The role of universal markedness in Hungarian gemination processes

LILLA MAGYAR
Massachusetts Institute of Technology
lilla@mit.edu

Introduction

Gemination in loanwords

- Singleton consonants may undergo gemination when borrow- ed into Hungarian from English, German (and occasion- ally, from French), for example: said [ʃæd] ‘shock’ ([Nádasdy 1989b], [Kertész 2006 (j)])
- Most common in word-final position in monosyllables
- Gemination depends on consonant class as well
- Voiceless obstruents are geminated more often than other con- sonants
- Nasals are geminated more often than liquids
- Voice fricatives are never geminated when there is no ortho- graphic geminate in the source word

Gemination in the native phonology

All consonants can be geminated.

<table>
<thead>
<tr>
<th>Consonant class</th>
<th>Singleton %</th>
<th>Geminate %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voiceless affricates</td>
<td>35%</td>
<td>65%</td>
</tr>
<tr>
<td>Voiceless stops</td>
<td>91%</td>
<td>9%</td>
</tr>
<tr>
<td>Voiceless fricatives</td>
<td>54%</td>
<td>46%</td>
</tr>
<tr>
<td>Voiced stops</td>
<td>67%</td>
<td>33%</td>
</tr>
<tr>
<td>Nasals</td>
<td>76%</td>
<td>24%</td>
</tr>
<tr>
<td>Liquids</td>
<td>72%</td>
<td>28%</td>
</tr>
<tr>
<td>Voiced fricatives</td>
<td>92%</td>
<td>8%</td>
</tr>
</tbody>
</table>

Comparison with wug data

Distribution of singletons and geminates is similar (with some differences):

<table>
<thead>
<tr>
<th>Consonant class</th>
<th>Singleton %</th>
<th>Geminate %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voiceless affricates</td>
<td>35%</td>
<td>65%</td>
</tr>
<tr>
<td>Voiceless stops</td>
<td>91%</td>
<td>9%</td>
</tr>
<tr>
<td>Voiceless fricatives</td>
<td>54%</td>
<td>46%</td>
</tr>
<tr>
<td>Voiced stops</td>
<td>67%</td>
<td>33%</td>
</tr>
<tr>
<td>Nasals</td>
<td>76%</td>
<td>24%</td>
</tr>
<tr>
<td>Liquids</td>
<td>72%</td>
<td>28%</td>
</tr>
<tr>
<td>Voiced fricatives</td>
<td>92%</td>
<td>8%</td>
</tr>
</tbody>
</table>

Testing Hypothesis 3

The model

- Based on the principle of Maximum Entropy
- Implemented in the UCLA Phonotactic Learner (Hayes & Wilson 2008)

Simulation 1

- Full monosyllables from the Hungarian Webcorpus used as training data
- Rhymes (short vowel and short consonant or geminate se- quences) from the wug test used as testing data

Strong positive correlation (r=0.86)

Simulation 2

- Rhymes from the corpus used as training data
- Final monosyllables of the Hungarian Webcorpus as testing data

Very weak positive correlation (r=0.25)

Simulation 3

- Rhymes from the corpus used as training data
- Final monosyllables extracted from the corpus as testing data

Very weak positive correlation (r=0.14)

Conclusions

- Native speakers’ geminate vs. singleton preferences in the wug test are mostly similar to what we find in loanwords, the native Hungarian lexicon, and also, across languages.
- The distribution of geminates and singletons in the corpus also line up with universal geminate markedness hierarchies.
- These are not complete rankings, but mostly implicational hi- erarchies.
- Voiceless obstruents are the most preferred geminates, less sonorous sonorant geminates are more common than least sonorous ones (except in the wug test), voice fricatives are very rare.
- Speakers do not generalise fine-grained distinctions.
- They seem to generalise based on broader categories, like consonant classes.
- It indicates that there may be bias for geminate markedness.

References


Acknowledgements

Thank you for comments and discussion: Adam Albright, Donca Steriade, Michael Keatonovics, Edward Fleming, Miklós Túrkmaczy, Katalin Mály, Csaba Oravecz, and audiences at the 22nd Manchester Phonology Meeting and Phonology Cir- cle at MIT.