Introduction

Factors affecting the well-formedness of English CVC syllables:

- **The OCP-place**: a gradient prohibition against homorganic consonants in the onset and the coda of the CVC syllables.
  
  E.g. gap vs. gap

  **HYPOTHESIS**: Syllables with C1 and C2 of the same place of articulation are underrepresented

- **The prominence alignment** between syllable stress, vowel height, and consonant place.

  **HYPOTHESIS**: Syllables that violate prominence alignment are underrepresented

Material:

- CMU pronunciation dictionary and CELEX lemma lexicon
- Only CVC syllables with primary or no stress
- Syllables - stressed and unstressed
- Consonants - coronals, dorsals, and labials
- Vowels - high (± reduced) and low (± mid)

Effect size evaluation:

- Observed frequency/Expected frequency ratio (O/E ratio):
  
  O/E ratio > 1.00 overrepresentation
  O/E ratio < 1.00 underrepresentation

Effect sizes calculated:

<table>
<thead>
<tr>
<th>Material</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>O/E ratio</td>
<td>25,888 CVC syllables from CELEX</td>
</tr>
<tr>
<td></td>
<td>83,794 CVC syllables from CMU</td>
</tr>
</tbody>
</table>

Results

1. Syllables that violate OCP-place are underrepresented:

   Onset-coda co-occurrences (OE values):

   - Labial and coronal consonants in stressed syllables
   - Labial and coronal consonants in unstressed syllables

   CMU

   - Labial: 0.46, 0.77, 1.23
   - Dorsal: 0.78, 0.42, 1.25
   - Coronal: 1.29, 1.20, 0.86

   CELEX

   - Labial: 0.42, 0.77, 1.23
   - Dorsal: 0.78, 0.42, 1.25
   - Coronal: 1.29, 1.20, 0.86

2. Syllables that violate consonant-stress alignment are underrepresented:

   - Labial and coronal consonants in unstressed syllables
   - Labial and coronal consonants in stressed syllables

   CMU

   - Labial: 0.27, 0.87, 1.68
   - Dorsal: 1.30, 0.79, 0.96
   - Coronal: 0.80, 1.14, 0.95

   CELEX

   - Labial: 0.28, 0.87, 1.68
   - Dorsal: 1.30, 0.79, 0.96
   - Coronal: 0.80, 1.14, 0.95

3. Syllables violating consonant-vowel alignment are underrepresented:

   - Low vowels with coronal consonants
   - High vowels with labial or dorsal consonants

   CMU

   - Labial: 0.68, 1.03
   - Dorsal: 0.96, 1.03

   CELEX

   - Labial: 0.65, 1.03
   - Dorsal: 0.96, 1.03

4. Syllables that violate vowel-stress assignment are underrepresented:

   - Low vowels in unstressed syllables
   - High vowels in stressed syllables

   CMU

   - Labial: 0.51, 1.63
   - Dorsal: 0.97, 1.68

   CELEX

   - Labial: 0.51, 1.63
   - Dorsal: 0.97, 1.68

Regession

Factors | Coefficient | Std. Error | t | p
-------|-------------|------------|---|---
Expected | 0.857 | 0.077 | 11.20 | <0.001
OCP | 0.991 | 0.204 | 4.709 | <0.001
Stressed/Unstressed | 1.070 | 0.209 | 7.723 | <0.001
Stressed/Unstressed Unstressed | -1.564 | 0.257 | -5.767 | <0.001
Lab&Dor onset/Unstressed | 0.820 | 0.251 | 3.288 | <0.001
Lab&Dor onset/Unstressed Unstressed | 0.733 | 0.264 | 2.641 | <0.001

R = 0.945 (F(6, 35) = 18.48; p < 0.001)

In CMU significant factors:

- Vowel-stress assignment
- OCP

No labial/dorsal in unstressed syllables

In CELEX significant factors:

- Vowel-stress alignment
- OCP

No labial/dorsal with high vowels

Conclusions

- The gradient OCP-place effect is active in all CVC syllables of English (not only in monosyllabic words)
- Pronomence alignment in CVC syllables:
  - Best stressed syllable – with low or mid vowels
  - Best unstressed syllable – with high or reduced vowels and coronal consonants
- Positional neutralization and augmentation for vowels
- Only positional neutralization for consonants

OT Analysis

Constraints (significant regression factors):

- OCP No homorganic consonants in onset and coda of the same syllable
- *a/ Low vowels are prohibited in unstressed syllables
- *H High vowels are prohibited in stressed syllables
- *aH No labial/dorsal onsets in unstressed syllables
- *aH No labial/dorsal onsets followed by high vowels
- *H No labial/dorsal onsets preceded by high vowels

Fath

- Markedness constraints ranked according to regression coefficients
- Markedness constraints unranked with respect to Fath

Multiple grammars

Measure of the model fit

Data | Precision | Recall | Markedness into frequency
-------|------------|-------|-----------------------
CMU | 0.75 | 0.88 | CELEX | 0.74 | 0.87

References:


