Gemination and Degemination in English Affixation

Lexical Strata, Variability, and Phonetic Evidence

Sonia Ben Hedia & Ingo Plag

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

48th Annual Meeting of the Societas Linguistica Europaea, 2 – 5 September 2015

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
  - Stress shift (Bauer, Lieber & Plag 2013 on *-able*)
  - Stress preservation (Collie 2008, relative frequency as a proxy for morphological segmentability)

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

- Serious implications for theories of morpho-phonology (Plag 2014)
  - Exception vs. rule
  - 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
  - Gemination: Longer duration than a singleton
  - 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*
  - ... is variable (*stalely* vs. *fully*, Bauer 2001, Bauer, Lieber & Plag 2013)
  - ... geminates (Lexical Phonology: level 2 affix)

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

Predictions

	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

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 /ɪm/ was investigated
- Manual segmentation and acoustic measurements in Praat (Boersma & Weenink 2014)

Methodology

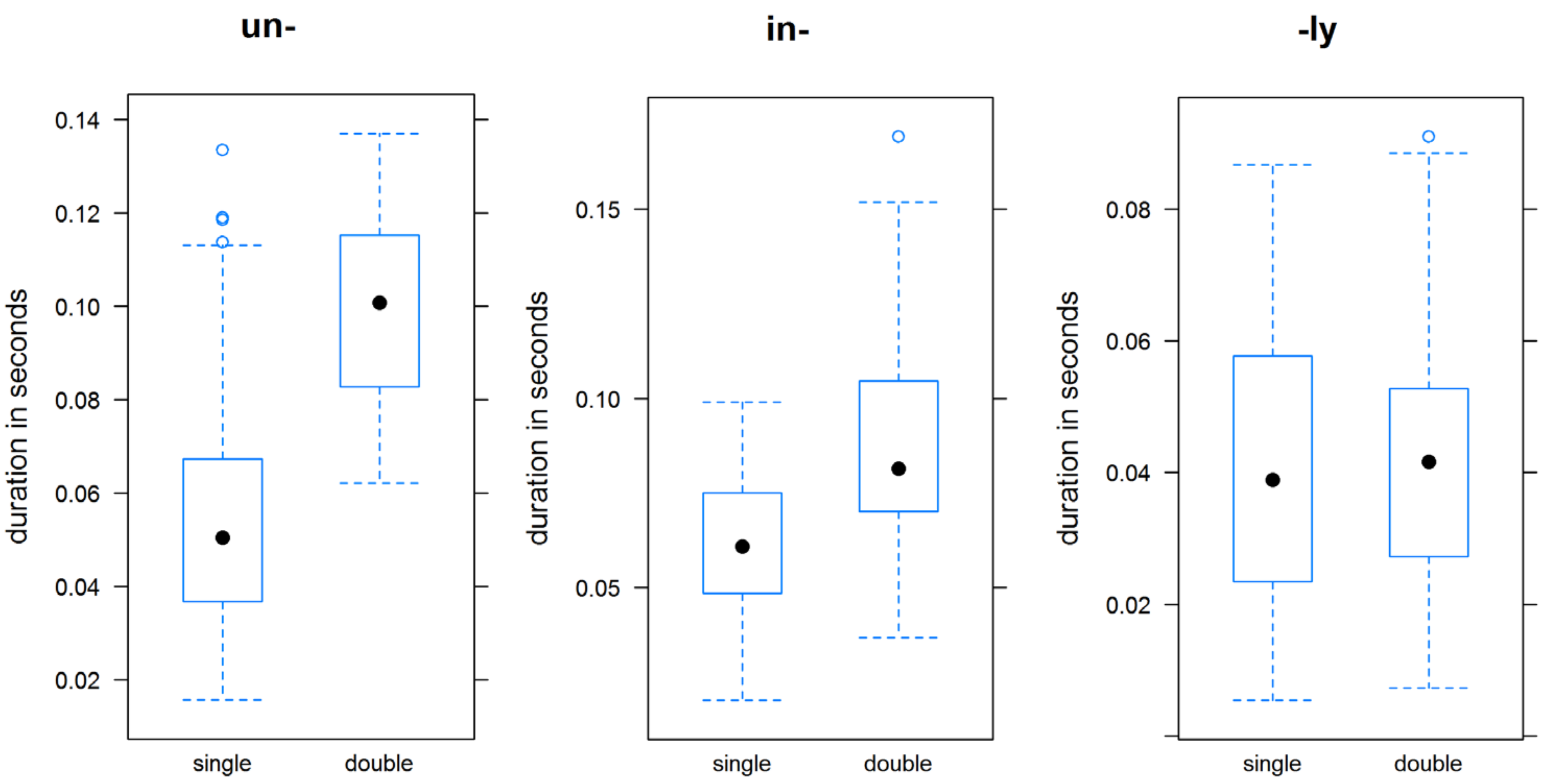
- Statistical Analysis: Multiple regression with **duration** as dependent variable and **number of consonants** (single vs. double) as predictor

- Coding of pertinent covariates:
    - Preceding Segment Duration
    - Preceding Segment
    - Following segment
    - Speech Rate
    - Stress
    - Syllabicity
    - Word Form Frequency
    - Relative Frequency
    - Affix
    - Semantic Transparency

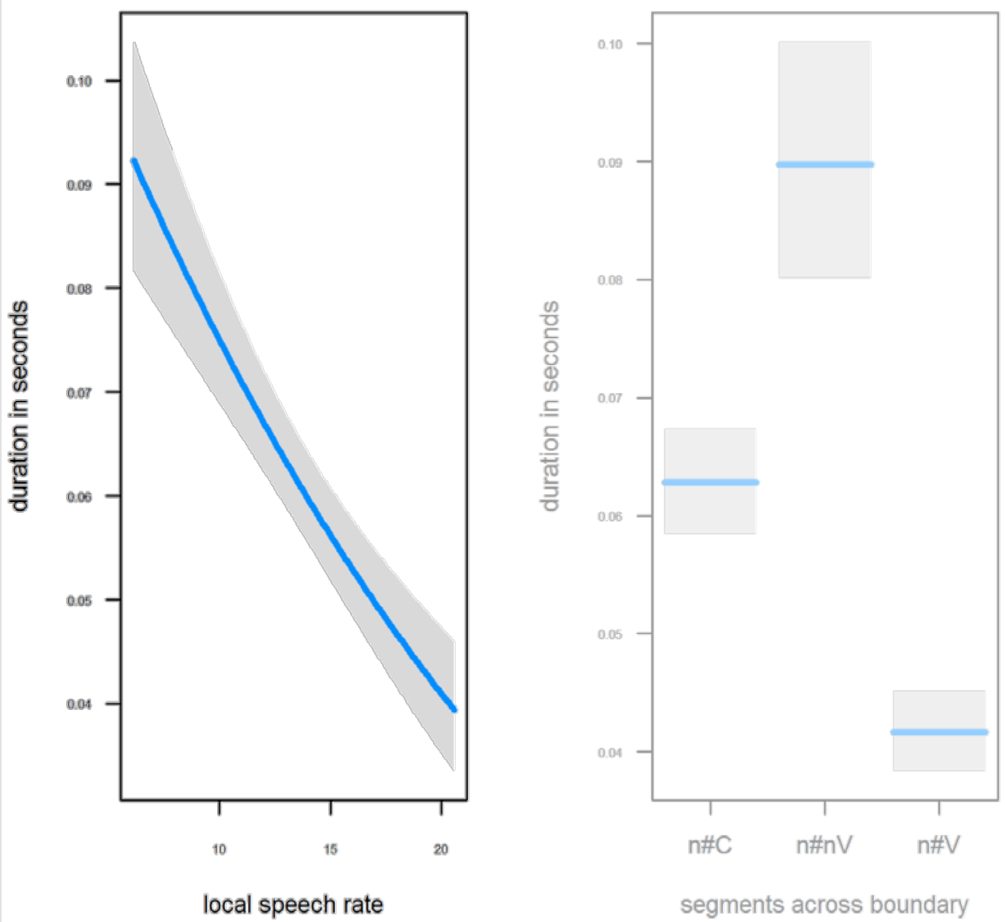
Overview of the data

	Double Consonant	**Single Consonant**	**Total per affix**
un-	22	136	158
in-	89	67	156
-ly	76	80	156

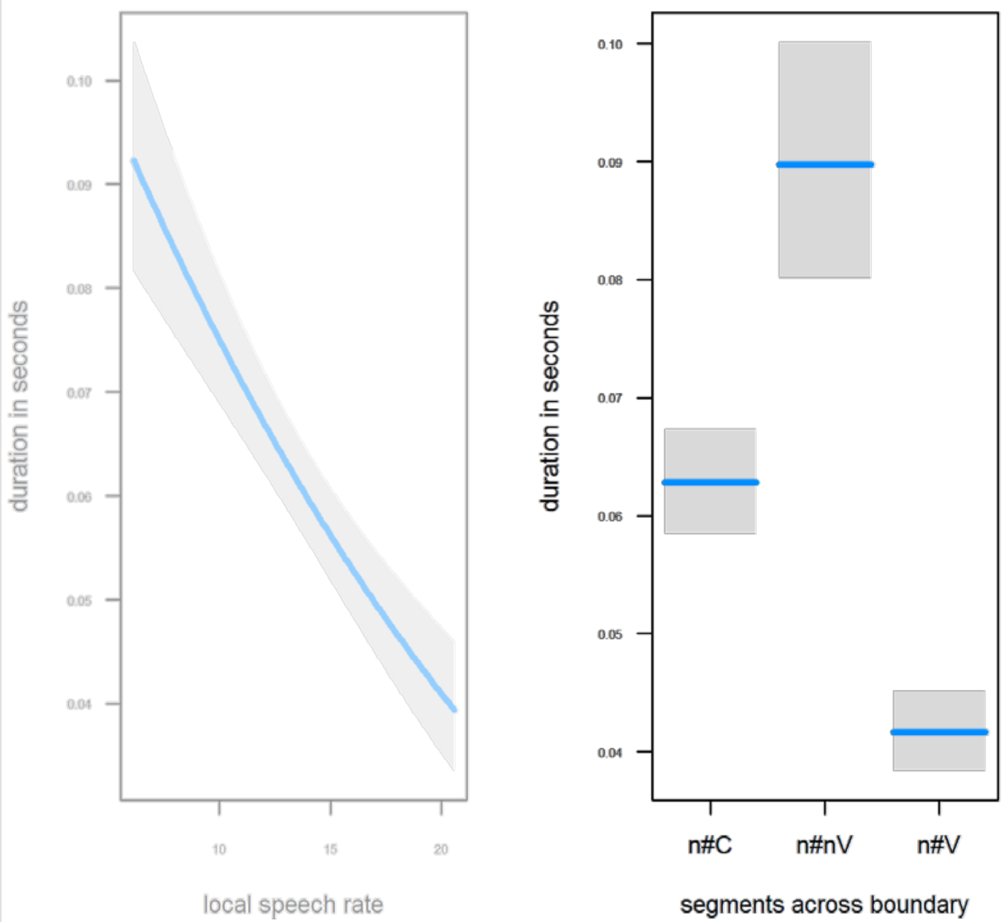
Results: Overview



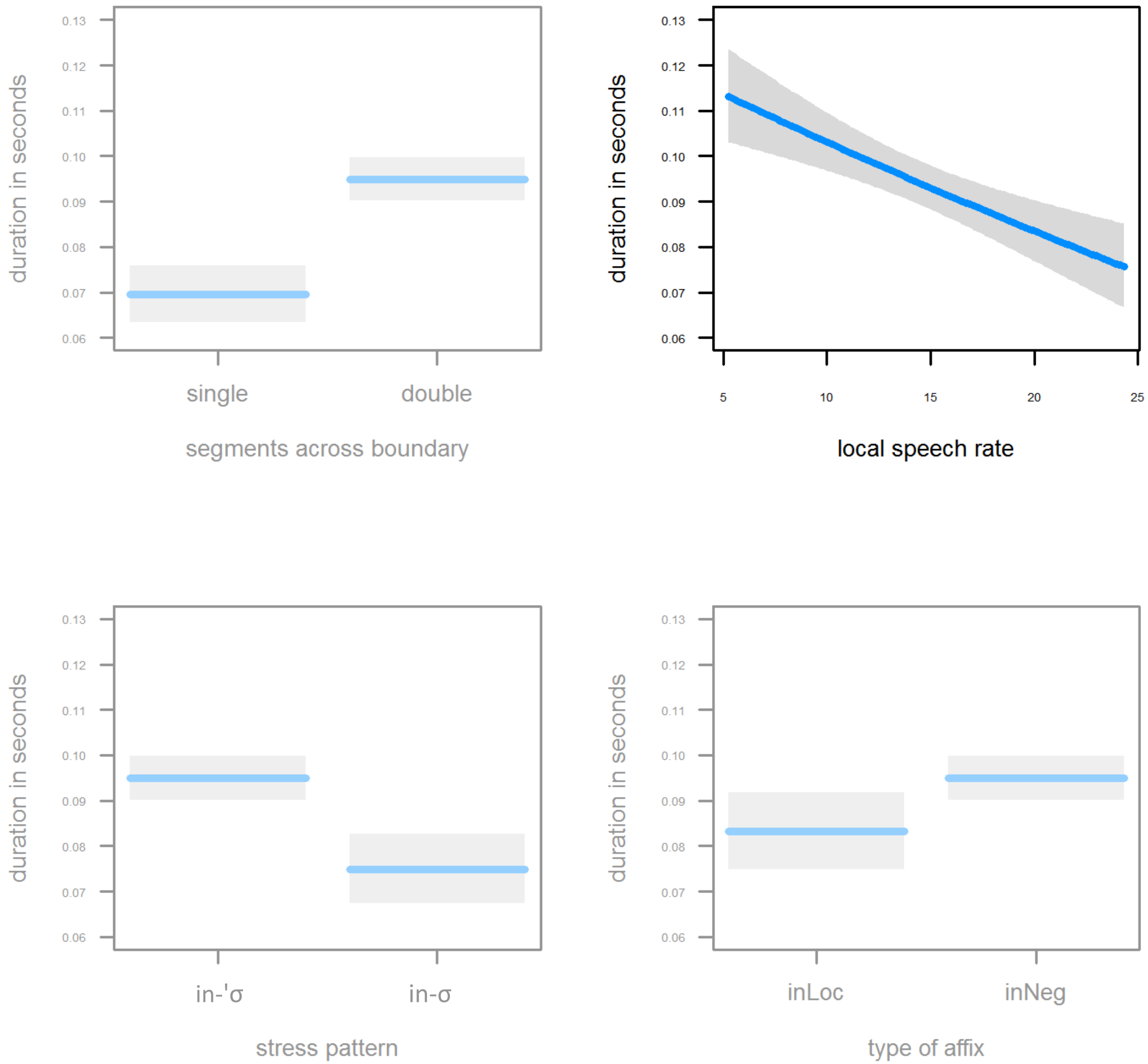
Results 1: *un-* geminates



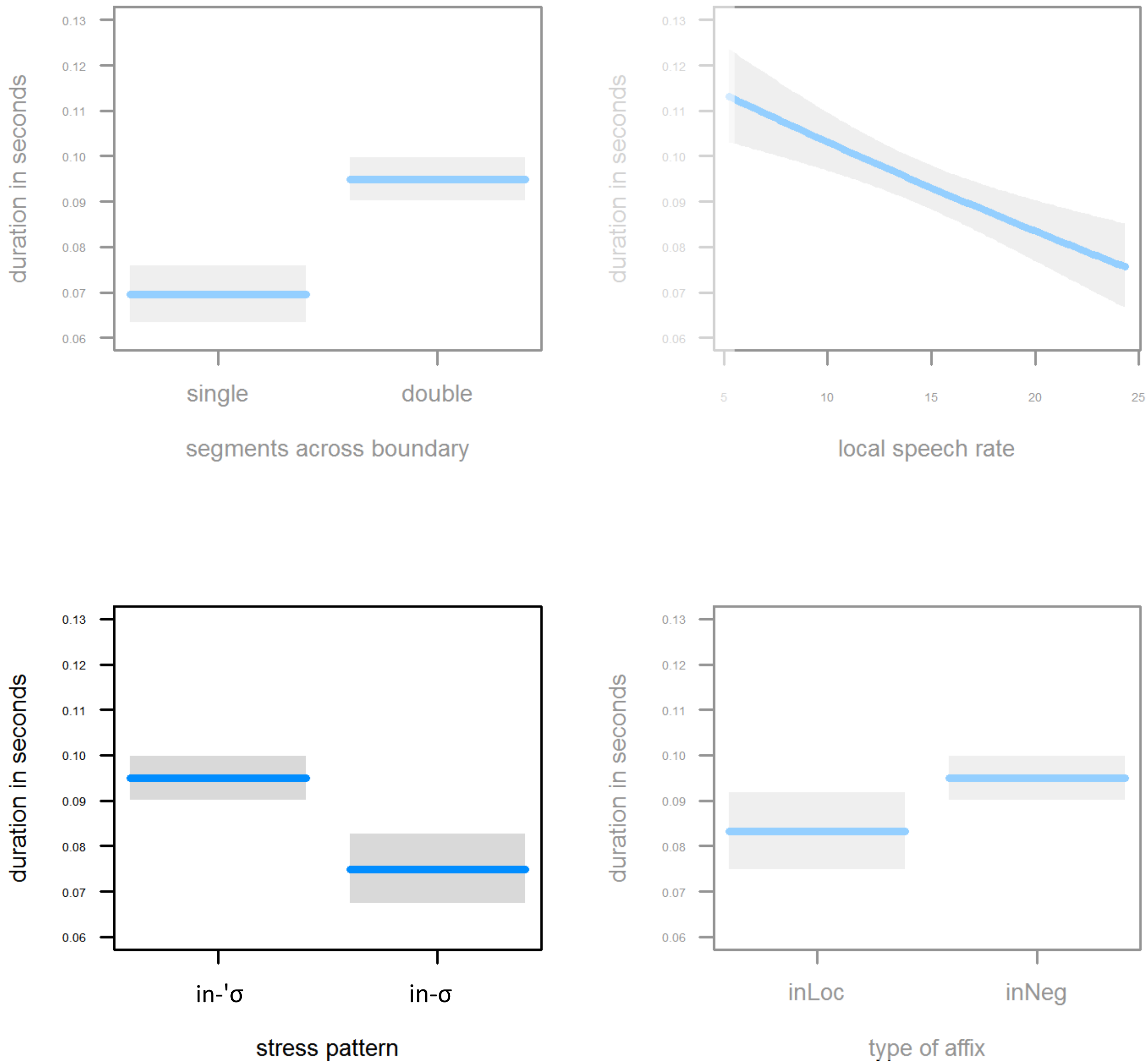
Results 1: *un-* geminates



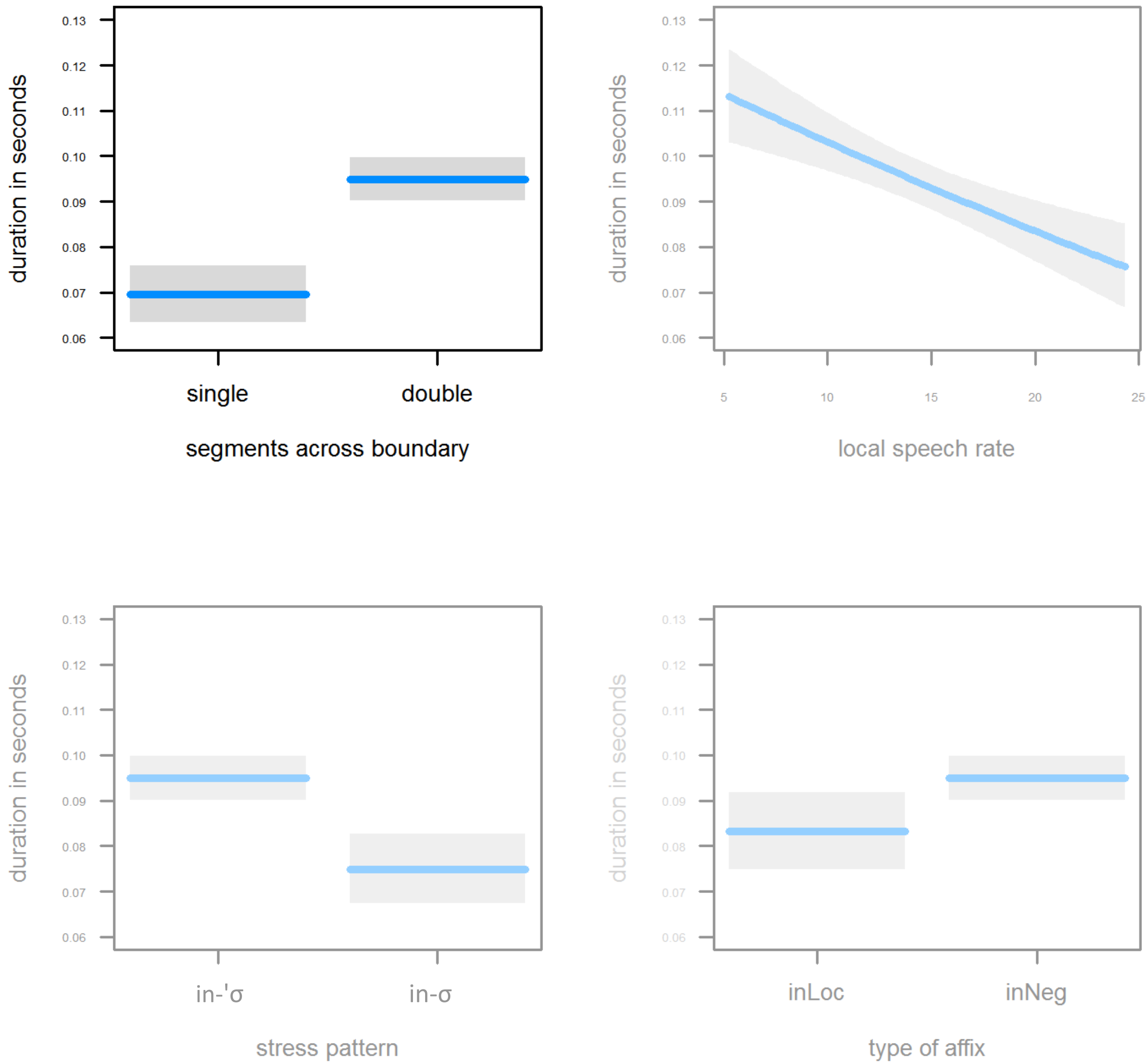
Results 2: *in-* geminates



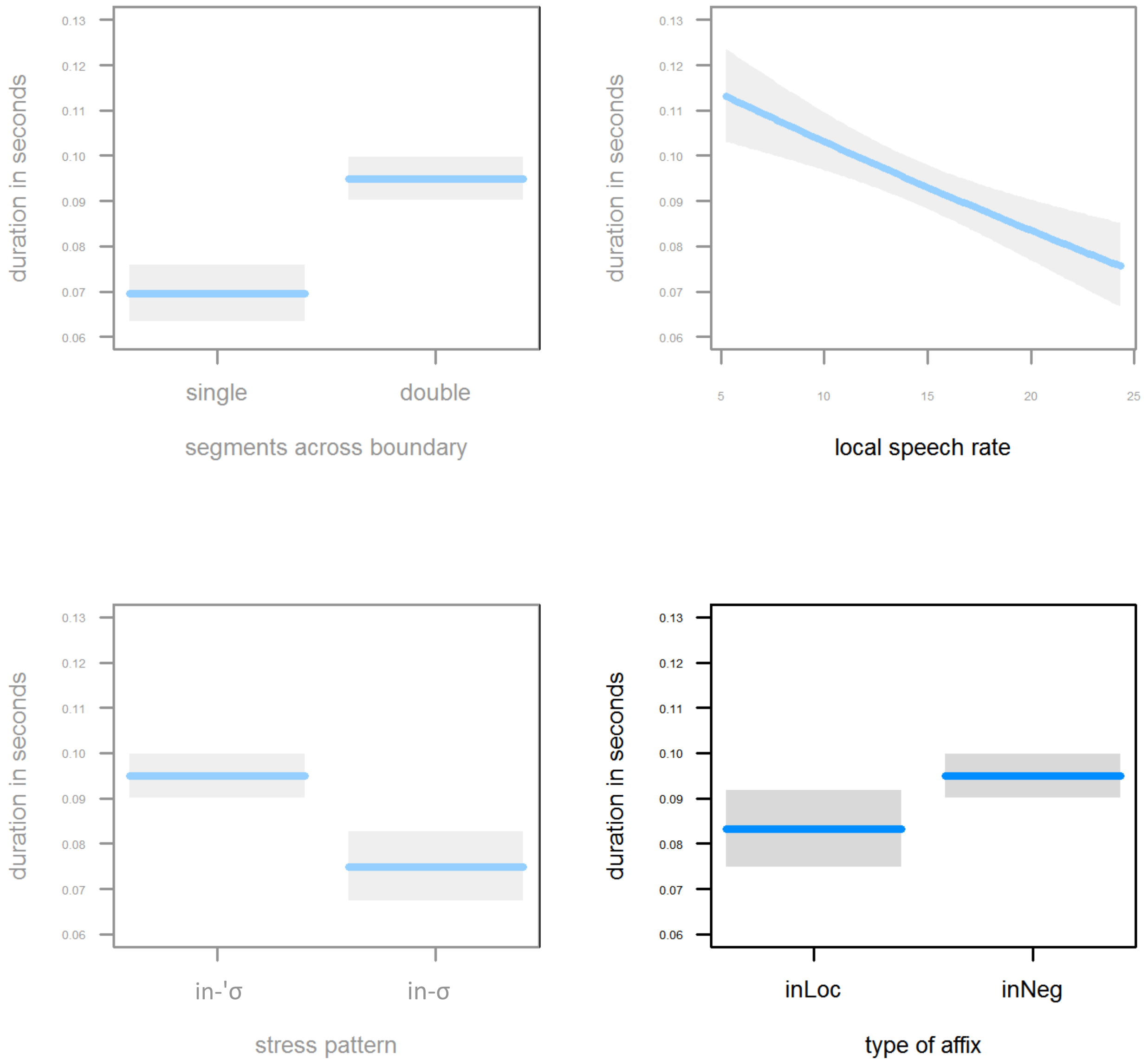
Results 2: *in-* geminates



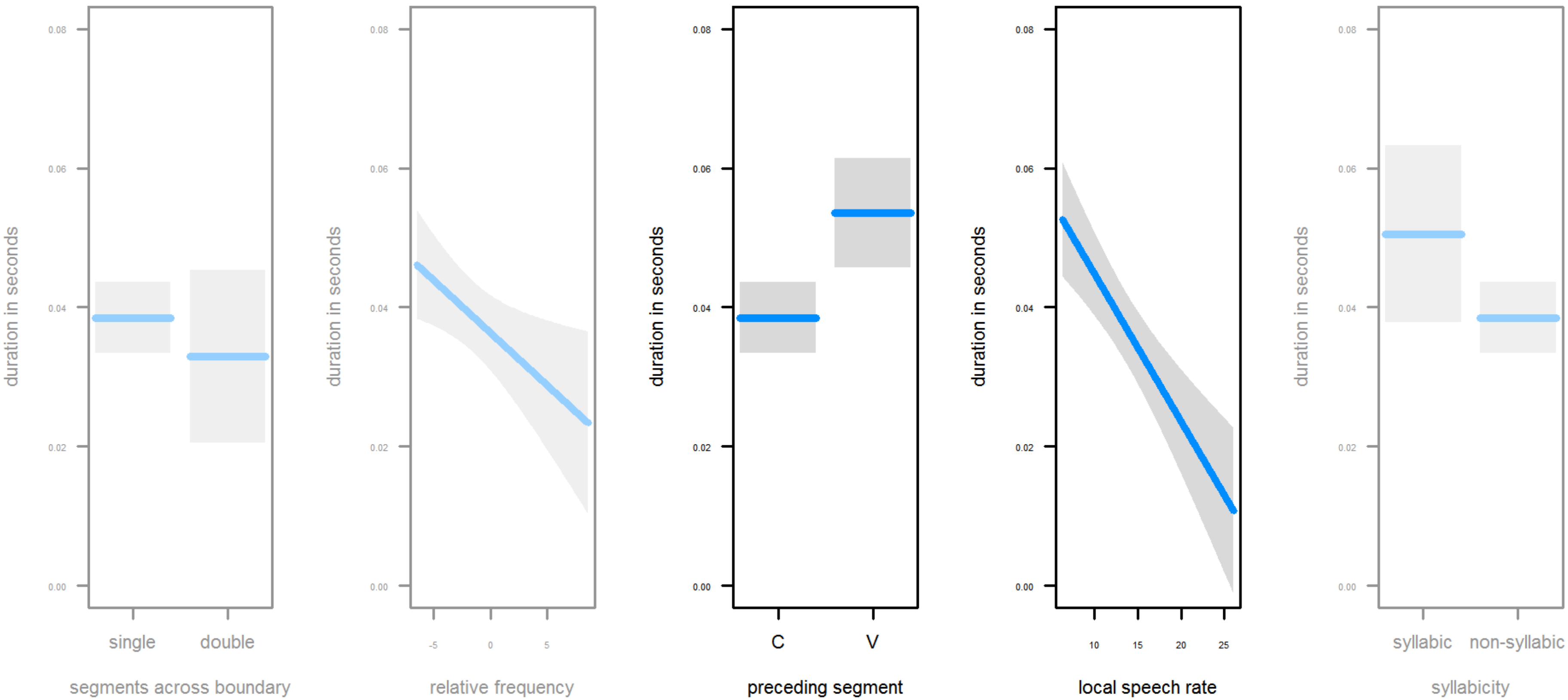
Results 2: *in-* geminates



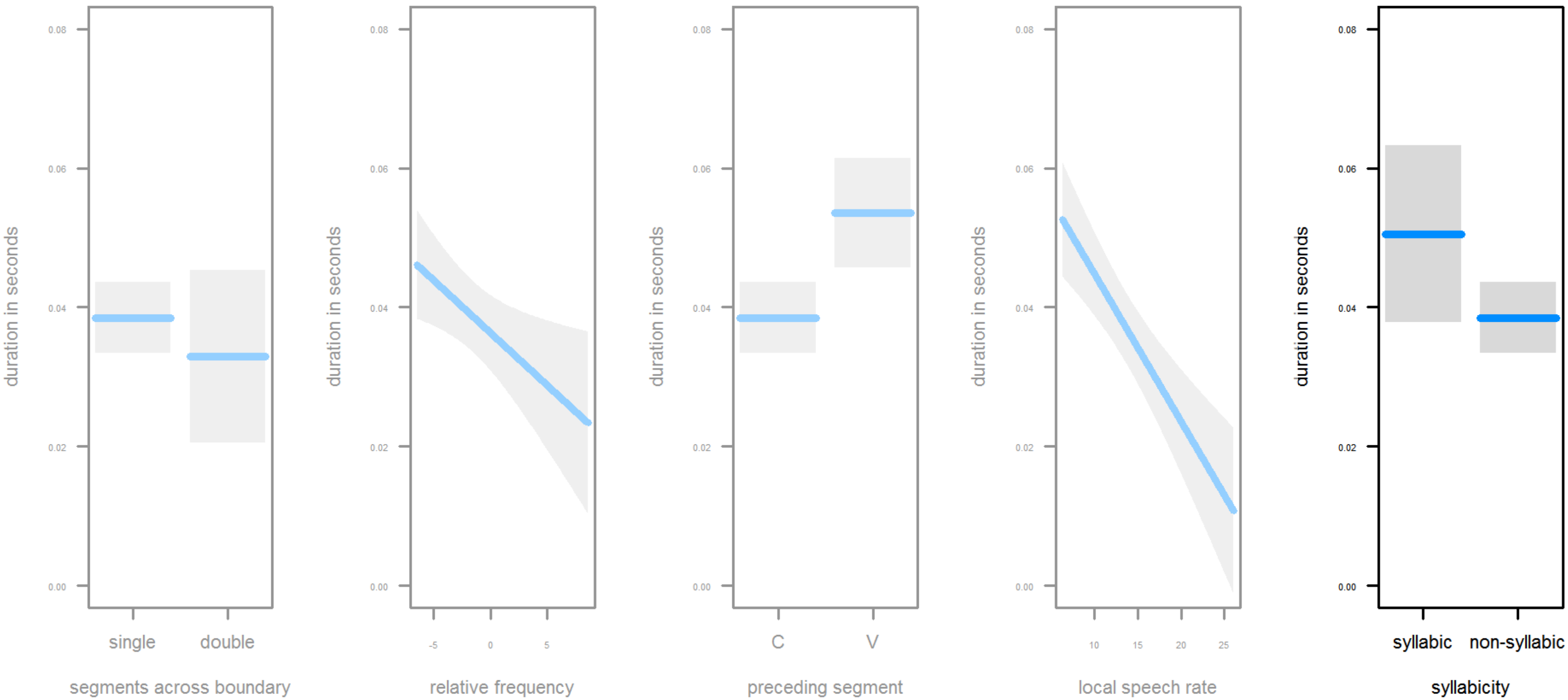
Results 2: *in-* geminates



Results 3: *-ly* does not geminate

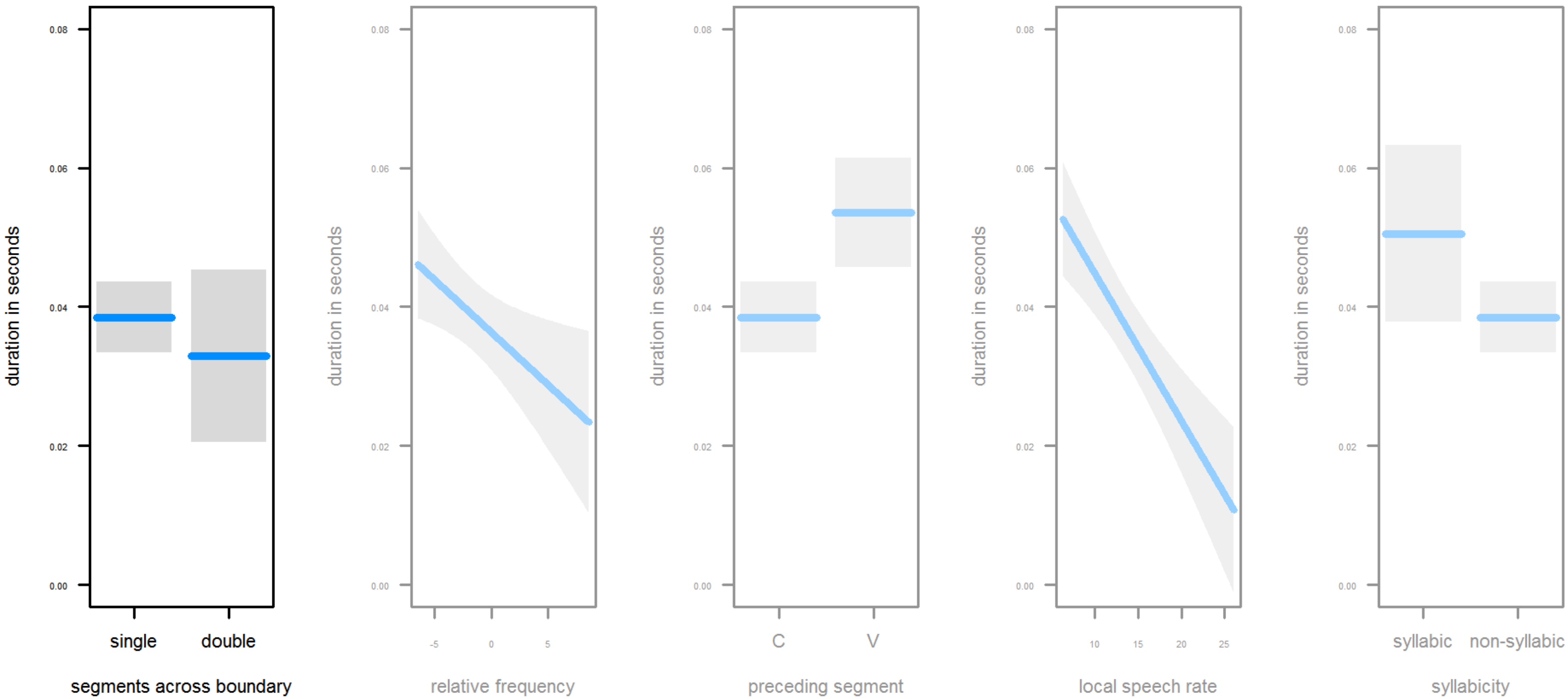


Results 3: *-ly* does not geminate

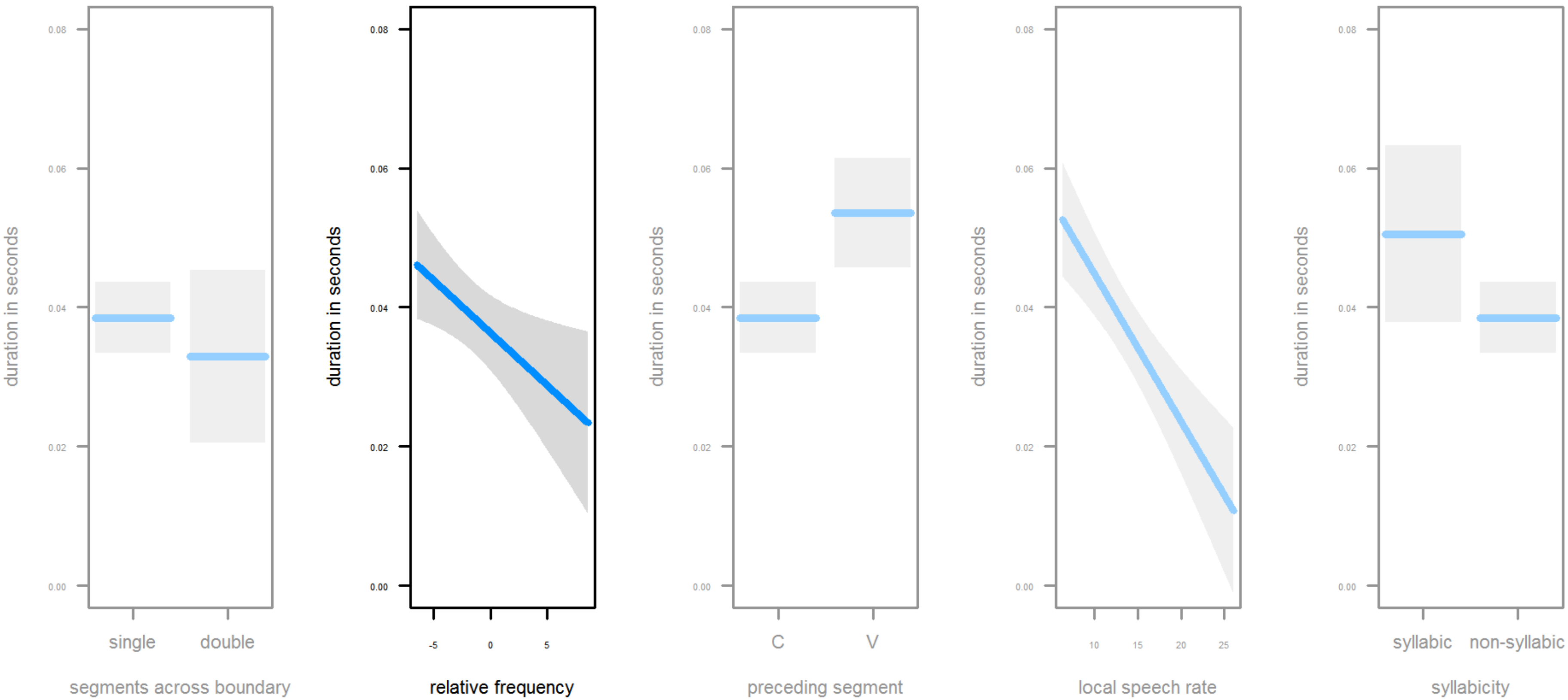


Additional covariate: Syllabicity (ment[l̩]y vs. ment[əl]y, odd[l]y)

Results 3: *-ly* does not geminate



Results 3: *-ly* does not geminate



Summary

- *un-* geminates: no surprise
- *in-* geminates: unexpected result
  - Effect of AFFIX: homophonous locative and negative *in-* prefixes are acoustically different
- *-ly* degeminates: unexpected result
  - 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
  - *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

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
```

im-model

```
#   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

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