Gemination and Degemination in English Affixation: a phonetic investigation of *dis-* and *–ly*

Sonia Ben Hedia

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

LabPhon 15, 13– 17 July 2016
Satellite Event: Reduction

(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*
dis-	*dis-satisfied*
-ly	*sole-ly, technical-ly*

- Phonetic correlates
  - Gemination: Longer duration than a singleton
  - Degemination: Same duration as a singleton
- Theoretical assumption: Degemination is affix- or stratum-dependent

Predictions: Lexical Phonology

	Level 1	Level 2
Morphological Process	in + numerous dis+ satisfied	
Phonological Process	i/n/umerous di/s/atisfied	
Phonetic Outcome	i[n]umerous di[s]atisfied	

Degemination

Predictions: Lexical Phonology

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

Degemination **Gemination**

Predictions: Morphological Separability

- Phonetic realization is dependent on morphological separability (e.g. Hay 2003, Smith et al. 2012)
- more separable → less reduction

> More separable complex words geminate.
>
> Less separable complex words degeminate.

- Separability:
  - Semantic Transparency: opaque vs. transparent
  - Type of Base: bound root vs. word
  - Relative Frequency: relative frequency of base and derivative

Empirical evidence?

- Three studies empirically investigated *in-* and *un-* in English
  - *un-* and *in-* geminate (Oh and Redford 2013, Kaye 2005, Ben Hedia & Plag 2015)

- No empirical study of *dis-* and *–ly*

This study

- Sample of *dis-* and *-ly*-affixed words with a double or a single consonant at the morphological boundary

- Data: Natural conversational speech from the Switchboard Corpus (Godfrey & Holliman 1997)

- Manual segmentation and acoustic measurements in Praat (Boersma & Weenink 2014)

Statistical Analsysis

- Multiple regression with **duration** as dependent variable and **environment** as predictor

dis-

Single		Double
s#C *(disfavor)*	s#V *(disambiguate)*	s#s *(dissatisfied)*

-ly

Single	Double	
single *(randomly)*	double *(really)*	syllabic-double *(mentally)*

Statistical Analsysis

- Multiple regression with **duration** as dependent variable and **environment** as predictor

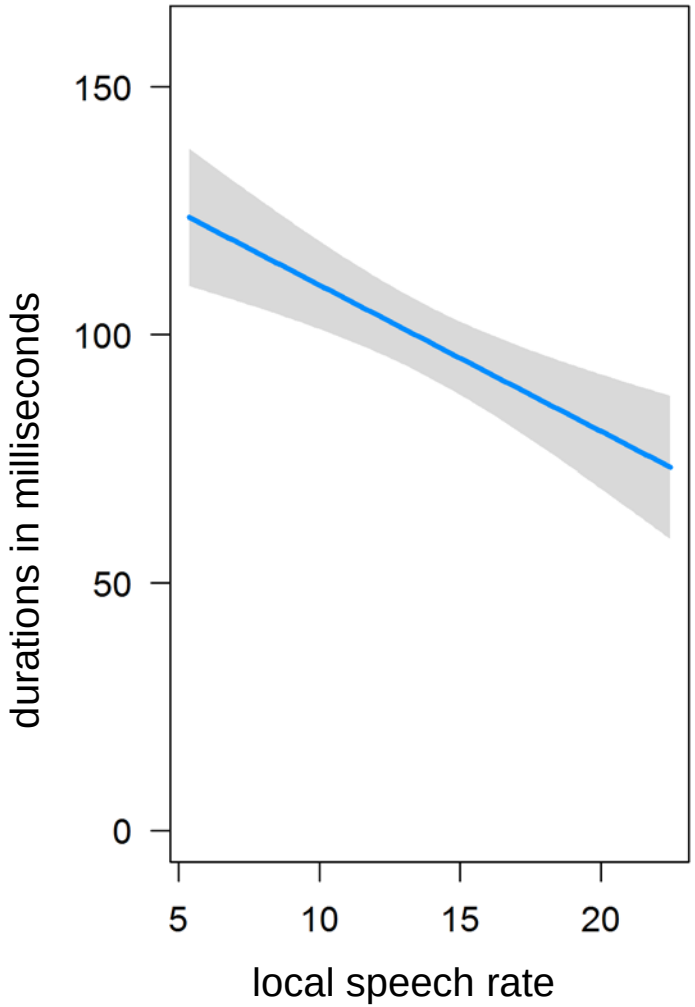
- Coding of pertinent covariates:
  - Preceding Segment Duration
  - Preceding Segment
  - Following segment
  - Speech Rate
  - Stress
  - Word Form Frequency
  - Relative Frequency
  - Semantic Transparency

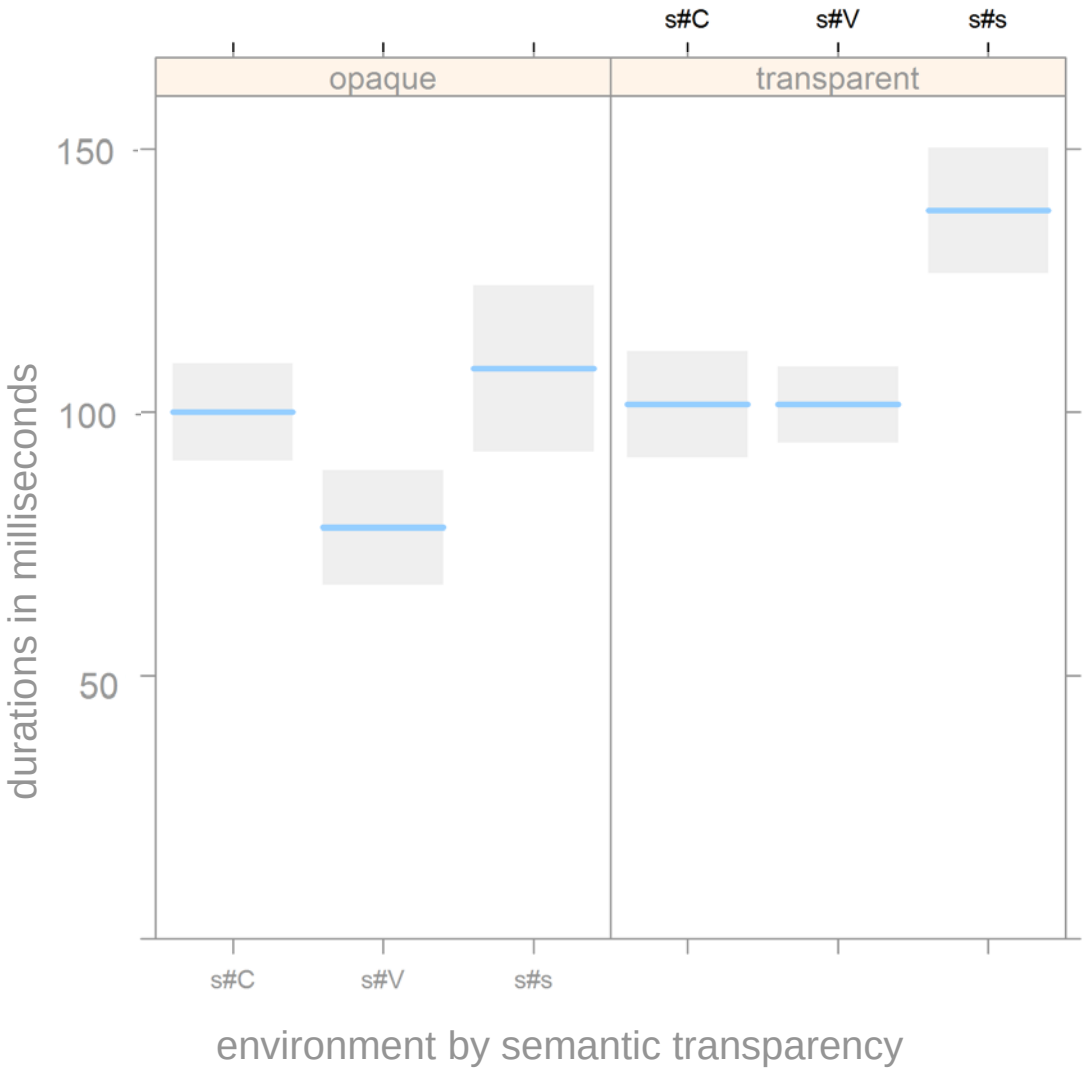
Statistical Analsysis

- Multiple regression with **duration** as dependent variable and **environment** as predictor

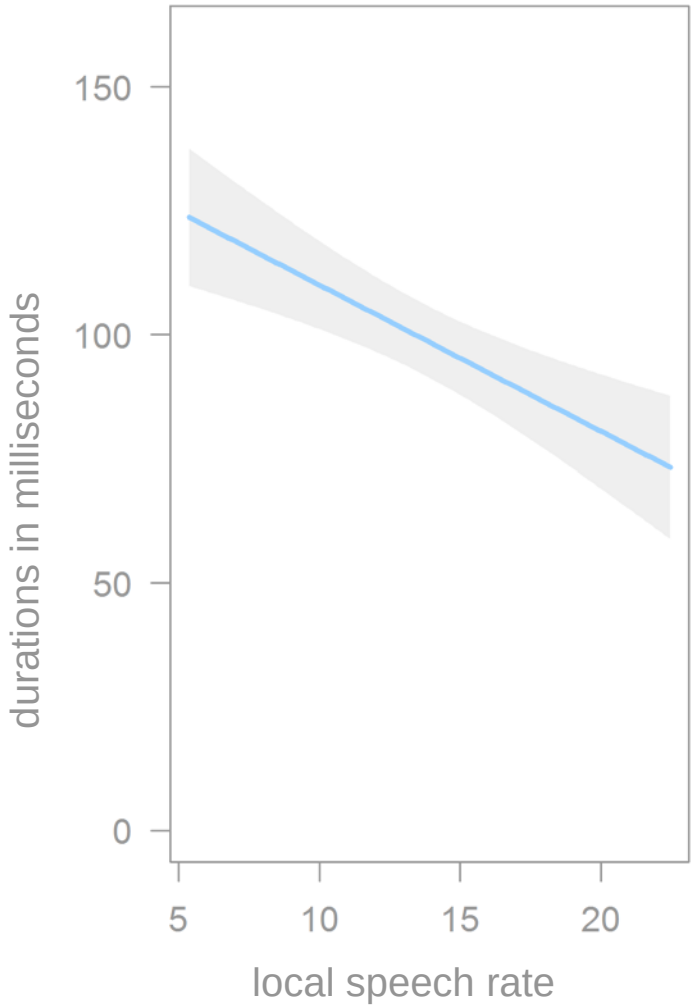
- Coding of pertinent covariates:
  - Preceding Segment Duration
  - Preceding Segment
  - Following segment
  - Speech Rate
  - Stress
  - Word Form Frequency
  - Relative Frequency
  - Semantic Transparency

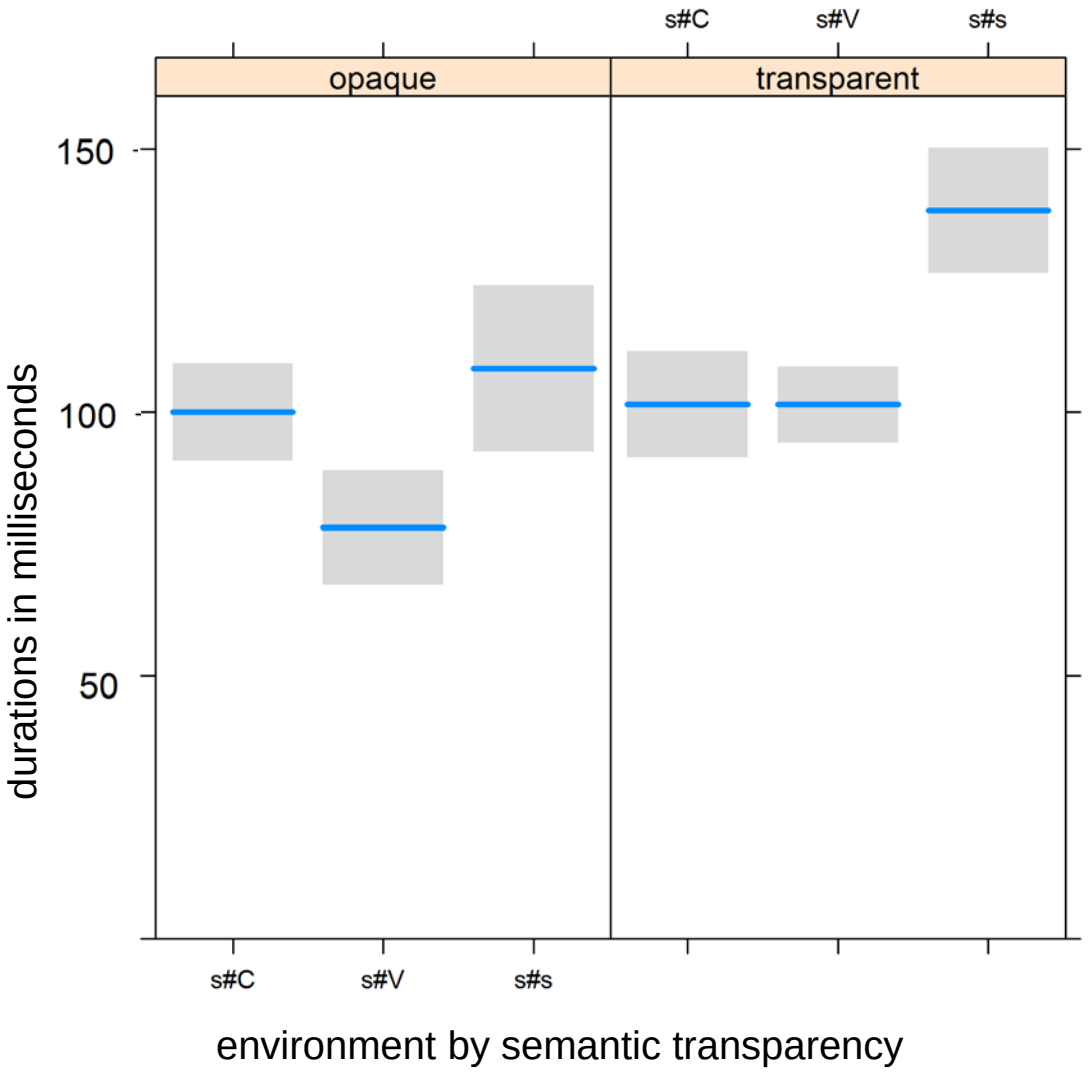
Results 1: *dis-*





Results 1: *dis-*

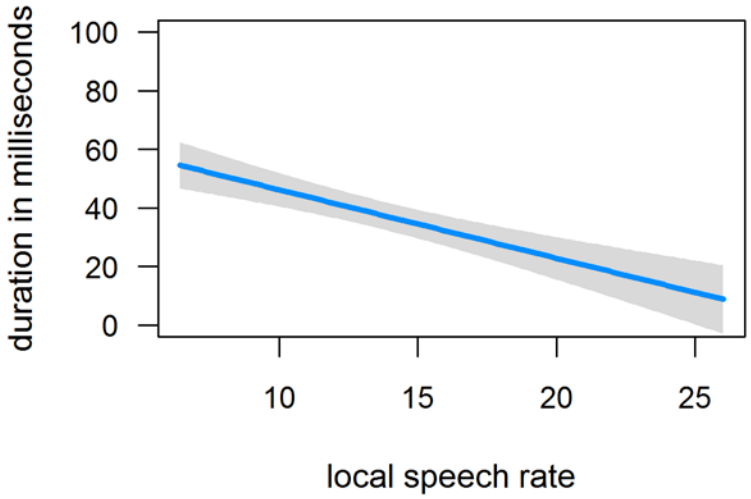


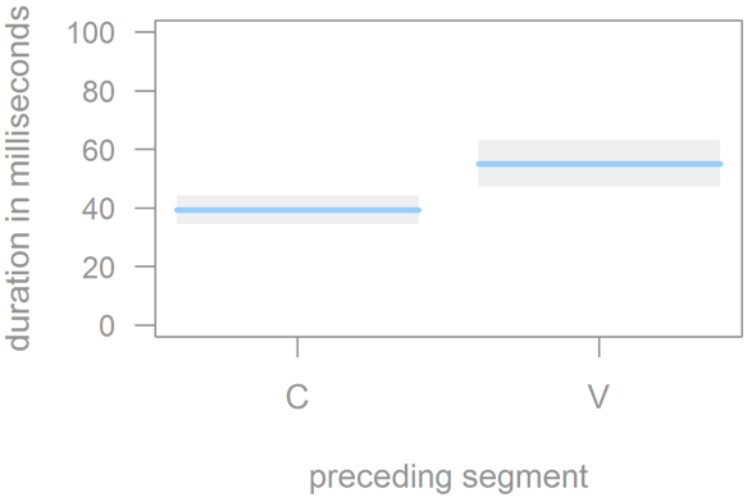


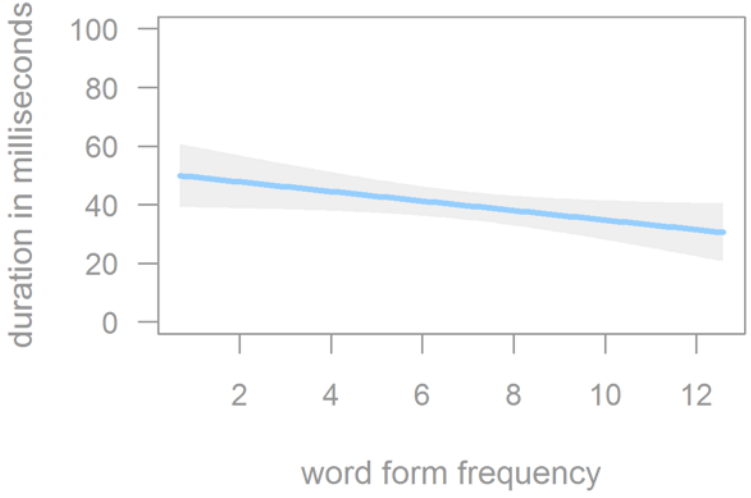
Summary: dis

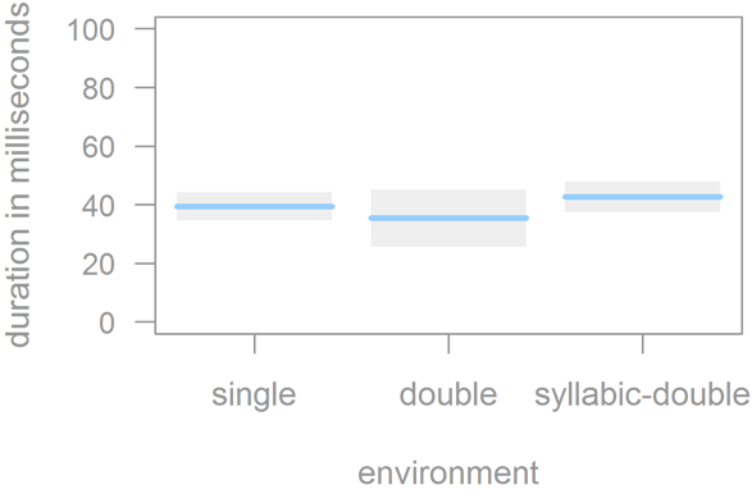
- *dis-* can geminate:
  - Interaction between Semantic Transparency and environment:
    - For transparent items *dis-* clearly geminates
    - For opaque items the /s/ in s#s is longer than in s#V but not longer than in s#C-structures

Results 4: *-ly* does not geminate

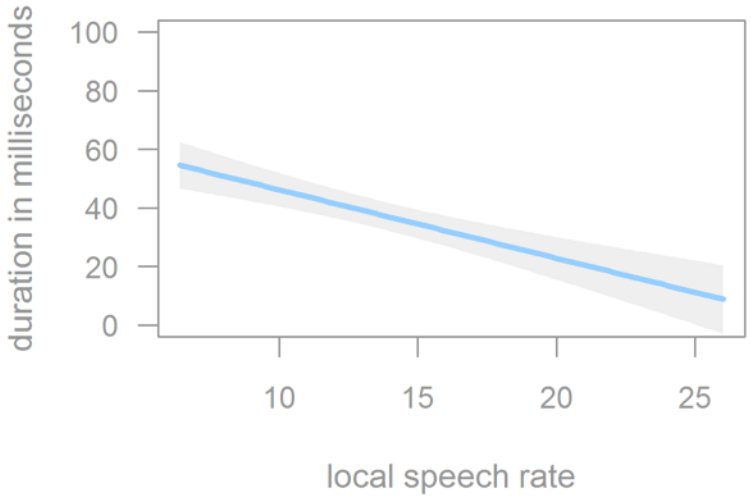


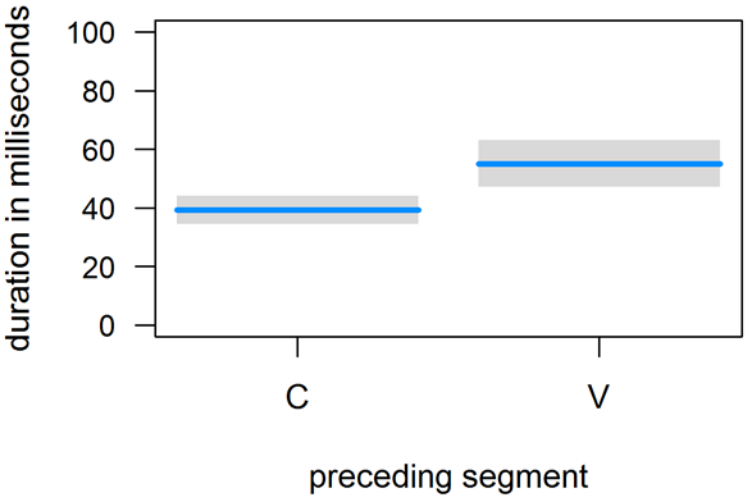


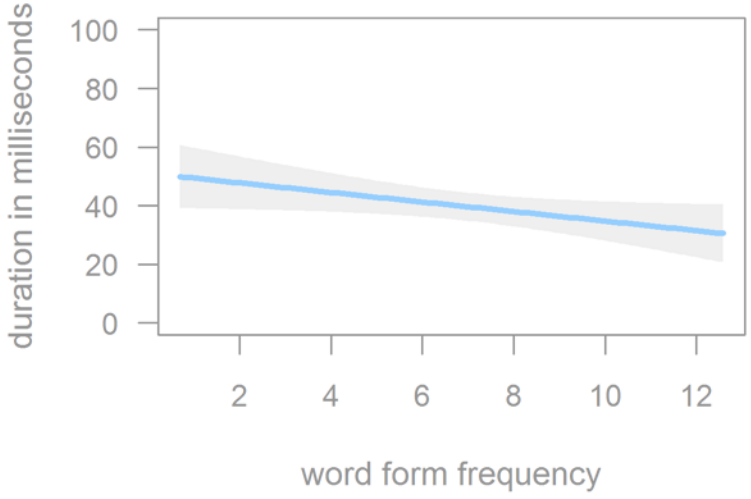


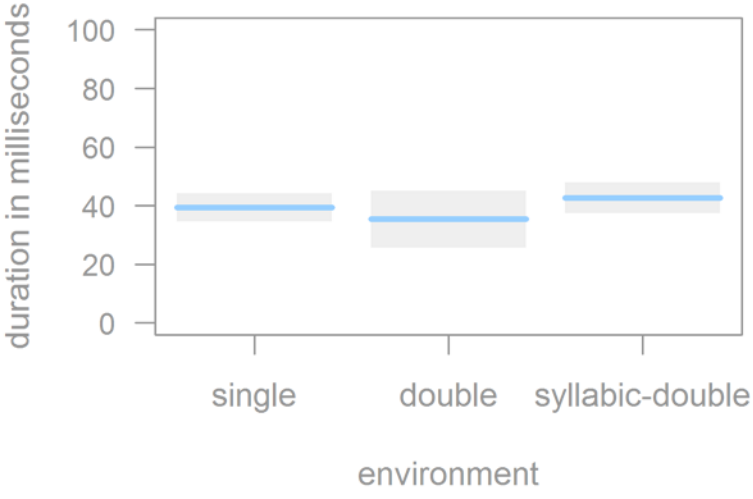


Results 4: *-ly* does not geminate

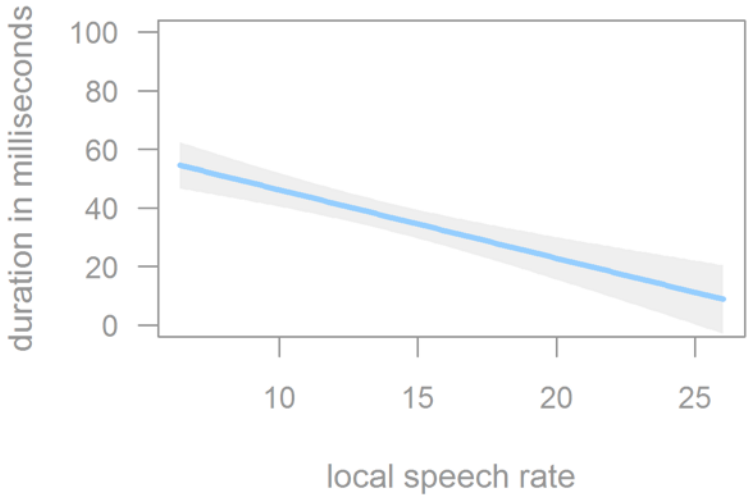


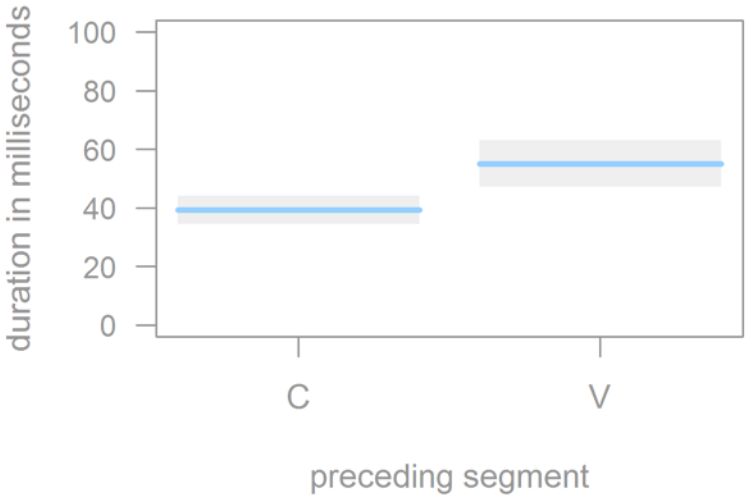


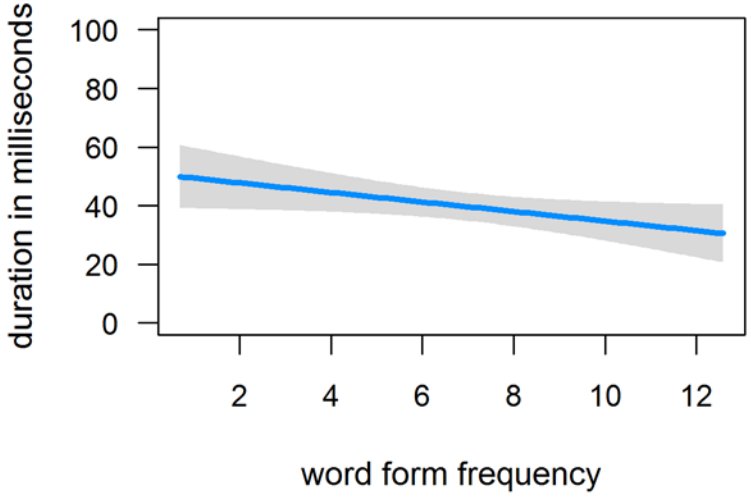


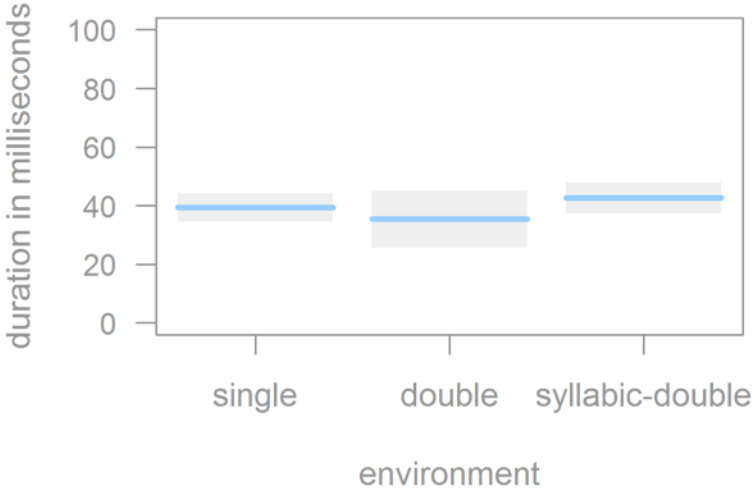


Results 4: *-ly* does not geminate

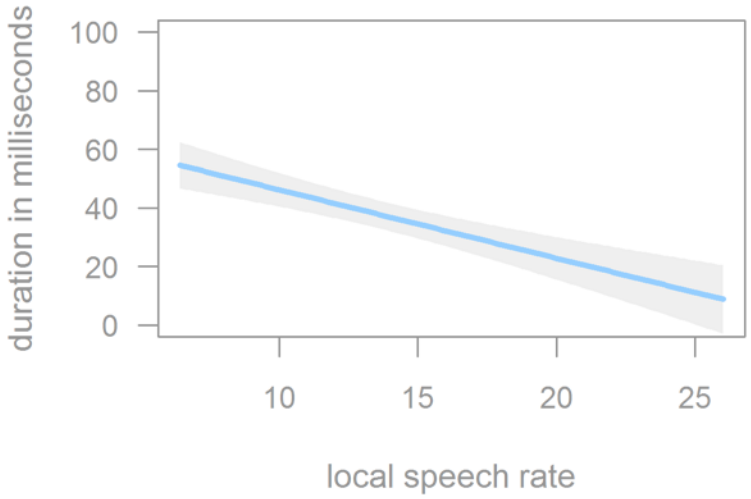


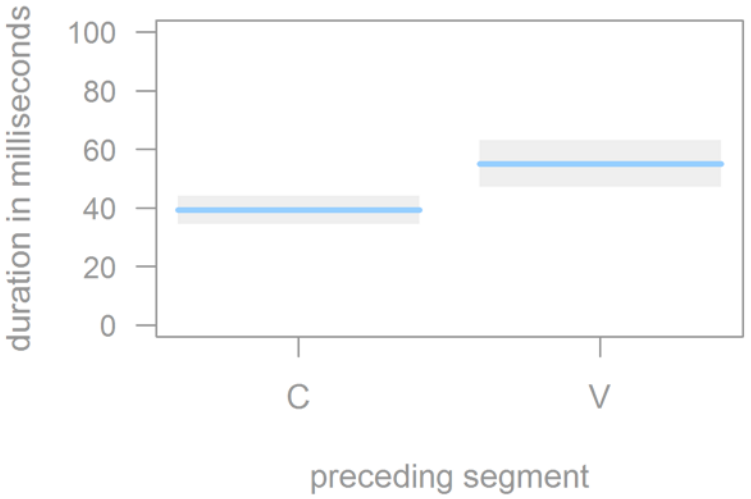


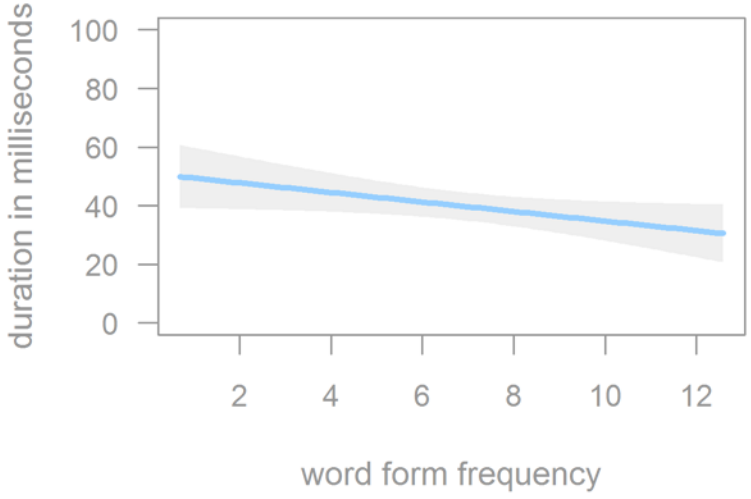


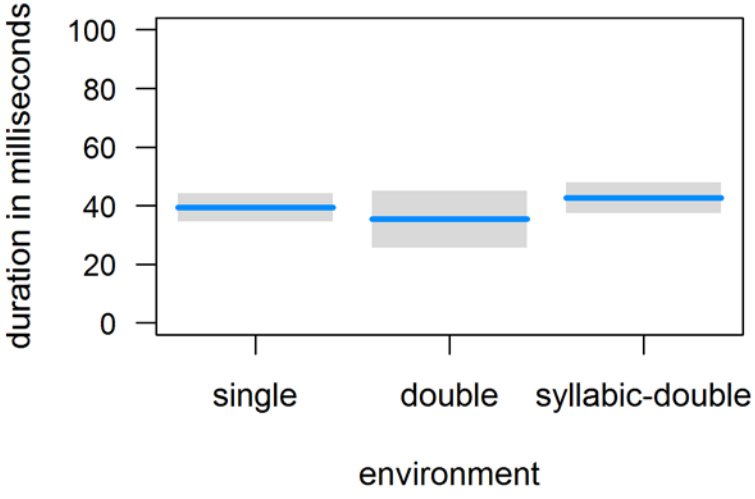


Results 4: *-ly* does not geminate









Implications

- Lexical Phonology makes wrong empirical predictions
- Gemination is not stratum-dependent, but affix-specific
- Morphological information is directly reflected in the speech signal
  - *dis-*
    Morphological separability is reflected in duration (cf. Hay 2007, Collie 2008, Ben Hedia & Plag 2016 on *in-*)
- 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!

Funding
Deutsche Forschungsgemeinschaft:
Forschergruppe 2372
- Grant PL151/8-1 ‘Morpho-phonetic Variation in English‘

References

Bauer, Laurie; Lieber, Rochelle; Plag, Ingo (2013): The Oxford reference guide to English morphology. Oxford: Oxford University Press (Oxford Linguistics).

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

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

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, Alan S. (2005): Gemination in English. In: *English Today* 21 (2), S. 43-55.

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.

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *The Behavioral and brain sciences, 22*(1), 1-38; discussion 38-75.

References

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, Ingo; Homann, Julia; Kunter, Gero (2015): Homophony and morphology: The acoustics of word-final S in English. In: *Journal of Linguistics.* Online verfügbar unter http://journals.cambridge.org/article_S0022226715000183.

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

```
lm(formula = bc ~ TransitionType + LocSpeech, data = unComplex2)

 Residuals:
   Min       1Q   Median       3Q      Max
-0.084297 -0.025824  0.000047  0.025345  0.114253

Coefficients:
                          Estimate    Std. Error   t value   Pr(>|t|)
(Intercept)                0.531198   0.015457  34.366    < 2e-16 ***
TransitionTypen#V         -0.047212   0.006890  -6.852   1.70e-10 ***
TransitionTypen#nV         0.049706   0.009800   5.072   1.13e-06 ***
LocSpeech                 -0.007540   0.001106  -6.814  2.08e-10 ***
  ---
  Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.03959 on 152 degrees of freedom
Multiple R-squared:  0.5706,              Adjusted R-squared:  0.5621
F-statistic: 67.33 on 3 and 152 DF,  p-value: < 2.2e-16
```

im-model

lm(formula = bc ~ NoCons + LocSpeech + StressPattern + Affix, data = imComplex3)

Residuals:

Min	1Q	Median	3Q	Max
-0.090887	-0.023970	-0.001624	0.024476	0.081057

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	0.3172056	0.0121191	26.174	< 2e-16	***
NoConsm#mV	0.0464675	0.0069756	6.661	4.75e-10	***
LocSpeech	-0.0034325	0.0007938	-4.324	2.77e-05	***
StressPatternbeforeUnstressed	-0.0355165	0.0076431	-4.647	7.29e-06	***
AffixinNeg	0.0204865	0.0074717	2.742	0.00685	**

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.03493 on 151 degrees of freedom
Multiple R-squared: 0.5263, Adjusted R-squared: 0.5137
F-statistic: 41.94 on 4 and 151 DF, p-value: < 2.2e-16

dis-model

lm(formula = AbsDur ~ TransitionType * SemanticTransparency + LocSpeech, data = dis_2)

Residuals:

Min	1Q	Median	3Q	Max
-0.045423	-0.015859	0.000122	0.015080	0.055498

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.1381778	0.0111967	12.341	< 2e-16 ***
TransitionTypes#V	-0.0218816	0.0074598	-2.933	0.00403 **
TransitionTypes#s	0.0082414	0.0092700	0.889	0.37579
SemanticTransparencytransparent	0.0014439	0.0068300	0.211	0.83294
LocSpeech	-0.0029549	0.0007211	-4.098	7.68e-05 ***
TransitionTypes#V:SemanticTransparencytransparent	0.0218454	0.0096749	2.258	0.02579 *
TransitionTypes#s:SemanticTransparencytransparent	0.0286366	0.0121318	2.360	0.01989 *

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.02254 on 118 degrees of freedom
Multiple R-squared: 0.3454, Adjusted R-squared: 0.3121
F-statistic: 10.38 on 6 and 118 DF, p-value: 3.239e-09

-*ly*-model

lm(formula = AbsDurCon ~ TransitionType + PrecSegVC + logWordFormFreqAllCoca + ´LocSpeech, data = lyComplex3)

Residuals:

Min	1Q	Median	3Q	Max
-0.039228	-0.013135	-0.000586	0.012467	0.044022

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	0.0812883	0.0078464	10.360	< 2e-16	***
TransitionTypedouble	-0.0039653	0.0047610	-0.833	0.406284	
TransitionTypesyllabic-double	0.0034212	0.0035686	0.959	0.339312	
PrecSegVCV	0.0156936	0.0044286	3.544	0.000531	***
logWordFormFreqAllCoca	-0.0016310	0.0007767	-2.100	0.037461	*
LocSpeech	-0.0023351	0.0004379	-5.332	3.65e-07	***

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.01822 on 145 degrees of freedom
Multiple R-squared: 0.2615, Adjusted R-squared: 0.2361
F-statistic: 10.27 on 5 and 145 DF, p-value: 1.918e-08

Overview of the data

	Double Consonant	Single Consonant	Total per affix
dis-	24	104	128
-ly	81	73	154

Overview of the data (types)

	Double Consonant	Single Consonant	Total per affix
un-	6	95	101
in-	16	67	83
dis-	9	55	64
-ly	77	73	150