Gemination and Degemination in English Affixation

A Phonetic Investigation of Lexical Strata, Semantics and Decomposability

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

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

Affixation	
Prefixation *in* as in *in#numerous* *un* as in *un#natural* *dis* as in *dis#solve*	Suffixation *ly* as in *real#ly*
Compounds *fun# name*	
Word Boundaries *He wanted some # more.*	

Morphology:

Gemination only occurs at morphological boundaries

The phonological level

Two identical underlying segments

u /n#n/ atural

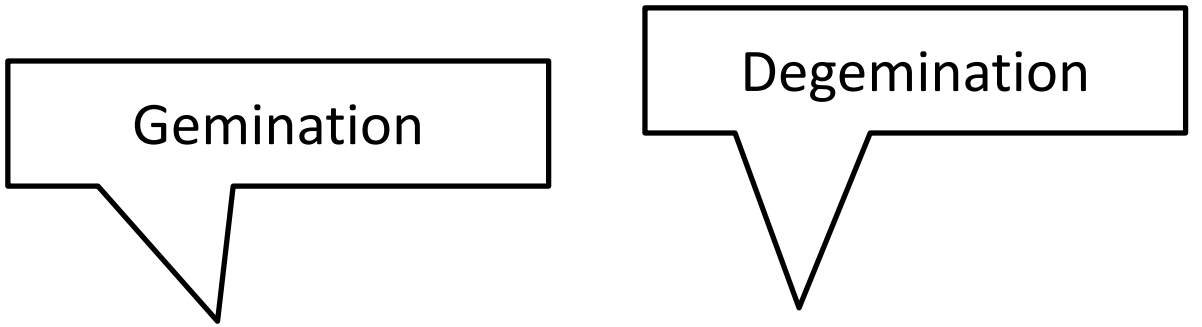
fu /n # n/ame

so /m # m/ ore

Phonology:

Sequence of two identical underlying segments

The phonetic level



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

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

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

Phonetics:

Duration as a phonetic correlate for gemination/ degeminati on

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

	Level 1	Level 2
Morphological Process	in + possible	un + predictable
Phonological Process		
Phonetic Outcome		

Weak boundary

Lexical Phonology

	Level 1	Level 2
Morphological Process	in + possible	un + predictable
Phonological Process	i/m/possible	
Phonetic Outcome		

Assimilation

Lexical Phonology

	Level 1	Level 2
Morphological Process	in + possible	un + predictable
Phonological Process	i/m/possibl	
Phonetic Outcome	i[m]possible	

Strong boundary

Assimilation

Lexical Phonology

	Level 1	Level 2
Morphological Process	in + possible	un + predictable
Phonological Process	i/m/possible	u/n/predictable
Phonetic Outcome	i[m]possible	

Assimilation **No assimilation**

Lexical Phonology

	Level 1	Level 2
Morphological Process	in + possible	un + predictable
Phonological Process	i/m/possible	u/n/predictable
Phonetic Outcome	i[m]possible	u[n]predictable

Assimilation **No assimilation**

Lexical Phonology

	Level 1	Level 2
Morphological Process	in + possible	un + predictable
Phonological Process	i/m/possible	u/n/predictable
all morphological information is erased		
Phonetic Outcome	i[m]possible	u[n]predictable

Assimilation | **No assimilation**

Lexical Phonology and Gemination

Weak boundary

strong boundary

	Level 1	Level 2
Morphological Process	in + numerous	un + natural sole + ly
Phonological Process	i/ /umerous	u/ /atural so/ /y
Phonetic Outcome	i[ ]umerous	u[ ]atural so[ ]y

Lexical Phonology and Gemination

	Level 1	Level 2
Morphological Process	in + numerous	un + natural sole + ly
Phonological Process	i/n/umerous	
Phonetic Outcome		

Deletion of one segment

Lexical Phonology and Gemination

	Level 1	Level 2
Morphological Process	in + numerous	un + natural sole + ly
Phonological Process	i/n/umerous	
Phonetic Outcome	i[n]umerous	

Degemination

Lexical Phonology and Gemination

	Level 1	Level 2
Morphological Process	in + numerous	un + natural sole + ly
Phonological Process	i/n/umerous	u/nn/atural so/ll/y
Phonetic Outcome	i[n]umerous	

Degemination **No deletion**

Lexical Phonology and Gemination

	Level 1	Level 2
Morphological Process	in + numerous	un + natural sole + ly
Phonological Process	i/n/umerous	u/nn/atural so/ll/y
Phonetic Outcome	i[n]umerous	u[n:]atural so[l:]y

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):

_un-_ geminates,
_in-_ degeminates

"In general such prefixes result in a doubled conso
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]eeper)." (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)

item	Word Frequency	Base Frequency	Relative frequency
uninhabitable	224	39	224: 39 = 5.74

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)

item	Word Frequency	Base Frequency	Relative frequency
uninhabitable	224	39	224: 39 = 5.74

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)

item	Word Frequency	Base Frequency	Relative frequency
uninhabitable	224	39	224: 39 = 5.74
immoral	390	4994	390: 4994 = 0.078

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 categerocial
- Gemination depends on the affix
  - *un-* and *–ly* geminate
  - *in-* degeminates
- 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*, *immigrational*)

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 morpho-phonological 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?

in-	*un-*	*-ly*
Kaye 05: speaker-dependent gemination Oh and Redford (2012): type-dependent gemination	Kaye 05: gemination Oh and Redford (2012): gemination	No evidence

- 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 (orthograhic) 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 /ɪm/ was investigated
- Manual segmentation and acoustic measurments 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‘

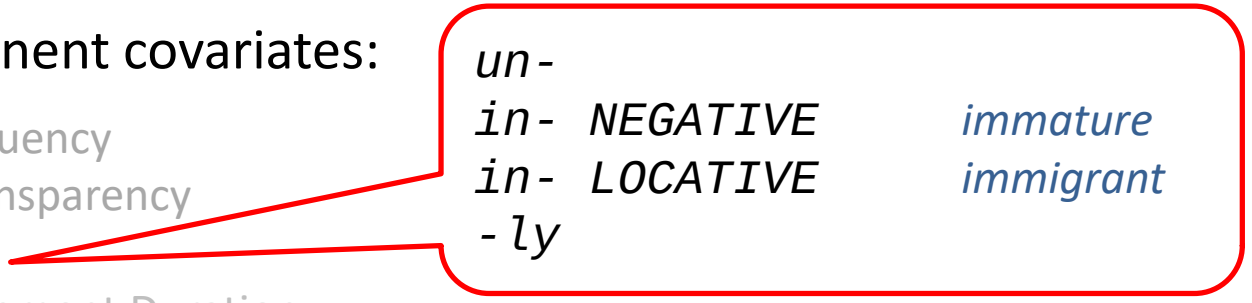
`opaque`:

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 **duration** as dependent variable and **number of consonants** (single vs. double) as crucial predictor

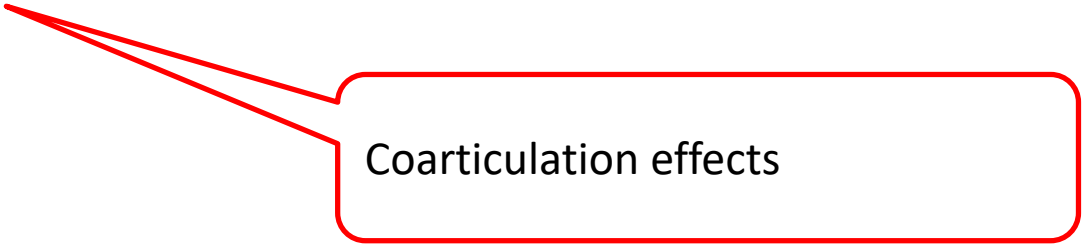
- Coding of pertinent covariates:
  - Relative Frequency
  - Semantic Transparency
  - Affix
  - Preceding Segment Duration
  - Preceding Segment
  - Following segment
  - Speech Rate
  - Position in utterance
  - Prosodic Structure
  - Syllabicity
  - Word Form Frequency

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



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.

$$\frac{\text{Number of segments}}{\text{word duration}}$$

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

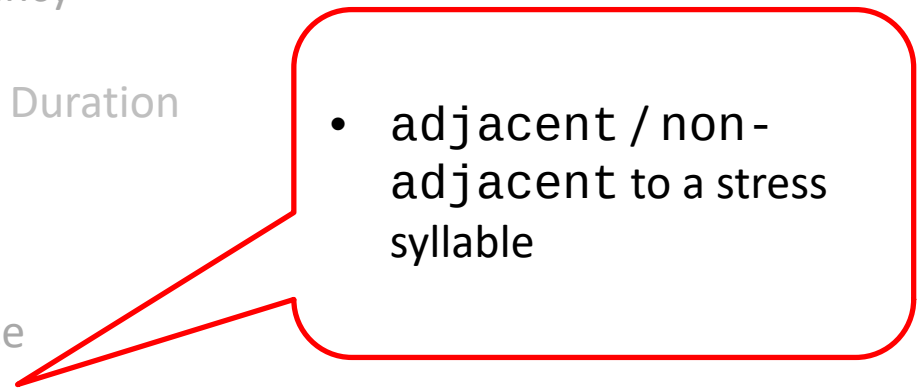
Final lengthening effect

`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



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

for *–ly:*

`syllabic` /l/ should be longer

ment[l̩]y ment[əl]y odd[l]y

Methodology

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

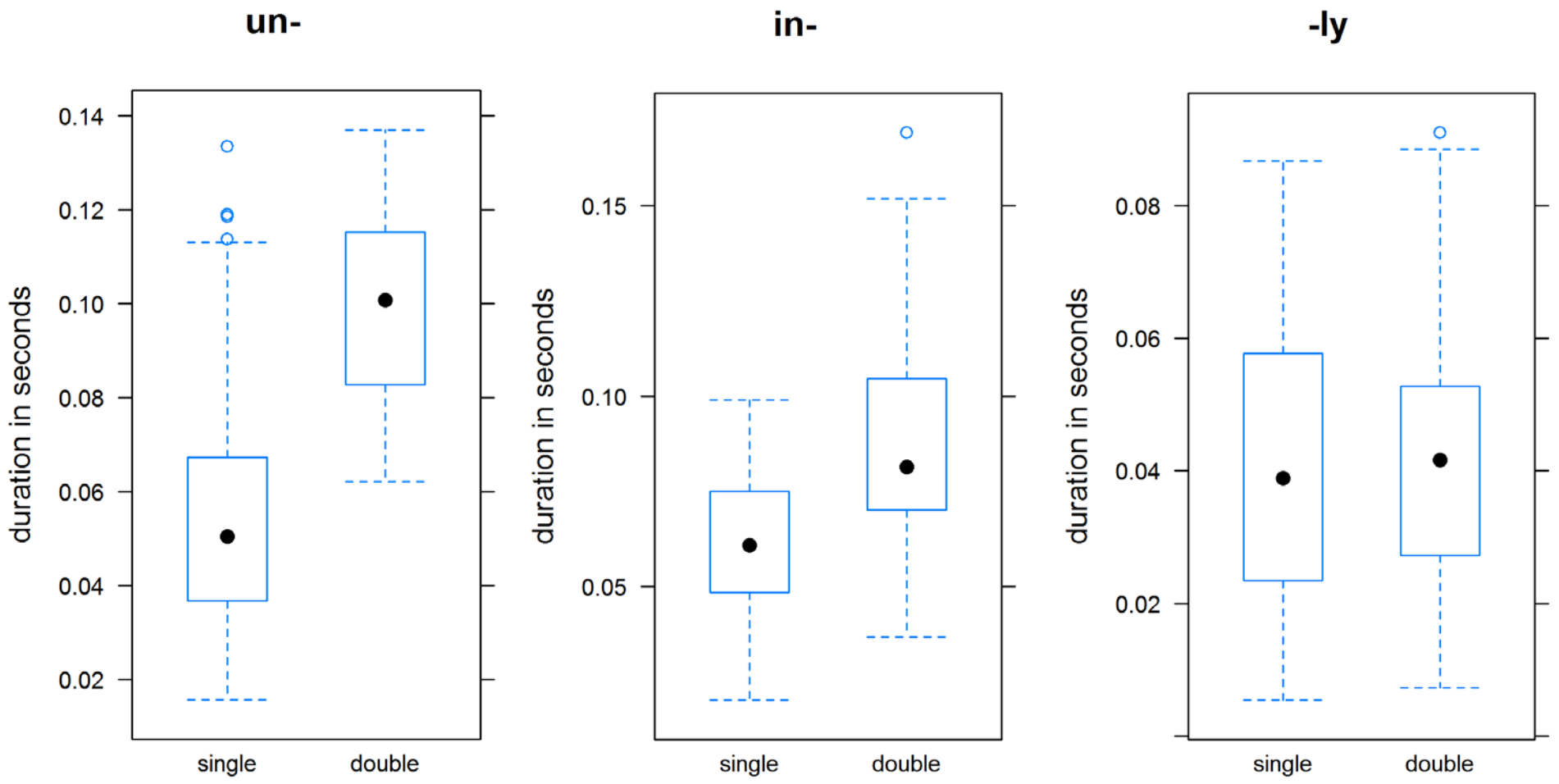
- Coding of pertinent covariates:
  - Relative Frequency
  - Semantic Transparency
  - Affix
  - Preceding Segment Duration
  - Preceding Segment
  - Following segment
  - Speech Rate
  - Position in utterance
  - Prosodic Structure
  - Syllabicity
  - Word Form Frequency

More frequent words are produced faster

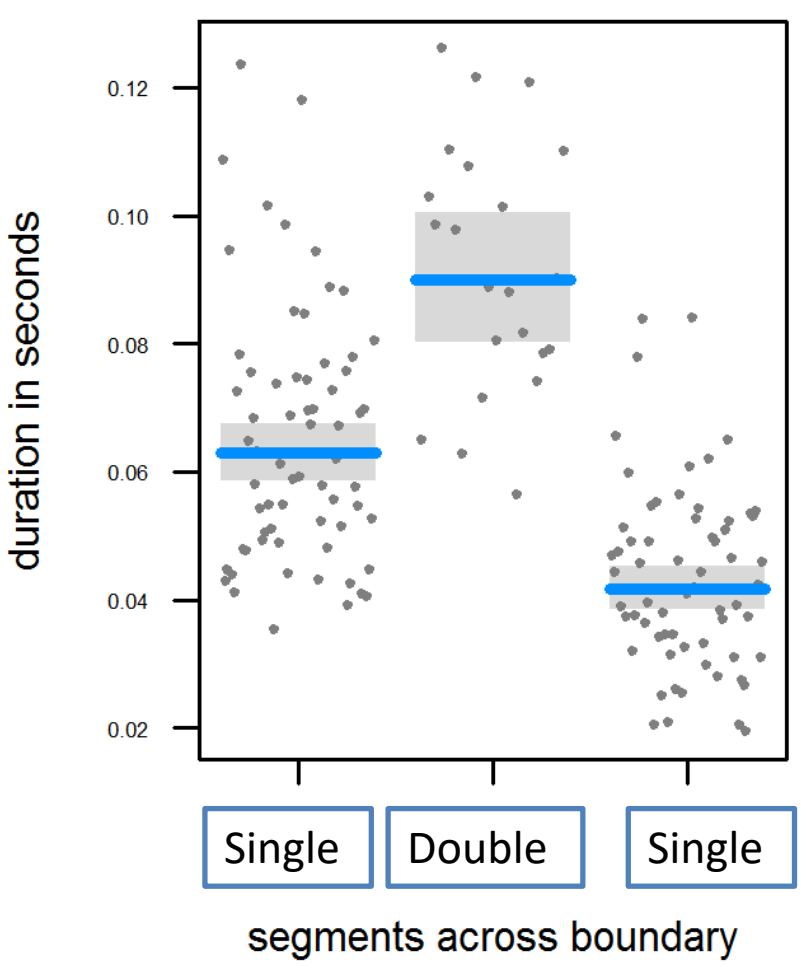
Overview of the data

	Double Consonant	Single Consonant	*Total per affix*
un-	22	136	*158*
in-	89	67	*156*
-ly	81	75	*156*

Results: Overview



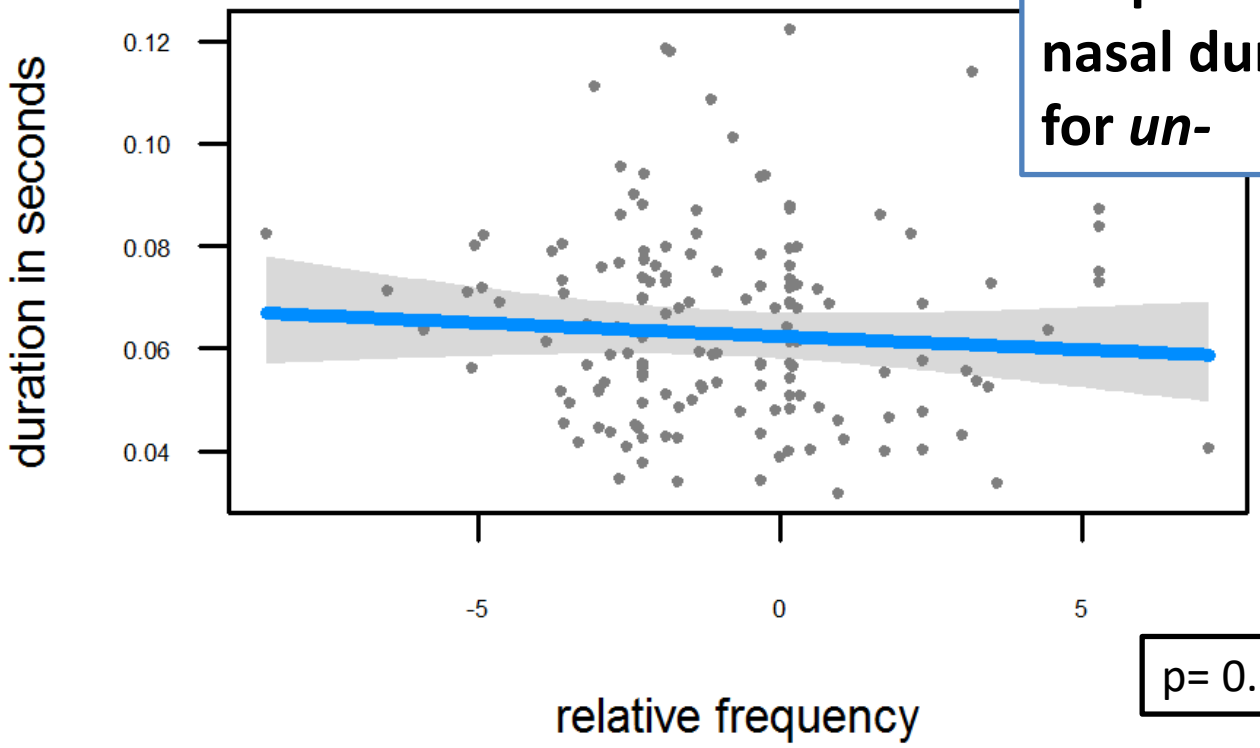
Results 1: *un-*



***un-* geminates**

p<0.05***

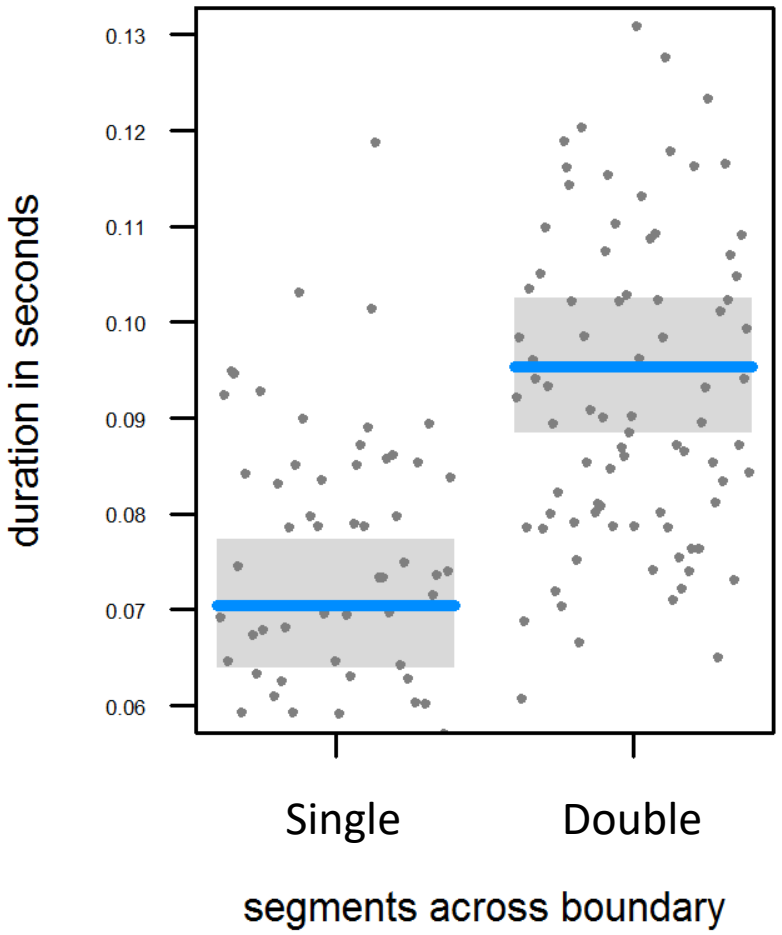
Results 1: *un-*



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

p= 0.379

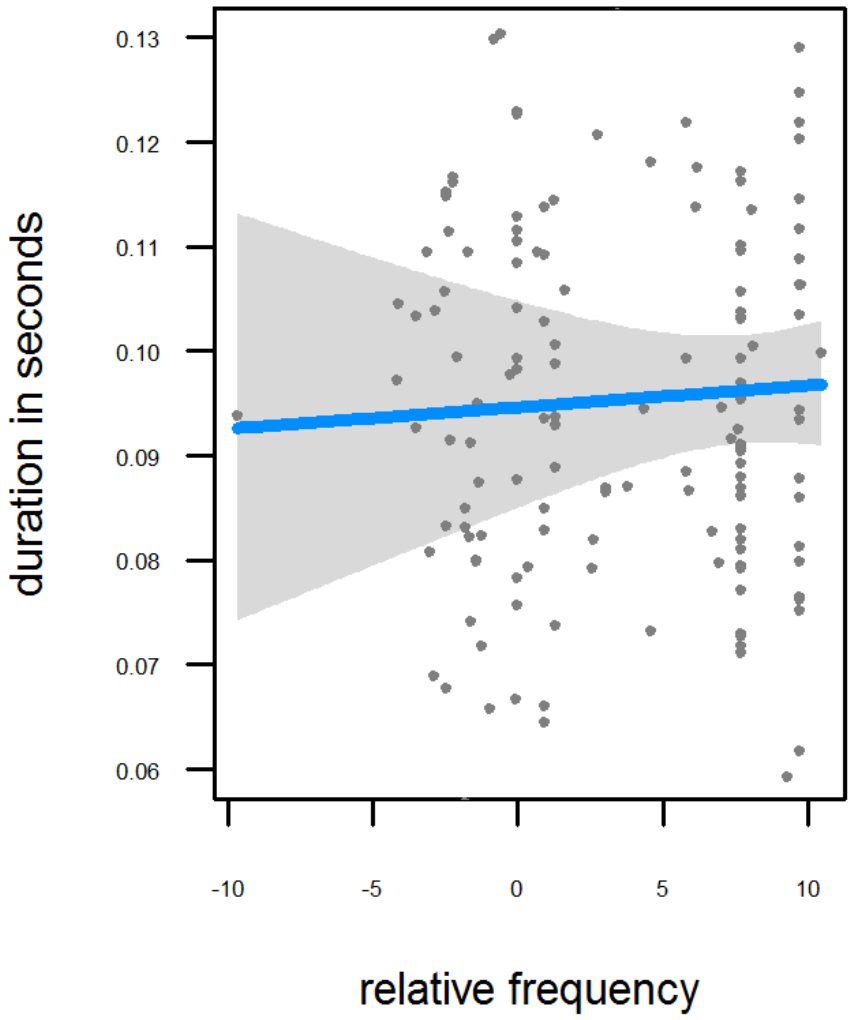
Results 2: *in-*



***in-* geminates**

p<0.05***

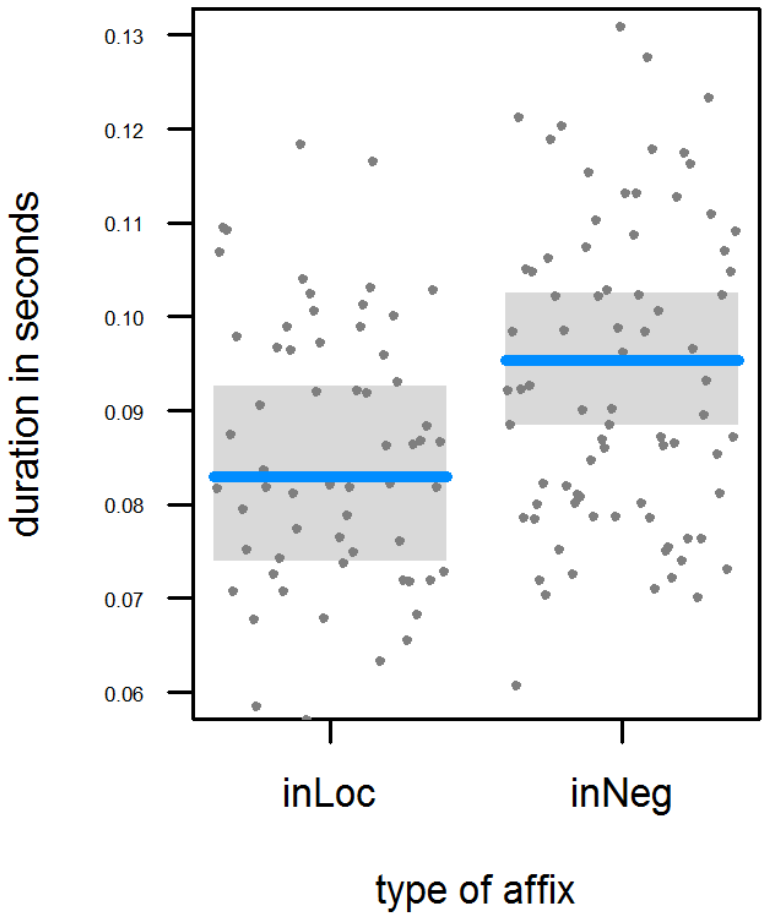
Results 2: *in-*



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

p= 0.707

Results 2: *in-*

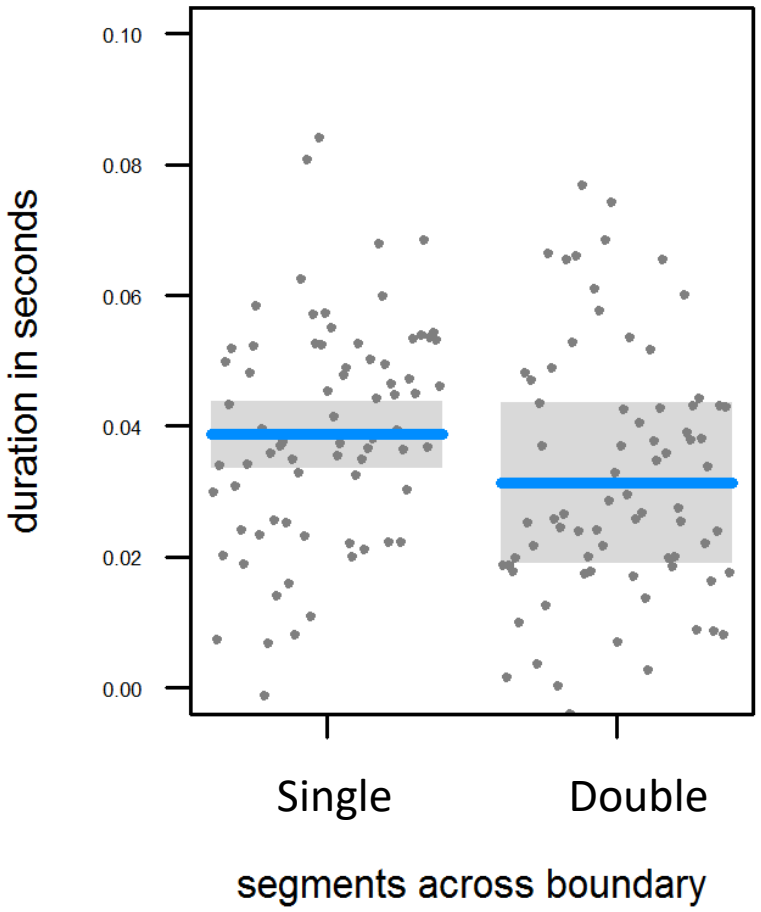


Effect of affix on nasal duration for *in-*

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

p<0.05***

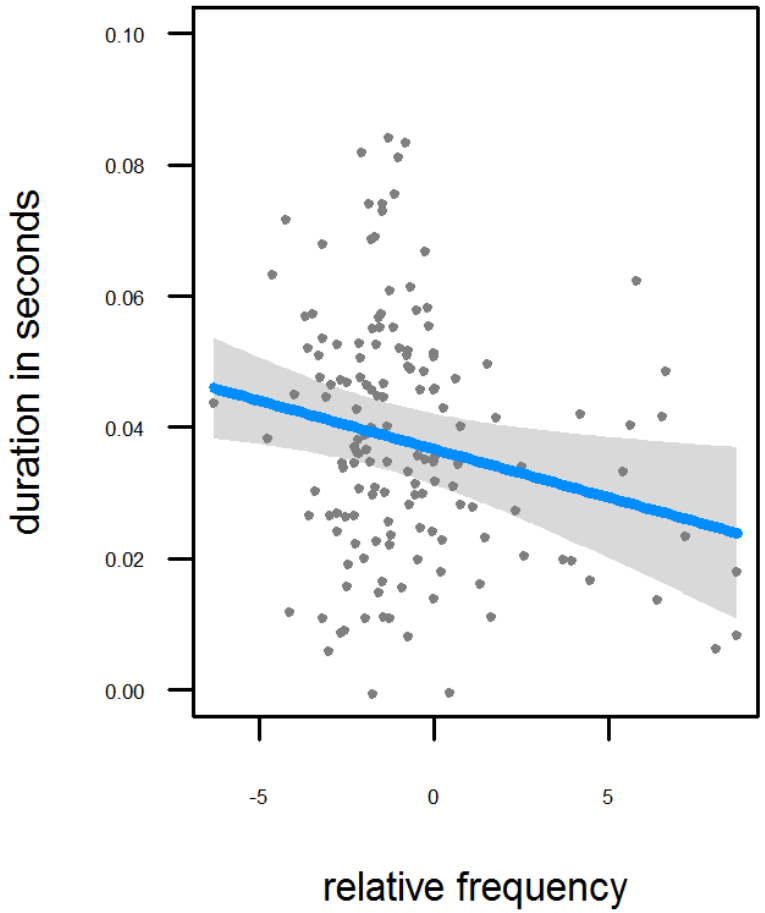
Results 3: *-ly*



-ly ***de*geminates**

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

un-	*in-*	*-ly*
-geminates - no effect of decomposabilty (relative frequency)	- geminates - no effect of decomposabilty (semantic transparency and relative frequency) - Nasal in negative *in-* is longer than in locative *in-*	-degeminates - effect of decomposabilty (relative frequency)

Our results and morpho-phonological theories

un-	*in-*	*-ly*
-geminates - no effect of decomposabilty (relative frequency)	- **geminates** - no effect of decomposabilty (semantic transparency and relative frequency) - Nasal in negative *in-* is longer than in locative *in-*	-**degeminates** - effect of decomposabilty (relative frequency)

Lexical Phonology:

- results for *un-* support theory BUT results for *in-* and *–ly* do not

Our results and morpho-phonological theories

un-	*in-*	*-ly*
-geminates **- no effect of decomposabilty (relative frequency)**	- geminates - **no effect of decomposabilty (semantic transparency and relative frequency)** - Nasal in negative *in-* is longer than in locative *in-*	-degeminates **- effect of decomposabilty (relative frequency)**

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

un-	*in-*	*-ly*
-geminates - no effect of decomposabilty (relative frequency)	- geminates - no effect of decomposabilty (semantic transparency and relative frequency) - **Nasal in negative *in-* is longer than in locative *in-***	-degeminates **- effect of decomposabilty (relative frequency)**

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 (*in-* and *–ly*)
- Morphological Segmentability Hypothesis is not supported by *un-*, and *in-* but by *–ly*
  - Might be due to type frequencies of *un-* and *in-*prefixed words
- Morphological information is directly reflected in the speech signal
  - *in-*: Homophonous affixes exhibit different acoustic properties (cf. Plag, Homann & Kunter 2015 on S)
  - *-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

Bauer, L. (2001). *Morphological productivity*. *Cambridge studies in linguistics: Vol. 95*. Cambridge England, New York: Cambridge University Press.

Bauer, L., Lieber, R., & Plag, I. (2013). *The Oxfort reference guide to english morphology*. *Oxford Linguistics*. Oxford: Oxford University Press.

Blazej, L. J., & Cohen-Goldberg, A. M. (2015). Can we hear morphological complexity before words are complex? *Journal of experimental psychology. Human perception and performance, 41*(1), 50–68.

Boersma, P. & Weenink, D. (2014). Praat: doing phonetics by computer. Retrieved from http://www.praat.org/

Cohen-Goldberg, Ariel M. (2013): Towards a theory of multimorphemic word production: The heterogeneity of processing hypothesis. In: Language and Cognitive Processes 28 (7), S. 1036–1064.

Collie, S. (2008). English stress preservation: the case for 'fake cyclicity'. *English Language and Linguistics, 12*(03), 505–532.

Cruttenden, Alan; Gimson, Alfred Charles (2014): Gimson's pronunciation of English. 8th ed. London, New York: Routledge.

Giegerich, H. J. (1999). *Lexical Strata in English: Morphological Causes, Phonological Effects*: Cambridge University Press.

Godfrey, John J.; Holliman, Edward (1997): Switchboard-1 Release 2. [Philadelphia, Pa.]: Linguistic Data Consortium.

Hay, J. (2007). The phonetics of 'un'. In J. Munat (Ed.), Studies in functional and structural linguistics: v. 58. Lexical creativity, texts and contexts (pp. 39–57). Amsterdam, Philadelphia: J. Benjamins Pub. Co.

Kaye, A. S. (2005). Gemination in English. *English Today, 21*(2), 43–55.

Kemps, Rachel J J K, Ernestus, M., Schreuder, R., & Baayen, R. H. (2005). Prosodic cues for morphological complexity: the case of Dutch plural nouns. *Memory & cognition, 33*(3), 430–446.

Kiparsky, Paul (1982): Lexical morphology and phonology. In: Linguistics in the morning calm. Selected papers from SICOL-1981. Unter Mitarbeit von The linguistic society of Korea. Seoul, Korea: Hanshin Pub. Co.

Mohanan, K. P. (1986). The theory of lexical phonology. Studies in natural language and linguistic theory: [v. 6]. Dordrecht, Boston, Norwell, MA: D. Reidel Pub. Co.; Sold and distributed in the U.S.A. and Canada by Kluwer Academic.

Oh, Grace E.; Redford, Melissa A. (2012): The production and phonetic representation of fake geminates in English. In: Journal of Phonetics 40 (1), S. 82–91.

Plag, I. (2014). Phonological and phonetic variability in complex words: An uncharted territory. *Italian Journal of Linguistics / Rivista di Linguistica.*

Plag, I., Homann, J., & Kunter, G. (2015). Homophony and morphology: The acoustics of word-final S in English. *Journal of Linguistics*.

R Development Core Team. (2014). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.r-project.org

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:

Variation depending on speaker, speech style, orthography….

"In a more formal, careful speech style, some nati
geminate some words, as Trask (*op. cit*.) notes. So
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..)

un-model

```
#   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 ***
#   TransitionTypen#nV  0.049355   0.009512   5.189 6.86e-07 ***
# TransitionTypen#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
```

im-model

```
#   lm(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

```
#   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

	Doubles	Singles
un-	5	94
in-	17	65
-ly	76	72

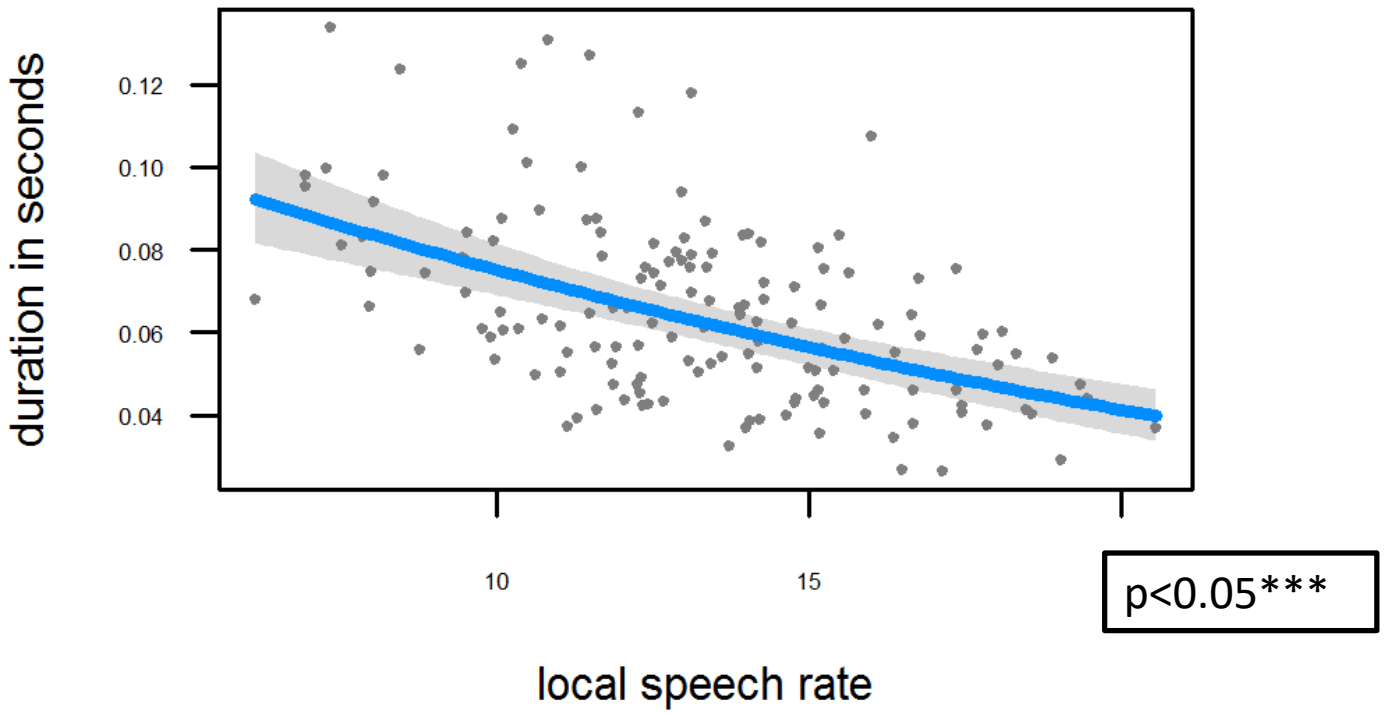
Assumptions about gemination in English

Bauer, Lieber, Plag (2013):

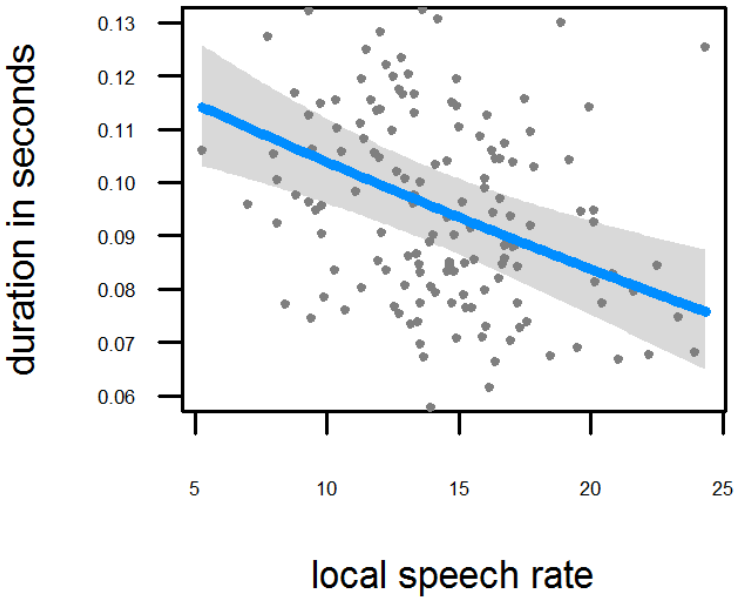
Variation depending on speakers, formality, speech tempo, types

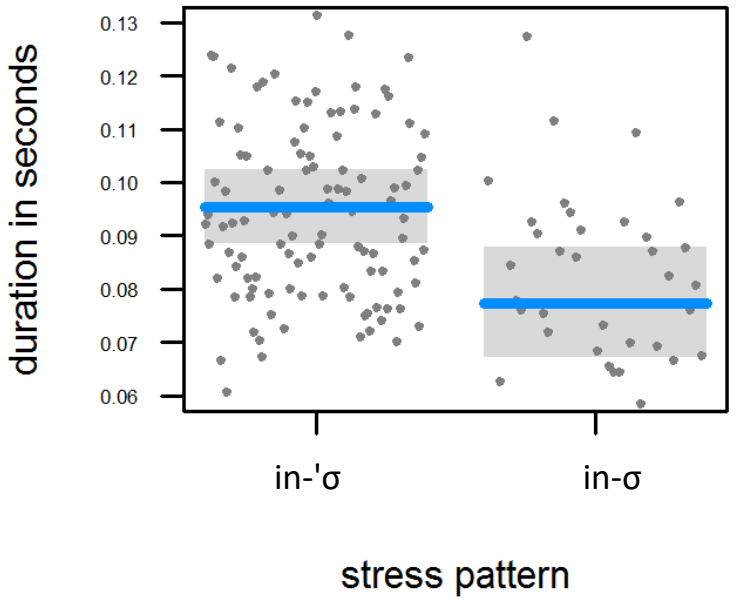
'Here we can observe non-uniform behavior of dif[illegible]d even of the same suffix with different speakers or [illegible] of formalilty, speech tempo, and so on. In the style r[illegible] pronouncing dictionaries, adverbial *–ly,* for examp[illegible]n *stalely* and *really*, but not with *fully* and *really*, nor with all suffixed bases ending in *–al* (*federally, globally, spiritually*), and variably with, for example *dully* and *wholly*.' (p. 169)

Results 1: *un-*


duration in seconds
0.12
0.10
0.08
0.06
0.04
10
15
local speech rate
p<0.05***

Results 2: *in-*

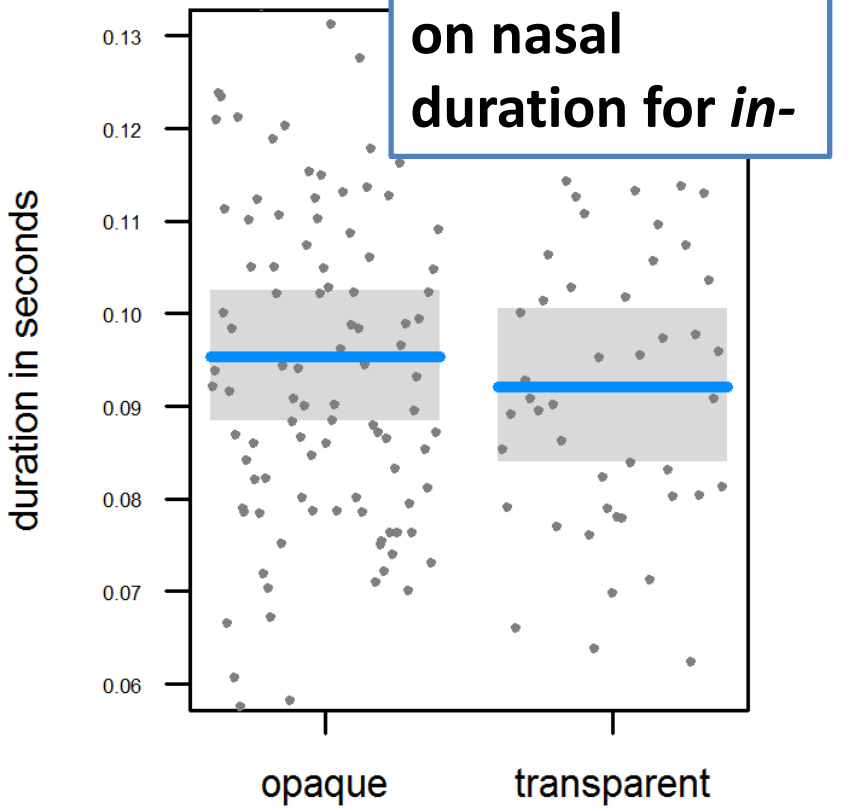




p<0.05***

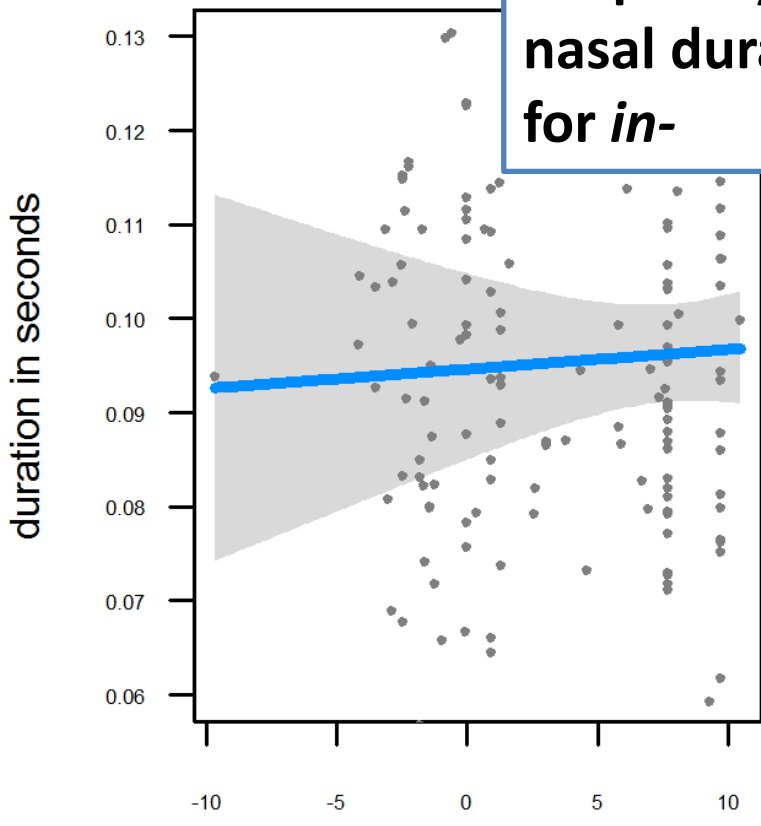
Results 2: *in-*

No effect of semantic transparency on nasal duration for *in-*



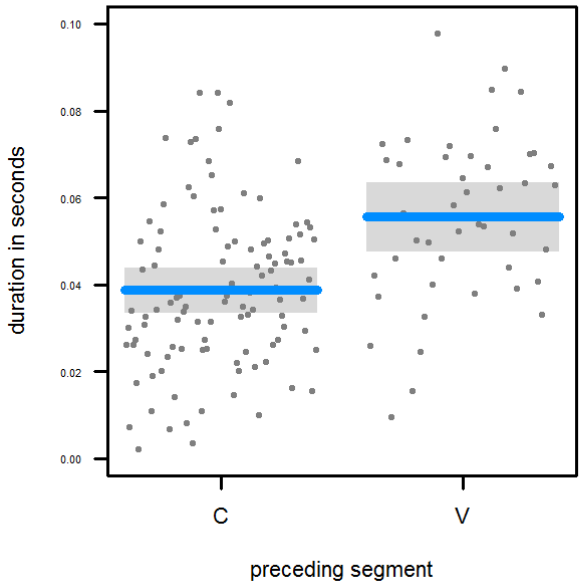
p= 0.51834

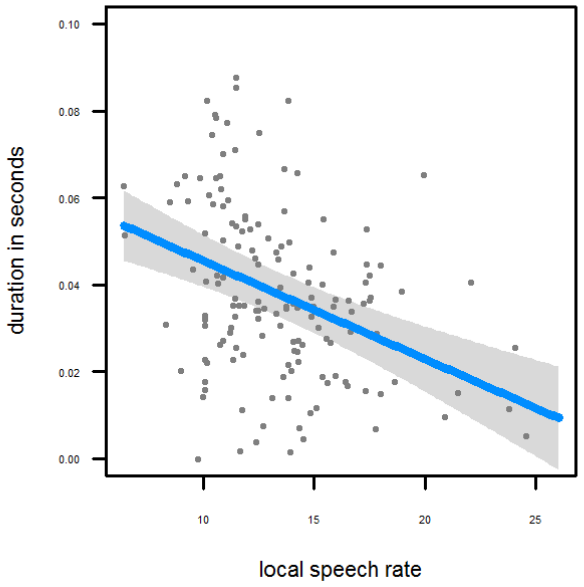
No effect of relative frequency on nasal duration for *in-*

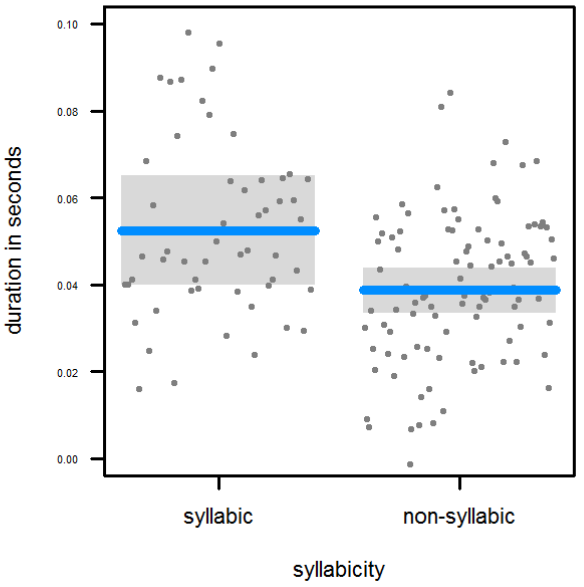


p= 0.707

Results 3: *-ly*







Additional covariate: Syllabicity (ment[l̩]y vs. ment[əl]y, odd[l]y)

p<0.05***