

Extrametricality and minimal word length in Hungarian

Stephen Grimes
stgrimes@indiana.edu
McWOP 11
November 4-6, 2005

1 Preliminaries

The status of extrametricality in Hungarian is matter of debate in the Hungarian phonology literature (cf. Siptár and Törkenczy, 2000) In this talk, I will both survey phenomena supporting extrametricality and also reanalyze data that have been used to argue against extrametricality. In addition, I will also examine new data that show how Hungarian's word minimality constraint interacts with extrametricality. In the end, I conclude that adopting extrametricality for Hungarian is compatible with the data presented.

Hungarian preliminaries

- Hungarian has binary length distinctions for vowels and consonants.
- Standard, Budapest Hungarian is generally assumed to have seven short-long vowel pairs.

(1)	Short Vowels	Long Vowels
	ü i u	u í ú
	ö e o	o é ó
	a	á

- The two low vowels e/é and a/á form a natural class with respect to several phonological processes, for example, Low Vowel Lengthening (cf. Vago, 1980). These two vowel pairs are also the only pairs that differ in *quality* as well and quantity.

Extrametricality preliminaries

- In the literature, extrametricality (Ito, 1989) and related phenomena can often be synonymous with extraprosodicity, Non-Finality constraints (Prince and Smolensky, 1993), Peripherality Condition (Hayes, 1980) or other so-called edge effects.
- In this talk, *extrametricality* is used to denote final consonant extrametricality, or even more precisely, extramoraicity. A word final consonant is not a member of the prosodic word domain. It is therefore not a member of the syllable and does not contribute to syllable weight.
- One effect of extrametricality is that we should observe an asymmetry in the possible syllable weights between word-final and non word-final positions. This is due to the prevention of word-final consonants from having moraic status generally attributed to a constraint on weight-by-position (Hayes, 1989).

2 A survey of arguments for extrametricality in Hungarian

Extrametricality issue #1: the lexicon

- Compare maximal syllables allowed in word-final and non word-final positions. There is an asymmetry between final and non-final syllable positions.
- As shown below in the first column of (2), long and short vowels are permitted word-finally before a single consonant.
- However, in the 2nd and 3rd columns of (2), excepting the low vowels, long vowels are not permitted to appear before a consonant cluster
- Comparing the bottom halves of the 1st and 3rd columns, we see that a tautosyllabic long vowel-consonant sequence is permitted word-finally, but not word-internally. This asymmetry in allowable weight disappears if the word-final consonant is not weight-bearing.

(2) Gaps in the lexicon of monomorphemic words (after Siptár and Törkenczy, 2000).

	VC#	VCC#	VC.C
i	hit 'belief'	ring 'sway'	inger 'stimulus'
ü	sün 'hedgehog'	csüng 'hang'	kürto 'funnel'
ö	sör 'beer'	gyöngy [d'önd'] 'pearl'	ördög 'devil'
e	nem 'gender'	szent 'saint'	persze [perse] 'of course'
u	fut 'run'	must 'grape juice'	undor 'digust'
o	lop 'steal'	gyors 'fast'	boglya [bogja] 'stack of hay'
a	hat 'six'	tart 'hold'	apró 'tiny'
í	sír 'grave'	----	----
u	bun 'sin'	----	----
o	bor 'skin'	----	----
é	kém 'spy'	érc 'ore'	érték 'value'
ú	rút 'ugly'	----	----
ó	kór 'disease'	----	----
á	láp 'marsh'	márt 'dip'	árpa 'barley'

Despite the asymmetry observed in (2), Siptár and Törkenczy do not consider it evidence for extrametricality because no vowel length alternations take place – these are static forms in the lexicon. They suggest that the data in (3), which show that superheavy syllables may surface when polymorphemic forms are considered, illustrate that the constraint on maximal syllable weight is a constraint on the morpheme, not the syllable.

(3) Polymorphemic forms

	VCC#	VC.C
í	sír-t	szív-tam
u	bun-t	bun-ben
o	fö-bb	bor-ben
é	kér-t	kér-ték
ú	túr-t	túr-nak
ó	kór-t	kór-nak
á	vár-t	vár-tam

- The data in (3) are apparently problematic for an extrametricality analysis because the expected asymmetry in syllable weight between word-final and non word-final positions disappears.
- The forms in (3) are permitted to have superheavy syllables due to a high ranked Output-to-Output or Paradigm Uniformity constraint that dominates the phonotactic constraint concerning syllable weight.

Extrametricality issue #2: syllable weight with respect to stress

Kerek (1971) reports facts concerning variable secondary stress and how it treats syllable weight.

- Primary stress in Hungarian always falls on the first syllable of the word.
- Secondary stress is typically assigned to odd-numbered syllables after the initial syllable (3rd syllable, 5th syllable, etc.)
- Stress shift: Secondary stress may optionally shift to away from a light syllable onto an adjacent heavy syllable. The example in (4) shows two possible stress patterns for a four-syllable word with a light penultimate syllable and heavy final syllable.

(4) **S** S **S** L S **H** ~ S S S L S **H** (bold face indicates stress)

What constitutes heavy syllables for stress shift?

(5)

	Light	Heavy
Word internal	CV	CVV, CVC, CVVC
Word final	CV, CVC	CVV, CVVC

Adopting extrametricality would cause the asymmetry between word internal and word final positions to disappear.

Extrametricality issue #3: Internal representation of geminates

- There is no reason to suppose that geminates in Hungarian are 'fake' geminates (cf. Hayes, 1986); they cannot be split by epenthesis nor do phonological processes act on one half of the geminate.
- Following Hayes, a 'true' geminate is represented underlyingly by a single root with a single mora attached.

(6) a. Light b. Heavy
 sok *sokk*
 'many' 'shock'

- If the form in (6a) were to receive a mora by weight-by-position, its representation would be indistinguishable from (6b).

3 Minimal word length

- The minimal word condition (Hayes, 1980; McCarthy and Prince, 1986/1996) grew out of research on metrical theory and templates. It assumes that prosodic words must contain a metrical foot, typically consisting of two syllables.
- Many languages allow monosyllabic words, and in these cases a so-called degenerate foot is allowed, a foot consisting of a single syllable, presuming this syllable is minimally bimoraic. This is the case in Hungarian.

(7) $\text{PrWd}_{\text{MIN}} = \mu\mu$ The minimal word in Hungarian is bimoraic.

- Function words are typologically immune from the condition in (7). The same is true in Hungarian, as pronouns, question words, and several separable verbal prefixes of the form CV would violate the minimal word condition.
- There are two content word exceptions: *fa* 'tree' and *ma* 'today'.
- Word minimality is not entirely without controversy – cf. final mid, rounded vowels must be long (cf. Törkenczy, 1994)
- Vacillation of high vowel length word-finally may occur in polysyllabic words, but remains long in monosyllabic words.

4 Minimal word length and extrametricality

I can now present additional arguments related to extrametricality, in conjunction with the minimal word condition. The key assumption here is that, under the assumption of extrametricality, a CVC form (where the vowel is short) **does not meet** the requirement of the minimal word condition.

Extrametricality/minimal word issue #1: Vowel length in monosyllabic words

It is my intuition that in comparing monosyllabic CVC and CVVC word forms, the vowel is found to be long surprisingly frequently! If this is the case, the scarcity of CVC words suggests their failure to meet the minimal word condition.

Frequency statistics

- My corpus is a dictionary (word list) compiled by András Kornai. This list was tagged for part of speech, and I selected nouns, verbs, adjectives, and adverbs to use as content words.
- Because I worried that only examining word forms (type frequencies) might distort the characteristics of the language, I also compiled word frequencies (token frequencies) to illustrate my point. These were from an online corpus (Halácsy et al., 2004)

How many "superheavy" – CVVC or CVCC – syllables should we expect?

(8) Percentages of words containing a CVVC or CVCC syllable

Two syllable word superheavy first syllable	Two syllable word superheavy final syllable	One syllable word superheavy (final) syllable
13.% of types 8.8% of tokens	34.1% of types 29.5% of tokens	73.6% of types 46.9% of tokens

- Consider the first column in (8). These words are almost exclusively compounds or multimorphemic words. Otherwise there virtually no superheavy word-internal syllables permitted.
- The fact that "superheavy" syllables are freely permitted word finally indicates that extrametricality is active.
- The higher rate of "superheavy" syllables in monosyllabic words is presumably due to the minimal word condition.

A closer look at the quality of the vowels

In monosyllabic words, many of the exceptions to the minimal word requirement under the addition of extrametricality contain "low" vowels: CaC or CeC.

The table in (9) shows that low vowel words constitute 87.8% of the token frequencies that violate the minimal word condition under extrametricality.

(9)

**Distribution of short vowels in
one syllable C₀VC words**

	Types	%	Tokens	%
a	72	26.1%	3372953	23.0%
e	64	23.2%	9487668	64.8%
i	21	7.6%	427409	2.9%
o	47	17.0%	621353	4.2%
u	21	7.6%	327504	2.2%
ö	34	12.3%	329757	2.3%
ü	17	6.2%	68976	0.5%
total	276	100.0%	14635620	100.0%

In the table in (10), monosyllabic words containing a non-low vowel are considered. In this case, 77% of the word types and 80.8% of the word tokens meet the bimoraic requirement under extrametricality.

(10)

One syllable content words

Low vowels excluded

	Type	%	Token	%
μ . C ₀ V	0	0%	0	0.0%
μ . C ₀ VC	140	23.0%	1,774,999	19.2%
$\mu\mu$.C ₀ VV	31	5.1%	1,187,687	12.8%
$\mu\mu$. C ₀ VVC, C ₀ VCC, or C ₀ VVCC	437	71.8%	6296365	68.0%

The prevalence of monosyllabic words with bimoraic weight under extrametricality is explained by the minimal word condition.

Extrametricality/minimal word issue #2: Germanic loan words

Nádasdy (1989) describes a case of "unmotivated" consonant gemination in recent borrowings (since 1750) primarily from German into Hungarian. Short consonants in German or English, languages which lack geminates, are borrowed as geminates in Hungarian. Some examples of this borrowing for monosyllabic words are given in (11), although the consonant lengthening process also applies to some consonants in longer words, both in word-final and word-internal positions.

(11) Examples of consonant lengthening

lakk	(<Ger. Lack)	'lacquer'
sokk	(<Ger. Schock)	'shock'
tipp	(<Ger. Tipp)	'idea'
meccs [mecc]	(<Eng. match)	'match'
blöff	(<Eng. bluff)	'bluff'

The pattern of borrowing monosyllabic words closed by a single consonant is described as follows:

- If the vowel is borrowed as long, no consonant lengthening occurs.
- If the vowel is borrowed as short, the consonant is always borrowed as a geminate.

Based on this generalization, germanic loanword phonology respects word minimality under extrametricality.

Extrametricality/minimal word issue #3: vowel length "alternations"

There is a closed class of words that alternate vowel length between nominative and suffixed forms.

(12)	Nominative	Accusative	
	víz	vizet	'water'
	nyár	nyarat	'summer'
	kéz	kezet	'hand'
	jég	jeget	'ice'

Here we can take the underlying form of the vowel to be short. Vowel lengthening presumably takes place due to the constraint on word minimality applying due to an extrametrical final consonant. Note, however, that these words must be lexically marked to allow for stem vowels that do not alternate in length.

5 Summary

The purpose of this talk has been to survey arguments for and against extrametricality in Hungarian. I hope to have advanced the argument for the existence of extrametricality by showing effects related to the minimal word condition.

The table in (13) summarizes the phenomena discussed in this talk.

(13)

Summary of evidence relating to extrametricality in Hungarian			
Evidence	Supports extrametricality?	Notes	Suggestive of diachronic extrametricality?
Asymmetry in allowable size of a maximal syllable word finally vs. word internally	yes	Confusion as to whether maximal syllable is a syllable or morpheme constraint	yes
Secondary stress weight sensitivities	yes		-
Geminate moraic representation	yes	Other geminate representations possible	-
Minimal word length – vowel length in CVC	yes	Approximately 20% of words remain unaccounted for	yes
Gemination in Germanic loan words	yes	Polysyllabic words may also have "unmotivated" gemination	yes
Vowel length alternations in nominal paradigm for a closed class of nouns	yes	Frozen vowel length "alternations"	yes

Final thought: Diachronic or synchronic?

- How can we tell whether extrametricality or word minimality are still active constraints in Hungarian?
- Lexicon may reflect output of formerly active constraints or rules
- Extremely recent English loanwords no longer undergo consonant germination seen in earlier Germanic loanwords.

References

- HALÁCSY, PÉTER, KORNAI, ANDRÁS, NÉMETH, LÁSZLÓ, RUNG, ANDRÁS, SZAKADÁT, ISTVÁN and TRÓN, VIKTOR 2004. Creating open language resources for Hungarian Paper presented at 4th International Conference on Language Resources and Evaluation (LREC2004).
- HAYES, BRUCE. 1980. A Metrical Theory of Stress Rules: MIT Ph.D. dissertation
—. 1986. Assimilation as spreading in Toba Batak. *Linguistic Inquiry*, 17.467-99.
—. 1989. Compensatory lengthening in moraic phonology. *Linguistic Inquiry*, 20.253-306.
- ITO, JUNKO. 1989. *Syllable Theory in Prosodic Phonology*. New York: Garland Press.
- KEREK, A. 1971. *Hungarian metrics: Some linguistic aspects of iambic verse: Uralic and Altaic Series*. Bloomington: Indiana University.
- MCCARTHY, JOHN J. and PRINCE, ALAN. 1986/1996. *Prosodic Morphology 1986*. New Brunswick, NJ: Rutgers University Center for Cognitive Science
- PRINCE, ALAN and SMOLENSKY, PAUL. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. Rutgers University Center for Cognitive Science Technical Report 2.
- SIPTÁR, PÉTER and TÖRKENCZY, MIKLOS. 2000. *The phonology of Hungarian*. Oxford: Oxford University Press.
- TÖRKENCZY, MIKLOS. 1994. A szótag [The syllable]. *Strukturális magyar nyelvtan, II kötet: Fonológia [Structural Hungarian Grammar, volume II: Phonology]*, ed. by Ferenc Kiefer, 272-392. Budapest: Akadémiai Kiadó.
- VAGO, ROBERT M. 1980. *The Sound Pattern of Hungarian*. Washington, D.C.: Georgetown University Press.

APPENDIX: Other idiosyncrasies with respect to low vowels

- Low vowel lengthening
- Qualitative as well as quantitative distinction
- Törkenczy (1994) notes that the low vowels 'é' and 'á' are in fact permitted in extra-heavy, monomorphemic syllables p.343 -> márt, férc
- Contradiction: Low long vowels can act as monomoraic as in (2), while in CaC or Cec words violating the minimal word condition, the short low vowels appear to act as long

This can be interpreted as Ident-Low-Vowel >> Phonotactics >> Ident-Vowel

Misc notes

Vowel Length alternations upon suffixation

két ~ ketto 'two'

jég ~ jegen 'ice'

Rebrus, pc: depends on whether CC sequence is falling or rising in sonority.

So, maximal syllable is VC] or VV], but VVC] seems dispreferred

(5)	a. Light	b. Heavy	c. Heavy	d. Superheavy
	s	s	s	s
	μ	μμ	μμ	μμμ
	šok	šok	a l	a l
	sok	sokk	ál	áll
	'many'	'shock'	'spurious'	'stand'

Possible Questions:

(1) Coerced moraicity to explain why CvC forms can surface as short.

This is good because they actually have duration

I cannot therefore distinguish between geminates and singletons