The Phonotactics of Hungarian

Miklós Törkenczy

Budapest
2004
For J & K

*Human existence de Selby has defined as ‘a succession of static experiences each infinitely brief’, a conception which he is thought to have arrived at from examining some old cinematographic films which belonged probably to his nephew. [...] Apparently, he had examined them patiently picture by picture and imagined that they would be screened in the same way, failing at that time to grasp the principle of the cinematograph.*

   Flann O’Brien: The Third Policeman

---

*For it is plain, that every word we speak is in some degree a diminution of our lungs by corrosion, and consequently contributes to the shortening of our lives.*

   Jonathan Swift: Gulliver’s Travels
Contents

Acknowledgements [5]
Abbreviations [6]
Transcription [8]

Chapter 1. Introduction: Aims, scope and coverage and layout
  1.1. Aims and scope [9]
  1.2. Principal claims [10]
  1.3. Chapter layout [15]

Chapter 2. Preliminaries
  2.1. Theoretical background: phonotactics in phonological theory [17]
     2.1.1. The division of labour between the phonotactic and the non-phonotactic aspects of phonology: levels [18]
     2.1.2. The domain of phonotactic constraints [21]
     2.1.3. Phonotactic strata [25]
     2.1.4. Degrees of phonotactic well-formedness, irregularities, accidental gaps [26]
  2.2. Framework and theoretical assumptions [28]

Chapter 3. ‘Static’ phonotactics: the phonological shape of the Hungarian word
  3.1 Introduction [36]
  3.2. Syllable structure: SSCs [37]
     3.2.1. The Hungarian syllable template: the basic syllable types [37]
     3.2.2. Onsets—word-initial clusters [39]
     3.2.3. Rhymes [48]
     3.2.4. Codas—word-final consonant clusters [51]
        3.2.4.1. Non-branching codas [51]
        3.2.4.2. Branching codas [51]
        3.2.4.3. Appendices [65]
  3.3. Transsyllabic constraints [75]
     3.3.1. Hiatus [75]
     3.3.2. Intervocalic consonant clusters [83]
        3.3.2.1. Two-member clusters [83]
        3.3.2.2. Clusters consisting of more than two members [101]
3.4. Morpheme structure: MSCs [105]
3.4.1. Domain-final open syllables and the minimal word/stem [105]
3.4.2. VVCC: the complexity of the rhyme [115]
3.4.3 Word-class-specific constraints: the phonotactics of verbs [121]
3.5. Sequence constraints [138]

Chapter 4. ‘Dynamic’ phonotactics: phonotactically motivated processes
4.1. Vowel–zero alternations [139]
4.1.1. Stem-internal unstable vowels: ‘epenthetic’ stems [139]
4.1.2. Stem-external vowel–zero alternations: stem-final unstable vowels and
‘linking’ vowels [144]
4.1.2.1 Stem-final vowel–zero alternations [144]
4.1.2.2 Suffix-initial vowel–zero alternation [145]
4.1.3. Lowering [153]
4.1.4. Analysis [163]
4.1.4.1. Syllabification—full vowels and defective vowels [163]
4.1.4.2. Major stems and ‘epenthetic’ stems—Type A and Type B
suffixes [167]
4.1.4.3. Syllabification and lowering [172]
4.1.4.4. The past suffix [180]
4.1.4.5. Analytic affixes and appendices [195]
4.1.4.6. OCP effects, residual problems [203]
4.1.4.6.1. ‘Epenthetic’ stems [203]
4.1.4.6.2. /t/-final stems [205]
4.2. Alternations involving consonants [212]
4.2.1. Alternating v-suffixes: -val/-vel, -vá/-vé [212]
4.2.2. h-alternations [219]
4.2.3. Fast Cluster Simplification [224]

Appendix A. The feature composition of underlying segments in Hungarian [230]
Appendix B. Lists [232]

Bibliography [245]
Acknowledgements

In thinking about, fumbling towards, preparing for, working on, writing and simply doing this dissertation I have been (undeservedly) lucky to receive ideas, reflections, wisdom, advice, instruction, criticism, friendship and various forms of help from a great number of people, which I would like to gratefully acknowledge here. While undoubtedly some have influenced my thinking and have helped me do this particular work more than others I find it impossible to come up with a fair ‘ranking’ so I will simply list the names in alphabetical order. The people I mention here were, are or have been my teachers, colleagues, students and/or friends. Without them I would never have come up with this (while, naturally, all the flaws and blunders are my very own): András T. László, Antal László, Bánréti Zsoltán, Bárányi Zsuzsa, Bartos Huba, Cser András, Jacques Durand, É. Kiss Katalin, Ferge Sándor, Gárdai Kinga, Morris Halle, John Harris, Harry van der Hulst, Kálmán László, Kassai Ilona, Jonathan Kaye, Michael Kenstowicz, Kiefer Ferenc, Kiss Zoltán, Komlósy András, Kornai András, Kürti (Hamp) Anna, Jean Lowenstamm, Lukács Ágnes, Marosán Lajos, Nádasdy Ádám, Novák Attila, Oravecz Csaba, Péter Mihály, Polgárdi Krisztina, Rebrus Péter, Catherine Ringen, Siptár Péter, Donca Steriade, Szende Tamás, Szentgyörgyi Szilárd, Szigetvári Péter, Trón Viktor, Robert Vago, Varga László, Wenszky Nóra, Cheryll Zoll.

I also want acknowledge the support of various grants and scholarships that I received while working on projects that lead to and/or eventually became parts of this dissertation: the Telegdi Zsigmond Grant (Center for Advanced Study in the Behavioral Sciences, September 1996 and December 1996,); the Short Study Grant to the Holland Institute of Linguistics, University of Leiden (Hungarian Soros Foundation 1998); the Széchenyi Professorship (1997-2000); the Fulbright Research Grant to the Massachusetts Institute of Technology (1999-2000); the Széchenyi Zsigmond Grant (2001-2004).

I would also like to thank my parents for always thinking I’m a clever boy.

And, finally and above all, everything is due to Júlia and Kristóf for absolutely everything.
Abbreviations

abl. ablative
acc. accusative
adj. adjective
ant anterior (feature)
ATR advanced tongue root
C consonant
CC cluster of two consonants
CCC cluster of three consonants
Co coda
camp. comparative
cond. conditional
cons consonantal (feature)
cont continuant (feature)
COR coronal node
dat. dative
def. definite conjugation
dim. diminutive
DOR dorsal node
ECH Educated Colloquial Hungarian
FCS fast cluster simplification
FSVS final stem vowel shortening
GP government phonology
imp. imperative
ind. indicative
indef. indefinite conjugation
iness. inessive
inf. infinitive
instr. instrumental
intr. intransitive
IPA International Phonetic Alphabet
ISVS internal stem vowel shortening
L laryngeal node
LAB labial node
lat lateral (feature)
LVL law vowel lengthening
MSC morpheme structure constraint
N nucleus or placeless nasal
nas nasal (feature)
NPA nasal place assimilation
O onset
OCP Obligatory Contour Principle
OP [+open,]
OVS object-verb-subject
The transcription symbols used in the dissertation are standard IPA except for the ones listed below (on the choice of symbols see Nádasdy and Siptár 1989):

<table>
<thead>
<tr>
<th>Present dissertation</th>
<th>IPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vowels:</td>
<td></td>
</tr>
<tr>
<td>ö</td>
<td>ø</td>
</tr>
<tr>
<td>öː</td>
<td>øː</td>
</tr>
<tr>
<td>ü</td>
<td>y</td>
</tr>
<tr>
<td>üː</td>
<td>yː</td>
</tr>
<tr>
<td>Consonants:</td>
<td></td>
</tr>
<tr>
<td>ţ</td>
<td>c</td>
</tr>
<tr>
<td>d̑</td>
<td>tʃ</td>
</tr>
<tr>
<td>š</td>
<td>ş</td>
</tr>
<tr>
<td>ž</td>
<td>ʒ</td>
</tr>
<tr>
<td>ṭ</td>
<td>ts</td>
</tr>
<tr>
<td>č</td>
<td>tʃ</td>
</tr>
<tr>
<td>ľ</td>
<td>dʒ</td>
</tr>
<tr>
<td>n̑</td>
<td>n</td>
</tr>
</tbody>
</table>
Chapter 1.
Introduction: Aims, scope, coverage, and layout

1.1. Aims and scope

This dissertation\(^1\) has a dual aim. On the one hand, it is intended to give a comprehensive and
detailed account of the phonotactic pattern of Hungarian, i.e. to describe what is a
phonologically possible Hungarian word in terms of the strings of segments it consists of. On
the other hand, the ‘static’ phonotactic regularities discovered will be shown to be active in a
‘dynamic’ way as well by conditioning phonological processes, chiefly vowel–zero alternations
and consonant–zero alternations (which are phonologically universally prone to phonotactic
conditioning).

The dialect this study focusses on is Educated Colloquial Hungarian (ECH) (cf.
Nádasdy 1985), i.e. the ‘educated’ Budapest variety of Hungarian. In addition to native
speaker judgements that underlie all data and generalizations presented in this dissertation, the
description of the phonotactic phenomena discussed here is based on a computerized database
(cf. Kornai 1986) comprising phonological (and other types of) information concerning
approximately 80 000 lexical items. I shall always indicate it clearly whenever I discuss other
varieties of Hungarian and/or use other sources of data.

\(^1\)The dissertation is based on chapters Five (Phonotactics: Syllable Structure, pp. 95-153)
and Eight (Processes Conditioned by Syllable Structure, pp. 214-277) of *The Phonology of
Hungarian* (Siptár and Törkenczy 2000) a recent monograph I co-authored with Péter Siptár
published by Oxford University Press in the series *The Phonology of the World’s Languages.*
Naturally, the chapters and material included here in a revised and extended form are those
that have been written exclusively by the present author.
1.2. Principal claims

In this section I present an overview of the specific principal claims that I make and argue for in the dissertation. Since the bulk of the dissertation is of analytical nature, most of the claims are analytic ones and only some of them are theoretical. For the convenience of the reader I have thoroughly cross-referenced each claim to the sections of the dissertation where I discuss them in detail.

(i) There are two syllable templates in Hungarian, the core syllable template, which is restricted to Block 1 of the derivation and the extended syllable template, which only becomes available for syllabification in Block 2. The extended syllable template may contain an Appendix preceding the Onset and/or an Appendix following the coda. There are no sonority sequencing violations in the core syllable, but may such violations may occur in the extended syllable. See sections 2.2; 3.2.4.3.; 4.1.4.5.

(ii) Hungarian words may begin with more than one consonant, but there are no branching onsets in Hungarian (cf. section 3.2.2). Word-initial consonants in word initial consonant clusters are licensed by being syllabified into the Appendix (cf. sections 3.2.4.3.; 4.1.4.5.)

(iii) In Hungarian there are no regular intervocalic CCC clusters undivided by an analytic boundary (cf. section 3.3.2.2.). Monomorphemic words containing intervocalic clusters of more than two consonants phonologically are treated in Hungarian as if they were compounds, i.e. a morphologically unitary domain is phonologically analysed as if it were two independent domains (cf. section 4.1.4.5.).

(iv) Hungarian permits maximally binary branching codas. Word-final consonant clusters consisting of more than two consonants and some word-final two member clusters are licensed as coda + appendix. The identity of these ‘extra’ word-final
consonants in word-final consonant clusters that are licensed by being syllabified into the Appendix is determined by morphology (the analytic suffixes -j, -d, -sz) and lexical marking. See section 3.2.4.

(v) Language-specific variations on the universal sonority hierarchy may exist, but only inasmuch as different language particular settings of sonority distance between segment classes are possible, language particular sonority ‘reversals’ are unpermitted (cf. section 2.2.)

The sonority hierarchy for Hungarian is

- stops, affricates < fricatives < < nasals < < liquids
- l << r << j

(where < is a smaller sonority distance than < < )

This manifests itself in the licensing constraints on coda clusters and interconstituent clusters (cf. sections 3.2.4.2. and 3.3.2.1.)

(vi) Codas licensed by government (right-to-left) and/or root or place binding. See section 3.2.4.

(vii) In Hungarian heteromorphemic hiatus is not permitted across a synthetic domain edge. Tautomorphemic hiatus is restricted by universal constraints (such as the Obligatory Contour Principle) and language-specific ones (two alternative hypotheses are offered: one disallows hiatus altogether, the other restrict it two some maximally two-member vowel sequences). See section 3.3.1.

(viii) The Syllable Contact Law is inoperative in Hungarian in general (cf. section 3.3.2.) although the phonotactics of verbs is skewed towards it in terms of type frequency (3.4.3).

(ix) Interconstituent clusters are licensed by government (right-to-left or left-to-right), root-binding, and Sp-licensing (cf. 3.3.2.) - constrained by the antilabial constraint (Chapter 3, (43), the antifrictative constraint (Chapter 3, (52) and the antipalatal
constraint (Chapter 3, (45)).

(x) \( /v/ \) is phonotactically asymmetrical: it behaves as an obstruent when in coda position, but it behaves as a sonorant when in onset position (Chapter 3, (51)).

(xi) The minimal word/stem in Hungarian is bimoraic (Chapter 3, (62))

(xii) In addition to SSCs, MSCs also determine the Hungarian phonotactic pattern. The relevant MSCs determine the distribution of long vs. short vowels domain finally (Chapter 3, (58), (62) (the minimal word/stem constraint), (65), (66), (67)) and the distribution of long vs. short vowels before consonant clusters (Chapter 3, (70ab) (*VVCC))

(xiii) There are phonotactic constraints in Hungarian that are independent of syllable or morpheme structure (cf. Chapter 3. (81)).

(xiv) The phonotactics of (monomorphemic) verbs is more restrictive than that of non-verbs. this manifests itself in the stricter conditions on government and binding in the case of codas (Chapter 3, (71), (73), (74), (75), (76)) and in frequency effects in the case of intervocalic clusters. See 3.4.3.

(xv) According to their phonological behaviour suffixes fall into the following classes:
   a. analytic (no vowel-zero alternation)
   b. synthetic
      Type A (phonotactically unmotivated vowel-zero alternation): they have an underlying initial full vowel (e.g. -\( V/k \) ‘pl’)  
      Type B (phonotactically motivated vowel-zero alternation): they are underlyingly consonant-initial (e.g. -\( t \) ‘acc’).

(xvi) Three types of vowel-zero alternation are distinguished: (i) stem-internal vowel-zero alternation, (ii) stem-final vowel-zero alternation (which is arbitrarily restricted to a few suffixes/stems and involves the deletion of the stem-final vowel)
and (iii) suffix-initial vowel–zero alternation. Of these (ii) is considered morphological (phonologically irregular). (i) and (iii) are analysed in detail. It is argued that (i) can be analysed as neither deletion nor epenthesis, but is the default spell-out (Default V) of defective vowels (i.e. vowels that only consist of a skeletal slot without any segmental melody), which are lexically present in the underlying representation of ‘epenthetic stems’ (Chapter 4, (22)). (iii) is analysed as the result of vowel deletion to repair hiatus (the rule Hiatus: Chapter 4, (28)) in Type A suffixes, and as the result of epenthesis (overparsing) by syllabification that inserts a defective vowel to break up an illicit cluster in Type B suffixes. There is no separate epenthesis rule: vowel–zero alternation is due to syllabification and default-spell-out.

(xvii) Syllabification is a right-to-left template-matching algorithm in Hungarian and is non-exhaustive (it can skip defective vowels (Section 4.1) but continuous (i.e. may reapply after each affixation and/or application of a phonological rule). A defective vowel can only syllabify in a singly closed syllable (Chapter 4.(23)). Syllabification is not cyclic in Hungarian (cf. 4.1.4.3.).

(xviii) A lowering stem/suffix underlyingly has a final floating [+ open₁] feature and a morpheme-final defective vowel V₂. Lowering is a process that spreads the floating [+ open₁] feature locally to a (full or defective) vowel which is incorporated into a syllable and is at the edge of a morpheme (Chapter 4.(33)). This spreading is a feature filling process.

(xix) The rules Hiatus and Lowering show derived environment effects, but they are not cyclic. Arguably, all Hungarian phonological rules apply non-cyclically, even the ones that only apply in derived environments (Section 4.1.4.3).

(xx) The past suffix is a /t/ whose root node is associated to a single timing slot followed by an empty timing slot (i.e. a timing slot devoid of melodic content) which is completely invisible to syllabification. The past suffix ends in a floating [+ open₁]
feature and a defective vowel $V_d$ because it is lowering. The empty timing slot of the suffix is filled by spreading from the root node of the preceding /t/ if the /t/ is preceded by a full vowel (/t/-spread: Chapter 4 (42)) There is no lexical degemination in Hungarian.

(xxii) Given the representations of ‘epenthetic’ stems, the accusative and the past tense suffix assumed, vowel-zero alternation in all these cases is the result of a unitary process: syllabification.

(xxii) The occurrence of variation in vowel-zero alternation is due to the existence of parallel stems (parallel underlying representations). See section 4.1.4.4.

(xxiii) The rules Hiatus (Chapter 4, (28)) and Default V (Chapter 4, (22)) only apply at Block 1.

(xxiv) Monoconsonantal analytic suffixes and irregular initial and final consonant clusters ones in monomorphic words syllabify (into the extended syllable) at Block 2

(xxv) The Obligatory Contour Principle is responsible for the special restrictions on the distribution of consonants flanking the unstable vowel of an ‘epenthetic’ stem (Section 4.1.4.6.1.).

(xxvi) The stem external vowel-zero alternation after /t/-final stems is due (a) to the morphological OCP-motivated rules (Chapter 4 (64) and (65)), which break up (64) and merge (65) the root nodes of a fake geminate and (b) lexical marking.

(xxvii) The difference between the behaviour of alternating (e.g. instrumental -val/-vel) and non-alternating v-suffixes (e.g. deverbal noun-forming -vány/-vény) does not derive from an underlying representational difference but from the fact that the former are synthetic and the latter are analytic. The alternation in alternating v-suffixes is due to the interaction between the rules C-spread (Chapter 4 (72)), which is a
generalised form of /t/-spread (Chapter 4 (42) and replaces it in the grammar, and v-delink (Chapter 4 (71)).

(xxviii) There is no non-postlexical degemination (what seems as degemination in the lexical phonology is really the lack of gemination (spreading)) (sections 4.1.4.4. and 4.2.1.).

(xxix) Productive $h$-alternation ([h]-[x]) is due to a rule that deletes the place node of an underlying voiceless dorsal fricative which is unspecified for [consonantal] in onset position (Chapter 4 (74)). Non-productive $h$-alternation ([h]-[o]) is non-phonological (lexical allomorphy).

(xxx) Fast Cluster Simplification is phonotactically motivated, but is unrelated to syllable or morpheme structure (Chapter 4 (82)).

1.3. Chapter layout

The chapter layout of the dissertation is organised so as to be suitable for the two major objectives. There are two central chapters (which take up the bulk of the dissertation): ‘Chapter 3. “Static” phonotactics: the shape of the Hungarian word’ which discusses syllable structure, morpheme structure, and ‘sequence’ phonotactics; and ‘Chapter 4. “Dynamic” phonotactics: phonotactically motivated processes’ in which phonotactically conditioned alternations are analysed. The two central chapters are preceded by a general introduction ‘Chapter 1. Introduction: Aims, scope, coverage, and layout’ and ‘Chapter 2. Preliminaries’, which is a more theoretically or non-analytically oriented chapter that focusses on two topics: the theory phonotactics in general and its place and status within (various) phonological framework(s) (‘2.1. Theoretical background: phonotactics in phonological theory’); and the general theoretical assumptions underlying the present work (‘2.2. Framework and theoretical assumptions ’). Finally, I have attached two appendices: Appendix A, which is intended to serve as reference on the featural composition of Hungarian segments assumed in the
dissertation, and Appendix B, in which I present exhaustive and near-exhaustive lists of lexical items that contain ‘interesting’ or unusual substrings of segments (and which is the result of many hours of frantic search in various databases and sources).
Chapter 2.
Preliminaries

2.1 Theoretical background: phonotactics in phonological theory

In this section I shall examine the central theoretical issues of phonotactic analysis in general, discuss the structuralist and the generative phonotactic tradition, highlight some problems/moot points, and identify the theoretical status of the phonotactic analysis presented in the dissertation.

Phonotactics is ‘the study of the positions occupied by phonological units relative to one another’.\(^1\) This definition is probably general enough to be acceptable for a phonologist of any theoretical affiliation with the qualification that the phonological unit referred to in the definition is assumed to be of segment size\(^2\) (although in principle the distribution of any phonological unit could be meant – such as the distribution of various types of syllables\(^3\)). Furthermore, it is also generally assumed that the largest unit within which segmental distribution is to be examined from a phonotactic point of view is the phonological word. The reason is that phonotactic regularities do not seem to apply across word boundaries and thus the main objective of a phonotactic analysis is to characterise what is a phonologically well-formed word\(^4\) (in other words, which string of segments can form a phonotactically

---

\(^1\)Algeo (1978: 206)

\(^2\)This naturally does not mean that the phonotactic statements expressing the regularities of the distribution of segment-sized units cannot refer to phonological units smaller than the segment (e.g. features, nodes in a feature-tree) or larger than the segment (e.g. the morpheme, the syllable).

\(^3\)In English, for instance, two unstressed syllables cannot begin a word.

\(^4\)There are aspects of phonological well-formedness that are not phonotactic in nature, for example, the well-formedness of words related to alternation patterns. A word may be phonologically ill-formed even though it is phonotactically well-formed. For instance, *săr\(\text{ar}t\) [ʃaːr̩t] ‘mud’ (accusative) is ill-formed because the vowel of the stem săr [ʃaːr] ‘mud’ shortens
2.1.1. The division of labour between the phonotactic and the non-phonotactic aspects of phonology: levels

Structuralist and post-structuralist (generative) phonological frameworks are typically multi-level in the sense that in order to account for what they consider to be phonological regularities (which may be (partially) different in different frameworks) they postulate (at least) two levels of representation (the phonological/underlying and the phonetic/surface) and a mapping process that maps the phonological representation onto the surface one. Given that in a multi-level framework there are thus (at least) two alphabets (inventories) of segment-sized phonological units, this creates a level problem for phonotactics: which is the level at which phonotactic regularities should be analysed/stated? As the patterning of sounds, i.e. the distribution of (the properties) of sounds in natural language is the object of phonological study, how much and which aspect(s) of this distributional information is to be handled by the phonotactics and how much and which aspect(s) by the mapping component?

before suffixes like the accusative (*sárat [ʃarət]). However, there is nothing anomalous about the combination of segments in *sárat [ʃarət] – it is just as well-formed phonotactically as attested várat [vaɾət] ‘castle’ (accusative), [jaɾət] járat ‘passageway’. I will return to this problem below.

5This is not meant to imply that there may be no phonotactic regularities that are expressible with reference to the (adjacent) string of segments involved exclusively, i.e. without reference to a larger phonological unit within which the string in question is positioned (‘sequence constraints,’ see section 2.1.2).

6Note that this problem does not even arise in one-level models (such as (some versions of) Declarative Phonology) where phonology basically is phonotactics (cf. Bird and Ellison 1994, Novák 1998, Bird, Coleman, Pierrehumbert and Scobie 1992, etc.), and this can be seen as an advantage of these models. The same is true of other recent constraint-based models that dispense with the underlying representation (Burzio 1996, 2002). Although Optimality Theory (Prince and Smolensky 1993, McCarthy and Prince 1993, 1995) differentiates between input and output and is thus not a one-level model, it avoids the level problem since – because of the ‘Richness of the Base’ (Smolensky 1996) – it claims that there are no language-specific restrictions on the input, no phonotactics of the lexicon and the surface pattern emerges as a result of the interaction of Gen and Eval applied to any input (cf. Kager 1999, McCarthy
In structuralist phonology (e.g. Trubetzkoy 1939/1969, Harris 1951, Hockett 1958) phonotactics refers to the combination of phonemes, so phonotactic regularities are stated at the phonological (‘phonemic’) level and thus apply to the phonological representation. As morphophonological alternations are considered to lie outside phonology proper, the mapping between the phonological representation and the phonetic one consists in the application of allophonic statements. Thus, the division of labour between phonotactics and allophonic statements is this: allophonic statements express/incorporate distributional information about non-contrastive (redundant) features (more precisely: features that are always non-contrastive (redundant) in the system\(^7\) and phonotactic statements express the distribution of contrastive features.\(^8\)

As in generative phonology in general morphophonology (alternations) is part of phonology, and the mapping (the derivation of the surface representation from the underlying one via phonological rules) is allowed to manipulate distinctive as well as redundant features (in the same way\(^9\)), underlying representations are more abstract and the difference between underlying and surface phonotactics is significant. For instance, in Standard British English words cannot end in a consonant cluster that consists of a nasal followed by a non-coronal voiced plosive ([*mb#, *ŋg#]). However, there are morphemes that can surface with a final cluster of this kind, though not when they are word-final. In the latter case the voiced stop is not realised: bomb [bɒm] - bombard [bɒmbɔ:d], strong [strɔŋ] - strong [strɔŋɡ]. If we consider these to be regular (non-suppletive) alternations, underlying forms with final /mb, ŋg/ are postulated, from which the voiced stops are deleted by a phonological rule word-finally. The obvious question is: Are strings with final nasal+ non-coronal stop clusters phonologically well-formed in English? The problem is that underlyingly they are, while at the surface they

\(^7\)given the ‘once a phoneme, always a phoneme’ principle (e.g. Gleason 1955, Jones 1957)

\(^8\)In Trubetzkoy this is more subtle: importantly, phonotactics also determines/underlies the status of a phonological opposition, whether it is ‘constant’ or ‘neutralisable’ (cf. Trubetzkoy 1939/1969).

\(^9\)cf. Halle (1959)
are not. The early and classical generative\textsuperscript{10} approach to the problem was to confine phonotactics to the underlying representation (in the form of statements (redundancy rules) about morpheme structure (MSCs), see below) and the surface pattern was assumed to be accounted for by MSCs and phonological rules in conjunction. Surface phonotactics had no theoretical status in these models: ‘[…] the facts about permissible and impermissible sequences at the output level are non-significant, the fortuitous product of the restrictions on underlying representations plus the phonological rules’ (Sommerstein 1977: 193-194). It soon became clear, however, that this division of labour is fraught with serious deficiencies since surface phonotactic constraints do play a role in the mapping as well\textsuperscript{11} (‘derivational constraints’, cf. Kisseberth 1970, Shibatani 1973) and it was argued that surface phonotactic constraints are necessary, and it is the underlying phonotactic constraints that are redundant and have no theoretical status (e.g. Sommerstein 1974, 1977). A more recent formulation of the same stance is that phonotactic constraints (in the form of syllable structure constraints) are ‘everywhere rules’ (cf. Kenstowicz 1994) that are active throughout the derivation but do not apply to the underlying representation (trivially, since underlying representations are unsyllabified). This is (basically\textsuperscript{12}) the approach that we take in the present dissertation.

In Lexical Phonology (e.g. Kiparsky 1982\textit{ab}, 1985), this picture can become somewhat more intricate because the modular organisation of the lexical phonology/morphology makes it possible to express the well-known difference between the phonotactics of different phonological/morphological levels/strata, in particular, the fact that the phonotactics of root-level derivatives is typically the same as that of undervived words while the phonotactics of word-level derivatives is usually very different. This can be accommodated by allowing partially different phonotactic rules to apply in the different lexical levels/strata.

---

\textsuperscript{10}cf. Halle (1962), Chomsky and Halle (1968), Postal (1968), Anderson (1974)

\textsuperscript{11}Another (related) deficiency was that the approach lead to a ‘duplication problem’: sometimes the same generalisation had to be stated twice in the phonology: once as an MSC, once as a phonological rule, cf. Kenstowicz and Kisseberth (1977, 1979) for a detailed discussion.

\textsuperscript{12}Though we also claim that MSCs, i.e. phonotactic constraints whose domain is the morpheme (\textit{not} the feature-filling redundancy rule MSCs of SPE) are also necessary to account for the phonotactic pattern of Hungarian (see sections 2.1.2, 3.4).
The analysis given in the present dissertation follows this by postulating two syllable templates that express phonotactic restrictions, the ‘core syllable’ and the ‘extended’ syllable templates, and only allowing core syllables to be built in Block 1 of the derivation and restricting extended syllables to Block 2 (see sections 2.2, 4.1.4.5).

2.1.2. The domain of phonotactic constraints

We have pointed out above that the ultimate aim a phonotactic analysis is to characterise the phonologically well-formed word. However, this does not necessarily mean that (a) the basic domain within which the phonotactic constraints/rules of the theory are assumed to apply is the word, or (b) a theory must have just one basic domain (a single ‘structural base’) with reference to which all the significant phonotactic generalisations can be captured. Indeed, most frameworks derive phonological the well-formedness of the word from that of a smaller unit such as the morpheme or the syllable and many analyses recognise the fact that while the phonotactic grammaticality of a larger unit may be derivable from the phonotactic grammaticality of the smaller units from which it is composed, there may be phonotactic regularities specific to (and only statable with reference to) the larger unit too.

Structuralist phonology was not committed to a single most important structural base for phonotactic analysis. It was assumed that the identity of such a unit may even be language-specific: in some languages the syllable may be the most appropriate unit for the statement of phonotactic regularities, in others it may be the word or the morpheme (e.g. Trubetzkoy 1939/1969).

Early and classical generative phonology assumed that the morpheme was the only structural base: phonotactic statements were Morpheme Structure Rules/Constraints (MSRs/MSCs). In fact, (implicitly or explicitly) the morpheme was also considered to be the largest domain for phonotactic regularities: the well-formed word was assumed to be a concatenation of well-formed morphemes (and any restrictions on their concatenation were
considered to be morphological and not phonological in nature\textsuperscript{13}). Morpheme-based phonotactics followed from the assumption that phonotactic rules only applied to the underlying representation (see section 2.1.1) of morphemes listed in the lexicon (since the lexicon was considered to be a list of morphemes), and the assumption that phonological representations consisted of strings of segments (which were represented as one-column feature matrices) and the only groupings of these segments were morphological/syntactic in nature (whose boundaries were represented by boundary symbols like ‘#’, ‘+’ or ‘=’). Prosodic groupings of segments such as the syllable had no theoretical status, since syllable division is (almost always) predictable from the sequences of segments and thus was considered to have no place in underlying representations. Whatever regularities seemed to be expressible with reference to the syllable were considered to be expressible and were to be expressed with reference to the segment sequences themselves directly (since syllable division was predictable from the segment sequences anyway), cf. Chomsky and Halle (1968). The syllable as the domain of phonotactic constraints (and a legitimate prosodic unit of phonological representations to which phonological rules may be sensitive to) was reintroduced into generative phonology following arguments presented in Kahn (1976/1980)\textsuperscript{14} and became standard in Autosegmental Phonology and Metrical Phonology and generally in later phonological theories within the generative tradition (e.g. Steriade 1982, Clements and Keyser 1983, Goldsmith 1990, etc.).\textsuperscript{15} Kahn’s phonotactic arguments for the syllable concerned the

\textsuperscript{13}Independently of the domain issue, this assumption is clearly false. There are well-known cases of phonological conditions on affixation (e.g. in English deverbal noun-forming -\textit{al} and verb-forming -\textit{ize} can only attach to stems with non-final stress; deadjectival verb-forming -\textit{en} cannot attach to sonorant-final stems; etc, cf. Siegel 1974, Carstairs-McCarthy 1998, Raffelsiefen 1996; for a comprehensive review of interesting cases see the LINGUIST List < linguist@linguistlist.org> Vol-13-92) and of phonological conditioning of suppletive alternations (e.g. the -\textit{sz}- -(V)\textit{il} alternation in the 2nd sg. present indefinite (see Section 3.2.4.3)).

\textsuperscript{14}Kahn was not the only one (or the first one) to argue for the syllable as the appropriate phonotactic domain after SPE (cf. Fudge 1969, Brown 1969, for instance), but he was the most influential within the generative paradigm.

\textsuperscript{15}This includes (most versions of) Optimality Theory too. Classical Government Phonology (e.g. Kaye, Lowenstamm and Vergnaud 1990, Harris 1994) and its later offsprings, strict CV phonology (e.g. Lowenstamm 1996, Scheer 1999) and strict VC phonology (e.g. Szigetvári 1999, 2000b) seem counterexamples at first, but they are not since they all recognise a
relationship between word-initial/final clusters and word-medial clusters (in morphologically simple words). He pointed out that a possible medial consonant cluster can be analysed into a combination of a possible word-final consonant or consonant cluster plus a possible word-initial consonant or consonant cluster: hypothetical *atktin is not a possible word in English because /*tk#/ is not a possible final cluster and /*/#kt/ is not a possible initial cluster – by contrast, hypothetical atklin is a possible word because /#kl/ is a possible initial cluster (clue, clash, etc.). If the morpheme is the (only) domain of phonotactics, then this fact is nothing more than an accident (Kahn 1976: 57-58). However, if the basic domain of phonotactics is the syllable and if we assume that the set of word-initial clusters is coextensive with the set of syllable-initial clusters and the set of word-final clusters is coextensive with the set of syllable-final clusters, then the above observation follows from syllabification: *atktin is not a possible word because it is not syllabifiable (*a.tktin, *at.ktin, *atk.tin, *atkt.in). Atklin, on the other hand, is a possible word because it is syllabifiable since /kl/ is a possible syllable/word (at.klin).16

The syllable as a basic phonotactic domain has two additional advantages. The first concerns the ‘uniformity’ of the syllable. The syllable is obviously a more uniform structural base than the morpheme (or the word) in the sense that it displays a much more limited range of variation: morphemes (words) can differ from one another in a wider range of dimensions (length, number of vowels, etc) than syllables which have a uniform structure in the sense that

prosodic domain of phonotactic organisation (Onset-Rime doublets in classical Government Phonology, pairs of CV positions and pairs of VC positions in strict CV and strict VC Phonology, respectively), they just deny the existence of the syllable as a constituent. We will return to a real counterexample, Phonetically Grounded/Driven Phonology (e.g. Hayes 1996, Steriade 2000) later in this section.

16Note that the relationship between word-initial/final and word-medial clusters is more complex than Kahn (1967) suggested. It is a well-known fact that phonotactics is more ‘relaxed’ at the periphery of analytic domains (e.g. word-finally and word-initially) than medially (cf. e.g. Kenstowicz 1994, Harris 1994, Törkenczy and Siptár 1999ab and references cited therein). One consequence of this is that not all combinations of a well-formed syllable-initial cluster plus a well-formed syllable-initial cluster form well-formed medial clusters (see section 3.1). This, however, does not crucially change the force of the argument for the syllable vs. the morpheme as the basic domain of phonotactics since exclusively morpheme-based phonotactics would predict no systematic relationship between word-initial/final and word-medial clusters (which is an untenable claim).
they *always* consist of a single nucleus preceded and/or followed by an optional consonant (cluster). The uniformity of the syllable is a desirable property if it is seen as a building block from the well-formedness of which the phonotactic grammaticality of the word derives (Törkenczy 2000, 2001). Also, the uniformity of the syllable predicts that phonotactic statements can be stated more generally than those that can be formulated with reference to the morpheme.\textsuperscript{17}

The syllable as a prosodic unit is usually/often assumed to have an internal (Onset - (Nucleus - Coda)\textsubscript{Rhyme})\textsubscript{Syllable} structure.\textsuperscript{18} This internal organisation is advantageous since it expresses phonotactic tendencies, i.e. phonotactic generalisations/expectations follow from it, such as the lack (or the marked character) of phonotactic constraints between a consonant and a following vowel. Since the morpheme has no internal phonological organisation (if we do not recognise its organisation as a combination of syllables) and as a phonotactic domain it is just a morpheme size string of segments it suggests no comparable predictions.

We assume in the present dissertation that the status of the syllable as the basic phonotactic domain is firmly established and we will formulate the basic Hungarian phonotactic constraints as Syllable Structure Constraints (SSCs). However, there is evidence (Kaye 1974, Kenstowicz and Kisseberth 1977, Törkenczy 1994a, Booij 1995, 1999, Hammond 1997) that the phonotactic well-formedness of words also depends on constraints independent of prosodic structure (syllabic organisation). The relevant constraints are Morpheme Structure Conditions (MSCs) and sequence constraints. MSCs define possible morpheme shapes and may refer to categorial information (word classes). Thus, they can impose constraints on what is a possible morpheme, noun, verb, etc. in a given language. These constraints are different from classical generative MSRs/MSCs (e.g. Chomsky and Halle 1968): the recognition of the

\textsuperscript{17}Discovering the ‘universal’ (cross-language) properties of the syllable (e.g. Trubetzkoy 1939/1969, Kaye and Lowenstamm 1981, Blevins 1995) seems a more plausible task than characterising the general properties of the morpheme across languages. Note, however, that Trnka (1936) attempted to formulate ‘general laws of phonemic combination’ with reference to the morpheme – note also Trubetzkoï’s critique (Trubetzkoy 1939/1969: 244-247).

\textsuperscript{18}There are other views of intrasyllabic structure than this one (which we assume in the present dissertation). For some discussion of the organisation of the syllable and references see section 2.2.
There are recent arguments that even underlying morpheme structure constraints are necessary (Booij 1995, 1999, Hammond 1997).

A particular language may also have well-formedness conditions that constrain the combination of segments irrespective of their affiliation with prosodic or morphological units. These sequence constraints may state that a given (sequential) combination of segments (or features) XY is ill-formed regardless whether it is wholly contained within or cuts across structural units such as syllables or morphemes within the word or even across the word (for some examples of clear cases of sequence constraints in Hungarian see sections 3.3.2.2, 3.5 and 4.2.3. In this dissertation we argue that all the three kinds are necessary to account for the phonotactic pattern of Hungarian.²⁰

2.1.3. Phonotactic strata

It is a well-documented property of the lexicon of a natural language that it may have a stratified structure. Phonologically, stratification manifests itself in the fact that lexical items that belong to different lexical strata (sublexicons) may display (partially) different phonological regularities. Probably the best known examples are Japanese (Itô & Mester 1995) and English (Chomsky & Halle 1968). The stratum-specific phonological regularities may involve alternations (captured by phonological rules) or phonotactic patterns (captured by phonotactic constraints: SSCs, MSCs, or sequence constraints). These sublexicons exist

---

¹⁹There are recent arguments that even underlying morpheme structure constraints are necessary (Booij 1995, 1999, Hammond 1997).

²⁰Note that there are recent theories which argue that all phonotactic constraints are sequence constraints, i.e. local restrictions on segment combination (grounded in perception/perceptibility) and (prosodic) domains (such as the syllable) have no role in phonotactics (e.g Côté 2000, Steriade 2000, Rebrus and Trön 2002).
 synchronically —they are not simply etymologically identical/similar sets of words. Sublexicons/lexical strata may sometimes be identified independently of the phonological regularities, specifically, phonological rules and/or phonotactic constraints may be associated with a specific word-class. Stress-assignment in English, for instance, follows two patterns (the ‘verb pattern’ and the ‘noun’ pattern) rather than a single homogeneous one (cf. Chomsky and Halle 1968). The phonotactics of the Classical Arabic verb is different from that of the noun (cf. McCarthy 1981). A similar situation exists in Hungarian where the phonotactics of verbs is more restrictive than the phonotactics of non-verbs (cf. Trón & Rebrus 2000, 2001, Rebrus & Trón 2002, Törkenczy 2000, 2001; see section 3.4.3) and where word-class membership partially determines the distribution of stem-final floating [+ open,] (which causes Lowering, cf. sections 4.1.3 and 4.1.4.3).

It is a nontrivial problem how phonotactic grammaticality ratings are influenced by a more intricate structuring of the lexicon. Several scenarios seem possible (e.g. each sublexicon has its own phonotactic subgrammar, or there is a designated default sublexicon that determines the grammaticality ratings of unlisted strings, etc). We shall not pursue this issue here (it is beyond the scope of the present dissertation), but note that the phonotactic ‘space’ in a language may not be homogeneous, so phonotactic constraints must be permitted to be associated with a sublexicon.

2.1.4. Degrees of phonotactic well-formedness, irregularities, accidental gaps

It is usually taken for granted that native speaker judgements about the phonotactic well-formedness of strings of segments show a binary division of strings into well-formed and ill-

---

21The items in a given sublexicon may or may not be etymologically related, as is the case with the sublexicons in Japanese (Yamato, Sino-Yapanese, Foreign and Mimetic, cf. Itô & Mester 1995). ‘Foreign’ words of various origin may constitute a sublexicon in a language (as in Japanese), but that does not mean that all the words of foreign origin in that language necessarily belong to the foreign stratum/sublexicon since some words of foreign origin are not identifiable phonologically (e.g. tánc ‘dance’ in Hungarian).

22There may be other independently identifiable strata. Recently there has been claims that place-names form a phonotactic sublexicon in Hungarian, cf. Rebrus & Trón (2002).
formed. It has to be pointed out, however, that this is not necessarily true, and there exists some experimental evidence that native speakers can distinguish more than two degrees of phonotactic grammaticality (Ohala 1984, 1986, Scholes 1966, Greenberg & Jenkins 1964). Although the evidence is far from conclusive, there have been attempts to construct algorithms that assign strings to more than two levels of phonotactics well-formedness (Chomsky & Halle 1968, Scholes 1966, Greenberg & Jenkins 1964, Clements & Keyser 1983). While these algorithms often make wildly different predictions, they all agree in the central assumption (which they share with the usual approach that only distinguishes two levels of well-formedness) that while unlistedness in the lexicon (non-occurrence) does not necessarily entail phonotactic ungrammaticality, the more similar an item is to most of the listed items, the more well-formed it is (or in the case of two-level models: the more likely it is to be well-formed). There seems to be an agreement that accidental gaps exist and that the lexicon of a natural language contains (a small number of) items that are phonotactically irregular. The difference between a multi-level phonotactic approach (i.e. one that permits more than two degrees of well-formedness) and a two-level approach (i.e. one that permits only two degrees of well-formedness) is that in the former accidental gaps and irregular strings (like any occurrence or non-occurrence) may represent various degrees of phonotactic well-formedness while in the latter an accidental gap is perfectly well-formed (it has the same status as a systematic occurrence) and an irregular attested string is just as ill-formed as a systematic gap.

In this dissertation I shall take the conservative stance and assume that there are only two degrees of phonotactic grammaticality: well-formed and ill-formed, and that strings

\[\text{23}\] For instance, about the evaluation of the number of anomalies in a string: some suggest that the number of phonotactic violations in a string does not influence the well-formedness of the string (Chomsky & Halle 1968) others predict that the more violations a string has, the less well-formed it is (Greenberg & Jenkins 1964). For a detailed discussion and critique of the various algorithms see Törkenczy 1987, 2000ab.

\[\text{24}\] The only approach I know that denies the existence of accidental gaps (in fact, the relevance of the accidental-systematic distinction) is Rebrus & Trón (2002). Note that, although some string are claimed to be phonotactically better than others in Rebrus & Trón (2002), this does not imply different degrees of phonotactic well-formedness in the sense discussed above. What they mean is a difference in markedness which manifest itself in implication: the presence a ‘worse’ (i.e. more marked) string (universally) implies the presence of a better (i.e. less marked) one.
present in or absent from the lexicon may be well-formed or ill-formed\textsuperscript{25} (which is basically the classical structuralist position, cf. Fischer-Jørgensen 1952, Vogt 1954).

\section*{2.2. Framework and theoretical assumptions}

In this section I discuss the main theoretical assumptions underlying the description of Hungarian phonotactics presented in this book.\textsuperscript{26} These concern (i) the derivation and the relationship between morphological and phonological domains, (ii) the representation of segments, and (iii) the representation of syllable structure. Further discussion of some details appears in the analytical chapters where they are relevant to the issues at hand.

In this dissertation I assume that—as in other languages (e.g. English)—there are two kinds of morphological domains in Hungarian. We shall refer to the two kinds of domains as ‘synthetic’ and ‘analytic’.\textsuperscript{27} The distinction is crucial in (i) the relationship between morphological domains and syllable structure/phonotactics, and (ii) the derivation.

Analytic morphological domain boundaries are opaque to phonotactic constraints, in other words, phonotactic constraints do not apply across them (cf. Kaye, Lowenstamm and Vergnaud 1990). For instance, in Hungarian there are no phonotactic restrictions that constrain which consonants can be juxtaposed in a cluster $C_sC_p$ when $C_s$ is the last consonant of the first half of a compound word and $C_p$ is the first consonant of the second half of the compound. The restrictions one may find are purely accidental or non-phonological.\textsuperscript{28} Intervocalic $/kp/$,

\footnotesize
\textsuperscript{25}Naturally one expects the number of actually occurring irregular strings to be very low and the phonotactically permitted/delimited ‘space’ to be filled by strings represented in the lexicon—though it has to be admitted that this ‘expectation’ is not formally built in the framework adopted in the dissertation (or in any generative model).

\textsuperscript{26}These are essentially the same as those in Siptár & Törkenczy (2000).

\textsuperscript{27}The terms are borrowed from Government Phonology (cf. e.g. Kaye and Vergnaud 1990, Kaye 1995), but the distinction is traditional in different varieties of Generative Phonology. It is the same as that between ‘+$ ‘ boundary and ‘#$’ boundary affixation, or Level 1 and Level 2 affixation (cf. Harris 1994).

\textsuperscript{28}The few non-accidental regularities that can be found are due to postlexical assimilations such as Voice Assimilation and Nasal Place Assimilation, cf. Siptár & Törkenczy (2000).
Of course, words always form their own analytical domains: \( \text{Légy} \text{ “bátor”} \) “Be brave!”.

Some authors distinguish a third type, the quasi-analytic suffix which is “intermediate” between analytic and synthetic. See Rebrus & Törkenczy (1999), Rebrus (2000b), Törkenczy (1998b), Törkenczy & Siptár (1999a). In this dissertation I consider these suffixes idiosyncratic phonologically and suggest that the special alternations they are involved in are instances of suppletive allomorphy, see sections 3.2.4.3., 4.1.3., and 4.1.4.4.

This difference is important in Vowel Harmony because (most) analytic suffixes harmonize, but preverbs and compound members do not (cf. Siptár & Törkenczy 2000).

For instance, identical coda clusters are permitted monomorphemically and when the coda consists of a stem-final consonant and a consonant that belongs to a synthetic suffix. However, hiatus is possible monomorphemically (and across an analytic boundary) but not when one of the vowels is stem-final and the other is initial in a synthetic suffix. See section 4.1.4.2.
Lexical Phonology (e.g. Kiparsky 1982ab) assumed an interleaving of morphology and phonology and thus, both phonological processes and morphological operations were said to take place in the module they ‘belong to’. As the modules are ordered with respect to one another, both the phonological processes and the morphological operations in Block 1 have to precede those in Block 2. Because of the problem of violations of the affixal order predicted by level ordering and that of ‘bracketing paradoxes’ (cf. Aronoff 1976, Fabb 1988, Cole 1995), a different interpretation was proposed in Halle and Vergnaud (1987) and Halle and Kenstowicz (1991). In their view it is only the phonological processes that are assigned to the modules. All of morphology happens before phonology and each suffix is simply marked according to which block of rules it triggers. Thus, the order of morphological operations does not have to mirror the order of the modules. There is evidence of violations of level ordering in Hungarian. The suffix -hat/-het ‘may’ is a case in point. It can be attached without a linking vowel to any stem that ends in a single consonant: lop-hat ‘may steal’ (3sg indef.), döf-het ‘may thrust’ (3sg indef.), lát-hat ‘may see’ (3sg indef.), rak-hat ‘may put’ (3sg indef.), etc. The lack of phonotactic interaction between the stem-final and the suffix-initial consonant suggests that it is an analytic (Block 2) suffix. Yet it can be followed by a suffix such as the past tense suffix -(V)t(t), which is synthetic (Block 1) since the occurrence of its initial linking vowel depends on the last consonant of the stem (cf. section 4.1.4.4.): rohan-hat-oṭṭ ‘may run’ (3sg past indef.)—compare rohan-t ‘run’ (3sg past indef.). We adopt the view that morphology precedes phonology rather than being interleaved with it, and that the phonological rules belong to (ordered) lexical modules, but otherwise shall interpret derivation in a somewhat different way.33

We shall assume that the suffixes are marked according to whether they are analytic or synthetic. Analytic suffixes must be in a (dependent) domain which is different from that of the stem they are attached to. This domain may be monomorphemic or may contain synthetic suffixes as well. Block 1 rules will apply only within (dependent or independent)

33This interpretation owes very much to Government Phonology (cf. Harris 1994, Kaye 1995), but is very different from it in many respects (e.g. Government Phonology does not permit rule ordering, let alone blocks of phonological rules).
analytic domains (thus in a structure \([[[X]Y]]\), they may apply (independently) to X and Y). Block 1 rules show derived environment effects, but the derivation is not (necessarily) cyclic within the domain (cf. section 4.1.4.3). Following Cole (1995) we assume that derived environment effects (i.e. that a given rule does not apply within the morpheme, but does when the triggering environment is the result of affixation (of certain affixes)) are not (exclusively) the property of cyclic rules, so we shall refer to the Derived Environment Constraint instead of the Strict Cycle Condition. When all the Block 1 rules have applied, the whole word is subjected to the rules of Block 2. An extended syllable template (cf. Chapter 3 and section 4.1.4.5) is available when this happens and Block 2 rules are assumed not to be subject to the Derived Environment Constraint. A given rule may occur in both blocks or only one of them.

The feature geometry assumed is essentially that proposed in Clements and Hume (1995), compare Siptár & Törkenczy (2000) and Appendix A.

The view of syllable structure and syllabification taken here will be fairly traditional. I assume that syllable structure is not present underlyingly, but is built up by syllabification in the course of the derivation. Syllabification is seen as a template-matching algorithm (Itô 1986, 1989)—cf. section 4.1.4.1.

I assume that the segments belonging to a syllable are organized into the sub-syllabic constituents onset, nucleus, rhyme and coda. I also make the assumption that the constituents are hierarchically organized.

---

34 It is necessary to allow Block 1 rules to apply within a dependent analytic domain because it may contain a synthetic suffix.

35 I take no stand as to whether rules can be ‘turned on’ or only ‘turned off’ (cf. Mohanan 1986, Halle and Mohanan 1985, Borowsky 1986).


Some authors deny the validity of the principle as a universal (Clements and Keyser 1983, Davis 1985) and there are known counterexamples. However, it appears that the unmarked case is when the principle holds (e.g. Fudge 1987). Note that this does not mean that there may be no phonotactic constraints holding between a vowel and a preceding consonant; it only means that if such a constraint obtains, it is not a syllable structure constraint—it can easily be a constraint on morpheme shape, for instance (cf. Davis 1991, Booij 1995, 1999).

![Diagram](image)

Under this view, syllable well-formedness derives from the well-formedness of the subsyllabic constituents. Given the hierarchical structure in (2), no restrictions (or at least only weaker ones) are expected to apply between the constituents onset and rhyme than between the nucleus and the coda or within each (sub)constituent. This is sometimes referred to as the Principle of Free Cooccurrence (Kaye 1995) and appears to hold true of Hungarian. Furthermore, (in Hungarian and universally) constraints on syllable well-formedness seem to apply to subsyllabic constituents and not to the constituent ‘syllable’ itself. This has led some researchers (Aoun 1979, Kaye, Lowenstamm and Vergnaud 1990) to deny the existence of the syllable as a constituent altogether. As nothing seems to hinge on this matter, we take no theoretical stand and retain the syllable as a convenient way of referring to the combination of an onset and a rhyme.

I assume that all segments that are phonetically interpreted must be prosodically licensed (Itô 1986, 1989). The question is whether this assumption necessarily means that each segment that appears at the surface is affiliated to one of the subsyllabic constituents. The answer is very important in the analysis of the so-called edge effects, i.e. the special character of (certain) clusters at the edges of (certain) morphological domains. There are strict and permissive approaches to this problem. Under the strict view, edge effects must be accounted for by normal syllable structure (i.e. the answer to the question above is yes). Thus, no special syllable structures are postulated that are limited to domain edges. Government Phonology

---

38 Some authors deny the validity of the principle as a universal (Clements and Keyser 1983, Davis 1985) and there are known counterexamples. However, it appears that the unmarked case is when the principle holds (e.g. Fudge 1987). Note that this does not mean that there may be no phonotactic constraints holding between a vowel and a preceding consonant; it only means that if such a constraint obtains, it is not a syllable structure constraint—it can easily be a constraint on morpheme shape, for instance (cf. Davis 1991, Booij 1995, 1999).
exemplifies this approach. In the permissive approach edge effects are accounted for by special syllable structures that can only appear at domain edges. There are several variations: in some analyses the special syllable structures in question may contain an additional subsyllabic constituent such as the appendix (e.g. Fudge 1969, Fujimura 1979, Hulst 1984), other approaches permit direct licensing (i.e. unmediated by a subsyllabic constituent) by the syllable node in the special syllables (e.g. Steriade 1982, Clements and Keyser 1983), still others allow direct licensing of segments by prosodic nodes higher than the syllable at domain edges (Rubach and Booij 1990, Törkenczy 1994a). It is difficult (and not always possible) to find empirical differences between the various approaches.

In this dissertation I adopt the permissive approach and allow an extended syllable, i.e. one containing an appendix, in Block 2 (cf. Chapter 3 and section 4.1.4.5). Only the core syllable template shown in (2) is available for syllabification in Block 1.

Phonotactic constraints are often explainable with reference to sonority and the Sonority Hierarchy (e.g. Clements 1988, Vennemann 1988, Rice 1992). Despite the difficulties with the phonetic definition of the Sonority Hierarchy (Clements 1990, Laver 1994), I take it to be a well-established phonological relationship between classes of segments. We also assume that the Sonority Hierarchy is universal and is the following:

\[ (2) \quad \text{Sonority Hierarchy} \]

\[
\text{stops, affricates} \lessdot \text{fricatives} \lessdot \text{nasals} \lessdot \text{liquids} \lessdot \text{glides} \lessdot \text{vowels}
\]

Although the Sonority Hierarchy is universal, there has to be room for some language particular variation: sonority ‘reversals’ are not permitted (e.g. a language may not classify obstruents as less sonorous than nasals), but different language particular settings of sonority distance between segment classes are possible (e.g. a language may determine that the sonority distance between stops and fricatives is smaller than that between fricatives and nasals; cf. Steriade 1982, Hulst 1984). We assume that phonotactic constraints can refer directly to the

\[ ^{39}\text{In GP, instead of special structures, special segmental material (empty vowels) may appear at the edges of domains. See, Kaye, Lowenstamm and Vergnaud (1990), Kaye (1990). See also Burzio (1994) on the relationship between allowing special structures vs. special segments.} \]

\[ ^{40}\text{For arguments against language particular Sonority Hierarchies, cf. Clements (1990).} \]
Sonority Hierarchy. In order to account for sonority-based asymmetries of segment combination we shall borrow the term ‘government’ from Government Phonology (e.g. Kaye, Lowenstamm and Vergnaud 1990), and Rice (1992) and state:

(3)  

**Government**

A segment X governs an adjacent segment Y if X is less sonorous than Y.  

I take government to be asymmetrical, but not intrinsically strictly directional in all governing domains, i.e. it is always directional, but its direction may be fixed in some structural positions but free in others. I assume that government is universally left to right in onsets and right to left in codas. In Hungarian transsyllabic clusters, however, the directionality of government is not fixed (it may be left to right or right to left), cf. section 3.3. I follow Kaye, Lowenstamm and Vergnaud (1990) and assume that government applies between timing slots.

Following Rice (1992) I assume that there may be another asymmetrical relationship between adjacent segments, i.e. the relationship of ‘binding’. I follow (and generalize) Rice’s definition (compare Rice 1992):

(4)  

**Binding**

A bound segment contains dependent structure.

Thus, a bound segment contains structure that does not differ from that of the segment that

---

41Note that this does not mean that the Sonority Hierarchy is a primitive (a scalar feature, for instance). I assume that the Sonority Hierarchy is derived. I take no stand whether it is to be defined in terms of features (cf. Clements 1990) or structurally (cf. Kaye, Lowenstamm and Vergnaud 1990, Harris 1990, Rice 1992).

42There are important differences between GP’s and Rice’s interpretation of government. Our interpretation here is closer to Rice (1992).

43I do not take a stand as to the interpretation/derivation of sonority. For the sake of simplicity (3) can be interpreted as directly referring to (2)).

44Compare Kaye, Lowenstamm and Vergnaud (1990) who assume that government is strictly directional in all governing domains.
binds it (e.g. in a homorganic nasal+ stop cluster the nasal is bound by the stop). Binding can apply to various nodes of the feature tree, e.g. to the root node (‘root-binding’) or the place node (‘place binding’) for instance. (in the example above the nasal is ‘place-bound’; in a (true) geminate the first consonant is ‘root-bound’, i.e. it has the same structure from the root down as the second consonant). I assume that binding is strictly directional and is right to left.

The mora is not a primitive in the present treatment, but is considered to be derivative of syllable structure. It is only used as a unit of measuring syllable weight (which, incidentally, plays very little role in Hungarian phonology\(^{45}\), cf. section 3.4.1) and does not function as a subsyllabic constituent/timing unit. I shall use ‘bimoraic’ as a convenient label to refer to syllables that have a branching rhyme or/and a branching nucleus. For arguments against moraic syllable structure (as proposed in Hyman 1985, Hayes 1989) cf. Brentari and Bosch (1990), Davis (1990), Sloan (1991), Tranel (1991), Rialland (1993).

\(^{45}\)For a different view, cf. Vago (1989a, b, 1992)
Chapter 3.

‘Static’ phonotactics: the phonological shape of the Hungarian word

3.1 Introduction

The phonological or phonotactic well-formedness of a word can be seen as an interplay of two factors: a prosodic and a non-prosodic one. On the one hand, a phonologically well-formed word must be parsable into (well-formed) prosodic units. In this dissertation we accept the general assumption that the prosodic unit that is chiefly responsible for phonotactic well-formedness is the syllable.\(^1\) In addition, there are well-known examples of phonotactic constraints whose domain is a higher level prosodic unit such as the foot or the prosodic word.\(^2\) As the foot does not seem to play an important role in Hungarian, a phonotactically well-formed Hungarian word is a unit which is exhaustively parsable into well-formed syllables. Thus, the phonotactic well-formedness of a word is derivable from well-formedness conditions on syllables (Syllable Structure Constraints (SSCs)). This relation between the well-formedness of words and syllables, however, is not symmetrical: while it holds that a well-formed word consists of a string of well-formed syllables, it is not true that any string of well-formed syllables constitutes a well-formed (potential\(^3\)) word: there are transsyllabic constraints that obtain between syllables, or more precisely, between adjacent subconstituents of different syllables. These constraints do not refer to a prosodic unit higher than the syllable, but impose restrictions on the bonding of syllable edges (interconstituent clusters). In addition, as we have

\(^1\)Naturally, this is only true of frameworks that refer to the syllable (compare SPE, Government Phonology, Phonetically Grounded/Driven Phonology).

\(^2\)For instance, the distribution of /h/ in English cf. Anderson and Ewen (1987), Harris (1994).

\(^3\)Naturally, well-formedness conditions of any kind must define units that are potentially well-formed in the given language, in other words, they must not treat accidental gaps as ill-formed (cf. e.g. Chomsky 1964, Halle 1962, Vogt 1954).
pointed out above, a language may have constraints on prosodic structure that directly refer to prosodic units higher than the syllable (e.g. conditions on word minimality, etc.).

The phonotactic well-formedness of words may also depend on *Morpheme Structure Constraints (MSCs)* and *Sequence Constraints*, i.e. constraints independent of prosodic structure. We shall see that MSCs and sequence constraints partially determine the phonotactic pattern of Hungarian.\(^4\)

### 3.2. Syllable structure: SSCs

In this section we discuss the constraints that apply within (the constituents of) the syllable and define the syllable template in Hungarian.

#### 3.2.1. The Hungarian syllable template: the basic syllable types

If we disregard the possible complexity of the onset, the nucleus and the coda, Hungarian has the following syllable types:\(^5\)

---

\(^4\)See section 2.1.

\(^5\)There is a surprising degree of agreement about this: authors of very different theoretical backgrounds agree that (disregarding constituent complexity) these are the basic syllable types in Hungarian, cf. Rácz (1961), Kaye and Lowenstamm (1981), Kornai (1990a, 1994), though see Kassai (1981) who also permits syllables consisting of consonants only. Naturally, authors whose framework excludes one (or more) of these structures come to different conclusions (e.g. Kaye, Lowenstamm and Vergnaud 1990).
Blevins (1995) has a sixth parameter for Edge Effect that we discuss later.

There are alternative ways of expressing (more-or-less) the same typological distinctions as these parameters cf. e.g. Kaye and Lowenstamm (1981), Clements and Keyser (1983), Prince and Smolensky (1993).

Blevins (1995) proposes the following binary (YES/NO) parameters to account for language particular variation in syllable typology: Obligatory Onset, Coda, Complex Nucleus, Complex Onset, Complex Coda⁶ (note that there is no Onset parameter, i.e. languages cannot choose to have no onsets). (2) shows the parameter settings for Hungarian⁷ (disregarding the last three parameters which refer to constituent complexity):

(2) Obligatory Onset      NO
    Coda               YES

As complex nuclei occur in Hungarian (i.e. the Complex Nucleus parameter is set ‘YES’), the syllable inventory in (1) can be extended:

A comparison of (1) and (3) reveals that the distribution of long and short vowelled syllables

⁶Blevins (1995) has a sixth parameter for Edge Effect that we discuss later.

⁷There are alternative ways of expressing (more-or-less) the same typological distinctions as these parameters cf. e.g. Kaye and Lowenstamm (1981), Clements and Keyser (1983), Prince and Smolensky (1993).
Some authors do consider it as evidence (e.g. Kahn 1976/1980), but currently there seems to be an agreement among phonologists that the assumption that word-initial/final consonant clusters are necessarily complex onsets/codas is false, see, e.g. Kenstowicz (1994), Kaye (1992), Kaye and Lowenstamm (1981), Kaye, Lowenstamm and Vergnaud (1990), Harris (1994), Steriade (1982), Rubach and Booij (1990), Davis (1990), Törkenczy and Siptár (1999ab).

The three parameter settings discussed so far are fairly uncontroversial. What is more problematic is the setting of the remaining two ‘complexity’ parameters Complex Onset and Complex Coda. Hungarian words can begin and/or end with consonant clusters (e.g. prém ‘pelt’, ptózis ‘ptosis’, part ‘shore’, akt ‘nude’, etc.) but this is not necessarily evidence that these clusters are true onsets or codas. It is a well-known fact that word edges (or certain morphological domain-edges) license special syllable structures. Specifically, there may be consonants or consonant sequences at the edges of these domains that are not incorporated into the onset or the coda of the syllable whose phonetically realized nucleus is the first or the last one in the word, respectively. So the question is whether the consonant clusters that occur word-initially and word-finally in Hungarian are true complex onsets and codas (respectively) or they are ‘edge clusters’, i.e. clusters occurring at domain edges whose initial or final member(s) are licensed by some special mechanism limited to the edges of domains and not by an onset or a coda constituent dominating them. At this point we are not primarily concerned with the actual licensing mechanism, which will be discussed later, but the analysis of onsets and word-initial clusters (section 3.2.2) and codas and word-final clusters (section 3.2.4).

3.2.2. Onsets—word-initial clusters

As we have seen already, it is not compulsory for a Hungarian syllable to have a (filled) onset. Thus, both vowel-initial and consonant-initial syllables are possible. In principle, any

---

8Some authors do consider it as evidence (e.g. Kahn 1976/1980), but currently there seems to be an agreement among phonologists that the assumption that word-initial/final consonant clusters are necessarily complex onsets/codas is false, see, e.g. Kenstowicz (1994), Kaye (1992), Kaye and Lowenstamm (1981), Kaye, Lowenstamm and Vergnaud (1990), Harris (1994), Steriade (1982), Rubach and Booij (1990), Davis (1990), Törkenczy and Siptár (1999ab).
consonant may be syllabified into a simplex onset. Word-initial two-member and three-member consonant clusters occur—they are shown in Tables I and II. The question is whether these clusters realize branching onsets or not.

9Word-initially, palatal /t/ only occurs in a single morpheme tyúk ‘hen’, but we consider this accidental.
### Table I. Word-initial CC clusters

| p  | t  | t̊ | k  | b  | d  | d̊ | g  | ċ  | j  | f  | s  | š  | v  | z  | ž  | m  | n  | n̊ | l  | r  | j  | x  |
|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| p  | +  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| t  |    | +  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| t̊ |    |    | +  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| k  |    |    |    | +  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| b  |    |    |    |    | +  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| d  |    |    |    |    |    | +  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| d̊ |    |    |    |    |    |    | +  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| g  |    |    |    |    |    |    |    | +  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| ċ  |    |    |    |    |    |    |    |    |    | +  |    |    |    |    |    |    |    |    |    |    |    |    |    |
| j  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| f  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| s  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| š  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| v  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| z  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| ž  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| m  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    | +  |
| n  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| n̊ |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| l  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| r  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| j  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| x  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    | +  |

Accordingly, (some of) these clusters have been analysed as branching onsets. For a detailed analysis and specific constraints cf. Törkenczy (1994a).

TABLE II. Word-initial CCC clusters

<table>
<thead>
<tr>
<th></th>
<th>pr</th>
<th>tr</th>
<th>kr</th>
<th>kl</th>
</tr>
</thead>
<tbody>
<tr>
<td>s</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
</tbody>
</table>


All the words which begin with consonant clusters are loan words, but this fact does not in itself say anything about the status of the initial clusters: it is perfectly possible that the words in question are phonotactically just as ‘normal’ as any ‘native’ item in the lexicon. Indeed, Hungarian speakers can detect no difference between the well-formedness of a word such as próm ‘fur’ and róm ‘monster’. While there are few words beginning with three consonants, words beginning with two consonants cannot be said to be infrequent (though there are many more consonant-initial words that begin with a single consonant). Also, there appear to be phonotactic restrictions holding between the consonants making up the word-initial clusters. While some of these restrictions are unrelated to syllable structure (e.g. the ban on adjacent obstruents differing in voicing), others seem specific to this position and may be interpreted as holding between the members of a branching onset (e.g. the non-occurrence of geminates).  

10 Accordingly, (some of) these clusters have been analysed as branching onsets. For a detailed analysis and specific constraints cf. Törkenczy (1994a).
These constraints may be construed as evidence for the well-formedness of branching onsets. Nevertheless, we want to suggest that the setting for the Complex Onset parameter is in fact ‘NO’ in Hungarian, and all the clusters that occur word-initially are ‘edge clusters’. We assume that the non-final consonants in these clusters are licensed by a special mechanism restricted to domain edges, notably, they are syllabified into a subsyllabic constituent called ‘appendix’. Thus, they are represented as (4a) rather than (4b):

(4)  

Evidence for or against this position may be drawn from alternations/processes that are sensitive to syllable structure and phonotactic patterns. Syllable structure conditioned alternations (to be discussed in detail in Chapter 4) do not present conclusive evidence since there is no alternation in Hungarian that would require that the clusters in question should not be represented as branching onsets. However, it must be pointed out that the relevant alternations/processes (such as vowel - zero alternations (cf. Chapter 4) and Fast Cluster Simplification (cf. section 4.2.3)) never make a branching onset interpretation necessary - i.e. an analysis of these syllable structure sensitive processes is always perfectly compatible with an edge cluster (appendix) interpretation of these consonant sequences.

Phonotactic patterns are another possible source of evidence: if we could show that the ‘need’ to analyse consonant clusters or substrings of consonant clusters as putatively branching onsets only arises at domain edges, then we could see this as an argument against their branching onset status, as domain edges may license ‘special’ edge clusters. Given that the word-initial position is suspect (since the clusters occurring there may be edge-licensed as appendix+ onset), the most promising place to look for such evidence is medial. In
Long vowels other than é/á are not permitted before consonant clusters, but this constraint has nothing to do with syllable structure: all consonant clusters (whether they are potentially well-formed as branching onsets or not) behave in the same way: */i:kta/ and */i:kla/ are equally ill-formed. For details, see section 3.4.2.

At first sight, Hungarian seems to abound in word-medial -CCC- clusters. However, the main source of such clusters is analytic suffixation (e.g. [[kard]böl] ‘from (the) sword’, [[vers]röl] ‘about (the) poem’, [[elv]telen] ‘without principles’, etc.) and compounding (e.g. [[vers]láb] ‘(metrical) foot’, [[elv]társ] ‘comrade’, etc.). We have pointed out above (see Chapter 2) that clusters straddling the edge of an analytic domain do not say anything about the phonotactics of the language, they are ‘accidental’ in the sense that no phonotactic restrictions apply across analytic domain edges—the relevant consonants are just juxtaposed without any restrictions. Thus, ‘real data’ are monomorphemic items, or words with synthetic suffixation containing medial -CCC-. Interestingly, there are no examples in Hungarian of synthetic suffixation creating -CCC- clusters. There are monomorphemic words with -CCC- clusters in the language, but, significantly, their number is rather low, about 300 items in our database. Again, all the relevant words are loans, but, naturally, this does not in

---

11 Long vowels other than é/á are not permitted before consonant clusters, but this constraint has nothing to do with syllable structure: all consonant clusters (whether they are potentially well-formed as branching onsets or not) behave in the same way: */i:kta/ and */i:kla/ are equally ill-formed. For details, see section 3.4.2.

12 For a discussion of complex codas, see sections 3.2.4., 4.1.4.4 and 4.1.4.5.

13 Multiply suffixed past forms of cluster-final stems are the only exception (e.g. [fiŋktok] ‘fart’ (3pl past indef.)). Even these clusters are often broken up ([fiŋgotok]), cf. section 4.1.4. See also the behaviour of cluster-initial (-C,C,V...) suffixes in section 4.1.2.2.
itself say anything about their well-formedness in Hungarian (examples: *bisztró* ‘bistro’, *centrum* ‘centre’, *komplex* ‘complex’, *export* ‘id.’, *improvizál* ‘improvise (3 sg pres. indef.)’, *instancia* ‘instance’, *ostrom* ‘siege’, etc.). Furthermore, there are 95 types of clusters altogether that the approximately 300 tokens exemplify, but, typically, the number of tokens in a given type is extremely low (cf. section 3.3.2.2 for a full list of the relevant clusters). There are only 7 types with 10 or more tokens and the majority of types (n= 48) only have one token. This suggests that medial -CCC- clusters are special/irregular in Hungarian.

True, (monomorphemic) medial -CCC- clusters do display certain regularities (e.g. in a medial -CCC- cluster $C_aC_pC_r$, $C_p$ is never a sonorant$^{14}$), we claim that these regularities are accidental in Hungarian in that they only reflect some of the regularities of the source languages the relevant words were borrowed from.$^{15}$ More precisely, if a constraint obtaining between medial $C_aC_pC_r$ is non-accidental, then we have to do with either of the following two situations: (i) it is identical with a constraint obtaining between the consonants of a corresponding two-member medial cluster $C_aC_b$ and thus reduces to a constraint applying between a syllable-final consonant and the following syllable-initial one, i.e. it is an interconstituent constraint (e.g. there are no words with medial -tpC- in Hungarian, but there are no words containing medial -tp- either); (ii) it is an MSC or a sequence constraint and thus it has nothing to do with syllable structure at all (e.g. adjacent obstruents have to agree in voicing in Hungarian). Otherwise, all apparent medial -CCC-specific constraints are accidental, just ‘debris’ of the constraints that exist in the languages the particular words containing them were borrowed from. Of course, as (more-or-less random) ‘samples’ from the phonotactics of the source languages, which conform to universal phonotactic constraints themselves, these clusters may nevertheless reflect universal phonotactic regularities/tendencies.

Another argument for the special character of medial -CCC- clusters involves a comparison of medial -CC- clusters and -CCC- clusters. In a language that permits branching onsets we expect to find $-C_aC_pC_r$- clusters where $C_pC_r$ is a well-formed branching onset and

$^{14}$For a discussion of medial -CCC- clusters, see section 3.3.2.2.

$^{15}$Of course, this does not mean that an explanation of why these clusters are not repaired is not in order, see a possible explanation in section 4.1.4.5.
-C<sub>a</sub>C<sub>γ</sub>- is a permitted interconstituent cluster (-C<sub>a</sub>,C<sub>γ</sub>,-). And vice versa, in general, for every -C<sub>a</sub>C<sub>γ</sub>- cluster we should find a matching -C<sub>a</sub>C<sub>γ</sub>- cluster if the latter is a permitted interconstituent cluster. Of course, accidental gaps may exist, but this should be the general tendency. It is interesting to compare English and Hungarian since, in the literature English is generally taken to be a language that has branching onsets. As can be seen in (5), English is well-behaved with respect to the generalization above.

(5) 

<table>
<thead>
<tr>
<th></th>
<th>English</th>
<th>VC&lt;sub&gt;a&lt;/sub&gt;C&lt;sub&gt;γ&lt;/sub&gt;V</th>
<th>VC&lt;sub&gt;a&lt;/sub&gt;C&lt;sub&gt;γ&lt;/sub&gt;C&lt;sub&gt;γ&lt;/sub&gt;V</th>
</tr>
</thead>
<tbody>
<tr>
<td>-kt-</td>
<td>vector</td>
<td>electronic</td>
<td></td>
</tr>
<tr>
<td>-pt-</td>
<td>chapter</td>
<td>dioptry</td>
<td></td>
</tr>
<tr>
<td>*-tk-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>*-pk-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>*-tp-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>*-kp-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

Hungarian, on the other hand, is very different: some -C<sub>a</sub>C<sub>γ</sub>C<sub>γ</sub>- clusters corresponding to well-formed -C<sub>a</sub>C<sub>γ</sub>- are curiously missing:

(6) 

<table>
<thead>
<tr>
<th></th>
<th>Hungarian</th>
<th>VC&lt;sub&gt;a&lt;/sub&gt;C&lt;sub&gt;γ&lt;/sub&gt;V</th>
<th>VC&lt;sub&gt;a&lt;/sub&gt;C&lt;sub&gt;γ&lt;/sub&gt;C&lt;sub&gt;γ&lt;/sub&gt;V</th>
</tr>
</thead>
<tbody>
<tr>
<td>-kt-</td>
<td>akta ‘file’</td>
<td>spektrum ‘spectrum’</td>
<td></td>
</tr>
<tr>
<td>-pt-</td>
<td>kapta ‘(boot) last’</td>
<td>dioptria ‘dioptry’</td>
<td></td>
</tr>
<tr>
<td>-tk-</td>
<td>atka ‘mite’</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>-pk-</td>
<td>lepke ‘butterfly’</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>*-tp-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>*-kp-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

We can either say that the missing clusters are accidental gaps, or the other explanation is that complex onsets are ill-formed. We suggest that the latter interpretation is correct.

It follows from the irregular status of medial -C<sub>a</sub>C<sub>γ</sub>C<sub>γ</sub>- clusters that it is never necessary to syllabify two (or more) consonants into an onset in medial position. Therefore we claim that the consonant clusters that occur in word-initial position (the only position where
consonant clusters arguably look like complex onsets) do not form onsets, but are edge clusters. Thus, the setting of the Complex Onset parameter is ‘NO’ in Hungarian. In our interpretation the phonotactic restrictions word-initial clusters display are just (fragmentary) reflections of the constraints that apply in the source languages the relevant words come from, or those of the universal phonotactic regularities that the phonotactics of the source languages (and the phonotactics of any language) conforms to.

The fact that branching onsets are not permitted does not in itself explain the scarcity/irregularity of -CCC- clusters. The reason is that a -CCC- cluster can in principle be parsed exhaustively even if it does not contain a branching onset: it could consist of a complex coda and a following non-branching onset: -CC.C-. This raises the question whether complex codas are well-formed in Hungarian. If the answer is negative, it follows that medial -CCC- clusters are ill-formed (assuming that complex onsets are also ill-formed word-initially and word-medially). There are words ending in more than one consonant, but this does not in itself ascertain the status of these final clusters as complex syllabic constituents. We will return to this problem in section 3.2.4.

---

16Such as the absence of initial /sr/ (compare attested /šr/ sróf ‘screw’), cf. Törkeneczy (1994a) and Törkeneczy and Siptár (1999).

17Chiefly English and German (cf. Siptár 1980).

18GP has a ‘built-in’ negative answer to this question since the theory does not permit complex codas (it does not even have a coda constituent, cf. Kaye, Lowenstamm and Vergnaud (1990), Harris (1994)).
3.2.3. Rhymes

The rhyme may be branching or non-branching in Hungarian. Thus, the following rhyme templates are well-formed:

(7)  

\[
\begin{array}{ll}
\text{a.} & \text{R} \\
\text{b.} & \text{R} \\
\text{N} & \text{N} \\
\text{X} & \text{X X} \\
\text{c.} & \text{R} \\
\text{d.} & \text{R} \\
\text{N Co} & \text{N Co} \\
\text{X X} & \text{X X X} \\
\text{e.} & \text{R} \\
\text{f.} & \text{R} \\
\text{N Co} & \text{N Co} \\
\text{X X X} & \text{X X X X}
\end{array}
\]

In general, there is no restriction on nuclei in branching or non-branching rhymes in Hungarian: any vowel can occur in a closed or an open syllable:
It is apparent in (8) above that (i) any vowel quality is possible and (ii) long and short vowels equally occur both in branching and in non-branching rhymes, i.e. no rhyme-specific phonotactic statement is necessary.

We shall see in later sections that this is an oversimplification because (a) the distribution of vowels in stem/word-final open syllables is different from that in medial open syllables (see section 3.4.1) and (b) only a very limited set of long vowels can occur in word-medial closed syllables and word-final syllables closed by more than one consonant when these syllables are undivided by a morpheme boundary (see section 3.4.2). The constraints that (a) and (b) are due to are not SSCs strictly speaking because they refer to the phonological word or apply within the morpheme.

There is one phonotactic restriction, however, which seems specific to the rhyme. This constraint concerns the distribution of surface roundedness/labiality within the rhyme. Vowels preceding the nasal+ stop clusters /mp, mb/ must be rounded if the vowel and the entire consonant cluster are within the rhyme. Accordingly, while there are many words like
The word galamb ‘dove’ is problematic/exceptional if the constraint is taken to apply to the UR because the vowel of the final syllable is only rounded at the surface /galamb/ - [gɔlɔmb], see Törkenczy (1994).

lump ‘drunkard’, komp ‘ferry’, tömb ‘block’, domb ‘hill’, words like hypothetical *limp or *semb whose vowels are not rounded are unattested. By contrast, there are many words like ember ‘human being’, bimbó ‘bud’, lámpa ‘lamp’, nánber ‘hag’ in which the second member of the nasal+ stop clusters is not within the rhyme (em.ber, bim.bó, lám.pa, nán.ber) and thus the vowel is not required to be rounded. This constraint can be seen as evidence for the rhyme node in Hungarian. Note that it is ‘directional’. It is a constraint on vowels preceding /mp, mb/ and cannot be seen as a requirement that labiality/rounding has to be shared within a V+ nasal+ stop rhyme since any vowel quality is possible in rhymes containing non-labial nasal+ stop clusters: e.g. /nt/ bán ‘hurt’, csont ‘bone’, ment ‘save’, dönt ‘decide’, hint ‘sprinkle’; /ng/ ing ‘shirt’, ráng ‘jerk’, zeng ‘resound’, döng ‘buzz’, korong ‘disk’. It has to be pointed out that the status of this constraint is unclear. It is (almost) exceptionless, but it does not play an active role in the phonology. There are no alternations that it would condition, and no evidence is available concerning native speakers’ intuitions about the well-formedness of strings violating it.
3.2.4. Codas–word-final consonant clusters

In Hungarian the coda differs from the onset in that the former may branch. Thus, the setting of the Complex Coda parameter (Blevins 1995) is ‘YES’. The coda is maximally binary branching. Furthermore, complex codas may be morphologically complex (i.e. there are suffixes solely consisting of consonants syllabified into the coda).

3.2.4.1. Non-branching codas

Any underlying consonant may systematically occur in a non-branching coda.\(^{20}\)

3.2.4.2. Branching codas

In Hungarian the surface form of words may end in at most three consonants (\textit{hat} ‘six’, \textit{part} ‘coast’, \textit{szfinx} [sfiəks] ‘sphinx’). Nevertheless, we claim that the coda is maximally binary branching, and that the more complex clusters at the ends of words are not (exhaustively) syllabified into a single coda. Furthermore, not all word-final two-term clusters realize branching codas. Let us examine word-final two-term clusters first. The notation used in Table III is as follows: a blank space in an intersection of a row and a column means that the relevant cluster is unattested. Numbers occur at intersections when a given cluster is attested: 1 = a cluster that \textit{only} occurs undivided by a morpheme boundary (analytic or synthetic); 2 = a cluster that \textit{only} occurs when divided by a morpheme boundary (analytic or synthetic), 3 = a cluster that occurs \textit{both} monomorphemically and when divided by a morpheme boundary (analytic or synthetic). A box is struck out by dashes to indicate that the relevant cluster(s)

\(^{20}\)The surface realization of underlying consonants may be determined by syllabic constituency, see the behaviour of /x/ (section 4.2.2). Coda /j/ hardly ever occurs: \textit{bridzs} /briːz/, ?/briː/ ‘bridge (card game)’. 
is/are subject to (eliminated on the surface by) assimilations.

Table III. Word-final CC clusters

|   | p  | t  | t’ | k  | b  | d  | d’ | g  | t’ | č  | j  | f  | s  | š  | v  | ž  | m  | n  | n’ | l  | r  | j  | x  |
|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| p | 1  | 1  | 31 | -  | 3  | 1  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  |
| t | 3  | 1  | -  | -  | 31 | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  |
| t’| 1  | 1  | -  | -  | 31 | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  |
| k | 1  | 1  | -  | 1  | 31 | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  |
| b | -  | 1  | 2  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  |
| d | -  | 3  | -  | -  | -  | -  | 1  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  |
| d’| -  | 2  | 3  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  |
| g | -  | 3  | 1  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  |
| t’| 1  | 3  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  |
| č | -  | 3  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  |
| j | -  | 2  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  |
| f | 1  | -  | -  | 1  | 2  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  |
| s | 3  | 1  | -  | -  | 3  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  |
| š | 3  | -  | -  | -  | 3  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  |
| v | -  | 2  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  |
| z | -  | 2  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  |
| ž | -  | 3  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  |
| m | 1  | 1  | 1  | 2  | 2  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| n | -  | 3  | 1  | 1  | 1  | 3  | 3  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | -  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |
| n’| 2  | 3  | 2  | 1  | 1  | 3  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  |
| l | 1  | 3  | 1  | 1  | 1  | 1  | 3  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 3  |
| r | 1  | 3  | 1  | 1  | 1  | 3  | 3  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 3  |
| j | 1  | 3  | 1  | 1  | 1  | 3  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 3  |
| x | 1  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  | -  |

Examples: (two examples are cited if a given cluster occurs divided and undivided by a morpheme boundary): csepp ‘drop’; recept ‘receipt’; copf ‘plait’; biceps ‘biceps’, kapsz ‘get’ (2sg pres. indef.); taps ‘applause’, lopf ‘steal’ (sg imp. indef.); ott ‘there’, olvadt ‘molten’; Detk ‘place name’, pötty ‘dot’; Batyk ‘place name’, vágysz ‘desire’ (2sg pres. indef.); akt

21 None of these assimilations are related to syllable structure (i.e. they operate regardless of the syllabification of the cluster to which they apply) and some of them are postlexical (see Chapter 7). Note that Table III shows the inventory of word-final clusters after these assimilations have applied, i.e. clusters that are subject to assimilations appear in it in the assimilated form.
There is a weaker version of the principle that tolerates sonority plateaus (cf. e.g. Törkenczy 1994, Blevins 1995). Here we adopt the stronger version.
section. So let us state the following (partly universal partly language specific) constraint for branching codas in Hungarian:

(9) Hungarian Branching Coda Constraint

Branching codas must be licensed either by government or root-binding.

Given the assumption that the direction of government is universally right to left in a coda, (9) upholds the SSP and permits geminate codas because they contain a shared root node (the first X is bound).

Some of the attested clusters in Table III do not conform to (9). These are the following:

(10) a. ps, ts, ks, fs,

pj, kj, bj, gj, fj, vj, mj, rj,

bd, d’d, gd, jd

b. pt, kt,

tk, t’k, t’k, čk,

pf,

pš, kš,

kč, dv, dz, nl

There is an important subdivision within the set of clusters in (10). While the clusters in (10b) are always monomorphemic, those in (10a) are predominantly polymorphemic. Specifically, the latter type of clusters are the result of suffixation by 2sg definite imperative -d, 2sg indefinite imperative -j, or 2sg present indicative indefinite -sz. We shall discuss these suffixes in more detail below. They are peculiar in that they may be added to consonant-final stems freely, i.e. without regard to phonotactic constraints. The clusters in (10b), in contrast to those in (10a) are ‘lexically restricted’ in the sense that they only occur in a handful of words (all of which tend to be loans or place names). It has been argued (see Rebrus and Trón 2002) that

---

Note that (rarely) /gd, ps, ks, mj, rj/ also occur monomorphemically.
the phonotactics of place names in Hungarian is different from that of other monomorphemic words, i.e. they form/belong to a separate phonotactic stratum. We could accommodate this observation in the present framework by claiming that there is a special template available for place names (see also section 3.4.3).

The complete list of words with final clusters of the (10b) set that can be found in our database is shown in (11):

(11) pt recept ‘receipt’, korrupt ‘corrupt’
tk Detk ‘place name’
t’k Batyk ‘place name’
t’k barack ‘peach’, palack ‘bottle’, tarack ‘howitzer’
čk Recsk ‘place name’
pf copf ‘plait’
pš arabs ‘Arab horse’, taps ‘applause’
kš voks ‘vote’
kč Szakcs ‘place name’
dv kedv ‘mood’, nedv ‘fluid’, üdv ‘salvation’.
dz edz ‘train准则’, pedz ‘begin to understand’
nl ajánl ‘recommend’

The monomorphemic occurrences of some of the clusters in (10a) mentioned above are also only attested in very few stems:

24 Note that words in final /dz/ cease to be exceptional at the surface since they are pronounced with a geminate affricate [dːz]: e.g. edz [ɛdːz]. Ajánl is often pronounced with a final geminate or non-geminate l: [ɛjaːl] or [ɛjaːl], in which case the phonotactic violation is ‘repaired’.
The licensing of these clusters is discussed later in this section.

Presumably, native speakers are able to identify these items as ‘foreign/strange/non-Hungarian’ (unfortunately, no experimental evidence is available to test this prediction). This statement must be modified slightly if we decide to analyse the phonotactics of place names along the lines sketched above: if they are assumed to constitute a separate phonotactic stratum and thus there is a different template available to them, then one would expect native speakers to be able to identify them as ‘place names’ on the basis of their phonotactic shape. Even then, as opposed to the ‘general’ or ‘regular’ template, this special place-name template would be inert in the sense that it would not interact with the phonotactically motivated processes discussed in Chapter 4.
There are very few exceptions to this requirement, all of which (we claim) are phonotactically irregular: [ft, sk, zg, ź]. The following is a full list of the stems containing these clusters:

    zg rezg ‘vibrate’
    źb idősb ‘elder’, kevesb ‘fewer’, nemesb ‘nobler’

The irregularity of non-coronal obstruent clusters in the coda is confirmed by the fact that they are broken up by epenthesis if (synthetic) suffixation should create such a cluster while coronal clusters are not: compare /ʒiraft/ ‘giraffe’ (acc.) [ʒira:fo:t] and /fošt/ ‘excrement’ (acc.) [fošt] (verbs behave somewhat differently: compare [fošot:] ‘defecate’ (3sg past indef.), see sections 3.4.3 and 4.1.2.)

Note also that affricates are disallowed in coda obstruent clusters regardless of the place of articulation of the other consonant: *ft#, *st#, *št#, *fšt#, *sšt#, *ššt#.

Coda clusters containing sonorants are not constrained by the above requirements: e.g. halk ‘quiet’ lomb ‘foliage’, perc ‘minute’, lánc ‘chain’. Thus, obstruent clusters in branching codas have to obey stricter constraints than other clusters. This suggests that a minimum sonority distance requirement is at play here. Let us assume that there is a minimum of sonority distance that is normally required for government to license clusters as branching...

27 This item is only included because it is usually cited in the literature on Hungarian phonotactics. Actually, it only appears as a bound form before vowel-initial suffixes (as an allomorph of the ‘epenthetic’ stem rezeg, cf. Chapter 4) and thus it is not an exception. Rezg as a free morpheme is obsolete and/or poetic.

28 These obsolete forms are actually polymorphemic. They all contain a no longer productive suffix -b and are not used in ECH. The corresponding regular (attested) forms are idősebb, kevesebb, nemesebb.
Recall that, while individual languages may not reverse sonority relationships in the Sonority Hierarchy, they can have different sonority distance settings between segment classes in the hierarchy. Cf. section 2.2.

Glides are omitted because Hungarian has no glides in our analysis.

Suppose that the sonority distance settings for Hungarian consonants are the following (where $<$ is a smaller sonority distance than $<<$):

\[(14) \quad \text{Sonority Hierarchy: Hungarian}^{30} \]

\[
\text{stops, affricates} < \text{fricatives} < \text{nasals} < \text{liquids}
\]

Furthermore, let us assume that (15) constrains government in Hungarian:

\[(15) \]

\[
a. \quad \text{Government can apply if the sonority distance between the segments in a governing relationship is at least } S_{\text{min}} \\
b. \quad S_{\text{min}} = \text{ }<< \text{ or } >> \\
\text{(where } S_{\text{min}} \text{ is the minimum sonority distance)}
\]

(14) and (15) together leave all obstruent clusters unlicensed. Let us further assume that clusters with subminimal distance, i.e. clusters whose members are not equally sonorous but do not conform to (15), may be well-formed if licensed by some special provision in the grammar. Note that this special licensing may not derive from binding because the consonants the clusters discussed consist of do not necessarily share a COR node since they may differ in the value for [antior]: e.g. most [mošt] ‘now’. Furthermore, coda clusters with subminimal sonority distance are not licensed by virtue of their simply being COR. It is evident that they display the same directionality effects as the clusters that conform to the minimal sonority distance requirement. For instance, /tš/ is a COR cluster, but is not a well-formed branching coda because, albeit minimally, /š/ is more sonorous than /t/. This suggests that it is government that licenses these clusters, but if the distance is subminimal between the members of a cluster, then government is subject to the following constraint in a coda:

---

$^{29}$Recall that, while individual languages may not reverse sonority relationships in the Sonority Hierarchy, they can have different sonority distance settings between segment classes in the hierarchy. Cf. section 2.2.

$^{30}$Glides are omitted because Hungarian has no glides in our analysis.

Let us now examine word-final sonorant clusters. Some of them (notably liquid+ nasal clusters) are licensed by government. Not all of them occur, but we consider these gaps accidental. In the remaining types (nasal+ nasal, nasal+ liquid and liquid+ liquid clusters) the second consonant cannot govern the first one. This correctly excludes nasal+ nasal, nasal+ liquid codas, but incorrectly renders all liquid+ liquid codas inadmissible. Although the evidence is somewhat meagre (as the relevant clusters only occur marginally in a few stems), we claim that government can apply in some of the liquid+ liquid clusters (notably /rl, jl, jr/). This is accounted for if we assume a fine-tuned sonority scale


32With one possible exception: /jm/. See the discussion below.

33We have already seen that the single violation ajánl is irregular (and is normally repaired by the time it surfaces as [ajalː(ː)]). The other apparent violations are j-final and will be discussed later in this chapter.

34As pointed out above, the relevant clusters are rare. /rl/ and /jl/ only occur in görl ‘showgirl’, fájl ‘file’, geil [gejl] ‘nauseatingly sweet’; /jr/ does not occur at all. They might be considered exceptional, but there is evidence (independent of coda phonotactics) for the sonority relations in (17), see section 3.3.2.
in which different liquids represent different degrees of sonority along the lines described in Hooper (1978), Clements (1990). For instance:

(17)  l << r << j

Again, the sonorant coda clusters that appear to violate the directionality requirement (i.e. the SSP) are all j-final.

All sonorant+ obstruent codas are licensed by government. Nevertheless, some of these clusters are ill-formed. Let us examine nasal+ obstruent clusters first. The problem with these clusters is that in addition to the homorganic ones /mp, mb, nt, nd, n’t/, n’d/\(^{35}\) (e.g. kolomp ‘bell’, lomb ‘foliage’, ront ‘destroy’, rend ‘order’, ponty ‘carp’, gyöngy ‘pearl’), non-homorganic /n’t, n’d/ also occur (lány-t ‘girl’ (acc.), hány-d ‘throw’ (sg imp. def.). /n’t, n’d/ only occur polymorphemically. We have pointed out above that definite imperative -d can be added to stems without regard to any phonotactic restriction. Thus, (as we shall see later) the polymorphemic occurrences of /n’d/ are not licensed by being syllabified into the coda. One might want to argue that the same state of affairs applies to /n’t/ as well. This is not the case, however. /n’t/ is always the result of suffixation by the accusative suffix or the past tense suffix. These suffixes are unlike the imperative -d in that a ‘linking’ vowel appears before them to prevent certain consonant clusters from being derived.\(^{36}\) The behaviour of these (types of) suffixes with respect to nasal-final stems is shown (18):\(^{37}\)

\(^{35}\)mt, nk, ng/ also appear in Table III. Of these /mt/ only occurs in two exceptional items teremt ‘create’, vikomt ‘viscount’. The latter two clusters are not problematic because they actually surface as homorganic clusters: e.g. link [liŋk] ‘untrustworthy’, rang [rɒŋ] ‘rank’. [nk] and [ng] are not even non-homorganic underlingly since /n/, from which all surface reflexes of [ŋ] derive is unspecified for place (cf. Appendix A).

\(^{36}\)See the details in Chapter 4.

\(^{37}\)In fact there is a third type of suffix behaviour. Suffixes like the 2sg definite are always vowel-initial unless added to a vowel-final stem: compare nyom-od, bán-od, hány-od (Type A suffixes, cf. section 4.1.2.2).
Though nasal+ voiceless stop clusters are marked universally compared to nasal+ voiced stop clusters (see Kager 1999 and references therein).

(19) would in fact allow unattested coda /nt/ and /nd/, but this causes no problems because /n/ is underlyingly placeless and thus these clusters would surface as [nt] and [nd] as a result of nasal place assimilation (cf. e.g. Vago 1980, Siptár & Törkenczy 2000).

We claim that this difference is attributable to the fact that accusative -t is syllabified into the coda while imperative -d is licensed in a different way. Given this assumption and because the well-formedness of nasal+ stop clusters obviously does not depend on the voicing of stops, /nt, nd/ have to be considered licensed codas. Since nasal+ stop clusters are licensed by government, we have to assume that they also have to meet an additional requirement which filters out those which are ill-formed. This constraint has to disallow non-homorganic nasal+ stop clusters, but permit /nt, nd/. (19) achieves this result:

(19) In a nasal+ stop coda cluster C1C2, C1 must be place-bound unless both of them are COR

It is not obvious whether the same constraint holds for nasal+ affricate clusters or not. Hungarian only has COR affricates, therefore /m/+ affricate codas are predicted to be ill-formed. This prediction appears to be true: there is a single exception /mc/, which only occurs in the morphologically complex form tere[mč] ‘create’ (2sg imp. def.).

Note, however, that the constraint extended to nasal+ affricates would permit /nt’, n’č, n’j, nj/ as coda clusters, but they do not occur. Of these, /nj/ is permitted since /nc/ occurs (e.g. mancs ‘paw’) and it is unlikely that the voicing difference should entail a difference in well-formedness. The others do not occur because an MSC (which bans preconsonantal nasals with an independent place specification) excludes them within the morpheme, and they may not result from suffixation

---

38 Though nasal+ voiceless stop clusters are marked universally compared to nasal+ voiced stop clusters (see Kager 1999 and references therein).

39 (19) would in fact allow unattested coda /nt/ and /nd/, but this causes no problems because /n/ is underlyingly placeless and thus these clusters would surface as [nt] and [nd] as a result of nasal place assimilation (cf. e.g. Vago 1980, Siptár & Törkenczy 2000).

40 On the syllable structure of this form, see Section 4.2.4.3.
because there are no suffixes that consist of an affricate. For this reason, permitting them causes no problems and therefore we extend (19) to cover nasal+ affricate codas too:

(20) In a coda cluster \( C_1 \) \( C_2 \)

\[
\begin{array}{c}
\text{root} \\
[+ \text{nas}] \\
[-\text{nas}] \\
[-\text{cont}] \\
\end{array}
\]

\( C_1 \) must be place-bound unless both of them are COR

The distribution of final nasal+ fricative clusters is governed by a slightly different constraint. Disregarding /v/ for the moment (whose sonority ranking, as we shall see, is problematic), only those nasal+ fricative clusters whose first member is /n/ and whose second member is a coronal seem to be well-formed. Although /mf, ms, mz, n’s/ are attested, they only occur (i) in polymorphemic clusters whose second consonant is 2sg present indefinite -sz, which always behaves in a special way (/ms/ nyom-sz ‘push’; /n’s/ hánysz ‘throw’), or (ii) in a few irregular stems. The following is an exhaustive list of these stems:

(21)

\[
\begin{align*}
mf & \quad \text{tromf ‘trump’} \\
mz & \quad \text{nemz}^{41} \quad \text{‘beget (3sg pres. indef.)’} \\
mž & \quad \text{tömzs ‘lode’}
\end{align*}
\]

Given the MSC referred to above, (22) accounts for the observed distribution:

(22)

In a coda cluster \( C_1C_2 \) where \( C_1 \) is [+ nasal] and \( C_2 \) is [-son, + cont], \( C_2 \) must be COR.

In the above discussion of branching codas we have disregarded /v/. As argued in Siptár & Törkenczy (2000), /v/ is a ‘two-faced’ consonant: it behaves as a sonorant in onsets, but as

---

41This item is obsolete in ECH as a free form.
an obstruent in codas (this is encoded by its unspecification for the feature [son], see Appendix A). We propose that despite its asymmetrical behaviour in onsets and codas (cf. Siptár & Törkenczy 2000 section 4.1.1) /v/ always has the sonority ranking of a fricative. Given the constraints discussed in the present section and the above assumption about its ranking in the sonority hierarchy, its distribution in final clusters is the expected one. Thus, government does not allow it to cooccur with obstruents as the first or the second element in coda clusters since the sonority distance between fricatives and other obstruents is subminimal. The few exceptional stems in which it does cooccur with obstruents in final position are listed in (11). Because of the directionality of government /v/+ sonorant codas are ill-formed. In codas /v/ is permitted following a sonorant if the sonorant is a liquid or a nasal since government can apply (e.g. ötv ‘argument’, elv ‘principle’, ölyv ‘hawk’) and (22) does not constrain nasal+ /v/ coda clusters as /v/ is unspecified for [son] (e.g. hamv ‘ash’, ellenszenv ‘antipathy’, könyv ‘book’).

We have not examined the distribution of preconsonantal /j/ yet. Törkenczy (1994a) claims that there are two constraints that apply to preconsonantal /j/ in branching codas; one requires that obstruents following /j/ must be coronal and the other excludes palatal consonants after /j/. In our view, these constraints are untenable because they make unmotivated and unnecessary distinctions between equally well-formed /jC/ clusters. For instance, /jn/ and /jm/ conform to the above constraints and thus are judged well-formed as opposed to /jn\²/, which violates them, and is therefore supposed to be ill-formed. This seems to make the right prediction since /jn/ and /jm/ are attested, but /jn\²/ is not. However, this difference is not really significant since the only stems in which the first two clusters are attested are kombájn

It would be undesirable to allow the sonority ranking of a segment to vary depending on the position it occurs in. The sonority ranking of /v/ as a fricative in onsets generally does not cause problems because we argue that branching onsets do not occur in Hungarian. /v/, however, does behave in a special way in onsets in interconstituent clusters, see Section 3.3.2.1.

Subminimal government is excluded because /v/ is not COR.

/v/ is extremely rare in final clusters containing sonorants other than /r/. The following is a complete list: elv, nyelv ‘language’, ölyv, hamv, -szenv ‘feeling’ < bound morpheme>, enyv ‘glue’, könyv. Of these hamv is obsolete as a free form. In the present analysis this is seen as an accident.
‘combine-harvester’ and slejm ‘phlegm’. In fact, preconsonantal /j/ is rare in final (nongeminate) clusters other than /jt/.\(^{45}\) Furthermore, /jt/ is the only cluster whose coda status can be tested in alternations: as the accusative and the past suffix (which can be realized as [t]) attach to /j/-final stems without an epenthetic vowel (sóhaj-t [šo:ja:t] ‘sigh’ (acc.), búj-t [bu:jt] ‘hide’ (past)), /jt/ must be a possible coda (cf. section 4.1.4). In order to avoid making untestable well-formedness distinctions within the set of final /jC/ clusters we claim here that all of them are well-formed and no constraint applies specifically in this environment.

3.2.4.3. Appendices

In the discussion of final clusters in the previous section we have disregarded the final clusters that contain the consonants /d, j, s/ when they realize the definite imperative, the indefinite imperative, and the 2sg present indefinite suffix respectively. Final clusters containing these suffixes often violate government (e.g. lópsz ‘steal’ (2sg pres. indef.)) and/or other constraints applying within the coda (nyomd ‘push’ (sg imp. def.)). In general, there are no phonotactic constraints applying between these suffixes and the final consonant of the stem they are attached to. This is completely true of definite imperative -d and indefinite imperative -j. These suffixes may be added to any stem. The gaps in /j/ or /d/ final clusters in Table III are not due to phonotactics: they are either (i) accidental (there are no verb stems ending in affricates or /x/, so /ťj, čj, jy/ and /xd/, /xj/ do not occur), or (ii) the result of assimilations (e.g. [šd]

\(^{45}\) /j/ frequently occurs in the final polymorphemic clusters /jd, js, jč/ but they all contain analytic suffixes (2sg definite imperative -d, indefinite imperative -j, or 2sg present indicative -sz). The following is a complete list of stems with final jC clusters other than /jt/: /jp/ selyp ‘lisper’; /jk/ sejk ‘sheik’, sztrájk ‘strike’, -ajk ‘lip < bound form> ’, hüvelyk ‘thumb’; /jd/ fajd ‘grouse’, gajd ‘hubbub’, majd ‘later’, ofszájd ‘offside’; /jg/ caig ‘cheap cloth’; /jt/ Svájc ‘Switzerland’; /jű/ dölyf ‘arrogance’; /jš/ fádervejsz ‘talcum powder’, hajsz < interjection> ; /jš/ Majs < place name> ; /jv/ ólyv ‘hawk’; /jž/ csuszpájz ‘a kind of vegetable dish’ , rajz ‘drawing’, spáţ ‘pantry’; /jž/ pajzs ‘shield’; /jm/ slejm ‘phlegm’; /jn/ kombájn ‘combine harvester’; /jl/ fájl ‘file’; /jh/ bolyh ‘fluff’. The items selyp and bolyh are only included because they are cited in the literature-in ECH they are bound forms that only occur before vowel-initial suffixes; as free forms they are obsolete.

\(^{46}\) Even if there were such underlying forms, they would be eliminated by assimilation (cf. Siptár & Törkenczy 2000).
In striking contrast, there are just a handful of irregular monomorphemic items (listed in (12)) and none containing suffixes other than -d, -sz and -j that are in violation of sonority sequencing. The suffix -sz behaves similarly, albeit to a somewhat limited extent. Final clusters containing it may violate sonority sequencing (e.g. /ps/ kapsz ‘get’ (2sg pres. indef.), /t's/ vágysz ‘desire’ (2sg pres. indef.), /ks/ raksz ‘put’ (2sg pres. indef.), but there is a phonotactic(ally motivated) phenomenon that concerns -sz: it cannot be attached to a stem that ends in a [+ strident] consonant; instead the allomorph -Vl is selected (e.g. tesz-el ‘put’ (2sg pres. indef.), néz-el ‘look’ (2sg pres. indef.), keres-el ‘search’ (2sg pres. indef.) and not *te[s], *ne[s], *kere[šš]. The allomorphy is certainly phonologically motivated, but we assume that it is not related to syllable structure: -Vl is selected even when the cluster that would result from suffixation with -sz is a well-formed coda/final cluster such as geminate [s], e.g. tesz-el (compare geminate [d] and [j] that can be the result of suffixation with definite imperative -d and indefinite imperative -j respectively: add [od:] ‘give’ (2sg imp. def.), fejj [fėj:] ‘milk’ (2sg imp. indef.)).

Another aspect of the independence of these suffixes can be seen if we examine word-final clusters that consist of more than two consonants. There are extremely few monomorphemic words that end in more than two consonants. (23) lists all the relevant items:

---

47 In striking contrast, there are just a handful of irregular monomorphemic items (listed in (12)) and none containing suffixes other than -d, -sz and -j that are in violation of sonority sequencing.

48 [šš] final forms like kere[šš] do rarely occur, but they are obsolete/unusual in ECH.
Table IV. Polymorphemic word-final CCC clusters

<table>
<thead>
<tr>
<th></th>
<th>t</th>
<th>d</th>
<th>s</th>
<th>j</th>
</tr>
</thead>
<tbody>
<tr>
<td>dz</td>
<td>—</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mj</td>
<td>—</td>
<td>+</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>nj</td>
<td>—</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ns</td>
<td>—</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nl</td>
<td>—</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>lj</td>
<td>—</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rn</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rj</td>
<td>—</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rl</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>jj</td>
<td>—</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>js</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>jś</td>
<td>—</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>jn</td>
<td>—</td>
<td>—</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Examples: edzd ‘train’ (sg imp. def.), teremtsd ‘create’ (sg imp. def.), lengsz ‘swing’ (2sg pres. indef.), zengd ‘resound’ (sg imp. def.), pánzt ‘money’ (acc.), bontsd ‘open’ (sg imp. def.)
Table IV shows the final polymorphemic CCC clusters that appear at the surface (the notation is the usual one where the boxes struck out by dashes denote clusters that are/would be eliminated by assimilations). The number of underlying clusters would be higher than those shown in Table IV because there are processes that simplify consonant clusters (e.g. degemination (cf. Siptár & Törkenczy 2000) turns the underlying triliteral cluster /rrs/ into [rs] in varrsz [vɔr̩ʃ] ‘sew’ (2sg pres. indef.) and Fast Cluster Simplification (cf. section 4.2.3.) may delete the second consonant from underlying /Ngd/ in fingd [fiŋd] ‘fart’ (sg imp. def.)). The attested clusters either have one of the three suffixes -d, -j, -sz discussed above, or the accusative -t as their final element. Let us set aside the accusative for the moment and concentrate on the other three. The fact that they can create final CCC clusters by attaching to stems ending in branching codas further attests to their phonotactic independence. We can account for this independence by claiming that they are in fact not syllabified into the coda, but belong to a special subsyllabic constituent, the appendix. Thus, in Hungarian the extended syllable can have an appendix not only initially, but finally as well:

(24)

The extended syllable shown in (24) is restricted to the right edge of analytic domains, i.e. the appendix must be peripheral. Furthermore, it may only occur after a coda. Consonant clusters that are (partially) in the appendix are unconstrained by coda restrictions (e.g. may violate sonority sequencing and may consist of more than two consonants). Regularly, on the coda side only analytic suffixes may be in the appendix. -d and -j are clearly syllabified into the appendix. We have seen that they can be added to any stem-final consonant and they can occur
as the last consonant in final CCC clusters, i.e. they can attach to any stem ending in a branching coda. The surface non-occurrence of some final C1C2C3 clusters where C1C2 is a possible coda and C3 is -d or -j is due to assimilations (e.g. [ltj] does not occur because /tj/ becomes [č] in imperatives (cf. Siptár & Törkenczy 2000)); or is unrelated to syllable structure (e.g. /mbd/ and /mbj/ do not occur because the morphology does not generate these combinations (there are no verb stems ending in /mb/, and -d and -j are verbal suffixes).

The appendix status of -sz is more problematic. We have seen that, modulo -Vl allomorphy, there are practically no restrictions between the stem final consonant and -sz. This is what we expect of an appendix. However, an inspection of Table IV reveals that very few -sz-final CCC clusters are attested (only [ŋks/ and [nls]). This is unexpected even if one takes it into consideration that -sz can only be added to verbs. The reason why there are few -sz-final CCC clusters is that, typically, -sz attaches to stems that end in more than one consonant with a linking vowel (e.g. látz, ’látasz ’see’ vs. *osztasz, osztasz ’distribute’; adsz, *adasz ’give’ vs. *kezdsz, kezdesz ’begin’). Some stems allow forms with and without the linking vowel (e.g. fingasz, fingsz ‘fart’); others only have forms without it (e.g. varrsz but *varrasz ‘sew’).

The occurrence of accusative -t in final CCC clusters is a further complication. We have pointed out above that accusative -t syllabifies into the coda, and if it cannot, an epenthetic vowel occurs before it (compare leves-t ‘soup’ (acc.) and zsirá-f-o-t ‘giraffe’ (acc.)), i.e. there is phonotactic interaction between it and the stem-final consonant (see the details in Chapter 4). Since three-term codas are not allowed, we would expect it to be preceded by an epenthetic vowel after stems ending in consonant clusters. Table IV shows that this is not the case; it can appear as the last consonant in a final CCC cluster. It does not behave like an appendix, however: it never occurs without an epenthetic vowel after stem final clusters C1C2 if C2+ [t] is not a possible branching coda (compare farm-ot ‘farm’ (acc.) and konszern-t), i.e. there is a phonotactic interaction between the -t and the stem-final consonant. On the other hand, the epenthetic vowel is not always missing if C2 of the stem-final cluster plus [t] make a possible branching coda. Some of these cases have independent explanations (e.g. all /rn/-final nouns are ‘lowering stems’, which in itself requires a linking vowel to be present before the suffix even though /n’t/ is a possible coda: árny-at ‘shadow’ (acc.), szárny-at ‘wing’ (acc.), szörnyet ‘monster’ (acc.), and similarly with other combinations, e.g. fürj-et ‘quail’ (acc.), törzs-et ‘trunk’ (acc.), cf. section 4.1.3. on lowering stems). Others, however, have
no independent explanation: /rš/-final stems always have a linking vowel before the accusative, even if the stem is not lowering: e.g. bors-ot ‘pepper’ (acc.) although /rš/ and /št/ are well-formed codas.

Thus the general problem is that the phonotactic restrictions on the melodic content of final clusters and those on the complexity of final clusters seem to suggest conflicting classifications for 2sg present indefinite -sz and accusative -t. The former suffix is completely insensitive to the melodic content of the stem-final consonant (a typical appendix-like behaviour: csap-sz ‘hit’ (2sg pres. indef.), but usually cannot be attached to cluster-final stems without a linking vowel (oszt-asz ‘distribute’ (2sg pres. indef.)). The latter, on the other hand, is sensitive to the stem-final consonantal melody (a typical coda-like behaviour: ón-t ‘tin’ (acc.) vs. nyom-ot ‘trace’ (acc.)), but if the final consonantal melody is right, can be sometimes added to stems ending in a consonant cluster without a linking vowel (konszern-t). The question is how to explain the non-appendix like behaviour of -sz and the non-coda like behaviour of -t.

First, let us try to answer the first part of the question. The problem is that—contrary to our expectations—a linking vowel appears after cluster-final stems before the suffix -sz (which is assumed to syllabify as an appendix). It is significant that not only the presence, but the quality of the linking vowel is also unexpected. The normal linking vowels are mid e/ö/o (i.e. [-open]) and not low a/e (i.e. [+ open]). The latter quality is the one that we get after lowering stems and suffixes: e.g. fog-at ‘tooth’ (acc.), szög-et ‘nail’ (acc.), tök-ök-et ‘pumpkins’ (acc.) (compare non-lowering bog-at ‘knot’ (acc.), rög-öt ‘clod’ (acc.)). Here, however, the lowered quality is not due to the stem but to the suffix itself. Verb stems are never lowering, and the quality of the linking vowel before -sz is always (unexpectedly) low after all the stems that it can follow (compare mond-ok ‘say’ (1sg pres. indef.) and mond-asz ‘say’ (2sg pres. indef.). Also, a low linking vowel normally does not alternate with zero: it

49Cf. Rebrus (to appear). The past suffix also behaves similarly to the accusative (cf. section 4.1.2.2.).

50Because of its behaviour after cluster-final stems, in some treatments -sz is regarded as a quasi-analytic suffix, see section 4.1.2.2 and Rebrus (2000), Rebrus and Törkenczy (1998).

51Underlying mid e is eventually phonetically implemented in ECH as low [ɛ] (cf. Siptár & Törkenczy 2000).
is present after lowering stems even after stems whose final consonant could form a well-formed branching coda with the suffix: e.g. *fal-at* ‘wall’ (acc.), *has-at* ‘stomach’ (acc.), *vár-at* ‘castle’ (acc.) (compare non-lowering *hivatal-t* ‘office’ (acc.), *kas-t* ‘hive’ (acc.), *vér-t* ‘blood’ (acc.)). By contrast, the low linking vowel of *-sz* is not stable: it is not present after some stems (compare *hat-sz* ‘influence’ (2sg pres. indef.) and *tart-asz* ‘hold’ (2sg pres. indef.)) We suggest that considering the *-Vsz* variant of *-sz* as an instance of allomorphy is in harmony with these facts. Then the *-Vsz* variant appears in the lexicon along with *-sz* and thus the unexpectedness of the vowel quality is then just a lexical fact. The low initial vowel of the *-Vsz* variant behaves just like any low linking vowel, i.e. it is stable and does not alternate with zero. The fact that the *-Vsz* allomorph typically appears after cluster-final stems, i.e. the allomorphy is phonotactically conditioned, is on a par with the behaviour of *-Vl*, which is a variant of *-sz* after [+ strident] stems. Both can be seen as cases where morphology is dependent on phonological information. This interpretation makes it possible to maintain that *-sz* syllabifies as an appendix.

A possible answer to the second part of the question (which concerns the behaviour of *-t* after cluster final stems) is to claim that the reason why *-t* attaches without a linking vowel to stems like *konszern* is that in these words the first consonant of the final cluster is not in the coda but in the nucleus as shown in (25) (where only the relevant structure is displayed):

\[
\begin{array}{c}
\sigma \\
R \\
\text{O N C} \\
| \\
| \\
X X X X \\
\text{konszern}
\end{array}
\]

This syllabification would allow *-t* to attach to *konszern* without a linking vowel (i.e. into the

\[\text{See section 4.1.3 on lowering stems.}\]

\[\text{The *-Vsz* allomorph also appears after a more-or-less arbitrary set of stems ending in *t* preceded by a long vowel. All the stems ending in the verb-forming suffix *-i t* belong here. E.g. *alakí t-asz* ‘form’, *vét-esz* ‘err’, *fűt-esz* ‘heat’ (compare *lát-sz* ‘see’).}\]
coda of the stem-final syllable) in spite of the fact that codas are maximally binary branching. In other words, it would explain why -t appears to be insensitive to the number of stem-final consonants (in some stems) although it must be syllabified as a coda. There are, however, several problems, which make this explanation untenable.

First of all, a very complicated statement would be needed to specify the conditions in which a consonant can syllabify into the nucleus. The class of consonants that could syllabify in this way is not difficult to identify. They must be [+ sonorant], since -t always attaches with a linking vowel to C1C2-final stems if C1 is an obstruent, even if C2+t is a well-formed coda: *taps-t ‘applause’ (acc.) (compare kas-t ‘hive’ (acc.), most ‘now’), sze[ks]-et not *sze[ks]-t ‘sex’ (acc.) (compare szes-t ‘alcohol’ (acc.), liszt ‘flour’). Furthermore, [+ sonorant] segments could only be in the nucleus of a closed syllable otherwise we would predict that (i) any consonant can follow a sonorant within the same syllable and (ii) no interconstituent constraints refer to sonorant final syllables (these predictions are untrue; cf 3.2.4.2 and 3.3.2). It is a significant fact that additional conditions would also have to be stipulated since /r/ behaves differently from other sonorants when it precedes stem-final obstruents. As pointed out above, -t attaches to /rš, rs, rz/-final stems with a linking vowel (bors-ot ‘pepper’ (acc.), mersz-et ‘courage’ (acc.) borz-ot ‘badger’ (acc.)). The problem is that (i) /r/ does not behave in this way before other stem-final consonants (compare e.g. konszern-t), and (ii) other sonorants do not behave in this way before stem-final /š, s, z/ (compare e.g. konstans-t ‘constant’ (acc.), fajansz-t ‘faience’ (acc.), pénz-t ‘money’ (acc.)). Thus, the conditions on the hypothesized syllabification of sonorants into the nucleus can hardly be formulated with a sufficient degree of generality.

Furthermore, there is some degree of unpredictable variation. Some stem-final sonorant+ /s, š, z, ž, n, ʃ, l, r, j/ clusters show more than one kind of behaviour: in some stems they always take the suffix -t without a linking vowel, in others they always require a linking vowel, and there are stems that allow both variants:

54 Note that these are not lowering stems, thus the presence of the linking vowel cannot be attributed to a factor independent of the clusters examined.
A similar behaviour can be attested if we examine stems that end in geminates whose melodic content is such that after the corresponding stem-final non-geminate segments no linking vowel appears. These are the stems ending in /ss, űs, űn, n̩n̩, ıı, rr, jj/. Typically, no linking vowel appears after these geminates before -t: *idill-t ‘idyll’ (acc.), *finn-t ‘Finnish’ (acc.), *plüss-t ‘plush’ (acc.), *dzessz-t ‘jazz’ (acc.), etc. Most of the examples with a linking vowel are lowering stems and thus they are irrelevant to the issue at hand: e.g. *uijj-at ‘finger’ (acc.), *toll-at ‘feather’ (acc.), *gally-at ‘twig’ (acc.), etc. (cf. section 4.1.3 on lowering stems). Nevertheless, here too there is some idiosyncratic variation: *genny-t/genny-et ‘pus’ (acc.), *orr-t/orr-ot ‘nose’ (acc.), *bross-t/bross-ot ‘brooch’, etc. These facts suggest that idiosyncratic restrictions would have to be imposed on the incorporation of sonorants into the

\[\begin{array}{cccc}
n\ddash & \text{protestáns-t} & \text{revans-ot} & \text{briliáns-t/briliáns-ot} \\
n\ddash & \text{fajansz-t} & \text{sansz-ot} & \text{reneszánsz-t/reneszánsz-ot} \\
n\ddash & \text{csuszpájz-t} & \text{rajz-ot} & - \\
n\ddash & \text{páñz-t} & \text{bronz-ot} & \text{csimpánz-t/csimpánz-ot} \\
j\ddash & - & - & \text{pajzs-t/pajzs-ot} \\
\end{array}\]

(26) 55  


56 There are no stems ending in /ZZ/.

57 These three-term clusters actually surface as two-term as a result of Degemination (cf Siptár & Törkenczy 2000).

58 There are no (non-lowering) /ss, űs, űn, n̩n̩, ıı, rr, jj/-final noun stems that select the variant with the linking vowel only, unless we include examples like *mell-ot ‘breast’ (acc.), *szenny-ot ‘dirt’ (acc.) which cannot be identified as lowering stems on the basis of the quality of the linking vowel in ECH (though other dialects show that they are lowering stems, cf. 4.1.3). There are some comparable verb stems, however, which always require a linking vowel before the past tense suffix -(t)t (which behaves similarly to the accusative, cf. section 4.1.4.4): *hall-ott ‘hear’ (3sg past indef.), *hull-ott ‘fall’ (3sg past indef.), *kell-ett ‘have to’ (3sg past indef.), *vall-ott ‘confess’ (3sg past indef.) Note that *hull-t ‘fall’ (3sg past indef.) is a possible alternative form along with *hall-ott.
nucleus. Some stems would have to be marked as not allowing it and others as optionally allowing it.

Even if the complexities/difficulties pointed out above were disregarded, the most serious problem with the hypothesis is that, after the relevant clusters, -t behaves in the same way even if the vowel preceding the cluster is long: kombájn-t, protestáns-t ‘protestant’ (acc.), pánz-t ‘money’ (acc.) ([nst] or [nt’t]), fájl-t ‘file’ (acc.). etc. Therefore, if we maintained that in these stems the postvocalic consonant of the final cluster is in the nucleus, then we would have to allow ternary branching nuclei. In fact, in trying to avoid ternary codas we would end up creating ternary nuclei.

Because of the problems discussed above we consider the syllabification shown in (25) untenable and suggest that the behaviour of -t after cluster-final stems like konszern is due to the fact that the relevant stems are lexically marked so that they exceptionally allow the syllabification of -t into the appendix. These stems will then have no linking vowel before -t. The stems that show variation (e.g. briliáns) appear twice (marked and unmarked) in the lexicon of speakers who use both variants. The lexically unmarked cluster-final stems will always have a linking vowel before -t. Thus, all word-final clusters containing the suffixes -d, -j, -sz have the structure coda+ appendix and accusative -t can also syllabify as an appendix after some cluster-final stems. This makes it possible to maintain that the coda constituent is maximally binary branching in Hungarian, although word-final ternary clusters do occur. In addition to the restriction on its melodic/morphological content, the occurrence of the appendix is subject to the following general condition (which is a version of the Peripherality Condition, cf. Hayes 1995):

(27) The appendix (i) must be peripheral in an analytic domain and (ii) must not be adjacent to the nucleus.

59 This suggests that for some speakers even these stems can be non-variable, a prediction that appears to be true.
3.3. Transsyllabic constraints

Transsyllabic constraints are constraints applying between adjacent segments belonging to different syllables. Logically, transsyllabic constraints could refer to segment clusters of the following kinds:

(28)  
   a. V.V  
   b. C.C  
   c. V.C  
   d. C.V

(28a) shows two adjacent nuclei (hiatus), (28b) is a coda followed by an onset (interconstituent cluster), (28c) is a nucleus followed by an onset, and (28d) is a coda consonant followed by a nucleus. Out of these four possibilities (28d) appears to be universally excluded by the Maximal Onset Principle (cf. Blevins 1995) (or any equivalent mechanism designed to capture the fact that a prevocalic consonant syllabifies universally as an onset rather than a coda\(^{60}\)). There are no transsyllabic constraints applying between a vowel and a following non-tautosyllabic consonant in Hungarian (28c) (for the distribution of long vowels before consonant clusters see section 3.4.2). Let us examine the constraints applying in contexts (28ab).

3.3.1. Hiatus

Nuclei can be adjacent (hiatus may occur) with the following restrictions: The initial vowel of vowel-initial synthetic suffixes deletes when they are attached to vowel-final stems (compare **ház-on** ‘on (the) house’ and **kapu-n** ‘on (the) gate’, cf. section 4.1.4.2). Some of the

---

remaining vowel clusters are broken up by a (postlexical) process of hiatus filling (e.g. fáig [faːjig] ‘up to the tree’ cf. Siptár and Törkenczy 2000 and references cited therein). The rest of the hiatuses surface (e.g. kakaó [kɔkɔː] ‘cocoa’, csataordí tás [ˈʃoːorditːaʃ] ‘battle cry’).

Table V shows clusters of two vowels (vowels in hiatus) that occur in Hungarian. Table V abstracts away from hiatus filling, but is near surface in the sense that it shows vowel clusters that survive after the deletion of the initial vowels of vowel-initial synthetic suffixes in hiatus. Blanks indicate that the combination in question is not attested, stars mark vowel clusters that only occur when separated by an analytic boundary, and a given vowel cluster is spelt out if there is at least one monomorphemic stem in which it occurs.

Table V VV clusters

<table>
<thead>
<tr>
<th>i</th>
<th>ü</th>
<th>e</th>
<th>ö</th>
<th>u</th>
<th>o</th>
<th>a</th>
<th>i:</th>
<th>ü:</th>
<th>e:</th>
<th>ö:</th>
<th>u:</th>
<th>o:</th>
<th>a:</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>*</td>
<td>*</td>
<td>ie</td>
<td>*</td>
<td>iu</td>
<td>io</td>
<td>ia</td>
<td>*</td>
<td>*</td>
<td>ie</td>
<td>iő:</td>
<td>iu</td>
<td>io</td>
</tr>
<tr>
<td>ü</td>
<td>*</td>
<td>*</td>
<td>üe</td>
<td>*</td>
<td>üa</td>
<td>üe:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e</td>
<td>ei</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>eu</td>
<td>eo</td>
<td>ea</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>eo</td>
<td>ea:</td>
</tr>
<tr>
<td>ŏ</td>
<td>u</td>
<td>ui</td>
<td>*</td>
<td>ue</td>
<td>uu</td>
<td>uo</td>
<td>ua</td>
<td>ui:</td>
<td>*</td>
<td>*</td>
<td>uo</td>
<td>ua:</td>
<td></td>
</tr>
<tr>
<td>o</td>
<td>oi</td>
<td>oe</td>
<td>oo</td>
<td>oa</td>
<td>oe:</td>
<td>ooa:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>ai</td>
<td>*</td>
<td>ae</td>
<td>*</td>
<td>au</td>
<td>ao</td>
<td>ai:</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>ao:</td>
<td>*</td>
</tr>
<tr>
<td>i:</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ť:</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e:</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>ŏ:</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ů:</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>u:</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>o:</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a:</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


---

61 Only monomorphemic examples are given.

In most cases the difference between stars and blanks in Table V is phonologically accidental. The reason is that morphologically complex hiatuses only survive (i.e. escape deletion) if the two nuclei become juxtaposed as a result of (i) compounding (disznóölés ‘pig killing’), (ii) prefixation (by preverbs, e.g. beleönt ‘pour into’) or (iii) analytic suffixation (e.g. ollóért ‘for scissors’), i.e. when they are separated by an analytic domain edge, a boundary across which no phonotactic or syllable structure constraints hold (cf. section 2.2). In a few cases, blanks are due to a regularity which is unrelated to hiatus. Thus, the lack of stars in the rows for /o, ö/ (the non-occurrence of polymorphemic /o, ö/-initial vowel clusters) is due to the fact that /o, ö/ are not permitted at the end of an analytic domain in general (cf. section 3.4.). Vowel Harmony (cf. Siptár and Törkenczy 2000 and references cited therein) accounts for the absence of monomorphemic vowel clusters containing front rounded vowels and back vowels.  

Bearing in mind the above observations, the following regularities specific to hiatus can be observed in Table V. In hiatus

\[(29)\]
\begin{enumerate}
\item identical segments cannot occur.
\item a long vowel cannot be prevocalic.
\item the vowels /ö, öː, ü, üː/ cannot occur.
\end{enumerate}

No separate statement needs to be included in the grammar of Hungarian to account for (29a) since it can be explained with reference to a general constraint on the form of phonological representations, the Obligatory Contour Principle (OCP), which bans adjacent identical

[62]núansz ‘nuance’ and Szöul ‘Seoul’ are the only exceptions. Note also the exceptional behaviour of domain-final /o/ in ‘foreign compounds’ discussed below.
elements on the same tier.\textsuperscript{63} Vowel clusters containing identical segments within an analytic domain\textsuperscript{64} are impossible to represent since (30a) below is the representation of long vowels and (30b) is excluded by the OCP.

\begin{equation}
\begin{array}{c}
\text{(30)} \\
a. \ X \ X \\
\downarrow V_a \\
b. \ X \ X \\
\downarrow V_a \downarrow V_a
\end{array}
\end{equation}

Words like \textit{kiirt} [kiirt] ‘exterminate’, \textit{rakátaautó} [rak:ta:aut:o] ‘rocket car’, \textit{biliig} [biliig] ‘up to the chamber pot’, etc. are only apparent counterexamples since in them the vowel segments making up the relevant clusters are in different analytic domains, they are ‘fake’ geminate vowels (and can be represented as (30b) without violating the OCP): [[\textit{ki}][\textit{irt}]], [[\textit{ra}][\textit{káta}][\textit{aut}o]], [[\textit{bili}][\textit{ig}]].\textsuperscript{65} The same applies to vowel clusters consisting of identical long vowels as in [[\textit{lő}][\textit{gárt}]] ‘for juice’. In Table V the monomorphemic occurrence of /oo/ and /uu/ (e.g. \textit{vákuum} ‘vacuum’, \textit{zoológia} ‘zoology’) seem truly problematic as they appear to be in violation of (29a) (i.e. the OCP). However, the few words that contain them\textsuperscript{66} are either pronounced with single long (rarely short) vowels ([vakum/vakum, zoologia]) or, rarely, with fake geminates comparable with the cases described above ([vakuum, zoológia]). We assume that in the latter pronunciation these words have been reanalysed as containing more than one internal analytic domain: [[\textit{váku}][\textit{um}]], [[\textit{zoolo}][\textit{gi}a]] to avoid violating the OCP (cf. 4.1.4.5). Note that the latter form is problematic in another respect: it violates the language-specific constraint (58) against domain-final mid vowels (cf. 3.4.1). If the form is pronounced


\textsuperscript{64}We assume that the OCP does not hold across the edge of an analytic domain.

\textsuperscript{65}In fact, the fakeness of these geminates can even be heard in Hungarian as the two vowels are pronounced with distinct pulses, compare \textit{kiirt} [kiirt] vs. \textit{si rt} [\textit{Sr}t] ‘cried’. The difference between fake and true geminate vowels is even more apparent in the case of á and é: \textit{odaad} [odo\dot{a}d] ‘give over to’ vs. \textit{kád} [ka:d] ‘tub’, \textit{leesik} [Ies\dot{e}sk] ‘fall down’ vs. \textit{késk} [ke\dot{e}sk] ‘be late’ (cf. 4.1.4.5).

\textsuperscript{66}The following is a complete list of the relevant items in our database: \textit{kooperál} ‘cooperate’, \textit{koordináta} ‘coordinate’, \textit{zoológia, individuum} ‘individual’, \textit{vákuum}.
The behaviour of a - á and e - é is exactly like that of the other short - long pairs, which confirms that (despite the phonetic difference in quality between them) the members of these respective pairs should have identical underlying feature melodies, i.e. should be represented on a par with other short - long pairs.

These are often phonologically identifiable as compounds: e.g. kvázi-vörös (note the lack of vowel harmony), paraanyag [pərən'yag] ‘suberin’ (note the fake geminate and the lack of Low Vowel Lengthening in para-), paraszövet ‘phellem’ (note the lack of vowel harmony and that of Low Vowel Lengthening in para-). Curiously, in ECH short /o/ can occur finally in the first member of these structures even when they are transparently compound-like and other phonological phenomena (such as the lack of vowel harmony) mark them as compounds: e.g. pszicho-biológia [psihobioloːgia, *psihoːbioloːgia] ‘psychobiology’, pszeudo-főnév [pseudofőːneːv, *pseudofőːneːv] ‘pseudo-noun’. It is unclear why (58) (an otherwise very active constraint which even loans have to conform to) can be violated in just these cases.

We have assumed that in Hungarian the members of all long – short vowel pairs are melodically identical and their only difference is that the root of a given feature tree is associated to one timing slot in the short member and two timing slots in the long one. Given this assumption, (in addition to clusters of identical vowels) the OCP should also exclude vowel clusters in which a short vowel and its long counterpart combine in any order (e.g. ií, íi). This prediction is borne out: hiatuses of this type do not occur, either (when undivided by an analytic domain edge). 67 Naturally, the OCP does not exclude these clusters if the vowels they consist of belong to different analytic domains: [[ki][í r]] ‘write out’, [[le][őr]] ‘burn down’, [[si][íg]] ‘up to the ski’, [[rú][ad]] ‘put on’, [[fő][őrdőg]] ‘chief devil’, [[ki][í r]] ‘write out’ etc.

---

67 The behaviour of a - á and e - é is exactly like that of the other short - long pairs, which confirms that (despite the phonetic difference in quality between them) the members of these respective pairs should have identical underlying feature melodies, i.e. should be represented on a par with other short - long pairs.
(29a) is clearly a systematic (i.e. non-accidental) OCP-based constraint. By contrast, the interpretation of (29b) and (29c) is less obvious. The reason is that although these constraints are almost exceptionless, it is difficult to say if they account for accidental or systematic gaps. (29b) is a better candidate for a systematic regularity, because there are sporadic examples in which an original prevocalic long vowel shortens in loan words adopted into Hungarian (e.g. [buik] and not *[buík] (from English Buick [bjuːk])). (29c) may well be an accident due to several (unrelated) factors: Vowel Harmony (see above), the relative infrequency of front rounded vowels and accidents of borrowing (most monomorphemic items with hiatus are loans).

(i) One possible analysis is that we take hiatus to be well-formed in general and consider (29bc) to be systematic regularities. Then they can be expressed as (31ab):

(31)

\[
\begin{align*}
\text{a. } & \ast N \quad N \\
& X \quad X \quad X \\
\text{b. } & \ast N \quad N \\
& X \quad (X) \quad X \quad (X) \\
\end{align*}
\]

Condition: the segment associated with either (or both) roots is LAB

\[\text{\textsuperscript{68}The full list of exceptions is } k\ddot{a}osz \text{ in the case of (29b) and entellekt\ddot{u}el \text{ ‘intellectual’, meni\ddot{e}tt, m\ddot{u}ezzin \text{ ‘muezzin’, habit\ddot{u}\ddot{e}, n\ddot{u}ansz, enteri\ddot{o}r \text{ ‘interior’, exteri\ddot{o}r \text{ ‘exterior’, mili\ddot{o} in the case of (29c).}\n\]

\[\text{\textsuperscript{69}Similarly, Zooey (from J. D. Salinger’s novel Franny and Zooey) is pronounced [zui] rather than *[zu:i] (note that this example obviously is not a spelling pronunciation).}\]
As there are no monomorphemic examples of vowel clusters consisting of more than two nuclei,\(^7^0\) in analysis (i) a separate constraint is needed to exclude \(*VVV:\)

(32) \[\begin{array}{c}
* N \\
\end{array} \begin{array}{c}
N \\
\end{array} \begin{array}{c}
N \\
\end{array} \]

(ii) It is notable, however, that the number of monomorphemic items actually surfacing with hiatus (i.e. unrepaired by hiatus filling) is low and that most of the relevant items are loans. The number of monomorphemic items with underlying hiatus found in the database used is \(n = 1311\). Most of these items, however, contain /i/ or /iː/ in hiatus (\(n = 1075\)) and surface with compulsory hiatus filling (cf. Siptári and Törkenczy 2000). Thus, there are only 236 monomorphemic items actually surfacing with hiatus. This may suggest an alternative analysis in which hiatus is ill-formed:

(33) \[\begin{array}{c}
* N \\
\end{array} \begin{array}{c}
N \\
\end{array} \]

In this analysis, (31ab) are accidental and not part of Hungarian phonology at all and (32) is redundant. Note that there is another way to express that hiatus is ill-formed: we could change the obligatory onset parameter (2) to ‘YES’. The ill-formedness of hiatus would follow from this parameter setting. Well-formed onsetless syllables would still exist, but they would be limited to initial position in an analytic domain: e.g. itt ‘here’, ár ‘price’, etc. This can be interpreted as an extension or generalization of the Peripherality Condition (Hayes 1980): exceptional syllable structure is permitted at the edge of a domain, not only in the sense that extra material can be added to the basic syllable template (peripheral extrasyllabic consonants may occur or peripheral consonants may be syllabified into a special constituent (the appendix)), but also in the sense that subminimal syllables may be licensed in peripheral position. Under analysis (ii) what remains to be explained is why violations of (33) are not

---

\(^7^0\)paranoia ‘id’ [parənojə] does not violate (30). dauer ‘perm’ ([døjer] in ECH) is the single exception, but in substandard Hungarian even this word is pronounced [døjjər].
One might suggest a third, sonority-based alternative to account for the predominance of hiatuses containing an /i/ melody. One could say that only vowel clusters containing a [-open₂] (high) melody are allowed (hiatuses permitted by this constraint may or may not be seen as subject to (31ab) so this analysis does not say anything about the status of (31ab)). Hiatus is otherwise disallowed. This can be given a sonority interpretation: as [-open₂] vowels are usually assumed to be less sonorous than [+ open₂] (nonhigh) vowels (e.g. Goldsmith 1990, Laver 1994) we can say that hiatus is only allowed in Hungarian if there is a sonority difference between the members of the vowel cluster (this presupposes that there is no sonority difference between [+ open₂] (low) vowels and [-open₂] (nonlow) vowels). It does not matter which of the two vowels is more sonorous: e.g. kiabál ‘shout’, mozaik ‘mosaic’ (note that even these hiatuses would be subject to postlexical Hiatus Filling [kijobal, mozojik]). The constraint then would be:

(i) Hiatus is asymmetric: one of the vowels must be governed.

Unfortunately this analysis makes wrong predictions about the well-formedness of clusters both of whose vowels are high. These would be ruled out by (i) because there is no sonority difference between the vowels. This is wrong since hiatuses in which both vowels are [-open₂] are well-represented within the set of hiatuses one of whose vowels has a [-open₂] specification: e.g. fiú, július, jezsuita, etc.
3.3.2. **Intervocalic consonant clusters**

In this section I discuss the constraints that apply to transsyllabic consonant clusters.

3.3.2.1. **Two-member clusters**

We have seen that any consonant can occur in a simplex onset or coda in Hungarian. Nevertheless, not all possible combinations of a nonbranching coda and a nonbranching onset can occur within a word: there are transsyllabic constraints specific to this context. Table VI shows the attested intervocalic two-member consonant clusters.\(^7^2\) Table VI abstracts away from allophonic differences (hence the lack of [ŋ, ű] for instance). The notation used is the usual one: a blank space in an intersection of a row and a column means that the relevant cluster is unattested; a star (*) in a box indicates that the relevant cluster only occurs when the two consonants are separated by an analytic morphological domain boundary; a cluster is spelt out if it is attested in monomorphemic items and the number of such items in the database is \(n > 15\); numbers have been used to indicate the number of monomorphemic items in the database when the cluster in question is attested in monomorphemic items and the number of such items is \(n \leq 15\). A box containing a spelt out cluster, a star or nothing is struck out by dashes to indicate that the relevant cluster(s) is/are subject to (eliminated on the surface by) assimilations.\(^7^3\)

---

\(^7^2\)Table VI is the result of a computer search in the database used (cf. Chapter 1) with some additions since the database does not contain inflected items.

\(^7^3\)None of these assimilations are related to syllable structure (i.e. operate regardless of the syllabification of the cluster to which they apply) and some of them are postlexical (Siptár & Törkenczy 2000).
### Table VI Intervocalic CC clusters

| p | t | tʰ | k | b | d | dʰ | g | tʰ | č | j | f | s | š | v | ž | m | n | nʰ | l | r | j | x |
|---|---|----|---|---|---|----|---|----|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| p | pp | pt | * | 10 | **** | 15 | 5 | - | 2 | ps | 8 | * | * | * | 3 | * | pl | pr | 2 | * | |
| t | * | tt | - | tk | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | tv | 6 | 3 | - | tl | tr | 5 |
| tʰ | 2 | * | tʰtʰ | 11 | * | - | 1 | * | 1 | - | 4 | * | 4 | 1 | * | 1 | |
| k | * | kt | kk | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| b | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| d | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| dʰ | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| g | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| tʰ | * | * | tʰk | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| č | * | * | čk | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| j | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| f | * | * | * | 8 | 3 | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| s | 15 | st | 2 | sk | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| š | šp | št | 6 | šk | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| v | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| ž | * | * | * | 4 | 4 | 2 | 4 | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * |
| m | mp | 2 | 1 | mb | 2 | * | * | * | 2 | 15 | 1 | 1 | 4 | 3 | 5 | 5 | mm | 8 | 3 | ml | 3 | 3 | 3 |
| n | * | nt | * | nk | * | nd | * | ng | ntʰ | nč | 9 | nf | ns | nš | nv | nz | 4 | * | nn | * | nl | * | * | 3 |
| nʰ | * | * | nᵗʰ | * | * | * | nᵈʰ | * | * | * | 1 | 1 | 6 | * | * | * | * | nⁿʰ | 1 | * | * | 6 |
| l | 9 | lt | 1 | lk | 6 | ld | 3 | lg | ltʰ | lč | If | 2 | lš | lv | 2 | 1 | lm | ln | * | ll | 9 |
| r | rp | rt | rtʰ | rk | rb | rd | 7 | rg | rtʰ | rč | rs | rs | rš | rv | rz | rž | rm | rn | * | rl | rr | rr | rx |
| j | 1 | jt | jk | 4 | jd | * | 7 | 2 | 4 | 2 | js | 1 | 4 | 9 | * | 6 | 9 | 4 | jl | 4 | 11 | 11 |
| x | * | * | * | 1 | * | * | * | * | * | * | * | * | * | * | * | * | * | * | * | 1 | 1 | * | * | 1 |

Examples (a question mark marks items in which the relevant cluster may contain an analytic domain boundary): *szappan* ‘soap’, *kapta* ‘(boot) last’, *lepe* ‘butterfly’, *kapca* ‘foot-cloth’,
Of course, some of these clusters are subject to assimilations, consequently, there are completely unattested clusters at the surface, e.g. */dp/* is subject to regressive voicing assimilation and thus [dp] is unattested: szabadpiac ‘free market’ [sobôtpijot].
The lack of branching onsets is not the only reason why the SCL does not hold in Hungarian: there are interconstituent clusters with rising sonority that would not be well-formed onsets even if branching onsets were permitted, e.g. /tl/ *katlan ‘cauldron’, /kn/ *szoknya ‘skirt’, /lm/ *lakmusz ‘litmus’, /zn/ *vēna ‘thin’, /nl/ *tőyleg ‘really’, etc.
an intervocalic cluster: systematic gaps do occur.

Let us examine Table VI and disregard clusters containing /x/ and /v/ for the moment (we shall discuss their behaviour separately at the end of this section). Then, it can be seen in the table that the greatest variety of monomorphemic intersyllabic clusters are of the type in which there is a *sonority difference* between the two consonants making up the cluster. In general, hardly any special restrictions (pertaining to place of articulation, for instance) apply to clusters of this type. Therefore, we assume that the primary source of licensing for interconstituent clusters is government:

(35) An interconstituent cluster whose member consonants are in a governing relationship (right-to-left or left-to-right) is well-formed.

As Table VI shows, however, (i) not all interconstituent clusters whose member consonants are equally sonorous are ill-formed; and (ii) not all interconstituent clusters whose member consonants have different sonority are well-formed. (i) suggests that government is not the only way in which interconstituent clusters can be licensed: geminates and some stop+ stop clusters, for instance, are well-formed and thus must be licensed by some other special means of licensing (note that (35) does not imply that clusters whose member consonants are not in a governing relationship are necessarily ill-formed). There are two types of interconstituent clusters that statement (ii) holds true of: clusters consisting of fricatives and stops (in either order) and nasal+ stop clusters. Let us disregard the latter type for the moment and focus our attention on the former: some intervocalic fricative+ stop and stop+ fricative clusters are not well formed. A possible way to handle this problem is to say that in Hungarian the sonority difference between fricatives and stops is not great enough for government to apply. Recall that the sonority distance settings for Hungarian consonants and the minimum sonority distance requirement are as follows (cf. section 3.2.4.2.):

---

76 Nasal+ obstruent clusters are an obvious counterexample, cf. the discussion below.
(36) a. *Sonority Hierarchy: Hungarian*

stops, affricates \(<\) fricatives \(<\) nasals \(<\) liquids

b. \(S_{\text{min}} = < < \text{ or } > >\)

Thus, no governing relationship can obtain between the consonants in fricative+ stop and stop+ fricative clusters in general, consequently special provisions must be made to license those clusters of this type that are well-formed. This would explain why—like other clusters whose members are equally sonorous (e.g. stop+ stop or nasal+ nasal clusters)—only some fricative+ stop and stop+ fricative clusters are well formed.

Let us now examine the non-analytic interconstituent clusters that are unlicensed by government. We have noted above that some of these clusters are well-formed. Three types of behaviour may be distinguished.

First, all intervocalic geminates are well-formed (the lack of monomorphemic /v:/ and /ž/ is an accidental gap). Geminates obviously cannot be licensed by government if we assume that government applies between timing slots (cf. 2.2) because the two adjacent timing slots have the same segmental content and thus are equally sonorous. Following Rice (1992) we assume that the licensing of geminates is due to the fact that they have shared structure (specifically, a single root node) and thus can be attributed to binding.\(^{77}\)

(37) *Interconstituent binding*

An interconstituent cluster C1C2 where C1 (the coda consonant) is root-bound is licensed.

The second type of well-formed clusters that seem unlicensed by government is liquid+ liquid combinations. All possible combinations of the three segments involved (/l, r, j/) seem to be permitted.\(^{78}\) Although all three are COR, their licensing cannot be attributed to binding (place-binding) because (a) /j/ is [-anterior] and thus does not share its place node (or even COR

---

\(^{77}\) Compare Kaye, Lowenstamm and Vergnaud (1990) who assume that the first slot in a geminate is empty and governed by the second.

\(^{78}\) The clusters /lr/ and /lj/ are subject to assimilations that are unrelated to syllable structure.
node) with /l, r/ and (b) there are ill-formed interconstituent clusters whose member consonants share a place node (e.g. */fp, st/, etc.). The licensing of liquid+ liquid combinations ceases to be a problem if we assume the fine-tuned sonority scale introduced in section 3.2.3:

\[(38) \quad \text{\textit{l < < r < < j}}\]

Given (38), the licensing of liquid+ liquid interconstituent clusters can be simply attributed to government.

The third group of well-formed interconstituent clusters not licensed by government consists of some nasal+ nasal clusters and some obstruent+ obstruent clusters. This group differs from the previous two in that only some of these clusters are well-formed.

An examination of Table VI shows that in (nongeminate) stop+ stop interconstituent clusters the second consonant must not be labial. Clusters whose second consonant is labial are ill-formed/unattested whereas those ending in coronals or velars are well-formed. The working of this ‘antilabial constraint’ can be seen in the following examples:

\[(39) \quad \text{stop + stop}\]

\[
\begin{align*}
\text{pt} & \quad *\text{tp} \quad \text{kaptcha ‘(shoemaker’s) last’} \\
\text{pk} & \quad *\text{kp} \quad \text{lepke ‘butterfly’} \\
\text{bd} & \quad *\text{db} \quad \text{labda ‘ball’} \\
\text{kt} & \quad = \quad \text{tk} \quad \text{akta ‘document’, atka ‘mite’}
\end{align*}
\]

The same constraint seems to hold in affricate+ stop, nasal+ nasal and stop+ fricative clusters. Consider the following examples:

\[(40) \quad \text{a. affricate + stop}\]

\[
\begin{align*}
\text{čk} & \quad *\text{čp} \quad \text{kecske ‘goat’} \\
\text{t'k} & \quad *\text{t'p} \quad \text{lecke ‘homework’}
\end{align*}
\]

\[79\]Stop+ affricate clusters are omitted because there are no labial affricates in Hungarian.
b. **nasal + nasal**

\[ \text{mn} \quad *\text{nm} \quad \text{himnusz ‘hymn’} \]

c. **stop + fricative**

\[ \text{pš} \quad *\text{pf} \quad \text{tapsi ‘bunny rabbit’, †cupfol}^{30} \]

\[ \text{kš} \quad *\text{kf} \quad \text{kuksol ‘hide, cower’, †bukfenc} \]

\[ \text{ps} \quad *\text{pf} \quad \text{apszis ‘apse’, †cupfol ‘pluck’} \]

\[ \text{ks} \quad *\text{kf} \quad \text{taxi ‘id.’, †bukfenc ‘somersault’} \]

There are a very small number of exceptions to the ‘antilabial’ constraint in the types of combinations examined. The complete list of the (monomorphemic) exceptional items (stems) I have found is as follows: /tʰp/ *pitypang ‘dandelion’, *pitypalatty ‘quail’s song’; /gβ/ *rogbi ‘rugby’; /gβʃ/ *bugyborǻk ‘bubble’, *lögyböl ‘handwash in water without rubbing’ < not in ECH>; /pf/ *cupfol ‘pluck’, *copfoss ‘pigtailed’; /tʃ/ *platform ‘stand’; /tʃʃ/ *fityfirítty ‘imp’; /kf/ *bakfis ‘young girl’, *bikfic ‘kid’, *bukfenc ‘somersault’, *pakfson ‘German nickel-silver’, *ukmukfük ‘in a jiffy’.

Fricative+ stop clusters behave in a more complex way: the ‘antilabial’ constraint does not work when the first consonant is coronal (41b), but it does if it is non-coronal (41a):

(41) **fricative + stop**

a. \[ \text{ft} \quad \text{fk} \quad *\text{fp} \quad \text{afta ‘thrush’, cafka ‘whore’} \]

b. \[ \text{sk} = \text{sp} \quad \text{viszket ‘itch’} = \text{aszpik ‘jelly’} \]

\[ \text{st} = \text{sp} \quad \text{posztó ‘felt’} = \text{aszpik ‘jelly’} \]

\[ \text{št} = \text{šp} \quad \text{este ‘evening’} = \text{püspök ‘bishop’} \]

\[ \text{šk} = \text{šp} \quad \text{eskü ‘oath’} = \text{püspök ‘bishop’} \]

\[ \text{zd} = \text{zb} \quad \text{gazda ‘master’} = \text{azbeszt ‘asbestos’} \]

We can account for these regularities if we assume that a special kind of licensing (call it Sp-}

---

\(^{30}\)The symbol † marks attested, but phonotactically ill-formed items.
licensing) is needed in order for an interconstituent cluster to be well-formed if it is not licensed by government or binding. A given language may or may not allow Sp-licensing to apply. We assume that in Hungarian, in general, Sp-licensing is granted to interconstituent clusters, i.e. it can license coda-onset clusters that are unlicensed by government or binding. Note that this type of licensing is not derivable from government if government is solely based on sonority; nor does it derive from binding. Sp-licensing is thus stipulative and its conditions are language specific.

\[(42) \quad \text{Sp-licensing} \]
\[\text{Sp-licensed interconstituent clusters are well-formed.}\]

Hungarian, however, imposes certain constraints on Sp-licensing, i.e. it disallows Sp-licensing in some interconstituent configurations. These constraints are discussed and formalised below:

The ‘antilabial’ effects are due to the following condition on Sp-licensing:

\[(43) \quad \text{In an interconstituent cluster } C_1C_2, \text{ LAB consonants cannot Sp-license the preceding consonant. Condition: } C_1 \neq [\text{COR, } + \text{ cont}] \]

As is expected, no antilabial effects can be detected when an interconstituent cluster is licensed by government or root-binding (note that in principle the clusters in question could display antilabial effects because there are labial consonants in the various manner classes that appear in the second position). Consider the following examples:

---

\[81\text{It could not be interpreted as binding even if we assumed that coronals are placeless because there are well-formed and Sp-licensed interconstituent clusters not containing a coronal: e.g. /pk/ as in lepke ‘butterfly’. /pk/ is not licensed by government (there is no sonority difference between the segments), and cannot be licensed by binding since /p/ has its own independent place specification.}\]

\[82\text{Hopefully, further research will be able to derive (some of) the effects due to Sp-licensing from general principles.}\]
C+ liquid and C+ affricate clusters are disregarded because there are no labial liquids or affricates.

(44)  

a. stop + nasal

etnikum ‘ethnic group’ = ritmus ‘rhythm’
bodnár ‘cooper’ = ködmön ‘sheepskin coat’

b. nasal + stop

cinke ‘titmouse’ = lámpa ‘lamp’
fondorlat ‘devious trick’ = bomba ‘bomb’

c. liquid + stop

boldog ‘happy’ = kolbász ‘sausage’
árkád ‘arcade’ = Árpád < name>

d. affricate + nasal

fecni ‘slip of paper’ = kecmereg ‘crawl’
kalucsni ‘galosh’ = pacsmag ‘suspicious concoction’

e. fricative + nasal

disznó ‘pig’ = pászma ‘ray of light’
vézna ‘thin’ = zuzmó ‘lichen’

f. nasal + fricative

unszol ‘urge’ = tánfereg ‘loiter’
emse ‘sow’ = kámfor ‘camphor’

(g. liquid + fricative

válság ‘crisis’ = delfin ‘dolphin’
persze ‘of course’ = férfi ‘man’

83C+ liquid and C+ affricate clusters are disregarded because there are no labial liquids or affricates.
94

h. liquid + nasal

málna ‘raspberry’ = elme ‘mind’
barna ‘brown’ = lárma ‘noise’

In addition to the interconstituent clusters that are licensed by government or root-binding there is another group of clusters which could in principle display an ‘antilabial’ effect, but do not: all clusters consisting of fricatives and/or affricates appear to be ill-formed. There are very few words containing non-analytic fricative/affricate+ fricative/affricate clusters. The following is an exhaustive list of the exceptional/irregular items:

/sť/ disziplinn ‘subject’, proszciónium ‘fore-stage’, reminiscencia ‘memory’, /fs/ ofszájd ‘off-side’, /šť/ aszfált ‘asphalt’, atmoszfára ‘atmosphere’, blaszfémia ‘blasphemy’, foszfát ‘phosphate’, foszfor ‘phosphore’, /šľ/ násfa ‘pendant’. Assuming that affricates are contour segments that contain the feature [+ continuant], the ‘antifricative’ constraint can be interpreted as a ban on the occurrence of the features [-son, + cont] under both root nodes in an interconstituent cluster. Since, trivially, the constraint only holds if the cluster is not licensed by government or binding, it can be built into Sp-licensing:

(45) A [-son, + cont] segment cannot Sp-license another [-son, + cont] segment in an interconstituent cluster.

This ‘antifricative’ constraint can mask the effect of the ‘antilabial’ constraint (cf. Törkenczy 1998)—this is why fricative/affricate+ fricative/affricate clusters do not display an ‘antilabial’ effect.

A further constraint can be identified if we examine the clusters that contain palatals. It seems that /t’, d’, n’/ make an interconstituent cluster ill-formed irrespective of whether they occur in the first or the second position if the two consonants are not in a governing relation. Consider the following examples (in 46 below the first column of clusters has clusters whose

---

84 We have disregarded triliteral clusters (there are four stems containing them in the database: o[p]szcán ‘obscene’, excentrikus ‘eccentric’, transzcendens ‘transcendent’, excellenciás ‘excellency’ (see Törkenczy 1994a) ) and clusters containing /v/. Cf. the discussion below.
members are in a governing relationship and the second shows clusters whose members are not):

(46)  
\[
\begin{align*}
\text{d'm} & \quad *\text{d'd} & \quad \text{hagyma 'onion'} \\
\text{n'd'y} & \quad *\text{n'n} & \quad \text{ke[n]gyel 'stirrup'} \\
\text{n'n} & \quad *\text{n'n} & \quad \text{ponyva 'canvas'} \\
\text{rt'y} & \quad *\text{št} & \quad \text{kártya 'card'} & \quad (\text{†ostya 'wafer'}) \\
\text{rd'y} & \quad *\text{žd} & \quad \text{bárgyú 'stupid'} & \quad (\text{†mezsgye 'border'}) \\
\text{rn'y} & \quad *\text{mn} & \quad \text{ernyó 'umbrella'} & \quad (\text{†nyimnyám 'weakling'})
\end{align*}
\]

The palatal liquid /j/ is unlike /t', d', n'/ in that it forms well-formed interconstituent clusters with any consonant irrespective of whether it occurs in the first place or the second place. This is to be expected, given that /j/ is at least minimally sonority-distinct from all the other sonority classes. Thus, all the examples below are well-formed:

(47)  
\[
\begin{align*}
\text{/j/} & \quad \text{se[j]pít 'lisp'} & \quad \text{gyapjú 'wool'} \\
& \quad \text{hajcsár 'drover'} & \\
& \quad \text{májszol 'munch'} & \quad \text{ifjú 'youth'} \\
& \quad \text{hájnal 'dawn'} & \quad \text{tőmjén 'incense'} \\
& \quad \text{kajla 'scatterbrained'} & \quad \text{varjú 'crow'} \\
& \quad \text{ká[j]ha 'stove'} & \\
\end{align*}
\]

Again, the 'antipalatal’ effects can be seen as a result of a constraint on Sp-licensing:

(48)  
\[
[\text{COR, } -\text{ant}] \text{ consonants cannot Sp-license another consonant in an inter constituent cluster.}
\]

There are few exceptions to (48). The following is an exhaustive list of occurring ill-formed items: /t'p/ \text{pity pang} ‘dandelion’, \text{pity palatty} ‘quail’s song’; /t'k/ \text{butykos} ‘bottle’, \text{bütykös} ‘knobbly’, \text{bütyköl} ‘repair’, \text{fiitykös} ‘stick’, \text{hetyke} ‘proud’, \text{lötyköl} ‘quickly wash’, \text{pityke}
The ranking of [h] (the realization of /x/ in the onset in Hungarian) in the sonority hierarchy is problematic. It is usually ignored in discussions of the sonority relations between segment classes (cf. for instance, Laver 1994, Ladefoged 1993, Steriade 1982, van der Hulst 1984, Anderson and Ewen 1987) or only mentioned in passing (e.g. Clements 1990 observes that ‘the sonority ranking of voiceless approximants is not well-established’ (p. 293)—Levin 1985 points out that [h] and [?] may function as obstruents (p. 65)). In the lack of (counter) evidence we simply stipulate that /x/ has the same sonority rank as a fricative.

In the discussion of interconstituent clusters so far we have disregarded those containing /x/ or /v/. As can be seen in Table VI, /x/ is free to occur as the second consonant, but is rare as the first consonant in an interconstituent cluster (note that it does occur in this position in a few words, e.g. ihlet ‘inspiration’). This distribution is not due to an interconstituent constraint. We assume that C+/x/ and /x/+ C clusters are licensed by government (i.e. that the sonority difference between /x/ and other sonority classes is sufficient for government to apply) and the scarcity of preconsonantal /x/ (regardless of whether the coda is part of an interconstituent cluster or not) is accidental. Note that even if the distribution were due to a constraint, it would be relevant to the coda position alone rather than the interconstituent domain.

The behaviour of /v/ is less straightforward. We shall see that it is just as ‘two-faced’ in its phonotactic behaviour in this position as it is with respect to voicing assimilation (cf. Vago 1980, Siptár & Törkenczy 2000, Siptár 1996). In order to see this let us examine what kind of behaviour we predict with respect to the interconstituent constraints discussed above (i) if /v/ is an obstruent (and has the sonority ranking of a fricative); and (ii) if /v/ has the sonority ranking of a non-nasal sonorant.

---

85The ranking of [h] (the realization of /x/ in the onset in Hungarian) in the sonority hierarchy is problematic. It is usually ignored in discussions of the sonority relations between segment classes (cf. for instance, Laver 1994, Ladefoged 1993, Steriade 1982, van der Hulst 1984, Anderson and Ewen 1987) or only mentioned in passing (e.g. Clements 1990 observes that ‘the sonority ranking of voiceless approximants is not well-established’ (p. 293)—Levin 1985 points out that [h] and [?] may function as obstruents (p. 65)). In the lack of (counter) evidence we simply stipulate that /x/ has the same sonority rank as a fricative.
First let us suppose that /v/ has the sonority ranking of a fricative. (49ab) show the predictions about the well-formedness of vC and Cv clusters respectively. Stars mark ill-formed clusters and ✓ marks well-formed ones. */✓ appears if some of the clusters within the class are predicted to be well-formed others are not.

(49) a. vC

<table>
<thead>
<tr>
<th>stop</th>
<th>affricate</th>
<th>fricative</th>
<th>nasal</th>
<th>liquid</th>
</tr>
</thead>
<tbody>
<tr>
<td>/v/</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

b. Cv

<table>
<thead>
<tr>
<th>stop</th>
<th>affricate</th>
<th>fricative</th>
<th>nasal</th>
<th>liquid</th>
</tr>
</thead>
<tbody>
<tr>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

If /v/ is a LAB fricative, then both Cv and vC interconstituent clusters are predicted to show antilabial effects. /v/+ LAB stop and stop+ /v/ clusters are expected to be ill-formed because they are not licensed by government (the sonority distance between stops and fricatives is too small) and Sp-licensing cannot apply since the second member of the interconstituent cluster is LAB. This prediction is only partly borne out: although /v/+ labial stop clusters do not occur, stop+ /v/ clusters are well-formed: e.g. udvar ‘courtyard’, rögvext ‘at once’, fegyver ‘weapon’, lekvár ‘jam’, borotva ‘razor’, kotyvaszt ‘concoct’ (cf. Table VI).

If /v/ is a fricative, then in accordance with (45), both vC and Cv clusters should display ‘antifricative’ effects. This is true of /v/+ fricative/affricate clusters, but (similarly to stop+ /v/ clusters) fricative+ /v/ clusters are well-formed: e.g. ösvény ‘path’, özvegy ‘widow(er)’, öszvér ‘mule’.

If /v/ is a fricative, then obstruent+ /v/ and /v/+ obstruent clusters are expected to show ‘antipalatal’ effects since government cannot license the relevant clusters and according to (48) Sp-licensing cannot apply. This again is only partly true since contrary to the prediction—palatal obstruent+ /v/ clusters are well-formed (e.g. fegyver ‘weapon’, kotyvaszt ‘concoct’)

If /v/ is a fricative, both /v/+ sonorant and sonorant+ /v/ clusters are predicted to be well-formed because these clusters are licensed by government This prediction is borne out.
To sum up, a fricative interpretation of /v/ makes correct predictions about the well-formedness of interconstituent clusters containing /v/ if (i) the other consonant in the interconstituent cluster is a sonorant and (ii) if /v/ occurs as C1 in an interconstituent cluster C1C2.

Let us now examine what predictions are made and whether they are borne out if /v/ is interpreted as a non-nasal sonorant. Let us assume that the sonority distance between /v/ and the other non-nasal sonorants is great enough for government to apply (i.e. ‘/v/ < liquids’). (50ab) show that under this interpretation all interconstituent clusters containing /v/ (vC and Cv alike) are predicted to be well-formed. The reason is that if /v/ has the sonority ranking of a non-nasal sonorant, then government would license all the clusters shown in (50), consequently binding and Sp-licensing would have no effect.

(50) a. vC

<table>
<thead>
<tr>
<th>stop</th>
<th>affricate</th>
<th>fricative</th>
<th>nasal</th>
<th>liquid</th>
</tr>
</thead>
<tbody>
<tr>
<td>/v/</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

b. Cv

<table>
<thead>
<tr>
<th>stop</th>
<th>affricate</th>
<th>fricative</th>
<th>nasal</th>
<th>liquid</th>
</tr>
</thead>
<tbody>
<tr>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

This prediction is not correct, however, since (as we have seen above) /v/+ fricative-affricate, /v/+ labial stop\(^{86}\) and /v/+ palatal obstruent clusters are ill-formed. This suggests that the sonorant interpretation of /v/ makes correct predictions if (i) /v/ occurs as C2 in an interconstituent cluster C1C2; and/or (ii) the other consonant in the cluster is a sonorant.

Thus, we are faced with a ‘sonority ranking paradox’: /v/ behaves as an obstruent when it occurs as the first member of an interconstituent cluster, but it behaves as a sonorant when it is the second member of an interconstituent cluster.\(^{87}\) The question is how to express this in

---

\(^{86}\)/v/+ stop clusters do not occur even if the stop is not LAB. In our interpretation this is accidental.

\(^{87}\)This is completely in agreement with the ambiguous nature of /v/ (cf. Vago 1980, Siptár & Tőrkcenczy 2000, Siptár 1996). It should be noted, however, that /v/+ liquid clusters are
terms of licensing. First of all, it is not possible for the same segment to have different sonority rankings depending on the position it occurs in, and we do not want to postulate two different underlying /v/-s (a sonorant and an obstruent). Furthermore, it does not help to assume that /v/ is ‘asymmetrical’ in the sense that although it is different in terms of sonority from both obstruents and sonorants—i.e. it is ‘closer’ to obstruents than to sonorants (obstruents < /v/ < sonorants), because this would still incorrectly predict antipalatal, antifricative and antilabial effects in intervocalic obstruent+/v/ clusters. The reverse, i.e. that it is closer to sonorants than to obstruents (obstruents < < /v/ < sonorants) does not help either, because it would remove /v/+ obstruent clusters from the purview of the constraints on Sp-licensing and, incorrectly, no antipalatal, antifricative and antilabial effects would be predicted. Thus, there seem to be two options: we can assume that (i) /v/ has the sonority ranking of a sonorant that is minimally sonority-distant from both the obstruents and the other sonorants (obstruents < < /v/ < sonorants) and stipulate that /v/ has to be Sp-licensed when it occurs in a coda which is part of an interconstituent cluster (even if it is licensed by government); or, alternatively, (ii) /v/ has the sonority ranking of a fricative, but is stipulated to be exempt from the restrictions on Sp-licensing in an onset which is part of an interconstituent cluster. The two solutions are equivalent in that both of them are stipulative. However, since the distribution of /v/ in branching codas suggests that it has the sonority ranking of a fricative (cf. section 3.2.4), we choose the latter solution and propose the following constraint:

(51) /v/ is Sp-licensed in an onset in an interconstituent cluster

(51) has the desired effect because while /v/ as C1 in an interconstituent cluster C1C2 remains rare and /v/+ nasal clusters do not occur. We consider this accidental.

The two underlying segments would be in complementary distribution, and other phonological processes involving [v] do not require such an analysis; see (cf. Siptár & Tőrkenczy 2000).

One might want to build (51) into the constraints on Sp-licensing by exempting /v/ from each of the relevant constraints. That, however, would unnecessarily complicate the constraints while leaving the analysis no less stipulative. It would be more interesting to derive the effect of (51) from the uniqueness of the representation of /v/, the fact that it is unspecified for [son]. We leave this problem for future research.
subject to the ‘antilabial’ and the ‘antipalatal’ licensing constraints, it is permitted to occur freely (i.e. unconstrained by these constraints) when it is C2 because it is licensed by (51) in that position. In its present form, (45) cannot prevent /v/+ fricative clusters from being Sp-licensed because /v/ is unspecified for [son] (i.e. it does not have [-son] feature that (45) crucially refers to). Thus, incorrectly, no ‘antifricative’ effects are predicted. This can be remedied by a minimal modification of (45):

(52) [-son, + cont] segments cannot Sp-license [+ cont] segments in an interconstituent cluster.

Now (52) can revoke Sp-licensing and /v/+ fricative clusters are correctly judged to be unlicensed. Note that the modification has no adverse effect—(52) still prevents interconstituent clusters consisting of fricatives and/or affricates (in any order or combination) from being Sp-licensed. The only difference is that (52) does not allow fricatives to Sp-license [+ cont] sonorants. The relevant clusters are well-formed (see Table VI), but they are licensed by government anyway and Sp-licensing is not necessary. Thus the change makes no difference here.

Finally, certain interconstituent clusters are ill-formed in spite of the fact that they appear to be licensed by the constraints discussed above. Specifically, non-homorganic nasal+ stop clusters are disallowed although they are licensed by government. This suggests that the licensing of nasals is subject to the following restriction:

(53) Coda nasals must be place-bound when followed by stops.

Note that (53) is not specifically an interconstituent constraint (see section 3.2.4) and that it also holds true of affricates (contour segments whose left ‘face’ is a stop). There are few exceptions to (53). What follows is a complete list of the exceptional items: /mt/ tamtam ‘tomtom’, nyomtar? ‘print’, /mk/ tömkeleg ‘abundance’, /md/ dumdum ‘id’, dínomdánom
‘merry-making’, /mɛ/ csámcseg ‘eat noisily’, csemcseg ‘eat noisily’.\(^{90}\)

3.3.2.2. Clusters consisting of more than two members

We have pointed out earlier that intervocalic clusters consisting of more than two consonants are irregular unless an analytic boundary breaks up the cluster.\(^ {91}\) There are such irregular items, but their number is relatively low. In the database there is just one item containing a five-member medial cluster ([ŋkšt] angström ‘id’) and there are only 23 monomorphemic stems with a four-member medial cluster. The following monomorphemic four-member clusters occur (the numbers in angled brackets indicate the number of monomorphemic items a given cluster occurs in):

\[^{90}\]These items may not be exceptional at all in that they probably contain internal domain boundaries that fall between the nasal and the stop. They are included here for the sake of completeness.

\[^{91}\]For the arguments see the discussion in section 3.2.2.
There are about 288 monomorphemic stems containing a three-member cluster in our database (e.g. *centrum* ‘centre’, *komplex* ‘complex’, *export* ‘id.’, *improvizál* ‘improvise’). There are 91 kinds of clusters in these items, but, typically, the functional load of a given cluster is extremely low: each type is only ‘utilized’ in a handful of morphemes. The following monomorphemic three-member clusters occur. The numbers in angled brackets indicate the number of monomorphemic stems a given cluster occurs in and in the case of [x] and [õ] noncontrastive differences are indicated. See Appendix B (i) for a complete list of the relevant stems.

\[\text{(54)}\]

\begin{tabular}{ll}
\hline
\text{pštr} & <1> \quad \text{obstruál} ‘obstruct’ \\
\text{pstr} & <1> \quad \text{absztrakt} ‘abstract’ \\
\text{jšr} & <1> \quad \text{lajstrom} ‘list’ \\
\text{kskl} & <1> \quad \text{exklúzi v} ‘exclusive’ \\
\text{kskr} & <1> \quad \text{krikszkraksz} ‘hardly legible script’ \\
\text{kspl} & <2> \quad \text{explicit} ‘explicit’, \text{explozi v} ‘explosive (adjective)’ \\
\text{kspr} & <1> \quad \text{expressz} ‘express’ \\
\text{kstr} & <4> \quad \text{dextrin} ‘id.’, \text{extra} ‘id.’, \text{extrán} ‘extreme’, \text{foxtrott} ‘id.’ \\
\text{nkst} & <1> \quad \text{gengszter} ‘gangster’ \\
\text{nštr} & <7> \quad \text{instruál} ‘instruct’, \text{demonstrál} ‘demonstrate’, \text{konstruál} ‘create’, \text{konstrukti v} ‘positive’, \text{menstruál} ‘menstruate’, \text{monstre} ‘grand’, \text{monstrum} ‘oversized creature/object’ \\
\text{nskr} & <1> \quad \text{szanszkrit} ‘Sanskrit’ \\
\text{rštl} & <1> \quad \text{vurstli}\textsuperscript{92} ‘amusement park’ \\
\hline
\end{tabular}

\textsuperscript{92} Usually with a three-member cluster in ECH: /ršl/ \textit{vursli}
The fact that these clusters are irregular does not mean that they do not display certain regularities. Figure (56) informally summarizes some of them:
In our analysis these regularities (and other possible ones crucially referring to medial -CCC-) are at least partly accidental in Hungarian. They are accidental inasmuch as they reflect a random set of the regularities of the source languages the relevant words were borrowed from, and may be systematic to the extent they result from (possibly universal) non-syllable-based restrictions (sequence constraints).

\(^{93}\)Notation: \(x>y\) ‘\(x\) is more sonorous than \(y\)’; \(x<y\) ‘\(x\) is less sonorous than \(y\)’; \(x=y\) ‘\(x\) and \(y\) are equally sonorous’.
3.4. Morpheme structure: MSCs

In this section we discuss phonotactic constraints that hold within the morpheme. These constraints may or may not be related to syllable structure.

3.4.1. Domain-final open syllables and the minimal word/stem

In section 3.2.3. above we pointed out that in general any of the underlying vowels can occur in nuclear position in a syllable. This is not true of open syllables in final position, or more precisely of open syllables at the right edge of a stem. In this position (underlyingly) high ([−open$_1$]), mid ([−open$_1$, + open$_2$]) and low ([+ open$_1$]) vowels behave differently. The restrictions are the following.

Short [−open$_1$, + open$_2$] vowels (/o, ö/) cannot occur in final position.

This constraint is usually assumed to hold in word-final position (cf. Nádasdy 1985, Nádasdy and Siptár 1994, Törkenczy 1994). However, it is really a constraint on the stem, because /o, ö/ cannot occur at the right edge of an internal analytic domain or immediately before a non-analytic suffix either.94

Assuming that unaffixed free morphemes are stems, and that such a stem plus an affix is also a stem, this constraint can be expressed as

94The interjection (no)no ‘well’ is a counterexample. Interjections in general do not seem to conform to the phonological constraints of the language (e.g. even syllabic consonants may occur in interjections, which otherwise are unattested in Hungarian: [pst] pszt! ‘hush!’). Note also the behaviour of foreign compounds, see 3.3.1.
This constraint holds regardless of the number of syllables the stem consists of: monosyllables and polysyllables behave in the same way. (58) is one of the few phonotactic constraints that has an active role in the phonology of Hungarian. The final vowels of loans ending in /o, ö/ are invariably lengthened in Hungarian (e.g. libretto [libretto]).

The behaviour of domain-final [-open₂] vowels is more complex. (59) shows the distribution of [-open₂] vowels at the end of an analytic domain in monosyllabic and polysyllabic words. As can be seen below, three types of items can be distinguished: words in the first column (marked VV) are always pronounced with a long final vowel and those in the third (marked V) invariably have a short final vowel. By contrast, the words in the column marked VV/V may have either long or short final vowels. All native speakers of Standard Hungarian agree in their treatment of the words in columns VV and V, but they may treat those in column VV/V in three different ways. Innovative speakers of ECH have short final vowels in these words (for them there is no difference between the words in VV/V and those in V). Conservative ECH speakers pronounce them with a long final vowel (i.e. for them there is a lexical difference between the words in VV/V and those in VV). For a third group of speakers (we shall call this group ‘intermediate’ ECH speakers) the final vowels in the words in column VV/V may be optionally long or short.

---

Note also that there are no cseh-type stems with an /o/ or /ö/ in the stem-final syllable (cf. section 4.2.2), which suggests that (58) is a constraint that applies at the surface.
There is another fact that (59) does not indicate: word-final long /i/ is extremely rare. There are only 9 such items in our database and with the exception of rí ‘cry’, sí, vizaví and zrí ‘trouble’ they are interjections or onomatopoeic words. We have no explanation for this depleted distribution and consider it an accident.
Domain final [-open₂] vowels are

i. long in monosyllables and short in polysyllables (Innovative ECH)
ii. long in monosyllables (Conservative/intermediate ECH)

As in (58), the domain in which (60) applies is the stem because [-open₂] vowels behave in the same way word finally, at the end of the non-final constituent of a compound, immediately preceding an analytic suffix and immediately preceding a non-analytic suffix. (61) shows this in the innovative ECH dialect (where final [-open₂] vowels are always short in polysyllables):

(61)

<table>
<thead>
<tr>
<th>word-final compound</th>
<th>analytic suffix</th>
<th>non-analytic suffix</th>
</tr>
</thead>
<tbody>
<tr>
<td>monosyllabic</td>
<td>b[u:]</td>
<td>b[u:]-nak</td>
</tr>
<tr>
<td></td>
<td>mďa-b[u:]</td>
<td></td>
</tr>
<tr>
<td>polysyllabic</td>
<td>ágy[u]</td>
<td>ágy[u]-talp</td>
</tr>
</tbody>
</table>

In accordance with the definition of the stem above, suffix-final [-open₂] vowels behave in the same way as [-open₂] vowels at the end of lexical stems: compare tetű [tetű] ‘louse’ and jőszí v-ű [jő:svű] ‘kind-hearted’.

In all the three dialects, monosyllabic words/stems are treated in the same way (i.e. they have long final [-open₂] vowels), thus the variation that distinguishes these dialects is confined to polysyllabic words/stems. This fact can be accounted for if we assume that there is a constraint that applies to all dialects and requires that the minimal word/stem should be bimoraic:

(62) \( \text{Stem/word}_{\text{min}} = \mu \mu \)

---


99 On the mora, cf. section 2.2.
We assume that all stems (affixed and unaffixed) in the lexicon have to conform to (62).\textsuperscript{100} It does not apply to affixes, function words, interjections and onomatopoeic words.\textsuperscript{101} The minimal stem/word constraint is trivially (vacuously) true of stems ending in [-open\textsubscript{1}, + open\textsubscript{2}] (mid) vowels: as stem-final short mid vowels are excluded in general, stems consisting of open monosyllables that end in a mid vowel cannot violate (62). As pointed out above, stems ending in [-open\textsubscript{2}] (high) vowels conform to (62). In innovative ECH the distribution of high vowels can be interpreted as the result of a constraint that bans stem-final long high vowels which is blocked if it should violate (62) (we shall return to the formalization of this constraint later).

Let us now examine the behaviour of [+ open\textsubscript{1}] (low) vowels in final position. Stems ending in low vowels also observe (62). There are a (small) number of function words and interjections (be ‘into’, de ‘but’, he ‘what? < I cannot hear you> ’, le ‘down’, ne ‘no(t)’, se ‘either’, te ‘you’, ha ‘if’, ja ‘Now I understand’, na ‘Come on!’) and two truly exceptional content words (fa ‘tree’, ma ‘today’) which violate it. Although low vowels are also subject to the minimal word/stem constraint (like high and mid vowels), their distribution in final position is different in several ways. Disregarding the exceptional words listed above, they pattern in the following way:

<table>
<thead>
<tr>
<th></th>
<th>stem-final</th>
<th>word-final</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>monosyllabic</td>
<td>polysyllabic</td>
</tr>
<tr>
<td>e, a</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>e:, a</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

(63) shows that (i) – unlike mid vowels—both long and short low vowels occur finally (e.g. csokoládé ‘chocolate’, teve ‘camel’, burzsoá ‘bourgeois’, apa ‘father’); (ii) – unlike high vowels in innovative ECH—both long and short low vowels occur finally in polysyllabic words: thus, \textit{modulo} the minimal word/stem constraint, the distribution of long and short low vowels is the same in word-final position; and (iii) – unlike high and mid vowels—low vowels behave

\textsuperscript{100}On the minimal word constraint in Hungarian, cf. Csúri (1990) and Törkenczy (1994a).

\textsuperscript{101}The full list of exceptions to (62) ending in non-low (i.e. [-open\textsubscript{1}]) vowels is \textit{ki, ki, mi, mi, ni, no, ti}, all of which are function words or interjections.
Some suffixes are exceptional in that they may be preceded by a and e, e.g. -ság/ség as in apa-ság ‘fatherhood’. For a discussion of low vowel alternations, see Vago (1980), Nádasdy & Siptár (1994, 1998), Siptár & Törkenczy (2000), Siptár (2002).

Word-final á is rare and final é is relatively infrequent. Not counting words containing the suffixes -vá/-vé, -ná/né, final á occurs in function words, abbreviations and interjections. Final é fares relatively better: disregarding suffixes that end in é, it occurs in about 130 stems all of which are loan words.

The question is how to make sense of this distribution. This problem is related to the analysis of stem-final é - e, á - a alternations. As pointed out above, of the [+ open₁] vowels only é and á can occur before suffixes. Let us assume that the underlying difference between é - e, á - a is only quantitative (cf. Vago 1980, Siptár & Törkenczy 2000, etc). Then, in an analysis in which all long vowels are underlyingly (pre)associated with two timing slots and short ones with a single timing slot, in principle, these alternations may be interpreted in two ways: as (i) the lengthening of underlying final short [+ open₁] vowels before suffixes; or (ii) the shortening of underlying final [+ open₁] vowels word-finally. The choice determines the underlying distribution of final é and á. If analysis (i) is chosen, the distribution of these vowels in final position is unconstrained by any restriction specific to [+ open₁] vowels: both long and short [+ open₁] vowels may occur in all final environments underlyingly. Under this analysis the lack/low number of tokens with final low vowels in some environments in (63) is either due to the minimal word/stem constraint, which is independent of [+ open₁] vowels (and rules out stems that are too short – this would apply to final short [+ open₁] vowels in monosyllables (fa and ma are irregular)), or is accidental (thus, the relative infrequency of final é and á need not be accounted for in the phonology). The surface lack of stem-final e and a before suffixes is due to a phonological rule that lengthens short low vowels stem-finally if

---

102 Some suffixes are exceptional in that they may be preceded by a and e, e.g. -ság/ség as in apa-ság ‘fatherhood’. For a discussion of low vowel alternations, see Vago (1980), Nádasdy & Siptár (1994, 1998), Siptár & Törkenczy (2000), Siptár (2002).

103 There are about 15 items; burzsoá ‘bourgeois’ and hajrá ‘spurt’ are the only content words with final á.

104 The same is true of a linear analysis where length is a feature [±long].
a suffix follows (‘Low Vowel Lengthening’, cf. Vago 1980a, Nádasdy and Siptár 1994, Siptár & Törkenczy 2000; compare Rebrus 2000a). Under this analysis the underlying distribution of [+ open₁] (i.e. low) vowels is analogous with that of [-open₂] (i.e. high) vowels in conservative ECH.

Under the shortening analysis only long low vowels occur stem (and word) finally at the underlying level, and the word-final short surface reflexes are derived by rule. According to this analysis there is an underlying constraint that is specific to [+ open₁] vowels: the short ones cannot occur finally. The minimal stem/word constraint would hold without exception. *fa* and *ma* conform to it *underlyingly*. Note, however, that (exceptionally?) the word-final low vowel shortening would not be blocked by the minimal word constraint (*fa* and *ma* surface with a short vowel) while it does seem to apply to high vowels in innovative ECH.

In this analysis the underlying distribution of [+ open₁] vowels is analogous with that of mid vowels. The problem with this analysis is a derivational one: how to prevent non-alternating final long á and é from undergoing word-final low vowel shortening (e.g. lé ‘juice’, csokoládé ‘chocolate’, burzsoá ‘bourgeois’). True, the number of these words is low, but, nevertheless, shortening has to be blocked somehow. One possibility is marking these words with arbitrary diacritics in the lexicon. The autosegmental notation allows for a distinction between alternating and non-alternating final long vowels representationally without having to resort to exception features or postulating extra underlying segments.

\[
\begin{array}{llll}
\text{a. } & \text{X} & \text{b. } & \text{X X} \\
\text{c. } & \text{X X} & \text{d. } & \text{X X} \\
\text{V} & \text{V} & \text{V} & \text{V}
\end{array}
\]

The representation (64a) would be used for non-alternating short low vowels (e.g. *hat* ‘six’, *nem* ‘gender’), (64b) for non-alternating long low vowels (*lé* csokoládé, *burzsoá*), and (64c)

---

105 This analysis has been proposed by e.g. Szépe (1969), Hetzron (1972), Abondolo (1988), Jensen and Stong-Jensen (1989a).

106 Or the diacritic may be ‘phonologized’: Abondolo (1988) postulates different underlying vowels in these items.

107 Similar representations have been proposed by Jensen and Stong-Jensen (1989a), Péter Rebrus (personal communication), and Ritter (1995); see also Rebrus (2000a).
or (64d) for alternating long low vowels (ápá, téve). In this case the alternation between long and short low vowels is handled by a rule that spreads the root node of a low vowel to an adjacent floating X slot stem-finally before a suffix (either to the left (if the representation is (64d)) or to the right (if the representation is (64c)). In this treatment, there is a constraint specific to [+ open,] vowels: those represented as (64a) cannot occur stem finally in the lexicon. ľa and ma would be exceptions to the minimal stem/word constraint and the scarcity of final low vowels represented as (64b) would be an accident. The underlying distribution of low vowels would be analogous to that of mid vowels.

It is not our main concern here to decide which analysis handles the low vowel alternations discussed best. The point is that phonotactically they are equivalent. The main difference between them is which other class of vowels the final low vowels are grouped together with: if the first analysis is chosen low vowels pattern with high vowels in the conservative dialect; in the latter they pattern with mid vowels. As high vowels in the innovative dialect behave differently from both, the choice between the alternative analyses cannot be made on phonotactic grounds. In the present dissertation I follow Siptár & Törkenczy (2000) and assume that the lengthening analysis is correct. Thus, long and short low vowels can occur freely stem-finally.

To sum up, mid vowels are constrained by (58) and (vacuously) the minimal word/stem requirement (62). (58) holds for all representations, derived or underived. The underlying distribution of low vowels is only constrained by the minimal word/stem requirement. High vowels behave like low ones in the conservative dialect: underlying representations have to conform to (62). In the innovative dialect, however, the minimal word/stem constraint plays an active role. As we have seen above, in this dialect, stem-final long high vowels are banned (cf. 65) unless the representation required by (65) should violate the minimal stem/word constraint (62).

---

In Optimality Theory (Prince and Smolensky 1993) such a relationship can be expressed in a straightforward manner: the minimal stem/word constraint dominates (65).

(65) 
\[
\begin{array}{c}
* N \\
\Lambda \\
X X \\
\vee \\
[-\text{open}_2] \\
\end{array}_{\text{stem}}
\]

In innovative ECH (65) is a static constraint in the sense that there are no alternations between long and short high vowels in this environment. In the intermediate dialect it is possible to argue that there is a rule (66) which shortens stem-final long [-open₂] vowels optionally because both alternants may surface:

(66) 
\[
\begin{array}{c}
N \quad N \\
/ \quad | \\
X X \quad - \quad X \quad / \quad _ {\text{stem}} \\
\vee \quad | \\
[-\text{open}_2] \quad [-\text{open}_2] \\
\end{array}
\]

This rule is blocked if the output should violate the minimal word/stem constraint, which acts as a filter, or a ‘derivational constraint’ (Kisseberth 1970). There is no alternation evidence in the innovative dialect: stem-final high vowels are simply always short except in monosyllables. There is no satisfactory way to express this relationship between (65) and (62) in the present framework. Restricting (65) to stems which are longer than monosyllabic ‘does the job’, but it should be noted that this way the minimal stem/word requirement is ‘built into’ this constraint (and thus is stated twice). ¹⁰⁹:

(67) 
\[
\begin{array}{c}
* \sigma \quad \sigma \\
\Lambda \\
N \\
\wedge \\
X X \\
\vee \\
[-\text{open}_2] \\
\end{array}_{\text{stem}}
\]

¹⁰⁹ In Optimality Theory (Prince and Smolensky 1993) such a relationship can be expressed in a straightforward manner: the minimal stem/word constraint dominates (65).
Given (67), long high vowels are excluded finally *only* in polysyllabic stems and final short high vowels are banned by the minimal stem/word constraint in monosyllabic ones.

To sum up, (i) all final vowels are subject to the minimal stem/word constraint, (ii) mid vowels are also constrained by (58), (iii) high vowels are input to the (optional) rule (66) in the intermediate dialect and (iv) have to meet (67) in the innovative dialect.
3.4.2. VVCC: the complexity of the rhyme

In the previous section we discussed the behaviour of open syllables and pointed out that the distribution of vowels is different in medial and final open syllables. Let us now examine the behaviour of closed syllables.

In general, any vowel seems to be possible in a closed syllable (cf. section 3.2.3). However, there are restrictions holding in this environment depending on (i) the position of the syllable in the word and (ii) the morphological complexity of the word. (68) shows the distribution of long and short vowels in word-final syllables closed by a single consonant, in word-final syllables closed by more than one consonant, and word-medially when these syllables occur monomorphemically, i.e. undivided by a morpheme boundary:
(68)\textsuperscript{110} abstracts away from a few exceptional items.

<table>
<thead>
<tr>
<th>VC##</th>
<th>VCC##</th>
<th>VC.C</th>
</tr>
</thead>
<tbody>
<tr>
<td>i \hit{‘belief’}</td>
<td>ring \‘sway’</td>
<td>\inger{‘stimulus’}</td>
</tr>
<tr>
<td>ü \sün{‘hedgehog’}</td>
<td>csüng \‘hang’</td>
<td>kür\tö ‘funnel’</td>
</tr>
<tr>
<td>ö \sör{‘beer’}</td>
<td>gyöngy \‘pearl’</td>
<td>ördö\g ‘devil’</td>
</tr>
<tr>
<td>e \nem{‘gender’}</td>
<td>szent \‘saint’</td>
<td>persze \‘of course’</td>
</tr>
<tr>
<td>u \fut{‘run’}</td>
<td>must \‘grape juice’</td>
<td>undor \‘disgust’</td>
</tr>
<tr>
<td>o \lop{‘steal’}</td>
<td>gyors \‘fast’</td>
<td>boglya \‘stack of hay’</td>
</tr>
<tr>
<td>a \hat{‘six’}</td>
<td>tart \‘hold’</td>
<td>apró \‘tiny’</td>
</tr>
</tbody>
</table>

| i: | sír \‘grave’ | - |
| ü: | bűn \‘sin’ | - |
| ö: | bőr \‘skin’ | - |
| e: | kán \‘spy’ | érc \‘ore’ | értők \‘value’ |
| u: | rút \‘ugly’ | - |
| o: | kör {‘disease’} | - |
| a: | láp \‘marsh’ | márt \‘dip’ | árpa \‘barley’ |

\textsuperscript{110}(68) abstracts away from a few exceptional items.


‘Epenthetic’ stem forms like pótlás [po:tlas] ‘replacement’, ólmoz [o:lmoz] ‘lead’ (verb), etc. are only apparent counterexamples since they do not contain a cluster underlyingly: /po tá: s\ö/ \textsuperscript{(cf. 4.1.4.2).}
Recall that there are no complex onsets in Hungarian (cf. section 3.2.2). As expected/predicted, the constraint discussed applies even in cases when the second member of the interconstituent cluster is more sonorous than the first one: hypothetical (monomorphemic) */mu*+tra/ and */mu*+rta/ are equally impossible.

The state of affairs described is reminiscent of that in English where word-final single consonants can be preceded by long vowels, but—disregarding clusters that contain Level 2 suffixes—word-final and some word-medial onsets may not. Of course there are important differences (in English all long vowels behave in the same way; there are complex onsets (hence some medial clusters may follow long vowels: April /æprɪl/); coronal clusters behave differently from non-coronal ones: pint /pɪnt/; etc). The analysis sketched above is thus analogous to those presented in Myers (1987), Borowsky (1988, 1989), Jensen (1993), Rubach (1996).

As can be seen in (68), (i) any short vowel is possible in a closed syllable, and (ii) of the long vowels, only /e:/ and /a:/ can occur in non-word-final closed syllables and word-final syllables closed by more than one consonant (cf. Törkenczy 1989). This poses two questions: (a) why is there a difference between word-final syllables closed by a single consonant and the other kinds of closed syllables?; and (b) why do /e:/ and /a:/ behave differently from the other long vowels?

One might want to answer question (a) by utilising the notion of extrasyllabicity/extrametricality. If we say that a single word-final consonant is extrametrical in Hungarian at the point where the constraint against long vowels (except /e:/ and /a:/) in closed syllables applies, then it is understandable why there is an asymmetry between word-final VVC sequences vs. word-final and word-medial VVCC sequences. In the first case the word-final consonant is extrasyllabic and therefore the word-final syllable is not closed: rút /ruː< t >/. If there is more than one word-final consonant, then rendering the final one extrametrical still leaves a closed syllable behind and thus the constraint on long vowels applies: *VVC< C> . A word-medial vowel followed by two consonants is necessarily subject to the constraint because word-medial consonant clusters are necessarily heterosyllabic and extrametricality cannot apply here because of the Peripherality Condition (Hayes 1980, 1982).

However, the discrepancy between the behaviour of final VVC vs. medial and final VVCC clusters disappears if polymorphemic clusters are considered as well. Any long vowel is possible before a cluster if there is a morpheme boundary after the vowel or between the

---

111 Recall that there are no complex onsets in Hungarian (cf. section 3.2.2). As expected/predicted, the constraint discussed applies even in cases when the second member of the interconstituent cluster is more sonorous than the first one: hypothetical (monomorphemic) */mu:rtα/ and */mu:trα/ are equally impossible.

112 The state of affairs described is reminiscent of that in English where word-final single consonants can be preceded by long vowels, but—disregarding clusters that contain Level 2 suffixes—word-final clusters and some word-medial ones may not. Of course there are important differences (in English all long vowels behave in the same way; there are complex onsets (hence some medial clusters may follow long vowels: April /æprɪl/); coronal clusters behave differently from non-coronal ones: pint /pɪnt/; etc). The analysis sketched above is thus analogous to those presented in Myers (1987), Borowsky (1988, 1989), Jensen (1993), Rubach (1996).
It makes no difference if the intervening morpheme boundary is the edge of an analytical (e.g. bűn-ben) or a non-analytical domain (bűn-t): the constraint only holds within morphemes. This suggests that the phenomenon discussed is not a constraint on the complexity of the rhyme, but rather on morpheme shape. In other words, it is an MSC and not an SSC. Thus, instead of an extrametricality/extrasyllabicity analysis of the type sketched above, we propose that the distribution of preconsonantal long vowel is simply governed by the MSC (70a) (for a full list of the exceptional morphemes that violate (70a), see note 110. The constraint seems to be even stricter if the tautomorphic consonants following the vowel form a geminate: in this case even /e/, a/ appear to be ill-formed (cf. Törkenczy 1989) —this constraint is formulated below as (70b).

\[(70)\]
\[a. \quad *VVCC\]
\[\text{domain: morpheme}\]
\[\text{condition: } VV \neq /e/, a/\]

---

113 Gloses: bőrben ‘skin’ (iness.), bűnben ‘sin’ (iness.), bűnt ‘sin’ (acc.), főbb ‘main’ (comp.), kőrt ‘ask’ (3sg past indef.), kőrták ‘ask’ (3pl past def.’), kórnak ‘disease’ (dat.), kőrt ‘disease’ (acc.), sírt ‘grave’ (acc.), szívtam ‘suck’ (1sg past), túrt ‘dig’ (3sg past indef.), túrtak ‘dig’ (3pl past indef.), várt ‘wait’ (3sg past indef.), vártam ‘wait’ (1sg past).

114 Which is further weakened by the fact that unlike in English—there is no phenomenon other than preconsonantal vowel length motivating word-final extrametricality.
There are very few true exceptions. An exhaustive list of suspicious monomorphemic items (found in our database) is given below:

115


- /i/ avîtt ‘old-fashioned’;

- /o/ -ôdzik < verbal ending> : e.g. vakarôdzik [vǒkoro:dzik] ‘scratch’;

- /õ/ -õdzik < verbal ending> : e.g. dörgôlôdzik [dôrgôlô:dzik] ‘rub against’, elôtt ‘before’, (kôttés ‘cake made of raised dough’).

In most of these morphemes the geminate is a sonorant, especially /l/, preceded by /a/. Note that most of the morphemes listed are usually pronounced in a ‘regularised’ form, i.e. with a non-geminate consonant following the long vowel: állat [a:lot], inkább [iŋka:b], majoránna [majora:nna], ĝjel [e:jel], etc. The verbal endings -ôdzik and -ôdzik also have alternative forms with intervocalic /z/, in which case they do not violate (70b): vakarôzik [vôkoro:zik], dörgôlôdzik [dôrgôlô:zik]. The truly exceptional items above, i.e. those that—in my native judgement/intuition—cannot be regularised have been emboldened. Of these, béllet is extremely rare/technical. The rest all have obstruent geminates after the long vowel. Köttés is only included for the sake of completeness: it is a dialectal word that is not part of ECH.

115There are very few true exceptions. An exhaustive list of suspicious monomorphemic items (found in our database) is given below:


- /i/ avîtt ‘old-fashioned’;

- /o/ -ôdzik < verbal ending> : e.g. vakarôdzik [vǒkoro:dzik] ‘scratch’;

- /õ/ -õdzik < verbal ending> : e.g. dörgôlôdzik [dôrgôlô:dzik] ‘rub against’, elôtt ‘before’, (kôttés ‘cake made of raised dough’).

In most of these morphemes the geminate is a sonorant, especially /l/, preceded by /a/. Note that most of the morphemes listed are usually pronounced in a ‘regularised’ form, i.e. with a non-geminate consonant following the long vowel: állat [a:lot], inkább [iŋka:b], majoránna [majora:nna], ĝjel [e:jel], etc. The verbal endings -ôdzik and -ôdzik also have alternative forms with intervocalic /z/, in which case they do not violate (70b): vakarôzik [vôkoro:zik], dörgôlôdzik [dôrgôlô:zik]. The truly exceptional items above, i.e. those that—in my native judgement/intuition—cannot be regularised have been emboldened. Of these, béllet is extremely rare/technical. The rest all have obstruent geminates after the long vowel. Köttés is only included for the sake of completeness: it is a dialectal word that is not part of ECH.
pairs whose members are considerably\textsuperscript{116} different in quality at the surface. One may try to explain the special behaviour of /ɛ:/ with respect to (70a) by connecting it with the special character of the pairs [ɔ – a:], [ɛ – e:].\textsuperscript{117} We assume in this dissertation (following Siptár & Törkenczy 2000) that [ɔ – a:] and [ɛ – e:] are underlyingly [DOR, + open1, + open2] and [COR, + open1, + open2], respectively. The surface quality differences (rounding (and height) in the case of [ɔ – a:] and height in the case of [ɛ – e:]) are the result of phonetic implementation conditioned by the underlying quantity difference: long /ɛ:/ is interpreted phonetically as mid ([e:]), short /a/ as rounded ([ɔ]). There is no theoretical reason not to do this the other way round (cf. Törkenczy 1994, Polgárdi 1997). One could assume that the underlying difference between the members of the pairs discussed is qualitative (/ɔ – a/, /ɛ – e/) and the surface difference in length is a matter of phonetic implementation which is conditioned by the underlying quality difference (/a/ and /ɛ/ will appear as long at the surface).

This move has advantages and disadvantages. On the positive side, the exceptional behaviour of [ɛ:] and [a:] with respect to (70a) is no longer a mystery: these vowels are not constrained by (70a) because (70a) is a constraint on long vowels and [ɛ:] and [a:] are not long underlyingly. (70a) could be restated without its condition and could be collapsed into a single constraint with (70b). However, in my view, the negative effects are more serious: (i) the vowel inventory would become asymmetrical (only high and some mid vowels would have long counterparts); (ii) it would be no longer possible to express Low Vowel Lengthening as a uniform process: it would have to be lowering or raising (depending on which vowel we take as underlying) for /ɛ/ – /e/, but rounding or unrounding for /ɔ/ – /a/; (iii) as /ɛ/ – /e/ and /ɔ/ – /a/ could not be analysed as length alternations, the alternations nyár ‘summer’ – nyar-at ‘summer’ (acc.), től ‘winter’ – tel-et ‘winter’ (acc.) could not be treated as the same process (Stem Vowel Shortening (cf. Siptár & Törkenczy 2000)) as that involving other vowels, e.g. víz ‘water’ – viz-et ‘water’ (acc.). Therefore, we shall not reanalyse the representation of the

\textsuperscript{116}There are small differences of height between the surface reflexes of the members of some other pairs, notably /ɔ: – o/ and /ō: – ō/, the long segments being slightly more closed than the short ones, cf. Siptár & Törkenczy (2000).

\textsuperscript{117}These pairs of vowels are unlike the rest in other ways as well, e.g. they are the only ones that alternate stem finally (cf. the discussion of Low Vowel Lengthening sections 3.4.1). It is an interesting idea to suppose that all these phenomena are related and may have a common explanation, but we will not pursue it in this dissertation.
vowels [ɔ – a:], [e – e:] in the way described above (for an additional piece of evidence, cf. section 3.3.1.). This means that the condition on (70a) remains a stipulation.

3.4.3 Word-class-specific constraints: the phonotactics of verbs

I have pointed out in section 2.1 that phonotactics may be specific to a subpart of the lexicon (a stratum/sublexicon) and that such a state of affairs occurs in Hungarian. In this section I shall discuss the phonotactic constraints that apply to a well-identifiable stratum of the Hungarian lexicon, the sublexicon of verbs.118 There is a phonotactic difference between verb stems and non-verb stems. A non-verb stem is not identifiable as a non-verb on the basis of the string of segments it consists of alone (by a native speaker), while a verb may be identified as a verb. This is a unidirectional relationship: some strings could not be monomorphemic verb stems, but there are strings that could equally be monomorphemic verb stems or non-verb stems: nyom ‘trace’, nyom ‘push’; part ‘shore’; tart ‘hold’; domb ‘hill’, but */domb/v.

The primary cue to phonotactic ‘verb-ness’ is the stem-final coda.119 We shall first consider monomorphemic free120 verb stems.

A monomorphemic free verb stem can end in any single consonant except /t', t', ć,

---

118 Note that it is no longer possible to add monomorphemic verb stems to the Hungarian lexicon. All new verbs (loans or otherwise) that enter the lexicon must have a verbal suffix (typically -ol/-el/-öl: sztrákol ‘go on strike’, downloadol [dounlo:dol] ‘download’). Thus, monomorphemic verb stems are a closed system.

119 It is often claimed that monomorphemic verb stems cannot end in a vowel in Hungarian (e.g. Vago 1980, 1989, Kiefer 1994). This is based on the fact that all monomorphemic verb stems that end in a vowel in isolation are ‘v-adding’ stems, i.e. they end in /v/ before a vowel-initial suffix: lő ‘shoot’ – ‘lőv-ök ‘shoot’ (1sg pres. indef.) (as opposed to vowel-final nominal stems, which may be ‘v-adding’ stems or ‘non-v-adding’ major stems: ló ‘horse’ – lov-ák ‘horses’, melő ‘job’ – melő-k ‘jobs’). If we assume that a stem-final /v/ underlies the [v] - ø alternation (as is usual in the relevant analyses), then it follows that, underlingly, verb stems can only end in a consonant. In the present analysis, however, we consider the [v] - ø alternation non-phonological (suppletive), cf. section 4.2.1. Consequently, it does not hold that all verb stems end in a consonant since both the /v/-final and the vowel-final alternants of the relevant verb stems are present in the lexicon.

120 This excludes -ik-verbs and defective verbs.

121 Of these the lack of final /j/ is not surprising since there are hardly any items in the lexicon with final short /j/: bridge /briːj/, ?/briːj/ ‘bridge (card game)’. Disregarding /j/, non-verbs/nouns can end in any single consonant, so the other gaps are specific to verb stems. Some generalisations could be made but it is unclear whether that the gaps in the case of verbs are accidental or not. For instance, we could claim that there is a constraint banning affricates at the end of a monomorphemic verb stem, but then it seems that long affricates (as opposed to short ones) are possible in the same position: metsz [mɛtʃ:] ‘cut’, edz [ɛdʒ:] ‘train’. However, there are only three monomorphemic verb stem that end in a long affricate: metsz, edz and pedz ‘begin to understand’. Since the geminates do exist, we take a ‘soft’ approach and consider that affricates are possible at the end of a verb stem (i.e. the gaps are accidental). Similarly, we take the lack of /t/, z/ accidental because their voiced/voiceless counterparts /d/, ʃ/, respectively, exist as non-branching codas in verbs (e.g. hagy ‘leave’, ás ‘dig’). The lack of /x/ is specific to the verb and it is excluded as a geminate coda as well. This is due to a constraint that excludes /x/ from the coda in the verb sublexicon:

(71) (verb sublexicon)

/x/ is not licensed in the coda.

Apart from the effect of (71) single codas are not radically more restricted in monomorphemic free verb stems than in non-verbs.

The situation is strikingly different if we examine verb stems that end in a consonant cluster. We claimed in section 3.2.4.3 that monomorphemic final three-term consonant clusters are exceptional/irregular. There are very few monomorphemic words that do contain these

---


122 Note that the affricate [dʒ:] is the surface reflex of the underlying cluster /dz/, cf. Siptár & Törkenczy (2000), Siptár (2002).

123 Geminate /z/ is also unattested in the coda generally, not just in verbs, geminate coda /x/ occurs only in non-verbs, e.g. pech /pɛx/ ‘bad luck’.
clusters (see the complete list in (23) above). It must be noted that none of these exceptional items are verbs. (72) below shows the complete list of two-term clusters (including geminates) that monomorphemic free verb stems can end in together with the number of stems (in angled brackets) that contain each type of final cluster:

(72) 

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Stems</th>
</tr>
</thead>
<tbody>
<tr>
<td>d</td>
<td>&lt;1&gt;</td>
</tr>
<tr>
<td>d'</td>
<td>&lt;2&gt;</td>
</tr>
<tr>
<td>g</td>
<td>&lt;3&gt;</td>
</tr>
<tr>
<td>t'</td>
<td>&lt;1&gt;</td>
</tr>
<tr>
<td>st</td>
<td>&lt;117&gt;</td>
</tr>
<tr>
<td>št</td>
<td>&lt;1&gt;</td>
</tr>
<tr>
<td>zd</td>
<td>&lt;2&gt;</td>
</tr>
<tr>
<td>žd</td>
<td>&lt;1&gt;</td>
</tr>
<tr>
<td>mt</td>
<td>&lt;1&gt;</td>
</tr>
<tr>
<td>(mz)</td>
<td>&lt;1&gt;</td>
</tr>
<tr>
<td>nt</td>
<td>&lt;100&gt;</td>
</tr>
<tr>
<td>nd</td>
<td>&lt;4&gt;</td>
</tr>
<tr>
<td>nz</td>
<td>&lt;1&gt;</td>
</tr>
<tr>
<td>ňg</td>
<td>&lt;51&gt;</td>
</tr>
<tr>
<td>(nl)</td>
<td>&lt;1&gt;</td>
</tr>
<tr>
<td>lt</td>
<td>&lt;14&gt;</td>
</tr>
<tr>
<td>ld</td>
<td>&lt;7&gt;</td>
</tr>
<tr>
<td>(lg)</td>
<td>&lt;1&gt;</td>
</tr>
<tr>
<td>l:</td>
<td>&lt;34&gt;</td>
</tr>
<tr>
<td>rt</td>
<td>&lt;8&gt;</td>
</tr>
<tr>
<td>rd</td>
<td>&lt;2&gt;</td>
</tr>
<tr>
<td>r:</td>
<td>&lt;2&gt;</td>
</tr>
<tr>
<td>jt</td>
<td>&lt;19&gt;</td>
</tr>
</tbody>
</table>


The parenthesized clusters in (72) above are only apparent (and are only included because they are cited in the literature). Nemz and uralg (the only stems with final /mz/ and /lg/) are obsolete in ECH as free stems and ajánl normally does not contain a final cluster [ʃaːl(ː)]. We have already pointed out in section 3.2.4 that teremt is irregular (and may not even be monomorphemic (?)).

---

124 For a complete list of verb stems representing the less populous types see Appendix Bii.
If we disregard these items, it becomes obvious that monomorphemic branching codas in the verb sublexicon are subject to much stricter constraints than (monomorphemic) branching codas in general. Here a monomorphemic branching coda must be a cluster that respects Sonority Sequencing (its right-hand term must be less sonorous than its neighbour on the left) and/or is homorganic (it is a partial or a full geminate). Specifically, such a cluster may be one of the following kinds: (i) it respects Sonority Sequencing, but is not strictly homorganic (/št, žd, jšt/), or (ii) it respects Sonority Sequencing and is strictly homorganic (/st, zd, nt, nd, nz, ɲg, rt, rd, lt, ld/), or (iii) it is unlicensed by Sonority Sequencing and is strictly homorganic (/dː, tː, dːtː, gː, lː, rːtː/). The generalisation that can be made is that clusters in set (i) are coronal while clusters in sets (ii) and (iii) are non-labial. This suggests that a branching coda in a monomorphemic free verb stem must be a coronal cluster if it is only licensed by government, but may be coronal or dorsal if it is licensed by place-binding or root-binding.

Assuming the general condition on the licensing of coda clusters (9 section 3.2.4) and the universal right-to-left direction of government in the coda, this is captured by the following constraints.  

(73)  

Branching Coda Constraint (verb sublexicon)

a. In a coda cluster C1C2 government can only apply iff C1 and C2 are both COR.

b. In a coda cluster C1C2 (root or place) binding can only apply if C1 and C2 are both COR or DOR

(73) does not permit a branching codas like /lk, rv/ (non-verb examples: halk ‘silent’, év ‘argument’) for verbs because they are unlicensed by government (they are non-coronal) and they are not licensed by binding (their terms do not share a root or a place node). Similarly, /pː, mb/ (non-verb examples: csepp ‘drop’, domb ‘small hill’) —although the do have a shared

---

125 This is true of regular non-verb stems too, but in the verb sublexicon it is also exceptionless.

126 Recall that the restriction that both members of the coda cluster should be coronal only applies in the case of subminimal government generally (i.e. outside the verb sublexicon) and not when the cluster is licensed by government (non-subminimal) or root-binding. Also, LAB segments are not excluded from binding generally. Compare (16) section 3.2.4.
What is special in this stratum is that there are no exceptions to the general constraint. Recall that we have considered words like *barack* exceptional, see section 3.2.4.

Root note and a shared place node, respectively— are not well formed branching codas in a verb because they are not licensed by government according to (73a) and they are also not licensed by binding since they are LAB (73b).

Not all the branching codas permitted by (73) are well-formed. Some further restrictions apply. It can be seen in (72) that the second position can only be filled by a sonorant it shares its root-node with the first (i.e. if it root-binds the first (= it is a geminate)): clusters like /j,l, rn, rn', ln/ (non-verb examples: *fajl* ‘file’, *konszern* ‘concern’, *szörny* ‘monster’) are not well-formed branching codas in this stratum although they are licensed according to (73) by government and/or place-binding. This can be stated as:

(74) (verb sublexicon)

In a coda cluster C1C2 if C2= [+ son], then C2 must root-bind C1.

Affricates only seem to be possible as geminates in the branching coda. They cannot occur in the first position in a branching coda when not root-bound, or in the second position when not root-binding the first: clusters like /t'k, rt', nt'/ (non-verb examples: *barack* ‘peach’, *perc* ‘minute’, *tánc* ‘dance’) are not well-formed branching codas. The lack of non-root-bound affricates (in the first position) is expected (and is not specific to the verb stratum)\(^{127}\) since they are not licensed by government in this position by the general coda constraint (9 section 3.2.4). The lack of non-root-binding affricates (in the second position) requires an additional constraint because clusters like /rt', nt'/ are licensed according to (73) by government and/or place-binding. This can be stated as:

(75) (verb sublexicon)

In a coda cluster C1C2 if C2 is an affricate, i.e. C2 branches under its root node, then C2 must root-bind C1.

The occurrence of palatals is extremely restricted in the branching coda of a free verb stem:

\(^{127}\) What is special in this stratum is that there are no exceptions to the general constraint. Recall that we have considered words like *barack* exceptional, see section 3.2.4.
only [+ continuant] palatals (/š, ž, j/) can occur and only when governed (i.e. in C1 position). [-continuant] palatals (/tɻ, dɻ, nɻ/) do not occur at all and [+ continuant] palatals cannot occur in C2 position or as geminates. Codas like /nɻtɻ, rɻ, rɻ, jɻ, tɻ; dɻ, nɻ; š, jɻ/ (non-verb examples: ponty ‘carp’, tárgy ‘object’, szörny ‘monster’, vers ‘poem’, pajzs ‘shield’, pötty ‘polka dot’, meggy ‘sour cherry’, genny ‘pus’, friss ‘fresh’, gally ‘branch) are not permitted. This can be stated as.\(^{128}\)

(76) (verb sublexicon)

In a coda cluster a [DOR, -ant] segment must be [-cont] and governed.

There are some branching codas that do not occur, but are permitted by (71), (73), (74), (75), (76). We consider these accidental gaps. Some of the accidentally missing coda clusters, although not found in monomorphemic free verb roots, actually occur undivided by a morpheme boundary finally in free suffixed verbs: e.g. [ŋk] kapar-unk ‘scratch’ (1pl pres indef.), [tː] vezet-ett ‘drive’ (3sg past indef.). By contrast, no morphologically undivided coda cluster occurs verb-finally that should violate the constraints discussed above (e.g. -bb [bː] could not be a verbal suffix; non-verb-example: comparative -bb nagy-obb ‘bigger’). This supports the accidental gap interpretation of the unattested clusters that are permitted by (71), (73), (74), (75) (76).

We have seen that the constraints that apply to branching codas in the verb sublexicon are more restrictive than the general constraints on branching codas. The same effect can be seen in the case of intervocalic clusters (albeit in a weaker form).

There are no monomorphemic verbs with an intervocalic consonant cluster consisting of more than three consonants. There are a few examples in our database of verb stems that seem to contain an intervocalic three-term cluster that is undivided by a morpheme boundary. An exhaustive list is given in (77):

\(^{128}\)Note that the specifications [+ continuant] and [-continuant] must be interpreted restrictively in the formulation of this constraint: it does not refer to palatal contour segments (palatal affricates), which contain both specifications, and whose distribution is constrained by (75).
### Glosses:

- **exponál** 'release the shutter of a camera', 
- **kasztrál** 'castrate',  
- **nosztrifikál** 'validate a foreign diploma',  
- **regisztrál** 'register',  
- **kimustrál** 'discard',  
- **ministrál** 'assist at mass',  
- **impregnál** 'impregnate',  
- **imprimál** 'pass the proofs for the press',  
- **improvizál** 'improvise',  
- **vámdliz**  
- **transzponál**  
- **inspiciál**  
- **inkriminál**  
- **pertraktál**  
- **kifi runcvancigol** (?)  
- **inkriminál**  
- **pertraktál**  

Note that in (77) all ‘monomorphemic’ verbs end in -\textit{vl} (or in one case), -\textit{vz}. These endings are (identical with) denominal verb-forming derivational suffixes that productively attach to clearly identifiable stems: e.g. \textit{kasza} ‘scythe’– \textit{kaszá-}l ‘scythe\_v’, \textit{dob} ‘drum\_N’ – \textit{dob-ol} ‘drum\_v’, \textit{gumi} ‘rubber’ – \textit{gumi-z} ‘rubberize’. Even though the verbs in (77) do not have such an easily identifiable nominal root+ verb-forming suffix structure (*expona – exponál ‘release the shutter of a camera’), it is reasonable to assume that they follow the same pattern: they consist of a nominal ‘phantom stem’ and a \textit{denominal} verb-forming derivational suffix.\textsuperscript{130}

Under this analysis there are no examples of monomorphemic verbs with three-member intervocalic consonant clusters. Recall that we consider monomorphemic intervocalic CCC clusters phonotactically ill-formed in general, and thus \textit{all} monomorphemic lexical items with an intervocalic cluster of more than two members are exceptional/irregular (cf. section 3.3.2.2). Thus, the restriction is not specific to verbs, but again, verb stems do not even offer

\textsuperscript{129}Glosses: \textit{exponál} ‘release the shutter of a camera’, \textit{kasztrál} ‘castrate’, \textit{nosztrifikál} ‘validate a foreign diploma’, \textit{regisztrál} ‘register\_v’, \textit{kimustrál} ‘discard’, \textit{ministrál} ‘assist at mass’, \textit{impregnál} ‘impregnate’, \textit{imprimál} ‘pass the proofs for the press’, \textit{improvizál} ‘improvise’, \textit{transzponál} ‘transpose’, \textit{inspiciál} ‘inspect’, \textit{inkriminál} ‘incriminate’, \textit{pertraktál} ‘discuss in detail’. \textit{Kifi runcvancigol} ‘figure out, calculate’ and \textit{vámdliz ‘?’} are only included because they can be found in the database. I have been unable to establish if \textit{f\'i runcvancig} and \textit{vámdliz} exist as nouns (they are not in the database) and what \textit{vámdliz} means.

\textsuperscript{130}They are like the ‘phantom stems’ in \textit{insist}, \textit{consist}, \textit{resist}, etc. in English, cf. Chomsky & Halle (1968), Aronoff (1976).
irregular counterexamples to the generalisation.

Let us examine two-member intervocalic clusters in verb stems. Table VII below shows the intervocalic CC clusters in monomorphemic verbs attested in the database and compares them to the attested CC clusters in monomorphemic non-verbs. The notation used is a variant of the usual one employed in this dissertation: a blank space in an intersection of a row and a column means that the relevant cluster is unattested or only occurs when the two consonants are separated by an analytic morphological domain boundary; a ‘+ ’ occurs if it is attested in monomorphemic items and the number of such items in the database is n > 15; numbers have been used to indicate the number of monomorphemic items in the database when the cluster in question is attested in monomorphemic items and the number of such items is n ≤ 15. Colons separate data about non-verbs from data about verbs: non-verb data appear on the left of the colon (non-verb:verb). The symbol ‘− ’ is used to indicate the lack of a cluster if the same cluster is attested on the other side of the colon (‘− :n’, or ‘n:: ’), i.e. if the cluster is attested in a monomorphemic verb or non-verb. For ease of reference intervocalic clusters that occur in monomorphemic verb stems (i.e. those with a number or a ‘+ ’ on the right of the colon: ‘:n’ or ‘:+ ’) have been encircled (in any colour). The circle is red if the number of such items in the database is n > 15 (i.e. if a ‘+ ’ appears on the right of the colon: ‘:+ ’); and the circle is black if the number of such items in the database is n ≤ 15 (i.e. if a number appears on the right of the colon: ‘:n’ ) while the number of non-verb items is greater than the number of verb items (i.e. ‘+ :n’ or ‘mn:n where m > n’). The circle is blue if the number of verb items is greater than the number of non-verb items and the number of non-verb items is zero (i.e. ‘− :n’); and the circle is green if the number of verb items is greater than the number of non-verb items and non-verb items do exist (i.e. ‘mn:n’ where m < n and m ≥ 0).

Some explanation is in order about the data included in Table VII before we examine it. I have included all verb stems in the database that can be regarded suspicious of having an internal monomorphemic intervocalic CC cluster. I have relied on following guidelines to judge what is suspicious:

(78) In order to qualify as a verb-stem-internal cluster (‘vsi-cluster’), an intervocalic cluster must meet one or more of the following partially overlapping, partially conflicting criteria:

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Description</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘− :n’</td>
<td>Non-verb</td>
<td>‘− :n’</td>
</tr>
<tr>
<td>‘+ :n’</td>
<td>Verb</td>
<td>‘+ :n’</td>
</tr>
<tr>
<td>‘mn:n’</td>
<td>Non-verb</td>
<td>‘mn:n’</td>
</tr>
<tr>
<td>‘+mn:n’</td>
<td>Verb</td>
<td>‘+mn:n’</td>
</tr>
</tbody>
</table>
(a) The VCCV string must be contained in a verb stem: the CC cluster cannot be divided or immediately preceded/followed by the edge of a productive suffix (and the stem must have more-or-less non-transparent, non-compositional meaning).

Example: this rules out /kt/ in buktat ‘flunk’ (because -tat is a productive suffix whose edge breaks up the cluster: buk-tat), and /kl/ in csuklik ‘hiccup’ (because -ik is a productive suffix whose edge immediately follows the cluster: csukl-ik).

(b) The verb-stem that contains the cluster may or may not have an internal stem, but if it does, then the pre-ending (ghost) stem (i) may only have bound forms (in the given meaning) or (ii) must be suppletive if one of its allomorphs is free.

Example: this rules out /pr/ in apr-i t ‘chop up’ (because the stem has a free allomorph apró ‘small’), /mp, st/ in komposzt-ál ‘compostv’ (because komposzt ‘compostN’ is a free morpheme), and /rm/ in karm-ol ‘clawv’ (because the stem has a free allomorph karom ‘clawN’).

(c) The cluster must be truly morpheme-internal (wholly contained) within a strict or extended verb stem.

Example: this rules out /ng/ in angoloz ‘study English’ (because, although the cluster is wholly contained within a monomorphemic stem, it is not a verb stem: [angol]N ozN)

(i) verb stem = strict verb root

The VCCV string must be strictly verb-root-internal: the CC cluster may not be preceded/divided/followed by the edge of even an unproductive suffix/ending.

Example: this accepts /lv/ in olvas ‘read’ (because it is a monomorphemic root and does not contain even a ghost stem), but rules out /pt/ in kaptat ‘climb with difficulty’ and /ks/ in maximál ‘maximize’ (because they
are extended verb stems, i.e. they contain ghost stems $kap$-$tat^{131}$, $maxim$-á$l$ or $maximá$-l$)

(ii) verb stem = extended verb root (the ghost stem must be verbal)

The CC cluster may be part of a ghost stem, but cannot be part of a non-verbal ghost stem: The cluster may be preceded/divided/followed by the edge of an unproductive suffix/ending iff the unproductive suffix/ending is deverbal (verb-.verb).

Example: this accepts /lv/ in olvas (because it is a monomorphemic root and does not contain even a ghost stem) and /pt/ in kaptat (because the ghost stem in $kap$-$tat$ is verbal, i.e. the unproductive verb-forming ending -$tat$ is a deverbal one), but rules out /ks/ in maximál (because the ghost stem in $maxim$-á$l$ or $maximá$-l is non-verbal i.e. the unproductive verb-forming ending -(V)$l$ is not a deverbal one)

(iii) verb stem = extended verb root (the ghost stem may be non-verbal)

The CC cluster may be part of a non-verbal ghost stem: The cluster may be preceded/divided/followed by the edge of an unproductive suffix/ending.

Example: this accepts /lv/ in olvas ‘read’, /pt/ in kaptat ‘climb with difficulty’ and /ks/ in maximál

I have included all suspicious/potential intervocalic vsi-clusters in Table VII. All verb stems containing a suspicious/potential intervocalic vsi-cluster must meet the criteria (78a, b) in order to qualify. Those that do have been sorted into three sets: ‘strictly suspicious’ (‘strict’), ‘moderately suspicious’ (‘moderate’) and ‘loosely suspicious’ (‘loose’). A strict intervocalic vsi-cluster is one that can be found in at least one verb that meets the criteria (78a, b, ci) (e.g. /lv/ olvas). A moderate intervocalic vsi-cluster is one that can be found in at least one verb that meets the criteria (78a, b, cii) (e.g. /pt/ kaptat). A loose intervocalic vsi-cluster is one that can

---

131I take the unproductive ending -$tat$ here to be different from the productive causative -$tat$/-$tet$ suffix.
be found in at least one verb that meets the criteria (78a, b, ciii) (e.g. /ks/ maximál).132 In Table VII, the markings for strict and moderate intervocalic vsi-clusters have been highlighted by shading.133 The examples that follow Table VII are always the ‘best’ ones in the sense that a strict example has been given for a cluster there is one, a moderate example has been given if there is not a strict example, but a moderate one exists and finally, a loose example is only provided if there are no strict or moderate examples. Strict examples have been set in bold, moderate examples have been underlined and endings have been CAPITALISED.

132These sets are in a relationship of inclusion: a strictly suspicious cluster is also moderately suspicious and loosely suspicious, etc. Since it makes no difference in the discussion we shall refer to the strict, moderate and loose sets as if they were complementary.

133The distinction between the moderate and the loose set is an extremely difficult (if not impossible) one to make (consistently) since only non-productive endings added to ghost stems qualify for being sorted into deverbal and non-deverbal. I took an (arbitrary,) authoritarian and procedural approach to this problem. I have sorted the suspicious verb-forming endings into three sets. I have considered an unproductive ending consistently deverbal if it is listed as such in all the four (very different) sources I have used (Kiefer 1998, Rebrus 2000a, Tompa 1961, Velcsovné 1988). Similarly, I have considered an unproductive ending consistently non-deverbal if it is listed as such in all the four sources. Finally, I have made no decision and judged an ending uncertain if it was not possible to make a decision. The latter state of affairs often occurs when the stem is bound: -kVdik in [tanár]-kodik, ‘work as a teacher’ is obviously non-deverbal since tanár ‘teacher’ is a noun, it is deverbal in [emel]-kedik, ‘rise’ since emel ‘lift’ is a verb, but in [incsel]-kedik, ‘tease’ it is uncertain since incsel- does not occur elsewhere/is bound. Similar examples: -i t: [meleg]_ADJ i t, ‘warm up’(meleg ‘hot’), [áll]_V i t, ‘make stop’ (áll ‘stop’), [sand]_V, ‘squint’, [kőzvet], ‘mediate’; -ul/-ül: [tan]-ul, ‘study’ (tan ‘teaching’), [nyom]-ul, ‘push ahead’ (nyom ‘push’), but [izg]-ul, ‘be anxious’.

Thus, -kodik/-kedik/-kődik, -i t, and -ul are judged uncertain. I have considered an item to contain a cluster that is a member of the moderate set if its ending was consistently deverbal, I have judged it loose if its ending was consistently non-deverbal or uncertain. The list of the relevant non-productive endings sorted is the following (only those variants have been included that have been found in the database in stems with a suspicious vsi-cluster): deverbal endings: -(k)ódik, -(k)őzik, -(k)őzik, -(t)at/-(t)et, -(ad/-ed, -an/-en, -ant/-ent, -asz/-eszt, -dogál/-degd, -dos, -ell/-all, -gat/-get, -int, -lal/-lel, -ő(d)zik, -ődik/-ődik, -ög/-eg/-ög, -ong/-eng/-öng; non-deverbal endings: -ál/-ől/-l, -ol/-el/-ől/-al/-l, -oz/-ez/-őz/-az/-z, uncertain endings: -(k)odik/-(k)ódik/-(o)dik, -(k)ozik/-(k)ezik/(k)őzik, -(ász(ik)/-ősz(ik), -(i)t, -(ul/-ül).
Table VII  Intervocalic CC clusters in monomorphic non-verbs and verbs\(^{134}\)

| p | t | t' | k | b | d | d' | g | t' | č | j | f | s | š | v | ž | m | n | n' | l | r | j | x |
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |

\(^{134}\)For an exhaustive list of verbs containing these clusters, see Appendix B (iii).
Note that frequency (type or token) has no (clear) formal/theoretical status in a generative model of phonology (derivational generative phonology, Optimality Theory or Government/CV/VC Phonology) — even though it is sometimes referred to in various generative analyses. Thus, inasmuch as the observations made here are correct and relevant to phonotactics, they actually point/lead outside the model used in this dissertation. Frequency is accommodated by other frameworks, e.g. in pre-generative (structuralist) phonology (e.g. Spang-Hanssen 1959) and recently, in laboratory phonology (e.g. Hay, Pierrehumbert and Beckman 2003, Pierrehumbert 1994) and the (re)emerging functional approach (Bybee 2001, Bybee & Hopper 2001, Rebrus & Trón 2002, Trón & Rebrus 2000, 2001).

It can be seen in Table VII that, although the set of intervocalic CC clusters that can be found in a monomorphemic verb stem (strict, moderate or loose) is similar to the set of intervocalic CC clusters in non-verb stems, verb stems are phonotactically more restrictive than non-verb stem in a ‘weak’ (frequency-related) sense. The intervocalic CC clusters found in verbs do follow the general pattern for interconstituent clusters (i.e. they conform to the general interconstituent constraints we have stated in section 3.3.2 ((35), (37), (42), (43), (48), (51), (52), (53)) and no additional, more restrictive phonotactic constraints can be formulated with a sufficient degree of generality in terms of natural classes. The higher restrictiveness of verb phonotactics manifests itself in the ‘population’ (type frequency) of clusters within the phonotactic space permitted by the general constraints.\(^{135}\)

The following generalisations can be made: (a) the number of the types of clusters that can be found in verbs is significantly lower than those found in non-verbs. Of the total 249 attested monomorphemic intervocalic CC clusters, 242 are attested in non-verbs and only 145 are attested in verbs; (b) It is extremely rare for an attested intervocalic CC cluster to be found in more verb stems than non-verb stems, i.e. \(nv < v\) (where \(nv\) is ‘non-verb’ and \(v\) is ‘verb’). Of the total 249 attested monomorphemic intervocalic CC clusters, there are only 12 clusters of this kind, 7 of which are only attested in verbs (numbers in parentheses refer to the \(nv:v\) ratio): /t\(\text{iv}\) (-:1), t\(\text{m}\) (-:1), sx (-:4), m\(\text{č}\) (-:2), j\(\text{p}\) (-:1), j\(\text{š}\) (-:1), xx (-:1)/, /d\(\text{ě}\) (2:5), gd (1:2), ě (1:9), m\(\text{ž}\) (2:3), jg (3:4)/, and the difference in favour of verbs is typically very small (8 times out of the 12 cases \(v > nv\) = 1); (c) By contrast, it is very often the case that an attested

\(^{135}\)Note that frequency (type or token) has no (clear) formal/theoretical status in a generative model of phonology (derivational generative phonology, Optimality Theory or Government/CV/VC Phonology) — even though it is sometimes referred to in various generative analyses. Thus, inasmuch as the observations made here are correct and relevant to phonotactics, they actually point/lead outside the model used in this dissertation. Frequency is accommodated by other frameworks, e.g. in pre-generative (structuralist) phonology (e.g. Spang-Hanssen 1959) and recently, in laboratory phonology (e.g. Hay, Pierrehumbert and Beckman 2003, Pierrehumbert 1994) and the (re)emerging functional approach (Bybee 2001, Bybee & Hopper 2001, Rebrus & Trón 2002, Trón & Rebrus 2000, 2001).
Stop+ stop is also populous in verbs provided that /t/ and /k/ combine. Note that this does not imply that there are no clusters in verbs that violate the Syllable Contact Law, just that those types of clusters are not populous in verbs.

Cluster is found in more non-verb stems than verb stems (or that it occurs in non-verb stems only), i.e. $nv > v$. Of the 249 attested intervocalic CC clusters 209 occur in a higher number of non-verb stems than verb stems, of which 104 are only attested in non-verbs. Also the difference in favour non-verbs is typically not small: in 92 cases out of the 209 attested clusters the difference in favour of non-verb stems is at least 10 ($nv - v \geq 10$, where non-occurrence is counted as zero); (d) There are a significantly greater number of ‘populous’ cluster types (i.e. a ‘populous’ cluster type is a cluster found in a relatively high number of stems) in the case of non-verbs than in the case of verbs. Of the 249 attested intervocalic CC clusters there are 124 relatively populous non-verb clusters (clusters with at least 10 tokens: $nv$: where $nv \geq 10$) and only 36 relatively populous verb clusters ($v$ where $v \geq 10$). Interestingly, the distribution of the populous intervocalic CC clusters in the phonotactic space is somewhat different for verbs and non-verbs. In the case of non-verbs the populous clusters are distributed more or less evenly in the phonotactic space delimited by the constraints stated in section 3.3.2 ((35), (37), (42), (42), (48), (51), (52), (53)). By contrast, in the case of verb stems, populous clusters tend to be geminates, $s/s+$ stop clusters and sonorant+ obstruent clusters, but not obstruent+ sonorant clusters (which can be populous in non-verbs). This may suggest that, although the Syllable Contact Law does not hold in Hungarian in general, the phonotactics of verbs is skewed towards it because the populous intervocalic clusters tend to conform to it.

These effects become even stronger the stricter one is in circumscribing what qualifies as an intervocalic cluster in a monomorphemic verb: of the 249 attested intervocalic CC clusters in Table VII 145 are attested in verbs under the loose interpretation. Of these only 84 are strict or moderate vsi-clusters (highlighted by shading in Table VII), of which only 19 are strict vsi-clusters. An exhaustive list of the strict vsi-clusters together with all the strictly suspicious stems is given in (79):

---

136Stop+ stop is also populous in verbs provided that /t/ and /k/ combine.

137Note that this does not imply that there are no clusters in verbs that violate the Syllable Contact Law, just that those types of clusters are not populous in verbs.

138Note that the highlighted numbers in Table VII are misleading to a degree because not all verb stems exemplifying a given type of moderate or strict cluster necessarily contain a moderate or strict vsi cluster. The number is highlighted if the cluster is attested in at least one moderately or strictly suspicious verb stem.
Many of the strict and moderate verb stems with an intervocalic geminate have an 'onomatopoeic' 'expressive' or 'mimetic' character (cf. Appendix B). We shall not pursue this idea here, but it is possible that these kinds of words have a special phonotactic makeup, perhaps they form another phonotactic stratum (cf. Fudge 1970, Itô & Mester 1995).

Thus, we have seen that the phonotactics of monomorphemic verb stems is more restrictive than the phonotactics of non-verb stems, which manifests itself (i) in the stricter constraints that apply to the coda and (ii) the frequency effects intervocalic consonant clusters display.

It would be desirable to connect (i) of the above to a difference between possible morphologically complex codas available for nouns and for verbs. If we examine word-final codas that consist of a stem-final consonant and a consonant that realises a synthetic suffix, we

---

139 Many of the strict and moderate verb stems with an intervocalic geminate have an 'onomatopoeic' 'expressive' or 'mimetic' character (cf. Appendix B). We shall not pursue this idea here, but it is possible that these kinds of words have a special phonotactic makeup, perhaps they form another phonotactic stratum (cf. Fudge 1970, Itô & Mester 1995).
find that the codas of this structure that can occur at the end of verbs is a subset of those that can occur at the end of nouns. The suffixes in question are the past tense suffix in the case of verbs and the accusative in the case of nouns. Both of them are realised as /t/ when they can form a branching coda with a stem-final consonant, otherwise they are preceded by a ‘linking’ vowel (see the details in section 4.1.2.2). The coda clusters that can be formed in this way are given below for nouns and for verbs:

<table>
<thead>
<tr>
<th>(80)</th>
<th>nouns (accusative)</th>
<th>verbs (past)</th>
</tr>
</thead>
<tbody>
<tr>
<td>st</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>št</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>nt</td>
<td>nt</td>
<td>nt</td>
</tr>
<tr>
<td>n't</td>
<td>n't</td>
<td>n't</td>
</tr>
<tr>
<td>lt</td>
<td>lt</td>
<td>-</td>
</tr>
<tr>
<td>rt</td>
<td>rt</td>
<td>-</td>
</tr>
<tr>
<td>jt</td>
<td>jt</td>
<td>-</td>
</tr>
</tbody>
</table>

As can be seen in (80) the phonotactics of morphologically complex branching codas are very similar in nouns and verbs except that it is more restrictive in verbs, because in a verb the first consonant must be a sonorant: /st, št/ can only occur as codas of this kind in a noun: csempész-t ‘smuggler’ (acc.) vs. csempész-ett, *csempész-t ‘smuggle’ (3sg past indef.); les-t ‘raised hide’ (acc.) vs. les-ett, *les-t ‘spy on sb.’ (3sg past indef.). Notice, however, that although the relationship between non-verb and verb phonotactics is the same in morphologically simplex branching codas and morphologically complex branching codas (in both cases it is the verb phonotactics that is more restrictive), in this case it is not possible to derive the behaviour of the morphologically complex branching codas in verbs from that of the morphologically simplex ones since /st, št/ are well-formed final codas in monomorphemic verbs: /st/ oszt ‘distribute’, /št/ fest ‘paint’ (and /st/ is even the most populous type, cf (72) above). Thus, the restriction on morphologically complex branching codas in verbs must be an extra stipulation.
3.5. Sequence constraints

We have noted that, in addition to SSCs and MSCs, the phonotactic pattern may also be determined by Sequence Constraints, i.e. constraints that apply irrespective of the syllabic (prosodic) or morphological affiliation of segments.

A sequence constraint briefly mentioned in section 3.2.4.2 that the segments of the segments /s, š, t’, č, z, ž, j/ should combine to form CC clusters (including fake geminates). A CC cluster that contains one of these segments in both its positions is excluded independently of the syllabic affiliation of the consonants the clusters consist of (cf. Törkenczy 1994a). Assuming that the segments involved are specified as [+ strident] (redundantly, by Stridency Spell-out, which specifies [COR, -son, + cont] segments (including affricates) as [+ strid]; cf. Siptár & Törkenczy 2000) this can be formulated as (81):

\[(81) \text{The feature } [+ \text{strid}] \text{may not be associated with adjacent segments that have independent root nodes}\]

Since the OCP bans adjacent identical feature specifications within a morpheme (81) rules out strident clusters other than true geminates (which have a single root node associated with two timing slots on the skeletal tier). Apart from a handful of exceptions (81) is only violated by (underlying) clusters occur when divided by an analytic domain boundary: [[húsz][centis]] ‘20-centimetre long’, [[tenyész][csődőr]] ‘stud’, [[hős][cincár]] ‘oak cerambix’, [[has][csikrás]] ‘stomach-ache’, [[kus][szerű]] ‘petty’, [[rács][szerkezet]] ‘grid structure’, [[gáz][ső]] ‘gas pipe’, [[kész][szori tás]] ‘handshake’, [[darázs][cső pés]] ‘wasp-bite’, etc. Since an analytical domain is the largest domain within which phonotactic constraints apply, (81) is truly a sequence constraint since it has to refer neither to the syllable to the morpheme as domains.

---

140 That is $7^2 = 49$ clusters some of which would undergo various assimilations involving voicing and place.

141 See section 3.2.4.2 for an exhaustive list.
Chapter 4.

‘Dynamic’ phonotactics: Phonotactically motivated processes

4.1. Vowel–zero alternations

Hungarian has an intricate system of vowel–zero alternations. We shall see as we proceed that not all of them can be analysed in the same way phonologically. Henceforward, we shall informally refer to any vowel that alternates with zero as ‘unstable’ and denote it with the symbol $V_o$. This term is meant to be neutral with respect to whether a given vowel–zero alternation is considered to be the result of epenthes, vowel deletion or some other phonological mechanism. In this section first we examine the distribution and the quality of unstable vowels and then we present an analysis of vowel–zero alternations.

The unstable vowels may occur stem-internally, i.e. inside a stem, and stem-externally, i.e. between a stem and a suffix.

4.1.1. Stem-internal unstable vowels: ‘epenthetic’ stems

The stem-internal occurrence of unstable vowels is restricted to a non-productive class of stems traditionally called ‘epenthetic’ (e.g. Vago 1980a), e.g. *bokor* ‘bush’ (compare *bokr-ok* ‘bush’ (pl.)). Although we shall refer to them by the traditional name, we make no claim here about (and actually will argue against) their epenthetic character. In general, the unstable vowel of these stems is phonetically expressed if they occur in isolation, or before consonant-initial

---

1 Although it is a non-productive class, it is rather populous: it contains about 150 verbs and 250 noun stems (some of which are morphologically complex). There are also a few adjectives and numerals that are ‘epenthetic’.
suffixes, but it does not appear at the surface if a vowel-initial suffix follows the stem. This is shown in (1):

(1)  
<table>
<thead>
<tr>
<th></th>
<th>C-initial suffix</th>
<th>V-initial suffix</th>
</tr>
</thead>
<tbody>
<tr>
<td>bokor</td>
<td>bokor-ban ‘bush’ (iness.)</td>
<td>bokr-ok</td>
</tr>
<tr>
<td>retek ‘radish’</td>
<td>retek-ben ‘radish’ (iness.)</td>
<td>retk-ek ‘radish’ (pl.)</td>
</tr>
<tr>
<td>kölyök ‘kid’</td>
<td>kölyök-ben ‘kid’ (iness.)</td>
<td>kölyk-ök ‘kid’ (pl.)</td>
</tr>
</tbody>
</table>

The unstable vowels in ‘epenthetic’ stems are regularly short and mid ([−open₁, +open₂]) whose frontness and rounding is determined by vowel harmony. Thus, they exhibit a ternary alternation: o/ö/e (see the examples in (1)) when they are phonetically realized. The front unrounded alternant is phonetically low, but this is due to phonetic implementation (see Siptár & Törkenczy 2000). There are only seven ‘epenthetic’ stems in which the height of the unstable vowel is irregular (high or low). These are shown in (2):

(2)  
|   | a  |  vacak ‘something worthless’ | vack-ot ‘something worthless’ (acc.) |
|   |    | kazal ‘haystack’ | kazl-at ‘haystack’ (acc.) |
|   |    | ajak ‘lip’ | ajk-at ‘lip’ (acc.) |
|   |    | fogazz ‘teethe!’ | fogz-ás ‘teething’ |
| i | Öriz ‘guard’ (3sg pres indef) | Örz-i ‘guards’ (3sg pres def) |
| ü | becssül ‘estimate’ (3sg pres indef) | becsül-és ‘estimate’ (noun) |
| u | bajusz ‘moustache’ | bajsz-ot ‘moustache’ (acc.) |

---

2 Some-vowel initial suffixes behave differently (e.g. terminative -ig, causal-final -ált: bokor-ig, bokor-ált and not *bokr-ig, *bokr-ált). These will be discussed in section 4.1.4.6.

3 In a few ‘epenthetic’ stems with a back unstable vowel, the vowel in the syllable preceding the unstable vowel is front i, í, or é e.g. szirm ‘petal’, szirm-ok ‘petal’ (pl.), ki noz ‘torture’ (3sg pres indef), ki nz-ás ‘torture’ (noun), cāoz ‘aim’ (3sg pres indef), cēdzás ‘aiming’. These behave exactly like other antiharmonic stems with respect to vowel harmony and have analogous underlying representations (i.e. they have an unlinked DOR that can link up to the unstable vowel if it is realized, cf. Siptár & Törkenczy 2000).
Stem-internal unstable vowels always occur in the last syllable of the stem, but the only vowel of a monosyllabic stem may not be unstable. If we examine the final $C_i V_u C_j$ string of ‘epenthetic’ stems (where $V_u$ denotes the unstable vowel, and $C_i$ and $C_j$ are consonants of any kind flanking the unstable vowel), it becomes clear that here the vowel-zero alternation is not phonotactically motivated. The reason is that the stem-final consonant cluster $C_i C_j$ that appears in the stem alternant whose unstable vowel is phonetically unexpressed is not always a phonotactically ill-formed final cluster (cf. Törkenczy 1992). Of course, there are consonants separated by a stem-internal unstable vowel that would make up an ill-formed cluster finally (e.g. the cluster /kr/ that occurs in bokr-ok, for instance), but this is not always the case. (3) shows some examples where the $C_i C_j$ cluster corresponding to the consonants flanking the unstable vowel in an ‘epenthetic’ stem is a possible word-final cluster:

(3)$^4$

<table>
<thead>
<tr>
<th>$C_i V_u C_j$</th>
<th>$C_i C_j$#</th>
</tr>
</thead>
<tbody>
<tr>
<td>viszonoz</td>
<td>vonz</td>
</tr>
<tr>
<td>viszonz-om</td>
<td>vonz</td>
</tr>
<tr>
<td>telek</td>
<td>halk</td>
</tr>
<tr>
<td>tell-ek</td>
<td>halk</td>
</tr>
<tr>
<td>füröd-nek</td>
<td>kard</td>
</tr>
<tr>
<td>fürd-ész</td>
<td>kard</td>
</tr>
<tr>
<td>szerez</td>
<td>borz</td>
</tr>
<tr>
<td>szerz-ünk</td>
<td>borz</td>
</tr>
<tr>
<td>torony</td>
<td>szörny</td>
</tr>
<tr>
<td>torny-ok</td>
<td>szörny</td>
</tr>
<tr>
<td>majom</td>
<td>slejm</td>
</tr>
<tr>
<td>majm-ok</td>
<td>slejm</td>
</tr>
<tr>
<td>bagoly</td>
<td>fogj</td>
</tr>
<tr>
<td>bagly-ok</td>
<td>fogj</td>
</tr>
</tbody>
</table>

Note that most of the $C_i C_j$ clusters in (3) are not only well-formed as word final clusters, but are perfectly well-formed syllable codas as well (/gi/ as in fogj is the only one that is not a well-formed branching coda, although it is a well-formed word-final cluster with /j/ in the

---

We have not included occurring \textit{irregular} CC\# clusters (cf. section 3.2.4.2) in (3) that also occur in C_1V_aC_j strings—but such a stem is perfectly possible phonologically: hypothetical \textit{penecs} could be an ‘epenthetic’ stem with \textit{penec}s- as an alternant before vowel-initial suffixes. Thus, we conclude that vowel-zero alternation is not phonotactically motivated stem-internally. Still, there are certain phonotactic restrictions on the shape of ‘epenthetic’ stems: (i) their unstable vowel is never preceded or followed by a consonant cluster (*CCV_aC, *CV_aCC);\(^6\) (ii) the consonants flanking the unstable vowel are never identical (*C_1V_aC_j if C_i\# C_j\#); and (iii) if they are both obstruents, they are either both voiced or both voiceless *C_1V_aC_j, if C_i, C_j = [-son] and only one of them has a laryngeal node).

It is a lexical property of a stem if it is ‘epenthetic’ or not, i.e. whether it has an unstable vowel or not. There are near-identical pairs of stems such that one of the members of a given pair has an unstable vowel where other member has a stable one. Compare the ‘epenthetic’ stems in (4a) with those in (4b) whose last vowel is stable:

\[^5\]We have not included occurring \textit{irregular} CC\# clusters (cf. section 3.2.4.2) in (3) that also occur in C_1V_aC_j strings. Several such examples exist: e.g. /tk/ \textit{retek}, \textit{rekt-ek}, \textit{Detk} < place name>; /čk/ \textit{mocsok} ‘filth’, \textit{mocsk-os} ‘filthy’, \textit{Recsk} (place name); /sk/ \textit{piszok} ‘dirt’, \textit{piszk-os} ‘dirty’, \textit{maszk} ‘mask’.

\[^6\]There is a single exception: the vowel of the denominal verb-forming suffix -Vz is unstable after the cluster-final stem \textit{hang} ‘sound’: \textit{hang-oz-tat} ‘proclaim‘ (3sg pres indef), \textit{hang-z-om} ‘sound‘ (1sg pres indef). The same suffix-initial vowel is always stable after other cluster-final stems: \textit{folt-oz-om} ‘patch’ (1sg pres indef), \textit{rend-ez-em} ‘put in order’ (1sg pres indef), \textit{bors-ozom} ‘pepper’ (1sg pres indef) and not *\textit{folt-om}, *\textit{rend-om}, *\textit{bors-om}.

\[^7\]It has been suggested (P. Rebrus, personal communication) that this latter constraint is more general and requires that C_i and C_j should not be homorganic. Given the feature system used in this dissertation, this claim is not true. In the following examples of epenthetic stems the consonants flanking V_a have the same place: /tV_a/ \textit{bátor} ‘brave’, /sV_a/ \textit{vászon} ‘canvas’, /tV_a/ \textit{ismáel} ‘repeat’, /zV_a/ \textit{közöl} ‘inform’ etc.
Both epenthesis and deletion have been proposed in the literature (cf. Vago 1980a, Jensen and Stong-Jensen 1988, 1989b (epenthesis); Kornai 1990 (vowel deletion)).

\[(4)\]

\[-C_iV_sC_j\]

| a. terem ‘hall’ | term-ek ‘halls’ | b. perem ‘edge’ | perem-ek ‘edges’ |
| vödör ‘bucket’ | vödr-ök ‘buckets’ | csődör ‘stallion’ | csődör-ök ‘stallions’ |
| szobor ‘statue’ | szobr-ök ‘statues’ | tábor ‘camp’ | tábor-ök ‘camps’ |
| torony ‘tower’ | torny-ök ‘towers’ | szurony ‘bayonet’ | szurony-ök ‘bayonets’ |

The data described above suggest that—unless arbitrary lexical marking is involved—stem-internal vowel-zero alternation is neither due to epenthesis nor to the deletion of a vowel represented on a par with vowels that do not alternate with zero.\(^8\)

The unstable vowel cannot be epenthetic because if the vowelless alternant of an ‘epenthetic’ stem is taken to be underlying, the epenthesis site cannot be predicted, i.e. the clusters that are supposedly broken up by epenthesis cannot be distinguished from those that are not (compare ‘epenthetic’ \(\text{torony}\) and non-alternating cluster-final \(\text{szörny}\)).

Stems like \(\text{csukl-ik} ‘hiccup’\) (pres 3sg indef), \(\text{bűzl-ik} ‘stink’\) (pres 3sg indef), \(\text{vedl-ik} ‘slough’\) (pres 3sg indef), etc. provide a further argument against the epenthesis analysis. These are bound stems that end in ill-formed coda clusters. They can only occur before (surface) vowel-initial suffixes, and their stem-final clusters are never broken up. Before (surface) vowel-initial suffixes, they look like ‘epenthetic’ stems (compare \(\text{csukl-ás} ‘hiccup’\) (noun) with ‘epenthetic’ \(\text{vezekl-és} ‘penitence’\)). However, their paradigms are defective in that they simply do not have the forms in which a suffix (or the lack of it) would render the stem-final cluster unsyllabifiable. Crucially, their stem-final clusters cannot be repaired by epenthesis (*\(\text{csukol-j} ‘hiccup!’\) (imp.)). In contrast to these defective stems, ‘epenthetic’ ones do have the corresponding forms (\(\text{vezekel-j} ‘repent!’\) (imp.)). If we want to distinguish defective stems from ‘epenthetic’ ones representationally, both types cannot be cluster-final, because then, the ill-formed final clusters that are to be broken up by epenthesis cannot be distinguished from those that are not.

On the other hand, the vowelless alternant of an ‘epenthetic’ stem cannot be derived by deletion from an underlying CVC-final form (where \(V\) is represented like other vowels)

---

\(^8\)Both epenthesis and deletion have been proposed in the literature (cf. Vago 1980a, Jensen and Stong-Jensen 1988, 1989b (epenthesis); Kornai 1990 (vowel deletion)).
because (a) it would not be possible to distinguish CVC-final stems that exhibit vowel–zero alternation from those that do not (compare ‘epenthetic’ torony and non-alternating szurony), and (b) it would not be possible to explain why the quality/quantity of the unstable vowel is predictable.

Any successful analysis will have to be able to make a three-way distinction between triplets of stems like torony - szurony - szörny, and must distinguish ‘epenthetic’ stems from ‘defective’ CC-final ones like csukl-.

4.1.2. Stem-external vowel–zero alternations: stem-final unstable vowels and ‘linking’ vowels

Several (types of) suffixes are involved in stem-external vowel–zero alternation. Given that we have defined stem-external vowel–zero alternation as occurring at the boundary between a stem and a suffix, there are two logical possibilities: (a) the stem-final vowel may be unstable, or (b) the suffix-initial vowel may be unstable. Both (a) and (b) occur in Hungarian.

4.1.2.1. Stem-final vowel–zero alternations


9Unstable vowels do not occur in non-initial position in a suffix. Noun-forming (derivational) -alom/-elem might be considered a suffix with an internal unstable vowel: vigilom ‘merry-making’, vigilam-at ‘merry-making (acc.)’, cf. víg ‘merry’; főelem ‘fear’, főelm-et ‘fear (acc.)’, cf. fél ‘be frightened’. This suffix, however, is no longer productive and the morphological complexity of the stems containing it has become obscured. Therefore we do not consider this ‘ending’ a suffix in the stem in which it occurs. (It would also be irregular if analyzed as a suffix since it is clearly derivational, not adjective-forming, but nevertheless it is lowering (cf. section 4.1.3 below)).

10The fact that the form is feketédik and not *feketádik is evidence that it is indeed the stem-final vowel that deletes and not the suffix-initial one. If the suffix-initial vowel had deleted,
‘become black’, cf. *fekete ‘black’) and the denominal noun-forming suffix -ász/-ész (erd-ész ‘forester’, cf. erdő ‘forest’, szől-ész ‘viticiculturist’, cf. szőlő ‘grape’). As can be seen, stems lose their final vowel before these suffixes. It is the special property of these suffixes (and not of the stems) that they cause the loss of the stem-final vowels since the same stems retain their final vowels before other vowel-initial suffixes. Consider the comparative forms barná-bb, lazá-bb, szomorú-bb, feketé-bb, etc. where the stem-final vowels are retained and the initial unstable vowel of the comparative suffix -V bb does not appear at the surface. In the case of stem-final vowel-zero alternation there is no restriction on the quality or quantity of the unstable vowel. It can be high (gömböly-ödik ‘become spherical’, cf. gömbölyű ‘spherical’), mid (fak-i t ‘make pale’, cf. fákó ‘pale’), or low (laz-ul cf. lazó), and it may equally be short or long. The unpredictability of the unstable vowel suggests that the mechanism responsible for this alternation is deletion.

4.1.2.2. Suffix-initial vowel-zero alternation

Suffix-initial vowel-zero alternation is more common and more complex than the stem-final one. It is not restricted to a handful of suffixes: many suffixes begin with an unstable vowel

\[\text{then the stem-final } e \text{ would have had to lengthen by Low Vowel Lengthening (cf. Siptár & Törkenczy 2000) giving } *\text{feketádik.} \]

11There are a few sporadic examples of stem-final vowel-zero alternation with other suffixes too, cf. Rebrus (2000a)

12Note that there are four stems that irregularly lose their stem-final vowels before the comparative suffix and some other vowel-initial suffixes: könnyű ‘easy’ (könny-ebb ‘easier’, könny-en ‘easily’), ifjú ‘young’ (ifj-abb ‘younger’, ifj-an ‘as a youth’), hosszú ‘long’ (hossz-abb, ‘longer’, hossz-an ‘at length’), lassú ‘slow’ (lass-abb ‘slower’, lass-an ‘slowly’). There are also three exceptional nominal stems that lose their final vowels before some vowel-initial suffixes: borjú ‘calf’ (borj-ak ‘calf’ (pl.)), varjú ‘crow’ (varj-ak ‘crow’ (pl.)), ifjú ‘youth’ (ifj-ak ‘youth’ (pl.)).

13Note that the adjective-forming suffixes discussed here (a) may not be added to any adjective and (b) after some vowel-final adjectives a -s- ([š]) is inserted before the suffix, e.g. olcsó-s-i t ‘make cheap’, karcsú-s-i t ‘make slim’, állandó-s-ul ‘become constant’.
(called a ‘linking’ vowel traditionally). Two types of suffixes may be distinguished depending on what motivates the suffix-initial vowel–zero alternation.

Type A. The initial unstable vowel of Type A suffixes is only unrealized when they are added to a vowel-final stem; a phonetically realized vowel is always present after a consonant-final stem, regardless of the identity of the stem-final consonant. Suffixes of this type include -ok/-ek/-ök ‘pl’, -on/-en/-ön ‘spr’, -om/-em/-öm ‘1sg poss’, etc. Verb roots end in a consonant before vowel-initial suffixes in Hungarian. Therefore, on the basis of the presence/absence of vowel-zero alternation alone, in principle, a verbal suffix with an initial stable vowel is indistinguishable from a Type A suffix added to a verb root, since the latter would also not display vowel-zero alternation. We shall consider vowel-initial verbal suffixes as Type A if their initial vowel is short o/e/ö (i.e. the typical unstable vowel), e.g. -ok/-ek/-ök ‘1sg pres indef’, -od/-ed/-öd ‘2sg pres def’, -om/-em/-öm ‘1sg pres def’. This is confirmed by their behaviour after a relative verb stem that ends in a vowel; compare lát-om ‘see’ (1sg pres def) and lát-ná-m ‘see’ (1sg pres cond def); ül-ök ‘sit’ (1sg pres indef) and ül-nék ‘sit’ (1sg pres cond indef). The behaviour of Type A suffixes is illustrated in (5):

(5)  
<table>
<thead>
<tr>
<th>C-final stem</th>
<th>pl</th>
<th>spr</th>
<th>1sg poss</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-final stem</td>
<td>lány ‘girl’</td>
<td>lány-ok</td>
<td>lány-on</td>
</tr>
<tr>
<td>V-final stem</td>
<td>holló ‘raven’</td>
<td>holló-k</td>
<td>holló-n</td>
</tr>
</tbody>
</table>

In these suffixes the presence of a realized unstable vowel is not motivated by the phonotactics of final consonant clusters, i.e. the vowel is not there to ‘repair’ otherwise ill-formed clusters. This can be seen in (6) where the realized suffix-initial unstable vowel appears to ‘break up’ well-formed word-final clusters all of which are also well-formed codas.

---

14 The literature is divided concerning the morphological affiliation of these vowels: there is no agreement as to whether linking vowels are part of the stem, or part of the suffix, or a separate entity between the stem and the suffix (see Antal (1977), Papp (1975) and Rebrus (2000a) and references therein). I will consider them to be part of the suffix.
In some marked cases the -t can also syllabify as an appendix, cf. section 3.2.4.3.

(6) $C_rV_nC_j$ - $C_iC_j$

dal-ok ‘song’ (pl.) halk ‘quiet’
bor-ok ‘wine’ (pl.) park ‘id.’
tan-ok ‘tenet’ (pl.) tank ‘id.’
kér-ed ‘ask’ (2sg pres. def.) téd ‘knee’
ostor-oz ‘whip’ (3sg pres. indef.) borz ‘badger’
csalán-os ‘nettle’ (adj.) gáláns ‘gallant’

As there is no phonotactic interaction between the consonants flanking the unstable vowel, it is reasonable to assume that the mechanism responsible for the vowel-zero alternation here is not epenthesis.

**Type B.** The vowel-zero alternation in Type B suffixes is phonotactically motivated. The accusative (−$V_r$) is the suffix that unquestionably belongs here. In this suffix the unstable vowel is phonetically unrealized if the suffixal consonant can syllabify as (part of) a well-formed coda, i.e. no linking vowel appears after vowels (7a), and stem-final consonants with which $t$ can form a branching coda (7b). Otherwise (7c), there is a vowel preceding the $t$ at the surface:  

(7) a. holló-t ‘raven’ b. ón-t ‘tin’ c. nyom-ot ‘trace’
kocsi-t ‘cart’ lány-t ‘girl’ pad-ot ‘bench’
tevé-t ‘camel’ dal-t ‘song’ kőp-et ‘picture’
kapu-t ‘gate’ sör-t ‘beer’ tök-öt ‘pumpkin’
fő-t ‘head’ baj-t ‘trouble’ hegy-et ‘hill’
tű-t ‘needle’ rő-t ‘gap’ i v-ét ‘arc’
anyá-t ‘mother’ kosz-t ‘dirt’ zsiráf-ot ‘giraffe’
sí-t ‘ski’ gőz-t ‘steam’ doh-ot ‘must’
menü-t ‘menu’ varázs-t ‘magic’ rab-ot ‘prisoner’

Interestingly, a linking vowel is present after an ‘epenthetic’ stem even if the final consonant

---

15In some marked cases the -t can also syllabify as an appendix, cf. section 3.2.4.3.
Type A suffixes typically do not cause this kind of variation in ‘epenthetic’ stems. The unstable vowel of the stem is not realized before such a suffix: bagly-ok ‘owls’ and not *bagoly-ok.

The geminates later degeminate by post-lexical degemination (cf. Siptár & Törkenczy 2000).

Some of the words in (8b) have alternative forms without the linking vowel, in which case the stem internal unstable vowel is phonetically expressed, e.g. bajusz-t/bajsz-ot, öböl-t/öbl-öt. The conditions on this variation are idiosyncratic and often unclear (some stems do not show any variation, others only with certain suffixes and not with others; there is variation across speakers, etc).\textsuperscript{16}

Disregarding the few exceptional cases discussed in 3.2.4.3, regularly, the accusative attaches to stems that end in a consonant cluster with a linking vowel (e.g. rajz-ot ‘drawing’ acc.). This follows from the fact that coda in Hungarian is maximally binary branching. It has to be pointed out, however, that stem-final geminates are peculiar in that some of them behave as if they were single consonants: there is no linking vowel after the stem-final geminates /ss, śś, zz, nn, n’n’, ll, rr, jj/, i.e. those geminates whose short counterparts can form a licit coda with /t/ (e.g. idill-t [idilt] ‘idyll’ (acc.), finn-t [fint] ‘Finnish,’ (acc.), plüss-t [plüšt] ‘plush’ (acc.), dzsessz-t [jɛst] ‘jazz’ (acc.)).\textsuperscript{17}

\begin{table}[h]
\centering
\begin{tabular}{ll}
\hline
(8) & a. non-epenthetic stems & b. ‘epenthetic’ stems \\
\hline
ön-t & haszn-ot ‘profit’ \\
lány-t & torny-ot ‘tower’ \\
dal-t & öbl-öt ‘bay’ \\
sör-t & ökr-öt ‘ox’ \\
baj-t & bagly-ot ‘owl’ \\
rős-t & - \\
kosz-t & bajsz-ot ‘moustache’ \\
gőz-t & - \\
varázs-t & - \\
\hline
\end{tabular}
\end{table}

\textsuperscript{16}Type A suffixes typically do not cause this kind of variation in ‘epenthetic’ stems. The unstable vowel of the stem is not realized before such a suffix: bagly-ok ‘owls’ and not *bagoly-ok.

\textsuperscript{17}The geminates later degeminate by post-lexical degemination (cf. Siptár & Törkenczy 2000).
Note that, although geminates are well-formed branching codas, there is a linking vowel present when the accusative is added to a t-final stem: bot-ot ‘stick’ (acc.), rát-et ‘meadow’ (acc.), öt-öt ‘five’ (acc.), and not *bot-t, *öt-t, *rát-t.

The past tense suffix also belongs to Type B in that the realization of its unstable vowel is also phonotactically motivated, but in some respects it behaves differently from the accusative. The differences are as follows:

(a) Its consonantal part is realized as geminate -tt when it is not adjacent to another consonant (regardless whether the preceding vowel is part of the stem or if it is an unstable vowel that appears at the surface: lő-tt ‘shoot’ (3sg past indef), dob-ott ‘throw’ (3sg past indef). This is usually explained as a result of the difference between the representation of the accusative and the past suffix (e.g. Vago 1980a).

(b) The conditions on when the unstable vowel appears at the surface after consonant-final stems are somewhat different: it can only be unexpressed if the stem-final consonant and -t can form a licit branching coda whose first term is a sonorant. Thus, as opposed to the accusative (9a), only a subset of the possible complex codas are available for syllabification for the past tense suffix (9b):

(9) a. **accusative**

<table>
<thead>
<tr>
<th>stem</th>
<th>meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>ón-t</td>
<td>‘smear’</td>
</tr>
<tr>
<td>lány-t</td>
<td>‘vomit’</td>
</tr>
<tr>
<td>dal-t</td>
<td>‘live’</td>
</tr>
<tr>
<td>sör-t</td>
<td>‘wait’</td>
</tr>
<tr>
<td>baj-t</td>
<td>‘blow’</td>
</tr>
<tr>
<td>rés-t</td>
<td></td>
</tr>
<tr>
<td>kosz-t</td>
<td></td>
</tr>
<tr>
<td>gőz-t</td>
<td></td>
</tr>
<tr>
<td>varázs-t</td>
<td></td>
</tr>
</tbody>
</table>

b. **past**

<table>
<thead>
<tr>
<th>stem</th>
<th>meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>ken-t</td>
<td>‘smear’</td>
</tr>
<tr>
<td>hány-t</td>
<td>‘vomit’</td>
</tr>
<tr>
<td>ē-t</td>
<td>‘live’</td>
</tr>
<tr>
<td>vár-t</td>
<td>‘wait’</td>
</tr>
<tr>
<td>fúj-t</td>
<td>‘blow’</td>
</tr>
<tr>
<td>kés-ett</td>
<td>‘be late’</td>
</tr>
<tr>
<td>csempész-ett</td>
<td>‘smuggle’</td>
</tr>
<tr>
<td>főz-ött</td>
<td>‘cook’</td>
</tr>
</tbody>
</table>

Note that ‘v-adding’ stems behave differently before the accusative and the past tense suffix. The final v of these stems does not appear before the past tense suffix (e.g. lő-tt ‘shoot’ (3sg past indef), fő-tt ‘cook’ (3sg past), rí-tt ‘cry’ (3sg past)), but it does before the accusative (e.g. lov-at ‘horse’ (acc.), köv-et ‘stone’ (acc.)). Cf. section 8.2.1 on ‘v-adding’ stems.
Another idiosyncratic property of the past tense suffix is that no linking vowel appears after two irregular (bound) verb stems: \textit{feküd- ‘lie’} and \textit{alud- ‘sleep’} (\textit{feküd-t ‘3sg past’, alud-t ‘3sg past’}) and after verbs that belong to a class of stems ending in \textit{ad/ed}: e.g. \textit{szalad-t ‘run’} (3sg past), \textit{marad-t ‘remain’ (3sg past), mered-t ‘stand out’ (3sg past), reped-t ‘burst’ (3sg past). These are pronounced with final geminate \textit{[tt]} as a result of Voicing Assimilation, cf. section Siptár & Tőrkczy (2000). Note that some verbs ending in \textit{ad/ed} do not belong to this class: e.g. \textit{ad-ott ‘give’} (3sg past), \textit{fed-ett ‘cover’} (3sg past), \textit{fogad-ott ‘receive’} (3sg past), \textit{szenved-ett ‘suffer’} (3sg past), \textit{tagad-ott ‘deny’} (3sg past). Essentially, the linking vowel does not appear if the final \textit{ad/ed} string can be regarded as a suffix (cf. Rebrus 2000a, compare Rebrus & Trôn in print).

It is not possible to test whether this behaviour of the past tense suffix is really different from that of the accusative since the latter may not be followed by another suffix.

\begin{tabular}{|l|l|l|}
\hline
& \textbf{3sg past indef.} & \textbf{3sg past def.} \\
\hline
nyom ‘push’ & nyom-ott & nyom-t-a \\
rak ‘put’ & rak-ott & rak-t-a \\
vés ‘chisel’ & vés-ett & vés-t-e \\
\hline
\end{tabular}

\textbf{The linking vowel is sometimes present after cluster-final stems even if a vowel-initial suffix follows the past suffix. Three types of behaviour may be distinguished: (i) after some CC-final stems the linking vowel is optional (11a); (ii) after others it is compulsory (11b); and some CC-final stems do not permit a linking vowel in this environment (11c):}

\textsuperscript{19}Another idiosyncratic property of the past tense suffix is that no linking vowel appears after two irregular (bound) verb stems: \textit{feküd- ‘lie’} and \textit{alud- ‘sleep’} (\textit{feküd-t ‘3sg past’, alud-t ‘3sg past’}) and after verbs that belong to a class of stems ending in \textit{ad/ed}: e.g. \textit{szalad-t ‘run’} (3sg past), \textit{marad-t ‘remain’ (3sg past), mered-t ‘stand out’ (3sg past), reped-t ‘burst’ (3sg past). These are pronounced with final geminate \textit{[tt]} as a result of Voicing Assimilation, cf. section Siptár & Tőrkczy (2000). Note that some verbs ending in \textit{ad/ed} do not belong to this class: e.g. \textit{ad-ott ‘give’} (3sg past), \textit{fed-ett ‘cover’} (3sg past), \textit{fogad-ott ‘receive’} (3sg past), \textit{szenved-ett ‘suffer’} (3sg past), \textit{tagad-ott ‘deny’} (3sg past). Essentially, the linking vowel does not appear if the final \textit{ad/ed} string can be regarded as a suffix (cf. Rebrus 2000a, compare Rebrus & Trôn in print).

\textsuperscript{20}It is not possible to test whether this behaviour of the past tense suffix is really different from that of the accusative since the latter may not be followed by another suffix.
Verb stems may end in clusters that are not well-formed codas. With one exception (metsz ‘etch’ (3sg pres indef)) they are bound stems. These stems are defective (cf. section 4.1.1) and typically belong to the -ik class of verbs that have the suffix -ik in present indefinite 3rd person singular instead of zero (cf. Károly 1957, Hetzron 1975, Törkenczy 2000b, 2001b, 2002ab, Rebrus and Törkenczy 1999). Examples include bűzd-ik ‘stink’, vedl-ik ‘slough’, áraml-ik ‘flow’, vonagl-ik ‘writhe’, habz-ik ‘foam’, játasz-ik ‘play’, etc. (It must be pointed out that not all -ik verbs end in clusters that are ill-formed as a coda, e.g. álmod-ik ‘dream’, hull-ik ‘fall’.

<table>
<thead>
<tr>
<th>(11)</th>
<th>3sg past indef</th>
<th>3pl past indef</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. fing ‘fart’</td>
<td>fing-ott</td>
<td>fing-ott-ak / fing-t-ak</td>
</tr>
<tr>
<td>mond ‘say’</td>
<td>mond-ott</td>
<td>mond-ott-ak / mond-t-ak</td>
</tr>
<tr>
<td>told ‘lengthen’</td>
<td>told-ott</td>
<td>told-ott-ak / told-t-ak</td>
</tr>
<tr>
<td>b. fest ‘paint’</td>
<td>fest-ett</td>
<td>fest-ett-ek *fest-t-ek</td>
</tr>
<tr>
<td>látsz-ik ‘seem’</td>
<td>látsz-ott</td>
<td>látsz-ott-ak *látsz-t-ak</td>
</tr>
<tr>
<td>csukl-ik ‘hiccup’</td>
<td>csukl-ott</td>
<td>csukl-ott-ak *csukl-t-ak</td>
</tr>
<tr>
<td>old ‘solve’</td>
<td>old-ott</td>
<td>old-ott-ak *old-t-ak</td>
</tr>
<tr>
<td>c. küld ‘send’</td>
<td>küld-ött</td>
<td>küld-t-ek *küld-ött-ek</td>
</tr>
<tr>
<td>kezd ‘begin’</td>
<td>kezd-ett</td>
<td>kezd-t-ek *kezd-ett-ek</td>
</tr>
<tr>
<td>fedd ‘scold’</td>
<td>fedd-ett</td>
<td>fedd-t-ek *fedd-ett-ek</td>
</tr>
</tbody>
</table>

The phonological shape of a given verb stem only partly determines its behaviour. The linking vowel must be present after the stem-final cluster if it ends in a t (e.g. fest-ett-ek) or if it is not a possible branching coda (e.g. csukl-ott-ak). Otherwise, it is phonologically unpredictable whether a linking vowel is compulsory, optional, or disallowed after a given CC-final verb stem in this environment. Indeed, after some stem-final clusters (e.g. /ld/), all three kinds of behaviour are attested, i.e. some stems have both forms (áld-t-ak/áld-ott-ak ‘bless’ (3pl past indef.)) while others have only one, either with or without the linking vowel (küld-t-ek/*küld-ött-ek ‘send’ (3pl past indef.) vs. *száguld-t-ak/száguld-ott-ak ‘speed’ (3pl past indef.)). This suggests that phonotactically the strings VC1C2VtV and VC1C2tV are both well-formed provided that C2 = /t/ and C1C2 is a well-formed coda. Individual CC-final stems that end in a well-formed coda\(^{21}\) whose second term is not /t/ must bear an arbitrary lexical mark as to

\(^{21}\)Verb stems may end in clusters that are not well-formed codas. With one exception (metsz ‘etch’ (3sg pres indef)) they are bound stems. These stems are defective (cf. section 4.1.1) and typically belong to the -ik class of verbs that have the suffix -ik in present indefinite 3rd person singular instead of zero (cf. Károly 1957, Hetzron 1975, Törkenczy 2000b, 2001b, 2002ab, Rebrus and Törkenczy 1999). Examples include bűzd-ik ‘stink’, vedl-ik ‘slough’, áraml-ik ‘flow’, vonagl-ik ‘writhe’, habz-ik ‘foam’, játasz-ik ‘play’, etc. (It must be pointed out that not all -ik verbs end in clusters that are ill-formed as a coda, e.g. álmod-ik ‘dream’, hull-ik ‘fall’,
csikland-ik ‘tickle’, etc.) There are two defective stems that are not -ik verbs (both are used in the definite conjugation only): kől- ‘doubt’ and sí nyl- ‘suffer’; e.g. kől-em (1sg pres def), sí nyl-i (3sg pres def). Special thanks are due to Attila Novák and Nóra Wenszky for these two elusive items.

22Except some final CCC clusters whose final consonant is syllabified into the appendix (cf. section 3.2.4.3.) and also CCC clusters in general that are divided by an analytic morphological domain boundary (cf. section 3.2.2.). On internal CCC clusters created by the past suffix cf. section 4.1.4.4.
suffixes.

Disregarding the special case of lowering (to be discussed in section 4.1.3 below), in suffix-initial vowel-zero alternation, the quality of the unstable vowel is identical to that involved in stem internal vowel-zero alternation. There is a single suffix the height of whose unstable vowel is irregularly high instead of mid: possessive 1pl -unk/-ünk (e.g. bot-unk ‘our stick’ vs. fű-nk ‘our tree’).

4.1.3. Lowering

We have seen above that the quality of the unstable vowel is regularly mid both stem-internally and suffix-initially. In the latter case, however, the linking vowel is sometimes low instead of mid (compare lowered ház-ak ‘house’ (pl.) with unlowered gáz-ök ‘gas’ (pl.)). We shall refer to this phenomenon as the lowering of unstable vowels. The back alternant of a lowered linking vowel is low a [ɔ] instead of o [œ]. The front alternant is low [ɛ] whose lowness is not the result of phonetic implementation (which interprets mid e as low as well, cf. Siptár & Törkenczy 2000), but is due to lowering. Since lowering makes the linking vowel low, rounding harmony cannot apply to it because Hungarian has no low front rounded vowels (compare lowered szüz-ek ‘virgin’ (pl.) with unlowered bűz-ök ‘smell’ (pl.)). Therefore, the lowered linking vowel only shows the binary alternation a/e instead of the usual unlowered ternary one o/e/ö (cf. Siptár & Törkenczy 2000). In ECH the lowering effect on the quality of the linking vowel can be detected after stems whose last nucleus is back or antiharmonic (e.g. ház-ak, hid-ak ‘bridge’ (pl.)) and after stems whose last nucleus is a front rounded vowel (e.g. szüz-ek, tőgy-ek ‘udder’ (pl.)). After a (non-antiharmonic) stem whose last nucleus is a front unrounded vowel, this lowering effect can only be detected in dialects that retain the distinction between mid [ɛ] and low [ɛ] at the surface; compare kép-et [keːpet] ‘picture’ (acc.) and gyeıp-ét [dɛ`pet] ‘lawn’ (acc.). This distinction is lost in ECH. The lowered and the unlowered front unrounded linking vowel are equally realized as [ɛ]: [keːpet, dɛ`pet]. Thus, vowel quality is not a clue for lowering after these stems. There is, however, a lowering effect that is (partially) independent of vowel quality. After stems that cause lowering (henceforward ‘lowering stems’), Type B suffixes occur with a (lowered) linking vowel even if the stem ends
in a consonant that they can normally attach to *without* a linking vowel. Thus, the following generalization seems to hold:\(^{23}\)

(12) Lowering requires a phonetically expressed unstable vowel.

This is shown with the accusative in (13):\(^{24}\)

<table>
<thead>
<tr>
<th>(13)</th>
<th><strong>normal stem</strong></th>
<th><strong>lowering stem</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>ón-t</td>
<td>tehen-et ‘cow’</td>
<td></td>
</tr>
<tr>
<td>lány-t</td>
<td>hány-at ‘how many’</td>
<td></td>
</tr>
<tr>
<td>dal-t</td>
<td>hal-at ‘fish’</td>
<td></td>
</tr>
<tr>
<td>sör-t</td>
<td>ár-at ‘price’</td>
<td></td>
</tr>
<tr>
<td>baj-t</td>
<td>haj-at ‘hair’</td>
<td></td>
</tr>
<tr>
<td>rés-t</td>
<td>has-at ‘stomach’</td>
<td></td>
</tr>
<tr>
<td>kosz-t</td>
<td>mesz-et ‘lime’</td>
<td></td>
</tr>
<tr>
<td>gőz-t</td>
<td>mőz-et ‘honey’</td>
<td></td>
</tr>
<tr>
<td>varázs-t</td>
<td>darázs-at ‘wasp’</td>
<td></td>
</tr>
</tbody>
</table>

Note that this lowering effect makes it possible to detect lowering even when the quality of the linking vowel is not a clue, i.e. in ECH after stems whose last nucleus is a front unrounded vowel, e.g. *mesz-et, méz-et*, provided they also end a consonant that could form a licit branching coda with a following *t*.

Normally, both Type A and Type B suffixes attach to vowel-final stems without a linking vowel. Given (12), however, a linking vowel is expected to occur after vowel-final lowering stems as well. (14) shows the behaviour of the plural and the modal (both Type B)

\(^{23}\)There are a few irregular stems that *are* lowering, but nevertheless, the accusative attaches to them without a phonetically expressed unstable vowel, e.g. *báj ‘charm’ (báj-ak ‘pl’, báj-t (acc.)), szakáll ‘beard’ (szakáll-ak (pl.), szakáll-(a)t (acc.)), (hegy)oldal ‘(hill)side’ ((hegy)oldal-ak (pl.), (hegy)oldal-t (acc.)). For a complete list cf. Papp (1975).

\(^{24}\)The past morpheme does not show this effect. This, however, is due to the fact that there happen to be no lowering verb stems.
after vowel-final lowering stems:\(^{25}\)

\[\begin{array}{ll}
\text{plural} & \text{modal} \\
\text{városi-ak} & \text{sátáni-an} \\
\text{pesti-ek} & \text{emberi-en} \\
\text{szomorú-ak} & \text{szomorú-an} \\
\text{keserű-ek} & \text{keserű-en} \\
\text{bántó-(a)k} & \text{bántó-(a)n} \\
\text{sértő-(e)k} & \text{sértő-(e)n}
\end{array}\]

As can be seen, the generalization expressed in (12) holds for the items in (14). The only special property of vowel-final lowering stems is that the linking vowel of the plural and the modal suffix is optional after mid vowels. Note that there is no linking vowel if the stem-final vowel is \(e, a, é, á\): fekete ‘black’: feketé-k, feketé-n; durvá ‘rough’: durvá-k, durvá-n; ordenáré ‘vulgar’: ordenáré-k, ordenáré-n; burzsoá ‘bourgeois’: burzsoá-k, burzsoá-n.\(^{26}\) Arguably, this is independent of lowering because the lack of a low linking vowel here can be viewed as an OCP effect. Given the representations of the relevant vowels (cf. Appendix A), forms like *feketé-ek or *durvá-ak would be OCP violations. Thus, one could suggest that, with some idiosyncratic variation, generally, (12) also holds for vowel-final stems unless another constraint (such as the OCP) is violated.

However, the behaviour of some other suffixes with initial unstable vowels (such as the accusative and the superessive) appears to be different in that they attach to the same stems without a linking vowel:


\[^{26}\text{The length of the stem final vowels in forms like feketé-k, durvá-k etc. is due to Low Vowel Lengthening (cf. Siptár & Törkenczy 2000).}\]
It has to be pointed out that adjectives are not turned into nouns by zero derivation, but some adjectives have lexicalized nominal counterparts. Thus, not every adjective has a corresponding (homophonous) noun.

This line of reasoning has been pointed out to me by Péter Rebrus (personal communication).

Note that it is not possible to analyse sósat as só-s-o-at (where -o is a deadjectival noun-forming derivational suffix) because derivational suffixes are non-lowering unless they are

\[
\begin{array}{ll}
\text{accusative} & \text{superessive} \\
\text{városi-t} & \text{városi-n} \\
\text{pesti-t} & \text{pesti-n} \\
\text{sátáni-t} & \text{sátáni-n} \\
\text{emberi-t} & \text{emberi-n} \\
\text{szomorú-t} & \text{szomorú-n} \\
\text{keserű-t} & \text{keserű-n} \\
\text{bántó-t} & \text{bántó-n} \\
\text{sértő-t} & \text{sértő-n} \\
\end{array}
\]

One may try to salvage (12) as a general statement in the following way. The stems in (14) are all adjectives. If we assume that (in contrast to the plural and the modal) case endings can only attach to nouns, then the stems in (15) must be nominal stems since the superessive and the accusative are case endings. Most adjectives are lowering, but only some nouns are, and usually a noun stem corresponding to a lowering adjectival stem is non-lowering.\(^{27}\) Compare, for instance, vörös-ek ‘red’ (pl.) and vörös-ök ‘communist’ (pl.), komikus-ak ‘comical’ (pl.) and komikus-ök ‘comedian’ (pl.), szárnyas-ak ‘winged’ (pl.) and szárnyas-ök ‘poultry’ (pl.), etc. One could claim then that the stems in (15) are simply non-lowering nouns that have corresponding adjectives that are lowering. Given this assumption, their behaviour in (15) conforms to (12).\(^{28}\)

Unfortunately, however, the argument above is untenable. (i) There is evidence that case endings can attach to adjectives since they can follow (overt) adjective-forming derivational suffixes: e.g. só-s-at ‘salty’, (acc.) (where -s is a denominal-adjective forming derivational suffix).\(^{29}\) (ii) The initial unstable vowel of some suffixes is unexpressed even after

\(^{27}\)It has to be pointed out that adjectives are not turned into nouns by zero derivation, but some adjectives have lexicalized nominal counterparts. Thus, not every adjective has a corresponding (homophonous) noun.

\(^{28}\)This line of reasoning has been pointed out to me by Péter Rebrus (personal communication).

\(^{29}\)Note that it is not possible to analyse sósat as só-s-o-at (where -o is a deadjectival noun-forming derivational suffix) because derivational suffixes are non-lowering unless they are
vowel-final stems that are unquestionably adjectives (and lowering). Consider the comparative suffix \(-V_{\nu}bb\), which is a Type A suffix just like the plural (e.g. piros-abb), and can only attach to adjectives, but whose unstable vowel never appears phonetically after lowering (or non-lowering) vowel-final stems (compare szomorú-abb, keserű-abb, *szomorú-abb, *keserű-ebb with szomorú-ak, keserű-ek, *szomorú-k, *keserű-k).

To sum up, (12) must be restricted to consonant-final stems:

\[
(16) \text{Lowering requires a phonetically expressed unstable vowel.}
\]

Condition: \(V_{\nu}\) is preceded by a C

After vowel-final lowering stems, suffixes with an unstable initial vowel behave idiosyncratically. Some suffixes attach to these stems with a (lowered) linking vowel, others without a linking vowel. As it is unpredictable which suffix will behave in which way, lexical marking must be involved. It must be pointed out, though, that the behaviour of linking vowels after vowel-final stems is not completely unrelated to lowering since —although the linking vowel does not always appear after a vowel-final lowering stem — if a linking vowel does appear after a vowel-final stem, it is always a lowered one (i.e. there is never a linking vowel after a vowel-final non-lowering stem).

Nominal vowel-final stems are non-lowering in general. The only exceptions are férfi ‘man’ and -fi ‘-man’ (as in e.g. hadfi ‘warrior’): fárfi-ak ‘man’ (pl.) in which the linking vowel is present (compare fárfi-t (acc.)).

According to its source, two types of lowering may be distinguished (cf. Rebrus and Polgárdi 1997):

(a) The source may be the preceding stem, i.e. after stems belonging to an arbitrary (closed) class (that of ‘lowering stems’, cf. Vago 1980a) the immediately following linking vowel is low. Data in (17) show that it is indeed the stem that is the source of lowering in this case, since the linking vowel of one and the same initial suffix shows up as mid after non-

adjective-forming (see the discussion at the end of this section).
lowering stems and as low after lowering ones:  

(17)  

<table>
<thead>
<tr>
<th></th>
<th>plural</th>
<th>accusative</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. <strong>normal stems</strong></td>
<td>bot</td>
<td>bot-ok</td>
<td>bot-öt</td>
</tr>
<tr>
<td></td>
<td>gyep</td>
<td>gyep-ek</td>
<td>gyep-öt</td>
</tr>
<tr>
<td></td>
<td>köd</td>
<td>köd-ök</td>
<td>köd-öt</td>
</tr>
<tr>
<td>b. <strong>lowering stems</strong></td>
<td>fog</td>
<td>fog-ak</td>
<td>fog-at</td>
</tr>
<tr>
<td></td>
<td>kép</td>
<td>kép-ek</td>
<td>kép-öt</td>
</tr>
<tr>
<td></td>
<td>szög</td>
<td>szög-ek</td>
<td>szög-öt</td>
</tr>
</tbody>
</table>

Some suffixes cause lowering as well, i.e. there are suffixes that turn any stem into a lowering one: *gáz-ok-at* ‘gas-pl-acc’, *bűn-öm-et* ‘sin-my-acc’. Compare *hat-od-ot* ‘sixth’ (acc.) where the fraction forming suffix *-od/-ed/-öd* does not cause lowering. We shall return to the problem of multiple suffixation and lowering below.

(b) Some suffixes appear to be ‘self-lowering’ in that the source of the lowering of the suffix-initial unstable vowel is the suffix itself. The suffixes involved are *-sz/-asz/-esz* (2sg pres. indef.), *-ni/-ani/-eni* (inf.) and *-lak/-lek/-alak/-elek* (1sg 2sg/pl), *-nak/-nek/-anak/-enek* (3pl pres. indef.), and *-na/-ne/-ana/-ene* (cond.). The initial linking vowel of these suffixes is low even after stems that are demonstrably non-lowering. This is shown in (18) where the non-self-lowering suffix *-k/-ok/-ek/-ök* (1sg pres. indef.) is included to show the contrast with their behaviour:

---

30 The linking vowel of the superessive behaves differently: it remains mid (and displays a ternary alternation) even after lowering stems: *ház-on* ‘house’ (spr.), *mé-en* ‘honey’ (spr.), *szűz-ön* ‘virgin’ (spr.).

31 Here, the lowering effect can only be detected in dialects that retain the distinction between mid [ɛ] and low [ɛ].
These suffixes (and -tok/-tek/-tök/-otok/-etek/-ötök (2pl pres. indef.) which behaves in the same way with respect to vowel–zero alternation, but is not self-lowering) have been analysed to form a separate class intermediate between analytic and synthetic (quasi-analytic suffixes). The special behaviour of these suffixes is considered idiosyncratic (non-phonological allomorphy) in the present analysis. For a different approach cf. Rebrus & Törkenczy (1999), Rebrus (2000b), Törkenczy (1998b), Törkenczy & Siptár (1999a).

In general, the initial unstable vowel of these suffixes appears if the stem ends in more than one consonant. Note, however, that they often behave idiosyncratically. In some forms the linking vowel is (unexpectedly) optional (e.g. mond-(a)sz vs. sért-esz). After some stems ending in a geminate, the linking vowel does not appear (e.g. áll-ni ‘to stand’ vs. hall-ani ‘to hear’). On the other hand, it does appear after some stems that end in a long vowel followed by a single consonant (bocsát-ani ‘to forgive’ vs. lát-ni ‘to see’). We argued in section 3.2.4.3 that—despite the phonological conditioning—the selection of the allomorphs of 2sg pres indef -sz is morphological rather than phonological. Similarly, we suggest that it is allomorphy that is involved in what appears to be low vowel–zero alternation in ‘self lowering’ suffixes. Considering the selection morphological has significant advantages. It makes it possible to maintain the generalization in (16) and this way phonological lowering will always have a local source: the preceding (relative or absolute) stem. Also, if the alternation is morphological, idiosyncrasies of the kind described above are more likely to occur.

Since ‘self-lowering’ is non-phonological, phonological lowering always spreads from the stem. Lowering can only influence the linking vowel immediately following the lowering stem. Consider nyolc-ad-ot ‘eighth’ (acc.) where nyolc is a lowering stem, but the accusative surfaces with an unlowered vowel (-ot). As we have pointed out above, the source of lowering may be the relative stem, i.e. some suffixes may be lowering, and can turn a non-lowering stem into a lowering one. While lowering is a lexical property of (nominal) stems, it is not (completely) unpredictable which suffix is lowering and which one is not. The claim

---

32 These suffixes (and -tok/-tek/-tök/-otok/-etek/-ötök (2pl pres. indef.) which behaves in the same way with respect to vowel–zero alternation, but is not self-lowering) have been analysed to form a separate class intermediate between analytic and synthetic (quasi-analytic suffixes). The special behaviour of these suffixes is considered idiosyncratic (non-phonological allomorphy) in the present analysis. For a different approach cf. Rebrus & Törkenczy (1999), Rebrus (2000b), Törkenczy (1998b), Törkenczy & Siptár (1999a).

33 See, however, the part-of-speech distribution of lowering to be discussed below.
in Vago (1980a) that lowering is positionally predictable, i.e. that the unstable vowel of first position suffixes (relative to the absolute stem) may or may not lower (depending on whether the absolute stem is lowering or not), but it is always low later (i.e. when not immediately adjacent to the stem) can be shown to be incorrect: for instance, the unstable vowel of the accusative does not lower after the fraction forming suffix: hat-od-ot ‘sixth’ (acc.), see Kornai (1990). According to Rebrus and Polgárdi (1997) derivational suffixes are non-lowering and inflectional ones are lowering. They point out contrasting pairs of examples like un-tat-om ‘make sb bored’ (1sg pres) and un-t-am ‘be bored’ (1sg past). In the former word, the causative derivational suffix (-tat) does not make the unstable vowel of the personal suffix low, but in the latter, the inflectional past tense suffix does. While it is true that inflectional suffixes are all lowering, the correlation between lowering and derivational suffixes is more complicated. Superficially, it seems that Rebrus and Polgárdi’s claim does not hold for all derivational suffixes: they may be lowering or non-lowering—compare só-s-ak ‘salty’ (pl.) (where -s is a denominal adjective forming derivational suffix) and harc-os-ok ‘warrior’ (pl.) (where -Vs is a noun-forming derivational suffix). This, however, is due to a factor independent of the inflectional/derivational character of suffixes. It can be explained with reference to the part-of-speech distribution of lowering. Note that the derivational suffix above that appears to be lowering is adjective-forming. In fact, all the derivational suffixes that lower are adjective-forming as well (e.g. tanul-ékony-a ‘teachable’ (pl.), ír-ott-at ‘written’ (acc.)).

As adjectives are generally lowering, we claim that these suffixes lower because they are adjective-forming. The distribution of lowering items in other word classes is the following. Nouns and pronouns include a large (but closed) set of lowering stems. It is unpredictable which nominal stems are lowering so they must be marked as such in the lexicon. New items

---

34 This amounts to the claim that all suffixes are lowering.

35 The lowering effect of inflectional suffixes that are never followed by another suffix (e.g. case endings) never actually manifests itself, but we assume that they are lowering as well.

36 The privative suffix -(V)tlan/-tlen/-talan/-telen and -van/-ven ‘-ty’ are exceptional. The former is adjective forming, but does not lower: e.g. tanul-atlan-ok ‘uneducated’ (pl.), and the latter is derivational, not adjective forming, but does lower: e.g. nyolec-van-at ‘eighty’ (acc.).

37 There are a few (irregular) exceptions, e.g. nagy-ok ‘big’ (pl.), agg-ok ‘old’ (pl.).
entering the language are invariably of the unmarked, non-lowering type (e.g. *szponzor-ok* ‘sponsors’, *kűr-ök* ‘free exercises [in figure skating]’). Absolute verb stems are never lowering. The attested distribution of lowering in stems and suffixes follows if we assume that they are non-lowering by default, but some are marked as lowering individually/idiosyncratically, while others are assigned lowering status by two morphological redundancy rules:

(19) a. Inflectional suffixes are lowering.
    b. Adjectives are lowering.

In addition, stems that belong to certain (more-or-less irregular) morphological classes are always lowering. For instance nominal ‘v-adding’ stems and ‘shortening’ stems all lower: e.g. *ló* ‘horse’ *lov-at* ‘horse’ (acc.), *madár* ‘bird’ *madar-at* ‘bird’ (acc.).³⁸ Thus, (19) may be extended with rules referring to specific morphological classes. To sum up, lowering is partly predictable on a morphological basis.

Trón & Rebrus (2001) argue that there is a correlation between the complexity or ‘markedness’ of stem-final clusters and the lowering character of (noun) stems. They claim that ‘the more marked a certain cluster is, the more likely it is that it is lowering’ (p. 25). While this seems to be true in some cases (e.g. all noun stems that end in */rj/* (which is definitely marked since it monomorphemically violates Sonority Sequencing) lower: *fürj-et* ‘quail’ (acc.), *sarj-at* ‘descendant’ (acc.)), it is far from being exceptionless³⁹ and would require a rather delicate (arbitrary?) ranking of constraints into a markedness or complexity scale to make it work. Stem-final */kt/* (*akt-ot* ‘nude’ (acc.)) */tk/* (*barack-ot* ‘peach’ (acc.), *palack-ot* ‘bottle’ (acc.), *tarack-ot* ‘howitzer’ (acc.)), */pf/* (*copf-ot* ‘plait’ (acc.)), */pʃ/* (*taps-ot* ‘applause’ (acc.)) do not lower—although, presumably, they are complex/marked clusters. */dv/*

---
³⁸Note that ‘epenthetic’ stems are not involved in such an implicational relationship. They may be lowering or non-lowering: compare *marok* ‘fist’ *mark-om* ‘my fist’ with *farok* ‘tail’ *fark-am* ‘my tail’.

³⁹Note that this is not necessarily a problem for Trón & Rebrus (2001) since their approach does not work with black-and-white constraints, see also Rebrus & Trón (2002).
—also a marked cluster —sometimes lowers (*kedv-et* ‘mood’ (acc.), *nedv-et* ‘fluid’ (acc.)),

sometimes does not (*üdv-öt* ‘salvation’ (acc.)). The same is true of /ld/ —presumably a non-
complex/less marked cluster: *hold-at* ‘moon’ (acc.), *föld-et* ‘ground’ (acc.), but *zsold-ot*
’soldier’s pay’ (acc.). Finally, *all* monomorphemic CCC clusters, which are certainly marked
in Hungarian, are non-lowering: /mps/ *mumpsz-öt* ‘mumps’ (acc.), /nst/ *dunszt-öt* ‘steam’
(acc.), /rst/ *karszt-öt* ‘karst’ (acc.), /šč/ *borscs-öt* < Russian soup> (acc.).

Only an unstable vowel may be the target of lowering. Stable suffix initial vowels
never lower. As pointed out above, the superessive suffix is special because it is unstable, but
it does not lower. The three types of suffixes are illustrated in (20). The plural is lowerable
and unstable, the causal-final is non-lowerable and stable and the superessive is non-lowerable
and unstable:

(20)

<table>
<thead>
<tr>
<th></th>
<th>plural</th>
<th>causal-final</th>
<th>superessenve</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>normal stem</td>
<td>fá-k</td>
<td>fá-ért</td>
<td>fá-n</td>
<td>‘tree’</td>
</tr>
<tr>
<td>gáz-ok</td>
<td>gáz-ért</td>
<td>gáz-on</td>
<td></td>
<td>‘gas’</td>
</tr>
<tr>
<td>lowering stem</td>
<td>ház-ak</td>
<td>ház-ért</td>
<td>ház-on</td>
<td>‘house’</td>
</tr>
</tbody>
</table>

---

40 This is undetectable (or, equivalently, they are non-lowering) in ECH.

41 This is certainly related to the fact that all these words are recent loans and lowering
stems are a closed class.

42 The fourth type does not exist: there are no suffixes with an initial lowerable stable
vowel.
4.1.4. Analysis

4.1.4.1. Syllabification—full vowels and defective vowels

Hungarian vowel-zero alternation is partly due to an underlying difference between full vowels ($V_f$) and defective vowels ($V_d$). Defective vowels are empty in the sense that they only consist of a skeletal slot without any segmental melody. By contrast, full vowels minimally have a VOCALIC node (and a ROOT node). This is shown in (21) below (where non-essential structure between the ROOT and the VOCALIC node is suppressed and the symbol $\Delta$ denotes a structure of any complexity, or nil):

(21) a. full vowel  
\[
\begin{array}{c}
N \\
| \\
X \\
| \\
root \\
| \\
vocalic \\
\end{array}
\]

b. defective vowel  
\[
\begin{array}{c}
N \\
| \\
X \\
| \\
\end{array}
\]

Note that defective vowels (as opposed to full ones) will have to be marked underlyingly in some way to syllabify as nuclei since otherwise the syllabification algorithm will not be able to identify them as such. We simply assume that they are prelinked to nucleus nodes.

Defective vowels are not interpreted phonetically unless they receive a vocalic node (i.e. are turned into full vowels) in the course of the derivation. This default process only targets a licensed $V_d$, i.e. one that is incorporated into a syllable, and turns it into a minimally full vowel. Default $V$ achieves this by assigning [+ open$_2$] to licensed defective vowels. Higher nodes (including the vocalic node and the root node) are automatically appended to ensure well-formedness (Sagey 1986, Clements and Hume 1995). This is shown in (22), where

---

irrelevant structure has been suppressed.44

\[(22) \textit{Default V}\]

\[
\begin{array}{c|c|c}
\sigma & \sigma \\
\hline
N & N \\
X & X \\
\hline
\text{(root)} & \text{(vocalic)} & [+ \text{open}_2]
\end{array}
\]

Vowel Harmony and the other default processes apply to the output of Default V to derive the correct surface vowel quality.

In accordance with standard assumptions about prosodic licensing, we assume that prosodically unlicensed material does not receive phonetic interpretation (cf. for instance Selkirk 1981, Itô 1986, 1989). Thus, unsyllabified defective vowels might persist up to the level of surface representation without being phonetically realized (i.e. they need not be stray-erased). Alternatively, they may be assumed to delete at some point in the derivation (possibly at the end of Block 1 or lexical phonology). Note that, if unsyllabified $V_d$-s persist in the postlexical phonology, then the locality conditions of postlexical processes (e.g. Degemination, Voicing Assimilation) have to be determined in such a way that they ignore defective vowels since in a string $C_1V_dC_2$, $C_1$ and $C_2$ should count as adjacent for assimilation. In section 4.1.4.5. below we argue that they are invisible in Block 2. We express this by supulating that they are erased at the end of Block 1. This will also simplify the statement of postlexical phonological rules.

We also assume that defective vowels are restricted in occurrence compared to full ones. Notably, they can only occur in singly closed syllables. Thus, the following constraints

\[\text{44The parentheses enclosing the root node and the vocalic node in (22) indicate that these nodes have been automatically appended. In this figure encircling indicates that a node dominates no structure.}\]
are added to the well-formedness conditions defining the Hungarian syllable template:

\[(23)\] \begin{align*}
    \text{a. } & * V_d ]_o \\
    \text{b. } & * V_d CC ]_o
\end{align*}

Disregarding non-essential structure, full vowels minimally have a vocalic node. Minimally full vowels receive place and aperture values by vowel harmony and default processes (in the way described in Siptár & Törkenczy 2000, Siptár 2002)

In addition to the lexical difference between full and defective vowels, vowel–zero alternation is also due to syllabification. We follow Itô (1986, 1989) and assume that syllabification is a template-matching algorithm. Template matching is directional (left-to-right or right-to-left), maximal (i.e. the syllable template is filled up with segmental material maximally) and is constrained by the Onset Principle (i.e. onsetless syllables are avoided if possible). Syllabification and epenthesis are not separate processes in that syllabification can build degenerate syllables, i.e. syllables that contain nodes dominating empty X-slots. Thus, syllabification may overparse segmental material by inserting empty positions. However, we only allow overparsing by empty nuclear positions. (24) is intended as a language-specific restriction on syllabification in Hungarian:\(^{45}\)

\[(24)\] Empty onset or coda positions may not be created in the course of syllabification.

Syllabification may be non-exhaustive (cf. Hyman 1990, Kenstowicz 1994). We assume that this can happen under the special condition given below:

\[(25)\] \textit{Non-exhaustiveness:} Defective vowels may remain unparsed into syllables.

Thus, a representation is well-formed even if it contains unparsed defective vowels.

As a result of (24) and (25) syllabification will skip lexical empty nuclear positions (defective vowels) should it be impossible for them to syllabify in a singly closed syllable.

\(^{45}\)Compare Itô ’s analysis of Axininca Campa (Itô 1989) where syllabification inserts empty onset positions as well.
This can happen word-finally, prevocally, or before a single consonant followed by a full vowel:

\[
(26) \quad \begin{align*}
\text{a. } & \ldots CV_{t}CV_{d}\# \rightarrow \ldots \{CV_{t}C\}V_{d}\# \\
\text{b. } & \ldots V_{t}CV_{d}V_{f} \rightarrow \ldots V_{t}\{CV_{d}\}V_{f} \\
\text{c. } & \ldots V_{t}CV_{d}CV_{f} \rightarrow \ldots V_{t}\{CV_{d}\}V_{f}\{CV_{f}\}
\end{align*}
\]

As is shown in (26) (where the syllable edges are indicated by curly brackets) the empty nuclear positions in question remain unaffiliated syllabically. They are not ‘rescued’ by syllabification creating a coda position after them because this is excluded by (24).\(^{46}\) Non-exhaustiveness together with the Onset Principle (cf. Itô 1989) ensures that \(.V_{t}CV_{d}CV_{f}\ldots\) strings syllabify as in (26c) rather than like this \(*.V_{t}\{CV_{d}\}V_{f}\ldots\), i.e. since defective vowels may be left unparsed, it is more important to obey the Onset Principle than to parse a defective vowel.

An empty nuclear position created by syllabification is representationally identical with a lexically empty position. Both are defective vowels in the sense defined above and may be turned into full vowels by (22).

In Hungarian, syllabification proceeds from right to left and is continuous, i.e. it (re)applies after morphological and phonological operations. Resyllabification is permitted, i.e. prosodic structure is erased if the nucleus is deleted along with its X-slot (cf. Hayes 1989) and the coda of a stem-final syllable becomes available for (re)syllabification if a vowel-initial suffix is added (compare Levin 1985).\(^{47}\)

In addition to prosodic structure, vowel–zero alternation is also sensitive to morphological structure. In section 2.2. a distinction was made between analytic and synthetic suffixation and it was pointed out that phonotactic constraints do not apply across the boundary of an analytic domain. The distinction between these two kinds of suffixes is crucial in the interpretation of vowel–zero alternation. We assume that both Type A and Type B suffixes are

\(^{46}\)A constraint disallowing completely empty syllables (i.e. syllables that have a defective vowel and an empty onset and/or coda) would have the same effect as (24).

\(^{47}\)Analytic suffixes behave differently, see the discussion below.
Both these treatments are formulated in a Government Phonology framework.

4.1.4.2. Major stems and ‘epenthetic’ stems—Type A and Type B suffixes

In section 4.1 we saw that vowels can alternate with zero stem-internally, stem-finally and suffix-initially. Of these, stem-final vowel-zero alternation is phonologically irregular in that (i) only an arbitrary set of suffixes trigger it in all the stems to which they are attached (cf. section 4.1.2.1), and (ii) an arbitrary set of stems undergo it before an arbitrary set of suffixes (only some of which belong to the set referred to in (i), cf. section 4.1.2.1 Note 12). Therefore, we assume that stem-final vowel-zero alternation is essentially morphological and we shall disregard it in the analysis below.

The stems and suffixes showing regular (phonological) vowel-zero alternation have the following underlying representations. We claim that ‘epenthetic’ stems do not end in consonant clusters (CJC) as is usually assumed (e.g. Vago 1980a, Jensen and Stong-Jensen 1988, 1989, Törkenczy 1994a, 1995), but contain a defective vowel in their final syllable that ends in a single consonant: -CVC# (compare Törkenczy 1992 and Ritter 1995). Thus, the three-way distinction between the last syllables of the triplets described in section 4.1.1. is made representationally in the following way: -CVaC (torony ‘tower’), -CVbC (szurony ‘bayonet’), -CC (szörny ‘monster’). Type A suffixes have an underlying initial full vowel (e.g. -Vk ‘pl’) and Type B suffixes are underlyingly consonant-initial (e.g. -t ‘acc’).

Let us now examine the relationship between vowel-zero alternation and syllabification. Figure (27) shows how consonant-final non-lowering non-epenthetic stems (i.e. major stems) are syllabified when Type A suffixes (27a) and Type B suffixes are attached to them:49

48 Both these treatments are formulated in a Government Phonology framework.

49 In the transcriptions the vowels denoted by phonetic characters and the symbol \( V_f \) are all full vowels. \( V_f \) is a minimal full vowel (i.e. one that at least has a vocalic node (and a root node)). The real difference in syllabification is between all full vowels vs. \( V_d \).
The syllabifications follow from the Hungarian syllable templates and right-to-left template matching and non-exhaustiveness. Type A suffixes are insensitive to the identity of the stem-final consonant because they are vowel-initial and thus they can always form a well-formed syllable with the stem-final consonant. Type B suffixes, on the other hand, are phonotactically sensitive to the stem-final consonant. The reason is that they are consonant-initial and unsyllabifiable in themselves: accusative -t syllabifies with the stem-final consonant just in case they can form a licit coda (bort). Syllabification creates a degenerate syllable, i.e. a syllable with an empty nuclear position V if the cluster is not syllabifiable as the coda of the last syllable of the stem (képVₜ). Thus, the vowel-zero alternation (‘epenthesis’ in this case) is due to syllabification.

The behaviour of Type A suffixes after vowel-final stems shows that, in addition to syllabification, there is a rule which is responsible for the vowel-zero alternations. This rule eliminates hiatus by (i) deleting a defective vowel when it is adjacent to a full one, and (ii) deleting a full vowel (together with its X-slot) when it follows another full vowel. It is necessary to delete the X-slot too, because simply deleting the segmental melody would only turn a full vowel into a defective one. This rule can be formulated as in (28ab) where V is the vocalic node.

(28) 

\begin{align*}
(27) & \text{a. kep-Vₜk} \quad \{\text{ke:}\} \{pVₜk\} \quad \text{képek} \, \text{‘picture’ (pl.)} \\
& \text{bor-Vₜk} \quad \{\text{bo}\} \{rVₜk\} \quad \text{borok} \, \text{‘wine’ (pl.)} \\
& \text{b. kep-t} \quad \{\text{ke:}\} \{pVₜt\} \quad \text{képet} \, \text{‘picture’ (acc.)} \\
& \text{bor-t} \quad \{\text{bort}\} \quad \text{bort} \, \text{‘wine’ (acc.)}
\end{align*}
Of the four combinations VV does not arise.

On postlexical hiatus filling see Siptár & Tőrкenczy (2000).

In the discussion of syllabification we shall first temporarily disregard the representation and the syllabification of lowering stems and suffixes for expository reasons. We shall deal with lowering in detail later in section 4.1.4.3. Due to the representation of lowering stems/suffixes some of the syllabifications that follow will have to be modified.

We shall discuss the third type of suffix combination (TB+ TA) together with the past tense suffix later in section 4.1.4.4. The double consonant in (29) means a true geminate (i.e. two timing slots associated with a single root) and not adjacent identical melodies.
The syllabification of these forms is straightforward. In the second example a degenerate syllable is created because /kt/ is not a possible (regular) coda.

In the examples discussed above empty nuclear positions (defective vowels) are created in the course of syllabification. We have also pointed out, however, that defective vowels are underlingly present in ‘epenthetic’ stems. The vowel-zero alternation in epenthetic stems is not the result of overparsing by syllabification, but to the special constraints (23ab) on syllables whose nucleus is a defective vowel. Figure (30) shows how the words *szurony*, *szörny*, and *torony* syllabify in isolation (30a), when suffixed by Type A suffixes (30b) and Type B suffixes (30c):

\[
\begin{align*}
(30) \quad & \text{a. suron}^\nu & \{su\} \{ron^\nu\} \quad \text{szurony} \\
& \text{sörn}^\nu & \{sörn^\nu\} \quad \text{szörny} \\
& \text{torV}_d^n & \{to\} \{rV_d^n\} \quad \text{torony}
\end{align*}
\]

\[
\begin{align*}
& \text{b. suron}^\nu \text{-V}_k & \{su\} \{ro\} \{n^\nuV_k\} \quad \text{szurony-ok} \\
& \text{sörn}^\nu \text{-V}_k & \{sör\} \{n^\nuV_k\} \quad \text{szörny-ek} \\
& \text{torV}_d^n \text{-V}_k & \{tor\}V_d \{n^\nuV_k\} \quad \text{torny-ok}
\end{align*}
\]

\[
\begin{align*}
& \text{c. suron}^\nu \text{-t} & \{su\} \{ron^\nu t\} \quad \text{szurony-t} \\
& \text{sörn}^\nu \text{-t} & \{sör\} \{n^\nuV_d t\} \quad \text{szörny-et} \\
& \text{torV}_d^n \text{-t} & \{tor\}V_d \{n^\nuV_d t\} \quad \text{torny-ot}
\end{align*}
\]

Szörny can syllabify as a CVCC syllable because the final consonant cluster can form a licit coda. There is no difference between the syllabification of *szurony* and *torony* in isolation because the defective vowel in the latter can syllabify in a singly closed syllable (30a). However, the same two stems do not syllabify in the same way when suffixed with a Type A suffix. As can be seen in (30b) the last vowel of *szurony* can syllabify as the nucleus of an open syllable in *szuronyok* (since it is a full vowel). By contrast, the second vowel of *torony*
It is sometimes claimed (e.g. Vago 1980a, Törkenczy 1992) that alternative forms exist in the accusative if the last consonant of the ‘epenthetic’ stem can form a licit coda with the following \(-t\): e.g. öböl-t/öbl-öt ‘bay’ (acc.). As pointed out in 4.1.2.2, this is not true of all ‘epenthetic’ stems: only some show this variation (cf. Papp 1975). We assume that for those that do, there are two entries in the lexicon: an ‘epenthetic’ one and a major one. Given this assumption, syllabification will yield the alternative forms: \{ö\} {bV}_{d}t vs. \{öb\} {V}_{d}t. The selection of one or the other entry is often idiosyncratic. Moreover, different suffixes may select different entries: öböl-t/öbl-öt ‘bay’ (acc.) vs. öbl-ök but *öböl-ök ‘bay’ (pl.)

It is sometimes claimed (e.g. Vago 1980a, Törkenczy 1992) that alternative forms exist in the accusative if the last consonant of the ‘epenthetic’ stem can form a licit coda with the following \(-t\): e.g. öböl-t/öbl-öt ‘bay’ (acc.). As pointed out in 4.1.2.2, this is not true of all ‘epenthetic’ stems: only some show this variation (cf. Papp 1975). We assume that for those that do, there are two entries in the lexicon: an ‘epenthetic’ one and a major one. Given this assumption, syllabification will yield the alternative forms: \{ö\} {bV}_{d}t vs. \{öb\} {V}_{d}t. The selection of one or the other entry is often idiosyncratic. Moreover, different suffixes may select different entries: öböl-t/öbl-öt ‘bay’ (acc.) vs. öbl-ök but *öböl-ök ‘bay’ (pl.)

The syllabification of multiply suffixed ‘epenthetic’ stems is unproblematic and follows from the mechanism discussed above:
4.1.4.3. Syllabification and lowering

We noted in section 4.1.3. above that Lowering (a) determines the quality of unstable suffix-initial vowels and (b) interacts with syllabification (cf. (16)). We also saw that the source of lowering in the representation of a stem or a suffix may be an unpredictable ‘mark’ or one which is morphologically predictable (cf. (19)). We claim that, representationally, this mark consists of two distinct (though always cooccurring) characteristics. A lowering stem/suffix has a final floating [+ open,] feature and a morpheme-final defective vowel $V_d$ as shown in (32):

\[
\begin{array}{c}
\text{N} \\
\text{X} \\
\text{X} \\
\text{root} \\
\Delta \\
\text{[+ open,]}
\end{array}
\]

Thus, specially marked nouns and pronouns, all adjectives, and all inflectional suffixes end in a structure shown in (32). This means that some of the representations discussed above have to be modified because they are, or they contain, such formatives. For instance, szörny must be represented as /sörn'$V_d/' instead of /sörn'/ because it is an (unpredictably) lowering noun and bor-t must be /bort'$V_d/' and not /bort/ because the final suffix is inflectional and therefore lowering). Naturally, the syllabification of such forms is also different, but given our assumptions about the syllabification and the interpretation of defective vowels, these modified representations will not change the outcome of the derivations, i.e. the phonetically realized forms. In the two examples above, for instance, the final defective vowel does not syllabify because it cannot occur in an open syllable, so these forms are parsed as /\{sörn'\}V_d/ and /\{bort\}V_d/, and thus the correct surface forms ([sörn']) and [bort]) are derived. We shall discuss some more complex cases below.
We interpret lowering as a process that spreads the floating [+ open₁] feature locally to a (full or defective) vowel which is licensed (i.e. incorporated into a syllable) and is at the edge of a morpheme. Lowering applies regardless whether the licensed vowel is morpheme-initial or morpheme-final. The spreading process is local and non-iterative, i.e. it targets a single V. If the target is a full vowel, the floating feature can spread to its aperture node. In the case of defective target vowels, we assume that the nodes necessary for preserving well-formedness (e.g. root, vocalic, aperture) are automatically created in the course of the spreading to the empty skeletal position (cf. Sagey 1986, Clements and Hume 1995). This is indicated by parentheses enclosing the relevant nodes.

(33) Lowering

\[
\begin{array}{c}
\sigma \\
| \\
N \\
\text{X} \\
\text{(root)} \\
\text{(vocalic)} \\
\text{(aperture)} \\
[+ \text{open}_1]
\end{array}
\]

Condition: the target is peripheral in a morpheme

Whether the target of Lowering is \( V_\text{a} \) or \( V_\text{f} \), the output of the process is a structure shown in (34) (where irrelevant nodes are omitted). Crucially, this means that (in addition to its lowering effect) Lowering turns a defective vowel into a full one:
For expository purposes we shall use the following special symbols in the representations below: Subscripted ‘\( \text{OP} \)’ before \( V_d \) stands for the floating [+ open] feature. \( V_{\text{FOP}} \) denotes the lowered full vowel that is the result of (33). It must be borne in mind, however, that the linear representations used are just shorthand for the corresponding non-linear ones in the same way as phonetic symbols are for the appropriate feature trees.

Spreading is a feature filling process, therefore Lowering is blocked if the target vowel has an aperture specification which is incompatible with the feature that is being spread.

Let us now examine how lowering interacts with syllabification.\(^{55}\) As pointed out above, a word-final defective vowel is not realized phonetically because it cannot be syllabified. Therefore, in this position, a floating [+ open] feature does not surface since (33) cannot apply to a syllabically unparsed \( V_d \). (35) shows this with monomorphemic lowering stems and inflectional suffixes (Type A and Type B) attached to non-lowering stems:\(^{56}\)

\[
\begin{align*}
\text{fog}_{\text{OP}} V_d & \rightarrow \{\text{fog}_{\text{OP}}\} V_d \quad \text{[fog]} \quad \text{fog} \ ‘\text{tooth’} \\
\text{hal}_{\text{OP}} V_d & \rightarrow \{\text{hal}_{\text{OP}}\} V_d \quad \text{[hal]} \quad \text{hal} \ ‘\text{fish’} \\
\text{bor-\text{v}k}_{\text{OP}} V_d & \rightarrow \{\text{bor}_{\text{OP}}\} V_d \quad \text{[borok]} \quad \text{borok} ‘\text{wine’ (pl.)} \\
\text{bor-t}_{\text{OP}} V_d & \rightarrow \{\text{bort}_{\text{OP}}\} V_d \quad \text{[bort]} \quad \text{bort} ‘\text{wine’ (acc.)}
\end{align*}
\]

Recall that Type B suffixes (such as the accusative) always show up with a (lowered) linking vowel after lowering stems regardless whether the stem final consonant can or cannot form a syllable.

\(^{55}\)For expository purposes we shall use the following special symbols in the representations below: Subscripted ‘\( \text{OP} \)’ before \( V_d \) stands for the floating [+ open] feature. \( V_{\text{FOP}} \) denotes the lowered full vowel that is the result of (33). It must be borne in mind, however, that the linear representations used are just shorthand for the corresponding non-linear ones in the same way as phonetic symbols are for the appropriate feature trees.

\(^{56}\)The placement of \( \text{OP} \) relative to a syllable boundary is irrelevant and is not meant to indicate whether the floating feature is inside or outside a syllable.
licit coda with the suffixal -t. The reason is that lowering stems are vowel-final. In both cases the stem-final defective vowel can syllabify with the -t and, consequently, can be the target of Lowering, which turns it into a low full vowel (36a). This contrasts with the behaviour of -t after major stems where the defective vowel only appears as a result of overparsing by syllabification after stem-final consonants that cannot form a licit coda with the suffixal consonant (36b):

\[
\text{(36)} \quad \begin{align*}
\text{a.} & \quad \text{n'ak}_{\text{op}} V_d - t_{\text{op}} V_d & \Rightarrow & \{n'a\} \{kV_{\text{fop}} t_{\text{op}}\} V_d & [n'ok\text{t}] & \text{nyakat ‘neck’ (acc.)} \\
& \text{fal}_{\text{op}} V_d - t_{\text{op}} V_d & \Rightarrow & \{fa\} \{lV_{\text{fop}} t_{\text{op}}\} V_d & [f\text{lot}] & \text{f\text{alat ‘wall’ (acc.)} }
\end{align*}
\]

\[
\text{b.} \quad \text{bak}_{\text{op}} V_d & \Rightarrow \{ba\} \{kV_{\text{op}}\} V_d & [b\text{kot}] & \text{bakot ‘buck’ (acc.)} \\
& \text{dal}_{\text{op}} V_d & \Rightarrow \{dalt_{\text{op}}\} V_d & [d\text{lt}] & \text{dalt ‘song’ (acc.)}
\]

Thus, the generalization stated in (16) follows from the representation of lowering stems, the syllabification algorithm and the special constraints on the syllabification of defective vowels.

We have seen that Type A suffixes show up with a low linking vowel after lowering stems. Since Type A suffixes underlyingly begin with a full vowel, the stem-final \(V_d\) of lowering stems is deleted by Hiatus and the floating [+ open] feature of the stem can spread to the licensed suffix-initial \(V_c\):

\[
\text{(37)} \quad \begin{align*}
\text{fog-ak} & \text{ ‘tooth’ (pl.).} & \text{fog-atok} & \text{ ‘your tooth’} \\
/fog_{\text{op}} V_d - V k_{\text{op}} V_d/ & \Rightarrow & /fog_{\text{op}} V_d - V tok_{\text{op}} V_d/ \\
\text{syllabification} & \{fog_{\text{op}}\} V_d \{V, k_{\text{op}}\} V_d & \{fog_{\text{op}}\} V_d \{V, tok_{\text{op}}\} V_d \\
\text{Hiatus} & \{fog_{\text{op}}\} \{V, k_{\text{op}}\} V_d & \{fog_{\text{op}}\} \{V, tok_{\text{op}}\} V_d \\
\text{syllabification} & \{fo\} \{g_{\text{op}} V, k_{\text{op}}\} V_d & \{fo\} \{g_{\text{op}} V, tok_{\text{op}}\} V_d \\
\text{Lowering} & \{fo\} \{gV_{\text{fop}}, k_{\text{op}}\} V_d & \{fo\} \{gV_{\text{fop}}, tok_{\text{op}}\} V_d \\
& \text{[fog\text{atak}]} & \text{[fog\text{atok}]}
\end{align*}
\]

As pointed out above, a suffix-initial \(V_c\) does not lower (i.e. it cannot receive the spreading feature), if it has an aperture feature which is incompatible with the feature spread by Lowering. That is the reason why the suffix-initial vowels of two Type A suffixes,
superessive -on/-en/-ön and possessive 1pl -unk/-ünk, do not lower after lowering stems: e.g. *fal-on ‘on the wall’ and *fal-unk ‘our wall’. The initial V's of both these suffixes are underlyingly specified as [-open₁] (cf. Siptár & Törkenczy 2000) and thus cannot receive the spreading [+ open₁] feature.

Lowering may be unordered with respect to Hiatus as both possible orderings yield the correct results. Note that Lowering can spread [+ open₁] past an unlicensed V onto the closest potential target (the suffix-initial V) because defective vowels have no melodic structure. Compare the two ways of ordering Hiatus and Lowering in (37) and (38):

(38)  

*fog-ak ‘tooth’ (pl.)*

/\fog\_op\_V\_d\_V\_k\_op\_V\_d/

<table>
<thead>
<tr>
<th>Syllabification</th>
<th>Lowering</th>
<th>Hiatus</th>
</tr>
</thead>
<tbody>
<tr>
<td>{fog_op}_V_d{V_k_op}_V_d</td>
<td>{fog}_V_d{V_fo_k_op}_V_d</td>
<td>{fog}_V_d{V_fo_k_op}_V_d</td>
</tr>
</tbody>
</table>

Multiply suffixed forms of lowering stems are also derived in a straightforward manner. (39) shows how the accusative plural of a lowering major stem (*fog ‘tooth’*) and an ‘epenthetic’ lowering stem (*sátör ‘tent’*) is derived.

(39)  

*fog-ak-at ‘tooth’ (pl. acc.)*  
*sátör-ak-at ‘tent’ (pl. acc.)*

/\fog\_op\_V\_d\_V\_k\_op\_V\_d\_t\_V\_d/  
/\ša\_t\_V\_d\_r\_op\_V\_d\_V\_k\_op\_V\_d\_t\_op\_V\_d/

<table>
<thead>
<tr>
<th>Syllabification</th>
<th>Lowering</th>
<th>Hiatus</th>
</tr>
</thead>
<tbody>
<tr>
<td>{fog_op}_V_d{V_f_op}{k_op_V_d_op}_V_d</td>
<td>{ša}_{t_V_d_r_op}<em>V_d{V_f}</em>{k_op_V_d_op}_V_d</td>
<td>{ša}_{t_V_d_r_op}<em>V_d{V_f}</em>{k_op_V_d_op}_V_d</td>
</tr>
<tr>
<td>{ša}_{t_V_d_r_op}<em>V_d{V_f}</em>{k_op_V_d_op}_V_d</td>
<td>{ša}_{t_V_d_r_op}<em>V_d{V_f}</em>{k_op_V_d_op}_V_d</td>
<td>{ša}_{t_V_d_r_op}<em>V_d{V_f}</em>{k_op_V_d_op}_V_d</td>
</tr>
<tr>
<td>{ša}_{t_V_d_r_op}<em>V_d{V_f}</em>{k_op_V_d_op}_V_d</td>
<td>{ša}_{t_V_d_r_op}<em>V_d{V_f}</em>{k_op_V_d_op}_V_d</td>
<td>{ša}_{t_V_d_r_op}<em>V_d{V_f}</em>{k_op_V_d_op}_V_d</td>
</tr>
</tbody>
</table>

As can be seen in the derivations, syllabification is continuous, i.e. potentially, it reapplies after each phonological rule. Both Hiatus and Lowering show derived environment effects: the
former can only apply if its target is in another morpheme and the latter at the edge of a morpheme in the environment of another one. Neither applies intramorphemically. However, there is no evidence that their application is cyclic: they only ever need to apply once in the course of the derivation (naturally, they may have multiple targets). Rules of this kind challenge the traditional claim in Lexical Phonology that only cyclic rules are subject to the derived environment constraint on rule application. In fact, we know of no phonological rules in Hungarian that must be considered cyclic on grounds other than the derived environment constraint. Therefore—although the phonological rules belong to blocks (Block 1 and Block 2) and each suffix is marked according to whether it is analytic or synthetic—we assume that

---

57 Similar rules have been identified in a number of other languages, e.g. Finnish and Ondarroan Basque, cf. Hualde (1989) and Cole (1995).

58 Jensen and Stong-Jensen (1989a) argue for cyclic epenthesis in Hungarian to account for the behaviour of ‘epenthetic’ stems. However, their arguments do not contradict our claim because (i) essentially, they are based on their epenthesis process blocking in a non-derived context; and (ii) the arguments do not carry over to the present analysis because they are crucially dependent on the assumption that it is epenthesis that is responsible for the vowel–zero alternation in ‘epenthetic’ stems, Type A, and Type B suffixes alike. This is a view that we reject for the reasons discussed in 4.1.1.
It must be pointed out that as long as only synthetic suffixes are attached to the stem, it makes no difference if we assume that the whole ‘preassembled’ suffixed stem is subjected to the relevant phonological rules, or that the rules are (re)applied gradually (i.e. ‘cyclically’) as each suffix is considered, over the suffixed form. This can be seen in the derivation below, in which we adopt the Halle and Vergnaud (1987) approach to cyclicity:

\[
\text{fog}_{\text{OP}} \text{V}_{\text{d}} \text{V}_{\text{OP}} \text{V}_{\text{d}} \text{t}_{\text{OP}} \text{V}_{\text{d}}
\]

**cycle1**

syllabification: \{fog\}_{\text{OP}} \text{V}_{\text{d}} \text{V}_{\text{d}} \text{t}_{\text{OP}} \text{V}_{\text{d}}

Hiatus: n.a.

Lowering: n.a.

**cycle2**

syllabification: \{fog\}_{\text{OP}} \text{V}_{\text{d}} \{\text{V}_{\text{OP}} \text{k}_{\text{OP}} \text{V}_{\text{d}} \text{t}_{\text{OP}} \text{V}_{\text{d}}

Hiatus: \{fog\}_{\text{OP}} \text{k}_{\text{OP}} \text{V}_{\text{d}} \text{t}_{\text{OP}} \text{V}_{\text{d}}

Lowering: \{fo\} \{\text{g}_{\text{OP}} \text{V}_{\text{OP}} \text{k}_{\text{OP}} \text{V}_{\text{d}} \text{t}_{\text{OP}} \text{V}_{\text{d}}

**cycle3**

syllabification: \{fo\} \{\text{g}_{\text{OP}} \text{k}_{\text{OP}} \text{V}_{\text{d}} \text{t}_{\text{OP}} \text{V}_{\text{d}}

Hiatus: n.a.

Lowering: \{fo\} \{\text{g}_{\text{V}_{\text{FOP}}} \text{k}_{\text{OP}} \text{V}_{\text{OP}} \text{t}_{\text{OP}} \text{V}_{\text{d}}

[fogOKoK]

---

59 It must be pointed out that as long as only synthetic suffixes are attached to the stem, it makes no difference if we assume that the whole ‘preassembled’ suffixed stem is subjected to the relevant phonological rules, or that the rules are (re)applied gradually (i.e. ‘cyclically’) as each suffix is considered, over the suffixed form. This can be seen in the derivation below, in which we adopt the Halle and Vergnaud (1987) approach to cyclicity:

In the representations that follow the *irrelevant* suffix-final floating [+ open,] feature and \text{V}_{\text{d}} are disregarded.

61 Note that Hiatus must be non-iterative, because it would delete the suffix-initial \text{V}_{\text{f}} if it could apply to its own output.
creates the string -CV_{\text{f-op}}V_d+ C. Note that Hiatus cannot delete the stem-final V_d because the vowel sequence is not derived. Syllabification yields -\{CV_{f-op}\}\{V_dC\} and via Lowering the derived representation is the same as in the case of Type A suffixes: -\{CV\}_d\{V_{f-op}C\}. Thus, the prediction is that Type A and Type B suffixes behave in the same way when added to surface vowel-final lowering stems: a lowered linking vowel shows up before both types of suffixes.

This prediction is not borne out. Some suffixes never have a linking vowel after a surface vowel-final lowering stem (e.g. accusative, superessive, comparative), others do (e.g. plural, modal).\textsuperscript{62} This difference in behaviour only partially correlates with the distinction between Type A and Type B suffixes. In 4.1.3. above we pointed out that the unpredictability of behaviour indicates that lexical marking must be involved. We suggest that the source of this idiosyncratic behaviour is allomorphy. Surface vowel-final lowering stems have two lexical allomorphs: a ‘normal’\textsuperscript{63} one that (like all lowering stems) ends in a defective vowel, and another one whose final defective vowel is missing. By default, suffixes select the ‘normal’ allomorph. Some suffixes, however, are marked to select the other allomorph. When subjected to phonology, the former concatenations will surface with a lowered linking vowel while the latter ones will not have a linking vowel. Under this interpretation, a linking vowel is phonologically required after all lowering stems. Some lowering stems are special in that they have non-lowering lexical allomorphs as well,\textsuperscript{64} and some suffixes are morphologically irregular because they select the non-lowering allomorphs of these stems.

\textsuperscript{62}Modulo the OCP effect and the optionality of the linking vowel after mid vowels as discussed in 4.1.3.

\textsuperscript{63}I deliberately avoid using the word ‘regular’ here, since lowering stems are marked compared to non-lowering ones. ‘Normal’ is intended to mean ‘representing the norm for lowering stems’.

\textsuperscript{64}We have no explanation why all surface vowel-final lowering stems belong to this set. It must be pointed out, however, that the set contains some surface consonant-final lowering stems as well, cf. note 23. It is an interesting fact that, in contrast to consonant-final inflectional suffixes, vowel-final ones do not lower: lány-ai-m ‘my daughters’ and not *lány-ai-am. The latter is predicted if Hiatus and Lowering apply to the underlying representation */l\mathtt{an}^{\text{n}}\text{-ai}_{\text{op}V_dV_{m-op}V_d}V_d/, whereas the correct output is derived if the UR is /l\mathtt{an}^{\text{n}}\text{-ai}-V_{m-op}V_d/. It is as if a hiatus consisting of a full vowel and a defective one were dispreferred within a morpheme.
4.1.4.4. The past suffix

We saw in section 4.1.2.2. that the behaviour of the other Type B suffix, the past tense morpheme, is more complex than that of the accusative. This suffix displays vowel-zero alternation as well as an alternation involving its consonant(s): -Vtt-t (lop-ott (s)he stole’, fal-t (s)he devoured’, lop-t-am ‘I stole’). The length of the suffix-final consonant depends on the presence/absence of the linking vowel: it appears as a geminate after a phonetically expressed linking vowel (lop-ott). Recall, however, that the occurrence of the linking vowel depends on (i) the identity of the stem-final consonant (there is no linking vowel if nongeminate (!) t can form a licit coda with the stem-final consonant (fal-t)), and (ii) whether a vowel-initial (non-analytic) suffix follows (there is no linking vowel if it does (lop-t-am)).

The interdependence of the length of the suffixal consonant and the conditions on the occurrence of the linking vowel raises some questions about the representation of the past suffix. Since it is a Type B suffix, it is consonant-initial. When the suffix-initial linking vowel appears, it is a V that is the result of overparsing by syllabification. This is doubly problematic if we assume that the suffixal consonant is an underlying geminate. First, it is hard to see how the linking vowel could be absent after (some) consonant-final stems if the suffix is underlyingly -CC. As codas are maximally binary branching, it could not syllabify

---

65 In the following discussion we abstract away from the effects of postlexical Degemination (cf. Siptár & Törkenczy 2000, Siptár 2000, 2001b, 2002), which may shorten a geminate past -tt, compare Eve[t] körûtä. ‘(S)he ate some pears’ and Eve[t:] epret. ‘(S)he ate some strawberries’.

66 (i) Recall that only some licit codas are available for the past tense suffix to syllabify. I disregard this complication here (cf. 4.1.2.2.) and assume that there must be a stipulation specific to the past suffix that disallows its syllabification into a complex coda whose first term is an obstruent. It would be desirable to derive this effect from the representation of the past suffix and/or (more) general conditions on syllabification (cf. 3.4.3). At present, I do not see how this could be done. (ii) Note that, similarly to the accusative (cf. 4.1.2.2.), after t-final verbs a linking vowel appears even though geminate /tt/ is a well-formed coda: út-ótt ‘hit’ (3sg past indef) and not *út-t. For a discussion of this problem cf. 4.1.4.6.

67 This consonant would degeminate postconsonantally later. Since the past suffix must be distinguished from the accusative (whose suffixal consonant never shows up as a geminate), the former cannot be a single consonant underlyingly.
into the coda of the final syllable of a consonant-final stem regardless of the identity of the coda consonant—the expected string that results from syllabification would be C-V,CC. Second, given (23b), it is not even possible to overparse a final CCC string in this way, since a defective vowel is not licensed to occur in a doubly closed syllable (*C-V,CC). Furthermore, the non-occurrence of the linking vowel before vowel-initial suffixes (lop-t-am) would also be a problem. As the geminate could not syllabify as the onset of the syllable whose nucleus is the suffix-initial vowel (*{lop}V_d{tt-V,m}), the V_d that is the result of overparsing by syllabification preceding the past tense suffix would not be skipped since it could syllabify in a syllable closed by the first half of the geminate ({lo}{pV,t}{t-V,m}). This wrongly predicts that the linking vowel surfaces even before vowel-initial suffixes: *[lopottom].

To sum up, the past suffix behaves as a single /t/ in the derivation when the presence/absence of the linking vowel is determined by syllabification, but appears as a geminate if the linking vowel occurs at the surface. We can express this by assuming that the length of the suffixal consonant is the result of gemination. Since the past suffix has to be distinguished from similar suffixes (i.e. the accusative) whose suffixal consonant does not geminate in the same context, we suggest that its underlying representation is the following:

(40) N
    X X X
    t [+ open1,]

Thus, the past suffix is a /t/ whose root node is associated to a single timing slot followed by an empty timing slot (i.e. a timing slot devoid of melodic content). It ends in a floating [+ open1,] feature and a V_d because it is lowering (since it is an inflectional suffix: lop-t-am). We assume that an empty timing slot is completely invisible to syllabification: it may remain unparsed such that (i) it may be left ‘outside’ syllables (41a), or (ii) it may be ‘inside’ a syllable, but unassociated to a subsyllabic constituent (41b).\(^{68}\)

\(^{68}\)This presupposes that the syllabification algorithm looks at root nodes when the syllable trees are erected. Then, a timing unit without a root node is skipped (i.e. invisible). Note that defective vowels are different. They may be skipped by syllabification (because of the special constraints they are subject to), but they may not occur unparsed within a syllable because they
Empty timing slots that are unparsed at the end of the derivation are not interpreted phonetically. They become visible to syllabification if they receive content. Then, as other ordinary segments they will be (and must be) parsed. We suggest that this is what happens to the past tense suffix in some contexts. Specifically, its empty timing slot may be filled by spreading from the preceding segment. This process spreads the root node of the /t/ onto a following empty timing slot if the /t/ is preceded by a full vowel:

(42) /t/-spread

(42) applies after Default V (22) has applied. Note, however, that it does not have to be ordered with respect to (22). If we assume that (42) applies whenever it can, it will automatically only apply after (22) (if (22) does apply).

Figure (43) below shows the behaviour of the past suffix after stems ending in a single consonant when the stem-final consonant cannot form a licit coda with the suffixal consonant (43a), and when it can (43b):[^69]

[^69]: To simplify non-essential features of the derivations that follow, OP stands for floating [+ open₁] and $V_{FOP}$ is the lowered full vowel that results from spreading by Lowering.
(43) a. *lop-ott ‘steal’ (3sg past indef.) b. *fal-\textit{t} ‘devour’ (3sg past indef.)

<table>
<thead>
<tr>
<th>UR</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>X X X X X X X</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Syllabification</th>
</tr>
</thead>
<tbody>
<tr>
<td>σ</td>
</tr>
<tr>
<td>R</td>
</tr>
<tr>
<td>O N O N Co N</td>
</tr>
<tr>
<td>X X X X X X X X</td>
</tr>
<tr>
<td>l o p t OP</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hiatus</th>
</tr>
</thead>
<tbody>
<tr>
<td>n.a.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Lowering</th>
</tr>
</thead>
<tbody>
<tr>
<td>n.a.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Default</th>
</tr>
</thead>
<tbody>
<tr>
<td>σ</td>
</tr>
<tr>
<td>R</td>
</tr>
<tr>
<td>O N O N Co N</td>
</tr>
<tr>
<td>X X X X X X X X</td>
</tr>
<tr>
<td>l o p V\textsubscript{f} t OP</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Spread /t/</th>
</tr>
</thead>
<tbody>
<tr>
<td>σ</td>
</tr>
<tr>
<td>R</td>
</tr>
<tr>
<td>O N O N Co N</td>
</tr>
<tr>
<td>X X X X X X X X</td>
</tr>
<tr>
<td>l o p V\textsubscript{f} t OP</td>
</tr>
</tbody>
</table>


As can be seen in (43a) the suffixal consonant cannot syllabify into the coda of the stem-final syllable, therefore a degenerate syllable is created by syllabification. The licensed \( V_d \) of this syllable becomes a full vowel by Default V and thus the suffixal /t/ can spread to the empty X slot on its right (lop-ott). No degenerate syllable is created, however, if the suffixal consonant can form a coda with the stem-final one (43b). In this case /t/-spread cannot apply since its structural description is not met, and the past suffix surfaces as a nongeminate [t] (falt). Comparable forms of cluster-final stems (e.g. dong-ott ‘buzz’ (3sg past indef.), csukl-ott ‘hiccup’ (3sg past indef.)) derive like (43a).

The derivation of multiply suffixed forms of the same stems (i.e. when the past suffix is followed by a Type B suffix) is shown in (44).

(44) a. lop-t-am ‘steal’ (sg1 past) b. fal-t-am ‘devour’ (1sg past)

<table>
<thead>
<tr>
<th>UR</th>
<th>N</th>
<th>N</th>
<th>N</th>
<th>N</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>lop</td>
<td>t</td>
<td>OP</td>
<td>( V_f )</td>
<td>m</td>
<td>OP</td>
</tr>
</tbody>
</table>

Syllabification

\[
\begin{array}{cccccccc}
\sigma & \sigma & n.a. \\
\hline R & R \\
O & N & O & N & Co & N \\
X & X & X & X & X & X & X \\
l & o & p & t & V_f & m & OP \\
\end{array}
\]
Hiatus

<table>
<thead>
<tr>
<th>σ</th>
<th>σ</th>
<th>σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td>O</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>l o p</td>
<td>t</td>
<td>OP</td>
</tr>
</tbody>
</table>

Syllabification

<table>
<thead>
<tr>
<th>σ</th>
<th>σ</th>
<th>σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td>O</td>
<td>N</td>
<td>Co</td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>l o p</td>
<td>t</td>
<td>OP</td>
</tr>
</tbody>
</table>

Lowering

<table>
<thead>
<tr>
<th>σ</th>
<th>σ</th>
<th>σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td>O</td>
<td>N</td>
<td>Co</td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>l o p</td>
<td>t</td>
<td>Vₐ</td>
</tr>
</tbody>
</table>

Default

n.a.

Spread /t/

n.a.

Note that when Hiatus deletes the defective vowel before the full vowel of the Type B suffix, crucially, the /t/ can syllabify ‘across’ the empty timing slot as the onset of the initial syllable of the following suffix. Thus, /t/-spread cannot apply because its structural description is not met and the past suffix surfaces as a nongeminate [t] (*lop-t-am, *fāl-t-am*). The difference between the two stems is that in the case of *lop-t-am* the stem-final consonant is followed by an unsyllabified Vₐ which is the result of overparsing by an earlier round of syllabification.
This $V_d$ eventually cannot syllabify (because it is not licensed to occur in an open syllable) and is not interpreted phonetically.

‘Epenthetic’ stems whose final consonant cannot form a licit coda with the /t/ of the past suffix (e.g. forog ‘revolve’) behave similarly to the comparable major stems in (43a) and (44a). The only difference in their behaviour is due to the underlying defective vowel in the final syllable of ‘epenthetic’ stems. (45) shows the (intermediate) representation of forog-ott ‘revolve’ (3sg past indef.) which is the result of syllabification:

(45)

When Default V applies to this representation, it turns the licensed $V_d$ into $V_p$ and the /t/ can spread to the available empty position on its right. The unsyllabified $V_d$-s do not receive phonetic interpretation, thus the surface form is [forgot:]. (46) shows a multiply suffixed form of the same stem (forog-tam ‘revolve’ (1sg past.)) after Hiatus and syllabification (and Lowering).\textsuperscript{70}

(46)

\textsuperscript{70}Note that the defective vowel between the stem final consonant and the past suffix is the result of a round of syllabification before Hiatus because the two consonants cannot form a licit branching coda.
The stem-internal licensed $V_d$ becomes a full vowel when Default $V$ applies to this representation. As /t/-spread cannot apply, the surface form is [foroktóm].

Given our assumptions about syllabification, the prediction for ‘epenthetic’ stems that end in a consonant with which /t/ can form a licit coda (e.g. /rabVₘ,l/ ‘rob’, /sodVₘ,t/ ‘roll’, /ugVₘ,r/-/jumₘ,p/ ‘jump’, /omVₘ,l/-/ ‘collapse’) is that they should form their singly and multiply suffixed past forms like the ‘epenthetic’ stems discussed above: -C}Vₘ{CVₘ,t}Vₘ# (like forgott) and -{CVₘ,C}Vₘ{tVₘ,fᵒpₘ} (like forogtam). This prediction is only borne out in the case of some past forms of some of these ‘epenthetic’ verbs. In (47) below we have charted the possible singly and multiply suffixed past forms of representative ‘epenthetic’ stems that end in the right consonants for branching codas. The present form and the nominalized one are included for comparison. We have capitalized the forms that are not predicted given the representation of ‘epenthetic’ stems and the syllabification algorithm.

(47)

<table>
<thead>
<tr>
<th>stem</th>
<th>3sg past indef.</th>
<th>1sg past</th>
<th>1sg pres def</th>
<th>nominalized form</th>
</tr>
</thead>
<tbody>
<tr>
<td>rabVₘ,l ‘rob’</td>
<td>RABOL-T</td>
<td>rabol-t-am</td>
<td>RABOL-OM</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>rabl-om</td>
<td>rabl-ás</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ugVₘ,r- ‘jump’</td>
<td>-</td>
<td>ugor-t-am</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>ugr-ott</td>
<td>UGR-OTT-AM</td>
<td>ugr-om</td>
<td>ugr-ás</td>
</tr>
<tr>
<td>omVₘ,l- ‘collapse’</td>
<td>OMOL-T</td>
<td>omol-tam</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>oml-ott</td>
<td>OML-OTT-AM</td>
<td>oml-om</td>
<td>oml-ás</td>
</tr>
</tbody>
</table>

71The [k] is the result of Voicing Assimilation, cf. Siptár & Törkenczy (2000).

72/ugVₘ,t/ and /omVₘ,l/- are bound stems (of the -ik class): ugr-ik, oml-ik (3sg pres) vs. ugor-j, omol-j (imp.)

73The reason is that syllabification will overparse the string consisting of the stem-final consonant and the suffixal /t/ in spite of the fact that they could form a branching coda because the $V_d$ that occurs in the last syllable of the stem is disallowed in a doubly closed syllable (cf. the discussion of the accusative of ‘epenthetic’ nouns in section 4.1.4.2.).
It must be pointed out that (i) all these ‘epenthetic’ stems seem to have unexpected forms, sometimes as the only form at a given point in the paradigm, sometimes as an alternative to an expected one; (ii) the unexpected forms are not confined to the past paradigm; (iii) it is unpredictable which forms of which stems will be unexpected.\(^{74}\) We suggest that the reason for this complex state of affairs is that not all forms of these stems derive from the same underlying representation. Parallel underlying representations exist for these verbs (cf. Törkenczy 2002b), one of which is ‘epenthetic’.\(^{75}\) For instance, rabol has an underlying major stem too, which has a full vowel in the last syllable (CV.C), hence rabol-t (and rabol-om). It is unpredictable which forms are derived from which UR(s) and whether only one, or more than one parallel UR is available for the same form (as in omol-t/oml-ott).\(^{76}\) The parallel UR is not necessarily CV.C-final. Forms like ugr-ott-am and oml-ott-am are derivable neither from a CV.C-final nor from a CV.C-final UR. We propose that these forms derive from an underlying stem that ends in a CC cluster which is not a possible coda.\(^{77}\) Thus, some of the lexemes discussed show allomorphy to such an extent that they may have as many as three parallel UR variants from which the different forms are derived.

We noted above that the singly suffixed past forms of cluster-final stems can be handled in a straightforward way. Multiply suffixed cluster-final stems, on the other hand, present a problem.

Multiply suffixed forms of stems ending in clusters that are not well-formed codas

\(^{74}\) (i) The fact that variation should occur does seem to be predictable for a certain class of verbs when suffixed with a certain type of suffixes: epenthetic -ik verbs stems have alternative forms before -sz/-asz/-esz (2sg pres. indef.), -ni/-ani/-eni (inf.) and -lak/-lek/-alak/-elek (1sg* 2sg/pl*), -nak/-nek/-anak/-enek (3pl pres. indef.), -na/-ne/-ana/-ene (cond.) and -tok/-tek/-tök/-otok/-etek/-ötök (2pl pres. indef.), i.e. quasi-analytic suffixes (see note 32 above): e.g. füröd-ni/fürd-eni ‘bathe’ (inf.). See Rebrus & Törkenczy (1998, 1999), Rebrus (2000b), Törkenczy (2002b).

\(^{75}\) (ii) It is interesting to note that the nominalized form is always the expected one.

\(^{76}\) The fact that there is variation among native speakers as to which alternative forms they find acceptable confirms this interpretation.

\(^{77}\) Compare the almost identical boml-ott ‘unfold’ (3sg past indef.), which has no alternative *bomol-t.

\(^{77}\) That is, the UR of the stem of these forms is like the bound stem /čukl-/ csukl-ik ‘hiccup’ (cf. the discussion below), whose stem-final cluster is never separated by a vowel.
(i.e. defective stems, e.g. /čukl-/ csukl-ott-am ‘hiccup’ (1sg past), /bűzl-/ bűzl-ött-em ‘stink’ (1sg past), /vedl-/ vedl-ett-em ‘slough’ (1sg past), cf. Károly (1957), Hetzron (1975), Törkenczy (2000b, 2001b, 2002ab), Rebrus and Törkenczy (1999), derive in the following way:

(48) csukl-ott-ak ‘hiccup’ (3pl past indef.)

UR

<table>
<thead>
<tr>
<th>N</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>č</td>
<td>u</td>
</tr>
</tbody>
</table>

Syllabification

\[ \sigma \sigma \sigma \]

<table>
<thead>
<tr>
<th>R</th>
<th>R</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>O</td>
<td>N</td>
<td>Co</td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>č</td>
<td>u</td>
<td>k</td>
</tr>
</tbody>
</table>

Hiatus

\[ \sigma \sigma \sigma \]

<table>
<thead>
<tr>
<th>R</th>
<th>R</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>O</td>
<td>N</td>
<td>Co</td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>č</td>
<td>u</td>
<td>k</td>
</tr>
</tbody>
</table>

Lowering

\[ \sigma \sigma \sigma \]

<table>
<thead>
<tr>
<th>R</th>
<th>R</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>O</td>
<td>N</td>
<td>Co</td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>č</td>
<td>u</td>
<td>k</td>
</tr>
</tbody>
</table>
The interesting point in this derivation is the output of Hiatus. If syllabification applied to the output of Hiatus to syllabify the /t/ into the onset of the last syllable of the word, the rest of the word could not be syllabified. The defective vowel preceding the /t/ could remain unparsed, but the consonant before it could not be syllabified into the coda on its left since they do not make up a licit coda (*{čukl}V_{q}{t...}). Thus, the whole CV_{q} string before the /t/ would have to remain unparsed (*{čukl}IV_{q}{t...}). This is excluded by non-exhaustiveness, which we restate here in a stricter form.\(^{78}\)

\(^{78}\)Compare (25).
Non-exhaustiveness

Only defective material (i.e. defective vowels and empty positions) may remain unparsed into syllables.

Another option would be for syllabification to overparse the stem-final cluster, but this is not possible either, since overparsing is a structure changing operation and thus can only happen in a derived environment (\(\ast \{\ddot{c}u\} \{kV_{l}\}V_{d}\{t\ldots\}\)). Thus, syllabification cannot apply to the output of Hiatus and the derivation proceeds as shown in (48).

For multiply suffixed forms of cluster-final stems that end in a well-formed coda (e.g. dong-t-ak ‘buzz’ (3pl past indef.)), the syllabification algorithm predicts that that they should follow the derivation of lop-t-am (cf. (44a)). That is, after Hiatus the past /t/ syllabifies as the onset of the syllable whose nucleus is the full vowel of the suffix following it. The stem-final consonants can syllabify as a coda and the defective vowel following them remains unparsed; /t/-spread cannot apply. This is shown in (50):

(49) Non-exhaustiveness

(50) dong-t-ak ‘buzz’ (3pl past indef.)

---

79 This is why the stem-final clusters of these stems are never broken up and why these stems have a defective paradigm, cf. section 4.1.1.
The first [k] is the result of Voicing Assimilation.

The prediction is correct for the stem *dong*, but recall that there are other stems ending in a branching coda that (i) either have an alternative multiply suffixed past form alongside the expected one (e.g. *fing* ‘fart’: *fing-tak/FING-OTT-AK*, *told* ‘lengthen’: *told-t-ak/TOLD-OTT-

\[\text{[doŋktək]}^{80}\]

\[80\] The first [k] is the result of Voicing Assimilation.
The stems with /t/-final clusters will be discussed in section 4.1.4.6. These unpredicted forms (which are capitalized in the previous sentence) are always of the same shape: they have a linking vowel after the stem (and consequently a geminate /tt/). The unexpectedness of these forms consists in the unmotivated occurrence of the linking vowel after the stem. The V_d (which results from a previous round of syllabification) is eventually unparsed after some stems (e.g. dong: ..Ng}V_d{t...}—which is the predicted case—after others it may be parsed (e.g. fing: ..Ng}V_d{t... / ..N} {gV_d}... or must be parsed (e.g. old: ..l}{dV_d}...). I do not really have an explanation for these forms and can only offer some speculation as to why the defective vowel behaves in this way after these stems. First of all, obviously, lexical marking must be involved since all these stems have well-formed codas, and the defective vowel may be parsed or unparsed after the same coda clusters in different stems (compare fing and dong, old and told) and therefore, the occurrence of the linking vowel cannot be predicted on the basis of the melodic content of the coda clusters. It is certainly the stems that have to be marked in some way. Second, the reason why this differential behaviour is only observed after cluster-final stems must be related to the status of internal CCC clusters. We have pointed out in section 3.2.2. that, apart from sporadic irregular monomorphemic examples, internal CCC clusters only occur if they are not within the same analytic domain. There is one systematic set of counterexamples to this generalization: multiply suffixed past forms of verb stems that end in a branching coda, such as [doŋktok] (recall that the past suffix is synthetic). The internal CCC cluster of these forms always consists of a branching coda followed by an onset. However, one could argue that the data above suggest that internal branching codas are dispreferred. This would make the unexpected forms above the regular case, and the stems that allow the underparsing of a V_d after a branching coda would have to be lexically marked. In the present treatment we leave this question open.

To conclude, we summarize the different types of (singly and multiply suffixed past forms of) verb stems discussed in this section. Only those forms of the stems are included in (51) that are predicted on the basis of the UR identified. The notation is as follows: C_x is a

81The stems with /t/-final clusters will be discussed in section 4.1.4.6.

82Domain-internal CCC clusters cannot have a different structure (*C.CC) since branching onsets are disallowed in Hungarian. On monomorphemic words with internal CCC clusters cf. section 4.1.4.5.
consonant such that /Cₚt/ is a well formed coda; Cₚ is a consonant such that /Cₚt/ is not a well formed coda; CₚCₜ is a well-formed coda; and CₚCₜ is not a well-formed coda. The parenthesized question marks are meant to show our indecision about which of the forms syllabification should predict (both forms are attested!). CVₚC final stems are the ones that are traditionally called ‘epenthetic’ and CₚCₜ final stems are ‘defective’.

(51)

<table>
<thead>
<tr>
<th>stem-final string in UR</th>
<th>singly suffixed past form</th>
<th>multiply suffixed past form</th>
</tr>
</thead>
<tbody>
<tr>
<td>CV₁C</td>
<td>CV₁Cₜ</td>
<td>fal-t</td>
</tr>
<tr>
<td>CV₂C</td>
<td>CV₂Cₜ</td>
<td>lop-ott</td>
</tr>
<tr>
<td>CV₃C</td>
<td>CV₃Cₜ</td>
<td>ugr-ott</td>
</tr>
<tr>
<td>CV₄C</td>
<td>CV₄Cₜ</td>
<td>forg-ott</td>
</tr>
<tr>
<td>CC</td>
<td>C₃Cₜ</td>
<td>dong-ott</td>
</tr>
<tr>
<td>CC</td>
<td>C₄Cₜ</td>
<td>csukl-ott</td>
</tr>
</tbody>
</table>
4.1.4.5. Analytic affixes and appendices

When Block 1 syllabification happens and the alternations dependent on syllable structure are calculated, material in one (dependent or independent) analytic domain is not visible to that in the other. This can be seen in (52) below where ‘epenthetic’ stems are shown in isolation, followed by a vowel-initial analytic suffix (terminative -ig), and by a vowel-initial synthetic suffix (plural -V/k):

\[(52) \quad \begin{array}{lll} 
\# & _V\text{-initial analytic suffix} & _V\text{-initial synthetic suffix} \\
\text{bokor ‘bush’} & \text{bokor-ig} & \text{bokr-ok} \\
\text{retek ‘radish’} & \text{retek-ig} & \text{retk-ek} \\
\text{kölyök ‘kid’} & \text{kölyök-ig} & \text{kölyk-ök} \\
\end{array} \]

(52) shows that the underlying defective vowel of ‘epenthetic’ stems is phonetically expressed before terminative -ig (and other vowel-initial analytic suffixes)\(^{83}\) in spite of the fact that the stem-final consonant syllabifies as the onset of the suffix-initial syllable at the surface. We attribute this to Default V having applied in Block 1 (while syllabification applies in both blocks). This means that all the licensed V\(_c\)-s are turned into full vowels before Block 2 syllabification applies, which can then syllabify the stem-final consonant as an onset since the

---

\(^{83}\)Other suffixes of this type are causal-final -ért (bokor-ért ‘for the bush’), anaphoric possessive -é(bokor-é ‘that of the bush’), adverb-forming -ul/iil (bantu-ul ‘in Bantu’).
Recall that Block 1 rules apply within analytic domains and then the whole word is submitted to the Block 2 rules cf. section 2.2. The fact that the internal brackets are not shown in the Block 2 stage of the derivation is not meant to imply that they have been erased. It is simply that the derivation interprets the larger domain at this stage.

We assume that appendices are not maximized to the detriment of a preceding coda, hence *re. te.kröl.

Even sequences of identical vowels are possible under these conditions: kiismer /kiismer/ ‘learn all about’, taxiig /taksiig/ ‘up to the taxi’, bantuul /bantuul/ ‘in Bantu’.

Consonant-initial analytic suffixes (e.g. inessive -ban/-ben, dative -nak/-nek, ablative -töl/-töl, delative -röl/-röl etc.) behave in the same way, except that Block 2 syllabification cannot syllabify the stem-final consonant as (part of) the onset of the suffix-initial syllable (bo.kor.ban ‘in the bush’, re.tek.röl ‘about horseradish’).

We have noted (cf. sections 3.3.1 and 4.1.2.2.) that hiatus is possible morpheme-internally (kies /kieš/ ‘picturesque’), when the two vowels are in different independent and/or dependent analytic domains ([[ ki ] [esik]] /kiešik/ ‘fall out’ (verb), [[ kapu ] ig] /kapuig/ ‘up to the gate’), but is not possible when the second vowel is initial in a synthetic suffix. In the last case, Hiatus deletes the suffix-initial vowel (cf. 4.1.4.2.). This pattern can be accounted for if we assume that Hiatus is only a Block 1 rule (where it is subject to the derived environment constraint), and does not apply in Block 2:

---

84 Recall that Block 1 rules apply within analytic domains and then the whole word is submitted to the Block 2 rules cf. section 2.2. The fact that the internal brackets are not shown in the Block 2 stage of the derivation is not meant to imply that they have been erased. It is simply that the derivation interprets the larger domain at this stage.

85 We assume that appendices are not maximized to the detriment of a preceding coda, hence *re. te.kröl.

86 Even sequences of identical vowels are possible under these conditions: kiismer /kiismer/ ‘learn all about’, taxiig /taksiig/ ‘up to the taxi’, bantuul /bantuul/ ‘in Bantu’.
Compare the different behaviour of an empty skeletal slot, cf. 4.1.4.4.

The stem-final V cannot be deleted by Hiatus because Hiatus only applies in Block 1.

<table>
<thead>
<tr>
<th></th>
<th>Block 1</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Syllabification</td>
<td>Hiatus</td>
<td>Syllabification</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[k{i}{eš}]</td>
<td>n.a.</td>
<td>n.a.</td>
<td>[k{a}{pu}]{V,}k</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[k{a}{pu}]{V,}k</td>
</tr>
<tr>
<td>Block 2</td>
<td>Syllabification</td>
<td>n.a.</td>
<td>n.a.</td>
<td>[k{a}{puk}]</td>
</tr>
<tr>
<td></td>
<td>[k{i}{eš}]</td>
<td>[k{a}{pu}]{ig}</td>
<td>[k{a}{puk}]</td>
<td></td>
</tr>
</tbody>
</table>

Vowel-initial analytic suffixes can be used to argue for the stray erasure of defective material (defective vowels and empty skeletal slots) at the end of Block 1. We have seen above that Block 2 syllabification parses the last consonant of underlyingly consonant-final stems (e.g. *pad* ‘bench’) as an onset when a vowel-initial analytic suffix follows: {p{a}{dig}} *pad-ig* ‘up to the bench’. Lowering stems (e.g. *vad* ‘beast’—compare *vad-ak* ‘beast’ (pl.)) are expected to syllabify in a different way if defective vowels are visible in Block 2 derivation. The reason is that lowering stems end in a defective vowel (/vad_{op}V_{d}/), and defective vowels may only remain unparsed outside a syllable, i.e. an unparsed V_{d} cannot occur within a syllable that has a nucleus. Consequently, the consonant preceding the final V_{d} of a lowering stem cannot syllabify ‘across’ the V_{d} to become the onset of the analytic vowel-initial suffix; *vad-ig* ‘up to the beast’ is predicted to syllabify as {vad}V_{d}{ig}. However, for native speakers, there is no difference between the syllabification of *padig* and *vadig*—both syllabify the intervocalic consonant into the second syllable. In order to avoid the unnecessary and counterintuitive difference between the syllabification of these items, we shall assume that defective vowels are erased at the end of Block 1 derivation. The two words will then be identical when Block 2 syllabification happens and will syllabify in the same way: {p{a}{dig}}, {v{a}{dig}}.

Although syllabification applies in both blocks, it is subject to different conditions in them. Block 1 syllabification can build syllable structure on the segmental melody, but it is subject to the derived environment constraint, so it can only overparse it (i.e. insert defective vowels) in a derived environment created by a synthetic suffix. Furthermore,

---

87 Compare the different behaviour of an empty skeletal slot, cf. 4.1.4.4.

88 The stem-final V_{d} cannot be deleted by Hiatus because Hiatus only applies in Block 1.
Block 1 rules (including syllabification) must be allowed to apply to dependent analytic domains as well as non-dependent ones (i.e. the material in a dependent domain cannot ‘wait’ uninterpreted until Block 2 rules apply to the larger domain) because synthetic suffixes may follow analytic ones, and processes that target material within a domain consisting of a stem and a synthetic suffix also target that within a dependent domain consisting of an analytic suffix and a synthetic suffix. For instance, (i) overparsing by syllabification can take place in the accusative of nouns ending in an analytic suffix (such as -ság/-ség ‘-hood’ or deverbal noun-forming -vány/-vény): [ [ lány ] ság-ot ] ‘maidenhood’ (acc.) vs. [ [ lát ] vány-t ] ‘spectacle’; (ii) Lowering applies after analytic lowering suffixes (e.g. -van/-ven): [ [ hat ] van-at ]; (iii) Hiatus (which is Block 1 only since it does not delete the initial vowel of vowel-initial analytic suffixes) applies after a vowel-final analytic suffix such as diminutive -ka/ke: [ [ malac ] ká-k ] ‘piglets’ (where the length of the suffix-final vowel is due to Low Vowel Lengthening (cf. Siptár & Törkenczy 2000)).

---

89Block 1 rules (including syllabification) must be allowed to apply to dependent analytic domains as well as non-dependent ones (i.e. the material in a dependent domain cannot ‘wait’ uninterpreted until Block 2 rules apply to the larger domain) because synthetic suffixes may follow analytic ones, and processes that target material within a domain consisting of a stem and a synthetic suffix also target that within a dependent domain consisting of an analytic suffix and a synthetic suffix. For instance, (i) overparsing by syllabification can take place in the accusative of nouns ending in an analytic suffix (such as -ság/-ség ‘-hood’ or deverbal noun-forming -vány/-vény): [ [ lány ] ság-ot ] ‘maidenhood’ (acc.) vs. [ [ lát ] vány-t ] ‘spectacle’; (ii) Lowering applies after analytic lowering suffixes (e.g. -van/-ven): [ [ hat ] van-at ]; (iii) Hiatus (which is Block 1 only since it does not delete the initial vowel of vowel-initial analytic suffixes) applies after a vowel-final analytic suffix such as diminutive -ka/ke: [ [ malac ] ká-k ] ‘piglets’ (where the length of the suffix-final vowel is due to Low Vowel Lengthening (cf. Siptár & Törkenczy 2000)).

Verb stems that can appear as vowel-final have /v/-final allomorphs before underlyingly vowel-initial suffixes. We assume that both allomorphs are listed in the lexicon. This appears to violate non-exhaustiveness as formulated in (49). Note, however, that (49) was designed to prevent the resyllabification of an already syllabified form such as \{čuk\} \{IV_{\text{t}}\}..., while here we have to do with the underparsing of non-defective material in an unsyllabified form. Let us suppose that the vacuous application of the syllabification algorithm is the same as non-application. Then, the application of syllabification may mean the (i) full parsing/reparsing, (ii) overparsing, or (iii) underparsing of a string. In the case of the already syllabified string \{čuk\} \{IV_{\text{t}}\}... application would result in (ii) (*\{ču\} \{kV_{\text{d}}\} V_{\text{d}}\{t\}...) or (iii): *\{čuk\} IV_{\text{d}}\{t\}.... Of these, (ii) is excluded by the derived environment constraint and (49) is intended to exclude (iii). Non-application, however, would still ‘yield’ a licit syllabified form (\{čuk\} \{IV_{\text{t}}\}...). The case of monomorphemic clusters under consideration is different: here, non-application is not possible because it would leave the whole morpheme unsyllabified. Overparsing is excluded for the same reason as above, and thus the minimal underparsing of non-defective material is the only option left (\{fe:r\}j). Thus,
consonants can syllabify as an appendix in Block 2:

\[(56)\]

\[
\begin{array}{ccc}
\text{Block 1} & \text{Syllabification} & \text{Block 2} \\
\{\text{pre:m}\} & \{\text{fe:rn}\} & \{\text{pre:m}\} \\
\{\text{p:re:m}\} & \{\text{fe:rn}\} & \{\text{p:re:m}\} \\
\{\text{pre:m}\} & \{\text{fe:rn}\} & \{\text{pre:m}\} \\
\text{Default V} & \text{n.a.} & \text{n.a.} \\
\end{array}
\]

It is an advantage of this treatment that clusters containing a subsyllabic analytic suffix and identical monomorphemic clusters receive the same analysis in terms of syllable structure: compare kár-j {ke:rn} ‘ask’ (imp. indef.) and fárj {fe:rn} ‘husband’.

Appendices are thus available for syllabification for consonants peripheral in an (independent or dependent) analytic domain if they are left unparsed by Block 1 syllabification. It is to be noted, however, that the licensing of morpheme-internal clusters consisting of more than two consonants is still unaccounted for. In section 3.3.2.2. we argued that these clusters are irregular. Nevertheless, they are not broken up by overparsing and they are not simplified

\[
\text{the appropriate version of Non-exhaustiveness must ban the underparsing of non-defective material in the first case, but must permit it in the second one. (49‘) is a possible formulation:}
\]

\[(49‘)\] Non-exhaustiveness: Syllabification may leave phonological material unparsed. Non-defective material may only be left unparsed as a last resort (where defective material is V₄ or an empty timing slot).

This is obviously the kind of problem that could be given an optimality theoretic interpretation (there is a conflict of constraints: it is more important to at least partially parse a morpheme than to obey the requirement that bans the unparsing of non-defective material).

\[93\text{Internal clusters of more than three consonants cannot be analyzed as a coda+ onset sequence. Internal CCC clusters cannot be syllabified as a simplex coda plus an onset because the onset may not branch. In section 4.1.4.4. we saw that multiply suffixed past forms of cluster-final stems suggest that the well-formedness of domain-internal branching codas is questionable. They are probably ill-formed (or at least marked). Morpheme internally they certainly seem to be ill-formed since morpheme-internal C1C2C3 clusters where C1C2 could be a licit coda are just as irregular/rare as those in which it could not.}\]
by deletion in the lexical phonology. The fact that they are not overparsed is due to the derived environment constraint, but it is not yet clear how they are licensed, since, morphologically, they are not peripheral in an analytic domain. We suggest here that the reason is a mismatch between purely morphological domains and phonologically relevant ones (cf. Törkenczy & Siptár 1999). Although the words containing these clusters are monomorphemic, phonologically, they are treated in Hungarian as if they were compounds, i.e. a morphologically unitary domain is phonologically analyzed as if it were two independent domains. The actual point at which the division of the morphological domain is made may vary from speaker to speaker, but is always in the middle of the cluster. Thus, every word containing a cluster longer than two consonants has more than one (re)analysis.

(57)

\[
\text{templom} \quad \left[ \right. \text{tem} \left[ \text{plom} \right] \quad \text{or} \quad \left[ \text{temp} \right] \left[ \text{lom} \right] \\
\text{export} \quad \left[ \right. \text{ek} \left[ \text{sport} \right] \quad \text{or} \quad \left[ \text{eks} \right] \left[ \text{port} \right] \\
\text{puzdra} \quad \left[ \right. \text{puz} \left[ \text{dra} \right] \quad \text{or} \quad \left[ \text{puzd} \right] \left[ \text{ra} \right] \\
\text{asztma} \quad \left[ \right. \text{as} \left[ \text{tma} \right] \quad \text{or} \quad \left[ \text{ast} \right] \left[ \text{ma} \right] \\
\text{lajstrom} \quad \left[ \right. \text{laj} \left[ \text{strom} \right] \quad \text{or} \quad \left[ \text{lajš} \right] \left[ \text{trom} \right] \quad \text{or} \quad \left[ \text{lajšt} \right] \left[ \text{rom} \right] \\
\]

Block 1 syllabification can only partially syllabify the material in each independent analytic domain (\[[ \{ek\} \left[ s\{port\} \right] \]) or \([[ \{ek\} s \left[ \{port\} \right] \]) and Block 2 syllabification can incorporate the unsyllabified peripheral consonants into extended syllables (\[[ \text{kek}\left[ s\{port\} \right] \]) or (\[[ \text{eks}\left[ \{port\} \right] \]).

---

94Compare Kassai’s empirical results (Kassai 1999abc).

95A probable scenario is that native speakers try to syllabify these words in the usual way (\{laj\}ś{t\{rom\}}) and make the division somewhere in the unsyllabifiable portion (the number of ways depends on how many consonants would be left unsyllabified).


97Overparsing is excluded since the string within the domain is not derived. Note that in some cases both domains of the reanalysed word may be fully syllabified in Block 1, e.g. \[[ \{temp\} \left[ \{lom\} \right] \].
The fact that different native speakers may syllabify these words differently and even the same native speaker may find more than one syllable division possible shows that they are, or can be, reanalyzed as compounds in different ways. With some items, one syllabification is much more likely than the alternative one(s): e.g. most (if not all) speakers would syllabify asztma as /ast.ma/ rather than /as.tma/. This suggests that everybody analyzes this word as [ [ ast ] [ ma ] ], which is unexpected in the present account. In most cases, however, all the predicted syllabifications seem equally possible: lajástrom = lajš.trom = lajšt.rom.

The above treatment of monomorphemic words containing clusters longer than two consonants is compatible with all other facts of Hungarian phonology. Its weakness is that there is very little internal independent motivation supporting it. In principle, evidence might come from backness/frontness harmony. As compound members do not have to harmonize (cf. Siptár & Törkenczy 2000), we would expect that there should be disharmonic stems among those that contain these overlong clusters. This appears to be true (e.g. angström ‘id.’, ösztrogán ‘oestrogen’). It has to be pointed out, however, that (i) real disharmony is just as rare among these words as in words that do not contain clusters longer than two consonants (e.g. sofőr ‘driver’), and (ii) most of the words with CCC clusters whose vowels do not agree in backness contain /e/ as a non-harmonizing vowel, which is neutral. If /e/ is neutral, then these words are not disharmonic. Thus, the ‘evidence’ is inconclusive. It must be pointed out that the ‘evidence’ would not be better even if /e/ were harmonic. The reason is that we would then have a lot of disharmonic words with /e/ that do not contain a CCC cluster (e.g. betyár ‘highwayman’, haver ‘friend’), i.e. disharmony would be just as frequent in these words as it is in those that have a cluster consisting of more than two consonants (e.g. export ‘id.’, komplett ‘complete’).

---

98 Stress, for instance, is not a problem: both compound and non-compound words have initial stress.

4.1.4.6. OCP effects, residual problems

In this section we discuss some residual problems concerning the vowel-zero alternations analysed above.

4.1.4.6.1. ‘Epenthetic’ stems

In section 4.1.1. I argued that the phonological mechanism responsible for the vowel-zero alternation in ‘epenthetic’ stems is not epenthesis, i.e. it is not phonotactically motivated. Nevertheless, I pointed out that there are certain phonotactic restrictions that hold between the consonants flanking the defective vowel of these stems. These are static well-formedness constrains that disallow morpheme-shapes that do not conform to them. They are repeated in (58) below:

\[(58)\]
\[
\begin{align*}
&\text{a. } \#C_iVC_jC_i \text{ if } C_i = C_j \\
&\text{b. } \#C_iVC_jC_i \text{ if } C_i, C_j = [\text{son}], \text{ and only one of them has a laryngeal node}^{100}
\end{align*}
\]
\[
\text{c. } \#CCVC_i, \#CV_aCC
\]

Such constraints may appear unexpected since the consonants involved are non-adjacent. I suggest that all three constraints can be attributed to the transparency of the intervening defective vowel. Although the consonants flanking the defective vowel are not string-adjacent on the skeletal tier, their root nodes are adjacent since a \(V_d\) has no material below its X slot. Thus, constraints that apply to features and nodes below the skeleton can hold between consonants that are separated by a defective vowel.

Therefore, (58a) can be attributed to the OCP. On the root tier two identical consonants that are separated by a \(V_d\) would be a fake geminate (cf. e.g. Perlmutter 1995) and

---

\(^{100}\)There seems to be a single exception, the ‘epenthetic’ verb \(képez\) ‘train’ [keːpeːz]: \(képzés\) ‘training’ [keːbzeːʃ], cf. Novák (1999).
would be banned by the OCP intramorphemically. Thus, the underlying representation of an ‘epenthetic’ stem could not contain the string in (59a) (where $C_a$ is a consonantal root node dominating a particular feature tree):\(^{101}\)

\[(59)\]

\[
\begin{align*}
\text{a.} & \quad N \\
& \quad X \quad X \quad X \\
& \quad \quad X \quad X \\
& \quad C_a \quad C_a \\
\text{b.} & \quad N \\
& \quad X \quad X \quad X \\
& \quad C_a
\end{align*}
\]

The structure in (59b) could not occur in an ‘epenthetic’ stem either. Although (59b) conforms to the OCP and is well-formed, the defective vowel ‘embedded’ in the true geminate could never surface since Default V ((22)) could not apply to the $V_a$ of (59b) because of the No-Crossing Constraint (NCC)—a geminate integrity effect, cf. e.g. Kenstowicz and Pyle (1973), Schein and Steriade (1986), Yip (1987), Clements and Hume (1995). In order for (22) to specify the $V_a$ in (59b) association lines would have to cross, which is banned by the NCC (e.g. Goldsmith 1976). Thus, an ‘epenthetic’ stem that has identical consonants separated by a $V_a$ cannot be represented.\(^{102}\)

Although the transparency of $V_a$ plays a role in the other two phonotactic restrictions

\(^{101}\)This presupposes a ‘strict’ interpretation of the OCP in which it can be determined by inspecting a single tier whether a given configuration is an OCP violation or not (e.g. McCarthy 1988: ‘Adjacent identical elements are prohibited.’) rather than a ‘loose’ one under which the structural tier (such as the skeleton) to which the features/nodes concerned are anchored has to be examined as well (cf. Hewitt and Prince 1989: ‘No melodic element may be structurally adjacent to an identical element’). Under the loose interpretation (59) would not be an OCP violation because the two $C_a$-s are not structurally adjacent (they are tier-adjacent, but their structural anchors, the X-slots in this case, are non-adjacent).

\(^{102}\)It must be pointed out that, apparently, identical place nodes can occur on the two sides of a $V_a$ in an ‘epenthetic’ stem (cf. 4.1.1. note 7). I have no explanation for this.
(58bc) as well, they cannot be derived from general constraints like the OCP or the NCC. (58b) is identical with the constraint which requires that adjacent obstruents must agree in voicing (cf. Siptár & Törkenczy 2000), i.e. that they either must share a laryngeal node, or neither should have one. What it shows is that for this constraint, adjacency must be defined on the laryngeal tier. (58c) is more problematic. It is easy to see why ‘epenthetic’ stems cannot end in more than one consonant (*CV_aCC). (a) This string would be unsyllabifiable if the stem is in isolation, or if it is followed by a consonant-initial suffix because of the restriction on the syllabification of a defective vowel ((23b)); (b) in a hypothetical stem ending in this string, the first one of the two stem-final consonants would always syllabify as a coda when a synthetic vowel-initial suffix follows: ..CV_aC}{C-V.. (as onsets may not branch). This means that the V_a would always surface since it would be parsed in a closed syllable and Default V would apply to it—consequently such a stem would not show vowel-zero alternation, which is what ‘epenthetic’ stems do by definition. The problem is that there is no similar reason for why consonant clusters do not precede the V_a in ‘epenthetic’ stems (*CCV_aC). Epenthetic stems containing this string could be syllabified in isolation (..C}{CV_aC}) and would display vowel-zero alternation, i.e. the defective vowel would not surface when a vowel-initial synthetic suffix is attached to the stem (…CC}V_a{C-V..). It is not clear how the constraint could be explained. We tentatively suggest that it may derive from the constraint excluding three adjacent consonants in Hungarian (which would then have to be formulated in terms of root nodes rather than X-slots), but here we leave this question open.

4.1.4.6.2. /t/-final stems

In the discussion of the accusative and the past tense suffix, both of whose consonantal melody is /t/, and which may receive a defective vowel by overparsing (both are Type B suffixes), we noted that such a linking vowel appears even if the stem ends in /t/ in spite of the fact that

---

103 Unless the cluster is impossible as a coda.

104 In GP a principled explanation of similar phenomena in French was proposed by Charette (1990, 1991).
geminates are licit codas (cf. sections 4.1.2.2. and 4.1.4.4.). Consider the examples below:

\[(60)\]

<table>
<thead>
<tr>
<th>accusative</th>
<th>past (3sg indef)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rět-ot</td>
<td>vět-ett 'do wrong'</td>
</tr>
<tr>
<td>hat-ot</td>
<td>hat-ott 'effect'</td>
</tr>
<tr>
<td>szövet-ot</td>
<td>szövet-ett 'make weave'</td>
</tr>
<tr>
<td>lapát-ot</td>
<td>matat-ott 'fumble'</td>
</tr>
</tbody>
</table>

The question is why these forms have a linking vowel, i.e. why syllabification overparses the string consisting of the stem final consonant and the suffix when the suffixal consonant could be syllabified into the stem-final syllable as part of a well-formed branching coda: e.g. *rět-t 'field' (acc.), *vět-t 'do wrong' (3sg past indef), etc.

The accusative of /t/-final lowering stems (e.g. hát-at ‘back’ (acc.)) and the accusative forms or the singly suffixed past forms of Ct-final stems (e.g. ezüst-öt ‘silver’ (acc.), döreszt-ett ‘wake sb. up’ (3sg past indef)) do not require a special explanation. They behave like all the other lowering stems and cluster-final stems (see sections 4.1.4.3 and 4.1.4.4.). It is the accusative and the past forms of non-lowering non-cluster-final stems like those in (60) that are problematic.

It would be desirable to avoid stipulating constraints that are specific to these stems and are not directly related to the general syllable template, and to be able to motivate the occurrence of the linking vowel with some general principle. It is tempting to find this principle in the OCP. I shall discuss a possible OCP-based account, but will point out that—because of some arbitrary complexities of the data and certain theoretical difficulties—it is not possible to give an account which connects the OCP and syllabification/overparsing.

The fact that a linking vowel shows up in the accusative and the singly suffixed past forms of /t/-final stems can be interpreted as a repair of an OCP violation (which consists in the juxtaposition of two identical root nodes as a result of suffixation) if we assume that (i) the OCP is not only a constraint on lexical representations, but is effective in the derivation as well; and (ii) the OCP violation created by synthetic suffixation is not repaired by merging the identical root nodes into a true geminate. Given these two assumptions we could motivate why overparsing happens: fake geminates may not be parsed as a branching coda (e.g. re:t-t_{op}V_{d}
The only exception, some multiply suffixed past forms, will be discussed below.

By contrast, analytic suffixation may freely create (fake) geminates:

[meg:ta] 'prevent', [ad:ta] 'give' (imp. indef.), [bab:an] 'in (the) bean'. Note that true geminate consonants and fake ones are phonetically indistinguishable (both have a single release stage). Affricates are the only exception, because the first half of a fake geminate affricate may (optionally) be released too (e.g. [t'te]) while the first half of a true one may not (e.g. [t:] but *[t't]). Compare (true) viccel [vit:zl, *vit'tzl] 'joke', gleccser [gleczer, *glecczer] 'glacier' with (fake) bohóc-cipő [bohoct:ipő:, bohoct'ipő:] 'clown shoe', apacs csónak [apocsonak, apoccs:onak] 'apache boat'.

There are two problems, however: the first one concerns some data we have not examined yet and the second one is theoretical.

As shown in (60), the linking vowel is always present in the accusative and the singly suffixed past forms of /t/-final stems. The multiply suffixed past forms of /t/-final verbs (in which the past suffix is followed by another (vowel-initial) suffix) behave in a complex and ad hoc way.

The generalization is the following: in these forms the past suffix appears without a linking vowel if the stem ends in the string at/et:

<table>
<thead>
<tr>
<th>(61)</th>
<th>singly suffixed form (3sg past indef)</th>
<th>multiply suffixed form (1sg past)</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>ápolgat-ott</td>
<td>ápolgat-t-am</td>
<td>'nurse repeatedly'</td>
</tr>
<tr>
<td></td>
<td>emelget-ett</td>
<td>emelget-t-em</td>
<td>'lift repeatedly'</td>
</tr>
<tr>
<td></td>
<td>várat-ott</td>
<td>várat-t-am</td>
<td>'make wait'</td>
</tr>
<tr>
<td></td>
<td>dolgoztat-ott</td>
<td>dolgoztat-t-am</td>
<td>'make work'</td>
</tr>
</tbody>
</table>

105 The only exception, some multiply suffixed past forms, will be discussed below.

106 By contrast, analytic suffixation may freely create (fake) geminates: [meg:ta] 'prevent', [ad:ta] 'give' (imp. indef.), [bab:an] 'in (the) bean'. Note that true geminate consonants and fake ones are phonetically indistinguishable (both have a single release stage). Affricates are the only exception, because the first half of a fake geminate affricate may (optionally) be released too (e.g. [t'te]) while the first half of a true one may not (e.g. [t:] but *[t't]). Compare (true) viccel [vit:zl, *vit'tzl] 'joke', gleccser [gleczer, *glecczer] 'glacier' with (fake) bohóc-cipő [bohoct:ipő:, bohoct'ipő:] 'clown shoe', apacs csónak [apocsonak, apoccs:onak] 'apache boat'.

107 Note that the accusative may not be followed by another suffix.
b. faggat-ott  faggat-t-am  ‘interrogate’
dädgelet-ett  dädgelet-t-em  ‘pamper’
ugat-ott  ugam-t-am  ‘bark’
matat-ott  matat-t-am  ‘rummage’

Otherwise, the past suffix shows up with a linking vowel:

\[
\begin{array}{ccc}
\text{singly suffixed form} & \text{multiply suffixed form} & \text{gloss} \\
(3\text{sg past indef}) & (1\text{sg past}) & \\
\hline
a. & vakí t-ott & vakí t-ott-am & ‘blind’ \\
 & húsí t-ett & húsí t-ett-em & ‘cool’ \\
 & taní t-ott & taní t-ott-am & ‘teach’ \\
\hline
b. & bocsát-ott & bocsát-ott-am & ‘allow’ \\
 & tát-ott & tát-ott-am & ‘open wide’ \\
 & fút-ött & fút-ött-em & ‘heat’ \\
 & fut-ott & fut-ott-am & ‘run’ \\
 & köt-ött & köt-ött-em & ‘tie’ \\
 & süt-ött & süt-ött-em & ‘bake’ \\
\end{array}
\]

There are four exceptions to the generalization above. Two of them have a linking vowel in the multiply suffixed form although the verb stem ends in \textit{at/et} (63a), and the other two do not have a linking vowel in the same form although the verb stem does not end in \textit{at/et} (63b):

\[
\begin{array}{ccc}
\text{singly suffixed form} & \text{multiply suffixed form} & \text{gloss} \\
(3\text{sg past indef}) & (1\text{sg past}) & \\
\hline
a. & hat-ott & hat-ott-am & ‘effect’ \\
 & vet-ett & vet-ett-em & ‘sow’ \\
\end{array}
\]
It is difficult to make sense of this pattern. It is not clear why verbs ending in at/et should behave differently from other /t/-final verbs. Note that it is not the morphological make-up of the stems that distinguishes those in (61) from those in (62). While in many of the relevant verb stems the at/et string is (part of) a derivational suffix (cf. (61a); ápol-gat, vár-at, dolgoz-tat, etc.), no such morphological complexity is obvious in others (cf. faggat, ugit, matat (61b)) and, furthermore, there are stems that end in a suffix among the /t/-final stems that do not end in at/et (cf. (62a); vakít, hús-ít, tan-ít), too. Thus, the reason for the differential behaviour is not morphological. There appears to be an arbitrary division in the set of /t/-final verbs.

If we want to keep the OCP as an explanation we have to assume that the merging of identical root nodes juxtaposed by suffixation in Block 1 is possible for the set of stems that end in at/et. This merging, however, is only possible if another suffix follows the past suffix. The singly suffixed past forms of all at/et-final stems, including the ones that allow merging in their multiply suffixed forms, have a linking vowel.

Even if we make some provision for the above complications, there are also theoretical problems with the idea that the OCP drives overparsing here.

(i) It is difficult to conceive overparsing, i.e. the insertion of a defective vowel by syllabification as a process that repairs an OCP violation. The reason is that the ‘repair’ would not eliminate the OCP violation: as defective vowels do not have phonological material below their skeletal point, the two identical root nodes (that of the stem final /t/ and the suffix-initial one) whose skeletal points the V₄ separates as a result of overparsing would remain adjacent. This is a problem for all the three forms under consideration (the accusative, the singly and the multiply suffixes past forms of /t/-final stems). A possible way out is to say that here the ‘loose’ interpretation of the OCP is in force. In this case overparsing would be a repair since

---

108 -(V)gat/-(V)get is the frequentative/diminutive suffix, -(t)at/-(t)et is the causative suffix and -ít is a denominal/deadjectival verb-forming suffix.
it would separate the structural anchors (the X slots) of the identical root nodes. Thus, they
would no longer be in violation of the OCP (see note 101). This, however, would be in
contradiction with the way the OCP is supposed to work in ‘epenthetic’ stems: there, crucially,
the ‘strict’ interpretation is required (cf. 4.1.4.6.1). Allowing different interpretations of
the same supposedly general principle within the same language (‘“strict” in the lexicon, but
“loose” in the derivation’) would make the principle so unrestrictive that it would lose
much/all of its explanatory power.

(ii) It might be argued that the ‘loose’ interpretation is possible in the derivation
because overparsing is crucial in the elimination of the OCP violation. Default V will eliminate
the violation even in the strict sense of the OCP and the V<sub>d</sub> created by overparsing is necessary
for Default V to apply. It must be pointed out, however, that not all V<sub>d</sub>-s created by
overparsing can syllabify. As Default V does not apply to those that do not, the OCP will be
violated in the strict sense if the unsyllabified vowel is flanked by identical consonants. The
syllabification algorithm predicts this state of affairs in the multiply suffixed past forms of the
verb stems being discussed, which are supposed to derive like loptam as shown in (44) in
section 4.1.4.4. Here the V<sub>d</sub> created by overparsing between the verb stem and the initial /t/
of the past suffix eventually cannot syllabify. Thus, these forms violate the OCP in the strict
sense even after Default V applies.<sup>109</sup>

Thus, we conclude that OCP-motivated overparsing is not a tenable account of the
behaviour of the multiply suffixed forms of /t/-final stems. The OCP does play a role,
however, but the repair is not overparsing. Let us assume a strict interpretation of the OCP and
that in Hungarian it applies to underlying forms and to derived representations in Block 1 (but
not in Block 2 and postlexically). OCP violations can be repaired in two ways: epenthesis (64)
and merging (65) (where nonessential structure is suppressed and root<sub>v</sub> is the root node of a
vowel):

---

<sup>109</sup>The stray erasure of defective material at the end of Block 1 creates an OCP violation
in the loose sense too.
Both rules are triggered by the OCP. (64) inserts a full vowel, and thereby can eliminate a violation. (65) achieves the same by creating a true geminate. We assume that these rules do not apply across an analytic boundary and that they are ‘morphological’ in the sense that they apply to the precompiled stem before phonological rules apply. Only (64) can apply to the accusative and the singly suffixed forms of /t/-final stems because (65) requires that there should be an adjacent vowel root node after the root of the second consonant in the input. (64) and (65) are in an elsewhere relationship, with (65) being the more specific rule. This predicts that only (65) applies to the multiply suffixed past forms. This is correct for items like those in (61) but not for those in (62). Therefore the latter stems (and hat and vet) have to be marked in the lexicon so that (65) may not apply to them, in which case (64) will.\footnote{For a very different recent analysis of the same set of data, see Rebrus & Trón (in print).}
4.2 Alternations involving consonants

4.2.1. Alternating v-suffixes: -val/-vel, -vá/-vé

There are two suffixes (instrumental -val/-vel: só-val ‘with salt’ and translative -vá/-vé só-vá ‘(turn) into salt’) which begin with a [v] after vowel-final stems, but after stems ending in consonants, the segmental content of their initial consonant is identical with that of the stem-final consonant. The stem-final consonant and the initial consonant of the suffix are realized as a geminate only if the stem ends in a single consonant (cf. (66a)). These suffixes will be referred to as ‘alternating v-suffixes’. There are ‘non-alternating v-suffixes’ as well (such as -van/-ven ‘-ty’: hat-yan ‘sixty’, deverbal noun-forming -vány/-vény: lát-vány ‘sight’, deverbal adverb-forming lő-ve ‘in the state of being shot’, -va/ve: lop-va ‘stealthily’), which are [v]-initial after vowel-final stems, but whose initial /v/ is unchanged/retained even after consonant-final stems (cf. (66b)).

(66)\textsuperscript{111} a. alternating v-suffix

\[
\begin{array}{ccc}
\text{V}_- & \text{VC}_- & \text{CC}_- \\
nő-vel & csap-pal [p:] & domb-bal [mb] \\
Feri-vel & méz-zel [z:] & vers-sel [rš] \\
falu-val & lécel [t\textsuperscript{t}:] & lánz-cal [nt\textsuperscript{t}] \\
lő-vel & kar-ral [r:] & férj-jel [rj] \\
\end{array}
\]

b. **non-alternating v-suffix**

<table>
<thead>
<tr>
<th>V_</th>
<th>VC_</th>
<th>CC_</th>
</tr>
</thead>
<tbody>
<tr>
<td>lő-ve</td>
<td>lop-va [pv]</td>
<td>old-va [ldv]</td>
</tr>
<tr>
<td>ró-ve</td>
<td>néz-ve [zv]</td>
<td>zeng-ve [ngv]</td>
</tr>
<tr>
<td>nyű-ve</td>
<td>nyom-va [mv]</td>
<td>tart-va [rtv]</td>
</tr>
<tr>
<td>ri -va</td>
<td>zár-va [rv]</td>
<td>fest-ve [štv]</td>
</tr>
</tbody>
</table>

The crucial analytical problem is how to distinguish the alternating v-suffixes from the non-alternating ones. The classical generative analysis was to set up an abstract underlying segment (usually /w/) as the initial consonant of the alternating v-suffixes while the non-alternating ones were considered underlyingly /v/-initial (cf. Szépe 1969, Vago 1980a). Autosegmental representation makes it possible to avoid this excessive abstractness. While the non-alternating suffixes are underlyingly /v/-initial, the alternating ones can be assumed to begin with an empty position that receives melody (by spreading) from the final consonant of the stem to which the suffix is attached. The result of the spreading is a geminate (csap-pal [čap:λ]) which degeminates if the stem is cluster-final by an independently motivated process of general postlexical degemination that applies in the environment of another consonant (domb-bal (intermediate) domb:λ → [dombol]). After vowel-final stems the position remains empty and is later specified as /v/ by default (nő-vel [nő:vel]). This idea is pursued by Vago (1989).\(^{112}\)

Here, we propose a different analysis which obviates the need for the default rule and can explain some asymmetries in the working of the putative general degemination rule. We retain the basic idea of the autosegmental analysis, i.e. that the alternation is due to spreading melody from the stem-final consonant to a suffix-initial empty position, but we claim that the rule that spreads the stem-final melody is the same rule that applies to derive the geminate of the past tense suffix, i.e. we generalize /t/-spread (42) as (67), where C is any consonantal root:

---

\(^{112}\)Vago uses moraic syllable structure (i.e. there is no skeletal tier), consequently, for him, empty positions are empty root nodes and not empty timing slots.
As (67) spreads the root node of a consonant to a following empty position only if the consonant is preceded by a full vowel, no rule of degemination is needed to account for cases like *domb-bal* [domb-bal]. In these cases (67) does not apply since its structural description is not met. This would explain why the putative degemination process is compulsory in this case while in other cases it is often optional and/or speech-rate dependent. What is compulsory is really the lack of gemination (i.e. spreading); it is only postlexical degemination, which is necessary for independent reasons and does not apply to the case at hand, which can be optional (cf. Nádasdy 1989a, Siptár & Törkenczy 2000, Siptár 2000, 2001b, 2002).

Assuming that (67) is responsible for the ‘complete assimilation’, I now examine the two related questions: how are alternating and non-alternating v-suffixes to be distinguished and what is the source of the surface [v] in the alternating suffixes after vowel-final stems. In principle the two kinds of v-suffixes may be distinguished (i) representationally, (ii) in terms of domains (i.e. the analytic vs. synthetic distinction), or (iii) with reference to a combination of (i) and (ii).

Let us assume (following Szépe 1969, Vago 1980a, 1989) that non-alternating v-suffixes have an underlying initial /v/ (68a). Suppose that alternating v-suffixes are different in that they begin with an empty timing slot (68b).

The non-alternating v-suffixes are certainly analytic, since they may be attached to any stem, regardless of the identity or the number of the stem-final consonants (cf. (67b)). Alternating v-suffixes, on the other hand, do display phonotactic interaction with the stem-final consonant and the maximum number of consonants that can arise as a result of the affixation is two. This
suggests that they are synthetic. However, with these assumptions, i.e. that alternating v-initial suffixes are synthetic and begin with an empty timing slot, the analysis runs into serious difficulties. Recall that, as opposed to defective vowels, empty X-slots are invisible to syllabification until they receive segmental content, and as such they can float ‘inside’ a syllable, i.e. if an empty X-slot is preceded by a consonant and followed by a vowel, the consonant can syllabify ‘across’ the empty X-slot as an onset (cf. section 4.1.4.4). This invisibility is not a problem when an alternating v-suffix is attached to major stems like *csap.* The stem-final consonant could first syllabify as an onset to the suffixal vowel, but after the spreading, the stem-initial position would become visible and the resulting geminate would syllabify as an onset+ coda sequence (*csap.pal*). The problem arises when the alternating suffix is attached to an ‘epenthetic’ stem e.g. *bokor* ‘bush’. The (intermediate) representation of the suffixed form would be the following after syllabification:

```
(69) σ
          R
        / O N  Co N  O  N  Co  N
     X  X  X  X  X  X  X  X
b o k  r  a l
```

The stem-final consonant would syllabify as an onset (‘across’ the invisible X) and the stem-internal V of the ‘epenthetic’ stem would remain unparsed (since it cannot syllabify in an open syllable). The problem is that Default V (22) would not target the unparsed stem-internal V, and consequently (67) could not apply because its structural description is not met (the spreading consonant is not preceded by a V). Thus, the predicted surface form would be *[bokɔrɔl] instead of the correct [bokɔrɔl] (bokorrəl ‘bush’ (instr.)).

Assuming that alternating v-initial suffixes are analytic (and allowing (67) to apply both in Block 1 and Block 2) does not help either. The reason is that since Block 1 rules apply in the dependent domain too (cf. sections 2.2 and 4.1.4.5), the suffix-initial empty position would be deleted by the convention that defective material is erased at the end of Block 1 derivation (cf. 4.1.4.5). Thus, the (67) would have no chance to apply in Block 2.
In order to avoid these problems we propose that the difference between the two kinds of suffixes is not representational, but simply a difference of domains: both of them begin with an underlying /v/ (whose root node is associated to a timing slot) but alternating v-suffixes are synthetic while non-alternating ones are analytic ([ [ hat-val ] hat-tal ‘with six’ vs. [ [ hat ] van ] hat-van ‘sixty’). We maintain that the assimilation is the result of spreading by (67), but claim that the empty timing slot targeted by (67) in the alternating suffixes is not underlying but derived by (70):

\[(70) \quad v\text{-delink}\]

\[
\begin{array}{c|c}
\text{X} & \text{X} \\
| & \dagger \\
\text{C} & \text{v}
\end{array}
\]

v-delink is a Block 1 rule that feeds (67). As it is a Block 1 rule, it is subject to the Derived Environment Constraint. Therefore, it does not apply in monomorphemic words containing postconsonantal /v/, e.g. tviszt ‘twist’, szvetter ‘sweater’, özvegy ‘widow’, olvas ‘read’, szarv ‘horn’, könyv ‘book’. Neither can it apply to analytic /v/-initial suffixes (or in compounds whose second member is /v/-initial) since their initial /v/ is not postconsonantal within the analytic domain even when they are preceded by a consonant-final stem, e.g. [ [ hat ] van ], [ [ lop ] va ], [ [ ár ] [ vi z ] ] ‘flood’, [ [ át ] [ vág ] ] ‘cut through’. Assuming that the timing slot remains to be linked to the onset node after (70) delinks the /v/, i.e. that the onset ‘branch’ is only removed if the timing slot is also erased or if the nucleus is deleted (cf. Hayes 1989), (70) need not be ordered with respect to Default V (22). This is a crucial assumption, since if the X becomes dissociated from the syllable as a result of (70), then it becomes invisible to syllabification and we are back to the problem with the ‘epenthetic’ stems discussed above. This could only be remedied by ordering (70) after (22), because after Default V, ‘epenthetic’ stems can behave like major stems ending in VC (like csap as described above).

If the /v/ is underlying in alternating v-suffixes, there is no need for a default rule
to insert it after vowel-final stems.\textsuperscript{113} (70) simply does not delink it in these cases and thus C-spread is inapplicable (\textit{nõ-vel}).

Lowering stems behave just like non-lowering ones: (70) applies when they are followed by an alternating v-suffix; compare non-lowering \textit{csap (csap-ok ‘taps’) csa[p]al ‘with a tap’} and lowering \textit{fal ‘wall’ (fal-ak ‘walls’) fá[l:]al}. This suggests that (70) must be slightly modified to permit the delinking of the suffix-initial /v/ even if there is an intervening V\textsubscript{d} between it and the last consonant of the stem:

(71) \textit{v-delink}

\[
\begin{array}{c}
X \quad (V_d) \quad X \\
\mid \quad \dagger \\
C \quad v
\end{array}
\]

(67) must be modified in a similar way to optionally permit a V\textsubscript{d} before the source and the target of the spreading. Note that the spreading does not result in line-crossing because defective vowels do not have root nodes.\textsuperscript{114}

\textsuperscript{113}This is a desirable consequence since a default rule that inserts /v/ would only ever apply to the two alternating v-suffixes. This is hardly the general scope one would expect from a rule which in essence would mean ‘/v/ is the default consonant in Hungarian’. In Vago (1989) the default rule is made somewhat more general as it also applies in ‘v-adding’ stems too. These stems end in a vowel in isolation, but have a stem-final [v] before vowel-initial synthetic suffixes: \textit{lő ‘horse’} – \textit{lov-ak ‘horses’, lő ‘shoot’} (3sg pres. indef.) – \textit{lőv-ők ‘shoot’} (1sg pres. indef.). In the present treatment this is considered to be suppletive allomorphy, i.e. the phonology does not derive the allomorphs from a single underlying representation. There are only a small number of stems that show this /v/-\textaealternation (n= 19), which is sometimes (unpredictably) accompanied by changes in the quality and/or the quantity of the vowel in the stem-final syllable: compare \textit{l[õ:] ‘shoot’} (3sg pres. indef.) – \textit{l[õ]v-ők} (1sg pres. indef.): change in vowel quantity, but no change in quality, and \textit{f[õ:] ‘cook’} (3sg pres. indef.) – \textit{f[õ]v-õk} (1sg pres. indef.): no change in vowel quantity and quality; \textit{l[õ:]} – \textit{l[õ]v-ak ‘horses’}: change in vowel quantity, but no change in quality, and \textit{f[õ:] ‘lake’} – \textit{f[õ]v-ak ‘lakes’}: change in both vowel quantity and quality. Several forms have an alternative form in which the stem behaves as a regular vowel-final major stem: e.g. \textit{szó ‘word’} – \textit{szav-ak/szó-k ‘words’, falu ‘village’} – \textit{falv-ak/falu-k ‘villages’}, etc. All this suggests that the alternation these stems display is non-phonological.

\textsuperscript{114}In essence, the modified rules say that defective vowels are invisible to the two operations. Note that we could not attribute this invisibility to the convention that erases
defective material at the end of Block 1 because (i) the rules show derived environment effects and (ii) Block 2 application of (71) would neutralize the difference between alternating and non-alternating v-suffixes. This raises the question whether the convention should be ‘split’ in such a way that it could differentiate between defective vowels and empty slots. We do not pursue this option here.
4.2.2. *h*-alternations

There are two alternations involving the sound [h]. It can alternate (a) with zero: [h]-[ç] (e.g. *cseh [çɛ] ‘Czech’ vs. *cseh-es [çɛhes] ‘Czech-like’), or (b) with a voiceless velar fricative: [h]-[x] (*doh [dɔx] ‘musty smell’ vs. *doh-os [dɔhoš] ‘musty’). It is unpredictable whether a morpheme that has an allomorph with a final [h] (e.g. [çɛh-ɛs, doh-oš]) displays alternation (a) or (b). In all these words [h] appears prevocally/in onset position. Otherwise (preconsonantally and word-finally), we get zero in the former set of items (henceforward *cseh*-type words) and [x] in the latter set (henceforward *doh*-type words). This is shown in (73) below:

(73)  

<table>
<thead>
<tr>
<th><em>cseh type</em></th>
<th><em>doh type</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>cseh</td>
<td>[çɛ]</td>
</tr>
<tr>
<td><em>Czech</em></td>
<td>‘musty smell’</td>
</tr>
<tr>
<td>cseh-tól</td>
<td>[çɛtɔl]</td>
</tr>
<tr>
<td>‘Czech’ (abl.)</td>
<td>‘musty smell’ (abl.)</td>
</tr>
<tr>
<td>cseh-es</td>
<td>[çɛhes]</td>
</tr>
<tr>
<td><em>Czech-like</em></td>
<td>‘musty’</td>
</tr>
</tbody>
</table>

The context beyond the word is irrelevant, i.e. we do not get alternants with [h] preceding vowel-initial words in either type: *cseh asszonyok* [çɛzoson’ok, *çɛhɔs:son’ok] ‘Czech women’, *doh okozta* [dɔxokosta, *dohokosta] ‘musty smell caused [it]’.

[h] and [x] never contrast in Hungarian. [h] cannot occur as a geminate, only [x] can (e.g. *fach [fɔx] ‘pigeon-hole’, *pech [pɛx] ‘misfortune’, *ahhoz [ɔx:ɔz]116 ‘to that’, *Bachot [bɔxɔt] ‘Bach’ (acc.). [x] occurs preconsonantally and word-finally (e.g. *doh [dɔx], doh-tól [dɔxto:l] ‘from [the] musty smell’, *ihlet [ixlet] ‘inspiration’, *jacht [jɔxt] ‘yacht’), while [h] only occurs in prevocalic position (e.g. *doh-os [dɔhoš], cseh-es [çɛhes] *hol* [hol] ‘where’, *nőha*

---

115 We are abstracting away from two minor variations on the general pattern: the (postlexical) voicing of [h] between sonorants and vowels (e.g. *konyha [kɔn’ho] ‘kitchen’, *csehes [çɛhes] ‘Czech-like’), and the (postlexical) fronting of [x] after front vowels (e.g. *pech [pɛx] ‘misfortune’). See Siptár (1994b), Siptár Páter & Szentgyörgyi Szilárd (2002ab, to appear).

116 *Ahhoz* can also be pronounced [ɔhoz].

Traditionally, the cseh type is considered to be the native pattern (cf. Deme 1961). This assumption was taken over by most generative accounts of the phenomenon (e.g. Vago 1980a, Siptár 1994b, Törkenczy 1994a). As there are no systematic constraints on the occurrence of vowels preceding the [h] in cseh-type words, and stems can end in the same set of vowels that can occur before a [h], h-deletion (rather than h-insertion) was assumed to apply in these words (preconsonantally and finally, or (equivalently) in the coda). Because of the complementary distribution between [h] and [x], doh-type morphemes were assumed to end in the same underlying segment as cseh-type morphemes and therefore had to be marked in the lexicon so as not to undergo the deletion rule (i.e. the doh type was considered exceptional). In these treatments, typically, /h/ is assumed to be underlying and all instances of surface [x] are derived by rule (but, in principle, this could be the other way round).

We claim that in present-day ECH it is the doh type that is the systematic pattern (cf. Siptár 1998b) rather than the cseh type, which we suggest is not phonological (anymore) and is best considered as suppletive allomorphy. The principal reason is that it is the doh type that is productive in the sense that (a) in ECH there is a tendency for cseh-type morphemes to be reclassified as doh-type items while the reverse is unattested (e.g. mēh [meː]/[meːx] ‘bee’, but eunuch ‘id.’ [eunux], *[eunu]); and (b) new h-final items (loans and acronyms) are always of the doh type (e.g. Hezbollaḥ [ḥezbolːɔxː] ‘id.’, APEH [ɔpɛxː] < name of the tax office> , BAH [bɔxː] < name of an intersection in Budapest> , etc.). Thus, while doh-type morphemes are an open class, there is only a single lexical item (cseh) that consistently represents the cseh type for all ECH speakers. The rest of the morphemes that are traditionally considered to belong to the cseh class show variation across ECH speakers or even within the speech of one and the same speaker (juh ‘sheep’, mēh ‘bee’, cēh ‘guild’, dūh ‘anger’, rūh ‘scabies’, āhr ‘hunger’, olāh ‘Wallachian’), or have been reclassified as doh-type morphemes (mēh ‘womb’, ?keh ‘wheeziness’) or as vowel-final ones (plēh ‘tin’).\(^{117}\)

\(^{117}\)Of the items that show variation some may be doh type only or cseh type only for a particular ECH speaker. In my own speech most of them clearly belong to the doh class (e.g. mēh ‘bee’ [meːx], mēh-ek [meːhek] (pl.), mēh-et [meːhet] (acc.), mēh-tōl [mexːtɔlː] (abl.)). It can also happen that for the same speaker some forms of a given morpheme show doh-like behaviour while other forms of the same morpheme are cseh-like: for many ECH speakers diūh behaves in this way: diūh [dūx], diūh-tōl [dūxtɔlː, dūtɔːlː] (abl.), diūhroham [dÆrohom,
We conclude that only the doh-type alternation is phonological synchronically, and cseh-type morphemes have two underlying allomorphs, a consonant-final and a vowel-final one, whose selection is morphological. Doh-type stems, on the other hand, are always consonant-final. Given the complementary distribution of [h] and [x], a decision has to be made as to which of the two segments is underlying in these stems. We suggest that /x/ is the underlying segment since in this case the rule can be formulated as the delinking of the C-place node of /x/ in onset position. This is less arbitrary than the strengthening of /h/ into [x] since, if /h/ were taken to be underlying, and it were assumed to acquire DOR (and a C-place node) in the coda, then it would be impossible to identify the source of the DOR feature assigned by the rule as it is not (necessarily) present in the environment of the putative /h/-s in the relevant stems (cf. Siptár 1998b).\textsuperscript{118}

The ‘weakening’ rule can be formulated as follows:

\begin{equation}
\text{(74) Onset} \\
\quad \text{X} \\
\quad \left[ \text{son} \right] \\
\quad \text{C-place [+ cont]} \\
\text{DOR}
\end{equation}

\*düxrohom] ‘a fit of anger’. Note that ā- is a bound stem that only occurs with (some) derivational suffixes and in compounds (e.g. ā-es [e:heš] ‘hungry’, ā-ség [e:xšeɡ, e:šeɡ] ‘hunger’, ākop [e:xkop] ‘[go] hungry’). Inasmuch as keh occurs in ECH at all in isolation and before analytic suffixes, it is a doh-type stem (?keh [kex], ?keh-töl [kextö̊l] (abl.), keh-es [keheš] ‘wheezy’). Plēh behaves exactly like vowel-final vécé’loo’: compare plēh [pleh], plēh-k [plehk] (pl.), plēh-t [pleht] (acc.), plēh-töl [plehtö̊l] (abl.) vs. vécé [vet’e], vécé-k [vet’e:k] (pl.), vécé-t [vet’e:t] (acc.), vécé-töl [vet’e:tö̊l] (abl.). For most speakers oláh is vowel-final, but, exceptionally, (for some speakers) Type A suffixes attach to it with a linking vowel: oláh [ola], oláh-t [olat] (acc.), but oláh-ok [ola:ok] (pl.) (see Papp 1975).

\textsuperscript{118}It is to be pointed out that putative /h/ ~ [x] is also unnatural in the sense that it is ‘strengthening’ in the coda, i.e. in a lenition site. /x/ ~ [h] is somewhat better because here the ‘weakening’ often happens in intervocalic position, which is a typical lenition site (cf. Harris 1990, 1997). Note, however, that in this account, the latter event also happens in initial onsets where lenition typically does not take place. This is just as problematic as the strengthening in the coda in the alternative analysis.
The output of (74) is a placeless fricative. In order for [h] to be derived from the underlying /x/, we have to assume that a placeless [+ cons, -son, + cont] segment is phonetically implemented as [-cons, -son, + cont], i.e. an obstruent glide. This means that implementation in this case should be feature changing. In order to avoid this, we can assume that the segment underlying [h] and [x] is unspecified for [cons]. Given this assumption, (74) derives a placeless non-sonorant continuant (which is not specified for [cons]). The correct surface realizations are derived if we assume two implementation rules: (i) [-son, + cont, DOR] segments have to be implemented as [+ cons] and (ii) placeless non-sonorants as [-cons].

The non-application of (74) to a geminate /xx/ (i.e. the impossibility of surface [h:]) is an instance of geminate inalterability (cf. Kenstowicz and Pyle 1973, Perlmutter 1995). As the (74) explicitly refers to the timing tier, it is to be interpreted exhaustively, i.e. it does not apply to an input in which the segmental content is multiply linked to two timing slots (and is in coda and onset position at the same time), cf. Hayes (1986), Schein and Steriade (1986).

This behavior of geminate [x:] reveals another advantage of the /x/-based account over the /h/-based one. If /h/ were the underlying segment, surface [x:] would have to be derived from an underlying geminate /hh/. The rule that derives surface [x] from a coda /h/ would not apply to a geminate /hh/ because of geminate inalterability (the same reason as above). Note, however, that while non-application yields the correct surface result in the /x/-based account (since the unchanged underlying segmental melody is the attested surface melody), in the /h/-based account an extra rule (specific to /hh/) is needed to make sure that geminate /hh/ surfaces as [x:].

Thus, we conclude that /h/ is not an underlying segment in Hungarian. All instances of [h] are derived from /x/ (or more precisely /X/, which is unspecified for [cons]) by (74).

The rule is not postlexical because it does not apply across a word boundary (dohokosta [doxokosta, *dohokosta] ‘musty smell caused [it]’). Given our assumptions about Block 1 and Block 2 derivation, this means that (74) is a Block 2 rule since it has to apply in non-derived environments (as well), e.g. in holló [hol:o:] ‘raven’, tehén [těhěn] ‘cow’ and it is not a structure building rule. Assuming that (74) is a Block 2 rule makes the prediction that

---

119 This notation is just shorthand for a single root node associated to two timing slots.
[h] (not [x]) occurs preceding vowel-initial analytic (i.e. Block 2) suffixes (like -ig, -ért, -é, -ul/ül, cf. Section 4.1.4.5.) in doh-type words since the stem-final /x/ is syllabified as an onset before a vowel-initial suffix in Block 2. If (74) were a Block 1 rule, [x] would be expected before vowel-initial analytic suffixes in these words because at the point when it applies, the stem-final consonant is still in the coda. It must be pointed out that native speaker intuitions/judgements (including my own) are uncertain on this point. While there are speakers who (claim to) pronounce doh-ért ‘for musty smell’ as [doxe:rt], others feel/make no difference between the pronunciation of the h in bohém ‘bohemian’ (monomorphic), doh-ot (stem+ synthetic suffix) and doh-ért (stem+ analytic suffix). As no experimental evidence or large scale survey is available, we merely point out that the former pronunciation would be problematic since it would involve the application of a Block 1 rule in a nonderived environment.
4.2.3. Fast Cluster Simplification

Clusters consisting of more than two consonants may be simplified in fast speech (cf. Dressler and Siptár 1989, Siptár 1991, Törkenczy & Siptár 1999ab). Fast Cluster Simplification (FCS) is an optional postlexical deletion process that targets consonants flanked by consonants, i.e. it deletes the middle one of a sequence of three consonants, as the examples show below:

(75) lambda [lɔ̃mbə, lɔ̃mə] ‘id.’
asztma [ɔ̃stmə, osmo] ‘asthma’
röntgen [rɔ̃dʒɛn, rɔ̃gɛn] ‘X-ray’
dombtető [domptetɔː, domtɛtɔː] ‘hilltop’

The first approximation of the rule expressing this process can be stated as (76)\(^{120}\):

(76) \[X \ X \ X / \ X \ X\]

FCS is postlexical as it can apply in monomorphemic words (e.g. asztma, röntgen) and across any boundary including synthetic (77a) and analytic (77b) affix boundary, and that between words in compounds (77c) and phrases (77d) (in (77) below the hyphens in the orthographical forms indicate affix boundaries and those between the constituents of compounds, and do not appear in normal spelling):

---

\(^{120}\) I assume that by the time FCS applies postlexical Degemination has already applied, so at this stage words like [sakk] [tábla] ‘chessboard’, [test] [tartás] ‘posture’ [kis] [stil lá] ‘petty’ only have intervocalic CC clusters and thus they do not fall under the purview of FCS.
Apparently, the homorganicity requirement in nasal + stop clusters (within an analytic domain) is not surface true (i.e. the output of FCS may violate it), though note that [n] is required to assimilate if FCS should create a [n]+ non-homorganic stop cluster: röntgen [rön/, *röng], rendben [ren, *ren]. Curiously, [n], an allophone of /n/, does not assimilate under the same circumstances: feszengtem [fesentem, *fesentem], örjöntek [örjöntek, *örjöntek].

Fast Cluster Simplification does not apply to all CCC clusters. For instance, it does not apply to the clusters shown in (78) below:

(78) a. feszeng-t-em [fesentem, fesentem\footnote{121}] ‘feel uncomfortable’ (1sg past)
örjöng-tek [örjöntek, örjöntek] ‘go berserk’ (3pl past indef)

b. rend-ben [røndben, rømben] ‘in order’
paraszt-nak [pøøstnøk, pøøsnøk] ‘for a peasant’

c. lomb-korona [lombokronɔ, lomkoronɔ] ‘foliage of a tree’
test-nevelős [tøshnøveːs, tøshnøveːʃ] ‘PE’

d. dob-d ki [doptki, dopki] ‘throw (it) out’
most pedig [moøtpсидig, moøsp ødig] ‘and now’

The differential behaviour of words like those in (75) and (78) has been used to suggest that FCS is a syllable structure conditioned process. It has been claimed that it applies if C2C3 of a C1C2C3 cluster is not a well-formed onset (e.g. lambdá), but it does not if C2C3 is a well-formed onset (e.g. centrum). In order to account for this pattern one could assume that there is an optional postlexical resyllabification process that moves the last consonant of a branching
coda into the onset of the following syllable. This process would be subject to the general well-
formedness conditions and would be expected to block if the resulting onset is ill-formed—hence
the FCS effect (cf. Dressler and Siptár 1989, Siptár 1991, Ács and Siptár 1994). This
interpretation would be problematic for the present analysis since we claim that onsets may not
branch in Hungarian (see Section 3.2.2). It is to be pointed out, however, that this position can
be shown to be untenable (cf. Törkency and Siptár 1999ab): contrary to what is predicted by
the above interpretation, FCS is not possible if C3 is a continuant even if C2C3 is not a
possible branching onset (granting—for the sake of argument—that branching onsets exist in
Hungarian and assuming that (most) occurring word-initial clusters are well-formed onsets):

(79)  handle [hɒndle:, *hɔnle:]  ‘second-hand dealer’
pántlika [pɑntlika, *pɑnlika]  ‘ribbon’
kompjúter [kompjúter, *komju:ter]  ‘computer’
aktfotó [ɔktfotɒ, *ɔkfotɒ]  ‘nude photograph’
pemzli [pɛmzli, *pɛmli]  ‘brush’ (n.)
hangsor [hɒŋkʃɔr, *hɔŋʃɔr]  ‘sound sequence’

(79) shows that FCS does not apply if C3 is [+ cont] irrespective of the syllabic affiliation of
the consonants in the cluster: compare hangsor [hɒŋkʃɔr, *hɔŋʃɔr] and hangtalan [hɒŋktɔɭɔn,
hɒŋtɔɭɔn] ‘soundless, silent’. Therefore, we conclude that FCS is not sensitive to syllable
structure (and is not a problem for our claim that onsets may not branch in Hungarian).

Thus, the reason why FCS does apply to the relevant clusters in (75), but does not
apply to those in (78) and (79) is that in the former set of words the C3 of the C1C2C3
clusters is [-cont] while in the latter two it is [+ cont]. 122

122There are sporadic examples in which FCS seems to apply although C3 is a continuant:
e.g. szoftver [sofťvɛːr] ‘software’, szendvics [sɛndvɛʃ, sɛnvɪʃ] ‘sandwich’, testvér [teʃvɛr, teʃvɛr] ‘brother’, mumpsz [mʌmps, mums] ‘mumps’. There are two things to be
noted here: (i) Some of the forms that appear to show FCS are actually lexicalized and are not
the result of deletion at all. For instance, for many speakers mumpsz is [mʌmps] regardless of
the tempo. Note that the same cluster cannot be simplified in other items: kolompszó
[kolompsɔː, *kolomso] ‘sound of the cattle bell’; (ii) Most of the problematic examples that
we have found have [v] as C3. One could use this fact to suggest that FCS is gradient rather
than absolute: it is more likely to apply if C3 is [v] than with other continuants. Compare
There are two further conditions on the application of FCS. It does not apply if C1 is a continuant sonorant:

(80) talpnyaló [təlpnˈːəlːoː,*təlnˈːəlːoː] ‘lackey’
bazaltkő [bəzɔltkˈːoː,*bəzőlkːoː] ‘basalt stone’
partner [pɔrtnɛr,*pɔrner] ‘id.’
szerbtől [sərptːoːl,*sərtoːl] ‘from (a) Serb’
sejtmag [ʃɛjtmaːɡ,*ʃɛjmːɡ] ‘cell nucleus’
fajdkakas [fɒjdkːakoʃ,*fɒjkːakoʃ] ‘blackcock’

It is not just the continuancy of C1 that matters here: note that FCS can apply if C1 is [+ son, -cont] (e.g. röntgen [rɔndɡen, rɔŋɡen]) or if it is [-son, + cont] (e.g. asztma [ɔstmo, ɔsmo]).

FCS cannot apply either if C2 is a fricative or an affricate:

\[\text{domtető} \text{ ‘hilltop’, dombvidék ‘hilly region’ and dombról ‘from the hill’. In domtető FCS is definitely possible [domtetːoː], in dombról it is definitely not *[domroːl]; dombvidék [domvideːk] is intuitively somewhere in between. I leave this problem for further research.}\]

\[\text{Naturally, it can also apply if C1 is [-son, -cont]: receptkönyv [rɛtʃɛptkənˈv, rɛtʃɛpkoŋˈv] ‘book of recipes’}.\]

\[\text{In essence, this means that the target must be a plosive. It is not possible to test the behaviour of the various sonorants as C2 because they are either (i) unattested as C2, or (ii) if attested, are preceded by [+ cont] sonorants, which in itself blocks FCS (e.g. modernkedik [mɔɾdnkɛ dik, *mədørkɛ dik] ‘act modern’ (3sg pres.), szörnyben [sɔrnˈbɛn, *sɔrˈbɛn] ‘in (a) monster’, filmtől [*fɪltːoːl ‘from (a) film’). [ɛ] and [j] (which are—with a handful of exceptions—always the post-consonantal surface reflexes of imperative -j) present a further complication: they also seem to be omissible in environments where FCS otherwise blocks. Compare Dobj neki egy törölközöt! ‘Throw her/him a towel! ‘and Dobj rá egy törölközöt! ‘Throw a towel on her/him!’}.\]

\[\text{In fast speech deletion of C2 is possible in both cases ([dɔbɲɛki/dɔbɲki, dɔbɲraː/dɔbraː]) in spite of the fact that in the second FCS cannot apply since the target is followed by a continuant ([r]). We tentatively suggest that [ɛ] and [j] are deleted by a separate optional process that targets [ɛ] and [j] exclusively between any two consonants. Further empirical research is definitely in order.}\]
(81) szenvtelen [sɛntɛlɛn, *sentɛlɛn] ‘indifferent’
eksztázis [ɛktɛzɪʃ, *ektaziʃ] ‘extasy’
Amszterdam [ɔmɛstdɔm, *ɔmɛrdɔm] ‘Amsterdam’
inspekció [inspekiɔ:, *inspekiɔ:] ‘inspection’
obskurus [ɔpʃkuruʃ, *opkuruʃ] ‘obscure’
láncalp [læntˈtɛl)p, *læntˈtɛl)p] ‘caterpillar track’
táncdal [tændˈdɔl, *tændɔl] ‘popular song’
parancsnok [poronɛnɔk, *porɔnɔnk] ‘commander’
narancsból [norɔnˈboːl, *norɔnboːl] ‘from (an) orange’

To sum up, FCS (75) is subject to three conditions: (i) C2, the target consonant, cannot contain the feature specification [-son, + cont], (ii) C3 must be [-cont] and (iii) C1 cannot be [+ son, + cont]. All three conditions must be satisfied in order for FCS to apply. This can be stated as (82)

(82) \[
\begin{array}{c}
X & \varnothing & X & X \\
C_\alpha & C_\beta & [-\text{cont}]
\end{array}
\]

Condition: \( C_\alpha \neq [-\text{son}, +\text{cont}] \)
\( C_\beta \neq [+\text{son}, +\text{cont}] \)

Substrings of clusters longer than three consonants behave in the same way as clusters containing exactly three consonants. For instance, FCS cannot apply to the four term cluster in foxtrott [fɔkstrɔt] ‘foxtrot’ because of the two potential targets (C2 and C3 of C1C2C3C4), the first one cannot delete as it is [-son, + cont] ([*fɔktrot:]), and the second one cannot delete since it is followed by a [+ cont] segment ([*fɔksrot:]). By contrast, FCS can apply to C3 of the C1C2C3C4 cluster in karsztból ‘from (a) karst formation’ [kɔɾzdɔˈbɔːl, kɔɾzdɔˈbɔːl] since all the thee conditions are met (C2 cannot delete because it is [-son, + cont] and is preceded by a [+ son, + cont] segment [*kɔɾzdɔˈbɔːl]).

Like other fast-speech processes, the conditions are relaxed gradually and FCS
Because of the lack of evidence, the exact way in which it happens must be the subject of future research.

FCS is a phonotactically motivated process but it is not conditioned by syllable structure. This suggests/confirms that phonotactics cannot be reduced to SSCs or syllabic organisation. We have already noted that some phonotactic regularities cannot be accounted for without reference to MSCs (cf. section 3.4). FCS, however, is a phonotactically motivated process unrelated either to syllable structure or to morpheme structure. It is only expressible in terms of string adjacency.

Note that, although FCS is a phonotactically motivated process, it is not ‘teleological’ in the sense that it does not apply to avoid or repair an ill-formed string. The surface string that may not be an output resulting from FCS may be perfectly well-formed phonotactically (and attested otherwise): *eksztázis [ɛkstɑːziʃ] although akta ‘document’ [ɔktɔ]; and vice versa the surface string that may be the output of FCS may not be attested otherwise, i.e. when it is not the output of FCS: *feszengtem [fɛsɛŋtem] although otherwise *[ŋC] if C≠[k,g].

Cô té (2000) attempts to explain the constraints on this process (and ultimately all phonotactically motivated processes) with reference to perception cues.

---

125 Because of the lack of evidence, the exact way in which it happens must be the subject of future research.

126 A similarly phonotactically motivated, but not syllable structure conditioned process that involves vowels rather than consonants optionally shortens [a;, e;] before consonant clusters if the second consonant of the cluster following [a;, e;] is [−cont] (cf. Törkenczy & Siptár (1999a)). Compare (a) and (b) below:

a.  Márta < name> [maːɾtɔ, mɑɾtɔ]
    fõrges ‘worm-eaten’ [feːɾɡɛʃ, fɛɾɡɛʃ]

b.  Mátra < place name> [maːɾtɔ, *mɑɾtɔ]
    võgré ‘at last’ [veːɡɾɛ, *veɡɾɛ]
    vãdli ‘calf (of leg)’ [vaːdli, *vɑdli]
    hõtfõ ‘Monday’ [hɛtʃfɔː, *hɛtʃfɔː]
    vãnszorog ‘crawl’ [vaːnsɔɾɔɡ, *vɑnsɔɾɔɡ]

127 Cô té (2000) attempts to explain the constraints on this process (and ultimately all phonotactically motivated processes) with reference to perception cues.
Appendix A.

The feature composition of underlying segments

For ease of reference I list here the underlying segments in stems and suffixes in Hungarian together with their feature compositions that I assume in the analysis presented in the dissertation. The feature compositions shown here are identical with those in Siptár and Törkenczy (2000) and the reader is referred to that work for arguments for them. The reader must also bear in mind that the actual representations assumed are feature-geometrical (see section 2.2) and the matrices below are just shorthand for the corresponding feature-trees.\footnote{This is why in the case of affricates (which are assumed to be contour segments) the symbol ‘±’ appears in the line [cont]: this is meant to be interpreted as a ‘−’ in the ‘left half’ of the feature tree and as a ‘+’ in the ‘right half’ and not as a third ‘±’ feature value.}

Naturally, in the formulation of the rules and regularities discussed in the dissertation, clear and overt reference is made to the feature geometry whenever it is crucial. Note that length is not represented in the charts below since quantity is considered as a structural property (to be represented on the timing tier) and not a melodic one. The surface vowel-pairs [ɛ, eː] and [a, aː] are assumed to be melodically identical underlyingly, i.e. /ɛ/ and /a/ respectively.\footnote{For a discussion of some consequences of this assumption, see section 3.4.2.}

\textbf{(i) Consonants} (see Siptár and Törkenczy 2000: 93)

\begin{center}
\begin{tabular}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|}
\hline
& \textbf{LABIAL} & \textbf{DENTAL} & \textbf{PALATAL} & \textbf{VELAR} \\
\hline
\textbf{LAB} & & & & & & & & & & & & & & & \\
\textbf{COR} & & & & & & & & & & & & & & & \\
\textbf{DOR} & & & & & & & & & & & & & & & \\
\hline
\hline
\textbf{p} & b & f & v & m & t & d & s & z & tʰ & l & r & tʰ & dʰ & š & ž & č & j & nʰ & j & k & g & x \\
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\hline
\end{tabular}
\end{center}
(ii) **Vowels in stems** (see Siptár and Törkenczy 2000: 157)

<table>
<thead>
<tr>
<th></th>
<th>i</th>
<th>ü</th>
<th>ü</th>
<th>ö</th>
<th>o</th>
<th>e</th>
<th>a</th>
</tr>
</thead>
<tbody>
<tr>
<td>COR</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>•</td>
<td></td>
</tr>
<tr>
<td>LAB</td>
<td>•</td>
<td></td>
<td></td>
<td>•</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DOR</td>
<td>•</td>
<td></td>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
<td></td>
</tr>
<tr>
<td>[open&lt;sub&gt;1&lt;/sub&gt;]</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>[open&lt;sub&gt;2&lt;/sub&gt;]</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

(iii) **Vowels in suffixes** (see Siptár and Törkenczy 2000: 158)

<table>
<thead>
<tr>
<th>NON-ALTERNATING</th>
<th>ALTERNATING</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>e:</td>
</tr>
<tr>
<td>COR</td>
<td>•</td>
</tr>
<tr>
<td>LAB</td>
<td></td>
</tr>
<tr>
<td>DOR</td>
<td></td>
</tr>
<tr>
<td>open&lt;sub&gt;1&lt;/sub&gt;</td>
<td>–</td>
</tr>
<tr>
<td>open&lt;sub&gt;2&lt;/sub&gt;</td>
<td>–</td>
</tr>
</tbody>
</table>
Appendix B.

Lists

Note: exhaustive lists are marked [full]; the number of tokens (stems) in a given type appears in angled brackets: < >

(i) Monomorphemic stems containing three-member medial clusters

pšk  obskurus [full]
psl  kapszli, mopszli, snapszli [full]
pst  absztinens, szubsztancia [full]
pst’ obszcn [full]
ptr  dioptria [full]
ksh  exhibicionizmus, exhortáció, exhumál [full]
ksk  exkommunikál, exkurzió, exkuzál [full]
ksl  dakszli [full]
ksn  fakszni [full]
ksp  expanzió, expediál, expedi ció, experimentál, exponál, exponens, export, expozi ció, expozíció [full]
kst  eksztázis, extempore, extenzi v, exteríör, exterritoriális, szextáns, szextett, textil, textus [full]
kst’ excellenciás, excentrikus [full]
ltr  bliktri, direktrisz, doktriner, elektro-, oktrojál, spektrum [full]
fst  bifsztek [full]
skr  diszkrá [full]
skv  biszkvit, diszkvalifikál, moszkvai, paszkvillus [full]
spn  eszpresszó, veszprán [full]
sth  keshthelyi, poszthumusz [full]
stm  asztma [full]
str adminisztratív, aposztrofál, asztrahánprénés, asztrakán, asztro-, ausztrál, ballusztrád-moti vum, bisztró, esztireng, gazstronómia, illusztráció, illusztris, indusztrialis, kasztrál, katasztrális, katasztrófa, kosztros, magisztrátus, nosztrifikál, osztrák, ósztriga, paszträna, pesztra, pisztráng, plasztron, regisztrál [full]
štr alabástrom, capistráng, destruál, flastrom, istráng, kimustrál, klastrom, ministrál, mustra, ostrom, [full]
štv testvér [full]
vdb luftballon [full]
zdr puzdra [full]
mlb embléma [full]
mbr ambrózia, ámbra, embrió, membrán, rembrandt-film [full]
mfl pamflet, strimflí [full]
mpl gimpli, krumpli, komplekszum, komplett, komplex, komplexum, komplexus, komplé, komplikál, plömplöm, stempli, szimpla, templom, trampli, zempláni [full]
mpr impregnál, impresszárió, impresszió, impresszum, imprimatúra, imprimál, imporduktív, improvízál, kompresszor, kompromisszum, [full]
mpt szimptóma [full]
mst amszterdamí [full]
mzl pemzli [full]
nčk trencskót [full]
ndg avantgard, röntgen [full]
ndl handlé, mándli, pemdlizik, svindli, vámdlíz [full]
ndr pondró, slendrián, szalamandra, tundra, condra, hipochondria, mándruc, plundra [full]
ndv szendvics [full]
nfl infláció, influenza, konfliktus, konflis [full]
nfr infra- [full]
nsf transzformáció [full]
nsk ruszinszkói [full]
nsp transzpárens, transzponál, transzport [full]
nšp konspiráció, inspekció, inspiciál, inspiráció [full]
nst  dinsztel, dunsztol, falanszter  [full]
nšt  installáció, instancia, instál, konstans, konstantinápolyi, konstatál, konstábler, konstelláció, konsternáció, konstitúció [full]
nst’  transzcendens [full]
ntl  pántlika [full]
ntr  antracán, antracit, antropológia, centralista, centrifuga, centrum, dzsentri, filantróp, hőzentráger, intranzigens, intráda, intrika, koncentráció, kontra, kontraktus, kontraszt, kontrol, mizantróp, premontrei [full]
nt’v  kifi runcvancigol ? [full]
ńgl  Anglia, konglomerátum, ringli, ringlispí l, ringló, stangli, tinglitangli, zsonglőr [full]
ńgr  kongresszus, csongrádi, gangrána, kongregáció, kongrúa [full]
ńgv  disztvingvál, pingvin, szangvinikus [full]
ńkf  frankfurti [full]
ńkl  inklináció, inklúzi ve, konklávé, konklzió, onkli [full]
ńkp  pingpon [full]
ńkr  bankrott, inkriminál, konkrá, pankráció, szinkron [full]
ńkt  adjunktus, konjunktúra [full]
ńkt’  disztinkció, funkció, interpunkció, szankció [full]
ńkv  inkvizi ciô [full]
ńfr  wolfram [full]
ńft  elvtike [full]
ńgr  Belgrád [full]
ńkl  fóklór [full]
ńsk  szilszkin [full]
ńst  colstok [full]
ńtr  altruista, filtrál, poltron, ultra [full]
ńrl  kurbl, ferblí [full]
ńrd  gardrób, ordré [full]
ńrdv  mordvin [full]
ńrgl  kvargli, smirgli [full]
ńrkl  cirkli, verkli [full]
(ii) Monomorphemic free CC-final verb stems

Note: some if the items containing the codas /st, nt, l:/ are/may be analysed as polymorphemic

| d: | < 1> | fedd [full] |
| d': | < 2> | edz, pedz [full] |
| g: | < 3> | csügg, függ [full] |
| t': | < 1> | metsz [full] |
| st | < 117> | aggaszt, akaszt, alátámaszt, alvaszt, apaszt, áraszt, bágyaszt, bigyeszt, |
bomlaszt, borzaszt, ...

št  < 1>  fest [full]
zd  < 2>  kezd, küzd [full]
žl  < 1>  esd [full]
mt  < 1>  teremt [full]
(mz  < 1> ) nemz [full]
nt  < 100>  bánt, biccent, billent, bont, bosszant, bölint, bööffent, buffant, buggyant, csahint,…
nd  < 4>  csikland, fecskend, mond, örvend [full]
 nz  < 1>  vonz [full]
ng  < 51>  állong, bolyong, borong, borzong, búsong, csapong, csellegg, cseng, csüng, dong, dühöng, dülöng, eseng, feszeng, fetreng, forrong, hajlong, hőzöng, hullong, ing, jajong, kering, kong-bong, lappang, lázong, leng, lézeng, mereng, ömleng, ödöng, őrjöng, rajong, rång, ring, setteng, sikong, szállong, szorong, teng, teng-leng, tolong, töpreng, ujjong, verseng, villong, visong, zajong, zeng, zsibong, zsong, zsong-bong [full]
(nl  < 1> ) ajánl [full]
lt  < 14>  fájt, kelt, kiált, költ, olt, ölt, rikolt, sikolt, süvölt, tult, tőlt, üvölt, vált, volt [full]
ld  < 7>  áld, fold, küld, old, száguld, told, told-fold [full]
(lg  < 1> ) uralg [full]
l:  < 34>  áll, csekdyell, dall, drágáll, furcsáll, fuvall, gyengől, hall, hátall, hosszall, hull, javall, kall, kell, kevesell, későll, koráll, lövell, nagyoll, nehezell, pall, restell, rivall, rosszall, rühell, sarkall, sokall, sugall, száll, szégyell, szökkell, vall, várell, vákell [full]
rt  < 8>  árt, árt, gyár, irt, ki sárt, márt, sárt, tart [full]
rd  < 2>  hord, kárd [full]
r:  < 2>  forr, varr [full]
jt  < 19>  bujt, ejt, fejt, felejt, fojt, gyűjt, gyűjt, hajt, lejt, nyűjt, ojt, óhajt, rejt, sejt, sőhajt, sújt, szakajt, szalajt, veszejt [full]
(iii) Verb stems containing intervocalic vsi-clusters

Note:  **Strict** examples have been set in bold, **moderate** examples have been underlined and endings have been CAPITALISED.

(a) vsi-cluster= non-geminate

n= 735

\[\text{pt} \quad < \quad 5 \quad > \quad 
\text{cspTET, kapTAT, adoptÁL, akceptÁL, kooptÁL [full]} \\
\text{pf} \quad < \quad 2 \quad > \quad 
\text{cupfOL?, klopfOL [full]} \\
\text{ps} \quad < \quad 1 \quad > \quad 
\text{abszolvÁL [full]} \\
\text{pl} \quad < \quad 3 \quad > \quad 
\text{átcaplAT, táplÁL, kopLAL [full]} \\
\text{pr} \quad < \quad 3 \quad > \quad 
\text{deprimÁL, tőprENG, reprodukÁL [full]} \\
\text{tk} \quad \text{vetKÖZIK, követKEZIK, kotkodácsOL, hagyatKOZIK …} \\
\text{tv} \quad < \quad 1 \quad > \quad 
\text{ötvÖZ [full]} \\
\text{tl} \quad < \quad 1 \quad > \quad 
\text{csatlAKOZIK? [full]} \\
\text{tr} \quad < \quad 3 \quad > \quad 
\text{kotrÖDIK, fetrENG, patronÁL [full]} \\
\text{th} \quad < \quad 1 \quad > \quad 
\text{rothAD [full]} \\
\text{tk} \quad < \quad 3 \quad > \quad 
\text{bütykOL, lótykÖL, ügyKÖDIK [full]} \\
\text{tv} \quad < \quad 1 \quad > \quad 
\text{kotyvASZT [full]} \\
\text{t\textsuperscript{2}m} \quad < \quad 1 \quad > \quad 
\text{fitymÁL, [full]} \\
\text{kt} \quad < \quad 12 \quad > \quad 
\text{affektÁL, annektÁL, bakTAT, iktAT, diktÁL, kecsegTET, lükTET, nyugTAT?, paktÁL, praktizÁL, rugTAT, vágTAT [full]} \\
\text{kt\textsuperscript{2}} \quad < \quad 1 \quad > \quad 
\text{akceptÁL [full]} \\
\text{ks} \quad < \quad 2 \quad > \quad 
\text{fixi rOZ, maximÁL [full]} \\
\text{kš} \quad < \quad 1 \quad > \quad 
\text{kuksOL [full]} \\
\text{kv} \quad < \quad 1 \quad > \quad 
\text{likvidÁL [full]} \\
\text{km} \quad < \quad 2 \quad > \quad 
\text{lakmárOZIK, tukmÁL [full]} \\
\text{kl} \quad < \quad 9 \quad > \quad 
\text{akklimatizÁL, csakliZ, csiklandOZ, deklemÁL, ödekítÖDIK, öklEL, proklamÁL, reklamÁL, zaklAT [full]} \\
\text{kr} \quad < \quad 1 \quad > \quad 
\text{akkreditÁL [full]} \]
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Example Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>bz</td>
<td>képzEL, tobzÓDIK</td>
</tr>
<tr>
<td>bź</td>
<td>habzsOL, lebzsEL</td>
</tr>
<tr>
<td>bl</td>
<td>öblÍ T, táblábOL, szublimÁL</td>
</tr>
<tr>
<td>br</td>
<td>babrÁL, celebrÁL, ebrudAL, õbrED, fabrikÁL, vibrÁL, zabrÁL</td>
</tr>
<tr>
<td>dr</td>
<td>addresszÁL</td>
</tr>
<tr>
<td>d'b</td>
<td>lögybÖL</td>
</tr>
<tr>
<td>gd</td>
<td>szökdEL, szökdácsEL</td>
</tr>
<tr>
<td>gz</td>
<td>egzaminÁL, egzeci rOZ, egzisztÁL, rögzÍ T, vegzÁL</td>
</tr>
<tr>
<td>gn</td>
<td>agnoszkÁL, dezignÁL, ignorÁL, impregnÁL, regnÁL, stagnÁL, szignÁL</td>
</tr>
<tr>
<td>gl</td>
<td>fogLAL, tagLAL</td>
</tr>
<tr>
<td>gr</td>
<td>degradÁL, emigrÁL, integrÁL</td>
</tr>
<tr>
<td>tk'</td>
<td>dömöckÖL, evickÚ, kéredzKEDIK, lubickOL, packáZ, peckEL, pöckÖL</td>
</tr>
<tr>
<td>t'm</td>
<td>kecmerEG</td>
</tr>
<tr>
<td>čk</td>
<td>fröcskÖL, pacskOL, pocskondiáZ, tapicskÁL</td>
</tr>
<tr>
<td>čm</td>
<td>becsmérEL</td>
</tr>
<tr>
<td>fl</td>
<td>caflAT, bifláZ</td>
</tr>
<tr>
<td>fr</td>
<td>lófrÁL</td>
</tr>
<tr>
<td>sp</td>
<td>diszponÁL, koszpítOL</td>
</tr>
<tr>
<td>st</td>
<td>adjusztÁL, asszisztÁL, biTAT?, dehonesztÁL, desztillÁL, egzisztÁL, engesztEL, gesztikuLÁL, invesztÁL, közTET, magasztAL, marasztÁL, molesztÁL, neheztEL?, osztOZIK, piszerkÁL, predesztinÁL, pusztUL, tapasztAL, tiszTEL, vesztegEL, vesztEGET, vigasztAL</td>
</tr>
<tr>
<td>sk</td>
<td>agnoszkÁL, alkalmazKODIK?, bámészKODIK, baszkurÁL, császkÁL, csimpaszKODIK, ereszKEDIK, fuoreszkÁL, foszforeszkÁL, hunyászKODIK, iszkOL, kapaszKODIK, mászkÁL, maszki rOZ, merészKEDIK, motoszkÁL, nyerészKEDIK, piszkÁL, poroszkÁL, prüszkÖL, reszkET, rugaszKODIK, tuszkOL, viszkET</td>
</tr>
<tr>
<td>sm</td>
<td>eszmŰ, piszmOG, szöszmötÖL</td>
</tr>
<tr>
<td>sl</td>
<td>ézlEL</td>
</tr>
<tr>
<td>šp</td>
<td>aspirÁL, hospitÁL, náspángOL, prosperÁL, respektÁL, tespED</td>
</tr>
<tr>
<td>št</td>
<td>böstorKÖDIK, instÁL, istápol, karisztOL, kóstÁL, kóstOL, lósTAT, protestÁL</td>
</tr>
</tbody>
</table>
pustOL, restELL, rostokOL, sisterEG, testÁL, tüsténKEDIK [full]

şk < 10> ágasKODIK, áskálÓDIK, böbiskOL, diskurÁL, diskÁL, kapiskÁL, miskárOL, paskOL, roskAD, viaskKODIK [full]

şm < 5> ismer, ismáEL, kasmatOL, pusmOG, susmOG [full]

şl < 4> búslaKODIK?, koslAT, pislANT, pisiOG? [full]

şh < 4> keshED, kushAD, peshED, poshAD [full]

vl < 1> szí vLEL? [full]

zd < 3> ?buzdí T, mozdí T, rezdí T [full]

zg < 4> birizgÁL, izgAT, izgUL, igazGAT [full]

zv < 1> közvetí T [full]

žl < 2> ?hi zLAL, ?i zLEL [full]

žd < 2> ?pezsdiL, pezdí T [full]

žg < 2> babusGAT, rebesGET [full]

žm < 1> prézsmitÁL [full]

žl < 1> ?vizzslAT [full]

mp < 16> amputÁL, csimpaszKODIK, extemporÁL, hemperEG, hómpölyÖG, imponÁL, importÁL, kámpicsorODIK, kalimpÁL, kompenzÁL, korrumpÁL, pumpOL, sompolyOG, szimpatizÁL, temperÁL, tempi rOZ [full]

mt < 1> nyomtat? [full]

mb < 13> ambicionÁL, bömbÖL, csámborOG, dörömbÖL, gémberEDIK, himbÁL, imbolyOG, kombinÁL, különbÖZIK, rombOL, tombOL, zörömbÖL, zsémbEL [full]

mč < 2> csámcsOG, csembcsEG [full]

mf < 1> somfordÁL [full]

mž < 3> hemzsEG, karimzsÁL, körömzsÁL [full]

ml < 5> emlEGET, emlí T, kamél? , szemLÁL? , számLÁL? [full]

nt < 31> argumentÁL, bántalmAZ?, bünTET, fontOL, frekventÁL, garantÁL, henterEG, intež, integrÁL, interjívOL, internÁL, interpolÁL, interpretÁL, interveniÁL, intonÁL, kántÁL, kommentÁL, kontemplÁL, kurrentÁL, lamentÁL, orientÁL, parentÁL, plántÁL, plántÁL, poenti rOZ, prezentÁL, reprezentÁL, tántorí T, tántorOG, tünTET, vonTAT [full]
bánKÓDIK, berzenKEDIK, bosszanKODIK, ellenKEZIK, ficánkOL, hetvenKEDIK, konkurrÁL, kunkOG, kunkorODIK, lankAD, pironKODIK, rimánKODIK, settenKEDIK, sopánKODIK, tehánKEDIK?, tüstánKEDIK [full]

andalÍ T, apprehendÁL, bandukOL, csendÜL, gondOL, gondOZ, gründOL, indÍ T, indUL, kirándUL, komendÁL, kondicionÁL, kondoleÁL, kondUL, landOL, lendÜL, lendÍ T, menDEGÍL, ondolÁL, penderEDIK, pendÜL, pocskondiÁZ, rándÍ T, rándUL, rekomendÁL, rendEL, rendÜL, sandÍ T, skandÁL, spendi rOZ, sündörÖG, szenderEDIK, szundÍ T, szundikÁL, tendÁL, tündöklik, ugrándOZIK, undorÍ T, vindikÁL, zendÜL, zsendÍ T [full]

ácsingÓZIK, barangOL, böngÍSZ, csatangOL, csilingEL, dangubÁL, döngÖL, engED, engedelmesKEDIK, engeszEL, flangÁL, kurjonGAT, langallik, marcangOL, násángOL, pingÁL, pironGAT, rongÁL, szoronGAT, tágÁL, tengÖDIK [full]

boncOL, cincÁL, cincOG, hancúrOZIK, hencEG, koncOL?, koncentrÁL, kuncOG, poncOL, ráncigÁL, szincerizÁL, toloneC?, viháncOL [full]

dancsOL?, inceselKEDIK, koncsorOG, kuncsorOG, linesEL?, pancesOL?, trancsi rOZ [full]

dezinficiÁL, inficiÁL, informÁL, konferÁL, konferÁL, konfirmÁL, konfundÁL, tánferEG [full]

konzolidÁL, lanszi rOZ, unszOL, vonszOL [full]

invesztÁL, invitÁL, involvÁL, konveniÁL, konvergÁL, konvertÁL, konverzÁL, szenvED [full]

inzultÁL, kompenzÁL, kondenzÁL, konzultÁL [full]

avanzÁL, revanzÁL [full]

ki nlÓDIK?, szi nlEL? [full]

inhalAL [full]

szontyolODIK [full]

göngyÖL, kalangyÁL [full]

senvvED [full]

kolportÁL [full]
áltAT, alterÁL, erölTET, inzultÁL, költÖZIK, konzultÁL, kultivÁL, máTAT, öltÖZIK?, váltOZIK [full]

álmA KODIK, alkoT, alkuszizzk, bövelKEDIK, bujálKODIK, emelKEDIK, foglalKOZIK, gyalKOL, hivalKODIK, kalkulÁL, kalkulÁL, lelkendEZ, sivalKODIK, találKOZIK, veselKEDIK, viselKEDIK, vulkanizÁL [full]

kiakolbóDI T [full]

áltOZ, foltOZ, fuldoklik?, haldoklik?, koldUL, nyeldekEL, oldalOG, üldÖZ [full]

dádelGET?, dolgoZIK, hallGAT, latolGAT?, találGAT, vulgarizÁL [full]

lefalcOL? [full]

abszolvÁL, alVAD, devalvÁL, dvEZ, involvÁL, megnyilvánUL, olVAD, olvas [full]

alkalmAZ, bántalmAZ?, cirkalmAZ, dallmahODIK, diadalmasKODIK, fogalmAZ, forgalmAZ, sugalmAZ, szolmizÁL, tartalmAZ [full]

volna? [full]

cirpEL, hörpINT, hörpÖL, horpAD, horpasZT, interpellÁL, interpolÁL, terpESZT?, torpAD, torpAN [full]

abortÁL, amortizÁL, artikulÁL, csörTET, deportÁL, dezertÁL, disszetÁL, értekEZIK, értelmEZ, étesUL, exportÁL, fertőZ, firtAT, importÁL, irtóZIK, kertEL, kolportÁL, konvertÁL, portÁL, szorti rOZ, tartázTAT, tartOZIK, történik, törTET, zsörtölÖDIK [full]

fortyAN, fortyOG, hortyOG, szörtyÖG, szortyOG, vartyOG [full]

acsarKODIK, babirkÁL, birKÖZIK, böstöRKÖDIK, botorkÁL, butkOL, cirkÁL, cirkalmAZ, cirkulÁL, érKEZIK, fárkOZIK, fárkÉSZIK, hamarKODIK, horkAN, horkOL, iparkODIK?, ismerKEDIK?, kárKEDIK?, kurkÁSZ, marki rOZ, pirkAD, pisztérkÁL, pörkÖL, sarkALL, serKEN, serkENT, szemerkÁL, szerkESZT, szurkOL, zárKÖZIK [full]

dorbézOL, turbékOL, verbuvÁL, zurbOL [full]

cserdI T, cserdÜL, csikordUL, csordI T, csordOGÁL, csordUL, csurdI T, dördÜL, édekel, fordI T, fordUL, gardi rOZ, gördI T, gördÜL, hörđÜL, irdUL, kérEZ, koordinÁL, mordUL, ordiBÁL, perdI T, perdÜL, serdÜL, somfordÁL,
tördEL?, verdes?, zördÍ T, zördÜL [full]

rd < 1> meggárgyUL [full]
rg < 23> argumentÁL, dezorganizÁL, dörgÖL, gargalizÁL, gurguláZ, háborGAT, hánytorGAT, hergEL, horgAD, horgOL, kapirogÁL, kergET, kergÜL, konvergÁL, kuperGAT, kurGAT, nyargAL, organizÁL, organizáL, sürGET, tekergÖDZIK, vergÖDIK, zarGAT [full]
rt < 10> exorcizÁL, farcinÁL, giürÖL, hurcOL, percENT, percipiÁL, porcOG, sercEG, sercEN, varcOG [full]
rê < 1> szürcsÖL [full]
rf < 5> cserfEL?, köntörfalAZ?, morfondi rOZ, perfektuÁL, perforÁL [full]
rs < 3> forszí rOZ, persziflÁL, perszonifikÁL [full]
rš < 5> harsAN, harsOG, hersEG, másékEL, társalOG, versENG [full]
rv < 7> hervAD, hervASZT, nyervOG, örvendid?, sorvAD, sorvASZT, szervEZ [full]
rz < 4> borzONG, borzAD, borzOL, toporzáOL [full]
rž < 6> dörzÖL, horzOL, morzOL, perszEL, porzOL, torzsalKODIK [full]
rm < 14> alarné rOZ, derrMED, derrMÉZT, dörmÖG, fürmED, harmonizÁL, informál, körmÖL?, konfirmál, mormOG, mormOL, permutÁL, szármAZIK, zsurmOL [full]

rn < 2> garné rOZ, internÁL [full]
nr < 7> ernyED, ernyESZT, görnyED, kornyikÁL, szörnyED, vernyákOL, vernyOG [full]
r1 < 5> birLAL, ötEL?, porlAD?, siúrlÓDIK?, törlesZT [full]
rj < 15> burjánZik, erjED, erjESZT, gerjED, kimarjUL, kurjANT, kurjONGAT, örjÍ T, örjÖNG, sarjAD, sarjázik, sorjÁZ, terjED, terjENG, terjESZT [full]
rh < 1> archaizÁL [full]
jp < 1> selypí T [full]

it < 6> sajtol, rejtőZIK, fejtEGET, kojtol, kujtorOG, folyTAT [full]
jk < 3> hajkurtÁSZIK, sulykOL, vájkÁL [full]
jd < 11> bolydUL, bujdOKOL?, bujDOS?, bujDOSIK?, gajdOL, jajdUL, kajdÁSZ, sajdíl T, sajdUL, sejdií T, zajdUL [full]
jg < 4> bolyGAT, stáigerOL, totójGAT, zajiGAT [full]
js < 2> hajszOL, maíszOL [full]
(b) vsi-cluster= geminate

n= 281

p:  < 24>  csappAN, cseppEN, csippENT, cuppAN, cuppASZT, cuppOG, frappi rOZ, huppAN, klappOL, koppAN, koppANT, lappAD, lappang, leppEG, rëppEN, süppED, szëppEN, szippANT, töppED, toppAN, toppASZT, toppESZT, zëppOL, zëppAN [full]

t:  < 14>  csattAN, csattANT, csattOG, csëttEN, csëttENT, kattAN, pattAN, pattOG, ragadTAT, rettEG, rettEN, rettENT, suttOG, tetTET [full]

t’: < 16>  fittyED, karattyOL, kettyEN, kettyENT, kettyINT, kottyAN, lottyAN, pittyED, pottyAN, pottyANT, rittyENT, suttyAN, suttyANT, szottyAN, tottyAN, zöttyEN [full]

k:  < 35>  akkomodÁL, akkumulÁL, bukkAN, csökkEN, csökkENT, kukkAN, kukkANT, kukkOL, meghökkEN, meghökkENT, mükkan, nyékkEN, nyékkan, okküpÁL, pükkan, pükkanT, pükkanST, rekkEN, rikkAN, rikkANT, rokkAN, rükkOL, sikkAD, sikkANT, sikkASZT, smákOL, szikkAD, szökkEN, szökkASZT, tikkAD, tikkASZT, vakkAN, vakkANT, zökkEN, zökkENT [full]

b:  < 19>  csöbbAN, döbbEN, döbbENT, dobbAN, dobbANT, fellebbEZ, gubbASZT, lebbEN, lebbENT, lobbEN, lobbAN, lobbANT, meghibbAN, rebbEN, robbAN, robbANT, rokkAN, zsibbAD, zsibbASZT [full]

d:  < 2>  edDEGÁL, idDOGÁL [full]

d’: < 5>  biggyED, biggyESZT, buggyAN, buggyASZT, rogyAN [full]
aggASZT, aggAT, aggÓDIK, csüggED, csüggESZT, faggAT, guggOL, higgAD, higgASZT, nyaggAT, nyegeET, raggAT, szaggAT, szaggerÁL [full]
biccEN, biccENT, biccEL, döccEN, koccAN, koccANT, moccANT, pöccEN, pöccENT, ruccAN, spriccEL, truccOL [full]
becsicsENT, feccsENT, fröccsENT, loccsAN, loccsANT, reccsEN, reccsENT, traccesOL [full]
menedzsEL [full]
biccEN, biccENT, bliccEL, döccEN, koccANT, koccANT, loccsANT, raggAT, saggAT, szaggAT, szaggerÁL [full]
becsicsENT, feccsENT, fröccsENT, loccsAN, loccsANT, reccsEN, reccsENT, traccesOL [full]
menedzsEL [full]
 adresszÁL, asszimilÁL, bosszankODIK, bosszANT, csosszAN, csusszAN, deklasszÁL, disszertÁL, dresszi rOZ, masszi rOZ, nyisszAN, nyisszANT, passzÍ T, passzOL, pisszEG, pisszEN, prüsszENT, prüsszÖG, sisszENT, szisszEN, szusszANT, tüsszEN, tüsszENT, tüsszÖG [full]
tussOL, tess&EL? [full]
duzzAD, duzzASZT, duzzOG, izzAD, izzÍ T?, rezzEN, rezzENT, zizzEN [full]
brummOG, cammOG, hümmÖG, kommentÁL, kommunikÁL, nyámmOG, nyammOG, stimmÉL, zümmÖG [full]
dünnyÖG, fonnyAD, fonnyASZT, gunnyASZT, konnyAD, sunnyOG, szunnyAD, vinnyOG [full]
alliterÁL?, apellÁL, ballag, billeg, billEN, billENT, brilli rOZ, cizzellÁL, cseillENG, csillAN, csillG, csillAPÍ T, csillapODIK, csillOG, desziillÁL, düllED, düllESZT, füllED, füllENT, füllESZT, fullAD, fullASZT, gyullAD, illAN, illEG-billEG, ilLEGET, ilLESZT, kallÓDIK, kellET, pallóROZ, pillANT, pillÓZIK, vállal, villAN, villoG, villONG, züllESZT [full]
cserrEG, csörrEN, csörrENT, csurrAN, csurrANT, durrAN, durrANT, durrÖG, herrEG, kurrÖG, surrAN, surrÖG, vırrAD, vırrASZT, zörrEN, zörrENT [full]
süllyED, süllyESZT, ujjONG, vıjjÖG [full]
kehhENT [full]
Bibliography

Note: The symbol ♦ appears before titles that contain a detailed/comprehensive description of the phonotactic pattern of Hungarian.


Csúcs Sándor (1968) A magyar szókészlet finnugor elemei a statisztika tükrőben *Nyelvtudományi Közlemények* LXX, 392-6.


Deme László (1945) Statisztikai vizsgálatok a hangátvetés köröből *Magyar Nyelv* XLI, 55-60.


Horger Antal (1932): A mássalhangzók kieséséhez Magyar Nyelv XXVIII, 140-5.


Hyman, Larry M. (1984). On the Weightlessness of Syllable Onsets. Brugman, C; Maccaulay, Monica; Dahlstrom, Amy; Emanatian, Michele; Moonwoman, Birch; O’Connor, Catherine (eds.) *Proceedings of the Tenth Annual Meeting of the Berkeley Linguistics Society*, Berkeley: Berkeley Ling. Soc., Univ. of California 1-14


Kenstowicz, Michael and Charles Kisseberth (1979). *Generative Phonology*. New York:
Academic Press.


Kiparsky, Paul (1985). Some Consequences of Lexical Phonology. *Phonology Yearbook* 2,
85–138.


Historica 4.2. 151-180.
Mártonfi Ferenc (1977). Jövevánszavak hangalakjának meghatározása fonológiai
automatával. Általános Nyelvészeti Tanulmányok 12. 185-220.
Inquiry 12, 373– 418.
Inquiry 17, 207– 263.
pp. 84-108.
University Press.
McCarthey, John J. and Alan Prince (1993) Prosodic Morphology I: Constraint Interaction
and Satisfaction. Ms. University of Massachusetts, Amherst and Rutgers University.
Beckman, Laura Walsh Dickey and Suzanne Urbanczyk (eds.) Papers in Optimality
Mass.: Graduate Linguistic Student Association. 249-384.
Mester, R. Armin and Junko Itô (1989). Feature Predictability and Underspecification:
Mohanan, Karuvannur Puthanveettil (1985). Syllable Structure and Lexical Strata in
English. Phonology Yearbook 2. 139-155.
Mohanan, Karuvannur Puthanveettil (1986). The Theory of Lexical Phonology. Dordrecht:
Reidel.
Mikes Ferenc (1936). Magyar i-rógépbillentyűzet Gyorsí rástudomány XVI, 93-5.
Dressler, Hans C. Luschutzky Oskar E. Pfeiffer and John R. Rennison (eds.)
Theory 5, 485– 518.
Nádasdy, Ádám (1989a). The Exact Domain of Consonant Degemination in Hungarian in
Linguistica Hungarica 39, 195– 213.


Szőpe György (1967). A magyar generatív fonológia néhány kérdése. *Nyelvtudományi
Értesí tő. 58. 305-310.

Tompa József (1956). Mai tárgyragos szóalakjaink egy típusáról Magyar Nyelv LII, 177-


Törkenczy, Miklós (2000b). Absolute phonological ungrammaticality: defective paradigms in Hungarian. 8th Manchester Phonology Meeting, Manchester [handout].


Treiman, Rebecca and Catalina Danis (1988). Syllabification of Intervocalic Consonants. Journal of Memory and Language 27.1. 87-104


Vértész Edit (1952). Adalékok a magyar nyelv hangtani szerkezetéhez (Első közlemény). *Nyelvtudományi Közlönyek* LIV, 96-140.


