PHONOLOGICAL CONSTRAINTS IN
SERBO-CROATIAN SYLLABLE MARGINS

and

MARKEDNESS IN GENERATIVE PHONOLOGY

by

Steven Uzelac

B.A. (Hons.) Simon Fraser University, 1969

TWO EXTENDED ESSAYS SUBMITTED IN PARTIAL
FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF
MASTER OF ARTS
in the Department

of

Modern Languages

© STEVEN UZELAC 1971

SIMON FRASER UNIVERSITY

March, 1971
APPROVAL

Name: Steven Uzelac

Degree: Master of Arts

Titles of Extended Essays:
1. Phonological Constraints in Serbo-Croatian Syllable Margins
2. Markedness in Generative Phonology

Examining Committee:

__________________________
N. J. Lincoln
Chairman

__________________________
E. W. Roberts
Senior Supervisor

__________________________
B. Newton

__________________________
R. C. DeArmond

Date Approved: April 27, 1971

(11)
Acknowledgement

Work on the first essay, "Phonological Constraints in Serbo-Croatian Syllable Margins", was initiated under the supervision of Dr. Ross Saunders.
ESSAY I

Phonological Constraints in Serbo-Croatian

Syllable Margins

by

Steven Uzelac

(iv)
TABLE OF CONTENTS

0.0 Introduction .............................................................. 1
    0.1 Statement of purpose ............................................. 1
    0.2 Summary of Contents ............................................ 1
    0.3 Description of the Corpus ....................................... 2

1.0 Segmentation procedures ................................................ 2
    1.1 Interlude segmentation ........................................ 4
    1.2 Isolated Consonantal Microsegments ............................ 4

2.0 The phonological syllable of Serbo-Croatian ....................... 6
    2.1 Structure ........................................................ 6
    2.2 Inventory of Possible Manifestations ......................... 9

3.0 The Constraint Grammar ................................................. 11
    3.1 Components of a Constraint Grammar ........................... 12
    3.2 The Rule Set ................................................... 18

4.0 Observations on the Juxtaposition Rule ............................. 39
    4.1 The Rules ...................................................... 40

5.0 Concluding Observations ............................................... 44

Appendix ................................................................. 47

Notes ................................................................. 50

Bibliography .......................................................... 52
0.0 INTRODUCTION

0.1 Statement of Purpose

The purpose of this study is to describe the structure of margins of the Serbo-Croatian phonological syllable in terms of a constraint grammar, i.e., a set of rules which account for the positional and sequential constraints on the phonological primes within the limits of the structure of syllable margins, and then to examine the phonological phenomena that constitute the rules of this grammar.¹

0.2 Summary of Contents

There are five major sections in addition to the Introduction and an Appendix.

Section 1.0 contains a summary of the description of the segmentation of the Serbo-Croatian utterance into successively smaller size units; the utterance through phonemic phrases, macrosegments, phonological words, microsegments, and finally the phonological syllable.

Section 2.0 contains a description of the structure of the phonological syllable with particular attention to the structure of the various sizes of onset and codas.

Section 3.0 contains the constraint grammar. The first subsection describes the components of the grammar and the second contains the rule sets.

Section 4.0 contains an examination and discussion of the juxtaposition rules.

Section 5.0 contains the concluding observations of the phonological constraints with respect to their relative powers.

The Appendix contains a list of all the grammatical onsets and codas of Serbo-Croatian and a table in which the juxtaposition rules are ranked
by their relative powers.

0.3 Description of the Corpus

The material comprising the corpus in this study is taken from dictionaries, grammars, and linguistic studies of Serbo-Croatian. The pattern of native Serbo-Croatian phonemes is the main concern. Borrowings from other languages will only be considered if they conform to the pattern of Serbo-Croatian. If not, they will not be allowed to disrupt the pattern, but will be listed at the end of a pattern set and discussed. The same holds true for nonnative phonemes, which always occur in nonnative words, particularly /f/, which destroy the language pattern.²

1.0 Segmentation Procedures

The phonological syllable is the result of a segmentation procedure which is essentially an IC analysis such as originated by Hockett 1955 and modified for the Slavic languages by Kučera and Saunders.³ The phonological syllable is, "The smallest unit or recurrent phonemic sequences which make it possible to describe the distribution of segmental phonemes and configurative phonemic entities most economically."⁴

It is unnecessary to outline the entire segmentation procedure here because this paper differs in no essential features from the above mentioned. In summary, however, "the utterance is segmented successively into smaller units. From the utterance through phonemic phrases, macrosegments, phonemic measures, phonological words, and microsegments we arrive at the definition of the phonological syllable as the smallest unit of recurrent phonemic sequences. The constituents of the syllable are then defined as the "nucleus" - the irreducible minimum present in any recurrent phonemic unit and the remainder as the "margin" of the syllable. The margin in turn is broken down into an optional "onset", defined as that which occurs
between a disjuncture and the first following nucleus and an optional "coda", defined as that which occurs between the last nucleus and the following disjuncture. The margin between two nuclei uninterrupted by any occurrence of disjuncture is an "interlude" and these are subsequently divided into onsets and codas by a set of ordered rules."

Furthermore, "although syllable stress is a property of the entire syllable, it reaches a maximum of intensity in one portion of the syllable and for this reason it can be said to signal the nucleus of the syllable. Pitch also attains peaks of intensity in exactly the same location as does Stress. This permits us to characterize the nucleus of the phonological syllable as the bearer of pitch and stress for the syllable as a whole."6

The nucleus of a syllable in Serbo-Croatian may consist of any vowel /i,e,a,o,u/ or the liquid /r/ initially before a consonant or inter-consonantly, i.e.,

\[ r \rightarrow \text{nucleus} / \frac{\#}{C} - C \]

All other segmental phonemes / p,t,k,b,d,g,c,ć,d,ž,f,s,š,h,v,z,ž,r,l, š,ž,ć,ć,ı,j,m,n,ń/ may only constitute the margin or part of the margin.

The optional part of the syllable, the margin, may be divided into two parts as stated above, the onset (mO) and the coda (mC). Thus the syllable has the structure (mO) N (mC).

There exists however, the problem of interlude division when there is more than one nucleus between two successive disjunctures. The basis for interlude division shall rest on the assumption that internal syllables have the same structure as external syllables. Therefore the interludes will be segmented so as not to create any internal onset or coda that does not exist as an external onset or coda. When two alternate ways are found to be equally acceptable, i.e., they do not add to the list of external
onsets or codas, then the alternative that has a higher probability of occurrence in the language is preferred. The maximum segments that an external onset or coda is found to have is three. Therefore any interlude with more than six members would add a new addition to the list of external onsets and codas. The largest interlude, however, has only four members.

1.1 Interlude Segmentation

With the above criteria in mind the interludes are segmented and given in order of preference in the following list:

<table>
<thead>
<tr>
<th>Interlude type</th>
<th>Alternate divisions ordered by preference</th>
</tr>
</thead>
<tbody>
<tr>
<td>N C N</td>
<td>N C N N C N</td>
</tr>
<tr>
<td></td>
<td>N C N N C N</td>
</tr>
<tr>
<td>N C C N</td>
<td>N C C N N C N</td>
</tr>
<tr>
<td></td>
<td>N C C N N C N</td>
</tr>
<tr>
<td>N C C C N</td>
<td>N C C C N N C C N</td>
</tr>
<tr>
<td></td>
<td>N C C C N N C C N</td>
</tr>
<tr>
<td>N C C C C N</td>
<td>N C C C C C N</td>
</tr>
<tr>
<td></td>
<td>N C C C C C N</td>
</tr>
</tbody>
</table>

1.2 Isolated Consonantal Microsegments (ICM's)

As Kučera 1961, for Czech, and Saunders 1970, for Russian, had to deal with the problem of ICM's so must the problem be dealt with here for Serbo-Croatian.
The segmentation of complex phonological words yields two types of microsegments:

1. Word microsegments (capable of constituting simple phonological words).

2. Isolated microsegments (incapable of constituting simple phonological words). These are found to be prepositions in Serbo-Croatian.

Isolated microsegments are further subdivided into:

1. Isolated syllabic microsegments, which because they contain a syllable nucleus are of no special concern.

2. Isolated consonantal microsegments, which are composed of one consonant and are asyllabic.

In Serbo-Croatian there are two ICM's, /k/ 'to' and /s/ 'with' or 'from'.

There are four alternate ways of dealing with ICM's.

1. "Leave them as they are, i.e., as asyllabic residues." This would leave behind a residue unaccounted for in the language.

2. "Consider ICM's as special types of syllables." This would mean that in this case two obstruents; a stop and a fricative, would function as a syllable nucleus. This is unacceptable because it would allow a stop to function as a syllable peak.

3. "Consider the ICM and its following disjunctures as constituents of the set of the following syllable." This alternative would violate the hierarchy of levels, for the syllable boundary would then extend beyond the boundary of the microsegment. This also proves to be unsatisfactory. Saunders 1966, considered this proposal the best, but abandoned it for the above reason in Saunders 1970.

4. "Suppress internal disjuncture and consider the ICM as a constituent
of the onset of the following syllable."\textsuperscript{13}

This alternative is the least attractive. First it would eliminate the microsegment boundary and therefore the concept of the microsegment, secondly it would modify a general phonological rule in Serbo-Croatian. The rule states that any cluster of obstruents constituting a syllable margin or part of a margin and not separated by a nucleus will have the same quality of voice. This rule would be violated by such sequences as:

\# k + bratu \# 'to brother'

\# s + bratom \# 'with brother'

where the /k/ and the /s/ do not assimilate in voice to the next segment in normal tempo.

Of the four alternatives the first seems the most satisfactory. Therefore ICM's shall be regarded here as Hockett, Kučera and Saunders regard them, "as asyllabic residues or as Kučera calls them, special presyllabic segments."\textsuperscript{14}

2.0 The Phonological Syllable of Serbo-Croatian

2.1 Structure

The structure of the syllable shall now continue to be analysed down to the ultimate phonological units of this study, the phonemes of Serbo-Croatian.

2.11 The Nucleus

The nucleus of a phonological syllable in Serbo-Croatian always consists of only one segmental phoneme. Therefore there is no hierarchical structure for the nucleus.

2.12 The Margin

The margin of a phonological syllable of Serbo-Croatian is discontinuous, and as stated above is divided into an optional onset and an
optional coda. The onset and codas are of three sizes:

a. one member - one member onset \( \langle 1^0 \rangle \)  
   - one member coda \( \langle 1^c \rangle \)

b. two members - two member onset \( \langle 2^0 \rangle \)  
   - two member coda \( \langle 2^c \rangle \)

c. three members - three member onset \( \langle 3^0 \rangle \)  
   - three member coda \( \langle 3^c \rangle \)

2.13 The Structure Onsets and Codas

Onsets and codas shall not be regarded as unstructured strings of segmental phonemes, but shall be viewed as structural units so that larger onsets and codas are composed of smaller units of smaller onsets and codas. The following is the specific structure of the various onset and coda types proposed here for Serbo-Croatian.

\[
/0/ \rightarrow \langle 1^0 \rangle ; \langle 1^c \rangle \leftarrow /0/ \\
\langle 1^0 \rangle \& \langle 1^0 \rangle \rightarrow \langle 2^0 \rangle ; \langle 2^c \rangle \leftarrow \langle 1^c \rangle \& \langle 1^c \rangle \\
\langle 1^0 \rangle \& \langle 2^0 \rangle \rightarrow \langle 3^0 \rangle ; \langle 3^c \rangle \leftarrow \langle 1^c \rangle \& \langle 2^c \rangle
\]

A one member onset or coda is composed of a single segmental phoneme. A two member onset or coda is composed of two valid one member onsets or codas respectively. This is the only possible structuring for two member onsets and codas. A three member onset or coda is composed of a sequence of a valid one member onset or coda followed by a valid two member onset or coda respectively.

The above proposed structuring is based on the observation that certain parts of the strings of phonemes with the lists of valid onsets and codas displayed a greater power of recombination than others. This may be demonstrated by a representative list of valid three member onsets.
The recurring partial in all five cases is /s/ followed by a valid two member onset. By equating the power of recursion with structural units the structure of a three member onset involves the juxtaposition of a valid one member onset followed by a valid two member onset.

There are alternate ways of structuring three member onsets and codas. One is to regard them as a sequence of three valid one member onsets or codas.

\[ \langle 1^0 \rangle \& \langle 1^0 \rangle \& \langle 1^0 \rangle \rightarrow \langle 3^0 \rangle \quad ; \quad \langle 3^C \rangle \leftarrow \langle 1^C \rangle \& \langle 1^C \rangle \& \langle 1^C \rangle \]

This alternative, however, does not account for recurring partials and represents the structure as a mere string of phonemes. It therefore must be discarded. The second alternative is to regard them as a sequence of a valid two member onset or coda followed by a valid one member onset or coda respectively.

\[ \langle 2^0 \rangle \& \langle 1^C \rangle \rightarrow \langle 3^0 \rangle \quad ; \quad \langle 3^C \rangle \leftarrow \langle 2^C \rangle \& \langle 1^C \rangle \]

This alternative is also discarded. One reason is because it would add a new two member coda to the list of valid two member codas. It also creates a situation within the structure of three member onsets whereby more new and more complicated juxtaposition rules are created to block combinations that would occur at a lower structure than would be needed with the preferred structure.
2.2 **Inventory of Possible Manifestations**

Each onset and coda type has a number of potential manifestations based on the permutations of its structural units. For example, the potential number of two member codas is eight; the number of valid members of its first unit (4) multiplied by the number of valid members of its second units (2). This may be represented by a matrix with the first member on the horizontal axis and the second member on the vertical axis. However, only half of the potential number occur. The remainder do not occur for three possible reasons:

1. They are ungrammatical. They do not follow the sound pattern of Serbo-Croatian and are not employed in the language. An example is the two member potential onset /$d-$/.

2. They are ungrammatical but are employed by the language. An example is the two member potential coda /-lm/ found in the recent borrowing /film/.

3. They are grammatical but are fortuitously unemployed in the language (these are Fischer-Jørgensen's accidental gaps in the language$^{15}$). An example is the two member potential onset /bz-/.

Using Vogt's terminology four levels of grammaticality are posited$^{16}$:

1. **Actual clusters.** These clusters occur in the corpus and follow the sound pattern of Serbo-Croatian.

2. **Virtual Clusters.** These clusters do not occur in the corpus but follow the sound pattern of Serbo-Croatian.

3. **Marginal Clusters.** These clusters occur in the corpus but do not follow the sound pattern of Serbo-Croatian.

4. **Inadmissible Clusters.** These clusters do not occur in the corpus nor do they follow the sound pattern of Serbo-Croatian.
2.21 The Recovery of Missing Grammatical Clusters

The recovery of grammatical clusters, the virtual clusters, depends upon two assumptions:

1. There are certain groupings of segmental phonemes which are more natural than others. Without any claims to universal applicability natural sets appear to be yielded in Serbo-Croatian by the suspension of the distinctive opposition of voicing. The natural sets are:

\[
\begin{align*}
\text{P} &/p, b/ , \text{T}/t, d/, \text{K}/k, g/ \\
\text{C} &/c/, \text{č}/č, ď/ň/ \\
\text{F} &/f, v/, \text{S}/s, z/,  ľ/ľ/ \\
\text{M} &/m/, \text{N}/n/, \text{ň}/ň/ \\
\text{L} &/l/, \text{ř}/ř/, \text{ř}/ř/ \\
\end{align*}
\]

2. Analogical extension. If one member of a natural set occurs in environments '-g', '-r', '-s', then all members of the natural set will be allowed this distribution. Only a rule that involved a constraint on the feature voicing would disallow this extension, for example, a rule which specifies that a two member onset cannot consist of two obstruents of dissimilar voice would invalidate the analogical extension to /sd-/ from /zd-/.

After all the grammatical clusters have been recovered and marked in the matrix then the ungrammatical clusters will be represented by the unmarked cells. Rules must now be formulated to account for the nonoccurrence of these ungrammatical clusters. "This set of rules specifying the membership of clusters and the sequential constraints which make certain sequences of segmental phonemes in certain structures inadmissable, will be called a constraint grammar." 17
3.0 **The Constraint Grammar**

The constraint grammar has three major goals:

1. The assignment of each possible onset and coda to one of the levels of grammaticality.

2. The examination of the phonological constraints which allow only the grammatical clusters.

3. A measure of confidence in quantitative terms of grammaticality assignments.

3.01 **Grammaticality Assignments**

Grammaticality should not be equated with occurring in the dictionary. This definition of grammaticality has no predictive power and fails to provide an explanation of why certain ungrammatical clusters seem to be less grammatical than others. It also fails to provide any clues of future phonological change.

Grammatical should be equated with conforming to the sound pattern of the language. The assignment of clusters to grammatical and ungrammatical occurs as sets of constraint rules, which permit only grammatical sequences, are applied to all the manifestations of possible clusters. Those clusters which are blocked are tagged ungrammatical. Ungrammatical clusters are assigned to a level of grammaticality. If they do not occur in the corpus they are inadmissible. If they occur they are marginal. Likewise the grammatical clusters are assigned to a level of grammaticality; as virtual clusters if they do not occur in the corpus and as actual clusters if they do occur.

3.02 **Phonological Constraints**

The second goal of the constraint grammar is an examination of the phonological constraints. There are two aspects which are of importance:
1. The Relative Power of a rule. This is the percentage of potential onsets and codas tagged ungrammatical by this rule.

2. The domain of a rule. It is of interest to know if the domain of a rule is limited to onsets or codas or can apply to both. Rules limited to onsets are 'onset rules' and rules limited to codas are 'coda rules' and rules which apply to both onsets and codas are 'universal rules'.

3.03 Measure of Confidence in Grammaticality Assignments

A quantitative measure of confidence in grammaticality assignments is accomplished by a combination of rule overlapping and relative rule power within a rule set. Both rule power and rule overlapping are discussed in section 3.12.

3.1 Components of a Constraint Grammar

A constraint grammar is unlike most grammars in that most other grammars will operate on an inventory of phonological primes to generate only the grammatical structures while a constraint grammar acts as a battery of filters to eliminate the ungrammatical structures from an inventory of all potential structures. There are three main components to the proposed grammar:

1. The initial input; the inventory of segmental phonemes of Serbo-Croatian.
2. The set of rules which result in the output.
3. The list of grammatical onsets and codas.

3.11 The Inventory of Segmental Phonemes

The phonemes of Serbo-Croatian are regarded as bundles of distinctive features which are differentiated from all other phonemes of Serbo-Croatian by the same set of distinctive features.
The set of distinctive features used here are the (articulatory) features found in Chomsky-Halle 1968. These are preferable to Jakobsonian features because they are easier to work with, e.g., MRL (see 3.21)

\([-\text{syllabic}]\]

is used to separate all members of the syllable margin from members of the nucleus. In Jakobsonian features the same set is defined as:

\[ [+\text{ consonantal}] + [+\text{ vocalic}] + [-\text{ consonantal}] \]

The main disadvantage here is the greater number of features, fourteen instead of twelve Jakobsonian features, but this is compensated for by the simpler rules.

The following matrix is fully specified. The redundancies are brought out in the rules.

1. syllabic : nonsyllabic
2. sonorant : nonsonorant (obstruent)
3. consonantal : nonconsonantal
4. nasal : nonnasal
5. lateral : nonlateral
6. continuant : noncontinuant
7. anterior : nonanterior
8. coronal : noncoronal
9. delayed release : instantaneous release
10. distributed : nondistributed
11. voiced : nonvoiced
12. low : nonlow
13. back : nonback
14. high : nonhigh
|      | p | b | t | d | c | k | g | f | v | s | z | ı | h | n | n | r | l | i | j | i | e | a | o | u |
| syl  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| son  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| cons |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| nas  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| lat  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| cont |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| ant  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| cor  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| del r|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| dis  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| voice|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| low  |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| back |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| high |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
The preceding matrix is useful for distinguishing each phoneme from every other phoneme and establishing boundaries for phonetic variations of phonemes. For example, the phoneme /p/ has a range of possibilities that would never include the other phonemes /t/ or /b/.

However, the preceding matrix has one drawback for phonotactic or syntagmatic investigations. In Serbo-Croatian a natural class is formed by the phonemes /v, h, m, n, ň, r, l, ĩ, i, e, a, o, u/. Before any of these phonemes the feature voiced-nonvoiced is operative. All of these except /v/ are plus sonorant and therefore can be easily recovered. Therefore to facilitate the rules /v/ shall be regarded as plus sonorant in all future discussion and rules. /v/ is no longer a member of the set F3/f, v/. Two new sets are formed F3/f/ and /V/3/v/.

3.12 The Phonological Constraint Rules

It is seen in 2.13 that the structuring of onsets and codas consists of sequences of smaller onsets and codas. This results in an ordering of rules; those rules resulting in an inventory of valid onsets and codas of smaller size preceding those rules of the larger sizes of onsets and codas.

There are two types of rules in the constraint grammar:

1. Membership rules

2. Juxtaposition rules

Membership rules assign the membership of each structural unit as seen in 2.13. Juxtaposition rules then operate upon the membership of the structural units to disallow all ungrammatical clusters. For example, a membership rule assigns all the nonsyllabic phonemes of Serbo-Croatian to one member onsets. No juxtaposition rules apply because there is no sequence of structures. The grammatical one member onsets now serve as an
input to the membership rules for two member onsets which are composed of a sequence of two one member onsets. Membership rules for each unit of the structure assigns the membership to each unit. Juxtaposition rules operate on the potential number of two member onsets to block all ungrammatical two member onsets. The grammatical one member and two member onsets serve as an input to the membership rules of three member onsets. Membership rules for three member onsets assign the memberships of each structural unit. The potential number of three member onsets is acted upon by juxtaposition rules to block all ungrammatical clusters. Likewise for codas.

Rule sets are ordered. This is to prevent interference or overlapping of constraints at higher levels with those at lower levels.

Rules within a set are unordered. If the rules were ordered then overlapping of rules would diminish the power of a later rule which shared part of the same domain. Part of this study is to ascertain the total number of clusters constrained by each rule. This could hardly be done if the rules were ordered for the above reason. Furthermore the overlapping of unordered rules gives a measure of ungrammaticality. For example, a cluster constrained by four rules is considered less grammatical than a cluster constrained by only one rule. In addition, a cluster constrained by one rule is more likely to be vulnerable to future phonological change than is a cluster constrained by many rules. This could give one indication of the direction of phonological change. Another reason for creating an unordered set of rules is to allow the rule the opportunity to apply to every cluster in the set. By allowing this a rule may be ranked according to the percentage of clusters it constrains.

3.13 The Form of the Rules

Both membership and juxtaposition rules are written out in full
followed by a formulation of the rule statement. Since the conventions of the rule formulae are easily recoverable from the rule statements and the rules are given in exactly the same manner as Saunders 1970, there is no need to repeat an explication of them here. However, in brief, the membership rules designate the membership of each structural unit in distinctive features and are followed by a list of segmental phonemes that represent the membership of each structural unit as given by the said rule. The juxtaposition rules act as constraints using distinctive features for primes. The juxtaposition rules are followed by a statement of 'Onset Power' (OP) or 'Coda Power' (CP) and possibly 'Total Power' (TP) which signifies the number of onsets or codas or onsets and codas, respectively, tagged ungrammatical by this rule, and 'Relative Onset Power' (ROP) or 'Relative Coda Power' (RCP) or both plus 'Total Relative Power' (TRP) which are percentages of the total of possible onsets or codas or onsets and codas, respectively, tagged ungrammatical within the rule set. Each juxtaposition rule is followed by a list of the clusters it has tagged ungrammatical.

3.2 The Rule Set

The rules are presented by onset-coda levels. First the membership rules are given, followed by Universal Juxtaposition Rules (UR), i.e., rules that apply to both onsets and codas of the same cluster size. These are in turn followed by Onset Juxtaposition Rules (OR) and then by Coda Juxtaposition Rules (CR) of the same set. At the completion of a rule set a list is given of all the onsets and codas that remain untagged as ungrammatical and are therefore grammatical.

3.2.1 One Member Onsets and Codas

MR1 "The membership of one member onsets and codas consists of all none syllabic
segments.

\[
\langle 1^m \rangle \in [-\text{syllabic}]
\]

OP = 5 \hspace{1cm} ROP = 16.7\% \hspace{1cm} CP = 5 \hspace{1cm} RCP = 16.7\% \hspace{1cm} TP = 10 \hspace{1cm} RTP = 16.7\%

3.211 The list of grammatical one member onsets and codas.

<table>
<thead>
<tr>
<th>onsets (\langle 1^0 \rangle)</th>
<th>codas (\langle 1^C \rangle)</th>
</tr>
</thead>
<tbody>
<tr>
<td>p, b, t, d, k, g</td>
<td>p, b, t, d, k, g</td>
</tr>
<tr>
<td>c, č, ć, č, ř</td>
<td>c, č, ć, č, ř</td>
</tr>
<tr>
<td>f, v, s, z, š, č, h</td>
<td>f, v, s, z, š, č, h</td>
</tr>
<tr>
<td>m, n, ń, r, 1, ́, j</td>
<td>m, n, ń, r, 1, ́, j</td>
</tr>
</tbody>
</table>

3.22 Two member onsets and codas

3.221 Membership Rules

MR2 "The membership of the first constituent of a two member onset consists of all the valid one member onsets except \(/f, č, ř, x, 1, ́, n, ń, j/\)."\(^{19}\)

\[
\langle 2^0 \rangle_1 \ni \begin{cases} \text{sonorant} \\ \hspace{1cm} \underline{- \text{continuant}} \\ \hspace{1.5cm} \begin{cases} + \text{coronal} \\ + \text{delayed release} \\ - \text{distributed} \end{cases} \end{cases} \begin{cases} \hspace{1.5cm} + \text{continuant} \\ \hspace{3cm} + \text{anterior} \end{cases} \begin{cases} - \text{coronal} \\ - \text{syllabic} \\ + \text{sonorant} \\ - \text{coronal} \end{cases} \sim (-\text{consonantal}) \end{cases}
\]

\(P = 225\) \hspace{1cm} \(RP = 35.2\%\)

MR3 "The membership of the second constituent of a two member onset consists of all valid one member onsets except \(/c, f/\)."\(^{20}\)
As a result of MR2 and MR3 only sixteen of the valid one member onsets may serve as the first constituent of two member onsets and only twenty-three can serve as the second constituent. The inventory of potential two member onsets resulting from the permutations of the two constituents equals 368.

\[
\begin{align*}
\langle 2^0 \rangle_I & \quad \langle 2^0 \rangle_{II} \\
p, b, t, d, k, g & \quad p, b, t, d, k, g \\
c, \dd, s, z, \dd, \dd & \quad c, d, \dd, \dd, \\
v, h, m & \quad s, z, \dd, \dd, \\
m, n, \dd, r, l, i, v, h, j
\end{align*}
\]

MR4 "The membership of the first constituent of two member codas consists of /s, \dd, z, \dd/.

\[
\langle 2^C \rangle \quad \exists \quad \begin{align*}
- \text{sonorant} \\
+ \text{continuant} \\
+ \text{coronal}
\end{align*}
\]

\[
P = 525 \quad \text{RP} = 84\%
\]

MR5 "The membership of the second constituent of two member codas consists of /t, d/.

\[
\langle 2^C \rangle \quad \begin{align*}
- \text{sonorant} \\
- \text{continuant}
\end{align*}
\]
As a result of MR4 and MR5 only four of the valid one member onsets may serve as the first constituent of two member codas and only two may serve as the second constituent. The inventory of potential two member codas resulting from the permutations of the two constituents equals 8.

\[
\begin{align*}
\langle z^C \rangle_I & \quad \langle z^C \rangle_{II} \\
s, z, c, \breve{c} & \quad t, d
\end{align*}
\]

3.222 Universal Juxtaposition Rules for two member onsets and codas.

UR1 "When two member onsets or codas consist of nonsonorants only, both segments are either voiced or nonvoiced.

\[
\langle 2^M \rangle \neq \begin{cases} \text{sonorant} & \text{voiced} \\ \text{sonorant} & \text{nonvoiced} \end{cases}
\]

OP = 91 ROP = 24.9% CP = 4 RCP = 50% TP = 95 RTP = 25.3%

The following potential two member onsets and codas are tagged ungrammatical by this rule:

**Onsets:**
- \(pb, pd, p\breve{z}, p\acute{d}, p\acute{j}, pg\)
- \(tb, td, tz, t\breve{z}, t\acute{d}, t\acute{j}, tg\)
- \(cb, cd, cz, c\breve{z}, c\acute{d}, c\acute{j}, cg\)
- \(sb, sd, sz, s\breve{z}, s\acute{d}, s\acute{j}, sg\)
- \(\breve{sb}, \breve{sd}, \breve{sz}, \breve{s}\breve{z}, \breve{s}\acute{d}, \breve{s}\acute{j}, \breve{sg}\)
- \(\breve{cb}, \breve{cd}, \breve{cz}, \breve{c}\breve{z}, \breve{c}\acute{d}, \breve{c}\acute{j}, \breve{cg}\)
- \(\breve{kb}, \breve{kd}, \breve{kr}, \breve{kz}, \breve{kd}, \breve{k\breve{d}}, kg\)

**Codas:**
- \(sd, \breve{zt}\)
- \(\breve{sd}, zt\)
3.223 Juxtaposition Rules for two member onsets.

OR1 "Two member onsets cannot consist of geminates".

\[
\left< 2^0 \right> \neq \left< \begin{array}{c} \text{configuration} \\ \text{configuration} \end{array} \right>_{\text{I}} \ast \left< \begin{array}{c} \text{configuration} \\ \text{configuration} \end{array} \right>_{\text{II}}
\]

\[
\text{OP} = 15 \quad \text{ROP} = 4.1\%
\]

The following potential two member onsets are tagged ungrammatical by this rule:

\[\text{pp, tt, ss, \dd, kk, bb, dd, zz, \dd, \dd, gg, vv, hh, mm}..\]

OR2 "Two member onsets cannot consist of a plus anterior sonorant and a nonsonorant."

\[
\left< 2^0 \right> \neq \left< \begin{array}{c} \text{sonorant} \\ \text{anterior} \end{array} \right>_{\text{I}} \ast \left< \begin{array}{c} \text{sonorant} \\ \text{-sonorant} \end{array} \right>_{\text{II}}
\]

\[
\text{OP} = 28 \quad \text{ROP} = 7.6\%
\]

The following potential two member onsets are tagged ungrammatical by this rule:

\[\text{vp, vt, vs, \dd, \dd, \dd, \dd, \dd, \dd, \dd, \dd, \dd, \dd, \dd}..\]

OR3 "Two member onsets cannot consist of a nonanterior sonorant and a nonsonorant except /t, d/.

\[
\left< 2^0 \right> \neq \left< \begin{array}{c} \text{sonorant} \\ \text{-anterior} \end{array} \right>_{\text{I}} \ast \left< \begin{array}{c} \text{sonorant} \\ \text{continuant} \\ \text{-coronal} \\ \text{-delayed release} \end{array} \right>_{\text{II}}
\]

\[
\text{OP} = 12 \quad \text{ROP} = 3.3\%
\]

The following potential two member onsets are tagged ungrammatical by this rule:

\[\text{hp, hs, h\dd, h\dd, h\dd, h\dd, h\dd, h\dd}..\]

OR4 "Two member onsets cannot consist of a nonsonorant except a plus anterior noncorononal and a plus continuant nonsonorant."
The following potential two member onsets are tagged ungrammatical by this rule:

\[ \left< 2^0 \right> \neq \left< \begin{array}{c} - \text{sonorant} \\ + \text{anterior} \end{array} \right> \uparrow \times \left< \begin{array}{c} - \text{sonorant} \\ + \text{continuant} \end{array} \right> \]

\[ \text{OP} = 44 \quad \text{ROP} = 18.2\% \]

The following potential two member onsets are tagged as ungrammatical by this rule:

\[ \text{ts, tš, tz, tž,} \]
\[ \text{cs, cš, cz, cž} \]
\[ \text{ss, sš, sz, šž} \]
\[ \text{šš, šš, šž, šž} \]
\[ \text{čš, čš, čž, čž} \]
\[ \text{ks, kš, kž, kž} \]
\[ \text{ds, dš, dz, dž} \]
\[ \text{zs, zš, zz, zž} \]
\[ \text{žš, žš, žž, žž} \]
\[ \text{žš, žš, žž, žž} \]
\[ \text{gs, gš, gz, gž} \]

OR5 "Two member onsets cannot consist of a noncontinuant plus coronal nonsonorant and a plus coronal nonsonorant."

\[ \left< 2^0 \right> \neq \left< \begin{array}{c} - \text{sonorant} \\ - \text{continuant} \end{array} \right> \uparrow \times \left< \begin{array}{c} - \text{sonorant} \\ + \text{coronal} \end{array} \right> \]

\[ \text{OP} = 50 \quad \text{ROP} = 13.6\% \]

The following potential two member onsets are tagged as ungrammatical by this rule:

\[ \text{tt, ts, tš, tč, tc, td, tz, tž, td, tž} \]
\[ \text{ct, cs, cš, cč, cd, cz, cž, cd, cž} \]
\[ \text{dt, ds, dš, dč, dš, dd, dz, dž, dd, dž} \]
OR6 "Two member onsets cannot consist of a noncontinuant plus anterior nonsonorant and a plus anterior noncoronal nonsonorant."

\[
\langle 2^0 \rangle \not\sim \left\langle \begin{array}{c}
-\text{sonorant} \\
-\text{continuant} \\
+\text{anterior} \\
+\text{coronal}
\end{array} \right\rangle
\]

\[
\epsilon \not\sim \left\langle \begin{array}{c}
-\text{sonorant} \\
+\text{anterior} \\
-\text{coronal}
\end{array} \right\rangle
\]

\[\text{OP} = 10 \quad \text{ROP} = 2.7\%\]

The following potential two member clusters are tagged ungrammatical by this rule:

- \(pp, pb, tp, tb, cp, cb\)
- \(bp, bb, dp, db\)

OR7 "Two member onsets cannot consist of a noncontinuant plus anterior plus delayed release nonsonorant and a nonsonorant."

\[
\langle 2^0 \rangle \not\sim \left\langle \begin{array}{c}
-\text{sonorant} \\
-\text{continuant} \\
+\text{anterior} \\
+\text{delayed release}
\end{array} \right\rangle
\]

\[\text{OP} = 14 \quad \text{ROP} = 3.8\%\]

The following potential two member onsets are tagged ungrammatical by this rule:

- \(cp, ct, cs, \tilde{c}s, \tilde{c}c, \tilde{c}e, ck, cb, cd, cz, \tilde{c}z, cd, \tilde{c}d, \tilde{c}f, cg\)

OR8 "Two member onsets cannot consist of a noncoronal nonsonorant and a noncoronal nonsonorant."

\[
\langle 2^0 \rangle \not\sim \left\langle \begin{array}{c}
-\text{sonorant} \\
-\text{coronal}
\end{array} \right\rangle
\]

\[\epsilon \not\sim \left\langle \begin{array}{c}
-\text{sonorant} \\
-\text{coronal}
\end{array} \right\rangle
\]

\[\text{OP} = 16 \quad \text{ROP} = 4.3\%\]
The following potential two member onsets are tagged ungrammatical by this rule:

\[ \langle 2^0 \rangle \not\in \langle - \text{sonorant} \mid + \text{anterior} \mid + \text{coronal} \rangle \not\in \langle - \text{sonorant} \mid - \text{anterior} \mid + \text{coronal} \rangle \]

OP = 30  ROP = 8.1%

The following potential two member onsets are tagged ungrammatical by this rule:

\[ \langle 2^0 \rangle \not\in \langle - \text{sonorant} \mid - \text{continuant} \mid + \text{coronal} \mid + \text{delayed release} \mid - \text{distributed} \rangle \]

OP = 4  ROP = 1.1%

OR9 "Two member onsets cannot consist of a plus anterior plus coronal nonsonorant and a nonanterior plus coronal nonsonorant."

OR10 "Two member onsets cannot consist of a nonanterior noncoronal nonsonorant and a nonsontinuant plus coronal plus delayed release nondistributed nonsonorant."
The following potential two member onsets are tagged ungrammatical by this rule:

\[
k', k\hat{c}, \hat{c}, s, s\hat{c}
\]

OR11 "Two member onsets cannot consist of a nonsonorant except a plus continuant plus anterior and a plus consonantal nonanterior noncoronal sonorant."

\[
\begin{align*}
\langle 2^0 \rangle & \neq \langle \text{sonorant} \mid \text{continuant} \rangle_1 \text{ or } \langle \text{sonorant} \mid \text{consonantal} \rangle_1 \\
\langle \text{syrllabic} \mid \text{anterior} \rangle_1 & \neq \langle \text{sonorant} \mid \text{anterior} \rangle_1 \\
\langle \text{continuant} \mid \text{coronal} \rangle_1 & \neq \langle \text{nasal} \mid \text{coronal} \rangle_1
\end{align*}
\]

\[\text{OP} = 14 \quad \text{ROP} = 3.8\%\]

The following potential two member onsets are tagged ungrammatical by this rule:

\[
ph, th, ch, \hat{sh}, ch, kh, bh, dh, \hat{zh}, \hat{zh}, gh, vh, hh, \text{m}
\]

OR12 "Two member onsets cannot consist of a nonsyllabic plus anterior noncoronal and a plus anterior noncoronal sonorant."

\[
\begin{align*}
\langle 2^0 \rangle & \neq \langle \text{syllabic} \mid \text{anterior} \rangle_1 \text{ or } \langle \text{sonorant} \mid \text{anterior} \rangle_1 \\
\langle \text{syllabic} \mid \text{coronal} \rangle_1 & \neq \langle \text{sonorant} \mid \text{coronal} \rangle_1
\end{align*}
\]

\[\text{OP} = 8 \quad \text{ROP} = 2.2\%\]

The following potential two member onsets are tagged ungrammatical by this rule:

\[
pm, pv, bm, bv, vm, vb, mm, mv
\]

OR13 "Two member onsets cannot consist of a noncontinuant nonsonorant except a nonanterior noncoronal and a plus nasal nonanterior sonorant."

\[
\begin{align*}
\langle 2^0 \rangle & \neq \langle \text{sonorant} \mid \text{continuant} \rangle_1 \text{ or } \langle \text{sonorant} \mid \text{nasal} \rangle_1 \\
\langle \text{continuant} \mid \text{coronal} \rangle_1 & \neq \langle \text{nasal} \mid \text{coronal} \rangle_1
\end{align*}
\]

\[\text{OP} = 7 \quad \text{ROP} = 1.9\%\]
The following potential two member onsets are tagged ungrammatical by this rule:

\[ \text{pñ, tñ, cñ, čñ, bn, dn, ŋñ} \]

OR14 "Two member onsets cannot consist of a nonsyllabic plus anterior noncoronal nonnasal and a plus nasal sonorant."

\[ \left< 2^0 \right> \neq \left< \begin{array}{|c|} \hline \text{-syllabic} \\ \text{-nasal} \\ \text{+anterior} \\ \text{-coronal} \\ \hline \end{array} \right> \; \{ \} \left< \begin{array}{|c|} \hline \text{+sonorant} \\ \text{+nasal} \\ \hline \end{array} \right> \]

\[ \text{OP} = 9 \quad \text{ROP} = 2.4\% \]

The following potential two member onsets are tagged ungrammatical by this rule:

\[ \text{pm, pn, ph} \]

\[ \text{bm, bn, bh} \]

\[ \text{vm, vn, vñ} \]

The onset /pn/ is found to be marginal as a result of this rule.

OR15 "Two member onsets cannot consist of a noncontinuant plus delayed release nonsonorant and a plus nasal coronal sonorant."

\[ \left< 0 \right> \neq \left< \begin{array}{|c|} \hline \text{-sonorant} \\ \text{-continuant} \\ \text{+delayed release} \\ \hline \end{array} \right> \; \{ \} \left< \begin{array}{|c|} \hline \text{+sonorant} \\ \text{+nasal} \\ \text{+coronal} \\ \hline \end{array} \right> \]

\[ \text{OP} = 6 \quad \text{ROP} = 1.6\% \]

The following potential two member onsets are tagged ungrammatical by this rule:

\[ \text{cn, čn, ŋn} \]

\[ \text{čn, čñ, ŋñ} \]

OR16 "Two member onsets cannot consist of a noncontinuant plus coronal nonsonorant and a plus lateral non anterior sonorant."
The following potential two member onsets are tagged ungrammatical by this rule:

\[
\langle 2^0 \rangle \neq \langle - \text{sonorant} \left| - \text{continuant} \right| + \text{coronal} \rangle \quad \text{OP} = 5 \quad \text{ROP} = 1.3\%
\]

OR17 "Two member onsets cannot consist of a noncontinuant plus delayed release or nonanterior nonsonorant and a conconsonantal sonorant."

\[
\langle 2^0 \rangle \neq \langle \langle - \text{continuant} \left| \langle + \text{delayed release} \right| \left| - \text{anterior} \rangle \right. \rangle \quad \text{OP} = 7 \quad \text{ROP} = 1.5\%
\]

The following potential two member onsets are tagged ungrammatical by this rule:

\[
\langle 2^0 \rangle \neq \langle - \text{sonorant} \left| - \text{anterior} \right| + \text{coronal} \rangle \quad \text{OP} = 5 \quad \text{ROP} = 1.3\%
\]

OR18 "Two member onsets cannot consist of a nonanterior plus coronal nonsonorant and nonnasal plus continuant plus coronal sonorant."

\[
\langle 2^0 \rangle \neq \langle - \text{anterior} \left| + \text{coronal} \right| \rangle \quad \text{OP} = 7 \quad \text{ROP} = 1.9\%
\]

The following potential two member onsets are tagged ungrammatical by this rule:

\[
\langle 2^0 \rangle \neq \langle - \text{anterior} \left| + \text{coronal} \right| \rangle \quad \text{OP} = 5 \quad \text{ROP} = 1.5\%
\]

OR19 "Two member onsets cannot consist of a noncontinuant plus anterior plus delayed release nonsonorant and a plus lateral plus coronal sonorant."
The following potential two member onsets are tagged ungrammatical by this rule:

\[ \langle 2^0 \rangle / \langle - \text{sonorant} \\
- \text{continuant} \\
+ \text{anterior} \\
+ \text{delayed release} \rangle \notin \langle + \text{sonorant} \\
+ \text{lateral} \\
+ \text{coronal} \rangle \]

\[ \text{OP} = 2 \quad \text{ROP} = .6\% \]

OR20 "Two member onsets cannot consist of nonnasal sonorant and a plus coronal nasal."

\[ \langle 2^0 \rangle / \langle + \text{sonorant} \rangle \notin \langle - \text{nasal} \rangle \]

\[ \text{OP} = 4 \quad \text{ROP} = 1.1\% \]

The following potential two member onsets are tagged ungrammatical by this rule:

\[ \langle 2^0 \rangle / \langle + \text{sonorant} \rangle \notin \langle + \text{nasal} \\
- \text{anterior} \rangle \]

\[ \text{OP} = 3 \quad \text{ROP} = .8\% \]

OR21 "Two member onsets cannot consist of a sonorant and a nonanterior nasal."

\[ \langle 2^0 \rangle / \langle + \text{nasal} \rangle \notin \langle - \text{anterior} \rangle \]

The following potential two member onsets are tagged ungrammatical by this rule:

\[ \langle 2^0 \rangle / \langle - \text{nasal} \\
+ \text{anterior} \rangle \notin \langle + \text{sonorant} \\
- \text{anterior} \\
+ \text{coronal} \rangle \]

OR22 "Two member onsets cannot consist of a nonnasal plus anterior sonorant and a nonanterior plus coronal sonorant."
The following potential two member onsets are tagged ungrammatical by this rule:

\[ vl, vn \]

OR23 "Two member onsets cannot consist of a nonanterior sonorant and a nonconsonantal sonorant."

\[ \langle 2^o \rangle \; \langle + \; \text{sonorant} \rangle \; \langle - \; \text{anterior} \rangle \; \langle - \; \text{consonantal} \rangle \]

The following potential two member onset is tagged ungrammatical by this rule:

\[ hj \]

The following list contains all the grammatical two member onsets and codas. Virtual clusters are marked with an asterisk. This list is followed by a list of the marginal two member onsets and codas.

**Grammatical Two Member Clusters**

Onsets:
- pt, ps, p\(\tilde{b}\), p\(\tilde{c}\), pr, pl, p\(\acute{l}\), pj
- bd, bz\(\ast\), b\(\acute{z}\)\(\ast\), b\(\tilde{y}\)\(\ast\), br, bl, bl, bj
- tk, tr, tl, tm, tn\(\ast\), tj, tv
- dg\(\ast\), dr, dl, dm, dn, dj, dv
- cr, cm, cv
- sp, st, sk, sr, sl, s\(\acute{i}\), sm, sn, s\(\acute{n}\), sj, sv, sh
- zb, zd, zg, zr, zl\(\acute{z}\)\(\ast\), zm, zn, z\(\acute{n}\), zj, zv, zh\(\ast\)
Marginal onsets are all found only in borrowings and fall into two classes:

1. /pn/, which utilizes native Serbo-Croatian phonemes.
2. /st, fr, fl/ which makes use of the phoneme /f/ which is not native
to Serbo-Croatian. /pn/ is tagged ungrammatical by OR14. The second class /sf,fr,fl/ is marginal because /f/ is not permitted to enter into phonotactic combinations on the same level as native Serbo-Croatian phonemes because /f/ does not act like a native phoneme. It does not pattern like its voiced counterpart /v/. To allow a more general statement of the Serbo-Croatian sound pattern /f/ is not allowed to enter into the input of phonotactic combinations. It is blocked by MR2 and MR3. Therefore combinations containing /f/ are automatically termed marginal.

Marginal codas are all found only in borrowings. They contain only phonemes native to Serbo-Croatian. Their constituent phonemes are not allowed to enter into phonotactic combinations because they interfere with the general sound pattern of native Serbo-Croatian codas. They are blocked by MR4 and MR5.

3.3 Three Member Onsets and Codas

3.31 Membership Rules

MR6 "The membership of the first constituent of three member onsets is limited to those valid one member onsets which consist of:

a) a plus continuant plus coronal nonsonorant
b) a nonanterior noncoronal plus voice nonsonorant
c) a plus consonantal nonanterior noncoronal sonorant.

\[
\begin{align*}
&\text{sonorant} \\
&+ \text{continuant} \\
&+ \text{coronal} \\
&- \text{sonorant} \\
&- \text{anterior}
\end{align*}
\]
There are eight subrules for the membership of the second constituent of three member onsets.

MR7.1 "The membership of the second constituent of three member onsets consists of those valid two member onsets which consist of a noncontinuant nonanterior or plus coronal nonsonorant and a nonsyllabic plus continuant plus coronal sonorant."

MR7.2 "The membership of the second constituent of three member onsets consists of those valid two member onsets which consist of a nonsyllabic plus anterior, noncoronal sonorant or nonvoiced noncontinuant nonsonorant and a nonsyllabic plus continuant plus coronal sonorant."

```
\left< 3^0 \right>_{11}
```

```
\left< 3^0 \right>_{11}
```

```
\left< 3^0 \right>_{11}
```

```
\left< 3^0 \right>_{11}
```

```
\left< 3^0 \right>_{11}
```
MR7.2 yields: pr, mr, vr

MR7.3 "The membership of the second constituent of three member onsets consists of those valid two member onsets which consist of a nonanterior noncoronal nonsonorant and a nonsyllabic plus lateral plus anterior sonorant."

\[ \langle 3^0 \rangle_{11,13} \begin{cases} \text{- sonorant} \\ \text{- anterior} \\ \text{- coronal} \end{cases} \begin{cases} \text{+ syllabic} \\ \text{+ sonorant} \\ \text{+ lateral} \\ \text{+ anterior} \end{cases} \]

MR7.3 yields: kl, gl

MR7.4 "The membership of the second constituent of three member onsets consists of those valid two member onsets which consist of a nonsyllabic plus anterior noncoronal nonnasal sonorant or nonvoiced nonsonorant and a nonsyllabic plus lateral plus anterior sonorant."

\[ \langle 3^0 \rangle_3 \begin{cases} \text{- syllabic} \\ \text{+ anterior} \\ \text{- coronal} \\ \text{- nasal} \end{cases} \begin{cases} \text{+ sonorant} \\ \text{+ lateral} \\ \text{+ anterior} \end{cases} \begin{cases} \text{+ sonorant} \\ \text{- sonorant} \\ \text{- voice} \end{cases} \]

MR7.4 yields: pl, vl

MR7.5 "The membership of the second constituent of three member onsets consists of those valid two member onsets which consist of a noncontinuant plus coronal non delayed release nonsonorant and a nonsyllabic nonconsonantal or nonnasal plus anterior noncoronal sonorant."

\[ \langle 3^0 \rangle_{11,17} \begin{cases} \text{- sonorant} \\ \text{- continuant} \\ \text{+ coronal} \\ \text{- delayed release} \end{cases} \begin{cases} \text{+ syllabic} \\ \text{+ sonorant} \end{cases} \begin{cases} \text{(- consonantal)} \\ \text{(- nasal)} \end{cases} \begin{cases} \text{+ anterior} \\ \text{- coronal} \end{cases} \]
MR7.5 yields: tv, dv
tj, dj

MR7.6 "The membership of the second constituent of three member onsets consists of those valid two member onsets which consist of a nonsyllabic plus consonantal nonanterior noncoronal nonvoiced and a nonsyllabic nonnasal plus anterior noncoronal sonorant."

\[
\begin{array}{c|c|c}
\langle 3^0 \rangle & - \text{syllabic} & - \text{sylablic} \\
& + \text{consonantal} & + \text{sonorant} \\
& - \text{anterior} & - \text{nasal} \\
& - \text{coronal} & + \text{anterior} \\
& - \text{voice} & - \text{coronal}
\end{array}
\]

MR7.6 yields: kv, hv

MR7.7 "The membership of the second constituent of three member onsets consists of those valid two member onsets which consist of a nonsyllabic plus anterior noncoronal sonorant and nonsyllabic nonconsonantal."

\[
\begin{array}{c|c|c}
\langle 3^0 \rangle & - \text{syllabic} & - \text{consonantal} \\
& + \text{sonorant} & + \text{sonorant} \\
& + \text{anterior} & + \text{nasal} \\
& - \text{coronal} & - \text{anterior}
\end{array}
\]

MR7.7 yields: mj, vj

MR7.8 "The membership of the second constituent of three member onsets consists of those valid two member onsets which consist of a nonanterior noncoronal plus voiced nonsonorant and a nonsyllabic plus nasal nonanterior sonorant."

\[
\begin{array}{c|c|c}
\langle 3^0 \rangle & - \text{sonorant} & - \text{sylablic} \\
& - \text{anterior} & + \text{sonorant} \\
& - \text{coronal} & + \text{nasal} \\
& + \text{voiced} & - \text{anterior}
\end{array}
\]

MR7.8 yields: gn

\[P = 2500 \quad \text{RP} = 83.3\%\]
As a result of MR6 and MR7 only six of the valid one member onsets may serve as the first constituent of three member onsets and only twenty of the valid two member onsets may serve as the second constituent. The inventory of potential two member codas resulting from the permutations of the two constituents equal 120.

\[ \binom{3}{0} \]

\[ s, s, z, z, g, h \]

\[ \binom{3}{1} \]

\[ pr, pl, tv, tr, tj, kv, kr, kl \]

\[ dv, dr, dj, gn, gr, gl \]

\[ mr, mj \]

\[ vr, vl, vj \]

\[ hv \]

MR8 "The membership of the first constituent of three member codas consists of the valid one member codas which is nonsyllabic nonconsonantal."

\[ \binom{3}{C} \]

\[ \binom{3}{C} \]

\[ P = 96 \quad RP = 96\% \]

MR8 yields: \( j \)

MR9 "The membership of the second constituent of three member codas consists of the valid two member coda which consists of a plus continuant plus anterior plus coronal nonvoiced nonsonorant and a noncontinuant plus anterior plus coronal non delayed release nonvoiced nonsonorant."

\[ \binom{3}{C} \]

\[ \binom{3}{C} \]

\[ P = 3 \quad RP = 75\% \]
MR9 yields: st

As a result of MR8 and MR9 only one valid one member coda may serve as the first constituent of three member codas and only one valid two member coda may serve as the second constituent. This results in the possibility of only one three member coda /jst/.

3.33 Juxtaposition Rules for Three Member Onsets

OR24 "The juxtaposition of the first and second constituents of three member onsets cannot result in nonsonorants with different values of voice."

\[
\begin{align*}
\langle 3^0 \rangle & \neq \langle - \text{sonorant} \mid \alpha \text{voice} \rangle_1 \sqcup \langle - \text{sonorant} \mid \alpha \text{voice} \rangle_1 \\
\text{OP} & = 36 \quad \text{ROP} = 30\%
\end{align*}
\]

The following potential three member onsets are tagged ungrammatical by this rule:

- sdv, sdr, sdj, sgň, sgr, sgl
- šdv, šdr, šdj, šgň, šgr, šgl
- zpr, zpl, ztv, ztr, ztj, zkř, zkr, zkl
- žpr, žpl, žtv, žtr, žtj, žkv, žkr, žkl
- gpr, gpl, gtv, gtr, gtj, gkv, gkr, gkl

OR25 "The juxtaposition of the first and second constituents of three member onsets cannot result in a nonanterior and a sonorant."

\[
\begin{align*}
\langle 3^0 \rangle & \neq \langle - \text{anterior} \rangle_1 \sqcup \langle + \text{sonorant} \rangle_1 \\
\text{OP} & = 24 \quad \text{ROP} = 20\%
\end{align*}
\]

The following potential three member onsets are tagged ungrammatical by this rule:

- šmr, šmj, švr, švl, švj, šhv
- žmr, žmj, žvr, žvl, žvj, žhv
The juxtaposition of the first and second constituents of three member onsets cannot result in a geminate.

\[ \langle 3^0 \rangle \neq \langle - \text{anterior} \rangle \left[ \begin{array}{c} \text{ant} \text{anterior} \end{array} \right] \langle - \text{coronal} \rangle \langle (\_\_\_) \rangle \]

OP = 4        ROP = 3.3%

The following potential three member onsets are tagged ungrammatical by this rule:

\[ ggn, ggr, egl, hhv \]

OR27 "The juxtaposition of the first and second constituents of three member onsets cannot result in a nonanterior noncoronal and a noncoronal."

\[ \langle 3^0 \rangle \neq \langle - \text{anterior} \rangle \left[ \begin{array}{c} \text{anterior} \end{array} \right] \langle - \text{coronal} \rangle \langle (\_\_\_) \rangle \]

OP = 28        ROP = 23.3%

The following potential three member onsets are tagged ungrammatical by this rule:

\[ gpr, gpl, gkv, gkr, gkl, eg gn, ggr, egl, gmr, gmj, gvn, gvl, gvj, ghv, hpr, hpl, hkv, hkr, hkl, hgn, hgr, hgl, hmr, hmj, hr, hv, hml, hhv \]

OR28 "The juxtaposition of the first and second constituents of three member onsets cannot result in a nonanterior noncoronal and a coronal and a plus consonantal."

\[ \langle 3^0 \rangle \neq \langle - \text{anterior} \rangle \left[ \begin{array}{c} \text{anterior} \end{array} \right] \langle - \text{coronal} \rangle \langle (\_\_\_) \rangle \]
OP = 8  ROP = 6.7%

The following potential three member onsets are tagged ungrammatical by this rule:

\[ gt\text{r}, gr, g\text{d}v, g\text{d}r \]
\[ h\text{t}v, h\text{r}, h\text{d}v, h\text{d}r \]

OR29 "The juxtaposition of the first and second constituents of three member onsets cannot result in a nonanterior and a plus anterior noncoronal."

\[ \langle 3^0 \rangle \not\in \langle \text{anterior} \rangle_1 \langle \text{coronal} \rangle_1 \]

OP = 28  ROP = 23.3%

The following potential three member onsets are tagged ungrammatical by this rule:

\[ \text{špr}, \text{špl}, \text{šmr}, \text{šmj}, \text{švr}, \text{švl}, \text{švj} \]
\[ \text{žpr}, \text{žpl}, \text{žmr}, \text{žmj}, \text{žvr}, \text{žvl}, \text{žvj} \]
\[ \text{gpr}, \text{gpl}, \text{gmr}, \text{gmj}, \text{gvr}, \text{gvl}, \text{gvj} \]
\[ \text{hpr}, \text{hpl}, \text{hmj}, \text{hmr}, \text{hvr}, \text{hvl}, \text{hvj} \]

OR30 "The juxtaposition of the first and second constituents of three member onsets cannot result in a nonanterior plus coronal and a plus coronal and a non coronal."

\[ \langle 3^0 \rangle \not\in \langle \text{anterior} \rangle_1 \langle \text{coronal} \rangle_1 \]

OP = 8  ROP = 6.7%

The following potential three member onsets are tagged ungrammatical by this rule:

\[ \text{štj}, \text{šd}v, \text{šd}j \]
\[ \text{žtv}, \text{žtj}, \text{žd}v, \text{žd}j \]
OR31 "The juxtaposition of the first and second constituents of three member onsets cannot result in a nonanterior plus coronal and a nonanterior and an alpha anterior minus alpha coronal."

\[
\langle 3^0 \rangle \not\equiv \langle -\text{anterior} \mid +\text{coronal} \rangle \land \langle -\text{anterior} \mid -\text{coronal} \rangle
\]

OP = 6  
ROP = \% 

The following potential three member onsets are tagged ungrammatical by this rule:

\[\hat{\text{skv}}, \hat{\text{zgn}}, \hat{\text{shv}}\]

\[\hat{\text{kv}}, \hat{\text{ggn}}, \hat{\text{hv}}\]

The following list contains all the grammatical three member onsets and codas. Virtual clusters are marked by an asterisk.

**Grammatical Three Member Clusters**

**Onsets:**

\[\text{spr}, \text{spl}, \text{stv}, \text{str}, \text{stj}^*, \text{skv}, \text{skr}, \text{skl}, \text{smr}, \text{smj}, \text{svr}, \text{svl}, \]
\[\text{svj}, \text{shv}, \]
\[\text{zdv}, \text{zdr}, \text{zdr}, \text{zgn}, \text{zgr}, \text{zgl}*, \text{zmr}^*, \text{zmj}^*, \text{zvr}^*, \text{zvl}^*, \text{zvj}^*, \]
\[\text{zhv}^* \]
\[\hat{\text{str}}, \hat{\text{skr}}, \hat{\text{skl}}^*, \]
\[\hat{\text{dr}}, \hat{\text{zgr}}^*, \hat{\text{zgl}}\]
\[\text{gdj} \]
\[\text{htj}, \text{hdj}^* \]

**Codas:**

\[\text{jst} \]

4.0 **Observations on the Juxtaposition Rules**

The juxtaposition rules are discussed in this section in serial order. The emphasis of the discussion is on the nature of the constraints. The
constraints are in features which fall into four classes:

1. Major Class Features
   syllabic
   sonorant
   consonantal

2. Cavity Features
   nasal
   lateral
   anterior
   coronal
   distributed
   low
   back
   high

3. Manner of Articulation Features
   continuant
   delayed release

4. Source Features
   voice

4.1 The Rules

The most significant feature in each rule is underlined.

URL Universal rule one is the only universal rule. It is a source feature rule. It constrains nonsonorants of different voice. Its ROP is 24.9% and its RCP is 50% and its RTP is 25.3%. It is also the only constraint rule for codas.

OR1 Onset rule one is a mixed rule, i.e., a rule whose constraints
encompass features in more than one class. It blocks geminate nonsyllabics. Its ROP is 4.1%.

OR2 Onset rule two is primarily a major class feature rule. It blocks labial (plus anterior noncoronal) sonorants and nonsonorants. Its ROP is 7.6%.

OR3 Onset rule three is primarily a major class feature rule. It blocks the nonanterior sonorants and nonsonorants except /t/. Its ROP is 3.3%.

OR4 Onset rule four is primarily a manner of articulation rule. It blocks nonsonorants except labials and nonsonorant continuants. Its ROP is 18.2%.

OR5 Onset rule five is primarily a cavity feature rule. It blocks coronal noncontinuants nonsonorants and coronal nonsonorants. Its ROP is 13.6%.

OR6 Onset rule six is primarily a cavity feature rule. It blocks anterior noncontinuant nonsonorants and labial nonsonorants. Its ROP is 2.7%.

OR7 Onset rule seven is a manner of articulation feature rule. It blocks anterior affricates and nonsonorants. Its ROP is 3.8%.

OR8 Onset rule eight is a cavity feature rule. It blocks noncoronal nonsonorants and noncoronal nonsonorants. Its ROP is 4.3%.

OR9 Onset rule nine is primarily a cavity feature rule. It blocks anterior coronal nonsonorants and coronal nonsonorants.

OR10 Onset rule ten is primarily a cavity feature rule. It blocks nonanterior noncoronal nonsonorants and nondistributed affricates.
OR11 Onset rule eleven is primarily a major class feature rule. It blocks nonsonorants except /s/ and the sonorant /h/. Its ROP is 3.8%.

OR12 Onset rule twelve is a cavity feature rule. It blocks labials and labial (plus anterior noncoronal) sonorants. Its ROP is 2.2%.

OR13 Onset rule thirteen is a mixed rule in that it blocks noncontinuant nonsonorants except velars and a nonanterior nasal. This involves the manner of articulation feature continuant and the cavity feature nasal. Its ROP is 1.9%.

OR14 Onset rule fourteen is a cavity feature rule. It blocks a labial nonsyllabic and a nasal. Its ROP is 2.2%. It tags the cluster /pn-/ marginal. This cluster is found to exist only in loan words.

OR15 Onset rule fifteen is primarily a manner of articulation feature rule and secondarily a cavity feature rule. It blocks affricates and coronal nasals. Its ROP is 1.6%.

OR16 Onset rule sixteen is primarily a cavity feature rule. It blocks coronal noncontinuants and nonanterior laterals. Its ROP is 1.3%.

OR17 Onset rule seventeen is a major class feature rule. It blocks a number of nonsonorants and the nonsyllabic nonconsonantal. Its ROP is 1.9%.

OR18 Onset rule eighteen is a manner of articulation rule. It blocks nonanterior coronals and the continuant sonorant /r/. Its ROP is 1.1%.

OR19 Onset rule nineteen is a mixed rule. It is part cavity feature and part manner of articulation. It blocks anterior affricates and laterals. Its ROP is 5.5%.
OR20  Onset rule twenty is a cavity feature rule. It blocks nonnasal
sonorants and nasal sonorants. Its ROP is 1.1%.

OR21  Onset rule twenty-one is primarily a cavity feature rule. It blocks
sonorants and nonanterior nasals. Its ROP is .8%.

OR22  Onset rule twenty-two is a cavity feature rule. It blocks anterior
anterior nonnasal sonorants and nonanterior coronal sonorants. Its ROP
is .6%.

OR23  Onset rule twenty-three is a mixed rule. It is part major class and
part cavity feature rule. It blocks a nonanterior sonorant and a noncon-
onsonantal sonorant. Its ROP is .3%.

OR24  Onset rule twenty-four is a source feature rule. It blocks nonson-
orants of different voice values. Its ROP is 30%.

OR25  Onset rule twenty-five is primarily a cavity feature rule. It blocks
nonanteriors and sonorants. Its ROP is 20%.

OR26  Onset rule twenty-six is a mixed rule. It blocks geminates. However,
it may be said to block geminate nonanteriors. Its ROP is 3.3%.

OR27  Onset rule twenty-seven is a cavity feature rule. It blocks non-
anterior noncoronals and noncoronals. Its ROP is 23.3%.

OR28  Onset rule twenty-eight is a major class feature rule. It blocks
nonanterior noncoronals and coronals and plus consonantals. Its ROP is 6.7%.

OR29  Onset rule twenty-nine is a cavity feature rule. It blocks nonanteriors
and plus anterior noncoronals. Its ROP is 23.3%.
OR30 Onset rule thirty is a cavity feature rule. It blocks nonanterior and plus coronals and noncoronals. Its ROP is 6.7%.

OR31 Onset rule thirty-one is a cavity feature rule. It blocks nonanterior plus coronals and nonanterior plus alpha anterior minus alpha coronals. Its ROP is 5%.

5.0 Concluding Remarks

One very obvious fact is that all but one of the juxtaposition rules apply to onsets only. URL is the only juxtaposition rule needed for codas. Serbo-Croatian has twenty-five one member onsets and codas, but one hundred and twenty-two two member onsets to only four two member codas and thirty-five three member onsets to only one three member coda. It is clear that Serbo-Croatian distinguishes differences to a much greater extent at the beginning of syllables rather than at the end.

5.01 Feature Observations

The features may be ranked according to the number of clusters they constrain. This is done by adding the ROP of each rule. The ROP is added as a number not as a percentage. In each rule one feature is the dominant feature. If in 4.1 two features are underlined in a single rule the ROP is divided equally between them. Feature classes may be ranked in the same manner.

Features are ranked by cumulative relative power in the following order:

<table>
<thead>
<tr>
<th>Rank</th>
<th>Cumulative relative power</th>
<th>Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>57.4</td>
<td>anterior</td>
</tr>
<tr>
<td>2</td>
<td>56.0</td>
<td>coronal</td>
</tr>
<tr>
<td>3</td>
<td>55.3</td>
<td>voice</td>
</tr>
<tr>
<td>4</td>
<td>19.3</td>
<td>continuant</td>
</tr>
</tbody>
</table>
Feature classes are ranked by cumulative relative power in the following order:

<table>
<thead>
<tr>
<th>Rank</th>
<th>Cumulative Relative Power</th>
<th>Feature class</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>122.1</td>
<td>cavity</td>
</tr>
<tr>
<td>2</td>
<td>55.3</td>
<td>source</td>
</tr>
<tr>
<td>3</td>
<td>25.0</td>
<td>manner of articulation</td>
</tr>
<tr>
<td>4</td>
<td>23.3</td>
<td>major cavity</td>
</tr>
</tbody>
</table>

The features exhibiting the greatest constraining powers are anterior, coronal, and voice. Together these three features account for about 75% of the juxtaposition constraints in Serbo-Croatian. Anterior and coronal account for about half of the constraints. This high value is probably due to a loss of the feature palatalization in Serbo-Croatian which remained in some other Slavic languages such as Russian. Thus Serbo-Croatian depends more upon the place of articulation (cavity features) than on manner of articulation.

Voicing is of course a major constraint since clusters of obstruents cannot have different voice values within a microsegment, as is true for all Slavic languages. Seven features combine to form the remaining quarter of constraints. The highest, continuant, is separated from voice by
thirty-six points at 19.3. It is followed in turn by sonorant 14.7, consonantal 8.6, nasal 6, lateral 1.6, and distributed 1.1.

As classes cavity features account for about 9% of the juxtaposition constraints in Serbo-Croatian; source features about 25%; manner of articulation features 11% and major cavity features 10%.
APPENDIX

Summary List of Grammatical Onsets and Codas of Russian

In the following list virtual clusters are marked by an asterisk and marginal clusters are listed at the end of its respective set.

<table>
<thead>
<tr>
<th>One Member Onsets</th>
<th>One Member Codas</th>
</tr>
</thead>
<tbody>
<tr>
<td>p, b, t, d, k, g</td>
<td>p, b, t, d, k, g</td>
</tr>
<tr>
<td>c, č, č, č, ʃ</td>
<td>c, č, č, č, ʃ</td>
</tr>
<tr>
<td>f, v, s, z, ŋ, ŋ, h</td>
<td>f, v, s, z, ŋ, ŋ, h</td>
</tr>
<tr>
<td>m, n,  ŋ, r, l, ₁, j</td>
<td>m, n,  ŋ, r, l, ₁, j</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Two Member Onsets</th>
</tr>
</thead>
<tbody>
<tr>
<td>pt, ps, pČ, pč, pr, pl, pl, pj</td>
</tr>
<tr>
<td>bd, bz*, bž*, bž*, br, bl, bl, bj</td>
</tr>
<tr>
<td>tk, tr, tl, tm, tn, tj, tv</td>
</tr>
<tr>
<td>dg*, dr, dl, dm, dn, dj, dv</td>
</tr>
<tr>
<td>cr, cm, cv</td>
</tr>
<tr>
<td>sp, st, sk, sr, sl, s₁, sm, sn, sn, sj, sv, sh</td>
</tr>
<tr>
<td>zb, zd, zg, zr, z₁, z₁*, zm, zn, zn, z₁, zv, zh*</td>
</tr>
<tr>
<td>čp*, čk, čl, čm, čv</td>
</tr>
<tr>
<td>žb, žg*, žl*, žm*, žv*</td>
</tr>
<tr>
<td>kt, kČ, kr, kl, kl, km, kn, kn, kv</td>
</tr>
<tr>
<td>gd, gd, gr, gl, gl, gm, gn, gn, gv</td>
</tr>
<tr>
<td>vr, vl, vj</td>
</tr>
</tbody>
</table>
ht, hd*, hr, hl, hhl, hm, hv
mr, ml, ml, mj

Marginal two member onsets
pn
sf, fr, fl

Two member codas
st
zd
št
žd

Marginal two member codas
ps, kt
rc, lm, nt, nd, js

Three member onsets
spr, spl, stv, str, stj*, sky, skr, skl, smr, smj, svr, svl, svj, shv
zd, zdr, zdj, zgá, zgr, zgl*, zmj*, zmr*, zvr*, zvl*, zvj*, zrv*
št, škr, škl*
ždr, žgr*, žgl
gdj
htj, hdj*

Three member codas
jst

Juxtaposition Rules ranked by decreasing relative power

<table>
<thead>
<tr>
<th>Rank</th>
<th>Relative Power</th>
<th>Rule</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>30</td>
<td>OR24</td>
</tr>
<tr>
<td>Rank</td>
<td>Relative Power</td>
<td>Rule</td>
</tr>
<tr>
<td>------</td>
<td>---------------</td>
<td>---------------</td>
</tr>
<tr>
<td>2</td>
<td>25.3</td>
<td>Ur1</td>
</tr>
<tr>
<td>3</td>
<td>23.3</td>
<td>{Or27, Or29}</td>
</tr>
<tr>
<td>4</td>
<td>20.0</td>
<td>Or25</td>
</tr>
<tr>
<td>5</td>
<td>18.2</td>
<td>Or4</td>
</tr>
<tr>
<td>6</td>
<td>15.9</td>
<td>Or5</td>
</tr>
<tr>
<td>7</td>
<td>8.1</td>
<td>Or9</td>
</tr>
<tr>
<td>8</td>
<td>7.6</td>
<td>Or2</td>
</tr>
<tr>
<td>9</td>
<td>6.7</td>
<td>{Or28, Or30}</td>
</tr>
<tr>
<td>10</td>
<td>5.0</td>
<td>Or31</td>
</tr>
<tr>
<td>11</td>
<td>4.3</td>
<td>Or8</td>
</tr>
<tr>
<td>12</td>
<td>4.1</td>
<td>Or1</td>
</tr>
<tr>
<td>13</td>
<td>3.8</td>
<td>{Or7, Or11}</td>
</tr>
<tr>
<td>14</td>
<td>3.3</td>
<td>Or3</td>
</tr>
<tr>
<td>15</td>
<td>2.7</td>
<td>Or6</td>
</tr>
<tr>
<td>16</td>
<td>2.2</td>
<td>{Or12, Or14}</td>
</tr>
<tr>
<td>17</td>
<td>1.9</td>
<td>{Or13, Or17}</td>
</tr>
<tr>
<td>18</td>
<td>1.6</td>
<td>Or15</td>
</tr>
<tr>
<td>19</td>
<td>1.3</td>
<td>Or16</td>
</tr>
<tr>
<td>20</td>
<td>1.1</td>
<td>{Or10, Or18, Or20}</td>
</tr>
<tr>
<td>21</td>
<td>0.8</td>
<td>Or21</td>
</tr>
<tr>
<td>22</td>
<td>0.6</td>
<td>Or22</td>
</tr>
<tr>
<td>23</td>
<td>0.3</td>
<td>Or23</td>
</tr>
</tbody>
</table>
Notes

1. The methodology here is the same as originated by R. Saunders, 1970, in his doctoral dissertation Phonological Constraints in Russian Syllable Margins. (Brown University)

2. This is an extension of the attitude taken by Kučera 1958, where he discusses two phonemic systems in Czech. One system is composed of native phonemes and the other system has /g/ a nonnative phoneme as its only member. This allows Kučera to make more general statements about Czech phonology.

3. Hockett 1955, originates the procedure pp.43-51. Kučera 1961, applies the same methodology for Czech but arrives at a more exact definition of the phonological syllable, pp.43-84. Saunders 1970, again uses the same methodology along with Kučera's definition of a phonological syllable and applies it to Russian pp.4-47.

4. Kučera 1961. Also see Pike 1947, p.144, where the phonological syllable is defined as the basic structural unit which serves best as a point of reference for describing the distribution of the phonemes of the language in question. Neither Kučera nor Saunders makes note of Pike's definition.


9. Ibid.


12. Saunders, 1970, p.46. "In my article, 'Asyllabic Residues in Russian', CJO 11:2 (1966), I favored the third alternative whereby the ICM plus its following disjuncture were considered to be constituents of the following syllable. This alternative, however, comes into conflict with our notion of a hierarchy of size levels. Our position was that a unit of a given size level could not extend beyond the
boundaries of units of the next higher size level. If we incorporate the ICM plus its following disjuncture into the onset of the following syllable, we are doing precisely what we just established we could not do, i.e., extending a unit of one size level, the syllable beyond the boundary of a unit on the next higher size level, the microsegment. Therefore this alternative is unsatisfactory.

18. Hodge, 1946, includes /f/ in this class, as does Grozdić 1969. I, however, have found no evidence to support this claim, which I feel is made by analogy to /v/, and have therefore excluded /f/ from this class.
19. /f/ is excluded from this set for reasons given in 0.3.
20. Ibid.
21. For a contrast in Russian see Saunders 1970, section 5.0.
BIBLIOGRAPHY


Bach, E., 1968, 'Two proposals concerning the simplicity metric in phonology,' *Glossa* 2, 128-49.


Delattre, P., 1967, 'Acoustic or articulatory invariance?', *Glossa* 1, pp.3-25.


Greenberg, J. 1962, 'Is the vowel-consonant dichotomy universal?' *Word* 18, pp.73-81.

---------- 1965, 'Some generalizations concerning initial and final consonant clusters,' *linguistic* 18, pp.5-34.


Halle, M., 1962, 'Phonology in a generative grammar,' Word 18, pp.54-72.


-------- 1968, Introduction to Phonological Theory (Englewood Cliffs, N.J.)


Hodge, C. 1946, 'Serbo-Croatian Phonemes,' Language 22, pp.112-20.


King, H., 1964, 'English internal juncture and syllable division,' Languages and Linguistics (Ann Arbor) pp. 199-211.


Malmberg, B. 1955, 'The phonetic basis for syllable division,' Studia Linguistica 9, pp. 80-87.


Martinet, A., 1955, Economie des changements phonetique (Bern).

---------- 1960, Elements de linguistic generale. (Paris).


Paravić, 1904, Grammaire de la Langue Serbo-Croate (Paris).


Pike, K., 1947, Phonemics (Ann Arbor).


ESSAY II

MARKEDNESS IN GENERATIVE PHONOLOGY

by

Steven Uzelac
Table of Contents

0.0 Introduction  page 1
0.1 Source of data  page 1
0.2 Motivations of Marking Conventions  page 1
0.3 Purpose of this Study  page 1

1.0 The Marking Conventions  page 2
1.1 Categories of Marking Conventions  page 2
1.11 Segment versus Non-segment Convention  page 2
1.12 Segmental Structure Conventions  page 2
1.13 Naturalness Condition Conventions  page 3

2.0 Conclusions  page 10

Notes  page 11

Bibliography  page 13
0.0 Introduction

0.1 Source of data. The most comprehensive discussion of markedness in generative phonology is to be found in Chomsky-Halle 1968, and therefore forms the basis for the discussion in this study.

0.2 Motivation. The concept of markedness in a phonological theory derives directly from the hypothesis that, phonological "features have intrinsic content." That is to say, the theory of distinctive features, such as presented in Chomsky-Halle 1968, assumes that the phonology of a language should be analysed and described with close regard to its phonetic properties. For example,

"When a child has determined the phonetic representation of a form, he has thereby determined a great part of its phonological structure." Chomsky 1968: 299-300 gave a list of about forty binary distinctive features specified in largely articulatory terms which they propose to use to describe the phonology of any natural language. Some of these feature values, however, are believed to be more natural than others, i.e., natural in the sense of being more expected. Chomsky and Halle therefore propose a system of marking conventions whereby segments composed of more natural feature values are less marked and those containing more less-natural feature values are more marked. Only marked feature values contribute to the complexity of a grammar. This gives a quantitative measure of the naturalness of a system, i.e., of its deviation from the expected as defined. This means that matrices are no longer given in plus or minus feature values but in M (marked) and U (unmarked) feature values according to the dictates of the marking conventions.

0.3 Purpose of this study. The marking conventions shall then here be examined as to their types and the functions they perform and then conclusions shall be drawn as to their usefulness.
1.0 The Marking Conventions

1.1 I divide the marking conventions into three categories:

(i) segment versus non-segment,

(ii) segmental structure

(iii) naturalness of systems and rules

1.11 Marking convention (I) is segment versus non-segment:

(I) \([\text{useg}] \rightarrow [\text{-seg}]\)

This convention states that the most natural segment is a non-segment, which could be interpreted as saying that the most natural speech is no speech at all. In a milder form this can be stated as shorter utterances are more natural than longer ones. However, there may be another intention. On page 371 \([-\text{seg}]\) is an essential feature in distinguishing a boundary. On the same page Chomsky and Halle state that they regard boundaries as units in the string, "boundaries must be regarded as units in the string." They are required in Chomsky and Halle's stress rules. One known universal is that all languages have boundaries. Therefore it may be stated that the most natural thing about language is that it has boundaries.

1.12 Segmental structure is dealt with in conventions (II a b) and (III a c d). They state that the most natural segment to occur after a boundary is a C (consonant) and the segment that follows the C should most naturally be a V (vowel), i.e., + CV...+ is the most natural multi-segmental structure. This is a syntagmatic relationship based on the principle of maximal differentiation, i.e., C is at the opposite pole of V. Note that this may be why Chomsky and Halle reintroduce the feature \([\text{vocalic}]\) after earlier rejecting it in favour of \([\text{syllabic}]\). This reintroduction of \([\text{vocalic}]\) prevents syllabic liquids and nasals from being as natural as a vowel. These conventions, however, do not account for /a/ being the simplest entry in a lexicon, unless, priority is given to convention (I) \([\text{useg}] \rightarrow [\text{-seg}]\)
and C in convention (II a) \([\text{cons}] \rightarrow [\text{cons}] /+\) is written as \([-\text{seg}]\). But this would have the unfortunate consequence of calling \([-\text{seg}]\) a C in which case it would be the most unmarked segment, but /a/ is claimed to be the most unmarked segment, "It was observed previously that the least marked segment, given the conventions outlined here, is the vowel /a/."\(^{15}\)

The only solution is to interpret convention (1) as saying shorter utterances are more natural than longer ones, in which case /a/ = +V+ is more natural than +CV...+

1.13 The remaining conventions are naturalness conditions on systems and rules. They are used for three purposes:\(^{16}\)

(i) eliminating lexical redundancies. (ii) securing optimal or least marked phonemic systems. (iii) linking: simplification of the statement of natural rules, vis à vis segmental relationships.

1.131 Conventions such as

(iv) \([+\text{voc}] \rightarrow [+\text{son}]\) 
(vii) \([+\text{low}] \rightarrow [-\text{high}]\)\(^{17}\)

are lexical redundancy conventions. Convention (iv) will make \([+\text{voc}]\) segments unmarked for \([\text{son}]\) because that value of \([\text{son}]\) will always be predictable. Convention (vii) is a restriction based on a quasi-logical argument which follows from the intrinsicness assumption, i.e., the body of the tongue is not high and low at the same time. But this is a question of segmentation. What if a diphthong were to be regarded as a single segment? Notice on page 288 in table (8) where \([\text{T}] \rightarrow [\text{aey}]\). If it is to be assumed that \([\text{aey}]\) comprises only one segment, and it is less economical to assume otherwise, \((\text{seg} \rightarrow [\text{seg}] \text{[seg]}\) is a contrary and less natural process according to marking convention (1) \([\text{Useg}] \rightarrow [-\text{seg}]\), then the segment
[æɪ] can be regarded as being both [+low] and [+high]. 18

1.132 In conjunction with conventions (II c d) and (III b); conventions 
(V-XII) produce least marked vowel systems and with conventions (XIII-XXVII) 
produce the least marked true consonantal system. Conventions (XXVIII-
XXXIV) produce the least marked liquid and conventions (XXXIV-XXXIX) 
deal with glides.

1.1321 Optimal vowel systems. Table (7) gives a list of twelve vowels 
with their marked and unmarked values.

<table>
<thead>
<tr>
<th>(7)</th>
<th>a</th>
<th>i</th>
<th>u</th>
<th>æ</th>
<th>e</th>
<th>ə</th>
<th>o</th>
<th>u</th>
<th>t</th>
<th>ɛ</th>
<th>ø</th>
</tr>
</thead>
<tbody>
<tr>
<td>low</td>
<td>u</td>
<td>u</td>
<td>u</td>
<td>m</td>
<td>m</td>
<td>u</td>
<td>u</td>
<td>u</td>
<td>m</td>
<td>u</td>
<td>u</td>
</tr>
<tr>
<td>high</td>
<td>u</td>
<td>u</td>
<td>u</td>
<td>u</td>
<td>m</td>
<td>m</td>
<td>u</td>
<td>u</td>
<td>m</td>
<td>m</td>
<td></td>
</tr>
<tr>
<td>back</td>
<td>u</td>
<td>-</td>
<td>+</td>
<td>m</td>
<td>u</td>
<td>-</td>
<td>+</td>
<td>m</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>round</td>
<td>u</td>
<td>u</td>
<td>u</td>
<td>m</td>
<td>u</td>
<td>u</td>
<td>m</td>
<td>m</td>
<td>m</td>
<td>m</td>
<td></td>
</tr>
<tr>
<td>complexity</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

The complexity of a system is given in (8),

"The complexity of a system is equal to the sum of the 
marked features of its members." 19

/a/ is the least marked vowel followed by /i/ and /u/ which are only 
marked for one feature. /a/ must then be included in any optimal system.

Chomsky and Halle do not discuss an optimal two vowel system. Halle, in 
Halle 1970, however, recognizes the existence of the optimal two vowel 
 system /a-θ/. 20 He agrees with and quotes Roman Jakobson that /a-θ/ are, 
"die einzigen, die nirgends fehlen dürfen". 21 With the information given 
in table (7) (above) an arbitrary decision would have to be between /a,i/ 
and /a,u/ in choosing an optimal two vowel system. This raises the question 
of why Chomsky and Halle have not included /θ/ in table (7). One reason, 
of course, is that /θ/ is not accounted for in the marking conventions.

The conventions account only for peripheral vowels. Another reason is that 
Chomsky and Halle may not want to include /θ/, for if /θ/ were to be in- 
cluded and to form part of the optimal two vowel system then it would have 
to be incorporated in any larger system in lieu of further conditions, and
Chomsky and Halle do not include /a/ in any larger system. The proposed optimal three vowel system is /a i u/ supported by statistical criteria, "On the same basis we can say that the three-vowel system /a i u/ is the simplest possible, a conclusion that seems to be supported by its predominance over other three-vowel systems in the languages of the world." Condition (8), however, is unable to select an optimal five vowel system because there is no way to choose among /æ e o ü t/ all of which have a complexity of 2. Chomsky and Halle, however, claim that the optimal five vowel system is /a i u e o/. Therefore further conditions are added. Condition (9) eliminates /æ o/ from (7).

(9) "No vowel segment can be marked for the feature 'round' unless some vowel segment in the system is marked for the feature 'high'."

A further condition, (10), eliminates /æ/. (10) "Other things being equal, a system in which more features have only the specification u is preferable to a system in which fewer features have only the specification u*." This leaves /e o/ to join /a i u/ to produce the desired optimal five vowel system. Notice, however, the arbitrariness of conditions (9) and (10). The only motivation in formulating them is that they produce the desired five vowel system. In order to obtain the desired four vowel system which is hypothesized as /a i u æ/ Chomsky and Halle talk of imposing, parallel to (10), a, "symmetry' condition on systems which would be expressed in terms of specified features rather than in terms of markedness. This would be our analog to the traditional conception of 'filling holes in the phonological pattern'."

They do not develop this because they do not, "have a sufficient understanding of the empirical issues involved; in particular, we are not certain that the system /a i u æ/ occupies the privileged position that we have suggested it does."
Such a condition would also be necessary to produce an optimal six vowel system, which given condition (10), would be:

\[
\begin{array}{ccc}
\text{i} & \text{u} & \text{u} \\
\text{e} & \text{o} & \text{e} \\
\text{a} & & \text{a}
\end{array}
\]

and not the expected (given the optimal four and five vowel systems):

\[
\begin{array}{ccc}
\text{i} & \text{u} & \\
\text{e} & \text{o} & \\
\text{æ} & \text{a}
\end{array}
\]

Note that condition (9) is no longer operative because segments in the systems, /e/ and /o/, are marked for high. The same indeterminacy exists for vowel segments containing greater numbers of segments. It seems therefore that Chomsky and Halle are content to be able to identify just the three and five optimal vowel systems, which it seems all that they were concerned with identifying in any event.

**1.1322 Optimal consonantal systems.** According to the matrix (12) there are five equally least marked consonants. They are /p t k s n/. Notice that among the vowels /a/ was the least marked, being marked for \(\text{seg}^\text{+}\) only, but there is no such, "unmarked consonant". Notice that when Chomsky and Halle give examples of the simplest possible lexical entries of more than one segment they use any of the five optimal consonants and /a/, i.e., "/pata/, /tata/, /kata/, /sasa/, /nana/". If there were one consonant, /C^m/, that was marked for only \(\text{seg}^\text{+}\) then the simplest lexical entry of more than one segment would be /C^m+a/. But when it came to the simplest possible entry then there would be no way to choose between /C^m/ and /a/. So by making only /a/ marked for plus segment and the consonants all marked for at least one more segment no such problem is encountered.
This also leads to the conclusion that vowels are more natural than consonants. The five above consonants are all marked for at least one feature. Chomsky and Halle say that,

"In view of the fact that the unmarked nasal must be /n/ and the unmarked continuant /s/, the unmarked plosive - if there to be one - would have to be /t/. This conclusion, however, seems unacceptable to us; in particular, the choice of the dental over the labial plosive appears incorrect." 31

They do not carry the discussion any further to ascertain the least marked consonant or to explain their preference for a labial - which seems to be alluding to Jakobson 1968.31 When a system containing more than five consonants is attempted the same impossible situation is faced as was the case in trying to determine vowel systems with more than five segments. For example, the members of both /b d g z/ and /b c x f/ are only marked for one more feature.33 The first set naturally being the more 'natural'. Chomsky and Halle say,

"Such conventions as those proposed for vowel systems (see (9) and (10)) might provide the correct results in this case, but our understanding of the case is too rudimentary for a detailed proposal to be of any value." 34

1.1323 The Optimal liquid. "Conventions (XXIX) and (XXXI) specify that the unmarked liquid is a dental." 35 "Conventions (XXXIII) and (XXXIV) specify that the unmarked coronal liquid is lateral (i.e., /l/ rather than /r/), and continuant."36 Notice that Chomsky and Halle have finally come to a decision as to whether /r/ or /l/ is a continuant and with no apparent explanation. They had left the question open on page 318.

1.1324 The conventions are unable to produce an optimal glide.

1.1325 Now that such an optimal consonant and vowel system exists, we may ask of what use is it? If the purpose is to get as few marked features as possible in a matrix then marking is a device whereby the segments in an utterance are not only contrasted but display a degree of 'naturalness' in
inverse proportion to the number of marked features needed to distinguish them, i.e., the closer they approach the natural or optimal system the more natural they are. But of what use is this? This says nothing about the grammar. This becomes meaningless unless Chomsky and Halle were to state that languages strive for the unmarked state. Chomsky and Halle have been careful in avoiding such a statement, but without it the concept of markedness and optimal systems is worthless, except statistically. No one could defend the statement that languages strive for the unmarked with such a tremendous number of examples where languages change from what Chomsky and Halle call unmarked to marked. What Chomsky and Halle fail to notice is that there are two forces in language; one towards the least effort or least resistance, for ease of articulation; and the other towards more resistance in order to create meaningful contrasts. With this in mind, it becomes meaningless to speak of 'optimal' systems derived from a statistical survey. It might, however, be meaningful to speak of natural systems from a learning (child-learning for example) or physiological basis, but this already seems to have been attempted by Roman Jakobson in 1940.

1.33 Linking. Marking conventions are used to simplify the more 'natural' processes in some phonological rules so that less features have to be stated for the same rule. This is the linking function. For example, in Slavic, the First Velar Palatalization is formulated in the following rule:

\[
\text{(25) } [-\text{ant} ] \rightarrow \begin{bmatrix}
-\text{back} \\
+\text{cor} \\
+\text{del rel} \\
+\text{strid}
\end{bmatrix} / \begin{bmatrix}
-\text{cons}
\end{bmatrix}
\]

so that

\[
\begin{align*}
k & \rightarrow \breve{c} \\
g & \rightarrow \breve{j} \\
x & \rightarrow \breve{s}
\end{align*}
\]

However, by making use of the linking capacity of the marking conventions the rules may be reformulated as:
(26) \([-\text{ant}] \rightarrow [-\text{back}] / \overset{\text{-cons}}{[-\text{back}]}

The output of this rule is then linked to the output by rules (XXIII b), (XXVI a) and (XXVII c)\(^41\), i.e.

Convention (XXIII b)

\[[\text{ucor}] \rightarrow [\text{\textless cor}] / [\text{-\textless back} - \text{ant}]\]

makes the output [+cor]

Convention (XXVI a)

\[[\text{udel rel}] \rightarrow [+\text{del rel}] / [-\text{ant} [+ \text{cor}]]\]

makes the output [+del rel]

Convention (XXVII c)

\[[\text{ustrid}] \rightarrow [\text{\textless strid}] / [-\text{del rel} [\{ [+\text{ant}] [+\text{cor}] \}]\]

makes the output [+strid]

Rule (26) is now linked to the output. Therefore a process can be stated more economically, i.e., by using fewer features, in this case three fewer features, if it is compatible with the linking processes found in the marking conventions, thus showing a 'natural' process.

1.331 Linking rules are supposed to contribute to the naturalness of a grammar. If they do, then the measure is an artificial one.\(^42\) In 1.133 it was shown how a rule was shown to be more natural by extracting certain processes and calling these universal natural processes, thereby permitting rule (25) to be written with three less features as (26).\(^43\) But this is surely just a paper saving device. What difference does it make that certain marking conventions are applied or the 'natural' processes are left in the rule unless it can be shown that this is a necessary process?
Like the rules which produce optimal vowel and consonant systems, the linking rules are only observationally adequate because they are merely the product of a statistical analysis. If Chomsky and Halle were to achieve descriptive and explanatory adequacy then they must construct a theory to explain why a 'natural' system or process should occur; this theory could then logically be extended so changes in systems could be shown to be necessary and predictable. Chomsky-Halle 1968, however, are content to be confined to observational adequacy.

2.0 Conclusions

2.1 The marking conventions in Chomsky-Halle 1968 merely represent a statistical analysis of what is common to an unspecified number of languages.

2.2 The marking conventions found in Chomsky-Halle 1968 are limited to only observational adequacy because Chomsky and Halle put forward no explicit theory to explain:

(i) what naturalness is.

(ii) if the natural or unmarked segments necessarily manifest themselves at the expense of the unnatural or marked, i.e., is there a necessary change from the marked to the unmarked, and why?

(iii) if (ii) is true how can the reverse process be explained?

(iv) the prediction of change in (ii) and (iii).
Notes

1 A list of the marking conventions is found in Chomsky-Halle 1968: 404-407. Also see Postal 1968 Chapter 8 for a discussion of markedness.

2 Chomsky-Halle 1968: 400.

3 See Chomsky-Halle 1968: 293-389 for the list and discussion of the distinctive features used here.

4 Postal 1968: 56.

5 Chomsky-Halle 1968: 400-402.


8 Ibid.


10 Ibid.

11 This was pointed out to me by E.W. Roberts.


16 Chomsky-Halle 1968: 400-402.


18 See Roberts 1969-70"Phonology l"for a discussion of a different segmentation than found here.


20 Halle 1970: 95, 103.


23 Ibid.: 410.


25 Ibid.
26 Ibid.
27 Ibid.
29 Chomsky-Halle 1968:413.
31 Chomsky-Halle 1968:413.
32 See note 6.
34 Ibid.
35 Ibid.
36 Ibid.
37 See Roberts 1969-70: Phonology 2, for a discussion of this point.
38 Also see Miller 1951 for the results of a similar study.
42 See Chomsky-Halle 1968:401 for what a natural rule should be as opposed to an unnatural one. The only criterion of naturalness is the use of fewer features.
44 See Chomsky 1964:63 for a discussion of the three levels of adequacy: 1. observation, 2. descriptive, 3. explanatory.
Bibliography


