis. 2018. Universal markedness in gradable adjectives revisited. The morpho-semantics of the positive form in Arabic. *Natural Language & Linguistic Theory* 36(1). 131–147.
Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Ken Hale & Jay Keyser (eds.), *The view from building 20*. 111–176. Cambridge, MA: MIT Press.
Janků, Lucie. 2022. *The nanosyntax of Czech nominal declension*. Brno: Masaryk University dissertation.
Kennedy, Christopher. 1999. *Projecting the adjective: The syntax and semantics of gradability and comparison*. New York: Garland.
Kennedy, Christopher & Beth Levin. 2008. Measure of change: The adjectival core of degree achievements. In Louise McNally & Christopher Kennedy (eds.), *Adjectives and adverbs: Syntax, semantics, and discourse*. 156–182. Oxford: Oxford University Press.
Kennedy, Christopher & Louise McNally. 2005. Scale structure, degree modification, and the semantics of gradable predicates. *Language* 81. 345–381.
Křivan, Jan. 2012. Komparativ v korpusu: explanace morfematické struktury českého stupňování na základě frekvence tvarů [Comparative in the corpus: explanation of the morphematic structure of Czech gradation based on the frequency of forms]. *Slovo a slovesnost* 73. 13–45.
McCawley, James. 1968. Lexical insertion in a transformational grammar without deep structure. *CLS* 4. 71–80.
Merchant, Jason. 2015. How much context is enough? Two cases of span-conditioned stem allomorphy. *Linguistic Inquiry* 46. 273–303.
Moskal, Beata & Peter W. Smith. 2016. Towards a theory without adjacency: hyper-contextual VI-rules. *Morphology* 26(3). 295–312. doi:10.1007/s11525-015-9275-y.
Neeleman, Ad & Kriszta Szendroi. 2007. Radical pro drop and the morphology of pronouns. *Linguistic Inquiry* 38(4). 671–714.
Radkevich, Nina. 2010. *On location: The structure of case and adpositions*. Storrs, CT: University of Connecticut dissertation.
Ramchand, Gillian. 2008. *Verb meaning and the lexicon*. Cambridge: Cambridge University Press.

Siegel, Dorothy. 1977. The adjacency constraint and the theory of morphology. *Proceedings of NELS* 8. 189–197.

Smith, Peter W., Beata Moskal, Ting Xu, Jungmin Kang & Jonathan David Bobaljik. 2019. Case and number suppletion in pronouns. *Natural Language & Linguistic Theory* 37. 1029–1101.

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

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

Starke, Michal. 2018. Complex left branches, spellout, and prefixes. In Lena Baunaz, Karen De Clercq, Liliane Haegeman & Eric Lander (eds.), *Exploring nanosyntax*. 239–249. Oxford: Oxford University Press.

Svenonius, Peter. 2012. Spanning. Ms., CASTL.

Svenonius, Peter. 2016. Spans and words. In Daniel Siddiqi & Heidi Harley (eds.), *Morphological metatheory*. 201–222. Amsterdam: John Benjamins.

Vanden Wyngaerd, Guido. 2018. The feature structure of pronouns: a probe into multidimensional paradigms. In Lena Baunaz, Karen De Clercq, Liliane Haegeman & Eric Lander (eds.), *Exploring nanosyntax*. 277–304. Oxford: Oxford University Press.

Vanden Wyngaerd, Guido, Michal Starke, Karen De Clercq & Pavel Caha. 2020. How to be positive. *Glossa* 5(1). 23. 1–34.

Wasow, Thomas. 1977. Transformations and the lexicon. In Peter Culicover, Thomas Wasow & Adrian Akmajian (eds.), *Formal syntax*. New York: Academic Press.

Wiland, Bartosz. 2019. *The spell-out algorithm and lexicalization patterns. Slavic verbs and complementizers*. Berlin: Language Science Press.

Williams, Edwin. 2003. *Representation theory*. Cambridge, MA: MIT Press.