Gemination and Degemination in English Affixation

A Phonetic Investigation of Lexical Strata, Semantics and Decomposability

Sonia Ben Hedia

Collaborators
Ingo Plag (U Düsseldorf), Gero Kunter (U Düsseldorf)

Funding
Deutsche Forschungsgemeinschaft
• Grant PL151/8-1 ‘Morpho-phonetic Variation in English’
Consonant gemination

**Geminate** [Lat. Geminate ‘doubled’] (also double consonant, long consonant):
A consonant that is distinguished by another exclusively by its longer period of articulation. The difference between simple and long consonants is phonologically relevant in some languages, e.g. Italian, but not in others, e.g. English. (Bussmann et al. 1996)

Italian: 

- *fato* ‘destiny’
- *fatto* ‘fact’

- *papa* ‘father’
- *pappa* ‘mush’
Gemination in English

• Sequence of two identical segments only at morphological boundaries, e.g.
  
  un#natural
  fun # name
  some # more

• ‘fake geminates’, ‘geminates’, ‘morphological geminates’

• What happens on the phonetic level?
  • Longer duration than a singleton ( = ‘gemination’)?
  • Same duration as a singleton ( = ‘degemination’)?
Gemination in English

• What is the pattern of gemination in English?

• Is there variation?

  • If so, which factors influence the duration of consonant length on morphological boundaries?
Why look at morphological geminates?

• Gemination in English as a phenomenon at the interface of morphology, phonology and phonetics
## The morphological level

<table>
<thead>
<tr>
<th>Affixation</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Prefixation</strong></td>
<td><strong>Suffixation</strong></td>
</tr>
<tr>
<td><em>in</em> as in <em>in</em>n<em>umerous</em></td>
<td><em>ly</em> as in <em>real</em>ly</td>
</tr>
<tr>
<td><em>un</em> as in <em>un</em>natural</td>
<td></td>
</tr>
<tr>
<td><em>dis</em> as in <em>dis</em>solve</td>
<td></td>
</tr>
<tr>
<td><strong>Compounds</strong></td>
<td></td>
</tr>
<tr>
<td><em>fun</em># name</td>
<td></td>
</tr>
<tr>
<td><strong>Word Boundaries</strong></td>
<td></td>
</tr>
<tr>
<td><em>He wanted some <em>#</em> more.</em></td>
<td></td>
</tr>
</tbody>
</table>

**Morphology:**

Gemination only occurs at morphological boundaries.
The phonological level

Two identical underlying segments

Phonology:
Sequence of two identical underlying segments

u /n#n/ atural
fu /n # n/ame
so /m # m/ ore
The phonetic level

Gemination

*u[n:]atural* or *u[n]atural*

*fu[n:]ame* or *fu[n]ame*

*so[m:]ore* or *so[m]ore*

Degemination

*Phonetics:*

Duration as a phonetic correlate for gemination/degemination
Why look at morphological geminates?

- Gemination in English as a phenomenon at the interface of morphology, phonology and phonetics
  - **Morphology**: Gemination only occurs at morphological boundaries
  - **Phonology**: Two identical underlying consonants
  - **Phonetics**: Duration as a phonetic correlate for gemination/degemination

- Investigating gemination in English can shed light on theories of the morpho-phonological and the morpho-phonetic interface, e.g. Lexical Phonology
This study

• What determines (de-)gemination at affixational boundaries?

• Gemination in *un-, in-, -ly*
  – Do they geminate?
    • Is the [n] in *unnatural* longer than the [n] in *uneven*?
  – Which factors determine consonant length?
    • Frequencies, speech rate...

• Which implications for theories of morpho-phonology?
  – Lexical Phonology
  – Morphological Segmentability Hypothesis (Hay 2003)
Lexical Phonology

- Affixes belong to different Lexical Strata

- Level 1 (like *in-* ) affixes display a weak morphological boundary and a great degree of integration with the base

- Level 2 affixes (like *un-* and *-ly*) display a strong morphological boundary and a lesser degree of integration with the base

- Morphological information is only reflected on phonemic level, not in phonetic detail
### Lexical Phonology

<table>
<thead>
<tr>
<th></th>
<th>Level 1</th>
<th>Level 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morphological Process</td>
<td>in + possible</td>
<td>un + predictable</td>
</tr>
<tr>
<td>Phonological Process</td>
<td></td>
<td><strong>Weak boundary</strong></td>
</tr>
<tr>
<td>Phonetic Outcome</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Lexical Phonology

<table>
<thead>
<tr>
<th></th>
<th>Level 1</th>
<th>Level 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morphological Process</td>
<td>in + possible</td>
<td>un + predictable</td>
</tr>
<tr>
<td>Phonological Process</td>
<td>i/m/possible</td>
<td></td>
</tr>
<tr>
<td>Phonetic Outcome</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Lexical Phonology

<table>
<thead>
<tr>
<th></th>
<th>Level 1</th>
<th>Level 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morphological Process</td>
<td>in + possible</td>
<td>un + predictable</td>
</tr>
<tr>
<td>Phonological Process</td>
<td>i/m/possible</td>
<td></td>
</tr>
<tr>
<td>Phonetic Outcome</td>
<td>i[m]possible</td>
<td></td>
</tr>
</tbody>
</table>

**Strong boundary**

**Assimilation**
## Lexical Phonology

<table>
<thead>
<tr>
<th></th>
<th>Level 1</th>
<th>Level 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morphological Process</td>
<td>in + possible</td>
<td>un + predictable</td>
</tr>
<tr>
<td>Phonological Process</td>
<td>i/m/possible</td>
<td>u/n/predictable</td>
</tr>
<tr>
<td>Phonetic Outcome</td>
<td>i[m]possible</td>
<td></td>
</tr>
</tbody>
</table>

**Assimilation**

**No assimilation**
## Lexical Phonology

<table>
<thead>
<tr>
<th></th>
<th>Level 1</th>
<th>Level 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morphological Process</td>
<td>in + possible</td>
<td>un + predictable</td>
</tr>
<tr>
<td>Phonological Process</td>
<td>i/m/possible</td>
<td>u/n/predictable</td>
</tr>
<tr>
<td>Phonetic Outcome</td>
<td>i[m]possible</td>
<td>u[n]predictable</td>
</tr>
</tbody>
</table>

- **Assimilation**
- **No assimilation**
# Lexical Phonology

<table>
<thead>
<tr>
<th></th>
<th>Level 1</th>
<th>Level 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morphological</td>
<td>in + possible</td>
<td>un + predictable</td>
</tr>
<tr>
<td>Process</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phonological</td>
<td>i/m/possible</td>
<td>u/n/predictable</td>
</tr>
<tr>
<td>Process</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phonetic Outcome</td>
<td>i[m]possible</td>
<td>u[n]predictable</td>
</tr>
<tr>
<td></td>
<td></td>
<td>**all morphological</td>
</tr>
<tr>
<td></td>
<td></td>
<td>information is erased</td>
</tr>
</tbody>
</table>

**Assimilation**

**No assimilation**

17
# Lexical Phonology and Gemination

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

**strong boundary**
# Lexical Phonology and Gemination

<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</td>
</tr>
<tr>
<td><strong>Process</strong></td>
<td></td>
<td>sole + ly</td>
</tr>
<tr>
<td><strong>Phonological</strong></td>
<td>i/n/umerous</td>
<td></td>
</tr>
<tr>
<td><strong>Process</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Phonetic Outcome</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Deletion of one segment**
## Lexical Phonology and Gemination

<table>
<thead>
<tr>
<th></th>
<th>Level 1</th>
<th>Level 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morphological Process</td>
<td>in + numerous</td>
<td>un + natural sole + ly</td>
</tr>
<tr>
<td>Phonological Process</td>
<td>i/n/umerous</td>
<td></td>
</tr>
<tr>
<td>Phonetic Outcome</td>
<td>i[n]umerous</td>
<td></td>
</tr>
</tbody>
</table>

**Degemination**
# Lexical Phonology and Gemination

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

**Degemination**  
**No deletion**
## Lexical Phonology and Gemination

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

**Degemination**  
**Gemination**
Lexical Phonology: Predictions for gemination and degemination

- Gemination depends on the affix involved!
  - *Un-* as a level 2 affix geminates
  - *In-* as a level 1 affix degeminates
  - *-ly* as a level 2 affix geminates
Assumptions about gemination in English

Gimson’s Pronunciation of English (2014):

“In general such prefixes result in a doubled consonant when the prefix-final and the stem-initial consonants are identical, e.g. *unnecessary* is pronounced with a double length [n:]. (This rule does not apply to *in-* and its variants, so for example *illogical* is pronounced with only a single /l/).” (p. 248)

Cohen-Goldberg (2013):

“Similarly in English, although geminates are banned from monomorphemic words (*spaghe*[tt]i) and words containing less productive affixes (e.g. *in-*: *i*[n]umerable), they are allowed in words containing more productive affixes and compounds (e.g. *un-*: *u*[nn]ecessary; *boo*[kk]eep*er).” (p.1055 f.)
Morphological Segmentability Hypothesis

- Gradient decomposability of affixed forms is reflected in phonetic detail
- The more decomposable a word is, the less reduction
- Decomposability measured in relative frequency (whole word frequency : base frequency)
Relative Frequency

- ratio : whole word frequency / base frequency
- The more frequent the whole word and the less frequent the base, the less decomposable (high relative frequency)

<table>
<thead>
<tr>
<th>item</th>
<th>Word Frequency</th>
<th>Base Frequency</th>
<th>Relative frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>uninhabitable</td>
<td>224</td>
<td>39</td>
<td>224 : 39 = 5.74</td>
</tr>
</tbody>
</table>
Relative Frequency

- ratio: whole word frequency / base frequency
- The more frequent the whole word and the less frequent the base, the less decomposable (high relative frequency)
- The less frequent the whole word and the more frequent the base, the more decomposable (low relative frequency)

<table>
<thead>
<tr>
<th>item</th>
<th>Word Frequency</th>
<th>Base Frequency</th>
<th>Relative frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>uninhabitable</td>
<td>224</td>
<td>39</td>
<td>224: 39 = 5.74</td>
</tr>
</tbody>
</table>
Relative Frequency

• ratio : whole word frequency / base frequency
• The more frequent the whole word and the less frequent the base, the less decomposable (high relative frequency)
• The less frequent the whole word and the more frequent the base, the more decomposable (low relative frequency)

<table>
<thead>
<tr>
<th>item</th>
<th>Word Frequency</th>
<th>Base Frequency</th>
<th>Relative frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>uninhabitable</td>
<td>224</td>
<td>39</td>
<td>224: 39 = 5.74</td>
</tr>
<tr>
<td>immoral</td>
<td>390</td>
<td>4994</td>
<td>390: 4994 = 0.078</td>
</tr>
</tbody>
</table>
Morphological Segmentability Hypothesis

• Gradient decomposability of affixed forms is reflected in phonetic detail

• The more decomposable a word is, the less reduction

• Decomposability measured in relative frequency (whole word frequency : base frequency)

• Less reduction in the prefix *un*- when part of more decomposable word (Hay 07)
  • *un*- in more decomposable words longer than in less decomposable words
Morphological Segmentability Hypothesis and gemination

Gemination of *in*, *un*-, and *ly*-affixed word does not depend on the ________ involved. It depends on the individual word’s __________ which can be measured in a word’s __________________. The more decomposable a word is, the _________ is the duration of the nasal at the morphological boundary. The less decomposable a word is, the _________ the nasal.
Morphological Segmentability Hypothesis and gemination

Gemination of *in-*-, *un-* and *ly-* affixed word does not depend on the *affix* involved. It depends on the individual word’s *decomposability* which can be measured in a word’s *relative frequency*. The more decomposable a word is, the *longer* is the duration of the nasal at the morphological boundary. The less decomposable a word is, the *shorter* the nasal.
Summary: Predictions about gemination in English affixation

Lexical Phonology

- Gemination and degemination is categorial.
- Gemination depends on the affix:
  - *un-* and *–ly* geminate
  - *in-* degeminate
- Morphological information is only reflected on phonemic level, not in phonetic detail.

Morphological Segmentability Hypothesis

- Gemination is gradient.
- Duration of the boundary adjacent nasal depends on a word’s decomposability measured in relative frequency.
- Morphological information is reflected in phonetic detail.
Empirical evidence

• Only two studies empirically investigated *in-* and *un-* in English
• No study on -*ly*

• Kaye (2005): experiment with very few types, spoken in isolation

  *in-*  immature – mature  
  gemination (but somewhat variable by speaker)

  *un-*  unaimed – unnamed – named
  gemination
Empirical evidence

Oh and Redford (2013)

• Experimental study with four types for each prefix
  immovable, immoral, immemorial, immeasured
  unnoticed, unnamed, unnerve, unnail

• Comparison of durations with (assumed) phonological singletons with orthographic doubles (e.g. *immunity, immigational*)

Results

• both *im*- and *un*- geminate
  but not all *im*- prefixed words geminate

Problems

• A priori classification of stimuli as geminates or non-geminates
• Small set of types
• Stimuli only spoken in carrier sentence ‘I said ___ again’, asking for normal vs. careful speaking style
Empirical evidence and morphophonological theories

How does the empirical evidence (Kaye 05, Oh and Redford 2012) relate to the predictions made by Lexical Phonology and the Morphological Segmentability Hypothesis?

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><em>in</em>-</td>
<td><em>un</em>-</td>
<td><em>-ly</em></td>
</tr>
<tr>
<td>Kaye 05: speaker-dependent gemination</td>
<td>Kaye 05: gemination</td>
<td>No evidence</td>
</tr>
</tbody>
</table>

- Lexical Phonology: results for *un*- support theory BUT results for *in*- do not
- Morphological Segmentability Hypothesis: neither supported nor falsified since decomposability not tested
Back to our study

- Natural conversational speech
- Include –ly
- Test affix-dependent gemination
  - Do un-prefixed words geminate?
    - Is the [n] in *unnatural* longer than the [n] in *uneasy*?
  - Do in-prefixed words geminate?
    - Is the [n] in *innumerous* longer than the [n] in *inefficient*?
  - Do –ly-suffixed words geminate?
    - Is the [l] in *really* longer than the [n] in *clearly*?
- **AND** test the influence of decomposability on gemination
  - Is the [n]/[l] in more decomposable words longer than the one in less decomposable words
Methodology

• Switchboard Corpus (Godfrey & Holliman 1997)
  • 2430 two sided phone conversations among North American speakers, 240 hours of speech.
  • 3 million word tokens
• Sample of un-, in- and -ly-affixed words with a double or a single (orthographic) consonant at the morphological boundary
• ‘affixed‘: The base must be attested outside the derivative with a similar meaning (unfair, implicit – explicit, really)
• For each affix we sampled up to 160 words per category
• only one token of a given type by a single speaker
• For the prefix in- only the allomorph /im/ was investigated
• Manual segmentation and acoustic measurements in praat
Methodology

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

• Coding of pertinent covariates:
  • Relative Frequency
  • Semantic Transparency
  • Affix
  • Preceding Segment Duration
  • Preceding Segment
  • Following segment
  • Speech Rate
  • Position in utterance
  • Prosodic Structure
  • Syllabicity
  • Word Form Frequency
Methodology

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

- Coding of pertinent covariates:
  - Relative Frequency
  - Semantic Transparency
  - Affix
  - Preceding Segment Duration
  - Preceding Segment
  - Following segment
  - Speech Rate
  - Position in utterance
  - Prosodic Structure
  - Syllabicity
  - Word Form Frequency

Measure of gradient morphological complexity

The more frequent the derivative vis-à-vis the base, the less complex the word

- *happyness* - happy
  - *discernment* - discernment

- *government* - govern
  - *insane* - sane
Methodology

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

• Coding of pertinent covariates:
  • Relative Frequency
  • Semantic Transparency
  • Affix
  • Preceding Segment Duration
  • Preceding Segment
  • Following segment
  • Speech Rate
  • Position in utterance
  • Prosodic Structure
  • Syllabicity
  • Word Form Frequency

  **transparent:**
  Affix + Base = Derivative
  im + possible = impossible
  NEG + ‘possible’ = ‘not possible’

  **opaque:**
  im + mediately ≠ ‘at once’
Methodology

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

• Coding of pertinent covariates:
  • Relative Frequency
  • Semantic Transparency
  • Affix
  • Preceding Segment Duration
  • Preceding Segment
  • Following segment
  • Speech Rate
  • Position in utterance
  • Prosodic Structure
  • Syllabicity
  • Word Form Frequency
Methodology

• Statistical Analysis: Multiple regression with \textit{duration} as dependent variable and \textit{number of consonants} (single vs. double) as crucial predictor

• Coding of pertinent covariates:
  • Relative Frequency
  • Semantic Transparency
  • Affix
  • Preceding Segment Duration
  • Preceding Segment
  • Following segment
  • Speech Rate
  • Position in utterance
  • Prosodic Structure
  • Syllabicity
  • Word Form Frequency

Gemination may also affect the vowel preceding the geminated segment  
(e.g. Ridouane 2010, Miller 1987, Oh and Redford 2011)
Methodology

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

• Coding of pertinent covariates:
  • Relative Frequency
  • Semantic Transparency
  • Affix
  • Preceding Segment Duration
  • Preceding Segment
  • Following segment
  • Speech Rate
  • Position in utterance
  • Prosodic Structure
  • Syllabicity
  • Word Form Frequency

Coarticulation effects
Methodology

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

• Coding of pertinent covariates:
  - Relative Frequency
  - Semantic Transparency
  - Affix
  - Preceding Segment Duration
  - Preceding Segment
  - **Following segment**
  - Speech Rate
  - Position in utterance
  - Prosodic Structure
  - Syllabicity
  - Word Form Frequency

Coarticulation effects
Methodology

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

• Coding of pertinent covariates:
  • Relative Frequency
  • Semantic Transparency
  • Affix
  • Preceding Segment Duration
  • Preceding Segment
  • Following segment
  • **Speech Rate**
  • Position in utterance
  • Prosodic Structure
  • Syllabicity
  • Word Form Frequency

Speech rate directly influences the duration of a given segment.

\[
\text{Number of segments} \quad \text{word duration}
\]
Methodology

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

- Coding of pertinent covariates:
  - Relative Frequency
  - Semantic Transparency
  - Affix
  - Preceding Segment Duration
  - Preceding Segment
  - Following segment
  - Speech Rate
  - **Position in utterance**
  - Prosodic Structure
  - Syllabicity
  - Word Form Frequency

Final lengthening effect:
mid, end, before pause
Methodology

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

- Coding of pertinent covariates:
  - Relative Frequency
  - Semantic Transparency
  - Affix
  - Preceding Segment Duration
  - Preceding Segment
  - Following segment
  - Speech Rate
  - Position in utterance
  - Prosodic Structure
  - Syllabicity
  - Word Form Frequency

- **adjacent / non-adjacent to a stress syllable**
Methodology

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

• Coding of pertinent covariates:
  - Relative Frequency
  - Semantic Transparency
  - Affix
  - Preceding Segment Duration
  - Preceding Segment
  - Following segment
  - Speech Rate
  - Position in utterance
  - Prosodic Structure
  - Syllabic
  - Word Form Frequency

for **–ly**: syllabic /l/ should be longer

ment[l]y  ment[ə]l y  odd[l]y
Methodology

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

- Coding of pertinent covariates:
  - Relative Frequency
  - Semantic Transparency
  - Affix
  - Preceding Segment Duration
  - Preceding Segment
  - Following segment
  - Speech Rate
  - Position in utterance
  - Prosodic Structure
  - Syllabicity
  - Word Form Frequency

More frequent words are produced faster
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>un-</td>
<td>22</td>
<td>136</td>
<td>158</td>
</tr>
<tr>
<td>in-</td>
<td>89</td>
<td>67</td>
<td>156</td>
</tr>
<tr>
<td>-ly</td>
<td>81</td>
<td>75</td>
<td>156</td>
</tr>
</tbody>
</table>
Results: Overview

- **un-**
  - Duration: Single vs. Double

- **in-**
  - Duration: Single vs. Double

- **-ly**
  - Duration: Single vs. Double
Results 1: *un*-geminates

*un*-geminates

$p<0.05^{***}$

duration in seconds

Single | Double | Single

segments across boundary
Results 1: *un-*

No effect of relative frequency on nasal duration for *un-*

$p=0.379$
Results 2: *in-*

*in-* geminates

\[ p < 0.05^{***} \]

- Single
- Double

segments across boundary

duration in seconds
Results 2: *in-*

No effect of relative frequency on nasal duration for *in-*

\[ p = 0.707 \]
Results 2: *in*-.

Effect of affix on nasal duration for *in*-

The [n] in negative in (e.g. *impolite*) is longer than the [n] in locative in (e.g. *implant*)

\[ p < 0.05^{***} \]
Results 3: -/ly

-ly degeminates

p = 0.19141
Results 3: -/ly

Effect of relative frequency on consonant duration for -/ly

The higher the relative frequency (less decomposable), the shorter the [l]

$p<0.05^*$
Summary

- **un-** geminates
  - No effect of relative frequency

- **in-** geminates
  - No effect of relative frequency
  - Effect of AFFIX: homophonous locative and negative *in-* prefixes are acoustically different

- **-ly** degeminates
  - Effect of RELATIVE FREQUENCY: morphological segmentability influences phonetic implementation
Our results and morpho-phonological theories

How do the results relate to the predictions made by Lexical Phonology and the Morphological Segmentability Hypothesis?
## Our results and morpho-phonological theories

<table>
<thead>
<tr>
<th>$un$-</th>
<th>$in$-</th>
<th>$-ly$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-geminates</td>
<td>-geminates</td>
<td>-degeminates</td>
</tr>
<tr>
<td>- no effect of decomposability (relative frequency)</td>
<td>- no effect of decomposability (semantic transparency and relative frequency)</td>
<td>- effect of decomposability (relative frequency)</td>
</tr>
<tr>
<td></td>
<td>- Nasal in negative $in$- is longer than in locative $in$-</td>
<td></td>
</tr>
</tbody>
</table>
Our results and morpho-phonological theories

<table>
<thead>
<tr>
<th></th>
<th>un-</th>
<th>in-</th>
<th>-ly</th>
</tr>
</thead>
<tbody>
<tr>
<td>-geminates</td>
<td></td>
<td>- geminates</td>
<td>-degeminates</td>
</tr>
<tr>
<td>- no effect of decomposability (relative frequency)</td>
<td></td>
<td>- no effect of decomposability (semantic transparency and relative frequency)</td>
<td>- effect of decomposability (relative frequency)</td>
</tr>
<tr>
<td></td>
<td>- Nasal in negative in- is longer than in locative in-</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Lexical Phonology:
- results for *un-* support theory BUT results for *in-* and *-ly* do not
Our results and morpho-phonological theories

<table>
<thead>
<tr>
<th></th>
<th>un-</th>
<th>in-</th>
<th>-ly</th>
</tr>
</thead>
<tbody>
<tr>
<td>-geminates</td>
<td>-geminates</td>
<td>-degeminates</td>
<td></td>
</tr>
<tr>
<td>- no effect of decomposability (relative frequency)</td>
<td>- no effect of decomposability (semantic transparency and relative frequency)</td>
<td>- effect of decomposability (relative frequency)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Nasal in negative in- is longer than in locative in-</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Lexical Phonology:
- results for un- support theory BUT results for in- and -ly do not

Morphological Segmentability Hypothesis:
- not supported by un-, and in- but by -ly
## Our results and morpho-phonological theories

<table>
<thead>
<tr>
<th></th>
<th>un-</th>
<th>in-</th>
<th>-ly</th>
</tr>
</thead>
<tbody>
<tr>
<td>-geminates</td>
<td>-geminates</td>
<td>-degeminates</td>
<td></td>
</tr>
<tr>
<td>- no effect of decomposability (relative frequency)</td>
<td>- no effect of decomposability</td>
<td>- effect of decomposability</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(semantic transparency and relative frequency)</td>
<td>(relative frequency)</td>
<td></td>
</tr>
<tr>
<td>- Nasal in negative in- is longer than in locative in-</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Lexical Phonology:
- results for un- support theory BUT results for in- and –ly do not
- No support: morphological information is directly reflected in the speech signal

### Morphological Segmentability Hypothesis:
- not supported by un-, and in- but by –ly
Implications

• Lexical Phonology makes wrong empirical predictions (\textit{in-} and \textit{–ly})

• Morphological Segmentability Hypothesis is not supported by \textit{un-}, and \textit{in-} but by \textit{–ly}
  • Might be due to type frequencies of \textit{un-} and \textit{in-}-prefixed words

• Morphological information is directly reflected in the speech signal
  o \textit{in-}: Homophonous affixes exhibit different acoustic properties (cf. Plag, Homann & Kunter 2015 on \textit{S})
  o \textit{–ly}: Degree of morphological separability correlates with acoustic duration (cf. Hay 2007, Collie 2008)

• 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!
References


Assumptions about gemination in English

Kaye (2005):

“In a more formal, careful speech style, some native speakers may geminate some words, as Trask (op. cit.) notes. Some of these for some native speakers might, in fact, be spelling pronunciations. Thus, a word such as unknown may actually be pronounced by some with a geminated [nn] due to the pronunciation of its orthographic representation. A geminated [nn] in unknown, however, sounds awkward in my own speech, but there is always the possibility of a pragmatically based, purposeful gemination, i.e., for special effect.” (p...)
Assumptions about gemination in English

Kaye (2005:)

“In a more formal, careful speech style, some native speakers may geminate some words, as Trask (op. cit.) notes. Some of these for some native speakers might, in fact, be spelling pronunciations. Thus, a word such as unknown may actually be pronounced by some with a geminated [nn] due to the pronunciation of its orthographic representation. A geminated [nn] in unknown, however, sounds awkward in my own speech, but there is always the possibility of a pragmatically based, purposeful gemination, i.e., for special effect.” (p..)
# lm(formula = bc ~ TransitionType + LocSpeech + logRelFreq, data = unComplex2)
#
# Residuals:
#   Min        1Q    Median        3Q       Max
# -0.079017 -0.027548 -0.000418  0.024647  0.097765
#
# Coefficients:
#   Estimate Std. Error t value Pr(>|t|)
# (Intercept)         0.530718   0.015002  35.377  < 2e-16 ***
#   TransitionType\^nV 0.049355   0.009512   5.189 6.86e-07 ***
# TransitionType\^V -0.050646   0.006649  -7.617 2.85e-12 ***
# LocSpeech -0.007567   0.001067  -7.092 5.05e-11 ***
#   logRelFreq -0.001087   0.001231  -0.883    0.379
# ---
# Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
#
# Residual standard error: 0.03791 on 148 degrees of freedom
# Multiple R-squared:  0.6032,  Adjusted R-squared:  0.5925
# F-statistic: 56.24 on 4 and 148 DF,  p-value: < 2.2e-16
```r
# Im(formula = bc ~ NoCons + LocSpeech + StressPattern + Affix +
#     logRelFreq + MorphBound, data = imComplex4)
#
# Residuals:
#   Min        1Q    Median        3Q       Max
# -0.085471 -0.023571 -0.002151  0.023254  0.079917
#
# Coefficients:
#   Estimate Std. Error t value Pr(>|t|)
# (Intercept)             0.3162114  0.0117894  26.822  < 2e-16 ***
#   NoConsm#m  0.0439957  0.0078566   5.600 1.01e-07 ***
# LocSpeech  -0.0033194  0.0007864  -4.221 4.22e-05 ***
#   StressPatternstr-unstr -0.0312126  0.0086733  -3.599 0.000435 ***
#   AffixinNeg  0.0209547  0.0073654   2.845 0.005070 **
#   logRelFreq  0.0003406  0.0009046   0.376 0.707087
# MorphBoundtransparent -0.0053302  0.0082326  -0.647 0.518340
# ---
# Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
#
# Residual standard error: 0.03383 on 148 degrees of freedom
# Multiple R-squared:  0.5431, Adjusted R-squared:  0.5246
# F-statistic: 29.32 on 6 and 148 DF,  p-value: < 2.2e-16
```
# ly-model

```r
# 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>5</td>
<td>94</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>
Assumptions about gemination in English

*Bauer, Lieber, Plag (2013):*

‘Here we can observe non-uniform behavior of different suffixes and even of the same suffix with different speakers or in different levels of formality, speech tempo, and so on. In the style represented in pronouncing dictionaries, adverbial –ly, for example, geminates with *stalely* and *really*, but not with *fully* and *really*, nor with all suffixed bases ending in –al (*federally, globally, spiritually*), and variably with, for example *duddy* and *wholly.’ (p. 169)
Results 1: *un-*

\[ p < 0.05^{***} \]
Results 2: \textit{in-}

\begin{itemize}
  \item \textbf{local speech rate}
  \begin{itemize}
    \item \textbf{duration in seconds}
    \begin{itemize}
      \item \textbf{p<0.05***}
    \end{itemize}
  \end{itemize}
  \item \textbf{stress pattern}
  \begin{itemize}
    \item \textbf{duration in seconds}
    \begin{itemize}
      \item \textbf{p<0.05***}
    \end{itemize}
  \end{itemize}
\end{itemize}
Results 2: \textit{in-}

- No effect of semantic transparency on nasal duration for \textit{in-} 
  \[ p = 0.51834 \]

- No effect of relative frequency on nasal duration for \textit{in-} 
  \[ p = 0.707 \]
Results 3: -/y

Additional covariate: Syllabic (ment[l]y vs. ment[əl]y, odd[l]y)  p<0.05***