

## The split-comparative hypothesis and the morphology of comparative adverbs in Czech

PAVEL CAHA<sup>1</sup>, KAREN DE CLERCQ<sup>2</sup>, & GUIDO VANDEN WYNGAERD<sup>3</sup>

<sup>1</sup> *Masarykova Univerzita, Brno*

<sup>2</sup> *CNRS/LLF/Université Paris Cité*

<sup>3</sup> *KU Leuven, Brussels*

ABSTRACT

This paper investigates complex allomorphy patterns in comparative adjectives and adverbs in Standard Czech and two of its dialects. We formulate an account of the allomorphy (including microparametric variation) within the Nanosyntax framework (Starke 2018). We argue that in order to capture the facts, we need to adopt two proposals. The first one is a split CMPR hypothesis (Caha et al. 2019), according to which the comparative head splits into two independent projections, C1 and C2. The second idea is a split ADV hypothesis (hinted at in Bobaljik 2012), according to which there are two positions for an Adv head in the functional sequence.

KEYWORDS adjectives · adverbs · allomorphy · comparative · nanosyntax

### 1 INTRODUCTION

This article explores the morphosyntactic relationship between comparative adjectives and comparative adverbs in Standard Czech and two Moravian dialects (East and Central). The forms under investigation exhibit complex patterns of allomorphy, where the shape of the comparative marker is influenced by two factors. The first factor concerns the (arbitrary) class of the root: the comparative degree is expressed differently for different roots. The second factor is part of speech: comparative marking differs for adjectives and adverbs of the same root.

We argue that in order to capture these patterns, we need to adopt two proposals. The first one is the Split-Comparative Hypothesis (Caha et al. 2019). The hypothesis is depicted in (1).

- (1) The Split Comparative Hypothesis:  
 comparative = [[[positive] C1 ] C2]

The structure in (1) follows the standard idea that the comparative is derived from the positive (see, e.g., Bobaljik 2012, Grano & Davis 2018). The new thing about it is that the comparative meaning is not added as a single structural component, but is distributed across two different syntactic heads.

The second ingredient we need is a proposal for adverbs. In Bobaljik's (2012) seminal work, it was proposed that comparative adverbs can be formed in two different ways. In some languages, comparative adverbs are derived from positive-degree adverbs, as in (2-a). We call the relevant adverbial head LOADV, since it comes below the comparative. In other languages, comparative adverbs are derived from comparative adjectives, i.e., by turning a comparative adjective into an adverb, as in (2-b). We call the adverbial head HIADV, since it is found above the comparative. The two different structures in (2-a,b)

are functionally equivalent.

- (2) Two ways of forming a comparative adverb (Bobaljik 2012)
- a. [[[ A ] LOADV ] CMPR ]
  - b. [[[ A ] CMPR ] HIADV ]

The current paper proposes that the complex patterns of allomorphy in Czech can be explained under the idea that it has both the low and the high ADV head. We call this the split adverb hypothesis. We depict the idea in a simplified form in (3), noting that in the final version of our proposal, the CMPR head will have to be split into C1 and C2.

- (3) Our proposal (simplified version)
- [[[[ A ] LOADV ] CMPR ] HIADV ]

The paper is organised as follows. §2 discusses the formation of comparative adjectives in Czech, focussing on the fact that the shape of the comparative marker is influenced by the root. §2 also presents the account by Caha et al. (2019), which has been designed to explain how exactly the root determines the form of the suffix. In §3, we turn to the fact that the comparative marker is influenced not only by the root, but also by the adjective/adverb distinction. These patterns are rather complex and, as we argue, impossible to derive in the system presented in Caha et al. (2019), if only a single adverbial head is assumed. In §4, we show how these facts can be explained under the hypothesis that Czech has both a low and a high adverbial head. §5 discusses adverbs in the dialects. §6 concludes.

## 2 COMPARATIVE ADJECTIVES

In this section, we discuss comparative adjectives in Standard Czech. In doing so, we also provide the account of the allomorphic variation proposed in Caha et al. (2019).

The first thing to know is that Standard Czech comparatives are formed by two allomorphs, namely *ějš* or *š*. The allomorph *ějš* is illustrated in (4-a), *š* is in (4-b).<sup>1</sup>

- (4) Two allomorphs of the comparative in Standard Czech

	POS ADJ	CMPR ADJ	GLOSS
a.	chab -ý	chab - <b>ějš</b> -í	'weak'
b.	slab -ý	slab - <b>š</b> -í	'weak'
	root-AGR	root-CMPR-AGR	

The two adjectives in (4) are similar both in their phonology and their meaning. This indicates that the use of one or the other marker is not uniquely determined by the phonology or the meaning of the base. The decision which roots require *š* and which *ějš* is therefore an idiosyncratic property of the root.<sup>2</sup>

We illustrate the 'arbitrary' nature of allomorph selection in more detail in table (5). In this table, we put side by side adjectival roots that have similar phonology, yet they differ in their comparative allomorphs. In the table, the root on the left (e.g. *star* 'old' in the first row) always has similar phonology as the root on the right (e.g. *čir* 'pure'),

<sup>1</sup>There is a small class of adjectives that have a complex base ending in *k*, e.g., *leh-k-ý* 'easy', which is traditionally described as having a zero allomorph in the comparative (*leh-č-í* 'easier'). Such adjectives either lose the *k* in the comparative, in which case they have *š*, e.g., *tež-k-ý* ~ *těž-š-í* 'heavy ~ heavier', or they can preserve the *k*, as in the case of *leh-č-í* 'easier'. The latter class can be interpreted as having a zero-derived comparative, with the final agreement marker triggering palatalisation of the *k* suffix (*leh-k-í* → *leh-č-í*). However, as pointed out in Caha et al. (2019), this analysis is moot, since an underlying sequence of morphemes *leh-k-š-í* would also be realised as *leh-č-í* in Czech, as a result of palatalisation and degemination. We set these adjectives aside here.

<sup>2</sup>Křivan (2012) concludes that frequency plays a role, with more frequent roots showing preference for *š*, as opposed to *ějš*.

but the allomorphs differ (see Vyshnevskaya 2022 for similar observations concerning Ukrainian).

(5)

	POS	CMPR	GLOSS	POS	CMPR	GLOSS
a.	star-ý	star-š-í	‘old’	čir-ý	čir-ějš-í	‘pure’
b.	bohat-ý	bohat-š-í	‘rich’	svat-ý	svat-ějš-í	‘saint’
c.	tvrđ-ý	tvrđ-š-í	‘hard’	hrđ-ý	hrđ-ějš-í	‘proud’
d.	drah-ý	draž-š-í	‘expensive’	uboh-ý	ubož-ějš-í	‘poor’

It can be observed that each of the suffixes triggers particular consonant mutations at the end of the base. The comparative marker *š* palatalises velars (see (5-d)). The suffix *ějš* palatalises velars, dentals, alveolars and labials. Palatalisation of the preceding consonant is sometimes marked directly on the consonant (e.g. *ř* in (5-a) and *ž* in (5-d)). Other times, it is marked on the vowel (as *ě*), yet the effect is the same. For instance, the comparative *svat-ěj-š-í* on line (5-b) contains a palatalised root (IPA: svacejʃiː).

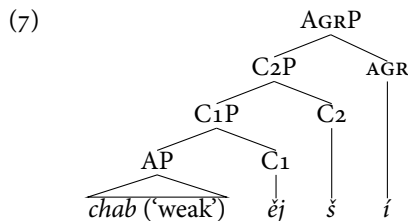
Let us now turn to the form of the two allomorphs, *ějš* and *š*. It has been argued in Caha (2017), Caha et al. (2019) that the comparative marker *ějš* should be decomposed into two independent morphemes, namely *ěj* and *š*. The first reason is that this explains why all comparative adjectives end in *š*, because after the decomposition, *ěj-š*, all comparatives contain the same morpheme, namely *š*. The second reason for decomposition is that each of the two morphemes can surface independently. We already saw this in (5), where *š* can appear without *ěj*. Conversely, *ěj* surfaces without *š* in Standard Czech comparative adverbs, see (6).

(6) Adverbs in Standard Czech

CMPR ADJ	CMPR ADV	GLOSS
chab-ěj-š-í	chab-ěj-i	‘weak’
rychl-ěj-š-í	rychl-ěj-i	‘fast’
červen-ěj-š-í	červen-ěj-i	‘red’
hloup-ěj-š-í	hloup-ěj-i	‘stupid’

These facts support the idea that *ěj* and *š* are independent morphemes.

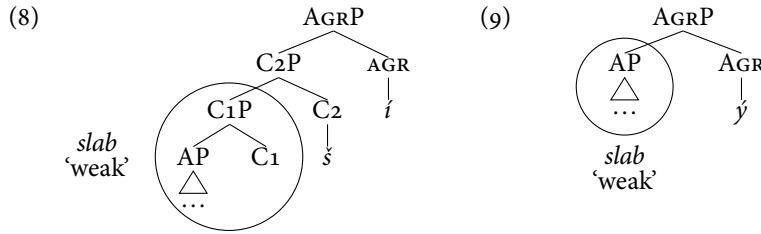
Caha et al. (2019) propose that the morphological complexity of the comparative marking indicates that the comparative meaning is distributed across (at least) two different syntactic positions. Caha et al. (2019) call these C1 and C2, as in (7). The morphemes *ěj* and *š* are the respective realizations of these heads.<sup>3</sup>



Let us now turn to the class of adjectives that only have *š* in the comparative. Building on Starke (2018), Caha et al. (2019) propose that lexical items (including roots) may lexicalise full phrases containing more than a single terminal node. Specifically, roots that combine with *š* are proposed to lexicalise not only the adjectival core of the meaning

<sup>3</sup>In the semantics literature, a complex meaning for the comparative has been proposed, for instance, in Kennedy & Levin (2008). In their approach, the adjective base corresponds to a measure function, from which the comparative is derived in two steps. First, a derived measure function is constructed, followed by the application of a POS head that turns the derived measure function into a property.

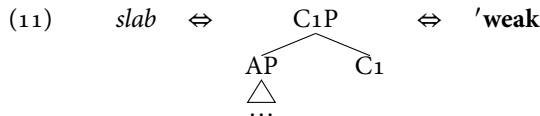
(that is, AP), but also C1. This is depicted in (8), where the root *slab* ‘weak’ lexicalises C1P, which contains C1 and AP. As a result, C1 is realised by the root, and the fact that the suffix *ěj* is absent with these roots follows automatically.



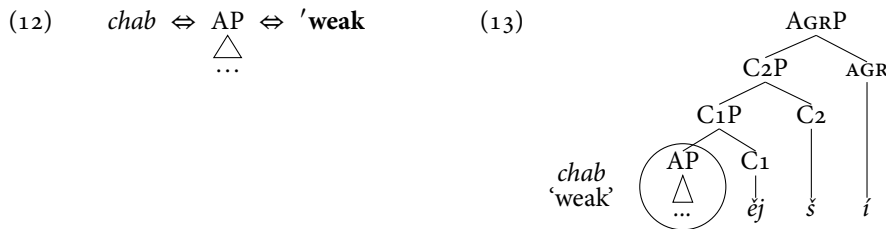
The proposal in (8) is compatible with the fact that the positive degree is realised by the same root *slab* ‘weak,’ as in (9). This is because of the fact that in the Nanosyntax approach, insertion needs to satisfy the Superset Principle:

- (10) *The Superset Principle* (Starke 2009)  
 A lexically stored tree matches a syntactic node iff the lexically stored tree contains the syntactic node.

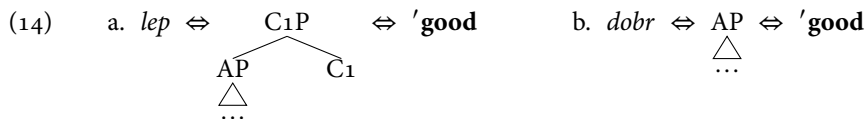
To see how this principle works, consider the lexical entry of the adjective *slab* ‘weak’ in (11). The entry links the phonology *slab* (on the left) with a syntactic tree (in the middle), and a concept (on the right):



The syntactic tree contained in the lexical item (11) is of size C1P, and it contains the AP as a subpart. It can therefore lexicalise either C1P, as shown in (8), or AP as in (9). If a root is lexically specified as just an AP, as in the entry for *chab* ‘weak’ (see (12)), it cannot realise C1P, because it does not contain C1P. As a result, C1 must be realised by *ěj*, see (13).



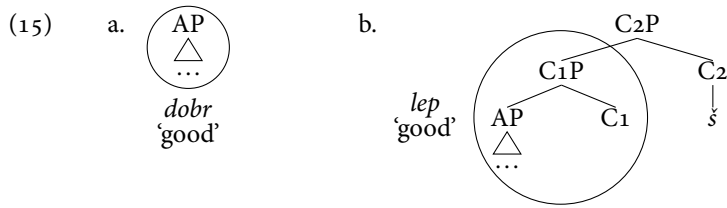
With suppletive adjectives such as *dobr-ý – lep-š-i* ‘good – better,’ this approach allows us to associate each root to a different syntactic tree as in (14). We assume that the adjectives are suppletive by virtue of being associated with the same concept.<sup>4</sup>



These lexical entries will give rise to the lexicalisation of the positive and the comparative as shown in (15), with the root *dobr* ‘good’ in the positive, and the root *lep* ‘bett’ in the

<sup>4</sup>Caha et al. (2019) use pointers as a way of linking two suppletive roots to each other. We do not discuss pointers here for reasons of space.

comparative.



The take-home message is that Caha et al. (2019) use root size to model the type of allomorph that each root combines with. This approach also provides a way to deal with root suppletion in an elegant way, since C1P-sized roots lexicalise different features in the positive and in the comparative. The analysis of suppletion correctly predicts that all suppletive adjectives will be in the *š-i* class (Caha et al. 2019). With this background, let us now turn to the discussion of comparative adverbs.

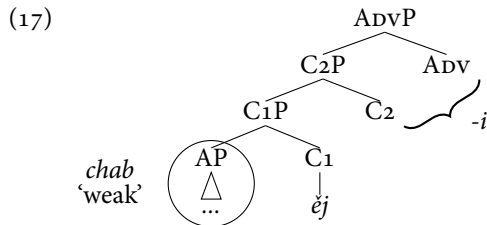
### 3 COMPARATIVE ADVERBS IN STANDARD CZECH

This section shows how adverb formation influences the marking of the comparative. Table (16) gives the standard forms of comparative adverbs corresponding to *ěj-š-i* adjectives. The table is repeated from (6) above, but with the positive forms added, for reasons to be made clear below.

(16) Adverbs of *ěj-š-i* adjectives in Standard Czech, positive included

CMPR ADJ	CMPR ADV	GLOSS	POS ADJ	POS ADV
chab-ěj-š-i	chab-ěj-i	'weak'	chab-ý	chab-ě
rychl-ěj-š-i	rychl-ěj-i	'fast'	rychl-ý	rychl-e
hloup-ěj-š-i	hloup-ěj-i	'stupid'	hloup-ý	hloup-ě

The table illustrates the fact that all the *ěj-š-i* adjectives form adverbs in *ěj-i*. Caha et al. (2019) propose an account where the structure of the adverb has both C1 and C2 (despite the absence of *š*). In addition, there is an ADV head above C2, see (17).



The reason why *š* is not present in (16) is because it is blocked by the adverbial ending *i*, which is a portmanteau that spells out the C2 projection along with the high ADV head. This is indicated by the brace in (17).

The Superset Principle implies that lexicalisation is subject to a constituency requirement. In (17), the C2 and ADV heads do not form a constituent which excludes the other heads. The reason for this is that we depict lexicalisations in an informal way using a brace. In actual fact, we assume that lexicalisation is driven by the Spellout Algorithm (Starke 2018). Application of this algorithm would require C1P to move from within the AdvP, so that C2 and Adv form a constituent. For reasons of space, we ignore these technical details and adopt the brace notation throughout. At the same time, the way we establish which lexical item lexicalises which head is fully compatible with such an algorithm and, in fact, relies on its existence.

The idea that the ending *i* is a portmanteau for C2 and Adv is supported by the fact that adverbs in the positive never have *i*; they have the adverbial ending *ě* instead. This

ending is different because it does not lexicalise C2. This can be seen in Table (16) in the final column.

The theory proposed in Caha et al. (2019) thus successfully deals with adverbs corresponding to *ěj-š-í* adjectival comparatives, using a structure with a high Adv head.

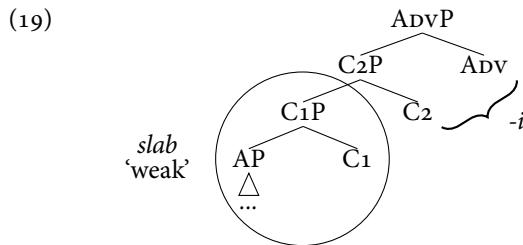
Let us now turn to the *š-í* class of comparatives. The comparative adverbs of the relevant adjectives are of two types, with the type determined by the root. The first (most common) type of adverb is shown in Table (18). The adverb is morphologically identical to the adverbs formed from *ěj-š-í* comparatives. The space between the root and *-š* in the first column indicates the absence of *-ěj*.

(18) Standard Czech comparative adverbs (*š-í* adjectives, Class 1)

CMPR ADJ	CMPR ADV	GLOSS	POS ADJ	POS ADV
slab -š-í	slab-ěj-i	'weak'	slab-ý	slab-ě
tuž -š-í	tuž-ej-i	'tough'	tuh-ý	tuz-e
tvrd -š-í	tvrd-ěj-i	'hard'	tvrd-ý	tvrd-ě

The most interesting fact is that the morpheme *ěj*, missing in the comparative adjective, is found in the adverb. This shows that its presence/absence is sensitive both inwards (to the identity of the root), and outwards (to the final morpheme of the comparative). Root-sensitivity must be invoked because the comparative *ěj* may be absent with only a subset of roots. Outwards sensitivity is revealed by the fact that when the comparatives in (18) are used as adverbs, *ěj* must appear again.

This pattern of distribution is not captured by the proposal in Caha et al. (2019). The prediction of their analysis is shown in (19). In this tree, the root realises [AP, C1] (recall the adjectival comparative in (8)). This leaves [C2, ADV] to be realised by the ending. These two features are predicted to be realised as *i*, precisely as in (17), wrongly predicting forms like \**slab-i* for the comparative adverb.



What is puzzling for this analysis is the fact that *ěj* appears in the adverbs corresponding to *š-í* adjectives. In the system proposed by Caha et al. (2019), it is expected that whenever a particular head can be spelled out by the root, then this type of realisation must be preferred to other ways of realizing it. This is ultimately the reason why comparative adjectives like *slab-š-í* 'weaker' lack *ěj*: since the root can spell out C1, *ěj* is not needed, recall (8). Therefore, the class of adverbs in (18) is problematic because *ěj* surfaces in the adverb with roots that can lexicalise C1.

The second class of comparative adverbs corresponding to *š-í* comparatives is shown in Table (20). This is a minor class, populated mainly by suppletive adjectival comparatives.

(20) Standard Czech comparative adverbs (*š-í* adjectives, Class II)

CMPR ADJ	CMPR ADV	GLOSS	POS ADJ	POS ADV
draž -š-í	draž -e	'expensive'	drah-ý	draž-e
lep -š-í	lep -e	'good'	dobr-ý	dobř-e
del -š-í	dél -e	'long'	dlouh-ý	dlouz-e

These adverbs come close to the predicted form, because they lack *ěj* just like the corresponding adjectives. With these roots, it appears that C1 is spelled out by the root both in the adjective and in the adverb.

However, the adverbs also show two unexpected features. First of all, the adverbial marker is not the predicted *i*, but *e*. This may initially look like the adverbial marker that we find in the positive degree, given in the final column in (20). However, we assume that these are different morphemes, since they interact differently with the final consonant of the root. This can be seen with the root *drah* ‘expensive’, where we get *draz* in the positive and *draž* in the comparative. It is known that in the Czech phonology, different endings trigger different palatalisations of the base, and since the palatalisations are different here, these are quite likely to be different *e*’s.

Another noteworthy fact is that the root always lengthens in the comparative adverb in (20), e.g., the adjectival root *drah* ‘expensive’ is realised as *draž* in the adverb (the accent on the vowel indicates length). Such lengthening looks like something specific to this class of comparative adverbs, and we shall capture it by proposing that the adverbial marker *e* is accompanied by an empty vocalic space  $\mu$ , which triggers the lengthening of the root.

To summarise, Caha et al. (2019) posited two classes of roots, yielding two different types of comparative adjectives: *š-í* vs. *ěj-š-í*. When comparative adverbs are introduced, it turns out that there are three classes of roots. This is because the *š-í* class of adjectives splits into two subclasses, where each subclass has a different way of forming the adverb, as we show in Figure 1, where the three arrows correspond to the three classes of roots. The notation  $\text{PAL}+\mu+e$  represents the fact that comparative adverb suffix *e* triggers palatalisation (PAL) and lengthens the preceding vowel ( $\mu$ ).

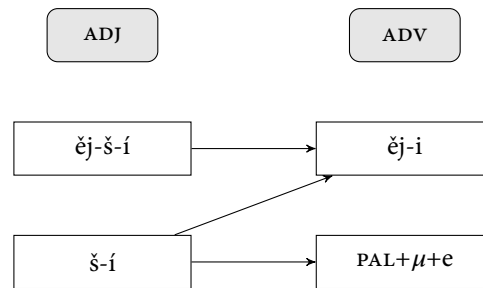


Figure 1: Adjective-adverb pairs in Standard Czech

The three classes are shown in a different format, with actual forms, in Table (21).

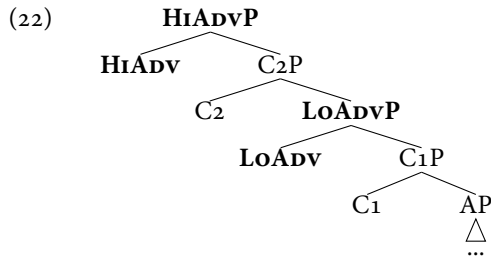
(21) Three classes of comparative ADJs/ADVs in Standard Czech

CMPR ADJ	CMPR ADV	GLOSS
chab-ěj-š-í	chab-ěj-i	‘weak’
slab-š-í	slab-ěj-i	‘weak’
draž-š-í	draž-e	‘expensive’

We can see that the system is descriptively simple (though analytically challenging, as we already mentioned). There are roots which always show *ěj* (top row, dark shade). Then there are roots which never show *ěj* (bottom row, light shade). And finally, there are ‘intermediate’ roots, which pattern with the *ěj*-less roots in the adjective, but with the other class in the adverb.

#### 4 DERIVING THE STANDARD PATTERN

In this section, we propose an account that delivers the correct pairing between the adjectival and the adverbial forms. The crucial idea is that adverbs have two additional heads compared to adjectives. Specifically, we propose a low adverb head (LOADV) in between C<sub>1</sub> and C<sub>2</sub>, and a high adverbial head (HiADV) above C<sub>2</sub>:



This structure is proposed mainly because it provides enough analytical space to account for Czech. However, Bobaljik (2012:§4.2) also provides some independent cross-linguistic evidence for two different positions of an adverbial head in a comparative, which we review briefly below.

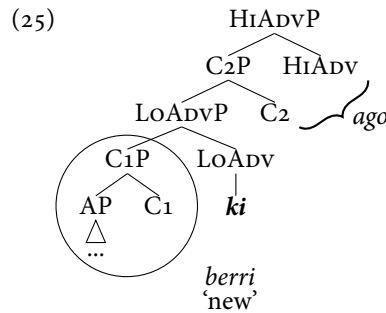
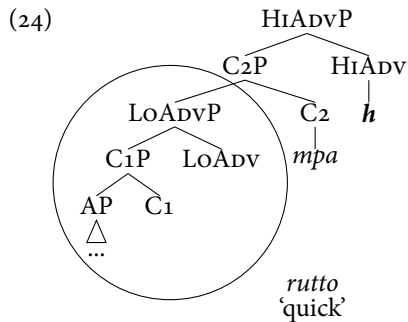
The evidence for the distinction between a low and a high Adv head comes from different affix orders in Karelian and Basque. We show this in Table (23).

(23) Karelian vs. Basque (Bobaljik 2012:118)

	Karelian 'quick'		Basque 'new'	
	POS	CMPR	POS	CMPR
ADJ	rutto	rutto- <i>mpa</i>	berri	berri - <i>ago</i>
ADV	rutto- <i>h</i>	rutto- <i>mpa-h</i>	berri- <i>ki</i>	berri- <i>ki-ago</i>

What we see here is that in Karelian, the adverbial suffix *-h* follows the comparative marker *-mpa*. On the other hand, the Basque adverbial suffix *-ki* precedes the comparative marker *-ago*. Following Bobaljik, we interpret this contrast as reflecting two different structural positions of the adverb head. Let us look into this in more detail.

There are reasons to think that Karelian *mpa* spells out C<sub>2</sub>, and therefore, that *h* occupies the high adverb head of (22). We depict our analysis in (24), where we assume that the remaining positions are realised by the root. We shall discuss the reasons for this analysis later on; for now, the main point is that if *mpa* realises C<sub>2</sub>, the position of *h* in the string reveals the need for an adverbial head above it.



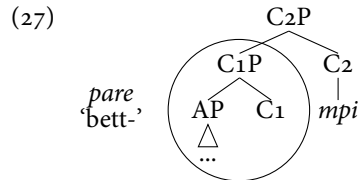
In (25), we show our analysis of the Basque comparative adverb 'newer'. Similarly to Karelian, the comparative marker *ago* spells out C<sub>2</sub>, which puts the adverbial *ki* into the LOADV head. In order to explain why the HiADV head is silent in Basque, we propose that *ago* spells out also HiADV. The root realises the remaining projections.



To repeat, what is crucial is that if the comparative markers in Basque and Karelian spell out C2, we can use the two different adverbial heads to explain the different orders: one higher and one lower than C2. Let us now turn to the evidence for the claim that the comparative markers spell out C2 in Karelian and Basque.

The analysis of Karelian *mpa* as C2 is suggested by the patterning of suppletive adjectives. Consider the forms of the suppletive adjective ‘good’ in (26).<sup>5</sup>

- (26) hyvä – pare-mpi  
 good bett-CMPR  
 ‘good – better’  
 (Karelian, Bobaljik 2012:116)



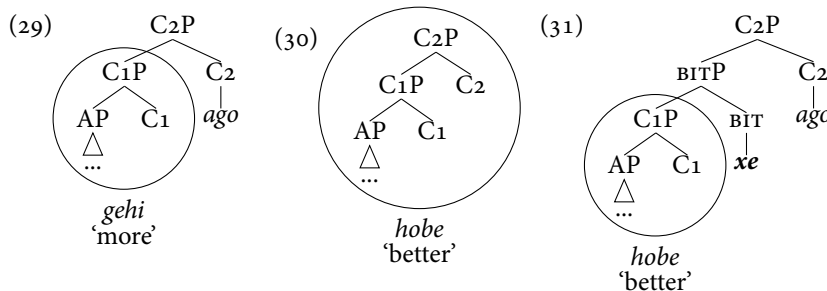
Recall now from (15) that Caha et al. (2019) posit slightly different feature specifications for pairs of suppletive roots. Following this approach, the suppletive comparative root *pare* ‘bett’ would be specified as lexicalising C1P, see (27). Spelling out C1 makes *pare* ‘bett’ different from the positive *hyvä* ‘good’, which spells out just AP. At the same time, specifying *pare* as C1P necessarily leads to the analysis of *mpi* as C2, because it comes on top of the suppletive root *pare* ‘bett’. If this is so, the Karelian adverb marker *h* must be hosted by an ADV head that is higher than C2, as posited in (24).

There are analogous reasons to think that the Basque comparative *ago* is also the realisation of C2. To see why *ago* is to be analysed as C2, consider the data in (28).

(28) Basque (Bobaljik 2012:156)

	POS	CMPR	‘a little more A’	GLOSS
a.	asko	gehi-ago	gehi-xe-ago	‘much’
b.	on	hobe	hobe-xe-ago	‘good’

First, we see that the suppletive form *gehi-ago* ‘more’ combines a suppletive root with an overt comparative marker. Assuming the same analysis of suppletion as for Karelian, we conclude that *gehi* must spell out C1P, which is why it differs from the positive *asko* ‘much’. If that is so, *-ago* must spell out C2, see (29).



A similar argument has been provided by Holaj & Starke (2019) on the basis of the suppletive pair *on* ‘good’ and *hobe* ‘better’ given on the bottom row of Table (28). Since *hobe* does not combine with any overt comparative marker, Holaj & Starke (2019) analyse it as lexicalising the whole C2P, see (30). However, in forms featuring the marker *xe* ‘a little bit’, *hobe* and *ago* do combine, yielding the form *hobe-xe-ago* ‘a little bit better’, recall (28). Holaj & Starke (2019) suggest that this can be understood if *xe* ‘a bit’ comes structurally in between C1 and C2, see (31). Its intervention prevents *hobe* from lexicalising C2, and

<sup>5</sup>It is not clear why ‘quick’ and ‘good’ use slightly different shapes of the comparative (*mpa* vs. *mpi*). A possible explanation is vowel harmony, but Bobaljik does not state this explicitly.

*ago* is therefore inserted. The suppletive root *hobe* realises just C1P in *hobe-xe-ago*. (This is allowed by the Superset Principle.)

In sum, the patterning of the suppletive adjectives in (28) tells us that *ago* is C2, and therefore, that the adverbial *ki* must be lower than C2, as proposed in (25).

However, we still need to determine the position of *ki* relative to C1. With this in mind, let us now turn to the adverbial form of the adjective ‘better,’ given in (32).

- (32) hobe-gi  
better-ADV

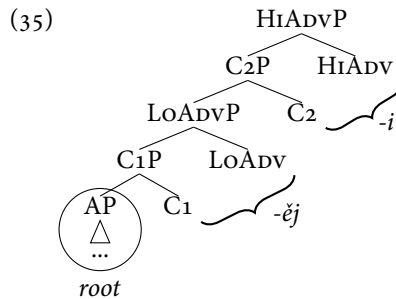
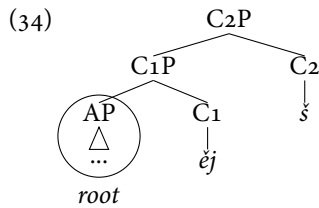
The relevant fact in (32) is that in the comparative adverb, the adverbial *ki* attaches on top of the suppletive form *hobe*; the form *\*on-gi-ago* (with *gi* attaching to the nonsuppletive root of the positive) is ungrammatical. This indicates that even though *ki* is lower than C2 (because it is lower than *ago*), it must be higher than C1, else it would not combine with *hobe*. These considerations locate *ki* unequivocally between C1 and C2, as proposed in (25).

In sum, Karelian and Basque each provide a piece of evidence for a particular location of an ADV head, yet the location is different in each case (Bobaljik 2012). The Basque adverbial marker *ki* is lower than C2, while the Karelian *h* is higher up. We therefore adopt a structure with two different ADV heads (ADVHI and ADVLO), which we shall now use to explain the patterns of allomorphy found in Czech adverbs. Recall that the patterns we want to capture are as in Table (33) (repeated from (21)).

(33) Three classes of comparative ADJS/ADVS in Standard Czech

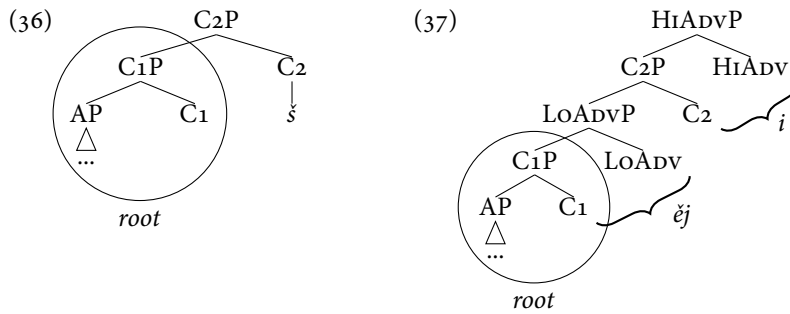
CMPR ADJ	CMPR ADV	GLOSS
chab-ěj-š-í	chab-ěj-i	‘weak’
slab-š-í	slab-ěj-i	‘weak’
draž-š-í	draž-e	‘expensive’

We begin with the *ěj-š-í* → *ěj-i* class on the top row. The roots of this class are lexically specified as APs. The structure of the comparative adjective is given in (34), and that of the comparative adverb in (35).



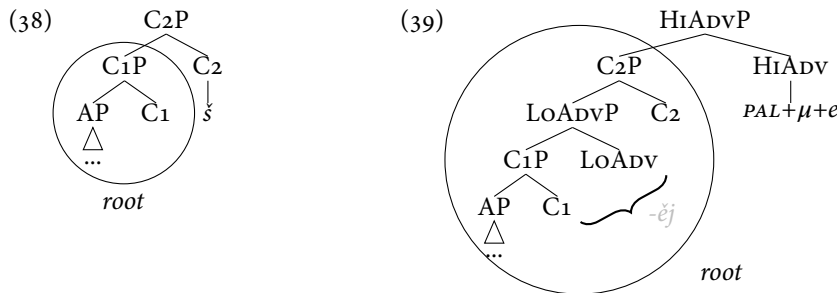
In (35), *ěj* spells out the LOADV head along with C1, as indicated by the brace. This is compatible with *ěj* lexicalising just C1 in the adjectival structure, since C1 is a proper subpart of a lexical entry of *ěj*, specified as [C1, LOADV]. The combination of [C2, HIADV] is realised as *i*.

Let us next turn to the *š-í* → *ěj-i* class (second row in (33)). This is the class where Caha et al. (2019) wrongly predicted adverbs like *\*slab-i* ‘more weakly’. Roots of this class are of size C1P. There is no need for *ěj* in the adjective, since the root spells out C1 (see (36)).



In the comparative adverb in (37), *ěj* is the only marker that can spell out LOAdv. On top of it, the marker *i* appears as the lexicalisation of [C2, HiAdv], exactly as in (35). This is then how the problematic class  $\check{s}\text{-}i \rightarrow \check{e}j\text{-}i$  arises: due to the additional LOAdv head, *ěj* is needed in the adverb, even though it is missing in the adjective.<sup>6</sup>

The third and final Standard Czech class is  $\check{s}\text{-}i \rightarrow \text{PAL}+\mu+e$ . It can be captured by associating the root to the adverbial C2P indicated in (39), where the root pronounces all the features except HiAdv, which is realised as  $\text{PAL}+\mu+e$ .<sup>7</sup>



In the comparative adjective in (38), the root cannot spell out C2P, since the C2P of (38) is not a sub-constituent of the root's lexical entry, which has the C2P tree as in (39). This is because the C2P of (38) lacks LOAdv. As a result, the root only spells out C1P, and C2 is spelled out as *š*.

Summarising the discussion of Standard Czech, we conclude that once the LOAdv head is added into the structure, it is possible to formulate an analysis that resolves the two puzzles we encountered. First, it provides a natural solution for the puzzling  $\check{s}\text{-}i \rightarrow \check{e}j\text{-}i$  class, where *ěj* emerges in the adverb due to the need to realise the LOAdv head. Second, it provides the analytical space for the two different comparative-adverb markers (*i* and  $\text{PAL}+\mu+e$ ), such that each marker has a different specification ([C2, HiAdv] for *i* vs. HiAdv for  $\text{PAL}+\mu+e$ ). The analysis also neatly captures the three classes of Standard Czech by showing that they correspond to three different root sizes: AP, C1P, and C2P.

This does not exhaust the logical options offered by our system: there could also be roots of size LOAdvP. Such roots would give rise to a fourth pattern:  $\check{s}\text{-}i \rightarrow i$ , i.e. the pattern that Caha et al. (2019) incorrectly predicted for the *slabý* type roots. This pattern does not exist in Standard Czech, and we treat this as an accidental gap. This decision is supported by the fact that roots of size LOAdvP exist in Czech dialects, as we discuss in

<sup>6</sup>In the adverb, the C1 feature could in principle be lexicalised by either the root or the *ěj* suffix. The spellout algorithm dictates that in such a case, the root only spells out AP. The *ěj* suffix lexicalises the remnant constituent [LOAdv [C1]], out of which the AP (lexicalised by the root) extracts. Technically, the derivation requires backtracking (Starke 2018).

<sup>7</sup>An anonymous reviewer asks what prevents a scenario where the root lexicalises the LOAdvP, combining with *i*, spelling out [C2, HiAdv]. The answer is that the root always tries to maximalise its lexicalisation potential, and only when a head cannot be lexicalised by the root must an affix be used. This follows from the Spellout Algorithm in Starke (2018), which determines the lexicalisation.

the following section.

## 5 DIALECT VARIATION

This section shows that a structure with two adverb heads also leads to a natural account of dialectal variation in comparative-adverb formation. Recall first that in the standard language, adverbs are marked by two different vocalic endings; see the leftmost column in Table (40).

(40) Dialectal Variation in Czech

Standard Czech	Apocope Dialects	
	East Morava	Central Morava
<i>i</i>	PAL+ $\mu$	
<i>ě</i>		PAL+ $\mu$

In a number of Czech dialects, these vocalic endings are absent (Balhar et al. 1997:379) and comparative adverbs are only marked suprasegmentally, namely by palatalisation and lengthening. Since the adverbial forms diachronically had a vocalic ending (Kosek 2014:95-6), we call these the Apocope Dialects. The loss of the final vowel affects the marking of the comparative degree in rather complex ways.

In this section, we zoom in on East and Central Morava dialects (Sections 5.1 and 5.2 respectively). Since these dialects lack the final vowel, they also lack the Standard-Czech distinction between *i* and *ě*, minimally on the surface. However, we argue that they also lack the distinction at a deeper level, in the sense that there has been paradigm levelling resulting in the loss of the Standard Czech distinction, as shown by the shading in (40). What Table (40) also intends to show is the different direction of the levelling: East-Morava has generalised the equivalent of standard *i* (the marker of [C<sub>2</sub>, H<sub>1</sub>ADV]), whereas Central-Morava has generalised the equivalent of standard *ě* (the realisation of H<sub>1</sub>ADV). This single point of difference interacts with the proposed structures in such a way that all the surface differences follow from this.

### 5.1 EAST MORAVA

In this section, we discuss adverbs in the East Morava dialect group. Our main point is to show that East Morava dialects have a uniform comparative-adverb marker, which is the counterpart of the standard *i*. The relevant facts are provided in Table (41) on the left. For ease of comparison, the classes are organised as in the standard language (we list the standard forms on the right).<sup>8</sup>

(41) East Morava dialect group (data from Hlubinková 2004, 2005)

East Morava			Standard Czech		
CMPR ADJ	CMPR ADV	GLOSS	CMPR ADJ	CMPR ADV	
tepl-ej-š-í	tepl-ěj	‘warm’	tepl-ej-š-í	tepl-ej-í	
ti(š) -š-í	tiš-ěj	‘silent’	tiš -š-í	tiš-ej-í	
dra(š) -š-í	draž	‘expensive’	draž -š-í	draž -e	

The table shows that there is no significant morphological difference between East Morava and the standard language regarding the adjectival comparatives. However, we can see that the adverbial forms lack the final vowel. The absence of the final vowel does not entail the absence of marking: we can see that adverbs in East Morava always lengthen

<sup>8</sup>The forms are given in the orthographic form, since the main focus here are not the phonetic details of the dialects, but the morphological structure of the form.

the final vowel, either of the suffix or the root (*tepl-ěj*, *draž*). Moreover, if they end in a non-palatal consonant (*drah-ý* ‘expensive’), they are palatalised (*draž*).<sup>9</sup>

Now since in East Morava, lengthening applies to all adverbs, this makes it tempting to propose an analysis with a uniform marking of comparative adverbs, which we show as  $PAL+\mu$  in Figure 2. This uniform adverbial marker always involves palatalisation in addition to lengthening. Palatalisation cannot be directly observed in the adverbs ending in *ěj*, since *j* is already palatal, but we can observe palatalisation whenever the base ends in a consonant that can be subject to it (e.g., *drah-ý* ‘expensive’ *draž* ‘more expensively’).

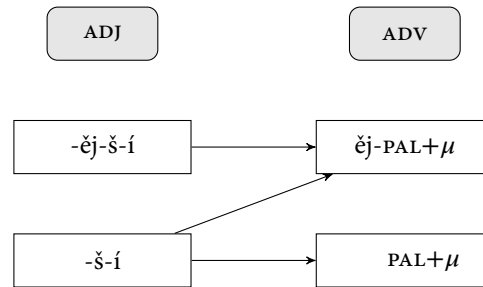
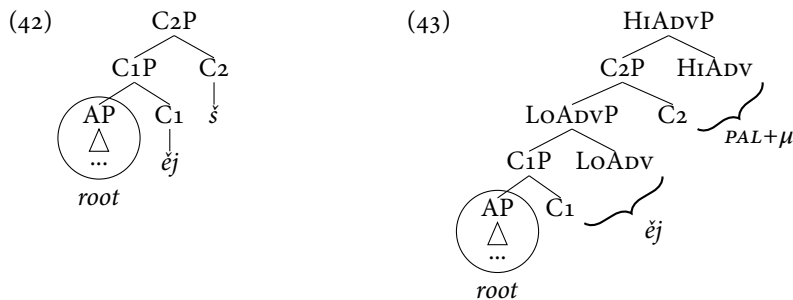


Figure 2: Adjective-adverb pairs in East Morava

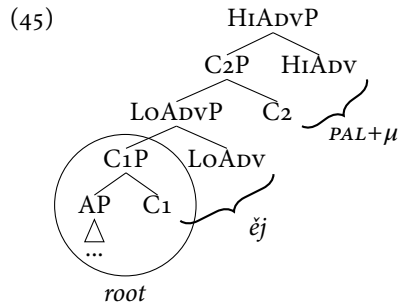
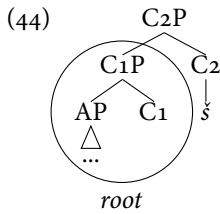
Other than having a single comparative-adverb marker ( $PAL+\mu$ ), the East-Morava dialects have a similar type of pairing between adjectives and adverbs as Standard Czech, with a clear intermediate class that lacks *ěj* in the adjectival comparative, but acquires it in the adverb. The main point of our analysis is to show that all these classes can be captured assuming a single comparative-adverb marker ( $PAL+\mu$ ) that lexicalises [C2, HiADV], which means that it has the same specification as the standard *i*.

Let us start the analysis with the *ěj-š-í* → *-ěj* class. This class is based on AP-sized roots. In (42), we see the adjective, and (43) gives the adverb. The lexicalisation is the same as in the standard language, the only difference being the phonological realisation of [C2, HiADV].

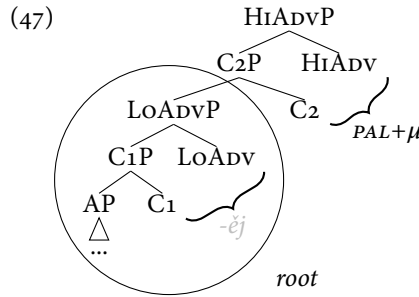
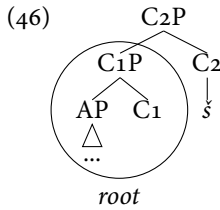


The analysis of the *š-í* → *-ěj* class is also analogous to the standard, and it is based on roots of the size C1P. The structures are shown below in (44) and (45).

<sup>9</sup>The lengthening of *ěj* to *ěj* is subject to variation within East Morava (short adverbial forms are also found). Hlubinková (2004) gives forms such as *tepl-ěj/tepl-ej* ‘warmer’. This contrasts with adverbs such as *draž* ‘more expensively’, which are never short (*\*draž*). We analyse this as a matter of phonology, because it is independently known that the palatalising *ě* often prevents lengthening in contexts where all other vowels show length alternations. For instance, Caha & Scheer (2008) observe that infinitives with a single vowel always lengthen (e.g., *pi-l* ‘drank’ has the infinitive *pi-t* ‘to drink’), but if that vowel is *ě*, it often fails to lengthen (*pě-l* ‘sang’ has the infinitive *pě-t* ‘to sing’). We therefore assume that morphologically, all comparative adverbs lengthen, but *ěj* adverbs may fail to show length for phonological reasons.



The final class of roots ( $\check{s}-i \rightarrow PAL+\mu$ ) could be of size C2P (as in the Standard), but since a dedicated HIADV marker (analogous to the standard  $\check{e}$ ) is missing in East Morava, a C2P size root would still pair with the uniform adverb marker  $PAL+\mu$ . For simplicity, we therefore assume that the roots in question are of the size LOADV, as we show in (46) and (47).



This concludes the analysis of the three classes in East Morava Czech. The main point was to show that if we assume a uniform non-segmental  $PAL+\mu$  ending, the analogue of Standard Czech  $i$ , we can easily capture the facts as they are found in East Morava dialects. Aside from the absence of a segmental adverb marker, East Morava differs from Standard Czech only in the size of the roots of the third class, which are of size LOADV rather than C2P. In the following section, we turn to Central Morava dialects. Our main idea is that the non-segmental ending in Central Morava is the analogue of the standard  $\check{e}$ , spelling out just HIADV.

## 5.2 CENTRAL MORAVA ADVERBS

In this section, we discuss comparative adverbs found in Central Morava Czech.<sup>10</sup> This dialect group differs from both the standard language as well as the East Morava dialects discussed in Section 5.1 in two important respects. First, it is the only dialect type that distinguishes four different classes of roots when it comes to comparative formation. Second, it maintains  $\check{s}$  in the comparative adverb in most of the classes.

The main set of facts that we aim to explain are given in Table (48). The table presents the four classes labelled with Roman numerals (I-IV), with Class IV exemplified by two adjectives.

<sup>10</sup>Similar facts are also found in West Bohemia.

## (48) Central Morava comparative adjectives and adverbs

	POS ADJ	CMPR ADJ	CMPR ADV	GLOSS
I.	chab-ý	chab-ěj-š-í	chab-ěj-š	‘weak’
II.	hrub-ý	hrub- š-í	hrub-ěj-š	‘rough’
III.	slab-ý	slab- š-í	sláb- š	‘weak’
IV.	drah-ý	draž- š-í	dráž	‘expensive’
	dobr-ý	lep- š-í	líp	‘good’

The first class is exemplified by the adjective *chab-ý* ‘weak.’ The comparative adverb *chab-ěj-š* illustrates that in this dialect, adverbs corresponding to *ěj-š-í* comparatives maintain both *ěj* and *š* in the adverb. Apparently, there is no overt adverbial ending, but this does not automatically mean that morphologically speaking, there is no adverbial marking. Possibly, the marking may be non-segmental, namely lengthening, which, however, fails to lengthen the palatalising *ě* (as discussed in footnote 9).

The second fact that makes Central Morava dialects different is that *š-í* comparatives have not two corresponding adverbs, but three. These correspond to classes II-IV in Table (48).

Class II has an adjectival *š-í* comparative paired with an adverb in *ěj-š*. The adverb thus looks the same as the adverb corresponding to an *ěj-š-í* adjective (as highlighted by the dark shading). We have already encountered this intermediate class before both in Standard Czech and in East Morava.

Another familiar pattern is found in Class IV. The adverb *dráž* ‘more expensively’ shows neither C1 *ěj* or C2 *š*, but it undergoes lengthening and palatalisation (which we take to be a non-segmental realisation of the comparative-adverb function). On the bottom row of Table (48), we also include the suppletive adverb *líp* ‘better,’ which belongs in the same class. We include it because this adverb makes it clear that there is no C2 ending *š*. This is a point that may be spurious in the case of *dráž* ‘more expensively,’ because one could argue that this is an effect of degemination, and that the underlying morphological structure is *dráž-š*. In this context, *líp* ‘better’ is a clearer instance of this class, since there is nothing phonologically odd about the ungrammatical *\*lípš* (certainly, there is no need for degemination here). Note, however, that the root of the adverb ‘better’ *líp* is lengthened (the comparative adjective being *lep*). Palatalisation has no effect on the labial.

The final class of *š-í* adjectives is represented by the root *slab-ý* ‘weak’ on row III. This root has the adverb *sláb-š*, which preserves the C2 marker *š*, while the root also lengthens. This is a new class, which is not found in Standard or East-Morava Czech.

Figure 3 summarises the four patterns.

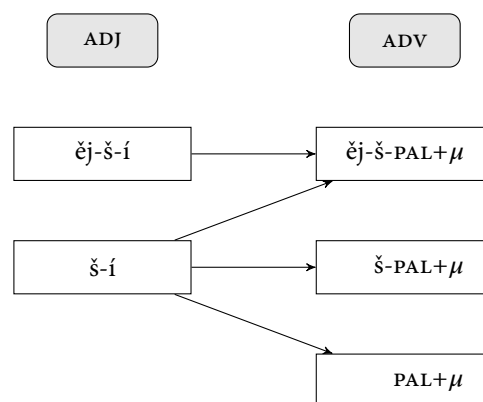
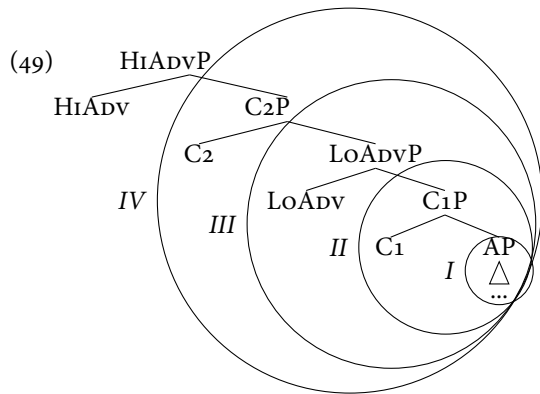


Figure 3: Adjective-adverb pairs in Central Morava

In the figure, we are again assuming that all classes have a uniform comparative-adverb marker, namely  $PAL+\mu$ . The evidence for this marker is direct in Classes III and IV of Table (48) (where we observe lengthening), but indirect in the case of Classes I and II. These classes have the  $\check{e}j-\check{s}$  ending, and  $\check{e}$  is immune to lengthening (as we have argued in footnote 9). We nevertheless assume that the same comparative-adverb marker is present here too, maintaining the idea that upon losing the segmental endings, the evidence for a distinction between two different adverbial endings is lost, and a single comparative-adverb marker appears in its stead. We shall argue that in Central Morava, this marker is the realisation of  $HiAdv$ , i.e., the counterpart of the Standard  $\check{e}$ . This idea is going to explain why comparative adverbs in this dialect acquire  $\check{s}$  in Classes I-III.

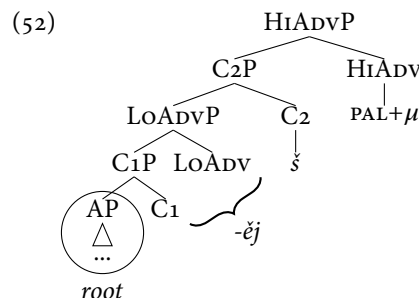
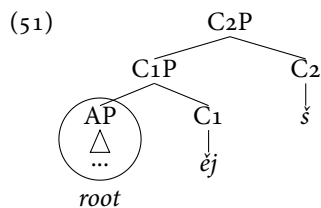
Figure 3 represents the most complex system found in the Czech dialects, and our goal will be to explain both the allomorphs found, and the pairing between adjectives and adverbs, which requires us to operate with four different classes of roots. Our idea is that the four classes of adjective-adverb pairs can be derived under the hypothesis that each class is based on a root of a different size, as shown in (49). The class numbers in (49) refer to the classes I-IV as given in (48), repeated for convenience in (50).



(50) Central Morava comparative adjectives and adverbs

	POS ADJ	CMPR ADJ	CMPR ADV	GLOSS
I.	chab-ý	chab- <b>ěj-š</b> -í	chab- <b>ěj-š</b>	'weak'
II.	hrub-ý	hrub- š-í	hrub- <b>ěj-š</b>	'rough'
III.	slab-ý	slab- š-í	sláb- š	'weak'
IV.	drah-ý	draž- š-í	draž	'expensive'
	dobr-ý	lep- š-í	líp	'good'

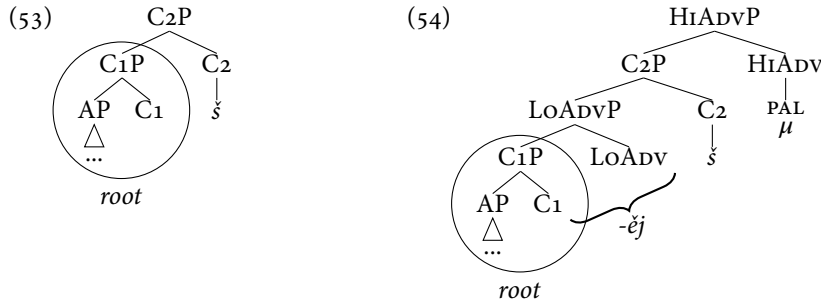
Let us now look at the individual patterns. We begin with the AP-sized roots (Class I). Such roots always need both  $\check{e}j$  and  $\check{s}$ , yielding the  $\check{e}j-\check{s}-i \rightarrow \check{e}j-\check{s}$  pattern. The structure of the comparative adjective is given in (51), C1 is spelled out by  $\check{e}j$ , and C2 by  $\check{s}$ .





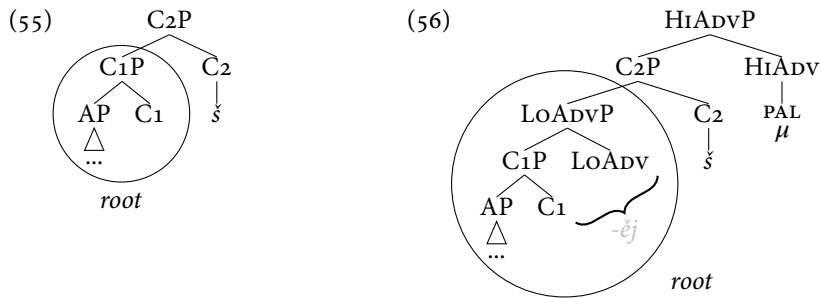
The structure of the comparative adverb is provided in (52). The suprasegmental ending only spells out HiADV, which leads to the emergence of *š* under C2. The morpheme *ěj* spells out C1 and LoADV, as it does in all other varieties.

Let us now turn to roots of size C1P (Class II). In the adjectival structure, these roots spell out C1P, and they therefore do not need *ěj*: their adjectival comparative is just *š-í*.



However, in the adverb structure, these roots still need *ěj*, because something must realise the LoADV head. The rest of the structure is spelled out as in (52). In sum, roots of the size C1P give rise to the *š-í* → *ěj-š* pattern, with *ěj* appearing due to the low adverb head.

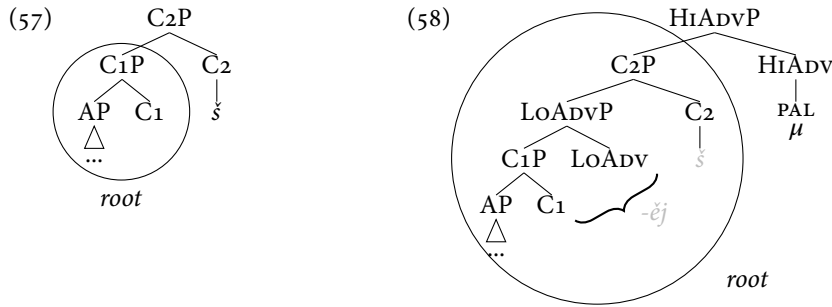
Roots of size LoADV (Class III) give rise to the *š-í* → *š* pattern. We show this in (55) and (56). Starting with the adverbial structure (56), we see that with these roots, there is no need for the marker *ěj*, since the root lexicalises the LoADV feature, which triggers the presence of *ěj* in the adverbs that have it. However, the root requires the presence of *š* to spell out C2, and it also lengthens, e.g. *slab-š-í* → *sláb-š* ‘weaker’.



The very same root can also spell out just C1P in the adjectival structure (55), because this is a sub-constituent of the LoADV; hence, spellout of C1P in (55) is allowed by the Superset Principle. The comparative marker *ěj* remains absent here too since the root lexicalises C1.

Finally, Class IV roots are of size C2P (where this C2P includes the LoADV head). They do not need either *ěj* or *š* in the comparative adverb, as shown in (58). They only undergo lengthening and palatalisation (recall from (48) the length difference between *draž-š-í* → *dráž* ‘more expensive’; or *lep-š-í* → *líp* ‘better’).<sup>11</sup>

<sup>11</sup>The raising of *e* to *i* under lengthening is productive in most Czech dialects, including Central Morava, and we are thus assuming that the adverb *líp* ‘better’ is a lengthened version of the adjectival root *lep*. Alternatively, we could also analyse *líp* as a suppletive adverbial root spelling out C2P (including the LoADV head), while *lep* would be just C1P.



In the comparative adjective in (57), the root spells out only C1P, and C2 is therefore spelled out as *š*. The root cannot spell out C2P in the adjectival structure (57), since this is not a sub-constituent of the root's lexical entry. The lexical entry of the root corresponds to the adverbial C2P, and since the adjectival C2P lacks LOADV, it is not a subconstituent of the adverbial one. As a result, the root can only spell out C1P in the adjective, and this leads to the emergence of *š* under C2 in (57).

### 5.3 SUMMARY

This section has shown that the system developed to handle Standard Czech (with two adverb heads and a split comparative) can be extended to capture various dialects of Czech. The main idea of the analysis is that due to the lack of a segmental ending, the dialects use a single adverbial ending. Our idea is that in some dialects (exemplified by East Morava), the single surviving ending is the counterpart of the standard *i*. In other dialects (Central Morava), it is the counterpart of *ě*. Different types of adverbial marking arise as an automatic consequence of this simple distinction. The final point we want to stress regarding this section is that the rich structure with two adverb heads and a split comparative allows us to capture the fact that there are four different classes of roots in Central Morava. The rather complex system of pairing between adjectival and adverbial comparatives in this dialect falls out as an effect of root size, with no need to say anything special beyond characterising the individual roots in terms of what features they are capable of realising.

## 6 CONCLUSIONS

In this article, we have investigated the complex system of pairing between comparative adjectives and comparative adverbs in Standard Czech and Czech dialects. We have provided a formal model of how this pairing is achieved for each variety, and we have also identified a plausible source of differences among the varieties: while Standard Czech has two different endings (*i* and *e*), the dialects only have a single ending. This single ending came to exist due to paradigmatic levelling on the one hand, and phonological erosion on the other, whereby the segmental vocalic adverbial endings got eliminated in the dialects. The difference in the direction of the levelling led to different systems of adverb formation.

We were able to achieve this result by adopting the following ideas.

- (59)
- a. The comparative meaning is distributed across two projections, C1 and C2
  - b. There is both a low and a high adverbial head
  - c. Lexicalisation targets non-terminals. Which lexical item realises which head is determined by the Spellout Algorithm of Starke (2018)
  - d. Roots belong to different classes. These can be modelled by attributing to each root a different number of functional features in the lexicon

Once these assumptions are adopted, the allomorphs that each root combines with in the adjectival and in the adverbial comparative are a function of the root's specification, in combination with the proposals (59-ac).

As a final note, we observe that in achieving this result, the proposal does not rely on complex context-sensitive rules. Consider, for instance, the distribution of *ěj* in adjective-adverb pairs such as *draž-š-í* — *dráž-e* 'more expensive' and *slab-š-í* — *slab-ěj-i* 'weaker.' On the surface, we observe the following: C1 (the position of exponence for *ěj*) is always realised as zero when attached to roots such as *drah*; but with the root *slab*, it is zero only if the form is an adjective, but not an adverb. Otherwise, C1 is realised as *ěj*. Thanks to phrasal lexicalisation, zero morphemes (with such complex distributions) play no role in our account, which we take to be a notable achievement of the analysis.

## ABBREVIATIONS

ADJ	adjective	CMPR	comparative
ADV	adverb	POS	positive

## ACKNOWLEDGMENTS

The authors would like to thank two anonymous JSL reviewers for helpful comments and suggestions. We also thank the participants of the Nanoseminar in the Fall of 2021, where parts of this material were presented, for their input. Guido Vanden Wyngaerd's research was supported by the Flemish Research Foundation FWO in the context of the REEDS research network on dialect syntax (FWO W002320N).

## CONTACT

PAVEL CAHA — [pavel.caha@phil.muni.cz](mailto:pavel.caha@phil.muni.cz) (corresponding author)  
 KAREN DE CLERCQ — [karen.de-clercq@cnrs.fr](mailto:karen.de-clercq@cnrs.fr)  
 GUIDO VANDEN WYNGAERD — [guido.vandenwyngaerd@kuleuven.be](mailto:guido.vandenwyngaerd@kuleuven.be)

## REFERENCES

- Balhar, Jan, Pavel Jančák & Dialektologické kolektiv Ústavu pro jazyk český AV ČR. 1997. *Český jazykový atlas [Czech language atlas], vol. 2*. Praha: Academia.
- 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, Karen De Clercq & Guido Vanden Wyngaerd. 2019. The fine structure of the comparative. *Studia linguistica* 73(3). 470–521.
- Caha, Pavel & Tobias Scheer. 2008. The syntax and phonology of Czech templatic morphology. In Andrei Antonenko, John Bailyn & Christina Bethin (eds.), *Annual workshop on Formal Approaches to Slavic Linguistics. The Stony Brook meeting 2007*, 68–83. Ann Arbor, MI: Michigan Slavic Publications.
- 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.

- Hlubinková, Zuzana. 2004. Stupňování adverbíí, zvláště ve východomoravských nářečích [Grading of adverbs, especially in East Moravian dialects]. *Linguistica Brunensia* 52. 109–114.
- Hlubinková, Zuzana. 2005. Tvoření deadjektivních adjektiv, zvláště ve východomoravských nářečích [Formation of deadjectival adjectives, especially in East Moravian dialects]. *Linguistica Brunensia* 53. 153–162.
- Holaj, Richard & Michal Starke. 2019. Who is afraid of a non-local wolf? Talk presented at CRAFF, Brno.
- 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.
- Kosek, Pavel. 2014. *Historická mluvnice češtiny I [Historical dictionary of the Czech language]*. Brno: Masarykova univerzita.
- 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.
- Starke, Michal. 2009. Nanosyntax: A short primer to a new approach to language. *Nordlyd* 36. 1–6.
- 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.
- Vyshnevskia, Anastasiia. 2022. There is more to the morphology of -š/-iš: the comparative in Ukrainian. *Linguistics in the Netherlands* 39. 263–278.