Gemination and Degemination in English Affixation
Lexical Strata, Variability, and Phonetic Evidence

Sonia Ben Hedia & Ingo Plag

17th International Morphology Meeting, 18–21 February 2016
Morpho-phonology as we know it

• Morpho-phonological alternations are categorical but may have lexical exceptions.

• The formal level of representation of morphemes is phonological in nature.

• Post-lexical phonology and phonetics have no access to lexical information.
Problems

• Morpho-phonological alternations are more variable than previously assumed, and governed by unexpected factors
  o Stress shift (Bauer, Lieber & Plag 2013 on -able)
  o Stress preservation (Collie 2008, relative frequency as a proxy for morphological segmentability)

• Subphonemic detail may reflect morphological information
  o Free vs. bound stems (Kemps et al. 2005, Blazej & Cohen-Goldberg 2015)
  o Different S morphemes (Plag, Homann & Kunter 2015)

• Serious implications for theories of morpho-phonology (Plag 2014)
  o Exception vs. rule
  o Lexical vs. post-lexical phonology (in linguistic theory, and in speech production models)
(De-)Gemination in English

• Sequence of two identical consonants across a morphological boundary

  un-
  un-natural

  in-
  in-numerous, im-material, il-logical, ir-resistable

  -ly
  sole-ly, technical-ly

• What happens at the segmental level? Gemination or deletion?

• Phonetic correlates
  o Gemination: Longer duration than a singleton
  o Degemination: Same duration as a singleton
Assumptions about gemination in English

• *un-* geminates

• *in-* degeminates

  (e.g. Cruttenden & Gimson 2014, Cohen-Goldberg 2014, Kiparsky 1982, Mohanan 1986)

• -*ly*
  
  o ... is variable (*stalely* vs. *fully*, Bauer 2001, Bauer, Lieber & Plag 2013)
  
  o ... geminates (Lexical Phonology: level 2 affix)

• General theoretical assumptions
  
  o Degemination is affix- or stratum-dependent
  
  o Degemination is a categorical morpho-phonological process with some lexical exceptions
## Predictions

<table>
<thead>
<tr>
<th></th>
<th>Level 1</th>
<th>Level 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Morphological</strong></td>
<td>in + numerous</td>
<td>un + natural sole + ly</td>
</tr>
<tr>
<td><strong>Process</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Phonological</strong></td>
<td>i/n/umerous</td>
<td>u/nn/atural so/l/y</td>
</tr>
<tr>
<td><strong>Process</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Phonetic Outcome</strong></td>
<td>i[n]umerous</td>
<td>u[n:]atural so[l:]y</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Degemination**

**Gemination**
Empirical evidence?

• Only two studies empirically investigated *in*- and *un*- in English

• Type-dependent (Oh and Redford 2013) and speaker-dependent (Kaye 2005) variation in degemination with *in*-prefixed words

  *immigrational* degemination
  *immemorial* gemination
  *immature* variable (by speaker)

• Problems: Only experimental data, only very small set of types

• No empirical study of *-ly*
This study

• What determines degemination at morphological boundaries? Three affixes: *un-*, *in-*, *-ly*

• Diagnostics: Acoustic duration

• Data: Natural conversational speech
Methodology

• Sample of un-, in- and -ly-affixed words with a double or a single consonant at the morphological boundary
• Switchboard Corpus (Godfrey & Holliman 1997)
• For the prefix in- the allomorph /im/ was investigated
• Manual segmentation and acoustic measurements in Praat (Boersma & Weenink 2014)
Methodology

• Statistical Analysis: Multiple regression with **duration** as dependent variable and **number of consonants** (single vs. double) as predictor

• Coding of pertinent covariates:
  • Preceding Segment Duration
  • Preceding Segment
  • Following segment
  • Speech Rate
  • Stress
  • Syllabic
  • Word Form Frequency
  • Relative Frequency
  • Affix
  • Semantic Transparency
Overview of the data

<table>
<thead>
<tr>
<th></th>
<th>Double Consonant</th>
<th>Single Consonant</th>
<th>Total per affix</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>un-</em></td>
<td>23</td>
<td>135</td>
<td>158</td>
</tr>
<tr>
<td><em>in-</em></td>
<td>89</td>
<td>67</td>
<td>156</td>
</tr>
<tr>
<td>-<em>ly</em></td>
<td>76</td>
<td>80</td>
<td>156</td>
</tr>
</tbody>
</table>
Results: Overview
Results 1: *un*-geminates
Results 1: *un-* geminates
Results 2: *in*-geminates

- **Segments across boundary**: The duration in seconds for single and double segments is compared.
- **Local speech rate**: The duration in seconds against the local speech rate is shown.
- **Stress pattern**: Comparison of in-σ and in-σ in duration.
- **Type of affix**: Comparison of inLoc and inNeg in duration.
Results 2: *in*-geminates

- **Segments across boundary:**
  - Single: Mean duration = 0.07 seconds
  - Double: Mean duration = 0.09 seconds

- **Local speech rate:**
  - Decrease in duration as speech rate increases

- **Stress pattern:**
  - *in-'σ*
    - Duration = 0.10 seconds
  - *in-σ*
    - Duration = 0.09 seconds

- **Type of affix:**
  - *inLoc*
    - Duration = 0.09 seconds
  - *inNeg*
    - Duration = 0.08 seconds
Results 2: *in-* geminates

- **Segments across boundary**: Duration in seconds for single vs. double segments.
- **Local speech rate**: Duration in seconds as a function of local speech rate.
- **Stress pattern**: Duration in seconds for *in-*σ vs. *in*-σ.
- **Type of affix**: Duration in seconds for *inLoc* vs. *inNeg*. 

![Graphs showing duration in seconds for different conditions and variables.](Image)
Results 2: *in*-geminates

- **Segments across boundary:**
  - Single: Duration in seconds range from 0.06 to 0.10.
  - Double: Duration in seconds range from 0.08 to 0.12.

- **Local speech rate:**
  - Duration in seconds decreases as speech rate increases.

- **Stress pattern:**
  - *in*-σ: Duration in seconds range from 0.07 to 0.09.
  - *in*-σ: Duration in seconds range from 0.08 to 0.10.

- **Type of affix:**
  - inLoc: Duration in seconds range from 0.08 to 0.10.
  - inNeg: Duration in seconds range from 0.07 to 0.09.
Results 3: -/y does not geminate
Results 3: -/y does not geminate

Additional covariate: Syllabic (ment[l]y vs. ment[əl]y, odd[l]y)
Results 3: -/y does not geminate
Results 3: -/y does not geminate
Summary

• *un-* geminates: no surprise

• *in-* geminates: unexpected result
  o Effect of AFFIX: homophonous locative and negative *in-* prefixes are acoustically different

• *-ly* degeminates: unexpected result
  o effect of RELATIVE FREQUENCY: morphological segmentability influences phonetic implementation
Implications

• Empirical facts contradict received wisdom for *in*- and *-ly*

• Lexical Phonology makes wrong empirical predictions

• Morphological information is directly reflected in the speech signal
  o *in* -: Homophonous affixes exhibit different acoustic properties (cf. Plag, Homann & Kunter 2015 on S)

• Challenges models of lexical phonology and models of speech production that state that post-lexical phonology has no access to morphological information (e.g. Lexical Phonology, Levelt, Roelofs & Meyer 1999)
Thank you very much for your attention!

We gratefully acknowledge financial support by
Deutsche Forschungsgemeinschaft (Grant PL151/8-1 ‘Morphophonetic Variation in English‘)
References


un-model

# Call:
# lm(formula = bc ~ TransitionType + LocSpeech, data = unComplex2)
#
# Residuals:
#    Min     1Q Median     3Q    Max
# -0.081237 -0.027028 -0.000937  0.025328  0.096961
#
# Coefficients:
# Estimate Std. Error t value Pr(>|t|)
# (Intercept)    0.581989   0.014676  39.655   < 2e-16 ***
# TransitionTypesingle-C -0.049389   0.009505  -5.196  6.59e-07 ***
# TransitionTypesingle-V -0.099885   0.009641 -10.360   < 2e-16 ***
# LocSpeech      -0.007646   0.001063  -7.196  2.83e-11 ***
# ---
# Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ‘ 1
#
# Residual standard error: 0.03788 on 149 degrees of freedom
# Multiple R-squared:  0.6011,  Adjusted R-squared:  0.5931
# F-statistic: 74.84 on 3 and 149 DF,  p-value: < 2.2e-16
# lm(formula = bc ~ NoCons + LocSpeech + StressPattern + Affix,
#    data = imComplex4)
#
# Residuals:
#   Min        1Q    Median        3Q       Max
# -0.081827 -0.023172 -0.002205  0.023101  0.083318
#
# Coefficients:
#   Estimate Std. Error t value Pr(>|t|)
# (Intercept)             0.2856713  0.0112978  25.286  < 2e-16 ***
#   NoConsdouble 0.0442330  0.0064822   6.824 2.08e-10 ***
#   LocSpeech -0.0032078  0.0007413  -4.327 2.76e-05 ***
#   StressPatternstr-unstr -0.0344743  0.0071455  -4.825 3.44e-06 ***
#   AffixinNeg 0.0196406  0.0069752   2.816  0.00553 **
#   ---
# Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
#
# Residual standard error: 0.0325 on 149 degrees of freedom
# Multiple R-squared:  0.5392,  Adjusted R-squared:  0.5268
# F-statistic: 43.58 on 4 and 149 DF,  p-value: < 2.2e-16
-ly-model

# lm(formula = AbsDurCon ~ NoCons + logRelFreq + PrecSegVC + LocSpeech +
#     Syllabic, data = lyComplex2)
#
# Residuals:
#  Min        1Q    Median        3Q       Max
# -0.046194 -0.013208 -0.001831  0.011909  0.045429
#
# Coefficients:
#   Estimate Std. Error t value Pr(>|t|)
# (Intercept)           0.0799558  0.0086899   9.201 3.41e-16 ***
#   NoConsdouble -0.0074318  0.0056623  -1.313 0.191410
# logRelFreq -0.0014775  0.0006016  -2.456 0.015219 *
#   PrecSegVCV 0.0168499  0.0047635   3.537 0.000542 ***
#   LocSpeech -0.0022602  0.0004393  -5.145 8.49e-07 ***
#   Syllabicnon-syllabic -0.0138244  0.0068922  -2.006 0.046726 *
#     ---
#    Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
#
# Residual standard error: 0.01876 on 146 degrees of freedom
# Multiple R-squared:  0.2435, Adjusted R-squared:  0.2176
# F-statistic: 9.398 on 5 and 146 DF,  p-value: 8.768e-08
# Types

<table>
<thead>
<tr>
<th></th>
<th>Doubles</th>
<th>Singles</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>un-</em></td>
<td>6</td>
<td>93</td>
</tr>
<tr>
<td><em>in-</em></td>
<td>17</td>
<td>65</td>
</tr>
<tr>
<td><em>-ly</em></td>
<td>76</td>
<td>72</td>
</tr>
</tbody>
</table>