Testing variability effects in Hungarian vowel harmony

Ágnes Benkő, Fanni Patay, Ágnes Lukács, Katalin Mády, Péter Rebrus, Miklós Törkenczy

ICSH13 29–30 June 2017, Budapest
Variation in backness harmony

We *have known (felt)* for a long time that

HVH is determined but also underdetermined by phonology

Phonologically identifiable zones of variation: [N], [BN], [BN⁺]

- **lexical variation:** *haver-ok* vs. *koncert-ek*
- **VACILLATION:** *fotel-ok* vs. *fotel-ek*

**Height Effect:** [Bi(:)]<[Be:]<[Bɛ]  x<y means: y is more “front” than x
- *papír-ok* < *tányér-ok, szatén-ek/ok* < *fotel-ek/ok, koncert-ek, haver-ok*

**Count Effect:** [BN]<[BNN⁺]
- *parti-ra* < *alibi-ra/re, horribilis-ra/re*
- *tányér-ra* < *klarinét-re/ra, protézis-re/ra*
### Problems with “traditional” approaches


- **impressionistic**
  - data based on introspection (*horr. dict.* hearsay)
- **categorical**
  - variation, neutrality, transparency, etc. considered Y/N
- **incidental**
  - gradience/variation considered random/inessential
- **unintegrated**
  - gradience/variation not built into phonological analysis
Empirical studies (Hayes & Cziráky Londe 2006)

We know (with more certainty) that

- The Height Effect & the Count Effect manifest themselves in (type) frequency (measured in backness/frontness ratio of stem types)
- Native speaker reaction (based on wug testing) matches the result of the corpus study

but: H & CL present a simplified picture

- 1 suffix (dative)
- multiple application of HE (HE & CE combined)
There is every indication that the patterning of variation in the zone is far richer

- stem-final consonants (Hayes et al. 2009)
- homogeneity of harmonic suffix behaviour: C-initial vs V-initial (R & T 2013)
- harmonic domain: gradience of morphological complexity (R & T 2017)
- multiple HE (HE & CE combined): [BN_xN_y] vs [BN_zN_w] (R & T 2016)
Corpus studies: variation as relative frequency

- type or token frequency (április problem)
- data sparseness problem
- ratio of front/back suffixed forms to all harmonically suffixed forms
- Frontness ratio

\[
F\text{-ratio} = \frac{\text{number of front suffixed forms}}{\text{number of front suffixed forms} + \text{number of back suffixed forms}}
\]
## Multiple Height Effect: [BNN]_ context

### Types of roots

<table>
<thead>
<tr>
<th></th>
<th>i(ː)</th>
<th>eː</th>
<th>ɛ</th>
</tr>
</thead>
<tbody>
<tr>
<td>i(ː)</td>
<td>[Bii] <em>alibi</em></td>
<td>[Bie] <em>klarinét</em></td>
<td>[Biɛ] <em>kabinet</em></td>
</tr>
<tr>
<td>eː</td>
<td>[Bei] <em>protézis</em></td>
<td>[Bee] <em>Athéné</em></td>
<td>[Beɛ] <em>konténer</em></td>
</tr>
</tbody>
</table>
Cumulative application: definition

let $x, y, z$ be neutral vowels

Cumulative interaction between Ns

(i) Height Effect for $N_2$: if $[Bx] \leq [By]$ then $[Bzx] \leq [Bzy]$

(ii) Height Effect for $N_1$: if $[Bx] \leq [By]$ then $[Bxz] \leq [Byz]$

(iii) transitivity: if $[Bx_1x_2] \leq [By_1y_2]$ and $[By_1y_2] \leq [Bz_1z_2]$ then $[Bx_1x_2] \leq [Bz_1z_2]$
Cumulative application: ordering

Orderings by transitivity are not indicated.
Cumulative application: F-ratios (Szószablya)

[Bee] vs. [Bεe]  ábécé (compound-like)  (otherwise) cumulativity holds

[Beε] vs. [Bεε]  Angelesban (1 token)
Locality

(iv) Locality Effect: if $[B_x] \leq [B_y]$ then $[B_{yx}] \leq [B_{xy}]$

3 additional ordered pairs: $[Bei] \not\leq [Bie] \quad [B\varepsilon i] \leq [B\varepsilon i] \quad [B\varepsilon e] \leq [B\varepsilon e]$
Other effects: V-final and C-final [BNN] roots

Internal harmonic consistency: $[Bxy\#] \approx [BxyC]$  
(where X≈Y means that the F-ratio of X is not significantly lower/higher than that of Y)

$[Bii\#] \approx [BiiC]$  
$[Bie\#] \approx [BieC]$  
$[Bi\varepsilon\#] \approx [Bi\varepsilon C]$  

$[Bei\#] \approx [BeiC]$  
$[Be\varepsilon\#] \approx [Be\varepsilon C]$  

$[B\varepsilon i\#] \approx [B\varepsilon iC]$  
$[Be\acute{e}\#] \approx [B\varepsilon eC]$  
$[B\varepsilon \varepsilon\#] \approx [B\varepsilon \varepsilon C]$
The Experiment

A sentence completion task with missing inflections

A pre-recorded sentence presented acoustically, in which the target inflection (and sometimes another syllable in the sentence) is masked with a cough

The context makes it clear which inflection is missing:

A trópusi kolibrinak/nęk kék a tollazata.

The tropical hummingbird.DAT blue plumage.
Benefits

• Allows us to examine the differences in variability in production without relying on metalinguistic awareness and conscious decision about the F/B variants.

• Crucially, it also allows us to collect data for stem+suffix combinations that are rarely or never attested in the corpus, thus providing new sets of data for systematically testing the above hypotheses.
Participants

21 adults: 14 women (19-66; 33 average), 7 men (26-67; 44 average)

18 right-handed, 2 left-handed, 1 ambidextrous

19 currently living in Budapest (10 born & raised)
Stimuli and procedure

104 stems, 4 suffixes (DAT, INS, ELA, ALL) - 416 sentences

74 [BN] (37 [Bɛ] + 37 [Be]); 3 [BBe], 3 [NBe]

(type [Bi] was disregarded: no variation, all are back)

24 [BNN] (6 [Bii], 6 [Bie], 3 [Bei], 6 [Bɛi], 3 [Bɛe])

(type [BNɛ] was disregarded: no variation, all are front)

Some types contain C-final and V-final stems

The target stems have frequencies of equal magnitude in the corpus, and the number of stems in each type corresponds to its frequency in the corpus
Stimuli and procedure

Two versions: each participant heard every target stem with two different suffixes (208 target sentences), and 208 filler sentences:

- no disharmonic stems
- no target suffixes
- 4-6 words long

Dependent variable: frontness of target suffix
Results by generalized type - Count Effect

Significant main effect of type \( (F(3, 76) = 29.34; \ p < .001) \); post hoc: BNN significantly different from all others \( (p < 0.001) \), no other pairwise differences are significant
Results by generalized type+suffix - Count Effect
Height Effect

Significant main effect of subtype ($F(3, 76) = 144.71; p<.001$);
post hocs: NBé = BBé, all other pairwise differences are significant ($p<0.001$)
Results by type
Cumulativity I.

Significant main effect of subtype ($F(3, 76) = 22.3; p<.001$); post hocs: $B_{ei} = B_{ii}$, all other pairwise differences are significant ($B_{ie} < B_{ei}$: $p<0.001$, others $p<0.01$)
Cumulativity II.

Significant main effect of subtype ($F(3, 76) = 94.08; p<.001$); post hocs: all pairwise differences are significant ($Beé < Béi: p<0.001$, others $p<0.01$)
F-ratio of C# and V# roots (internal consistency)

|      | webcorpus |          | experiment |          |     |  
|------|-----------|-----------|------------|-----------|-----|-----|-----|
|      | C# | V#   | diff. | C# | V#  | diff. | p   |
| Bii  | 0.51| 0.63 | -0.12 | 0.50| 0.76 | -0.25 | 0.002|
| Bié  | 0.75| 0.33 | 0.42  | 0.45| 0.21 | 0.24  | 0.001|
| Bei  | 0.70| 0.77 | -0.07 | 0.73| 0.86 | -0.13 | 0.065|
Comparison with the webcorpus results

<table>
<thead>
<tr>
<th>types</th>
<th>webcorpus</th>
<th>experiment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bi</td>
<td>Bé</td>
<td>Be</td>
</tr>
<tr>
<td>Bii</td>
<td>Bié</td>
<td>Bie</td>
</tr>
<tr>
<td>Béi</td>
<td>Béé</td>
<td>Bée</td>
</tr>
<tr>
<td>Bei</td>
<td>Beé</td>
<td>Bee</td>
</tr>
</tbody>
</table>
Conclusion

We examined the harmonic behaviour of existing words combined with four different suffixes, based on an experiment with real people, and we provided a statistical analysis.

The results were not affected by stem frequency (vs sparsity problem in corpora).

In a $[BN_1N_2]$ stem, the quality of $N_1$ and the quality of $N_2$ are equally important (i.e. Cumulativity).

What’s next: behaviour of the suffixes themselves.


Rebrus, Péter and Törkenczy Miklós. 2013. Magánhangzó-diszharmónia. Talk delivered to MTA Nyelvtudományi Intézet, Budapest on 22 October


This work has been supported by National Scientific Grant NKFI-119863 ‘Experimental and theoretical investigations of vowel harmony patterns’

Special thanks are due to Bertalan Polner for his invaluable help in the experiments.
Cumulativity I. with suffixes
Cumulativity II. with suffixes