



Memorial
University of Newfoundland

Department
of Linguistics

October 19, 2018

Morris Halle's *The Sound Pattern of
Russian*: The road not taken

B. Elan Dresher
University of Toronto

Daniel Currie Hall
Saint Mary's University

*Two roads diverged in a yellow wood,
And sorry I could not travel both
and be one traveler, long I stood
And looked down one as far as I could
To where it bent in the undergrowth;*

Robert Frost, 'The Road Not Taken' (1916)

Introduction

Morris Halle's *The Sound Pattern of Russian* (1959; henceforth *SPR*) sits at a significant fork in the road in the development of phonological theory.

It can be viewed from one perspective as the culmination of a tradition of phonological analysis associated with the Prague School, the last in a line of works that include Roman Jakobson's *Kindersprache, Aphasie, und allgemeine Lautgesetze* (1941), Jakobson, Fant & Halle's *Preliminaries to Speech Analysis* (1952), and Jakobson & Halle's *Fundamentals of Language* (1956).

From another perspective, it is the first major work in the new framework of Generative Phonology, a precursor to Chomsky & Halle's *The Sound Pattern of English* (1968).

Introduction

From his earlier work with Jakobson, Halle (1959) retained the notion of a branching tree that generates all and only the contrastive features of each Russian underlying segment.

SPR contains a novel argument for why it is necessary to generate feature specifications by means of such trees.

However, *SPR* is perhaps best known in the phonological literature for Halle's argument against the structuralist phoneme, based on his analysis of Russian regressive voicing assimilation (RVA).

Somewhat ironically, this analysis had the effect of devaluing the importance of the branching tree, and of contrastive features more generally.

Introduction

As a consequence, Chomsky and Halle's *The Sound Pattern of English* (1968) abandoned contrastive underspecification and feature hierarchies.

The result, in our view, was that generative grammar gave up some major insights of the Prague School phonologists, including Halle himself.

However, Halle could have taken a different path in *The Sound Pattern of Russian* that would have made all the difference.

Introduction

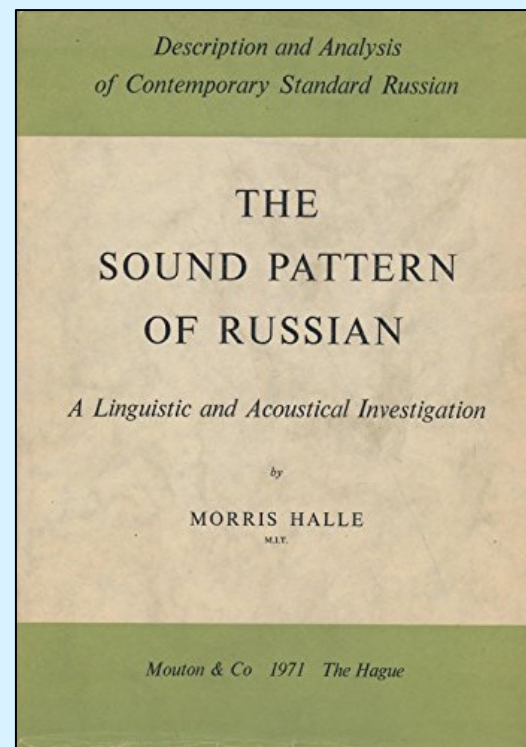
In this talk we will first look at the elements that underpinned Halle's famous argument, and the consequences it had for Generative Phonology.

In the second part, we will show how Halle could have easily taken another road.

We will propose that this other road is still there to be followed, and is still worth considering.

Halle (1959): The branching tree

On page 46 of *The Sound Pattern of Russian (SPR)* is Figure I-1, a magnificent tree diagram that shows the contrastive feature specifications of every phoneme of Russian.



On page 46 of *The Sound Pattern of Russian (SPR)* is Figure I-1, a magnificent tree diagram that shows the contrastive feature specifications of every phoneme of Russian.

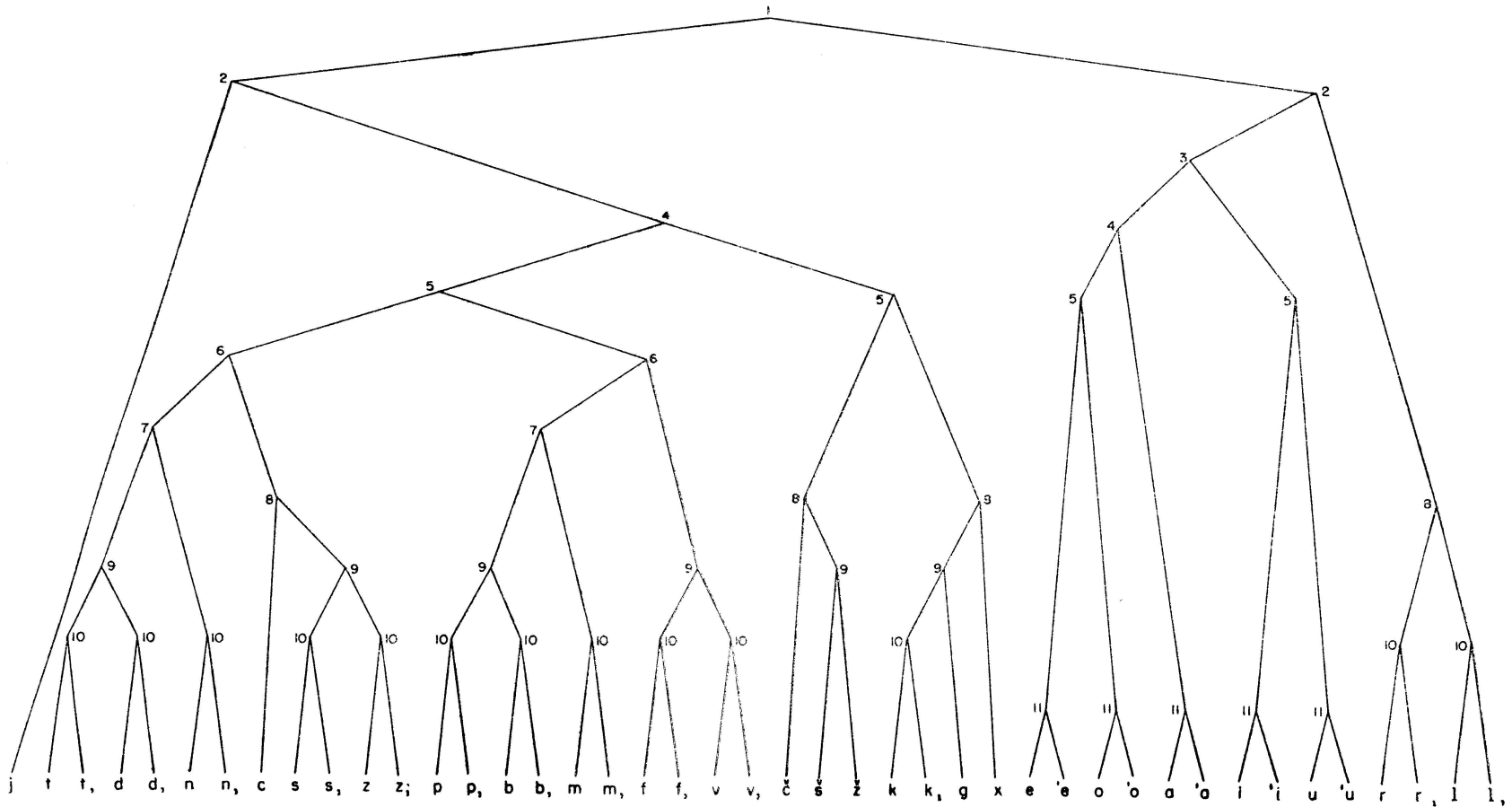


Fig. I-1. Branching diagram representing the morphemes of Russian. The numbers with which each node is labelled refer to the different features, as follows: 1. vocalic vs. nonvocalic; 2. consonantal vs. nonconsonantal; 3. diffuse vs. nondiffuse; 4. compact vs. noncompact; 5. low tonality vs. high tonality; 6. strident vs. mellow; 7. nasal vs. nonnasal; 8. continuant vs. interrupted; 9. voiced vs. voiceless; 10. sharpened vs. plain; 11. accented vs. unaccented. Left branches represent minus values, and right branches, plus values for the particular feature.

The highest feature is [\pm vocalic]: all the phonemes on the left in blue are [$-$ vocalic] glides and consonants, and the ones on the right in red are [$+$ vocalic] vowels and liquids.

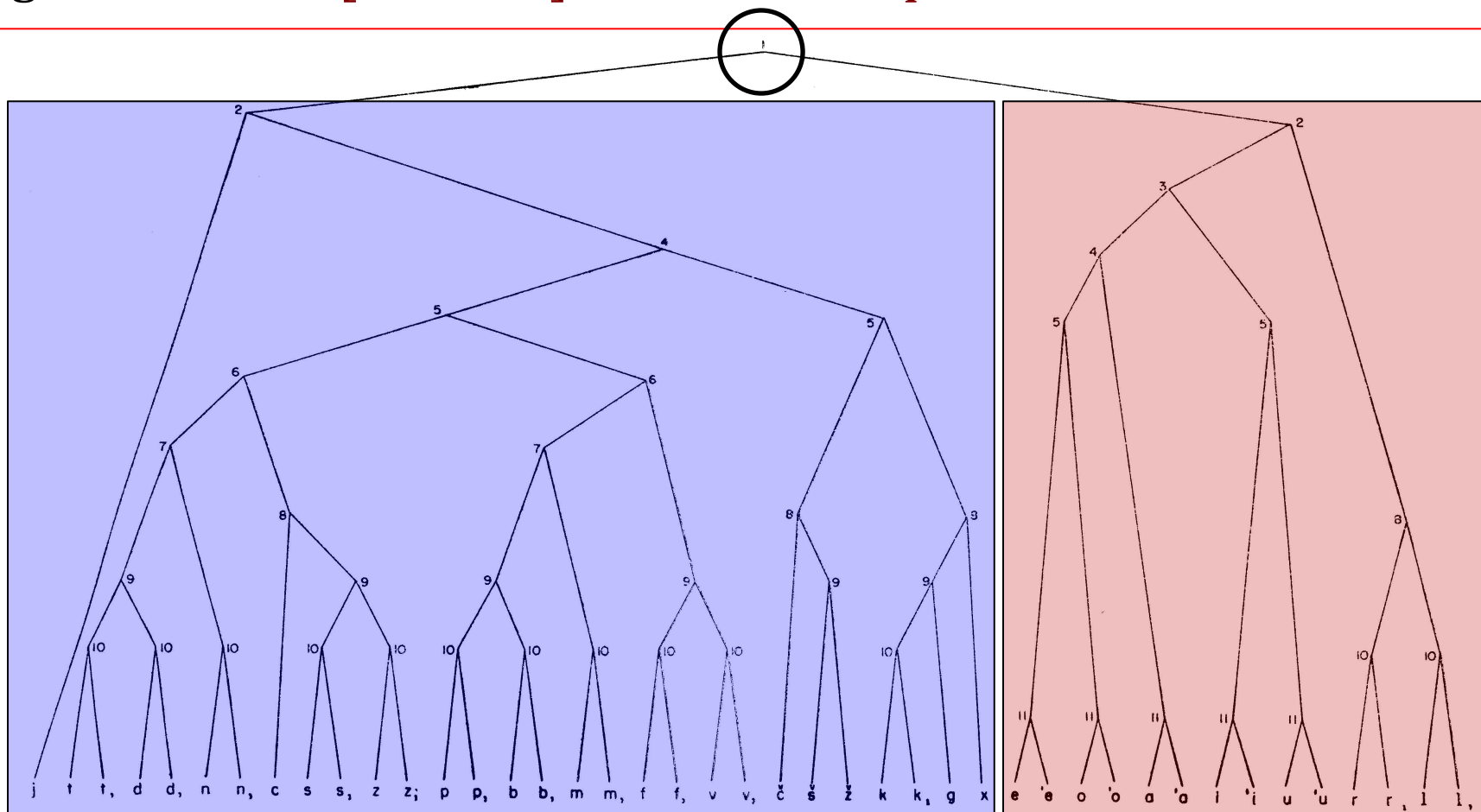


Fig. I-1. Branching diagram representing the morphonemes of Russian. The numbers with which each node is labelled refer to the different features, as follows: 1. vocalic vs. nonvocalic; 2. consonantal vs. nonconsonantal; 3. diffuse vs. nondiffuse; 4. compact vs. noncompact; 5. low tonality vs. high tonality; 6. strident vs. mellow; 7. nasal vs. nonnasal; 8. continuant vs. interrupted; 9. voiced vs. voiceless; 10. sharpened vs. plain; 11. accented vs. unaccented. Left branches represent minus values, and right branches, plus values for the particular feature.

The next feature is $[\pm\text{consonantal}]$, which is contrastive in both major branches of the tree.

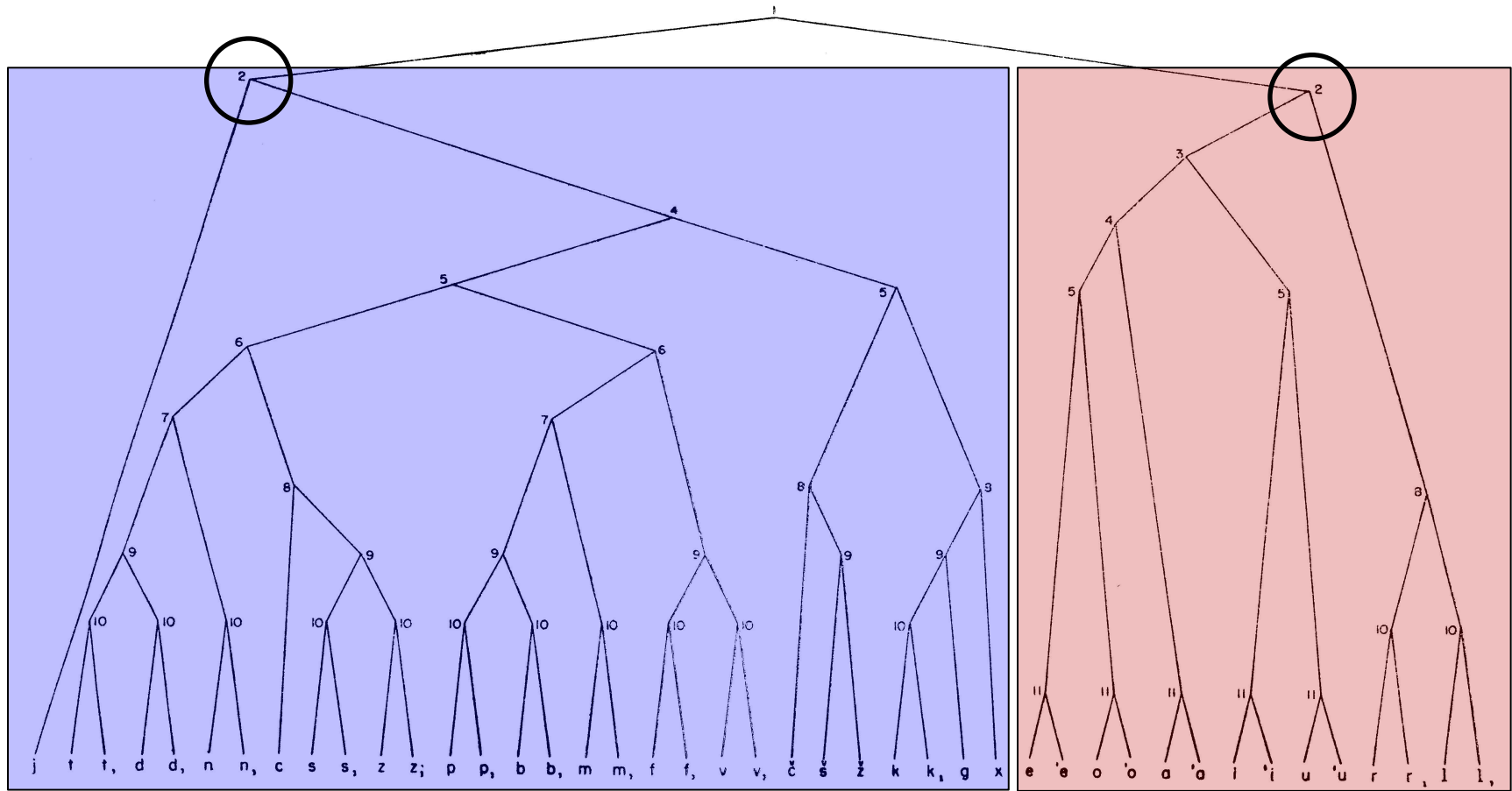


Fig. I-1. Branching diagram representing the morphemes of Russian. The numbers with which each node is labelled refer to the different features, as follows: 1. vocalic vs. nonvocalic; 2. consonantal vs. nonconsonantal; 3. diffuse vs. nondiffuse; 4. compact vs. noncompact; 5. low tonality vs. high tonality; 6. strident vs. mellow; 7. nasal vs. nonnasal; 8. continuant vs. interrupted; 9. voiced vs. voiceless; 10. sharpened vs. plain; 11. accented vs. unaccented. Left branches represent minus values, and right branches, plus values for the particular feature.

Looking first at the left branch of the tree, only the glide /j/ is contrastively [-consonantal]. As it is now unique, no further features are assigned to /j/.

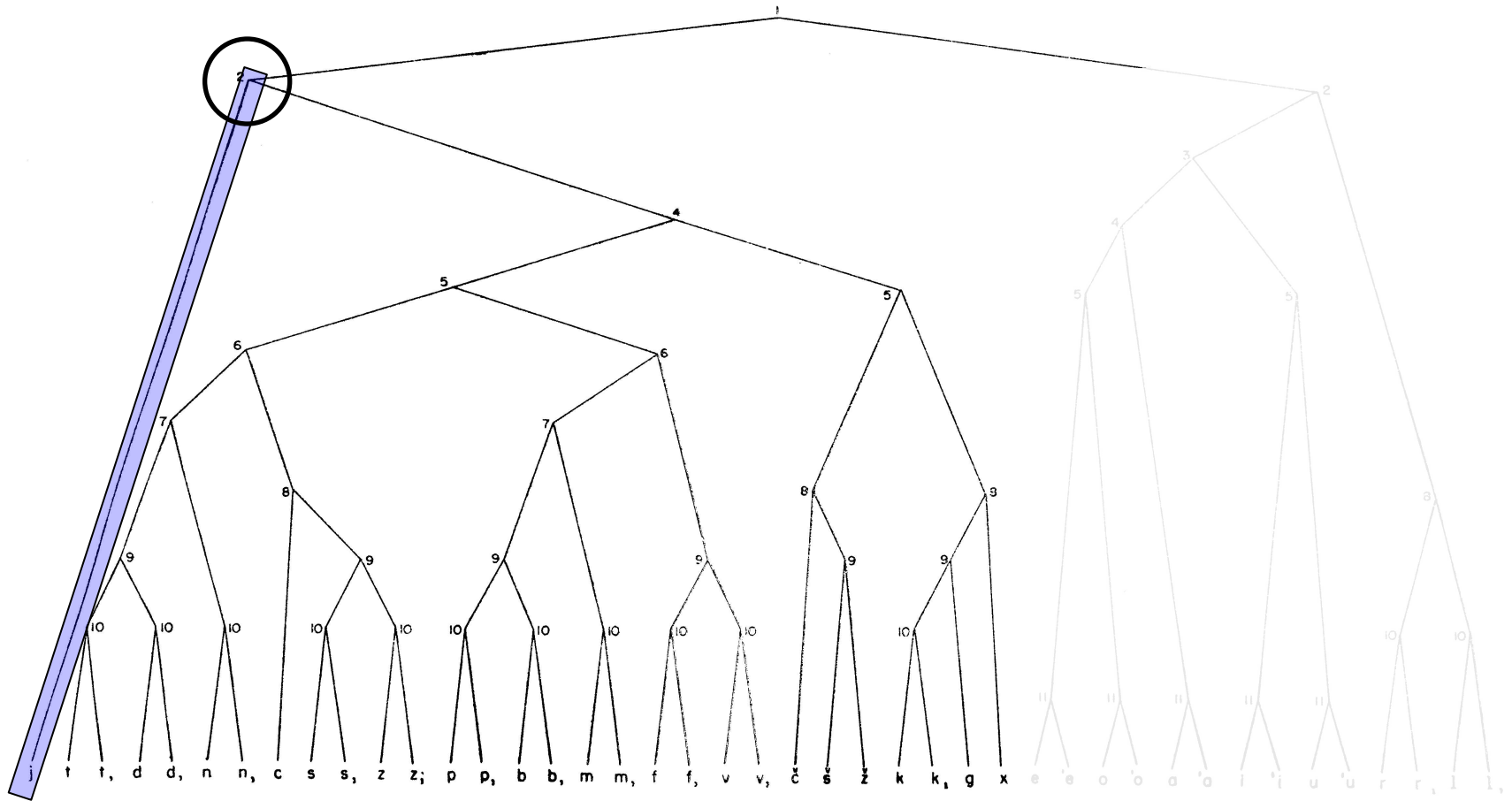


Fig. I-1. Branching diagram representing the morphemes of Russian. The numbers with which each node is labelled refer to the different features, as follows: 1. vocalic vs. nonvocalic; 2. consonantal vs. nonconsonantal; 3. diffuse vs. nondiffuse; 4. compact vs. noncompact; 5. low tonality vs. high tonality; 6. strident vs. mellow; 7. nasal vs. nonnasal; 8. continuant vs. interrupted; 9. voiced vs. voiceless; 10. sharpened vs. plain; 11. accented vs. unaccented. Left branches represent minus values, and right branches, plus values for the particular feature.

All the other segments are [+consonantal], and therefore need to be distinguished from each other by additional features.

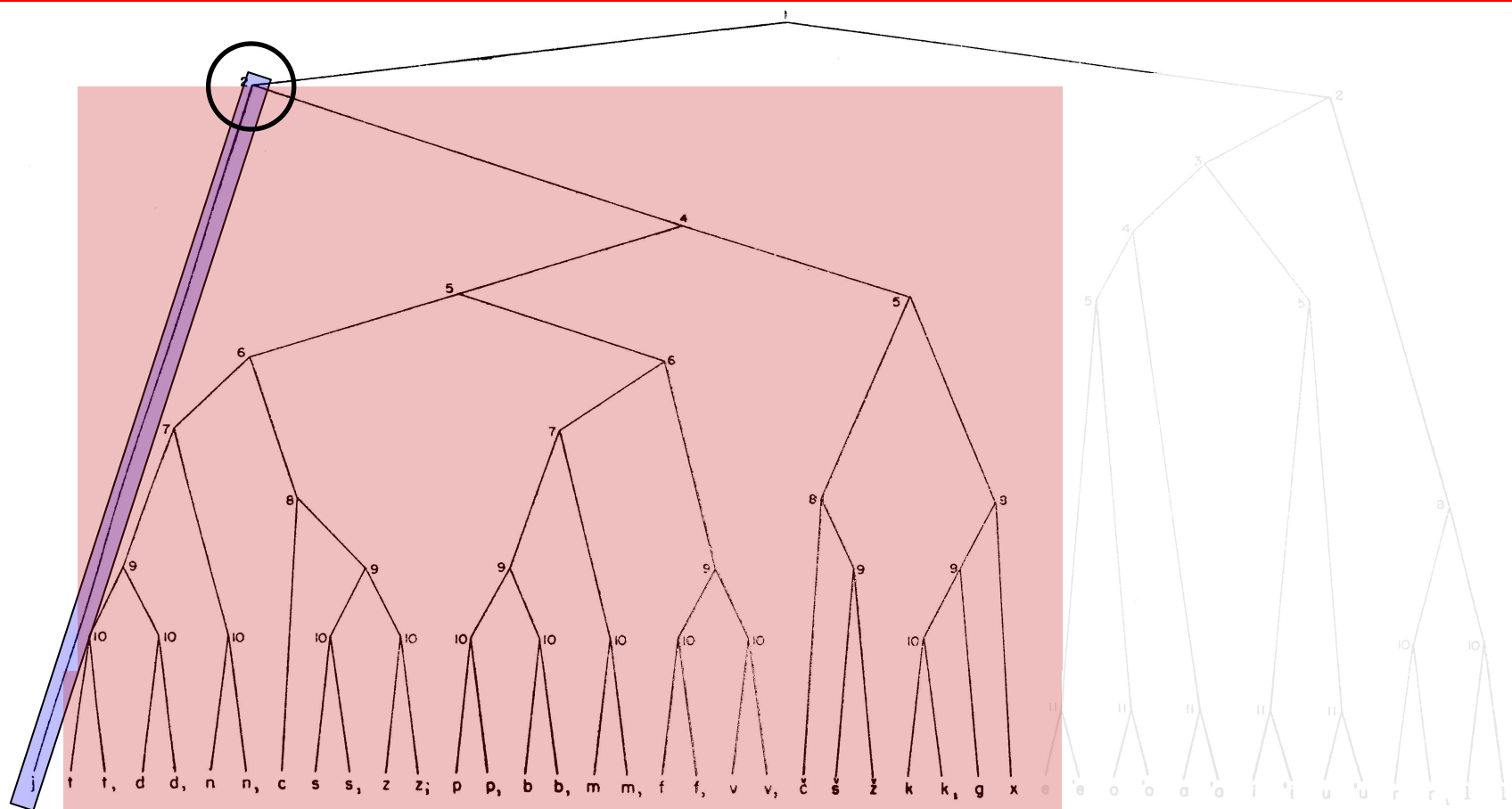


Fig. I-1. Branching diagram representing the morphemes of Russian. The numbers with which each node is labelled refer to the different features, as follows: 1. vocalic vs. nonvocalic; 2. consonantal vs. nonconsonantal; 3. diffuse vs. nondiffuse; 4. compact vs. noncompact; 5. low tonality vs. high tonality; 6. strident vs. mellow; 7. nasal vs. nonnasal; 8. continuant vs. interrupted; 9. voiced vs. voiceless; 10. sharpened vs. plain; 11. accented vs. unaccented. Left branches represent minus values, and right branches, plus values for the particular feature.

On the [+vocalic] side of the tree, **vowels** are [-consonantal] and **liquids** are [+consonantal].

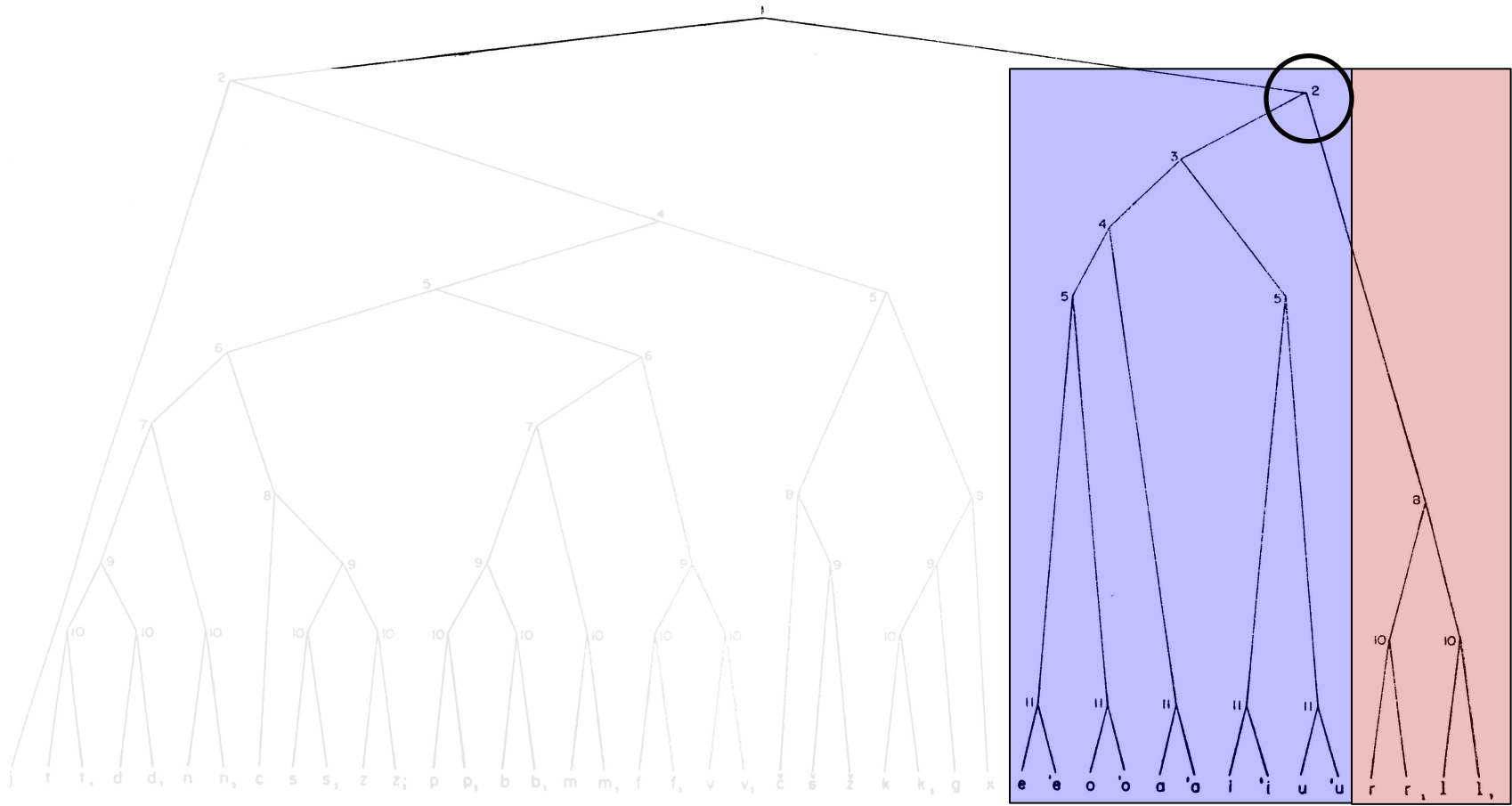


Fig. I-1. Branching diagram representing the morphemes of Russian. The numbers with which each node is labelled refer to the different features, as follows: 1. vocalic vs. nonvocalic; 2. consonantal vs. nonconsonantal; 3. diffuse vs. nondiffuse; 4. compact vs. noncompact; 5. low tonality vs. high tonality; 6. strident vs. mellow; 7. nasal vs. nonnasal; 8. continuant vs. interrupted; 9. voiced vs. voiceless; 10. sharpened vs. plain; 11. accented vs. unaccented. Left branches represent minus values, and right branches, plus values for the particular feature.

We continue dividing the tree by contrastive features until every phoneme has been uniquely distinguished.

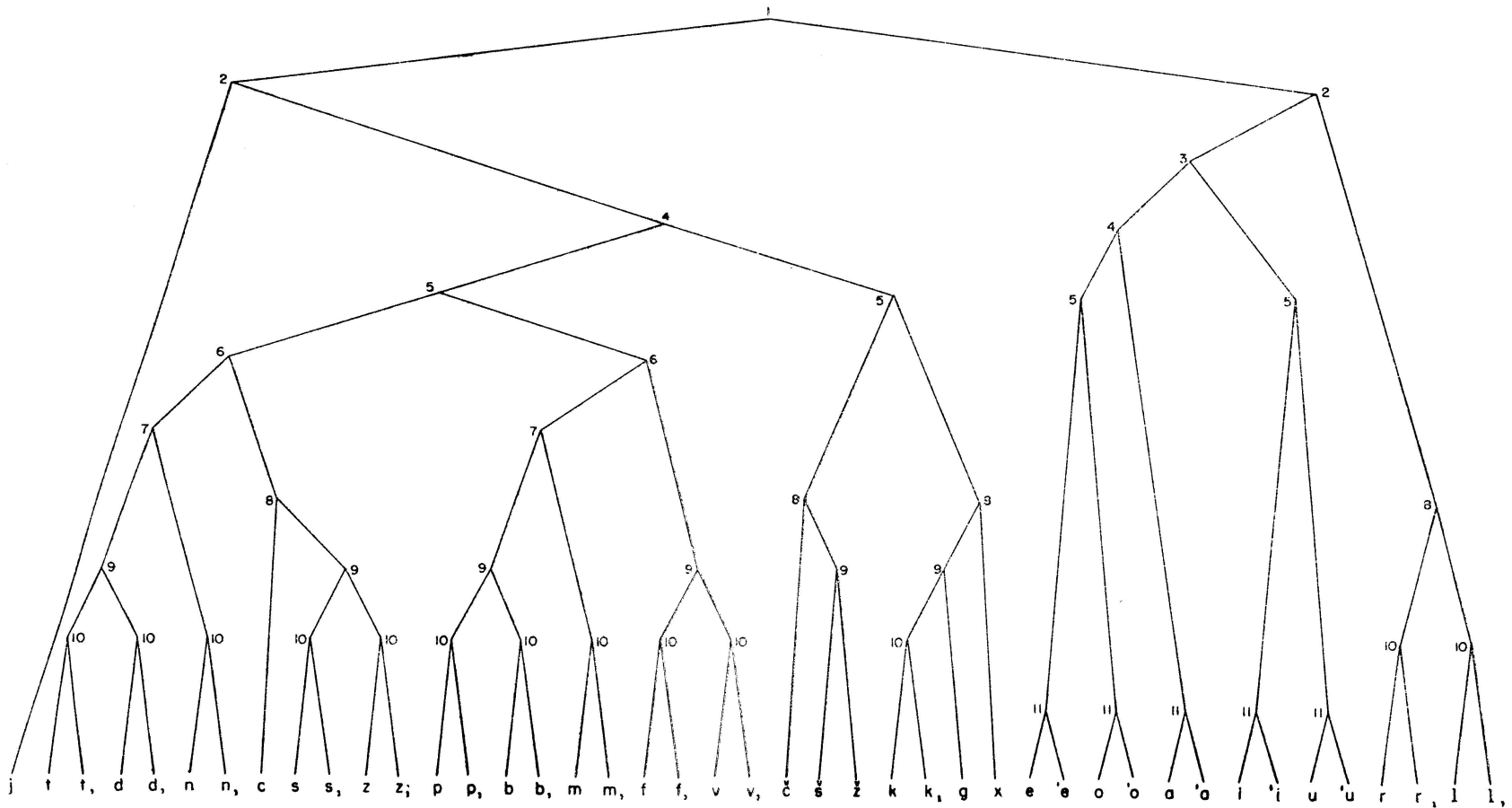


Fig. I-1. Branching diagram representing the morphemes of Russian. The numbers with which each node is labelled refer to the different features, as follows: 1. vocalic vs. nonvocalic; 2. consonantal vs. nonconsonantal; 3. diffuse vs. nondiffuse; 4. compact vs. noncompact; 5. low tonality vs. high tonality; 6. strident vs. mellow; 7. nasal vs. nonnasal; 8. continuant vs. interrupted; 9. voiced vs. voiceless; 10. sharpened vs. plain; 11. accented vs. unaccented. Left branches represent minus values, and right branches, plus values for the particular feature.

An argument for specification by branching trees

An argument for branching trees

Before continuing, we should ask why the branching tree occurs at all: why is it necessary to specify features in this way?

In *SPR* Halle makes an argument on behalf of branching trees; this is the first such argument we have found in the literature.

Halle (1959) argued that phonological features must be ordered into a hierarchy because this is the only way to ensure that segments are kept properly distinct.

The Distinctness Condition

Specifically, he proposed (1959: 32) that phonemes must meet the Distinctness Condition:

The Distinctness Condition

Segment-type {A} will be said to be different from segment-type {B}, if and only if at least one feature which is phonemic in both, has a different value in {A} than in {B}; i.e., plus in the former and minus in the latter, or vice versa.

This formulation is designed to disallow contrasts involving a **zero value** of a feature.

How do we establish contrasts?

Consider the typical sub-inventory /p, b, m/ shown below, and suppose we characterize it in terms of two binary features, [±voiced] and [±nasal].

In terms of full specifications, /p/ is [-voiced, -nasal], /b/ is [+voiced, -nasal], and /m/ is [+voiced, +nasal].

Which of these features is contrastive? Many people reason as follows:

	/p/	/b/	/m/
[voiced]	–	+	+
[nasal]	–	–	+

How do we establish contrasts?

We observe that /p/ and /b/ are distinguished only by [voiced]; so these specifications **must** be contrastive.

Similarly, /b/ and /m/ are distinguished only by [nasal]; these specifications must **also** be contrastive.

What about the uncircled specifications? These are predictable from the circled ones:







	/p/	/b/	/m/
[voiced]	⊖	⊕	+
[nasal]	−	⊖	⊕

How do we establish contrasts?

Since /p/ is the only [-voiced] phoneme in this inventory, its specification for [nasal] is predictable, hence redundant. We can write a rule or constraint:

Similarly, /m/ is the only [+nasal] phoneme, so its specification for [voiced] is redundant:

This is a still-popular way of thinking about contrastive specifications; we can call it the ‘minimal contrast’ (MC) approach (Padgett 2003, Calabrese 2005, Campos-Astorkiza 2009, Nevins 2010 explicitly, and many others implicitly).

	/p/	/b/	/m/
[voiced]			
[nasal]			

If [-voiced], then [-nasal]

If [+nasal], then [+voiced]

Minimal Contrast (MC)

According to the definition proposed by Nevins (2010: 98):

A segment S with specification $[\alpha F]$ is **contrastive** for F

if there is another segment S' in the inventory that is featurally identical to S , except that it is $[-\alpha F]$.

	/p/	/b/	/m/
[voiced]	⊖	⊕	+
[nasal]	-	-	+

The condition is met by [voiced] on /p/ and /b/.

Minimal Contrast (MC)

The condition is **not** met by [-nasal] on /p/, because there is no other segment in this inventory that differs from /p/ **only** by this feature.

	/p/	/b/	/m/
[voiced]	-	+	+
[nasal]	⊖	-	+

Minimal Contrast (MC)

The condition is **not** met by [-nasal] on /p/, because there is no other segment in this inventory that differs from /p/ **only** by this feature.

For this feature to be contrastive on /p/ there would have to be a voiceless nasal /ᵐ/ in this inventory.

	/p/	/b/	/m/	/ᵐ/
[voiced]	-	+	+	-
[nasal]	⊖	-	+	⊕

Minimal Contrast (MC)

The same would be required by Minimal Contrast to make [+voiced] contrastive on /m/.

	/p/	/b/	/m/	/ḿ /
[voiced]	–	+	+	–
[nasal]	–	–	+	+

Minimal Contrast (MC)

In the absence of voiceless nasal /m̥/, Minimal Contrast gives us the specifications below.

According to the Distinctness Condition, however, this set of specifications is **not** properly contrastive: Minimal Contrast is the wrong definition of contrast.

	/p/	/b/	/m/
[voiced]	–	+	
[nasal]		–	+

Let's see what the Distinctness Condition says about these specifications.

The Distinctness Condition

Segment-type {A} will be said to be different from segment-type {B}, if and only if at least one feature which is phonemic in both, has a different value in {A} than in {B}; i.e., plus in the former and minus in the latter, or vice versa.

According to the Distinctness Condition, /p/ is 'different from' /b/, because /p/ is [-voiced] and /b/ is [+voiced].

	/p/	/b/	/m/
[voiced]	-	+	
[nasal]		-	+

The Distinctness Condition

Segment-type {A} will be said to be different from segment-type {B}, if and only if at least one feature which is phonemic in both, has a different value in {A} than in {B}; i.e., plus in the former and minus in the latter, or vice versa.

According to the Distinctness Condition, /p/ is 'different from' /b/, because /p/ is [-voiced] and /b/ is [+voiced].

Similarly, /b/ is 'different from' /m/, because /b/ is [-nasal] and /m/ is [+nasal].

	/p/	/b/	/m/
[voiced]	-	+	
[nasal]		-	+

The Distinctness Condition

Segment-type {A} will be said to be different from segment-type {B}, if and only if at least one feature which is phonemic in both, has a different value in {A} than in {B}; i.e., plus in the former and minus in the latter, or vice versa.

But /p/ is **not** 'different from' /m/: where one has a feature, the other has no specification.

Therefore, these specifications are not properly contrastive.

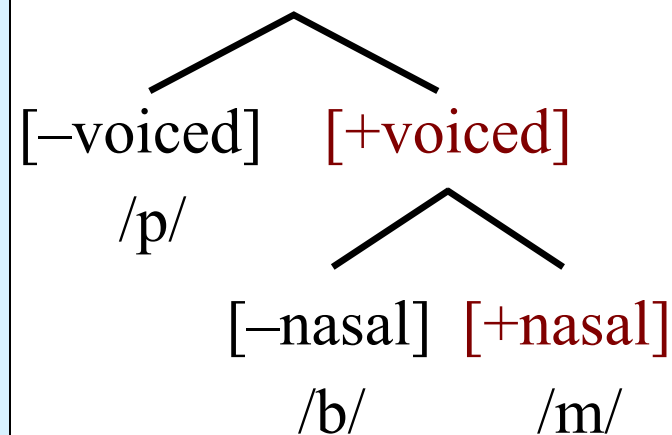
	/p/	/b/	/m/
[voiced]	–	+	
[nasal]		–	+

The Distinctness Condition

They violate the Distinctness Condition because no feature hierarchy yields this result.

If we order [voiced] > [nasal], we generate an 'extra' specification on /m/.

	/p/	/b/	/m/
[voiced]	–	+	+
[nasal]		–	+



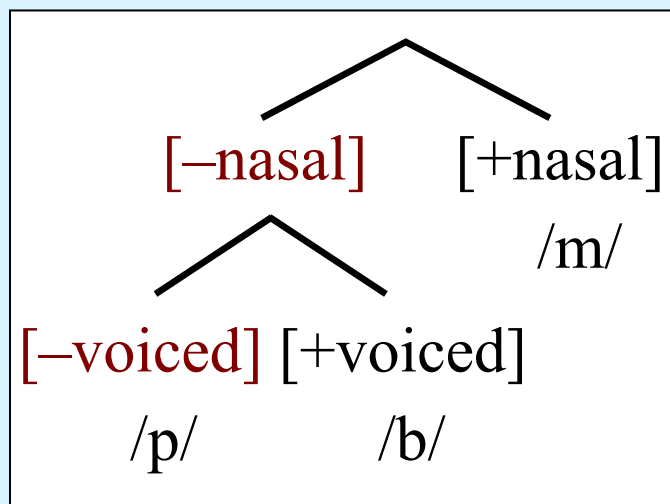
The Distinctness Condition

The specifications below violate the Distinctness Condition because no feature hierarchy yields this result.

If we order [voiced] > [nasal], we generate an 'extra' specification on /m/.

If we order [nasal] > [voiced], we generate an 'extra' specification on /p/.

	/p/	/b/	/m/
[voiced]	–	+	
[nasal]	⊖	–	+

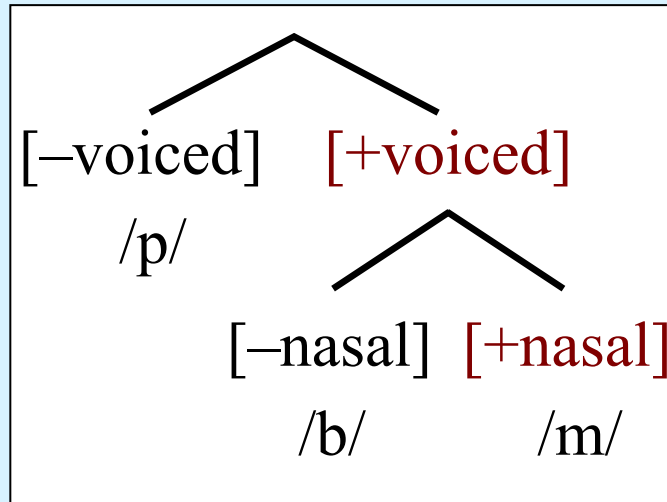


Contrastive ≠ Unpredictable

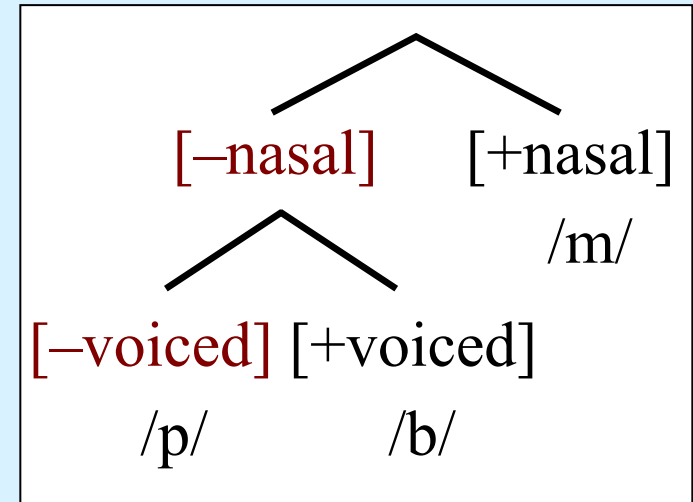
Either of the specifications below is properly contrastive.

Note that in a hierarchical approach, a contrastive feature is not necessarily unpredictable.

[voiced] > [nasal]



[nasal] > [voiced]



	/p/	/b/	/m/
[voiced]	-	+	+
[nasal]		-	+

	/p/	/b/	/m/
[voiced]	-	+	
[nasal]	-	-	+

Therefore, according to *SPR*, to ensure that all the phonemes of a language are distinct from one another, it is necessary that their feature specifications must be generable by a branching

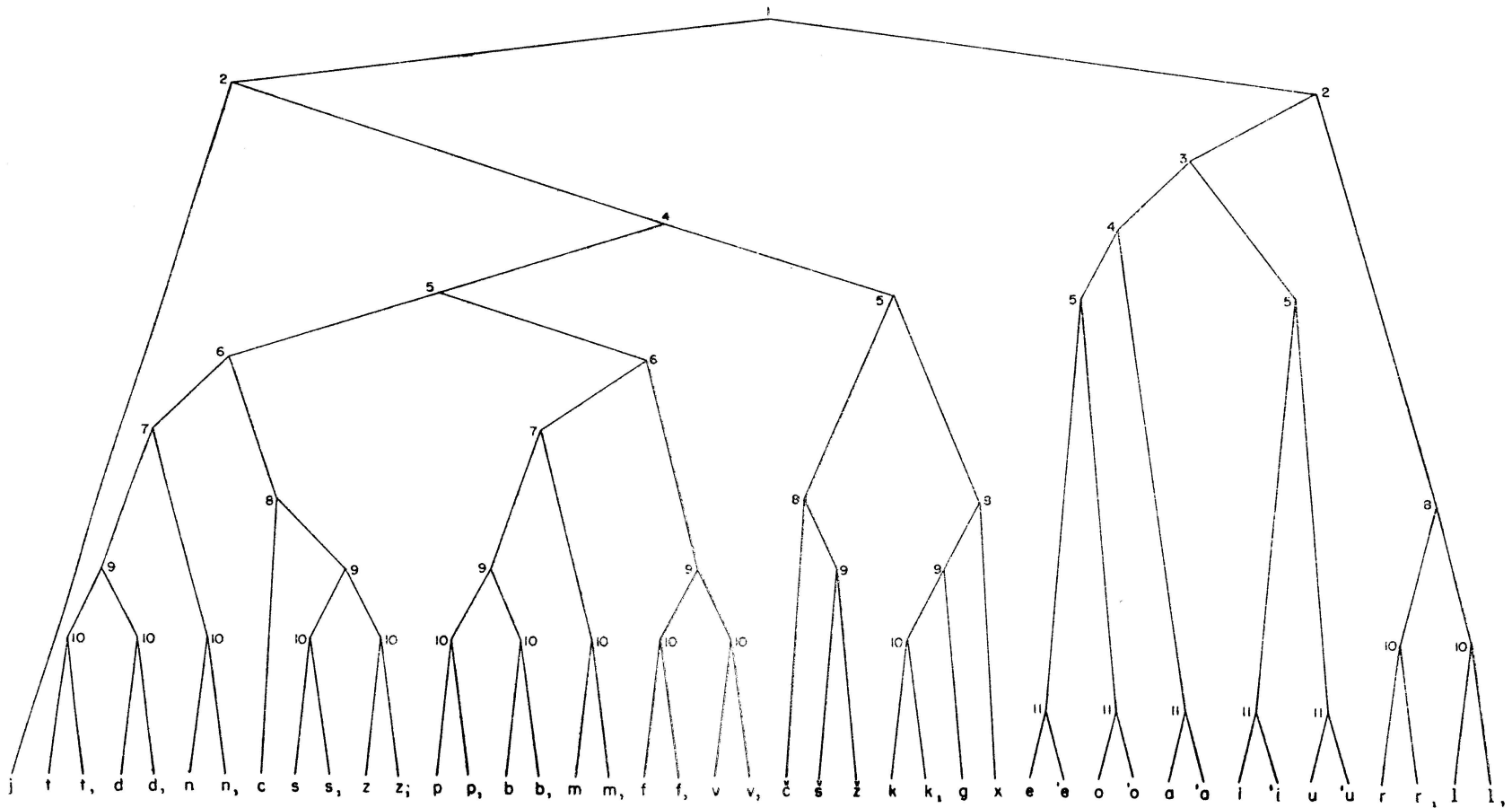


Fig. I-1. Branching diagram representing the morphonemes of Russian. The numbers with which each node is labelled refer to the different features, as follows: 1. vocalic vs. nonvocalic; 2. consonantal vs. nonconsonantal; 3. diffuse vs. nondiffuse; 4. compact vs. noncompact; 5. low tonality vs. high tonality; 6. strident vs. mellow; 7. nasal vs. nonnasal; 8. continuant vs. interrupted; 9. voiced vs. voiceless; 10. sharpened vs. plain; 11. accented vs. unaccented. Left branches represent minus values, and right branches, plus values for the particular feature.

The importance of feature ordering

Feature ordering

As the previous demonstration shows, in a hierarchical approach to feature specification the **ordering** of the features is crucial.

For a given inventory of segments, different feature orders can result in different contrastive specifications

The potentially dramatic effects of ordering on specification can be illustrated with one section of the Russian tree.

Effects of feature ordering

The segments in the red box are [-vocalic] and [+consonantal].

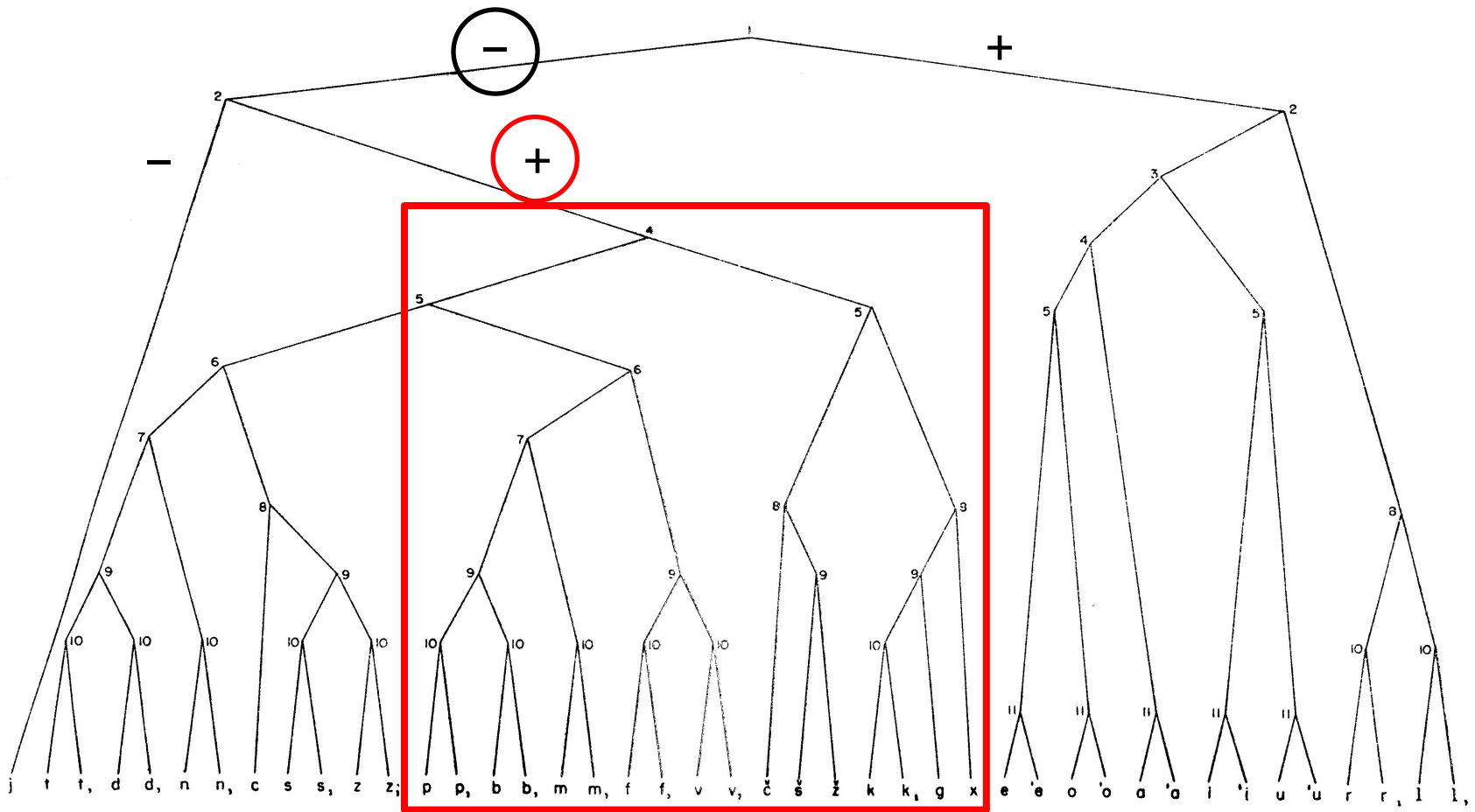
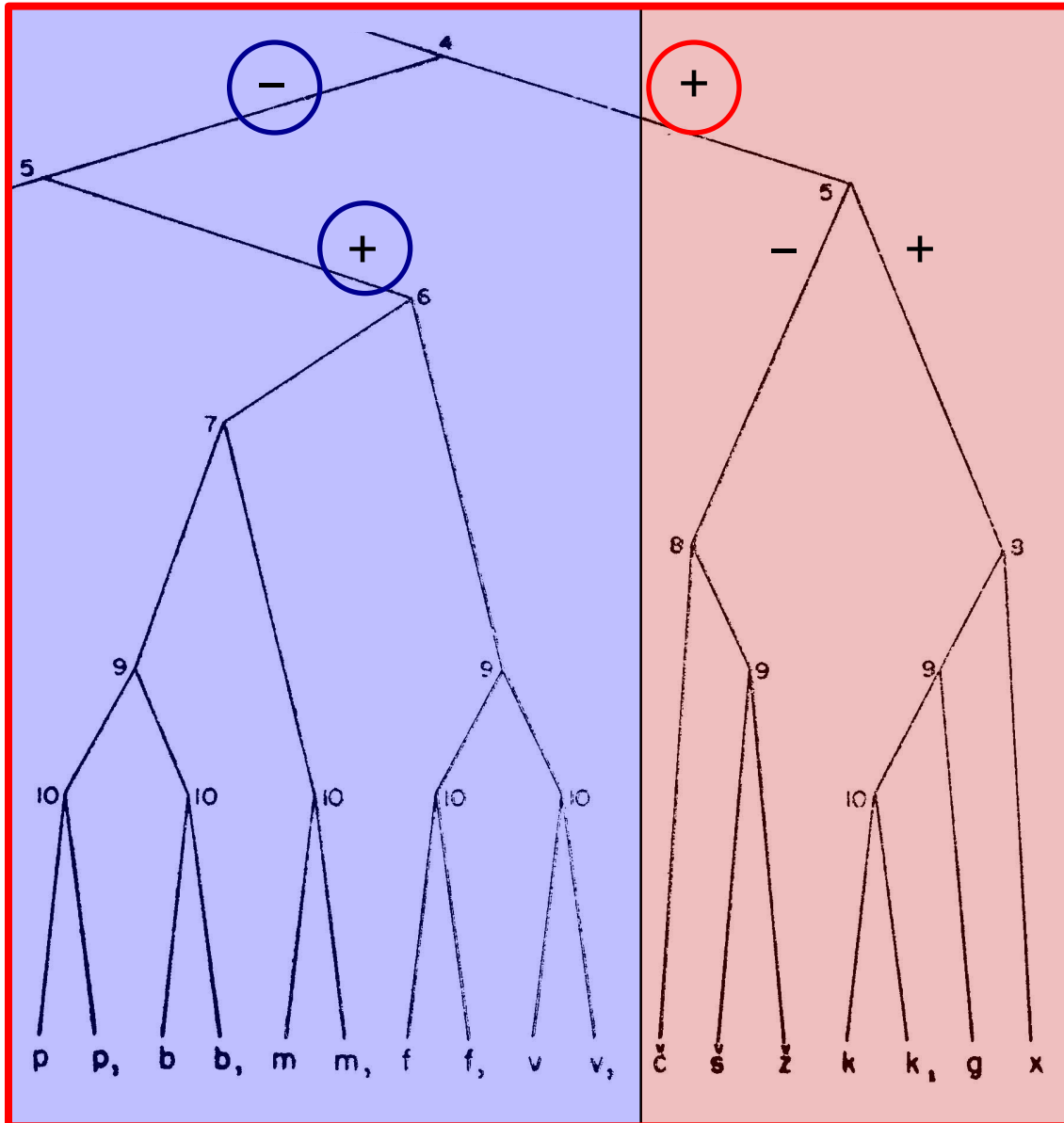


Fig. I-1. Branching diagram representing the morphemes of Russian. The numbers with which each node is labelled refer to the different features, as follows: 1. vocalic vs. nonvocalic; 2. consonantal vs. nonconsonantal; 3. diffuse vs. nondiffuse; 4. compact vs. noncompact; 5. low tonality vs. high tonality; 6. strident vs. mellow; 7. nasal vs. nonnasal; 8. continuant vs. interrupted; 9. voiced vs. voiceless; 10. sharpened vs. plain; 11. accented vs. unaccented. Left branches represent minus values, and right branches, plus values for the particular feature.

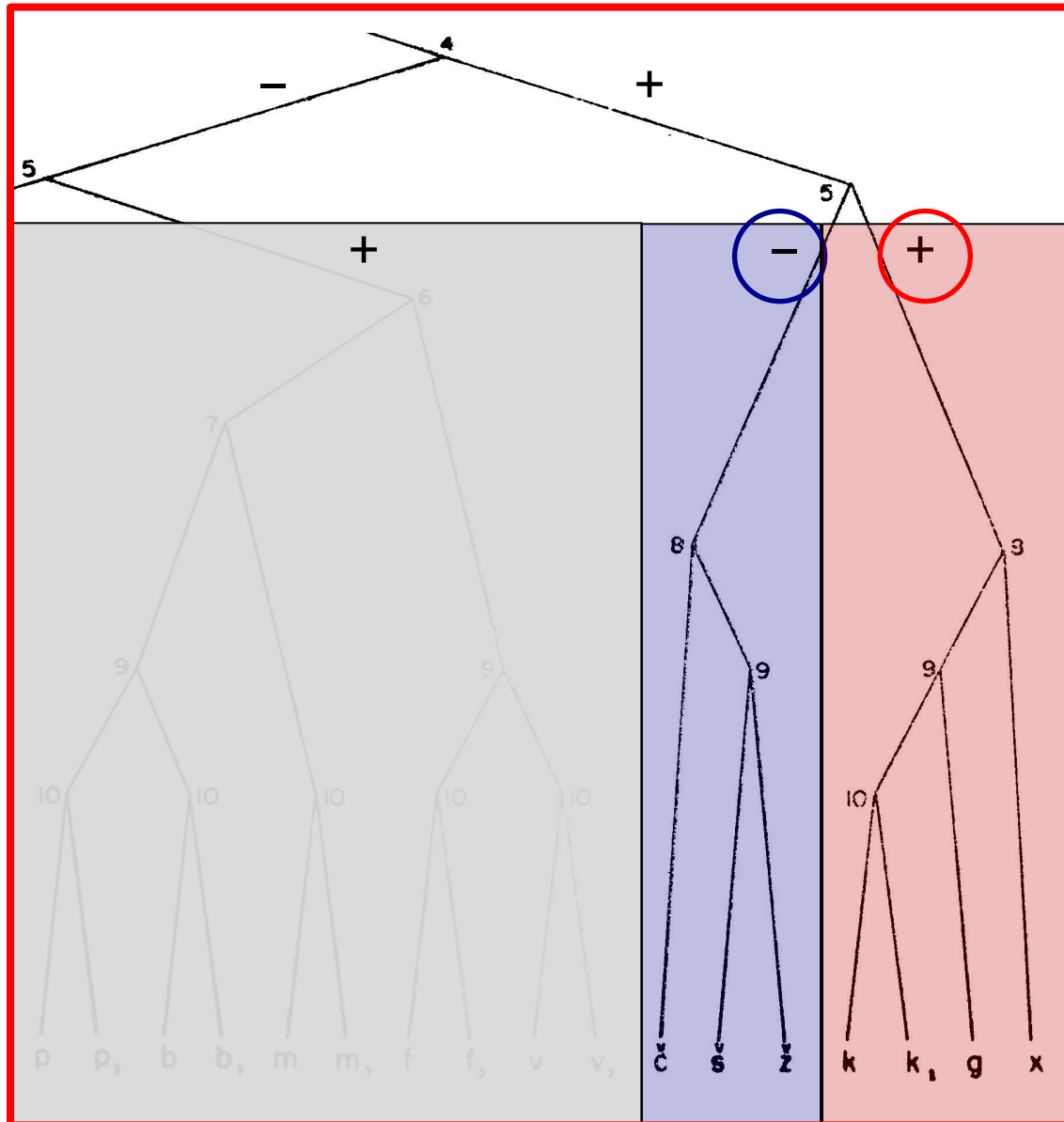
In a larger view:



In blue under [-compact] and [+low tonality] are the labial consonants (stops, nasals, and fricatives).

In red under [+compact] are the posterior coronal and velar consonants.

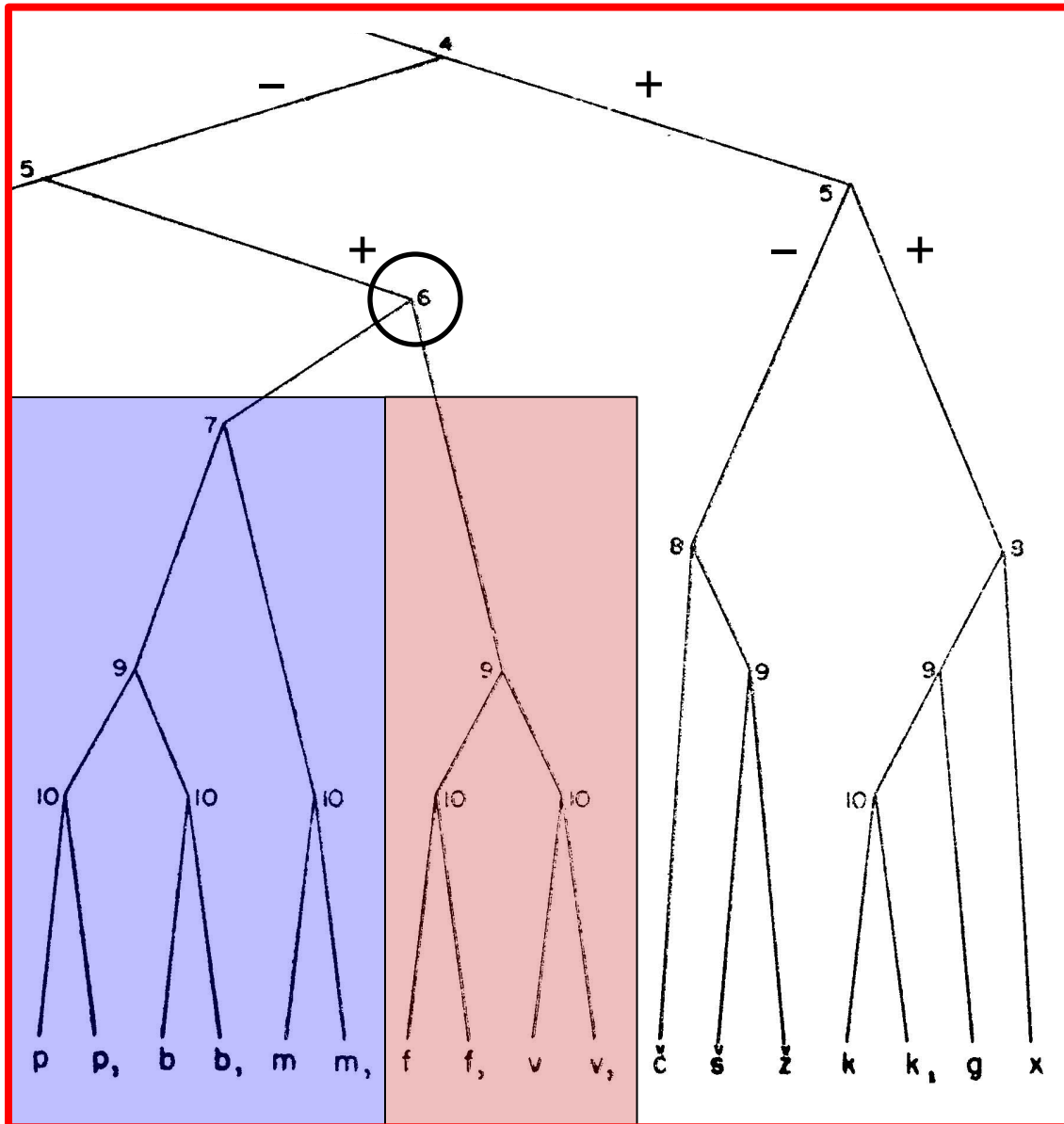
Effects of feature ordering



The posterior coronals
č, š, ž are [-low tonality]
(in blue).

The velars are [+low
tonality] (in red).

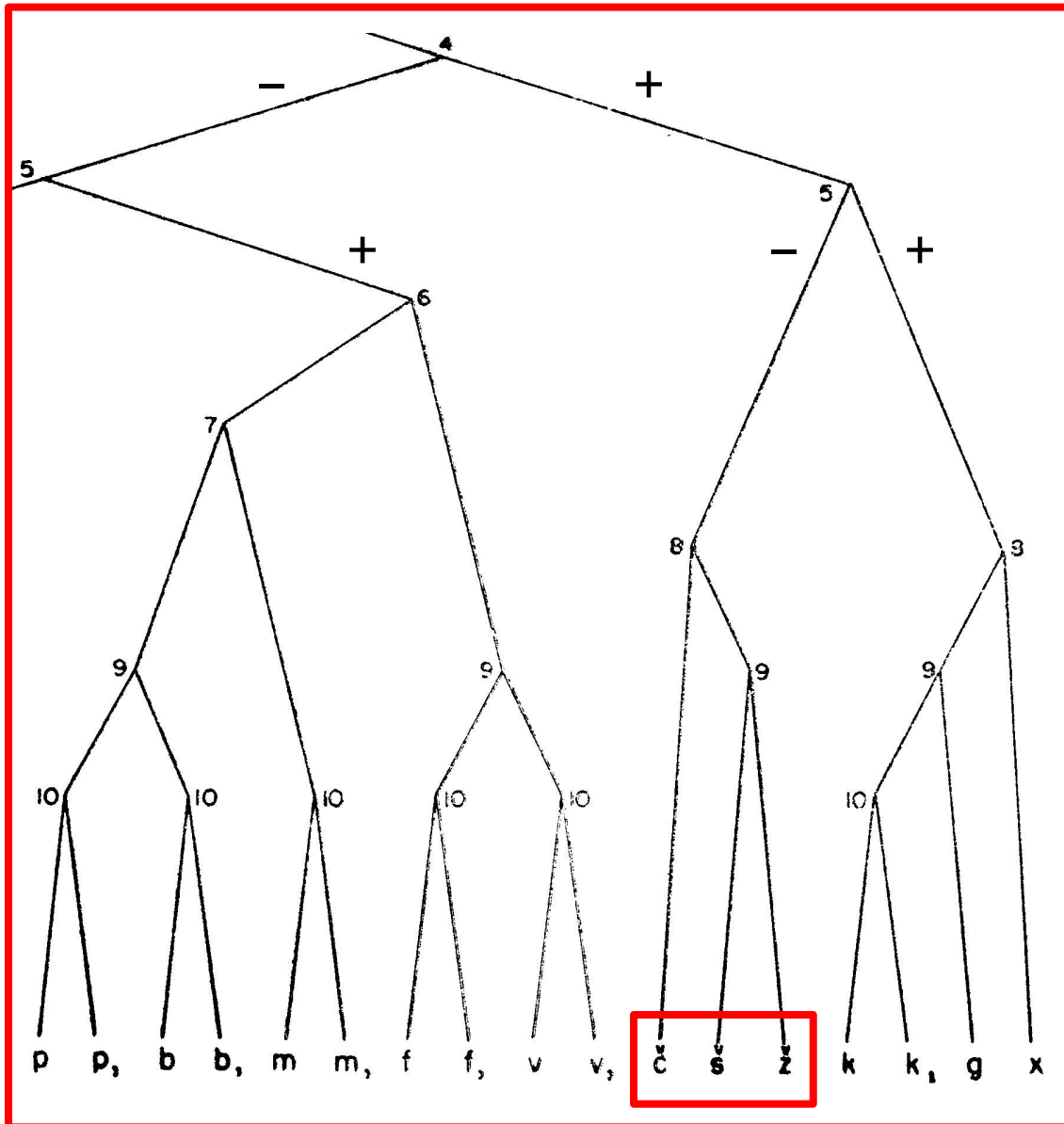
Effects of feature ordering



Feature 6 (circled) stands for [strident]. It applies within the labials to distinguish [-strident] stops from [+strident] fricatives.

Consequently, feature 8, [continuant], does not apply to the labials because the stops and fricatives have already been distinguished by [strident].

Effects of feature ordering



Perhaps unexpectedly, [strident] does not apply to the prototypically strident \check{c} , \check{s} , \check{z} (IPA /tʃ, ʃ, ʒ/) because they already form a separate group.

All these specifications could be altered if the features were ordered differently.

Rationale for feature hierarchies

Given the importance of the ordering of features in determining what the feature specifications are, it is important to know why Halle (1959) chose to order the features the way he did.

Halle (1959: 29–30) provides the rationale, in his Condition (5):

Condition (5)

In phonological representations the number of specified features is consistently reduced to a minimum compatible with satisfying Conditions (3) and (4).

Roughly speaking, Conditions (3) and (4) require that the phonological description meet basic conditions of adequacy (we will come back to this later).

Rationale for feature hierarchies: Minimality of specifications

That is, the main criterion for deciding on how to order features in *SPR* is to minimize the number of feature specifications.

We will call this the **Minimality Principle**, which can be restated as follows:

Minimality Principle for Feature Ordering

The criterion for ordering features into a hierarchy is to minimize redundancy in phonological representations and to maximize the amount of information conveyed by each feature.

Rationale for feature hierarchies: Minimality of specifications

Halle's concern with Minimality is reflected in his observation (1959: 44–5) that his analysis of Russian contains 43 phonemes specified by 271 feature specifications, or 6.3 distinctive feature statements per phoneme.

(Actually, by my count his analysis has 272 specifications—this will turn out to be significant!)

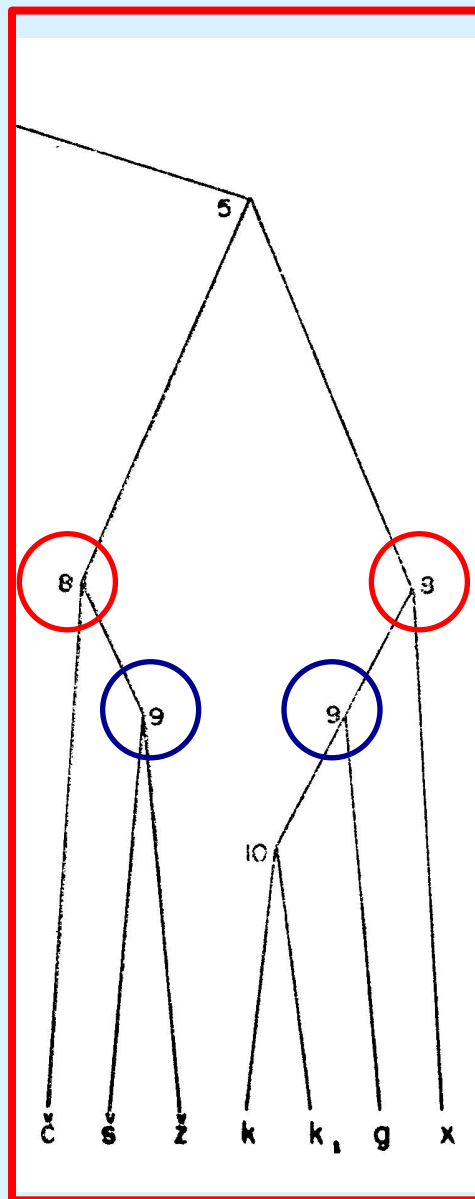
He compares 6.3 with the lower limit of $\log_2 43 = 5.26$ specifications, which would represent the most efficiently branching tree for 43 phonemes ($\log_2 43 = 5.43$ by my calculator).

The principle of Minimality can lead to feature orderings that may strike us as counter-intuitive, or orderings that do not closely reflect phonological patterning.

Effects of feature ordering

The ordering of two features in the part of the tree we looked at earlier had momentous consequences for the development of phonological theory.

These are features 8, $[\pm\text{continuant}]$, in the red circles, and 9, $[\pm\text{voiced}]$, in the blue circles.

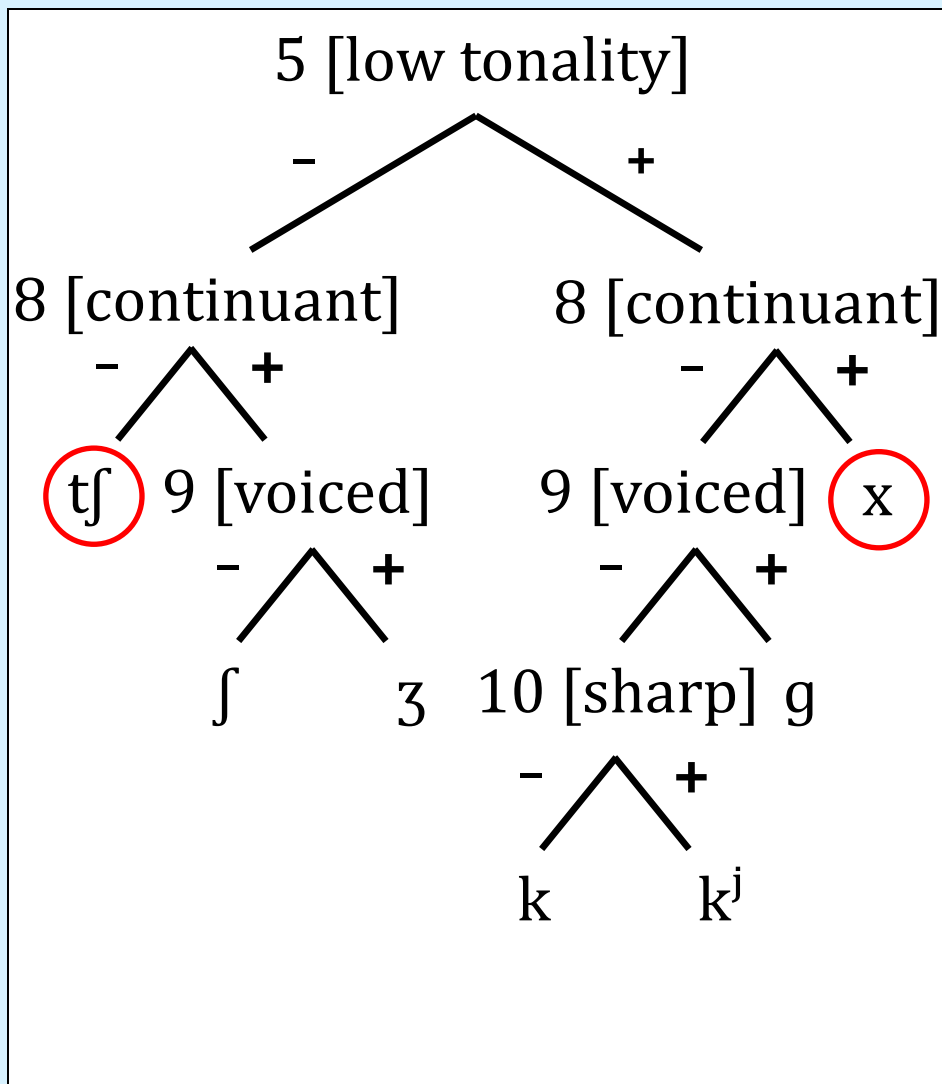


$[\text{continuant}]$ is ordered above $[\text{voiced}]$; every phoneme in this diagram has a specification for $[\text{continuant}]$.

The same is not the case for $[\text{voiced}]$!

The 'unpaired' phonemes

In a larger and more legible view:



In the ordering shown, /tʃ/ and /x/ are unspecified for [voiced].

But as Halle famously pointed out, these segments (as well as /ts/) behave phonologically like other voiceless obstruents with respect to voicing assimilation.

Derivation with ‘unpaired’ phoneme

In *SPR*, this is accounted for by the following rules:

Rule P 1b: Unless followed by an obstruent, /ts/, /tʃ/, and /x/ are voiceless.

Rule P 3a: If an obstruent cluster is followed [...] by a sonorant, then with regard to voicing the cluster conforms to the last segment.
Regressive Voicing Assimilation (RVA)

Derivation with 'unpaired' phoneme

An example is the derivation of [safxós] 'state farm' from /sovxoz/.
[voiced] is not specified on /x/ by the branching tree.

By Rule P1b its \emptyset is immediately changed to [-voiced].

Underlying	/s	o	v	x	o	z/
[voiced]	—		+	\emptyset		+
Rule P1b	/s	o	v	x	o	z/
[voiced]	—		+	—		+

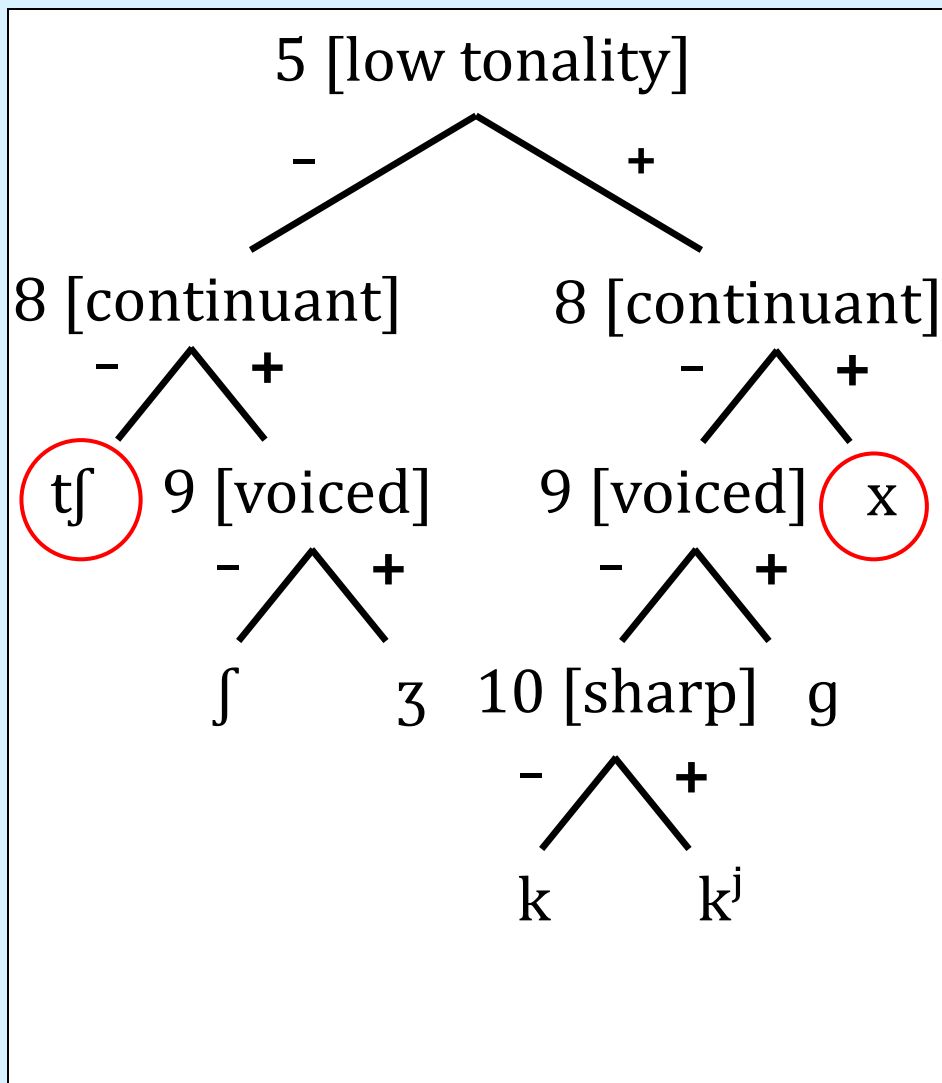
Derivation with 'unpaired' phoneme

Thereafter, /x/ is like any other contrastively voiceless segment, and it acts as such to condition RVA (Rule P3a) on /v/, which devoices to [f].

Other rules result in the surface form.

Underlying	/s	o	v	x	o	z/
[voiced]	—		+	∅		+
Rule P1b	/s	o	v	x	o	z/
[voiced]	—		+	—		+
Rule P3a	/s	o	f	x	o	z/
[voiced]	—		—	—		+
Other rules	/s	a	f	x	ó	s/
[voiced]	—		—	—		—

The 'unpaired' phonemes



So although 'unpaired' /tʃ, x, ts/ are not specified for [±voiced] underlyingly by the branching tree, they are assigned [-voiced] early in the derivation, and subsequently behave like other voiceless segments.

This analysis formed the basis of Halle's famous argument against the structuralist, or 'taxonomic', phoneme.

Halle's argument against the structuralist ('taxonomic') phoneme

The taxonomic phoneme

From the 1930s to the 1950s, American linguists who were followers of Leonard Bloomfield attempted to make the definition of the phoneme more precise than it had been.

In practice, this amounted to placing an increasing number of restrictions on the relationship between the phonemic level and the phonetic signal.

One such restriction concerned ‘phonemic overlapping’ (Bloch 1941): one sound (a phone) may not correspond to two different phonemes in the same context.

The taxonomic phoneme

For example, an underlying /tʃ/ can surface as [tʃ], or, when subject to Regressive Voicing Assimilation (RVA), as [dʒ].

Because there is no underlying */dʒ/ in Russian, both [tʃ] and [dʒ] are unambiguous allophones of a single phoneme /tʃ/. That's OK.

Consider now an underlying /t/: just like /tʃ/, it can surface as voiceless [t], or as [d] when voiced by RVA.

Phonemic Representation

/tʃ/

/t/

Phonetic Representation

[tʃ] [dʒ]

[t] [d]

The taxonomic phoneme

In this case, however, there is another source for [d] in Russian, namely underlying /d/.

/d/ is in contrast with /t/ in some positions, but the contrast is neutralized in RVA environments.

(The overlap is even worse than this, because RVA can devoice /d/ to [t] if the final obstruent in the cluster is voiceless; so both phonetic [t] and [d] could be allophones of either /t/ or /d/.)

Phonemic Representation

/tʃ/

/t/ /d/

Phonetic Representation

[tʃ] [dʒ]

[t] [d]

The taxonomic phoneme

Consequently, without additional information from the meaning or morphology, we wouldn't know which phoneme any given [t] or [d] belongs to in neutralization contexts.

Similarly, phonetic [k] or [g] could derive from either /k/ or /g/, [p] or [b] could belong to /p/ or /b/, and so on for every such pair of voiced and voiceless obstruents.

This type of overlapping was not allowed by most American structuralist phonologists in the 1950s.

Phonemic Representation

/tʃ/

/t/ /d/

Phonetic Representation

[tʃ] [dʒ]

[t] [d]



The taxonomic phoneme

The only solution, in this theory, is to require that every phonetic [t] must be an allophone only of the phoneme /t/, and every [d] is an allophone only of the phoneme /d/.

Phonemic Representation

/tʃ/

/t/

/d/

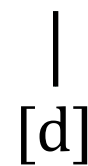
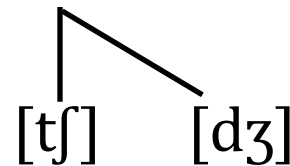
Phonetic Representation

[tʃ]

[dʒ]

[t]

[d]



The taxonomic phoneme

The only solution, in this theory, is to require that every phonetic [t] must be an allophone only of the phoneme /t/, and every [d] is an allophone only of the phoneme /d/.

The fact that *t* alternates with *d* must be expressed in a mapping between the morphophonemic (lexical) level and the phonemic level; overlapping is permitted between morphophonemes.

Morphophonemes

{tʃ}

{t}

{d}

Phonemic Representation

/tʃ/

/t/

/d/

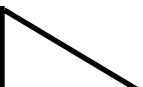
Phonetic Representation

[tʃ]

[dʒ]

[t]

[d]



The taxonomic phoneme

Morphophonemic
Representations: {tʃ} {t} {d}

*Morphophonemic
rules*

RVA 1



Phonemic
Reps.: /tʃ/ /d/ /t/

*Phonemic
rules*

RVA 2



Phonetic
Reps.: [tʃ] [dʒ] [d] [t]

Halle showed that the same rule of Regressive Voicing Assimilation (RVA) that applies in the morphophonemic component to change one morphophoneme into another (say, {t} into /d/)...

must apply again in the component that turns phonemes into allophones (e.g. /tʃ/ to [dʒ]).

This duplication is made necessary only to satisfy the constraints on overlapping applied to the phonemic level.

Against the taxonomic phoneme

Underlying Lexical
Representations: {tʃ} {t} {d}

*Phonological
rules*

RVA

Phonetic
Reps.: [tʃ] [dʒ] [d] [t]

Thus, the grammar can be simplified by rejecting the phonemic level and allowing a smooth transition from underlying lexical representations to surface phonetic representations, with no intermediate level (such as the old phonemic level) accorded special status

The demise of underspecification and the branching trees in generative phonology

The end of underspecification

Although not much (if at all) remarked upon at the time, there was a further far-reaching consequence of the *SPR* analysis:

The distinction between *contrastive* and *non-contrastive* features became unimportant as far as the workings of the phonology are concerned, as illustrated by the derivations shown earlier.

After all, if a phoneme (like /x/ or /tʃ/) is not assigned a contrastive feature by the branching tree, it can nevertheless acquire that feature in the course of the derivation, whenever it is needed.

The end of underspecification

This made the whole notion of contrastive underspecification vulnerable to arguments by Lightner (1963) and Stanley (1967).

Underspecification was abandoned altogether in Chomsky and Halle's *The Sound Pattern of English* (*SPE*, 1968), along with the branching trees that generate them.

The result was that language-particular feature contrasts did not play a role in the theory of generative grammar that developed from *SPE*.

Consequences of Minimality

As a result, the branching trees, or more properly, contrastive feature hierarchies, disappeared from generative phonology for a generation.

This whole sequence of events began with the decision to order features on the basis of the Minimality Principle; but Halle could have taken another path...

*And both that morning equally lay
In leaves no step had trodden black.
Oh, I kept the first for another day!
Yet knowing how way leads on to way,
I doubted if I should ever come back.*

Prague School phonology: Contrastive properties and the original rationale for ordering features into a hierarchy

Origins of contrastive feature hierarchies

Although the main criterion for ordering features in *SPR* is Minimality, that is, minimizing the number of feature specifications, this was not the original rationale.

The notion of specifying phonemes in terms of contrastive features ordered into hierarchies can be traced back to the work of the Prague School phonologists, Roman Jakobson and N. S. Trubetzkoy, in the 1920s and 1930s.

Though branching trees did not yet make an appearance, they implicitly underlie some of their analyses (Dresher 2009, 2016).

Active properties are contrastive

An idea that can be traced to the beginnings of modern phonology is that only some properties of a segment are *active*, or *relevant* (Trubetzkoy 1939) to the phonology, and these are the *distinctive*, or *contrastive*, properties.

An early expression of this idea can be found in Jakobson's (1962 [1931]) discussion of the difference between the Czech and Slovak vowel systems.

In this work, as well as in later publications, such as Jakobson & Lotz 1949 and Jakobson, Fant & Halle 1952, features are specified (or unspecified) in order to account for synchronic phonological behaviour, or patterns of loanword adaptation.

Another rationale for feature hierarchies: Phonological activity

It is thus **phonological activity** that determines what the features are, and how they are ordered, where feature activity can be defined as follows (based on Clements 2001: 77):

Feature activity

A feature can be said to be **active** if it plays a role in the phonological computation; that is, if it is required for the expression of phonological regularities in a language, including both static phonotactic patterns and patterns of alternation.

Another rationale for feature hierarchies: Phonological activity

That is, the original rationale for ordering features was not Minimality, but what we call the **Activity Principle**:

Activity Principle for Feature Ordering

The criterion for ordering features into a hierarchy is to reflect patterns of phonological activity in a language.

The original intuition behind feature hierarchies is that there is a connection between contrast and activity (Dresher 2009, 2015).

Another rationale for feature hierarchies: Phonological activity

In fact it can be shown that phonological activity still played an important role in choosing the ordering of features in *SPR*.

Recall that Halle's analysis requires 272 feature specifications.

There are other orderings that use the same features as Halle uses that result in far fewer specifications; however, one has to be willing to accept some strange-looking groupings of segments.

Thus, the first feature in Halle's ordering is [vocalic], which divides all phonemes into [-vocalic] glides and consonants and [+vocalic] vowels and liquids.

Notice that $[\pm\text{vocalic}]$ divides the inventory unequally: there are 29 $[-\text{vocalic}]$ phonemes on the left in blue, and 14 $[\text{+vocalic}]$ vowels and liquids in red.

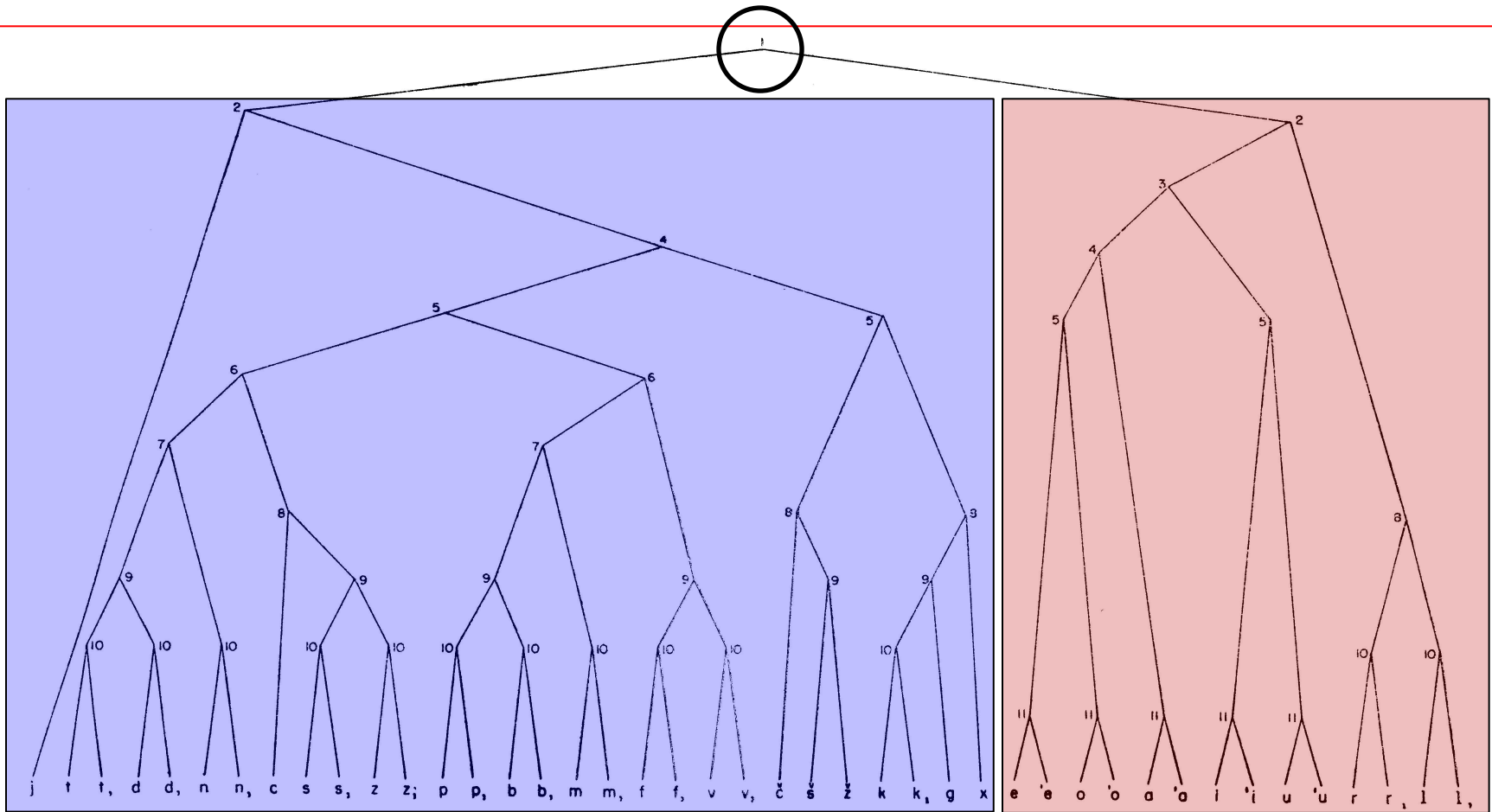
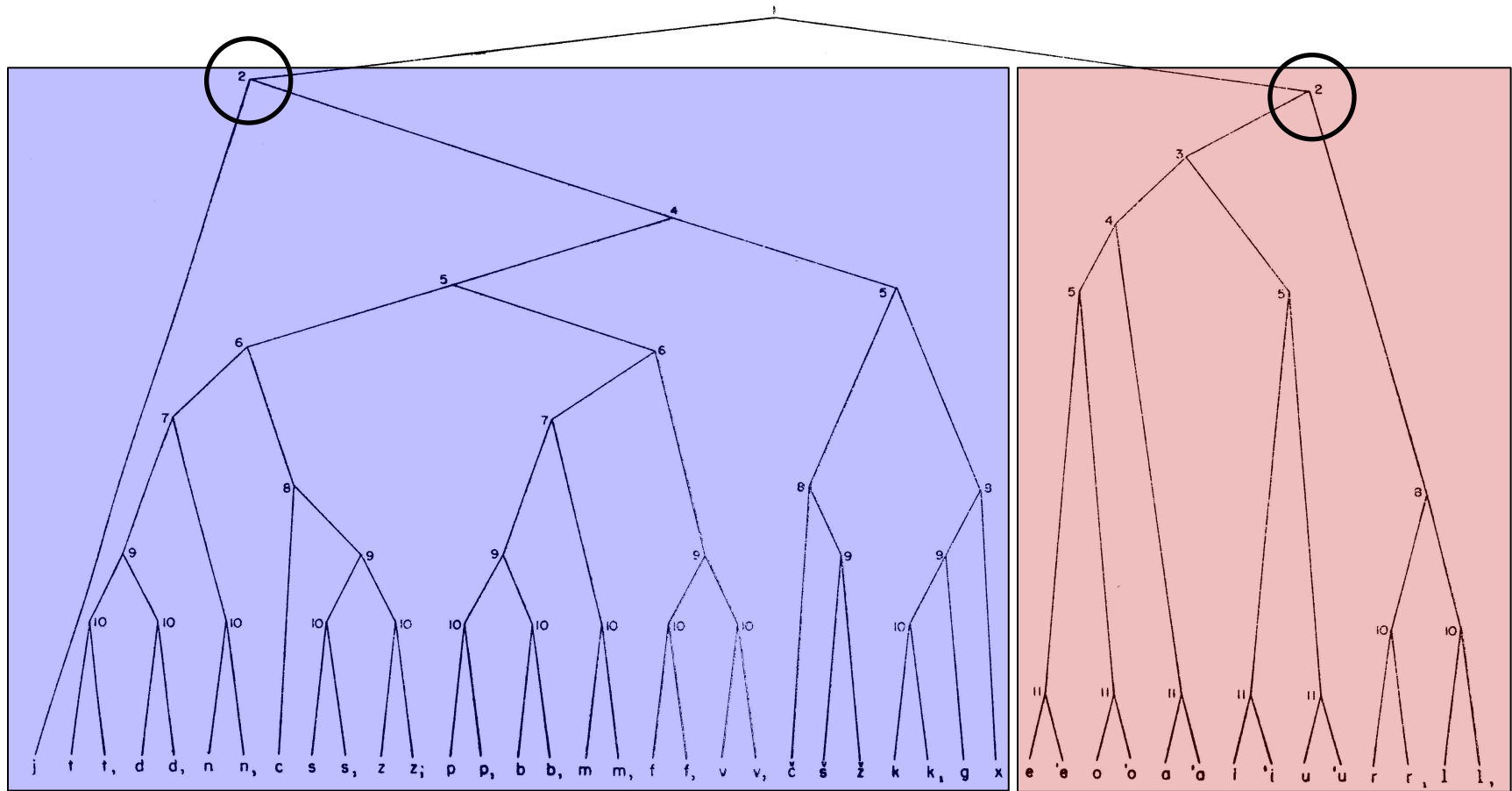


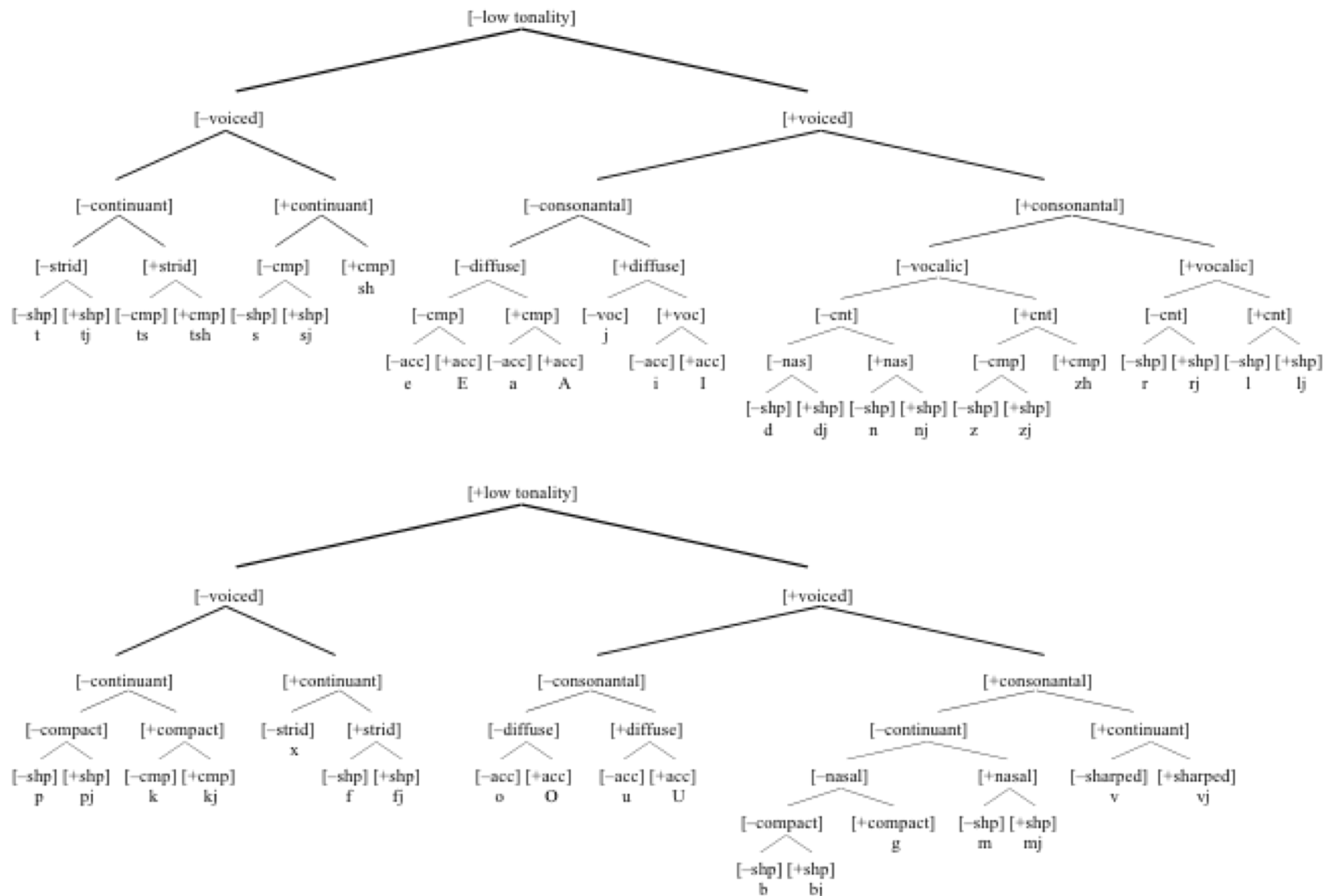
Fig. I-1. Branching diagram representing the morphonemes of Russian. The numbers with which each node is labelled refer to the different features, as follows: 1. vocalic vs. nonvocalic; 2. consonantal vs. nonconsonantal; 3. diffuse vs. nondiffuse; 4. compact vs. noncompact; 5. low tonality vs. high tonality; 6. strident vs. mellow; 7. nasal vs. nonnasal; 8. continuant vs. interrupted; 9. voiced vs. voiceless; 10. sharpened vs. plain; 11. accented vs. unaccented. Left branches represent minus values, and right branches, plus values for the particular feature.

The next feature, $[\pm\text{consonantal}]$, makes very unequal divisions: on the left, it divides between /j/ and the other 28 consonants; on the right, it opposes 10 vowels against 4 liquids.

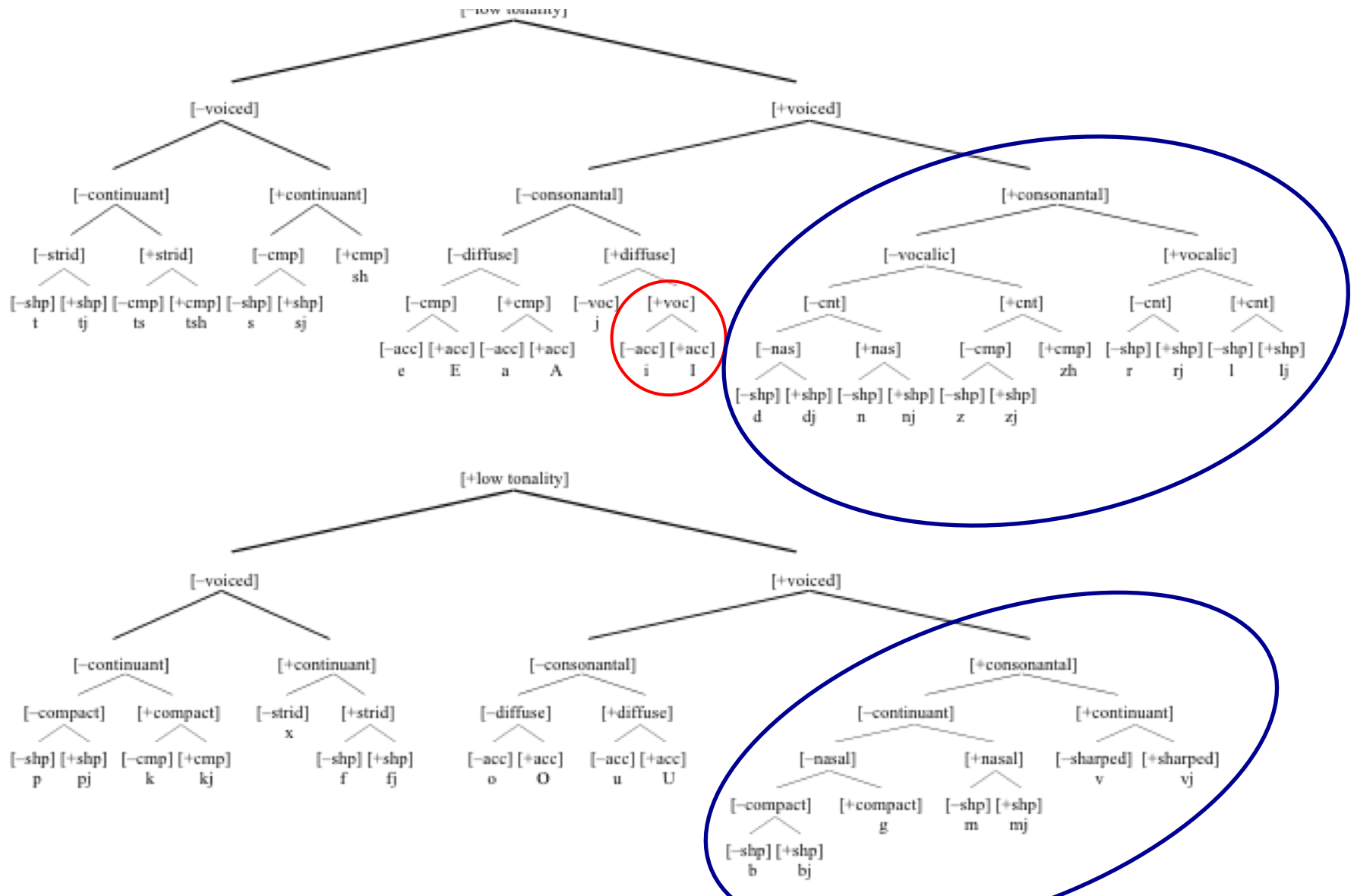


Unequal divisions make for relatively inefficient trees, and increase the number of contrastive specifications.

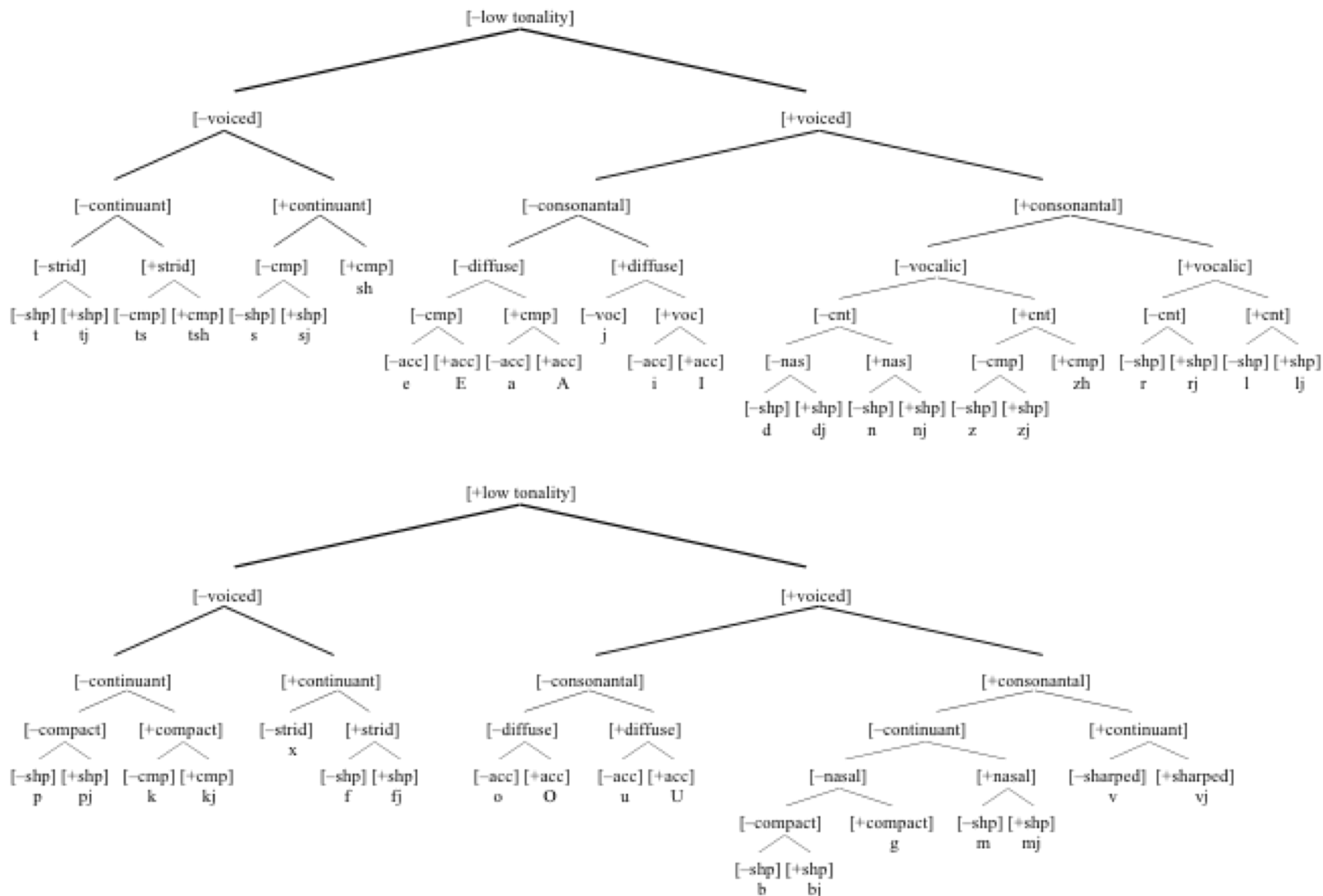
Here is a tree with 243 specifications, 29 fewer than Halle's.
 The top feature, [\pm low tonality], divides the tree 25 against 18.



Only 2 vowels (circled) are [+vocalic], and only some consonants are [+consonantal]. The groupings are not very natural.



The fact that Halle did not consider this ordering is evidence that he did not follow Minimality to the exclusion of Activity.



Rationale for feature hierarchies

Recall Halle's Condition 5, his version of Minimality:

Condition (5)

In phonological representations the number of specified features is consistently reduced to a minimum **compatible with satisfying Conditions (3) and (4)**.

Here is Condition (4) (Halle 1959: 24):

Condition (4)

The phonological description must be appropriately integrated into the grammar of the language. Particularly, in selecting phonological representations of individual morphemes, these must be chosen so as to yield simple statements of all grammatical operations ... in which they may be involved.

Rationale for feature hierarchies

Condition (5)

In phonological representations the number of specified features is consistently reduced to a minimum **compatible with satisfying Conditions (3) and (4)**.

In other words, the number of feature specifications is reduced to a minimum **as long as the resulting representations make sense in terms of the grammar of the language**.

That is, when assigning contrastive features to segments we are guided by Minimality as long as the result makes sense with respect to Activity.

The Contrastivist Hypothesis

The connection between contrast and activity is made explicit by what Hall (2007: 20) calls the **Contrastivist Hypothesis**:

The Contrastivist Hypothesis

The phonological component of a language L operates only on those features which are necessary to distinguish the phonemes of L from one another.

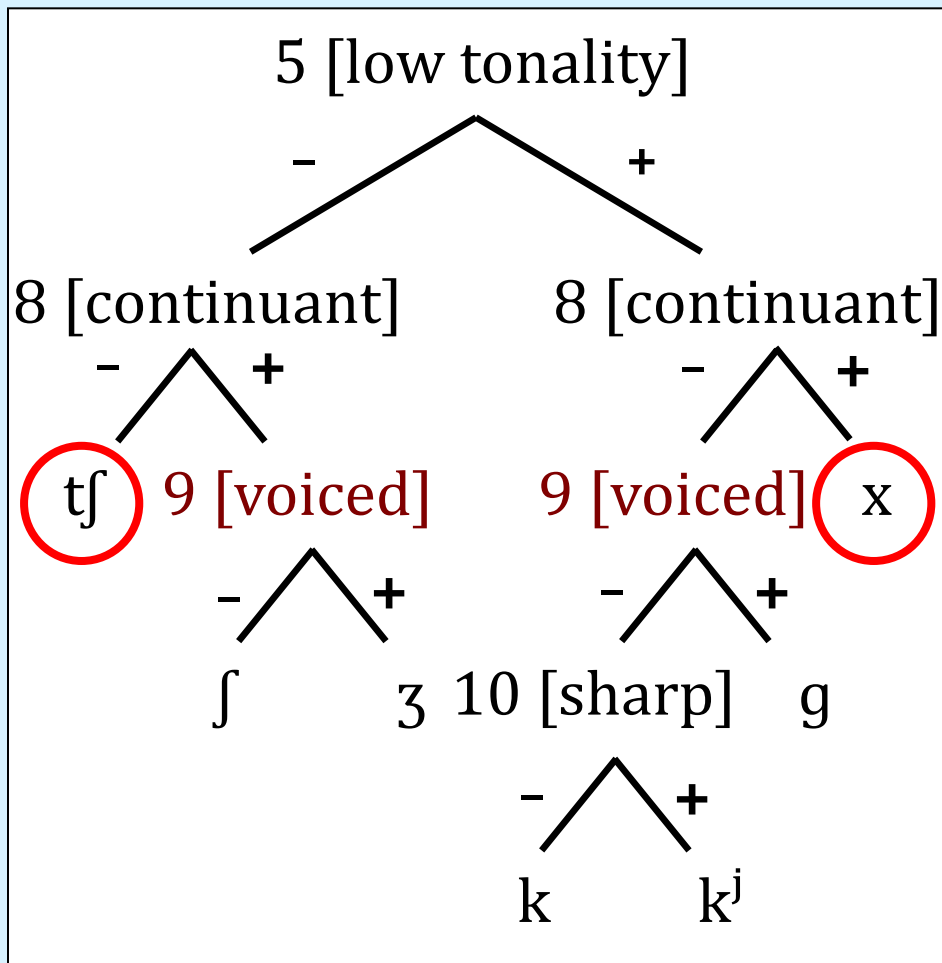
It follows that only **contrastive** features can be **active** in phonological processes (Dresher 2009; Mackenzie 2011, 2013).

Let us revisit the Russian example assuming now that Activity is our guiding principle.

Activity as the guiding principle
for ordering features:
The road not taken

The 'unpaired' phonemes redux

Halle's ordering in *SPR*

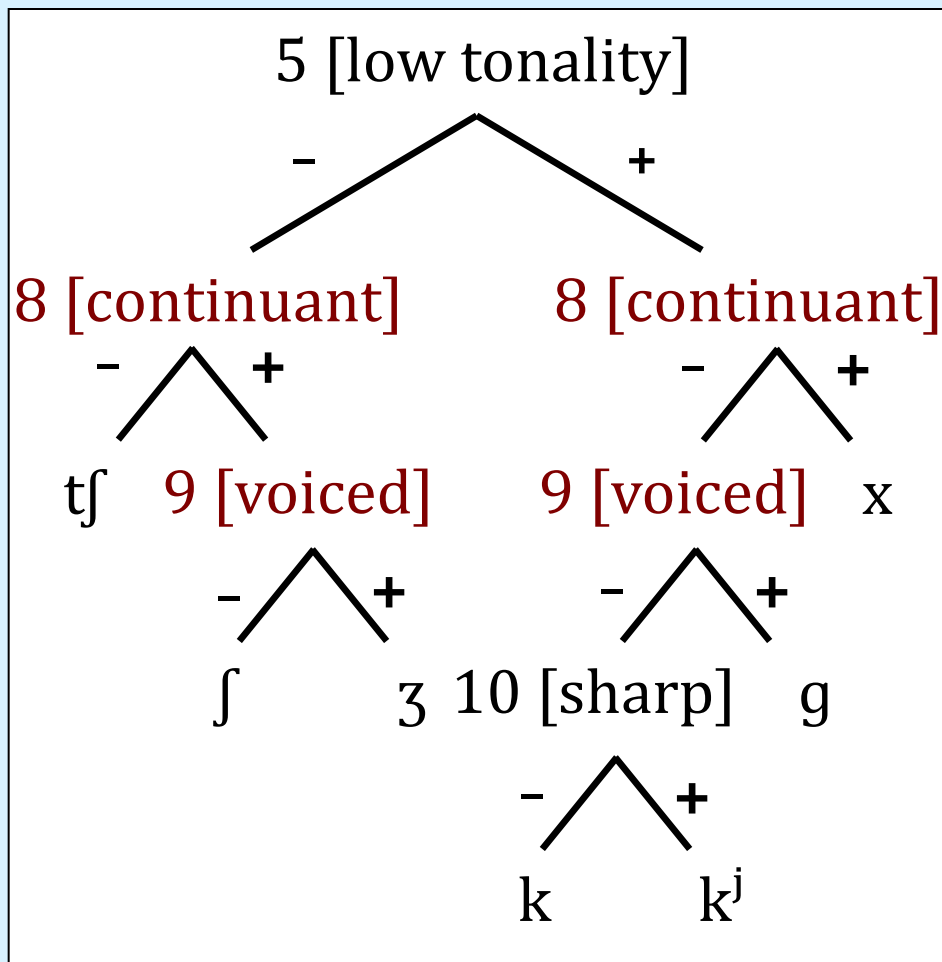


We have seen evidence from activity that the 'unpaired phonemes' /tʃ, x/ (and /ts/) must have a specification for [-voiced] with respect to the rule of RVA.

According to the Contrastivist Hypothesis, then, we conclude that these phonemes **must be** contrastively specified for that feature in the tree, contrary to the ordering in *SPR*.

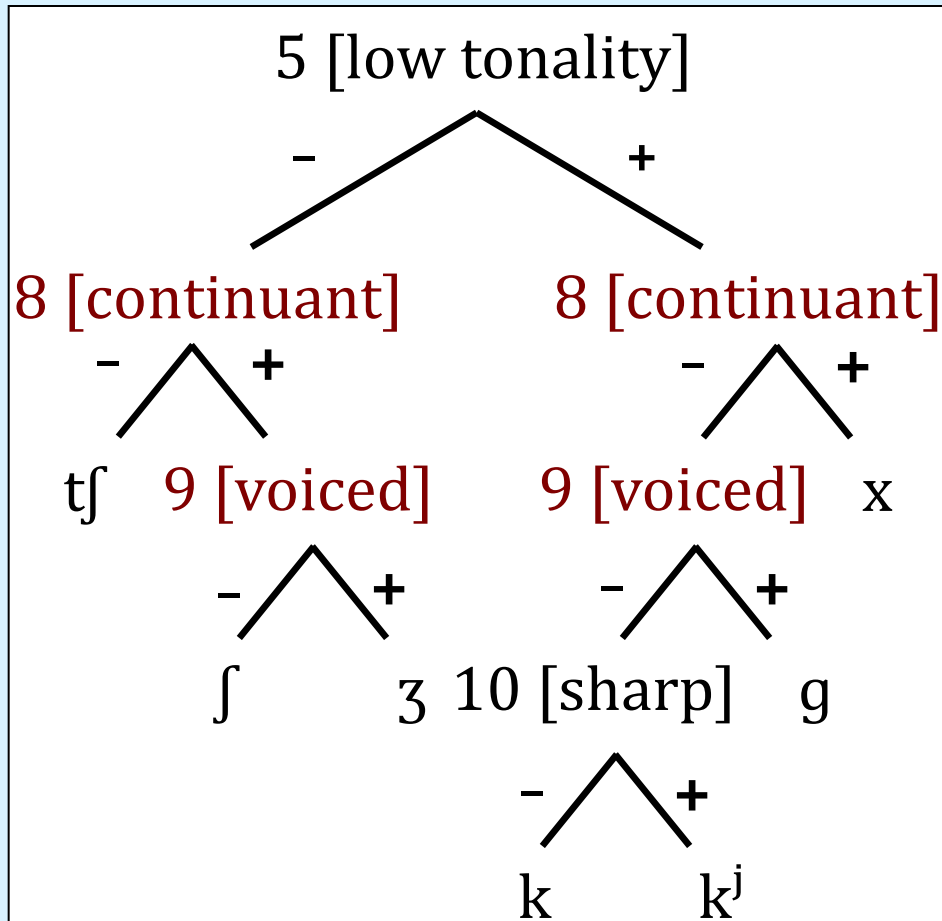
A minimal change in the ordering of [continuant] and [voiced] is enough to achieve this result and put this problem in a different light.

Halle's ordering in *SPR*

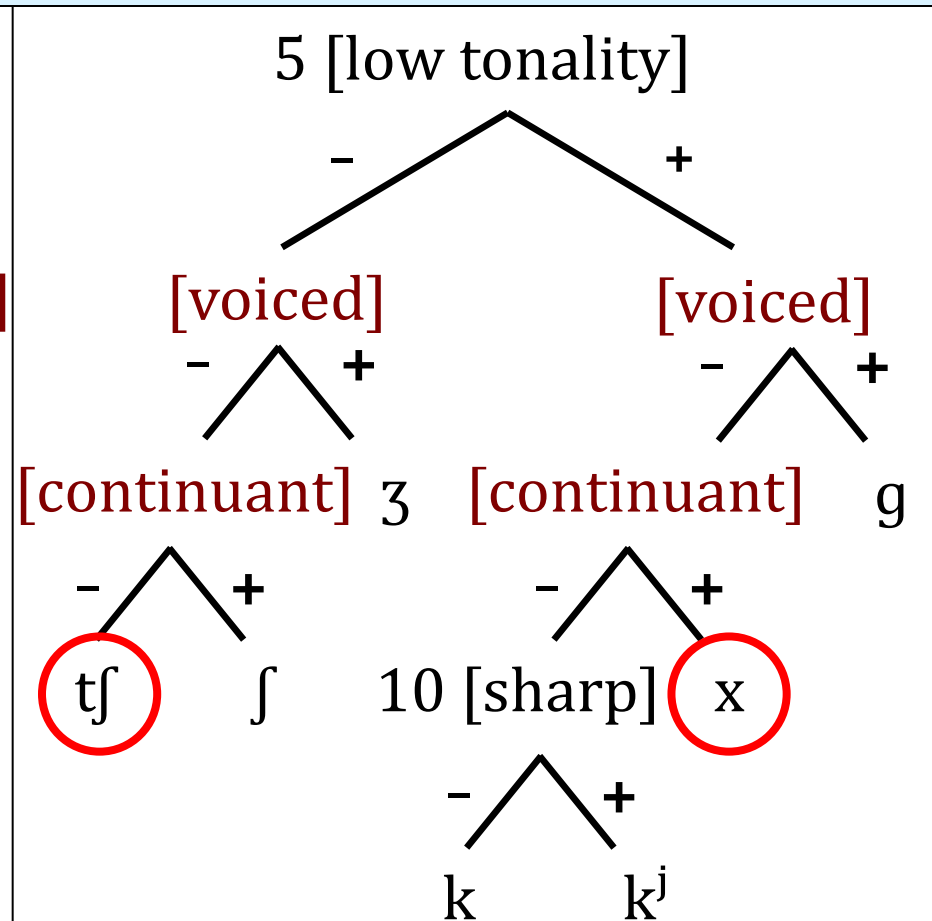


By ordering [voiced] slightly higher, the ‘unpaired’ phonemes become contrastively [-voiced], even though they have no voiced counterparts that are minimally different.

Halle's ordering in *SPR*

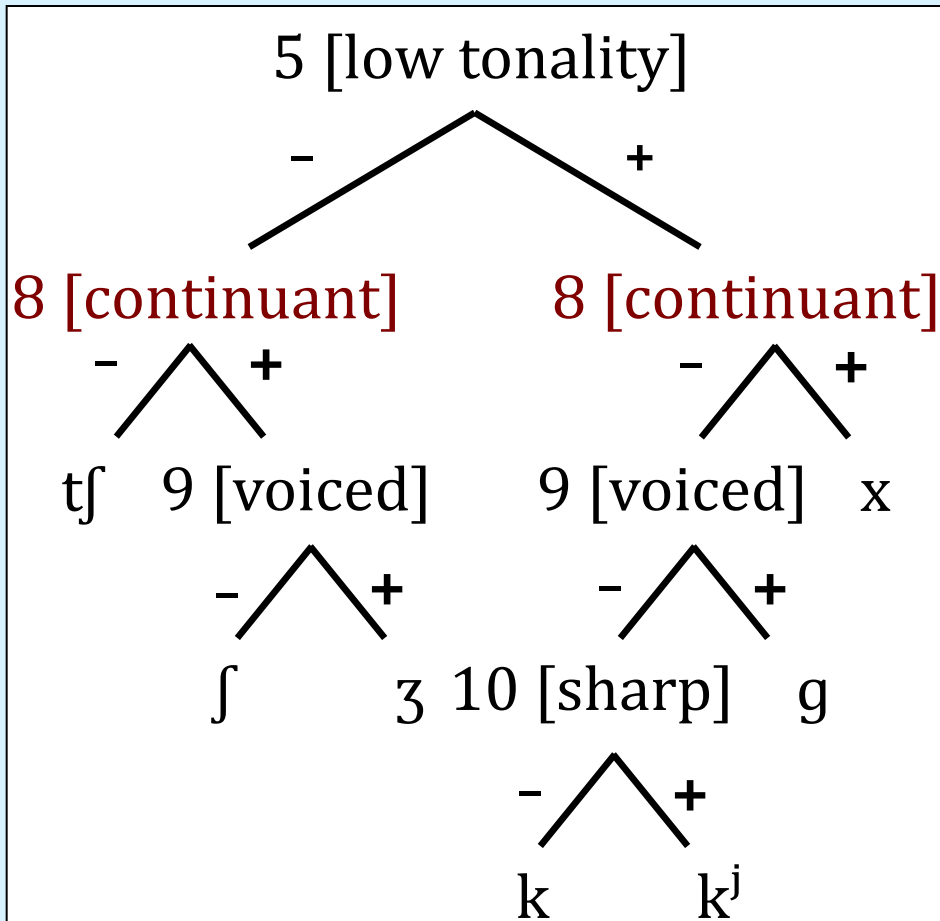


Revised ordering

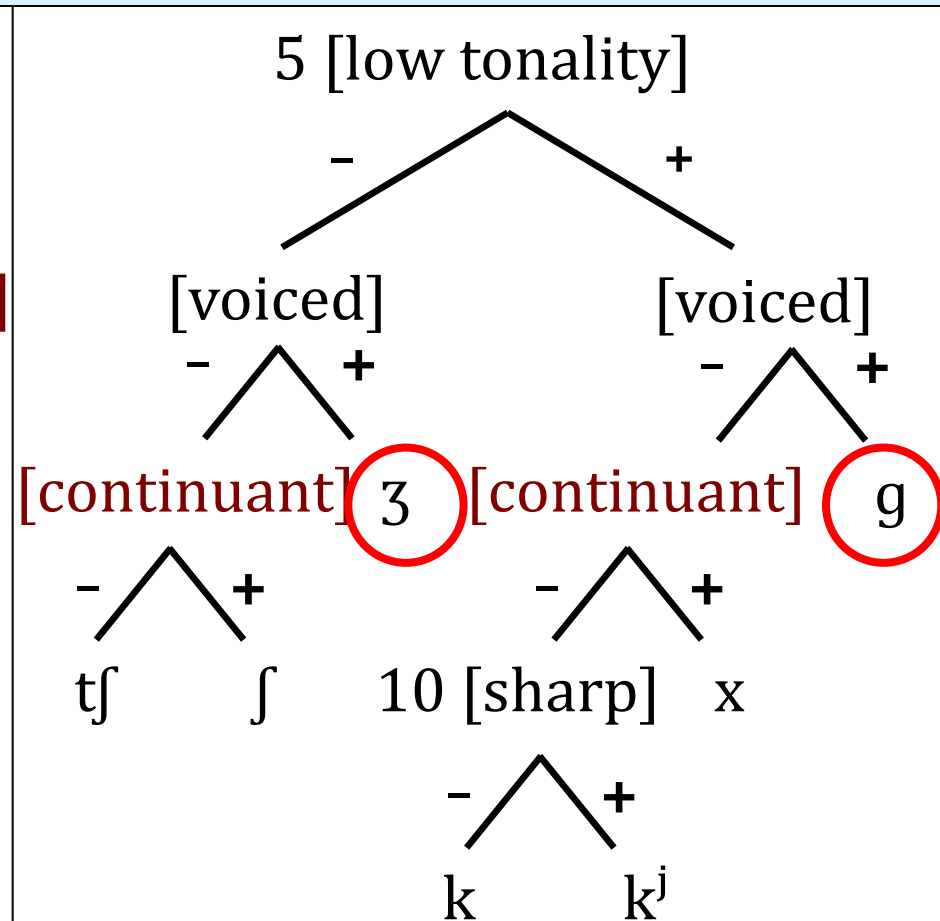


But the contrastive hierarchy forces a tradeoff: now the voiced consonants /ʒ/ and /g/ are unspecified for [continuant]. Is this a good result?

Halle's ordering in *SPR*



Revised ordering



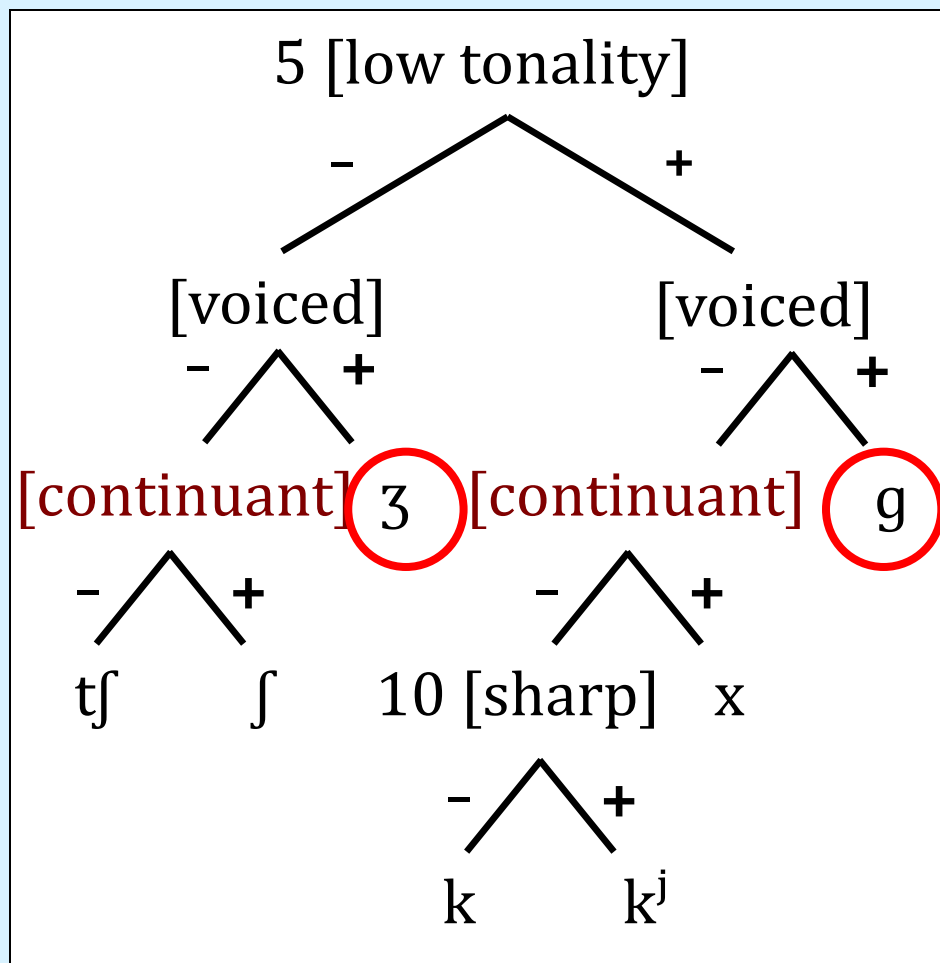
A revised Russian feature hierarchy

Dresher & Hall (2009) argue that there is circumstantial phonetic evidence that it is:

In some southern dialects of Russian and in Ukrainian, /g/ is realized as a voiced continuant [ɣ] or [ɦ].

This is a first indication that the status of /g/ as a stop may not be contrastively important; it is, however, consistently voiced.

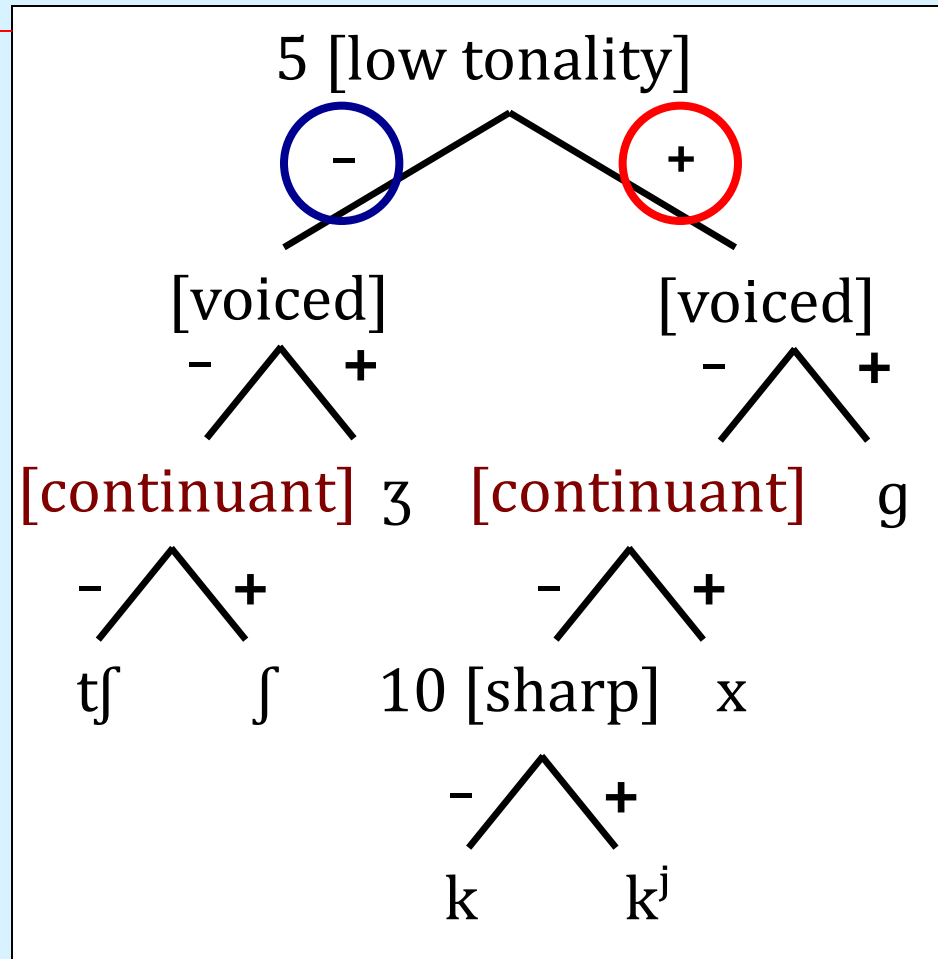
Revised ordering



There is also some (morpho)phonological evidence in the alternations resulting from the First Velar Palatalization; in terms of Halle (1959), the main change is in [low tonality]:

[+low tonality] → [-low tonality]

Revised ordering



Whereas continuant /x/ remains continuant [ʃ], and non-continuant /k/ remains non-continuant [tʃ], stop /g/ changes to fricative [ʒ].

[+low tonality] → [-low tonality]

/x/ → [ʃ]

[-voiced
+continuant]

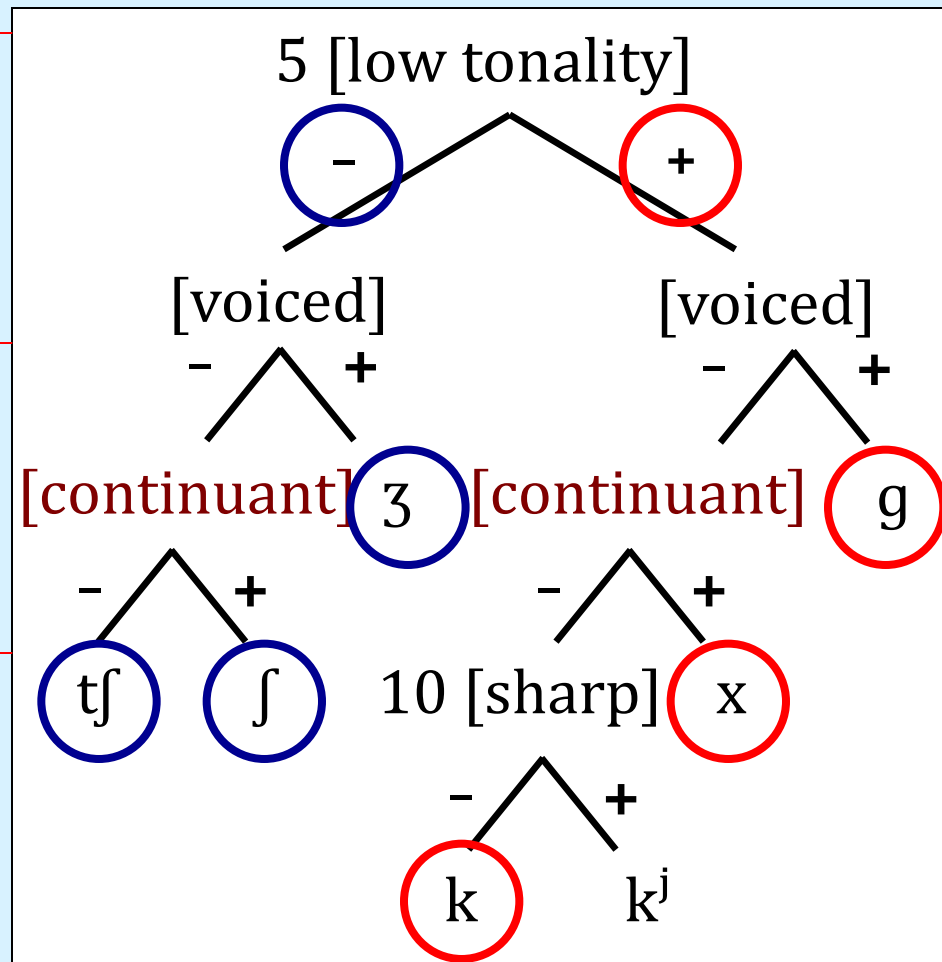
/k/ → [tʃ]

[-voiced
-continuant]

/g/ → [ʒ]

[+voiced
∅ continuant]

Revised ordering



Some examples are given below (Lightner 1965); see Radišić (2009) for a similar analysis of Serbian alternations.

Adjectives:

POSITIVE	COMPARATIVE	
tʲi x -ij	tʲi ʃ -e	‘quiet(er)’
zark k -ij	zart ʃ -e	‘hot(ter)’
dorog g -oj	doro ʒ -e	‘dear(er)’

Verbs:

3RD PL.		3RD SG.
max x -ut	ma ʃ -et	‘wave(s)’
pek k -ut	pe ʃ -et	‘bake(s)’

Denominal adjectives:

NOUN	ADJECTIVE	
ʃerepa x -a	ʃerepa ʃ -ij	‘turtle’/‘testudinian’
vol k	vol ʃ -ij	‘wolf’/‘lupine’
vra ǵ	vra ʒ -ij	‘enemy’/‘hostile’

Consequences of the reordering

This analysis suggests a different picture of phonological levels.

The Halle-Chomsky arguments against the structuralist phoneme, defined by a series of conditions as discussed in Chomsky (1964), still go through; but the duplication problem raised by RVA disappears.

This is because the rule applies one time to segments bearing a **contrastive** specification of [\pm voiced], whether the result is an already existing phoneme or a new allophone of a phoneme.

While the phonemic/allophonic distinction does not mark out a special level, the difference between contrastive and non-contrastive phonology does.

Contrast in phonology

Underlying Lexical
Representations: /tʃ/ /t/ /d/

*Contrastive
phonology*

RVA

Contrastive
Representations: |dʒ| |d| |t|

*Post-phonological
phonetic processes
(enhancement, etc.)*

Surface Phonetic
Representations: [dʒ] [d] [t]

Phonology proper is governed by the Contrastivist Hypothesis.

These forms are the output of the contrastive phonology, made up only of contrastive features.

The post-phonological component admits non-contrastive features, enhancement, etc. (Dyck 1995; Hall 2011).

The output of the contrastive phonology

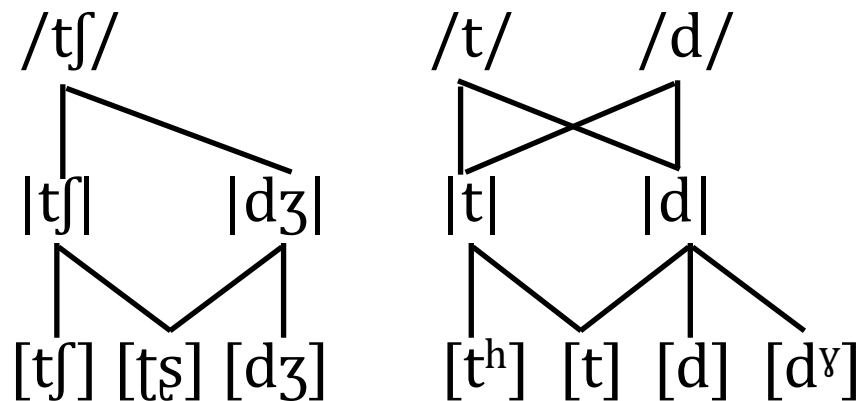
In this model, the elements in the output of the contrastive phonology, which I designate here between straight brackets | |, occupy an intermediate level between underlying lexical and surface phonetic representations.

Unlike the taxonomic phoneme, however, there are no constraints against overlapping or any other constraints on the relation between this level and phonetic representations.

Underlying Phonemic Reps.

Output of Contrastive Phonology

Phonetic Representations



The output of the contrastive phonology

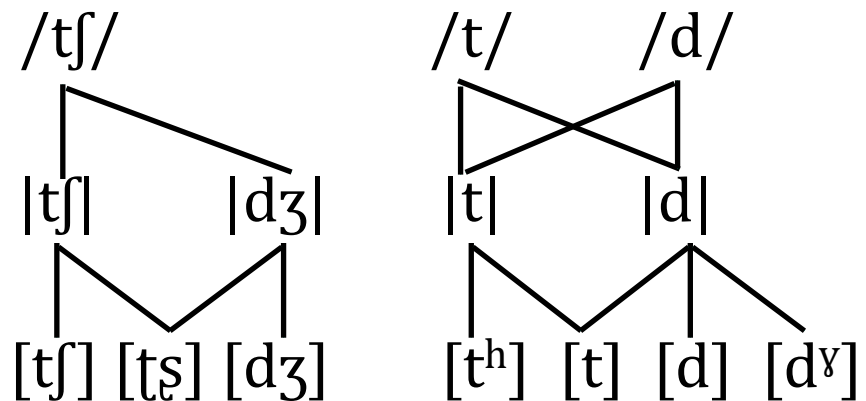
Therefore, all the arguments against the taxonomic phoneme would have still been the same.

What would be different is that between underlying lexical representations (now simply called phonemic) and the surface phonetic level (which has never been well defined) there would have been an intermediate level.

Underlying Phonemic Reps.

Output of Contrastive Phonology

Phonetic Representations



The output of the contrastive phonology

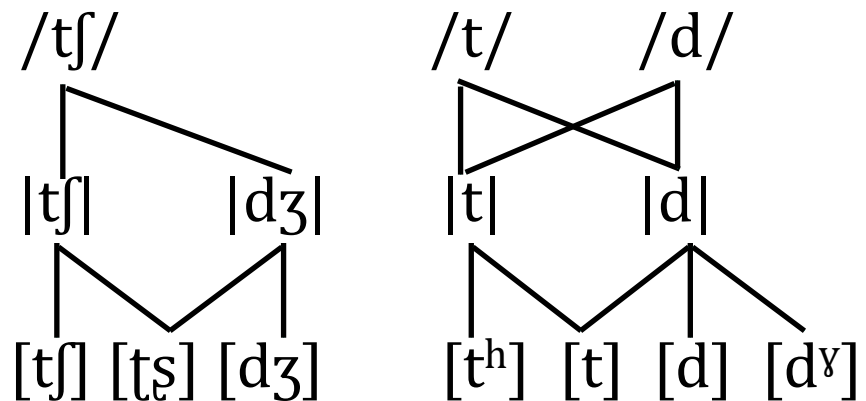
This level is not the old taxonomic phoneme, but is the output of the contrastive phonology, where only contrastive features are in play, and the input to enhancement and 'low-level' phonetics, where non-contrastive features are added.

This level resembles the level that divides lexical from post-lexical phonology, as it is commonly understood (Kiparsky 1982, 1985; Kaisse & Shaw 1985; Mohanan 1986).

Underlying Phonemic Reps.

Output of Contrastive Phonology

Phonetic Representations



Conclusion

In short, if Halle had more consistently favoured Activity over Minimality as the principle governing feature ordering in Russian:

- [voiced] would be ordered above [continuant];
- the ‘unpaired segments’ would be contrastively specified as [-voiced];
- the connection between contrast and phonological activity would be maintained;
- contrastive feature hierarchies (branching trees) would remain the way to generate contrastive representations.

In sum, on this path, language-particular contrast remains an important means of accounting for phonological patterning.

Conclusion

And here is a surprise:

Recall that I said that Halle's (1959: 44–5) report that his analysis of Russian requires 271 feature specifications appeared to be off by 1, and that he actually has 272 specifications.

It turns out, that when we reorder [voiced] and [continuant] as we suggested, without changing anything else in Halle's tree, the number of feature specifications comes to ... 271!

Coincidence? I don't know. But this means that ordering [voiced] over [continuant], the road not taken, would have been preferred even by Minimality.

Thank you!

*I shall be telling this with a sigh
Somewhere ages and ages hence:
Two roads diverged in a wood, and I—
I took the one less traveled by,
And that has made all the difference.*

References

- Bloch, Bernard. 1941. Phonemic overlapping. *American Speech* 16(4): 278–284.
- Calabrese, Andrea. 2005. *Markedness and economy in a derivational model of phonology*. Berlin: Mouton de Gruyter.
- Campos Astorkiza, Judit Rebeka. 2007. Minimal contrast and the phonology–phonetics interaction. Doctoral dissertation, University of Southern California.
- Chomsky, Noam. 1964. Current issues in linguistic theory. In Jerry A. Fodor and Jerrold J. Katz (eds.), *The structure of language*, 50–118. Englewood Cliffs, NJ: Prentice-Hall.
- Chomsky, Noam and Morris Halle. 1968. *The sound pattern of English*. New York: Harper & Row.

- Clements, G. N. 2001. Representational economy in constraint-based phonology. In T. Alan Hall (ed.), *Distinctive feature theory*, 71–146. Berlin: Mouton de Gruyter.
- Dresher, B. Elan. 2009. *The contrastive hierarchy in phonology*. Cambridge: Cambridge University Press.
- Dresher, B. Elan. 2015. The motivation for contrastive feature hierarchies in phonology. *Linguistic Variation* 15: 1–40.
- Dresher, B. Elan. 2016. Contrast in phonology 1867–1967: History and development. *Annual Review of Linguistics* 2: 53–73.
- Dresher, B. Elan and Daniel Currie Hall. 2009. Contrast in the Twentieth Century and Beyond. Presented at the 17th Manchester Phonology Meeting, University of Manchester, May 2009.

Dyck, Carrie. 1995. Constraining the phonology–phonetics interface, with exemplification from Spanish and Italian dialects. Doctoral dissertation, University of Toronto.

Frost, Robert. 1920. *Mountain interval*. New York: Henry Holt and Company.

Hall, Daniel Currie. 2007. The role and representation of contrast in phonological theory. Doctoral dissertation, University of Toronto.

Hall, Daniel Currie. 2011. Phonological contrast and its phonetic enhancement: Dispersedness without dispersion. *Phonology* 28: 1–54.

Halle, Morris. 1959. *The sound pattern of Russian: a linguistic and acoustical investigation*. The Hague: Mouton. Second printing, 1971.

Jakobson, Roman. 1941. *Kindersprache, Aphasie, und allgemeine Lautgesetze*. Uppsala: Uppsala Universitets Arsskrift.

Jakobson, Roman. 1962 [1931]. Phonemic notes on Standard Slovak. In Roman Jakobson, *Selected writings I. Phonological studies*, 221–230. The Hague: Mouton. [Published in Czech in *Slovenská miscellanea (Studies presented to Albert Pražak)*. Bratislava, 1931.]

Jakobson, Roman, C. Gunnar M. Fant and Morris Halle. 1952. *Preliminaries to Speech Analysis*. MIT Acoustics Laboratory, Technical Report, No. 13. Reissued by MIT Press, Cambridge, Mass., Eleventh Printing, 1976.

Jakobson, Roman and Morris Halle. 1956. *Fundamentals of language*. The Hague: Mouton.

Jakobson, Roman and John Lotz. 1949. Notes on the French phonemic pattern. *Word* 5: 151–158.

- Kaisse, Ellen M., and Patricia A. Shaw. 1985. On the theory of Lexical Phonology. *Phonology Yearbook 2*: 1–30. 305–328.
- Kiparsky, Paul. 1982. From cyclic to Lexical Phonology. In Harry van der Hulst and Norval Smith (eds.), *The structure of phonological representations (Part I)*, 131–176. Foris: Dordrecht.
- Kiparsky, Paul. 1985. Some consequences of Lexical Phonology. *Phonology Yearbook 2*: 85–138.
- Lightner, Theodore McGraw. 1963. A note on the formation of phonological rules. *Quarterly progress report* (Research Laboratory of Electronics, MIT): 68, 187–189.
- Lightner, Theodore McGraw. 1965. Segmental phonology of Modern Standard Russian. Doctoral dissertation, MIT.
- Mohanan, K. P. 1986. *Lexical Phonology*. Dordrecht: Reidel.

- Mackenzie, Sara. 2011. Contrast and the evaluation of similarity: Evidence from consonant harmony. *Lingua* 121: 1401–1423.
- Mackenzie, Sara. 2013. Laryngeal co-occurrence restrictions in Aymara: Contrastive representations and constraint interaction.. *Phonology* 30(2): 297–345.
- Nevins, Andrew. 2010. *Locality in vowel harmony*. Cambridge, MA: MIT Press.
- Padgett, Jaye. 2003. Contrast and post-velar fronting in Russian. *Natural Language and Linguistic Theory* 21: 39–87.
- Radišić, Milica. 2009. The double nature of the velar /g/ in Serbian. *Toronto Working Papers in Linguistics* 30: 91–103.
- Stanley, Richard. 1967. Redundancy rules in phonology. *Language* 43: 393–436.
- Trubetzkoy, N. S. 1939. *Grundzüge der Phonologie*. Göttingen: Vandenhoeck & Ruprecht.