Non-phonological factors of phonological variation
A large scale wug-experiment for Hungarian vowel harmony

Tamás Biró and Mihály Füredi

Eötvös Loránd University

OCP 13 (Old World Conference in Phonology), 16 January 2016.
Vacillating stems in Hungarian

Hungarian [±back] vowel harmony:

- *asztal* [ɒstɒl] ‘table’ + Dative –nAk = asztalnak.
- *függőny* [fygːɔɲ] ‘curtain’ + Dative –nAk = függőnynek.
- *fotel* [fotɛl] ‘armchair’ + Dative –nAk = fotelnak ~ fotelnek.

Backness = probability $P$ of B-suffix, influenced (at least) by:

- Stem’s vowel pattern: vacillating stems typically back $V^+ + \{ɛ, eː, i\}^+$
- Stem’s fine-grained structure of V and C qualities (e.g., Hayes et al., 2009)
- Stem’s semantic-stylistic properties. (e.g. Forró, 2013)
- Suffix (case)
- Speaker’s dialect (Blaho and Szeredi, 2013)
- Speech rate (Hetényi and Biró, Wednesday)
- What else?
Vacillating stems in Hungarian

Hungarian [±back] vowel harmony:

\[
\begin{align*}
    \text{asztal} & \quad [\text{asztál}] & \quad \text{‘table’} & \quad + \text{ Dative } -n\text{Ak} = \text{asztaln}ak. \\
    \text{függöny} & \quad [\text{függőny}] & \quad \text{‘curtain’} & \quad + \text{ Dative } -n\text{Ak} = \text{függöny}n\text{ek}. \\
    \text{fotel} & \quad [\text{fotel}] & \quad \text{‘armchair’} & \quad + \text{ Dative } -n\text{Ak} = \text{fotel}n\text{ak}\sim\text{fotel}n\text{ek}.
\end{align*}
\]

Backness = probability $P$ of $B$-suffix, influenced (at least) by:

- Stem’s vowel pattern: vacillating stems typically back $V^+ + \{ɛ, eː, i\}^+$
- Stem’s fine-grained structure of V and C qualities (e.g., Hayes et al., 2009)
- Stem’s semantic-stylistic properties. (e.g. Forró, 2013)
- Suffix (case)
- Speaker’s dialect (Blaho and Szeredi, 2013)
- Speech rate (Hetényi and Biró, Wednesday)

What else?
Vacillating stems in Hungarian

Hungarian [± back] vowel harmony:

\[
\begin{align*}
\text{asztal} & \quad [\text{ɒstɒl}] \quad \text{‘table’} \quad + \text{Dative } \rightarrow n\text{Ak} = \text{asztalnak}. \\
\text{függöny} & \quad [\text{fygːøn}] \quad \text{‘curtain’} \quad + \text{Dative } \rightarrow n\text{Ak} = \text{függönynek}. \\
\text{fotel} & \quad [\text{fotɛl}] \quad \text{‘armchair’} \quad + \text{Dative } \rightarrow n\text{Ak} = \text{fotelnak} \sim \text{fotelnek}.
\end{align*}
\]

Backness = probability \( P \) of B-suffix, influenced (at least) by:

- Stem’s vowel pattern: vacillating stems typically back \( V^+ \) + \{ɛ, ɛː, i\}^+
- Stem’s fine-grained structure of V and C qualities (e.g., Hayes et al., 2009)
- Stem’s semantic-stylistic properties. (e.g. Forró, 2013)
- Suffix (case)
- Speaker’s dialect (Blaho and Szeredi, 2013)
- Speech rate (Hetényi and Biró, Wednesday)

What else?
Overview

1. Wug-tests
2. Design
3. Results
4. Conclusions
Overview

1. Wug-tests
2. Design
3. Results
4. Conclusions
Wug-tests for Hungarian vowel harmony

Our starting point:


Women in the Middle Ages used *hádél* to wash clothing. Back then, *hádél* grew abundantly in the fields. It is very hard to find nowadays, but it is said that *hádélnak* or *hádélnek* had a wonderful fragrance. (Hayes and Londe, 2006:70)
Wug-tests (Berko, 1958):

- Have native speakers generate inflected forms of novel stems.
- Demonstrate productive morpho-phonological rules/patterns.
- As with any experimental design, can there be experimental artefacts?

“Frames and instructions were composed with the goal of encouraging the subjects to treat the stems as long-forgotten but authentic words of Hungarian, rather than as recent loans.” (Hayes and Londe, 2006:70)

- Intuition of some native speakers: old Hungarian words more likely to receive back suffixes than recent loans. Is it really so? (Cf. closed class of antiharmonic stems.)
- More generally, does the frame also influence the suffix choice?
- Are there other (non-phonological factors) affecting allomorphy?
Wug-tests Design Results Conclusions

Wug-tests for Hungarian vowel harmony

Wug-tests (Berko, 1958):

- Have native speakers generate inflected forms of novel stems.
- Demonstrate productive morpho-phonological rules/patterns.
- As with any experimental design, can there be experimental artefacts?

“Frames and instructions were composed with the goal of encouraging the subjects to treat the stems as long-forgotten but authentic words of Hungarian, rather than as recent loans.” (Hayes and Londe, 2006:70)

- Intuition of some native speakers: old Hungarian words more likely to receive back suffixes than recent loans. Is it really so? (Cf. closed class of antiharmonic stems.)
- More generally, does the frame also influence the suffix choice?
- Are there other (non-phonological factors) affecting allomorphy?
Overview

1. Wug-tests
2. Design
3. Results
4. Conclusions
Matched-pair design

- Reproducing Hayes & Londe, 2006 with different frames: contrasting old Hungarian to new foreign + observe further factors.

- Weather event, old Hungarian context:
  Each year in the Middle Ages, the population of the Great Hungarian Plain prepared for the arrival of the hádél. The hádélnak or hádélnek involved a sudden fall in temperature and much precipitation. We have to ascribe the extinction of more species [to] hádélnak or hádélnek.

- Weather event, new foreign context:
  Each year, the growing population of Antarctica prepares for the arrival of the hádél. The hádélnak or hádélnek involves a sudden fall in temperature and much precipitation. We have to ascribe the extinction of more species [to] hádélnak or hádélnek.
Matched-pair design

- Reproducing Hayes & Londe, 2006 with different frames: contrasting old Hungarian to new foreign + observe further factors.

- Weather event, old Hungarian context:
  Each year in the Middle Ages, the population of the Great Hungarian Plain prepared for the arrival of the *hádél*. The *hádél* involved a sudden fall in temperature and much precipitation. We have to ascribe the extinction of more species [to] *hádélnak* or *hádélnek*.

- Weather event, new foreign context:
  Each year, the growing population of Antarctica prepares for the arrival of the *hádél*. The *hádél* involves a sudden fall in temperature and much precipitation. We have to ascribe the extinction of more species [to] *hádélnak* or *hádélnek*.
Experimental material

- Reproducing Hayes & Londe, 2006 with different frames: contrasting old Hungarian to new foreign + observe further factors.
  - Self-coded. Snowball launched on Facebook, as well as [nyest.hu](http://birot.web.elte.hu/ragozas/).
- Number of participants: $N = 2999$ (frameset 1), $N = 689$ (frameset 2).

- Wug words: from earlier experiment
  - Minor adjustments: avoid phonemes unlikely in foreign words (e.g., [$n$]). All words with initial C (no need to adjust definite article).
  - Targets: 5 strongly vacillating (*hádél, poribit, kolén, vuszék, vánél*), 2 barely vacillating, dominantly back (*pozin, monil*).
  - Fillers: 3 non-vacillating back (*szandat, kánit, bortog*), 5 non-vacillating front (*zefét, petlér, fánedeg, luteker, kálendel*).
  - (Vacillating vs. non-vacillating: according to the 2006 study.)
Experimental material

- Reproducing Hayes & Londe, 2006 with different frames: contrasting old Hungarian to new foreign + observe further factors.
- Number of participants: $N = 2999$ (frameset 1), $N = 689$ (frameset 2).

- **Wug words:** from earlier experiment
  - Minor adjustments: avoid phonemes unlikely in foreign words (e.g., [ŋ]). All words with initial C (no need to adjust definite article).
  - Targets: 5 strongly vacillating (*hádél, poribit, kolén, vuszék, vánél*), 2 barely vacillating, dominantly back (*pozin, monil*).
  - Fillers: 3 non-vacillating back (*szandat, kánit, bortog*), 5 non-vacillating front (*zefét, petlér, fánedeg, luteker, kálendel*).
  - (Vacillating vs. non-vacillating: according to the 2006 study.)
Experimental material

- Reproducing Hayes & Londe, 2006 with different frames: contrasting old Hungarian to new foreign + observe further factors.
- Number of participants: $N = 2999$ (frameset 1), $N = 689$ (frameset 2).
- Wug words: from earlier experiment
  - Minor adjustments: avoid phonemes unlikely in foreign words (e.g., [ŋ]). All words with initial C (no need to adjust definite article).
  - Targets: 5 strongly vacillating (*hádél, poribit, kolén, vuszék, vánél*), 2 barely vacillating, dominantly back (*pozin, monil*).
  - Fillers: 3 non-vacillating back (*szandat, kánit, bortog*), 5 non-vacillating front (*zefét, petlér, fánedeg, luteker, kálendel*).
  - (Vacillating vs. non-vacillating: according to the 2006 study.)
Frames: inspired by earlier experiment. In each frameset,
- Targets – two domains: old Hungarian context (7 frames), and new foreign context (7 frames).
- Targets – 6+1 ontological categories (2 frames each): human, animal, plant, artefact, naturally occurring object, natural force (weather events) + personal name.

Motivation: relevant categories in developmental psychology (e.g., Keil 1979) and the cognitive science of religion (Boyer 1994). Different ontological categories subject to different folk-theories, different inferences, different association networks.

Fillers – 11 frames non-specified for domain, various or unclear for ontological category.

Similarly to Hayes and Londe (2006): type wug words twice, first in nominative case, then in dative case. Boring?
Matched-pair design

- “Proto-patterns”, such as FBFWFBBWFWBWFW, where W = target wug word, F = front filler, B = back filler. (Always start with FB or BF.)
- “Patterns”, such as FBNFHBOPFTBOFA, where N = personal name, H = human, C = weather condition, etc.
- A random back filler wug word for each B. A random front filler wug word for each F. A random target wug word for each N, H, etc.
- Even-numbered subjects: 4 new foreign domain frames, and 3 old Hungarian domain frames.
- Matched (odd-numbered) subject: same questionnaire, but mirrored for target frame domains.

<table>
<thead>
<tr>
<th>Subject</th>
<th>filler frame 2</th>
<th>filler frame 7</th>
<th>old H pn</th>
<th>filler frame 5</th>
<th>new F hum</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>2n</td>
<td>fr filler ww 3</td>
<td>ba filler ww 1</td>
<td>target ww 2</td>
<td>fr filler ww 2</td>
<td>target ww 6</td>
<td>...</td>
</tr>
<tr>
<td>2n + 1</td>
<td>filler frame 2</td>
<td>filler frame 7</td>
<td>new F pn</td>
<td>filler frame 5</td>
<td>old H hum</td>
<td>...</td>
</tr>
<tr>
<td></td>
<td>fr filler ww 3</td>
<td>ba filler ww 1</td>
<td>target ww 2</td>
<td>fr filler ww 2</td>
<td>target ww 6</td>
<td>...</td>
</tr>
</tbody>
</table>
Matched-pair design

- Within Experiment 1 (or within Experiment 2), contrast
  - for given target wug word, and ontological category,
  - dative suffix allomorph in old Hungarian context vs.
    dative suffix allomorph in new foreign context.
  - Subjects $2n$ vs. $2n + 1$: only difference is domain, all other factors (ontological category, fillers, order, etc.) being the same.

- Between Experiment 1 and Experiment 2, contrast
  - for given target wug word, and ontological category and domain,
  - dative suffix allomorph in Experiment 1 vs.
    dative suffix allomorph in Experiment 2.
  - Subjects $k^{(1)}$ vs. $k^{(2)}$: only difference is frame text, all other factors (ontological category, domain, fillers, order, etc.) being the same.
Overview

1. Wug-tests
2. Design
3. Results
4. Conclusions
Reproducing Hayes and Londe 2006

Overall backness of specific wug words (same ranks, larger values):

<table>
<thead>
<tr>
<th></th>
<th>Exp 1</th>
<th>Exp 2</th>
<th>Hayes &amp; Londe (*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>hádél</td>
<td>0.45</td>
<td>0.41</td>
<td>0.27</td>
</tr>
<tr>
<td>poribit</td>
<td>0.31</td>
<td>0.28</td>
<td>0.34</td>
</tr>
<tr>
<td>kolén</td>
<td>0.43</td>
<td>0.44</td>
<td>0.36</td>
</tr>
<tr>
<td>vuszék</td>
<td>0.59</td>
<td>0.57</td>
<td>0.42</td>
</tr>
<tr>
<td>vánél</td>
<td>0.54</td>
<td>0.54</td>
<td>0.45</td>
</tr>
<tr>
<td>pozin</td>
<td>0.94</td>
<td>0.94</td>
<td>0.92</td>
</tr>
<tr>
<td>monil</td>
<td>0.95</td>
<td>0.94</td>
<td>0.92</td>
</tr>
</tbody>
</table>

(NB: mo[n]il !)

E.g., based on H&L, one might think hádél ‘quite fronter’ than the rest (and so,... [phonological theory]...). Reproduction shows it is not necessarily so.

(*) [link](http://www.linguistics.ucla.edu/people/hayes/HungarianVH/HayesLondeHungarianWugTestData.txt)
 Depends on domain? old Hungarian vs. new foreign

- Matched-pair design with binary categorical outcome: McNemar’s $\chi^2$ test ($H_0$: same probabilities in the two conditions).
- **Bad news:** No significant difference in overall data.
  - Experiment 1: $\chi^2 = 0.2258$, df = 1, $p = .635$.
  - Experiment 2: $\chi^2 = 2.7589$, df = 1, $p = .097$.
- **Good news:** mutually neutralising significant results.
  - Personal names: backness oldH < newF.
    (Experiment 1: $p = .0011$; Experiment 2: $p = 0.024$)
  - Human made artefacts: backness oldH > newF.
    (Experiment 1: $p = .013$; Experiment 2: $p = 0.0016$)
  - Naturally occurring objects: backness oldH > newF.
    (Experiment 1: $p = .0006$; Experiment 2: $p = .058$)
  - Humans, animals, plants, weather events: n.s.
Wug-tests | Design | Results | Conclusions

Depends on domain? old Hungarian vs. new foreign

- Matched-pair design with binary categorical outcome: McNemar’s \( \chi^2 \) test (\( H_0 \): same probabilities in the two conditions).

- **Bad news:** No significant difference in overall data.
  - Experiment 1: \( \chi^2 = 0.2258, \text{df} = 1, p = .635 \).
  - Experiment 2: \( \chi^2 = 2.7589, \text{df} = 1, p = .097 \).

- **Good news:** mutually neutralising significant results.
  - Personal names: backness oldH < newF.
    (Experiment 1: \( p = .0011 \); Experiment 2: \( p = 0.024 \) )
  - Human made artefacts: backness oldH > newF.
    (Experiment 1: \( p = .013 \); Experiment 2: \( p = 0.0016 \) )
  - Naturally occurring objects: backness oldH > newF.
    (Experiment 1: \( p = .0006 \); Experiment 2: \( p = .058 \) )
  - Humans, animals, plants, weather events: n.s.
Dependence on various factors: logistic regression

Backness: \( P(\text{suffix} = [n\delta k] |...) = ? \)

E.g., backness of \( \text{hádél} \) in Experiment 1:

<table>
<thead>
<tr>
<th>Category</th>
<th>Count</th>
<th>Total</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>overall</td>
<td>1360</td>
<td>3072</td>
<td>.443</td>
</tr>
<tr>
<td>personal name</td>
<td>187</td>
<td>433</td>
<td>.432</td>
</tr>
<tr>
<td>old Hungarian personal name</td>
<td>87</td>
<td>217</td>
<td>.401</td>
</tr>
<tr>
<td>new foreign personal name</td>
<td>100</td>
<td>216</td>
<td>.463</td>
</tr>
<tr>
<td>artefacts</td>
<td>161</td>
<td>440</td>
<td>.366</td>
</tr>
<tr>
<td>old Hungarian artefacts</td>
<td>89</td>
<td>225</td>
<td>.396</td>
</tr>
<tr>
<td>new foreign artefacts</td>
<td>72</td>
<td>215</td>
<td>.335</td>
</tr>
<tr>
<td>weather event</td>
<td>217</td>
<td>445</td>
<td>.488</td>
</tr>
<tr>
<td>old Hungarian weather event</td>
<td>96</td>
<td>224</td>
<td>.429</td>
</tr>
<tr>
<td>new foreign weather event</td>
<td>121</td>
<td>221</td>
<td>.548</td>
</tr>
</tbody>
</table>
As the experiment proceeds...

nr: number of the item within questionnaire (NB: first two always fillers).

\[ S \sim ww + PS + PPS + nr + frame : exp + ww : sem \]

Coefficients for most levels of categorical variable nr are significant. (A non-significant model, \( p = .761 \), which can nonetheless be significantly improved by introducing \( PS:nr \) and \( PPS:nr \) interactions. Other models yield similar pictures.)
Further significant factors

1. The wug word.
   - The wug word’s pattern (e.g., Bé, Bii) in interaction terms, rather than the wug word itself: sometimes improves the glm model.
   - Wug word and ontological domain interaction: E.g., vuszék as an artefact (but also as a natural object) more likely to get front suffix ($p < .01$).

2. Priming: the suffix given by the subject for previous items (last two tested, both highly significant).

3. Those finishing the test: more back responses than those not finishing it. (Otherwise, unfinished questionnaires not included in statistics).

4. Sound symbolism: negative weather events more often back suffix than positive weather events (newF: $p = .046$; oldH: $p = .0005$).
Non-significant factors

Since we had the information, why not test these:

- Time elapsed since the beginning of the questionnaire (worse predictor than \( nr \) of item).
- Gender of the subject.
- Time of day.

**A note of caution:** A factor that has been *n.s.* may still prove significant in a repeated experiment (with larger sample). Still, we expect the effect to be small.

Moreover, a factor that is significant here, can be due to type I error.
Overview

1. Wug-tests
2. Design
3. Results
4. Conclusions
Conclusions

- Several native speaker’s intuition: words for old Hungarian objects more likely to get [+back] suffix than new foreign objects. ‘Folk-historical linguistics’? This intuition seems to be confirmed. Interestingly, opposite direction effect for personal names.

- For sure: backness of a wug word depends on frame! Effect is small, but highly significant when measured on a large sample.

- Exactly which (phonological, syntactic, semantic) aspects of the frame influence allomorphy, remains to be established. Likely influence of ontological category. Likely sound symbolism: increased backness if negative connotation.

- Significant priming effect detected.
Thank you for your attention!

Tamás Biró:
