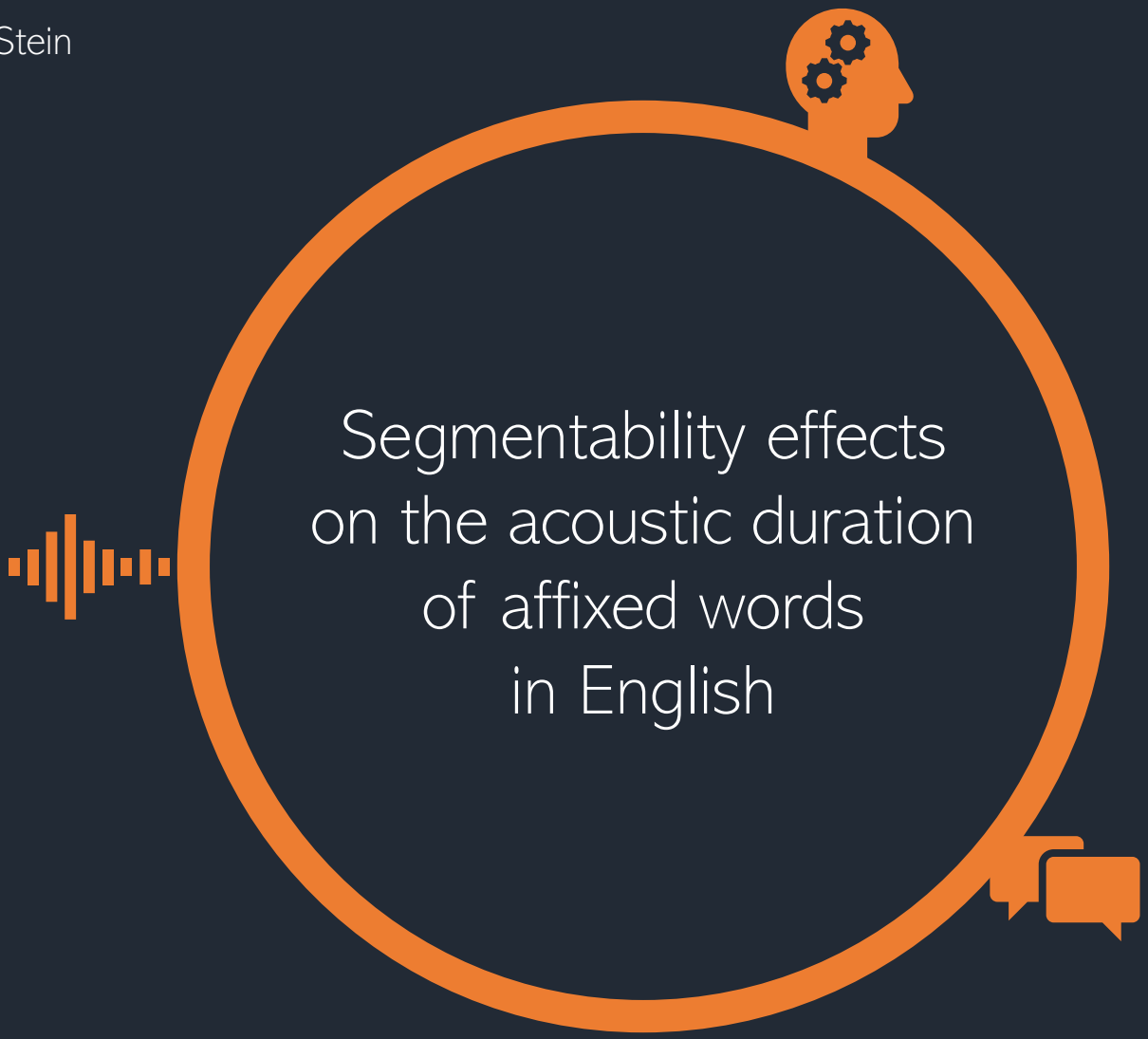
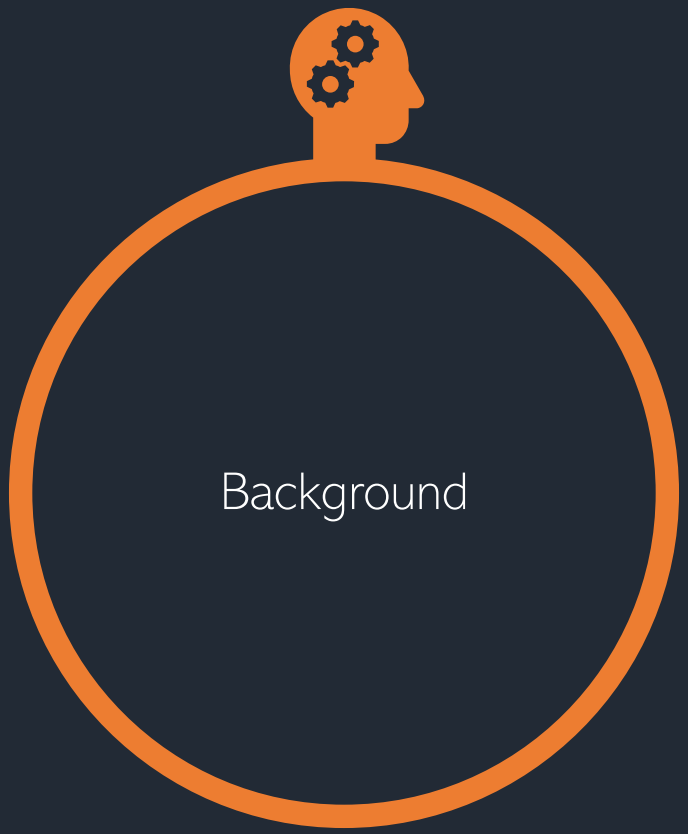
Interfaces of Phonetics, May 18–19, 2021

Simon David Stein
Ingo Plag

hhu Heinrich Heine Universität Düsseldorf
DFG Deutsche Forschungsgemeinschaft
FOR2373



Segmentability effects on the acoustic duration of affixed words in English


Background

Morphological segmentability

Morphological segmentability

the degree to which speakers can decompose a complex word into its constituents

Hay 2001, 2003, 2007

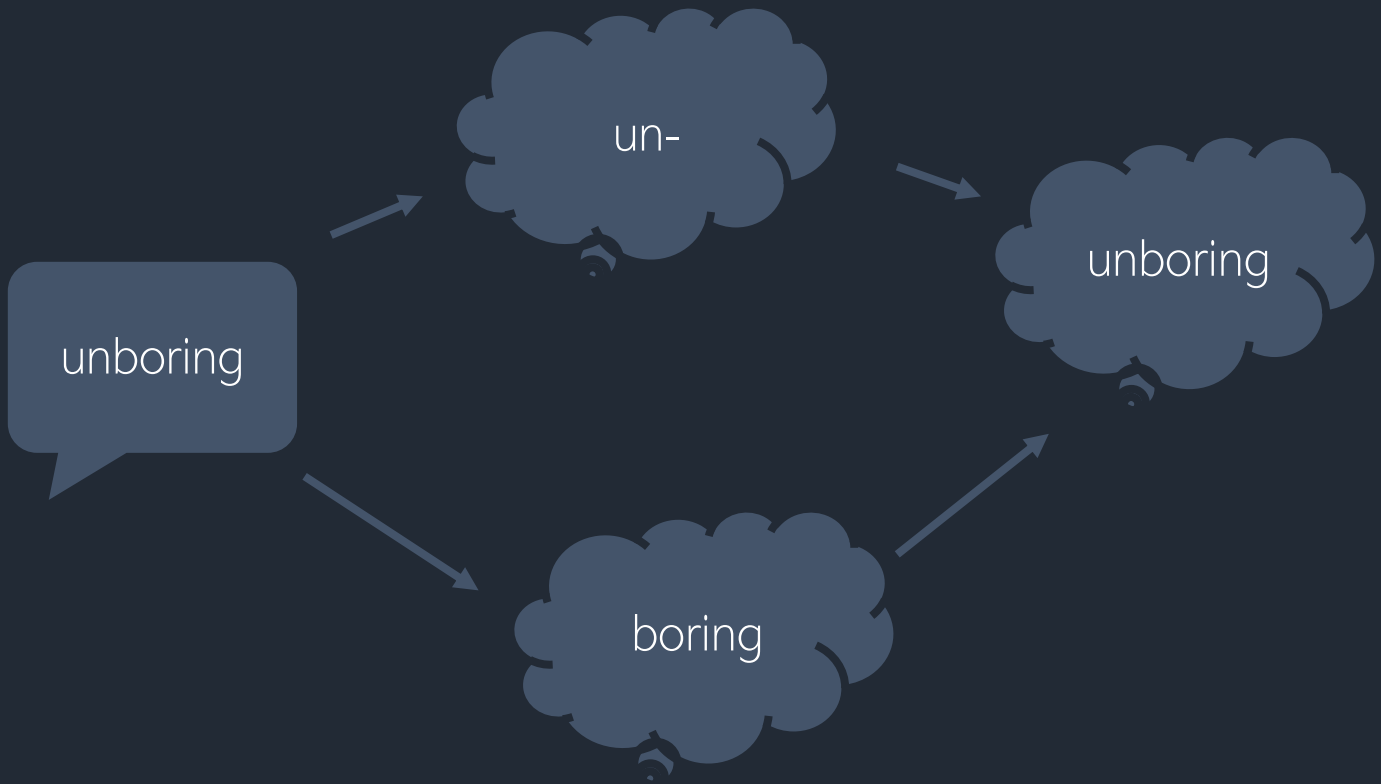
Morphological segmentability

the degree to which speakers can decompose a complex word into its constituents, operationalized as:

$$relative\ frequency = \frac{base\ frequency}{word\ frequency}$$

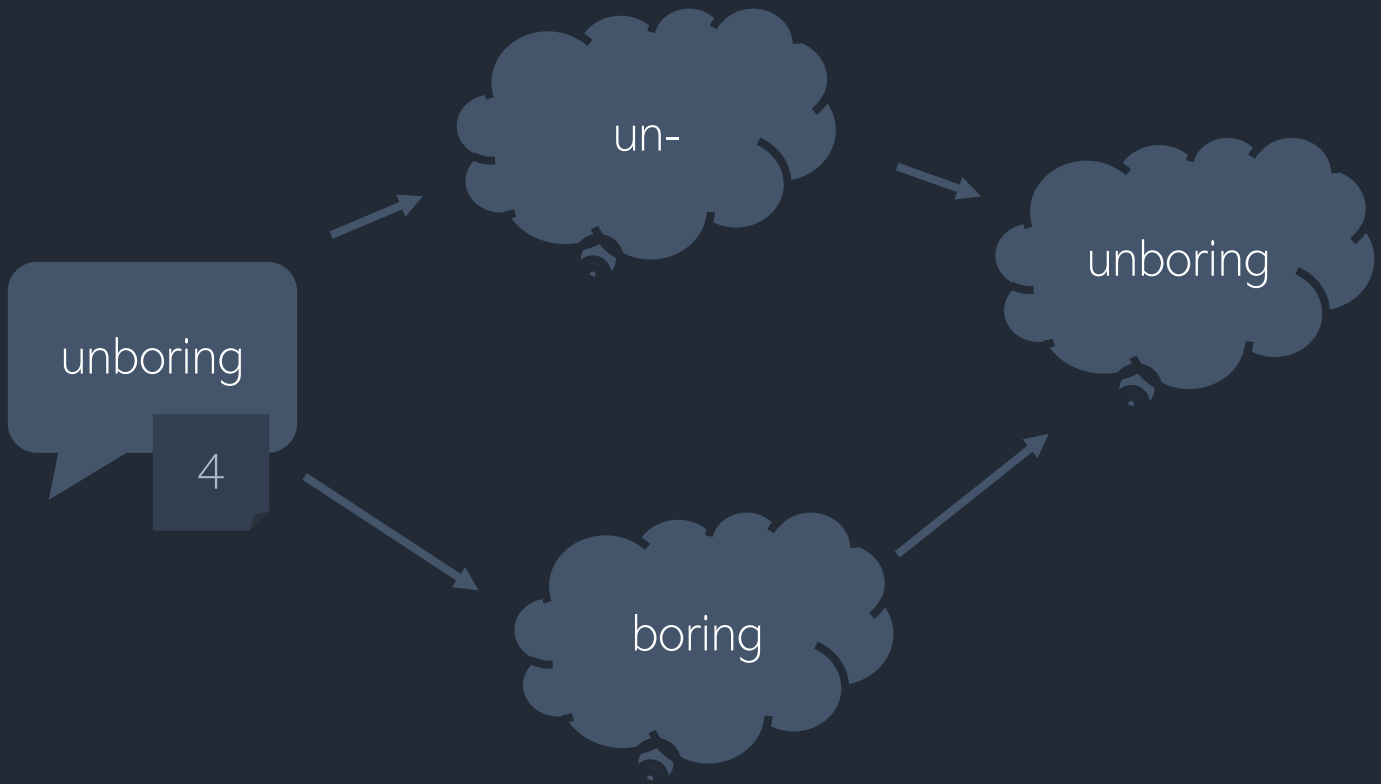
Hay 2001, 2003, 2007

Morphological segmentability



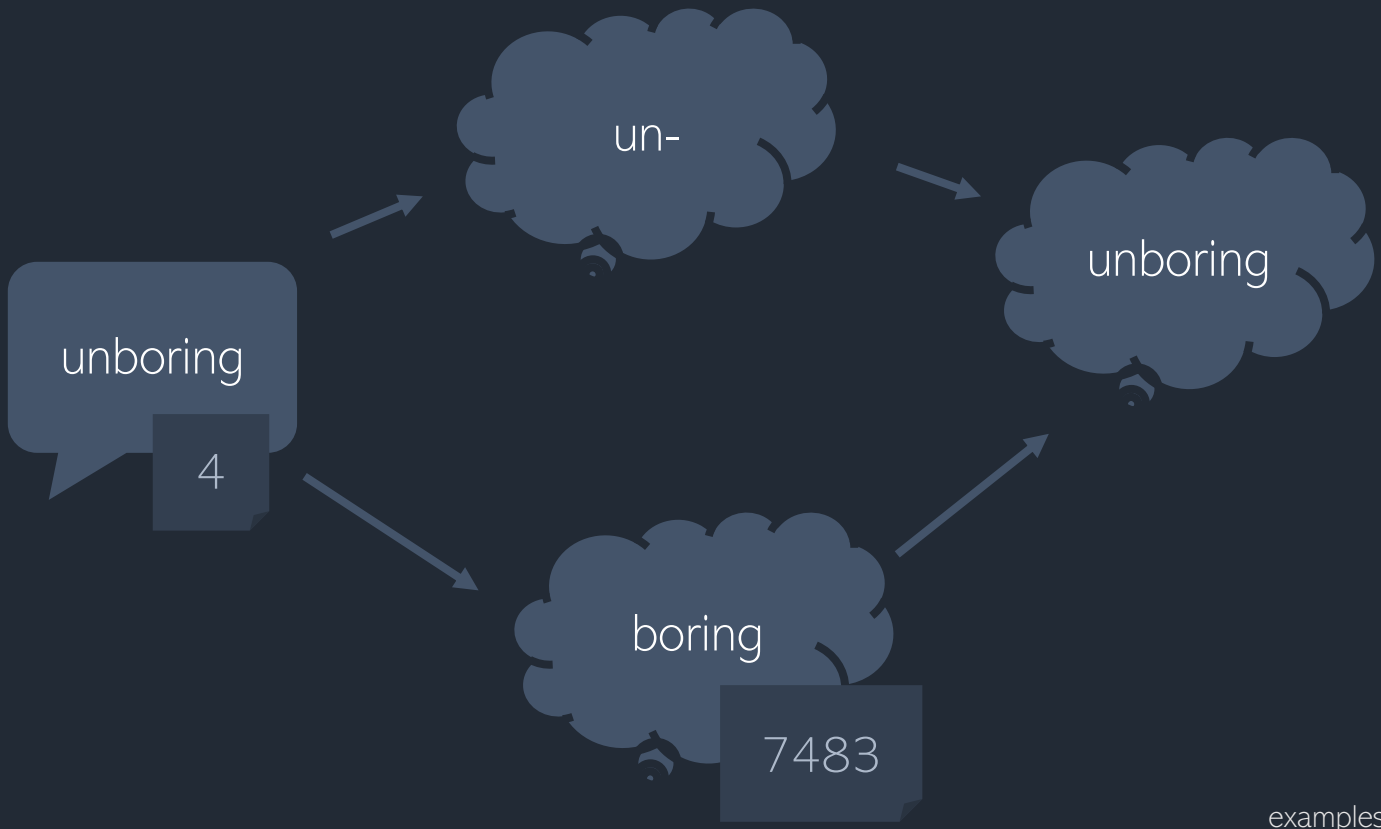
examples partly from Hay 2007
figure adapted from Hay 2001: 1045

Morphological segmentability



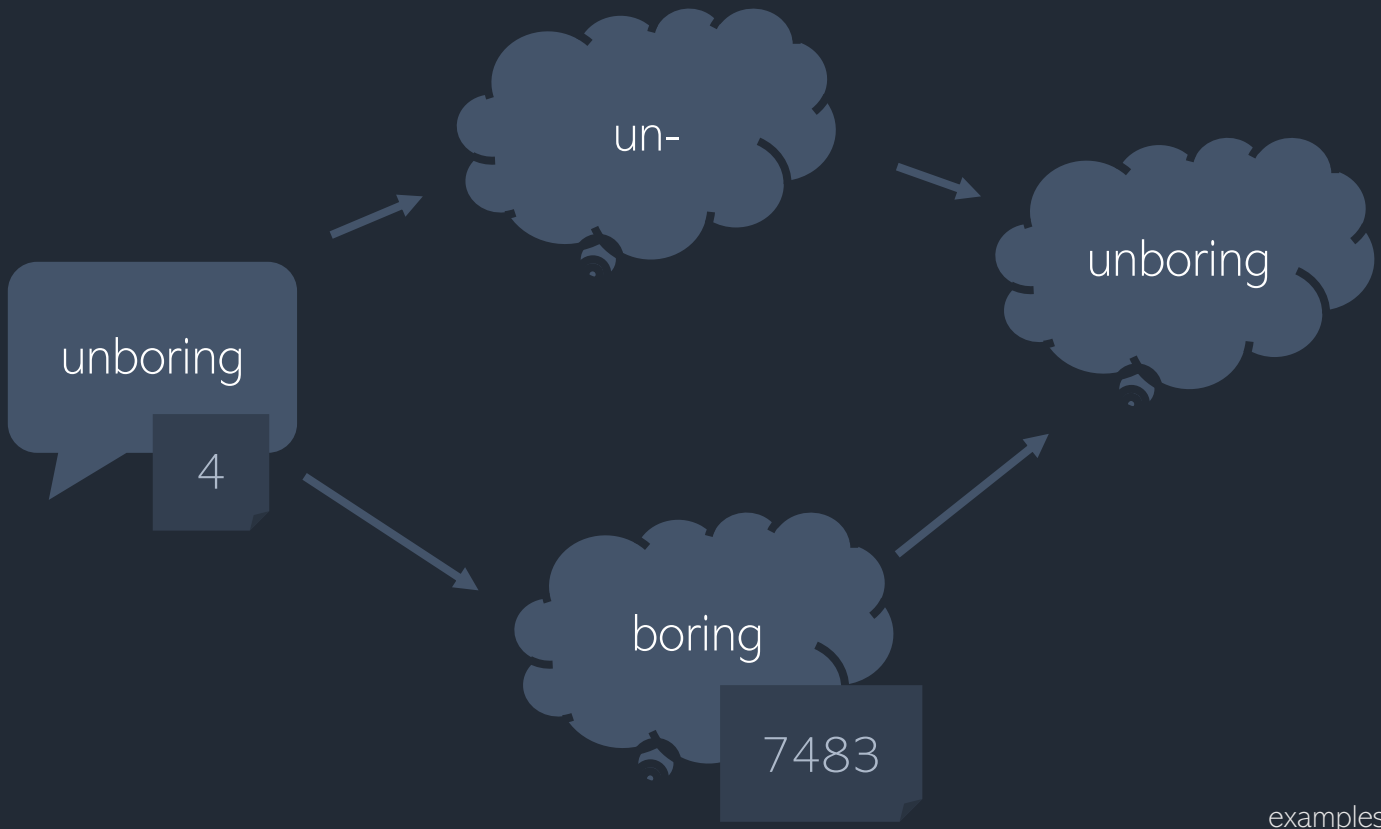
examples partly from Hay 2007
figure adapted from Hay 2001: 1045
frequencies taken from COCA, Davies 2008

Morphological segmentability



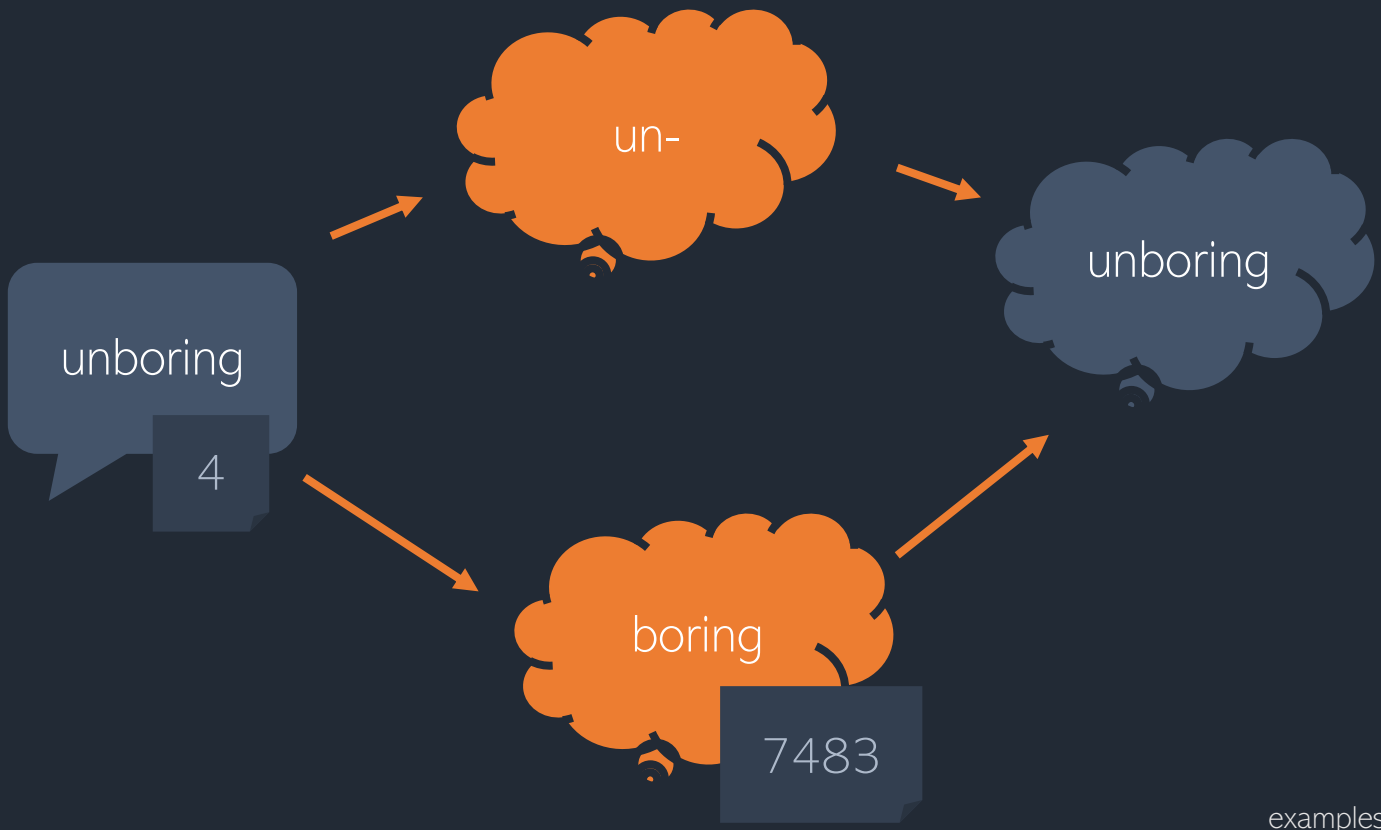
examples partly from Hay 2007
figure adapted from Hay 2001: 1045
frequencies taken from COCA, Davies 2008

Morphological segmentability
high



examples partly from Hay 2007
figure adapted from Hay 2001: 1045
frequencies taken from COCA, Davies 2008

Morphological segmentability
high



examples partly from Hay 2007
figure adapted from Hay 2001: 1045
frequencies taken from COCA, Davies 2008

Morphological segmentability

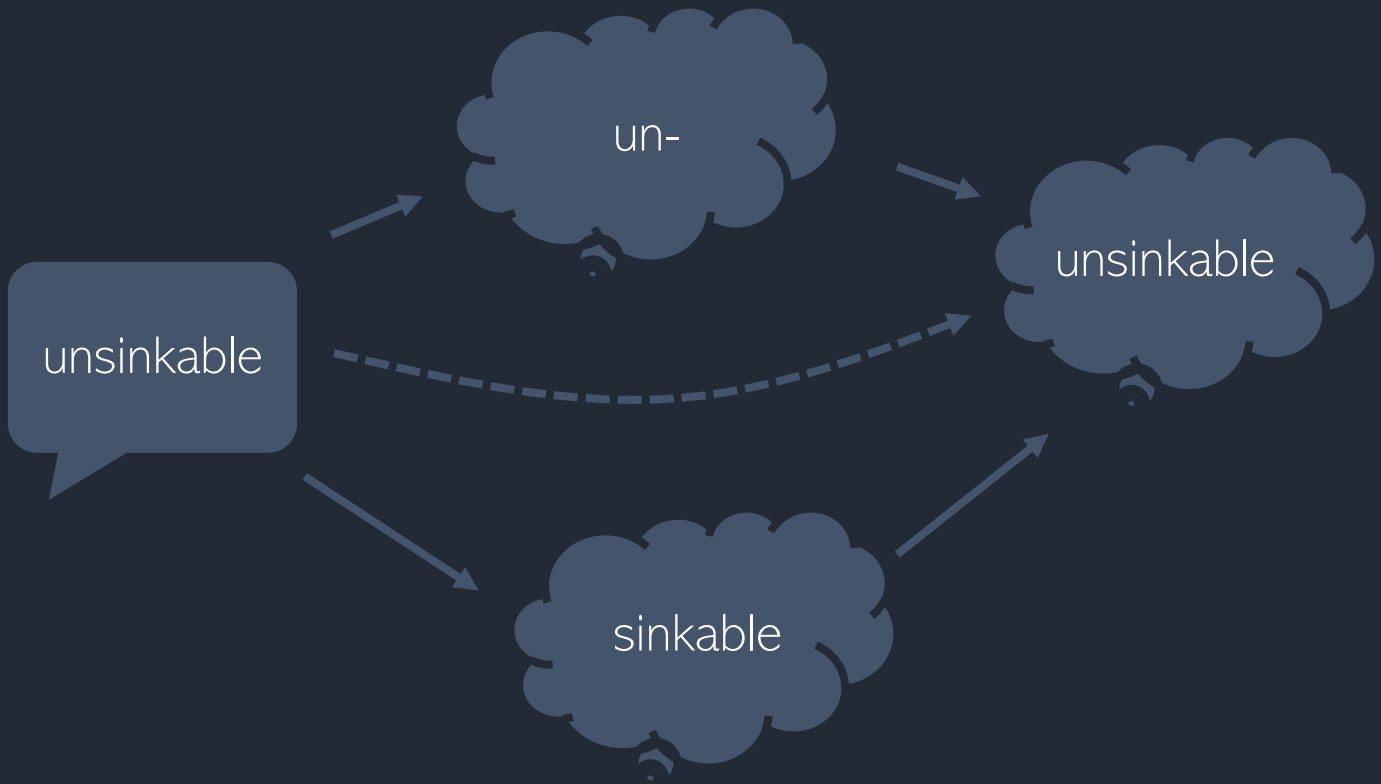


figure adapted from Hay 2001: 1045

Morphological segmentability

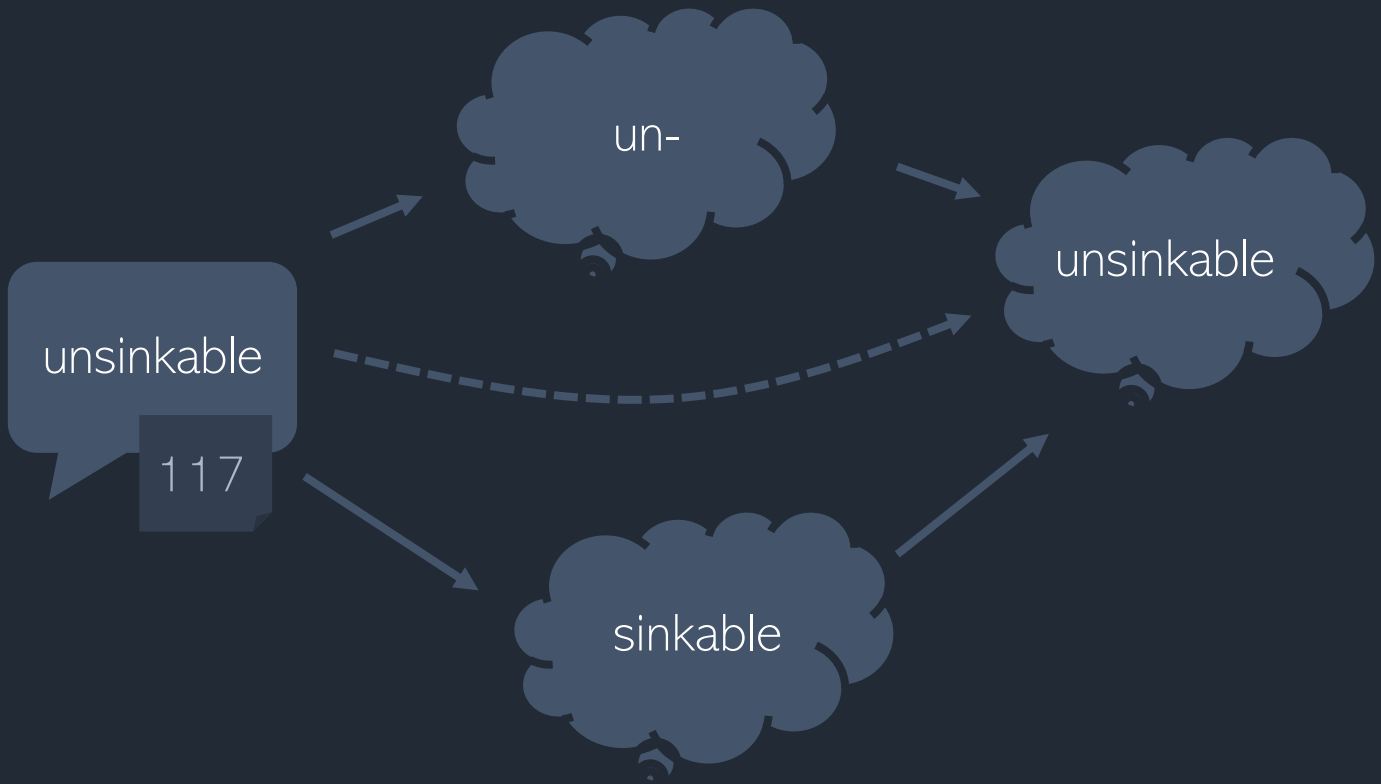


figure adapted from Hay 2001: 1045
frequencies taken from COCA, Davies 2008

Morphological segmentability

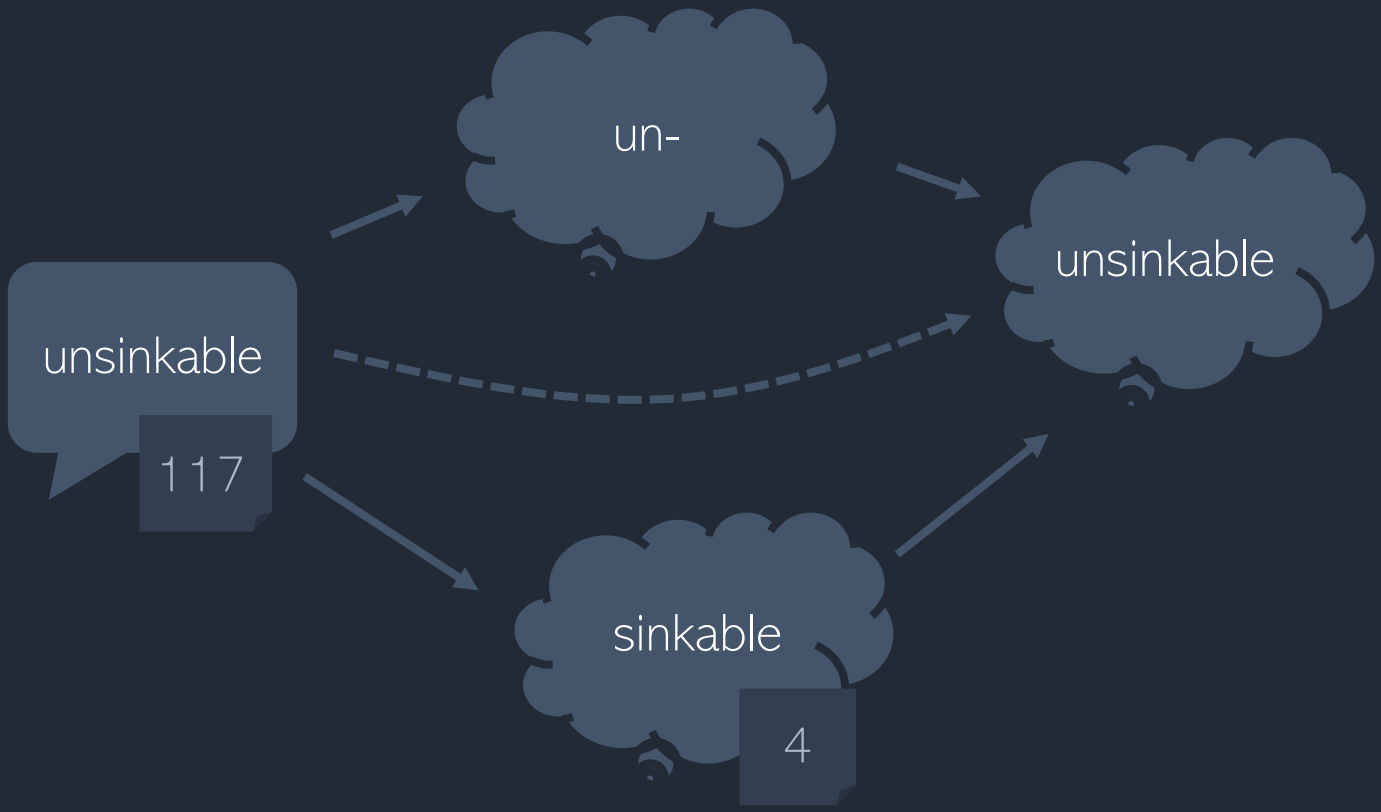


figure adapted from Hay 2001: 1045
frequencies taken from COCA, Davies 2008

Morphological segmentability
low

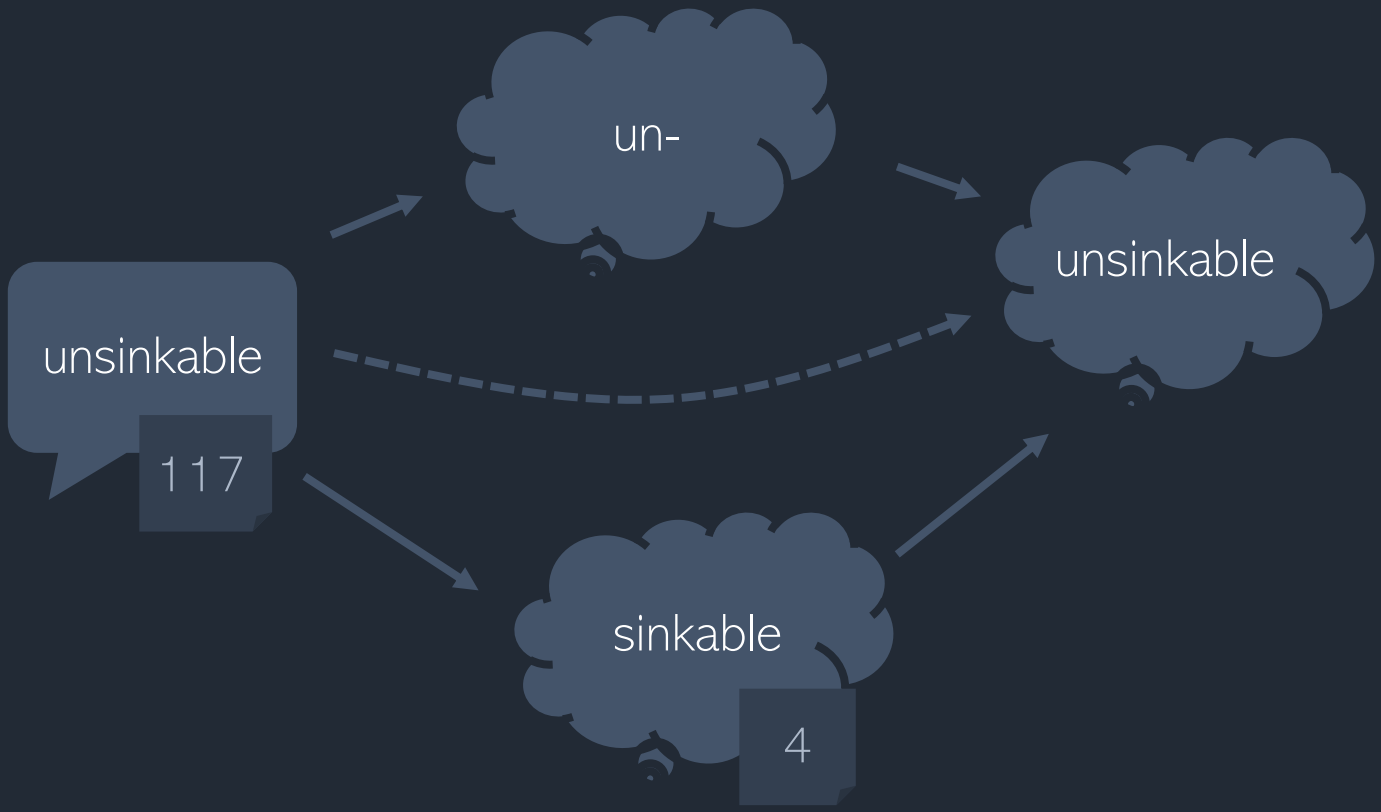


figure adapted from Hay 2001: 1045
frequencies taken from COCA, Davies 2008

Morphological segmentability
low

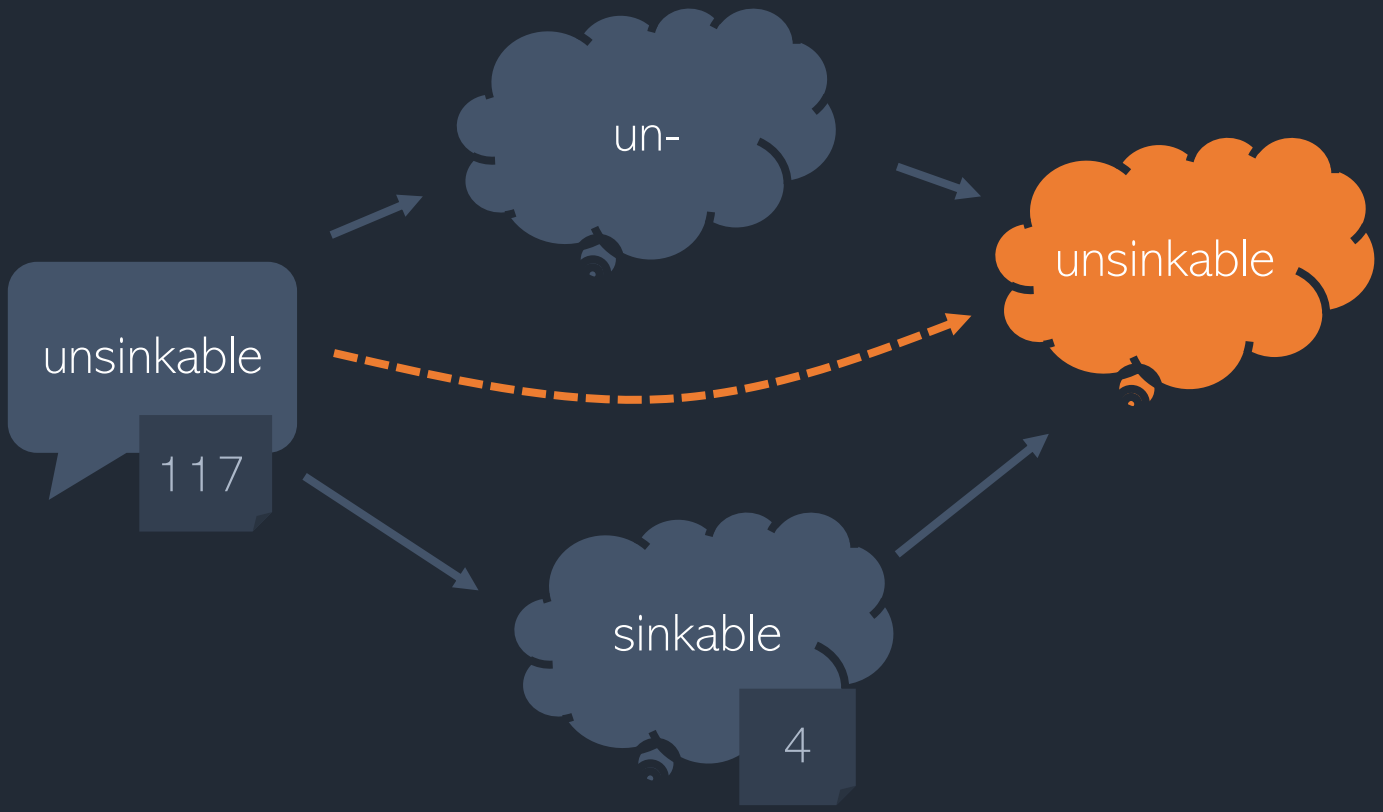


figure adapted from Hay 2001: 1045
frequencies taken from COCA, Davies 2008

The segmentability hypothesis

Hay 2001, 2003

The segmentability hypothesis

More segmentable words should be protected against reduction, i.e., longer in duration.

Hay 2001, 2003

Previous studies

Previous studies

Higher relative frequency has been found to be associated with:

Previous studies

Higher relative frequency has been found to be associated with:

no change in duration

Pluymaekers et al. 2005b
Plag & Ben Hedia 2018
Ben Hedia & Plag 2017
Zimmerer et al. 2014
Zuraw et al. 2020

Previous studies

Higher relative frequency has been found to be associated with:

no change in duration

Pluymaekers et al. 2005b
Plag & Ben Hedia 2018
Ben Hedia & Plag 2017
Zimmerer et al. 2014
Zuraw et al. 2020

longer durations

Plag & Ben Hedia 2018
Zuraw et al. 2020
Hay 2003
Hay 2007

Previous studies

Higher relative frequency has been found to be associated with:

no change in duration

Pluymaekers et al. 2005b
Plag & Ben Hedia 2018
Ben Hedia & Plag 2017
Zimmerer et al. 2014
Zuraw et al. 2020

longer durations

Plag & Ben Hedia 2018
Zuraw et al. 2020
Hay 2003
Hay 2007

shorter durations

Pluymaekers et al. 2005b
Schuppler et al. 2012

The suspect

Prosodic structure

Plag & Ben Hedia 2018

The suspect

Prosodic structure

Segmentability effects might be counteracted by a strong prosodic boundary:

Plag & Ben Hedia 2018

The suspect

Prosodic structure

Segmentability effects might be counteracted by a strong prosodic boundary:

- The weaker the prosodic boundary is, the less can higher relative frequency protect against reduction.

Plag & Ben Hedia 2018

The suspect

Prosodic structure

Segmentability effects might be counteracted by a strong prosodic boundary:

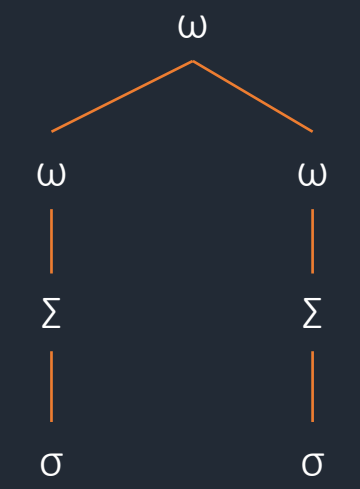
- The weaker the prosodic boundary is, the less can higher relative frequency protect against reduction.
- The stronger a prosodic boundary is, the more pre-boundary lengthening might cancel out reduction effects in barely segmentable words.

Plag & Ben Hedia 2018

Types of prosodic word integration

Raffelsiefen 1999

Types of prosodic word integration



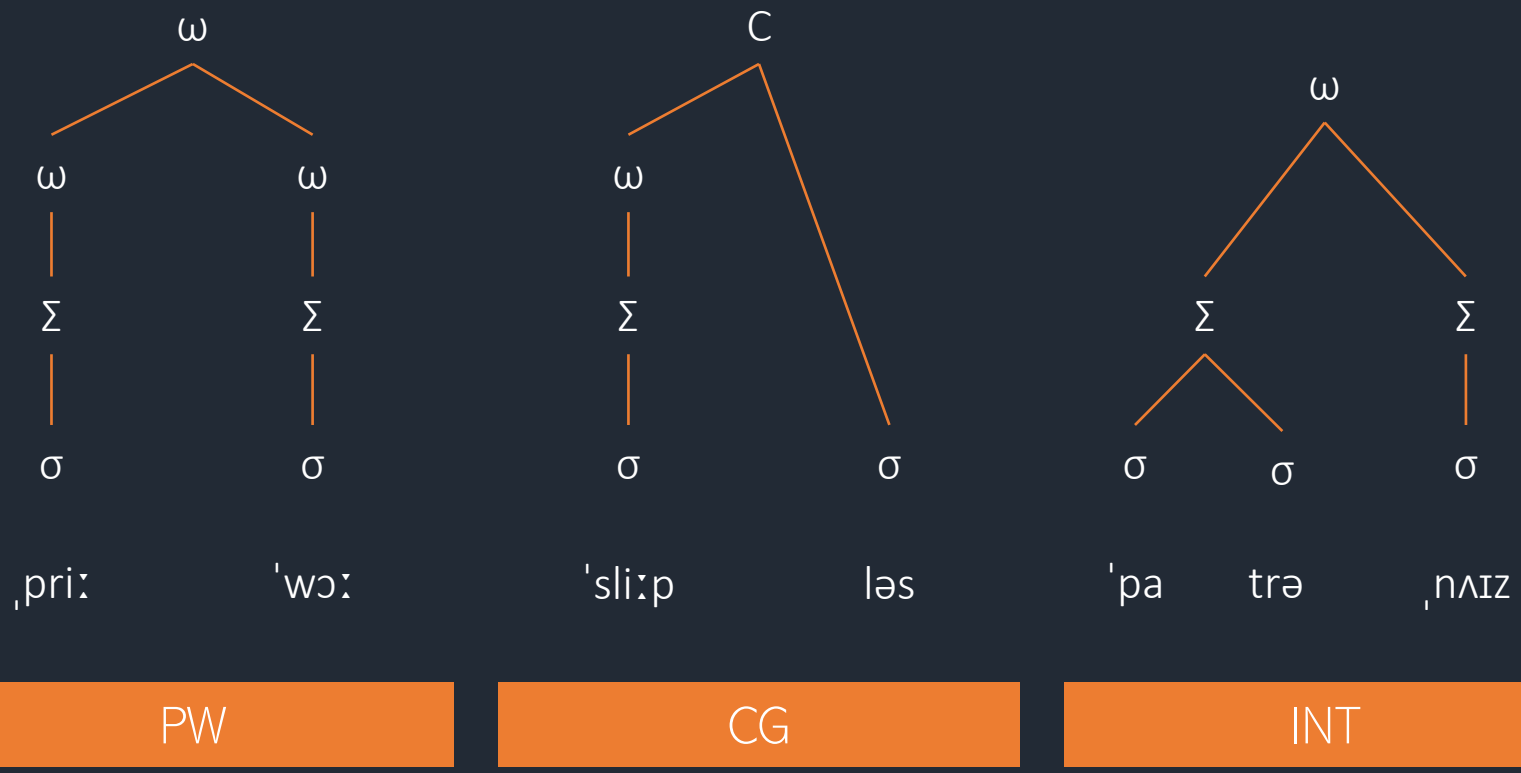
PW

Raffelsiefen 1999

Types of prosodic word integration

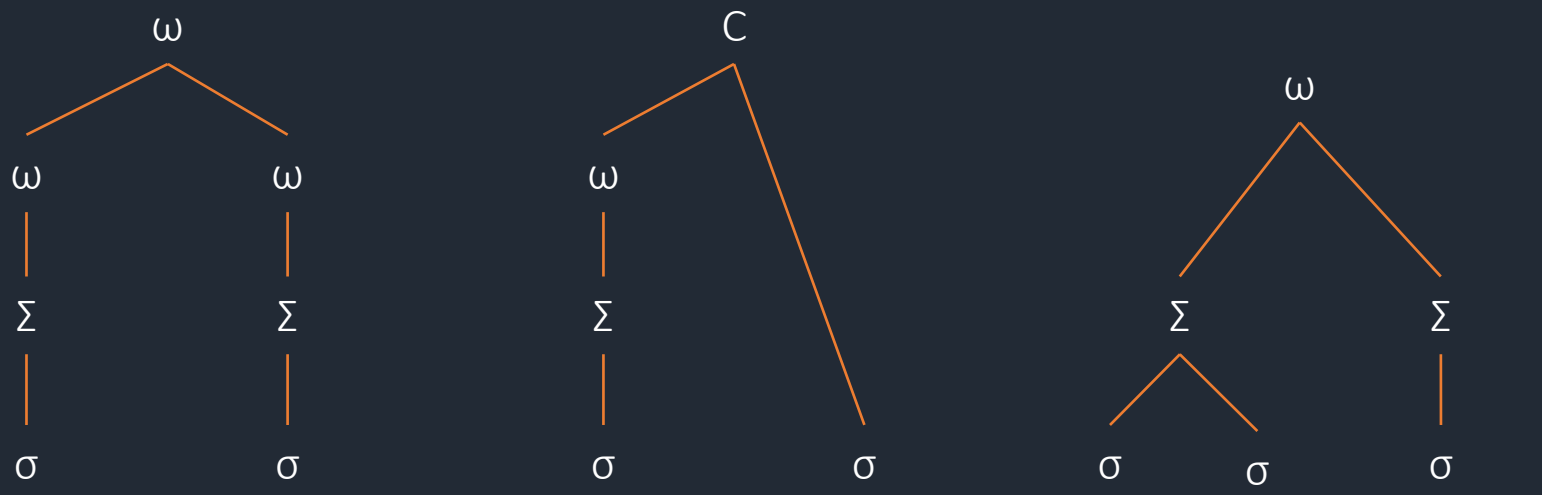
Raffelsiefen 1999

Types of prosodic word integration



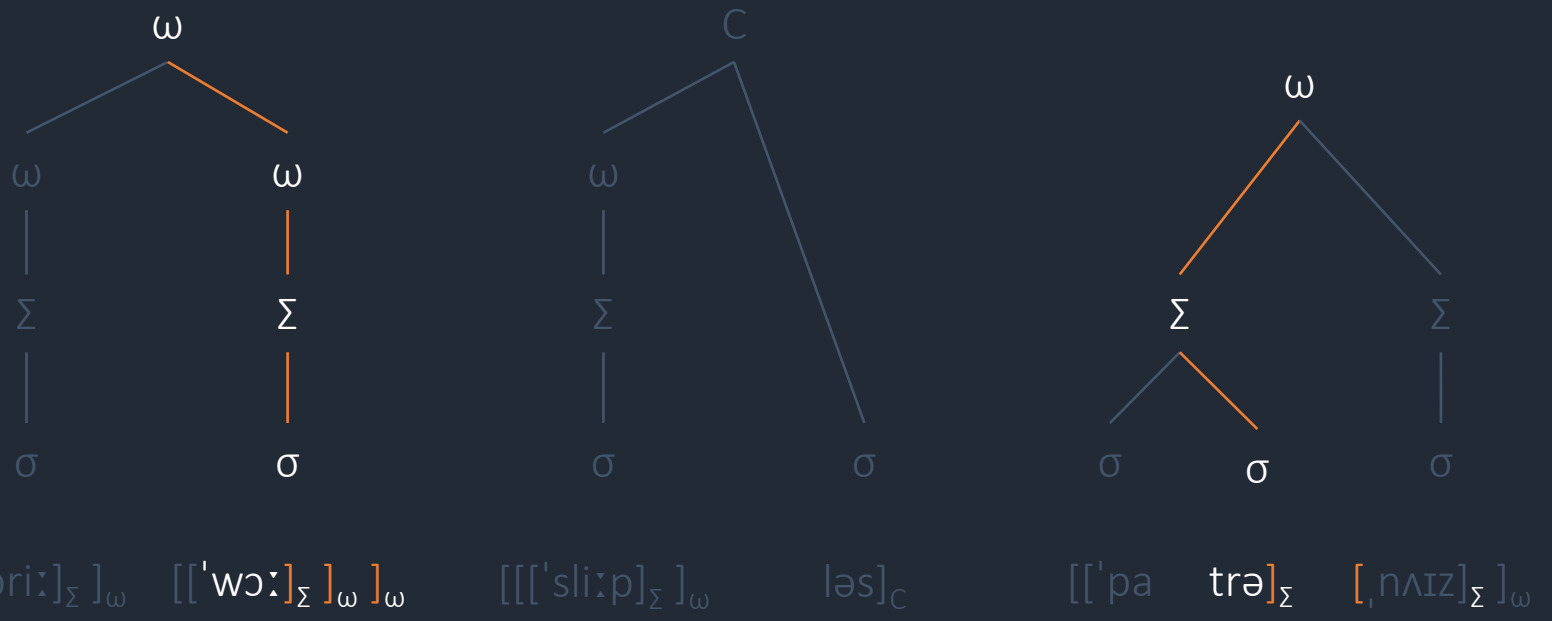
Raffelsiefen 1999

Types of prosodic word integration



Raffelsiefen 1999

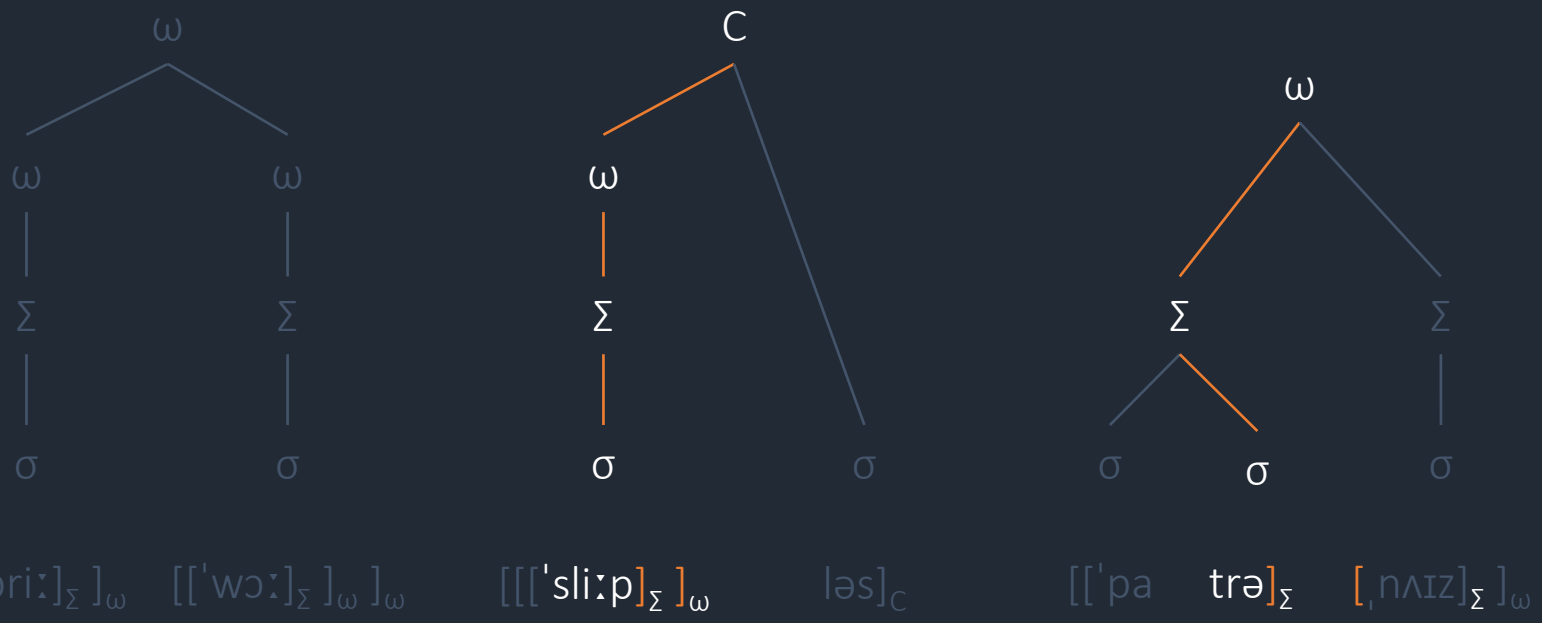
Types of prosodic word integration



PW | CG | INT

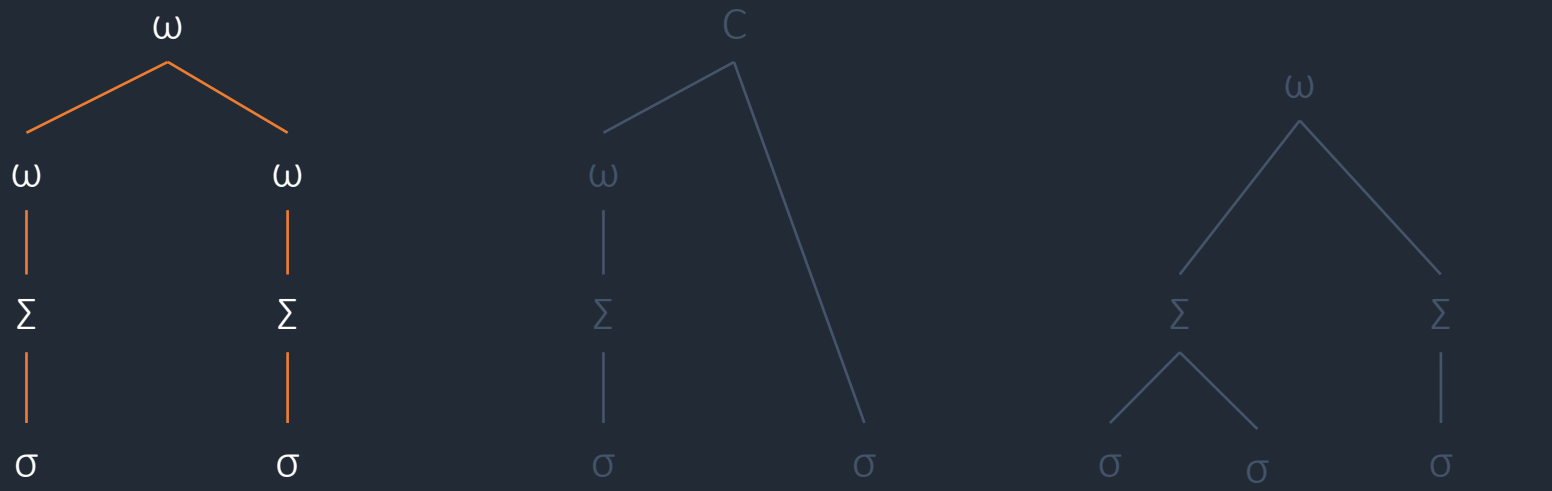
Raffelsiefen 1999

Types of prosodic word integration



Raffelsiefen 1999

Types of prosodic word integration



Raffelsiefen 1999

Expected prosodic lengthening hierarchies

Expected prosodic lengthening hierarchies

PW	INT	CG
base ‣ affix	affix ‣ base	affix ‣ base

X ‣ Y *X more lengthened than Y*

Expected prosodic lengthening hierarchies

PW	INT	CG	bases	affixes
base ‣ affix	affix ‣ base	affix ‣ base	PW ‣ CG ‣ INT	CG ‣ INT ‣ PW

X ‣ Y *X more lengthened than Y*

Summary of expectations

Summary of expectations

H_1 Higher relative frequency should protect against reduction, i.e., be associated with longer durations.

Summary of expectations

H_1 Higher relative frequency should protect against reduction, i.e., be associated with longer durations.

H_2 More prosodic integration should prevent relative frequency from protecting against reduction.

Summary of expectations

H_1 Higher relative frequency should protect against reduction, i.e., be associated with longer durations.

H_2 More prosodic integration should prevent relative frequency from protecting against reduction.

H_3 Pre-boundary lengthening should follow the expected prosodic lengthening hierarchies.


Methods

Data

	PW		CG		INT	
	tokens	types	tokens	types	tokens	types
Audio BNC	*dis-, in-, pre-, un-*		*-ness, -less*		*-ation, -ize*	
	1602	170	529	55	4168	220
QuakeBox	*dis-, un-, re-*		*-ness, -ment*		*-ation, -able, -ity*	
	684	69	441	37	1145	76
ONZE	*dis-, un-, re-*		*-ness, -ment*		*-ation, -able, -ity*	
	810	84	745	48	1556	125

Coleman et al. 2012, Walsh et al. 2013, Gordon et al. 2007

Data

	PW		CG		INT	
	tokens	types	tokens	types	tokens	types
Audio BNC	*dis-, in-, pre-, un-*		*-ness, -less*		*-ation, -ize*	
	1602	170	529	55	4168	220
QuakeBox	*dis-, un-, re-*		*-ness, -ment*		*-ation, -able, -ity*	
	684	69	441	37	1145	76
ONZE	*dis-, un-, re-*		*-ness, -ment*		*-ation, -able, -ity*	
	810	84	745	48	1556	125

Modeling

mixed-effects regression with random intercepts for word type

Coleman et al. 2012, Walsh et al. 2013, Gordon et al. 2007

Model structure

duration difference ~ (1 | Word) +
relative frequency · type of morpheme +
relative frequency · prosodic category +
prosodic category · type of morpheme +
speech rate +
number of syllables +
bigram frequency +
mean biphone probability +
corpus

Model structure

duration difference ~ (1 | Word) + relative frequency · type of morpheme +
relative frequency · prosodic category +
prosodic category · type of morpheme +
speech rate +
number of syllables +
bigram frequency +
mean biphone probability +
corpus

duration difference
residuals of a linear model observed duration ~ baseline duration

Model structure

duration difference ~ (1 | Word) +
relative frequency · type of morpheme +
relative frequency · prosodic category +
prosodic category · type of morpheme +
speech rate +
number of syllables +
bigram frequency +
mean biphone probability +
corpus

Model structure

duration difference ~ (1 | Word) + relative frequency · type of morpheme +
relative frequency · prosodic category +
prosodic category · type of morpheme +
speech rate +
number of syllables +
bigram frequency +
mean biphone probability +
corpus

$$relative\ frequency = \frac{base\ frequency}{word\ frequency}$$

Model structure

duration difference ~ (1 | Word) + relative frequency · type of morpheme +
relative frequency · prosodic category +
prosodic category · type of morpheme +
speech rate +
number of syllables +
bigram frequency +
mean biphone probability +
corpus

type of morpheme
affix or base

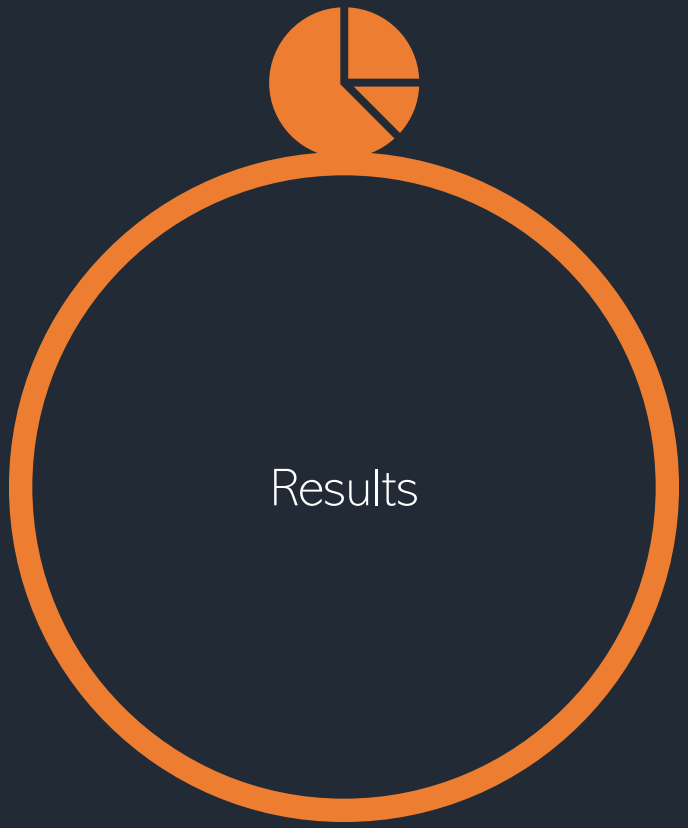
Model structure

duration difference ~ (1 | Word) + relative frequency · type of morpheme +
relative frequency · prosodic category +
prosodic category · type of morpheme +
speech rate +
number of syllables +
bigram frequency +
mean biphone probability +
corpus

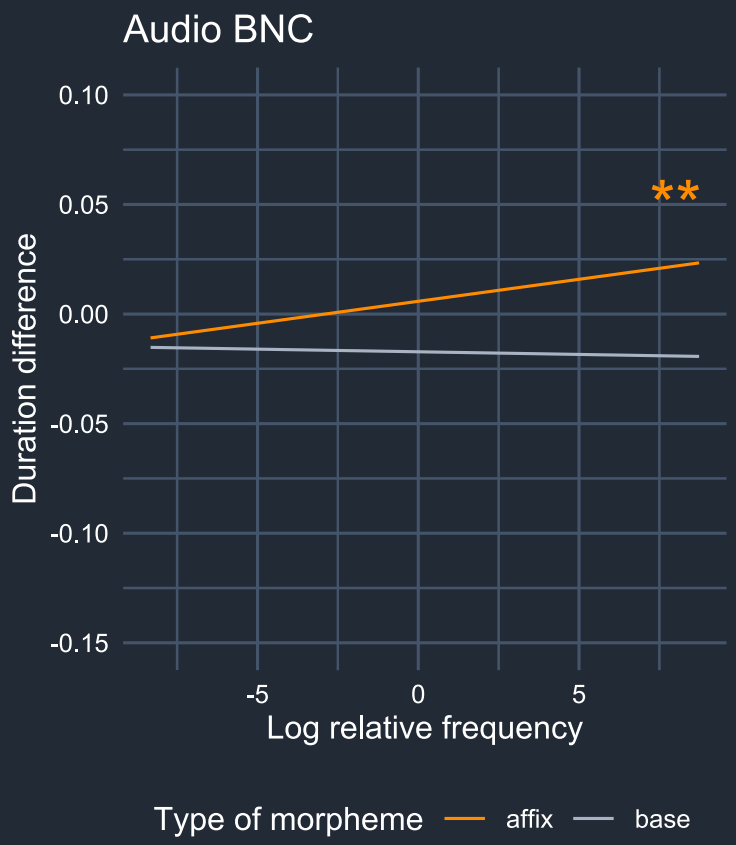
prosodic category
PW, CG, INT

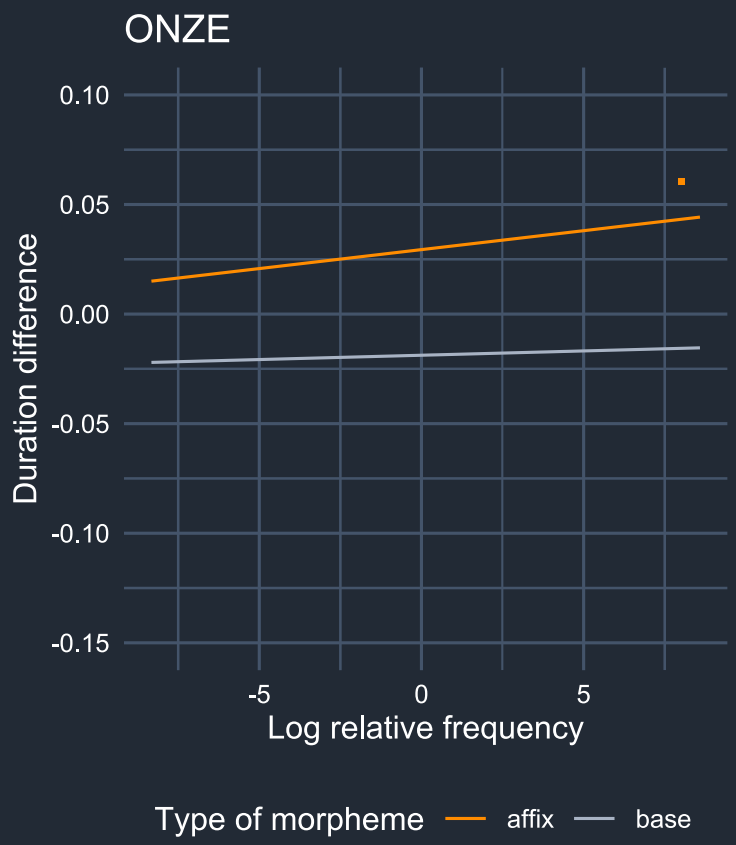
Model structure

duration difference ~ (1 | Word) +
relative frequency · type of morpheme +
relative frequency · prosodic category +
prosodic category · type of morpheme +
speech rate +
number of syllables +
bigram frequency +
mean biphone probability +
corpus


Results

Relative frequency · Type of morpheme





Relative frequency · Prosodic category

In general, prosodic word structure is not a gatekeeper for relative frequency effects.

Prosodic category · Type of morpheme

	PW	INT	CG	bases	affixes
EXP	base ‣ affix	affix ‣ base	affix ‣ base	PW ‣ CG ‣ INT	CG ‣ INT ‣ PW

X ‣ *Y* *X more lengthened than Y*
X = *Y* *no difference*

Prosodic category · Type of morpheme

	PW	INT	CG	bases	affixes
EXP	base ▸ affix	affix ▸ base	affix ▸ base	PW ▸ CG ▸ INT	CG ▸ INT ▸ PW
BNC	base ▸ affix	affix ▸ base	base ▸ affix	PW ▸ CG ▸ INT	INT ▸ PW = CG

X ▸ Y *X more lengthened than Y*
X = Y *no difference*

Prosodic category · Type of morpheme

	PW	INT	CG	bases	affixes
EXP	base ‣ affix	affix ‣ base	affix ‣ base	PW ‣ CG ‣ INT	CG ‣ INT ‣ PW
BNC	base ‣ affix	affix ‣ base	base ‣ affix	PW ‣ CG ‣ INT	INT ‣ PW = CG
QKE	base ‣ affix	affix ‣ base	base ‣ affix	PW ‣ CG = INT	INT ‣ PW = CG

X ‣ Y *X more lengthened than Y*
X = Y *no difference*

Prosodic category · Type of morpheme

	PW	INT	CG	bases	affixes
EXP	base ‣ affix	affix ‣ base	affix ‣ base	PW ‣ CG ‣ INT	CG ‣ INT ‣ PW
BNC	base ‣ affix	affix ‣ base	base ‣ affix	PW ‣ CG ‣ INT	INT ‣ PW = CG
QKE	base ‣ affix	affix ‣ base	base ‣ affix	PW ‣ CG = INT	INT ‣ PW = CG
ONZ	base ‣ affix	affix ‣ base	base ‣ affix	PW ‣ CG = INT	INT ‣ PW ‣ CG

X ‣ Y *X more lengthened than Y*
X = Y *no difference*

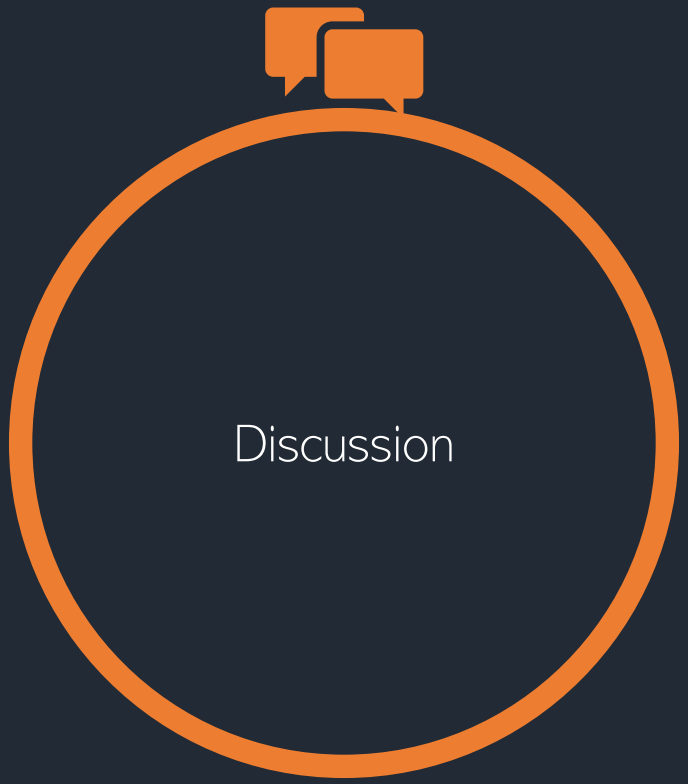
Prosodic category · Type of morpheme

	PW	INT	CG	bases	affixes
EXP	base ‣ affix	affix ‣ base	affix ‣ base	PW ‣ CG ‣ INT	CG ‣ INT ‣ PW
BNC	base ‣ affix ✓	affix ‣ base ✓	base ‣ affix ✗	PW ‣ CG ✓ ‣ INT ✓	INT ‣ PW = CG ✗
QKE	base ‣ affix ✓	affix ‣ base ✓	base ‣ affix ✗	PW ‣ CG ✓ = INT ?	INT ‣ PW ✓ = CG ✗
ONZ	base ‣ affix ✓	affix ‣ base ✓	base ‣ affix ✗	PW ‣ CG ✓ = INT ?	INT ‣ PW ✓ ‣ CG ✗

X ‣ Y *X more lengthened than Y*
X = Y *no difference*

Prosodic category · Type of morpheme

In general, prosodic boundaries fail to account consistently for durational differences.


Discussion

H_1 partial support

Higher relative frequency can be associated with lengthening, but often isn't.

Hay 2003, Pluymaekers et al. 2005b, Hay 2007, Schuppler et al. 2012, Zimmerer et al. 2014, Ben Hedia & Plag 2017, Plag & Ben Hedia 2018, Zuraw et al. 2020, Bowden et al. 2010

H_1 partial support

Higher relative frequency can be associated with lengthening, but often isn't.

- If we take the effect seriously, it implies that morphological information is sometimes still reflected at the subphonemic level.

Hay 2003, Pluymaekers et al. 2005b, Hay 2007, Schuppler et al. 2012, Zimmerer et al. 2014, Ben Hedia & Plag 2017, Plag & Ben Hedia 2018, Zuraw et al. 2020, Bowden et al. 2010

H_1 partial support

Higher relative frequency can be associated with lengthening, but often isn't.

- If we take the effect seriously, it implies that morphological information is sometimes still reflected at the subphonemic level.
- However, our study "replicates" the mixture of effects and null effects.

Hay 2003, Pluymaekers et al. 2005b, Hay 2007, Schuppler et al. 2012, Zimmerer et al. 2014, Ben Hedia & Plag 2017, Plag & Ben Hedia 2018, Zuraw et al. 2020, Bowden et al. 2010

H_1 partial support

Higher relative frequency can be associated with lengthening, but often isn't.

- If we take the effect seriously, it implies that morphological information is sometimes still reflected at the subphonemic level.
- However, our study "replicates" the mixture of effects and null effects.
- Additional analyses suggest that positive relative frequency effects on duration only emerge in the presence of word frequency effects.

Hay 2003, Pluymaekers et al. 2005b, Hay 2007, Schuppler et al. 2012, Zimmerer et al. 2014, Ben Hedia & Plag 2017, Plag & Ben Hedia 2018, Zuraw et al. 2020, Bowden et al. 2010

H_1 partial support

Higher relative frequency can be associated with lengthening, but often isn't.

- If we take the effect seriously, it implies that morphological information is sometimes still reflected at the subphonemic level.
- However, our study "replicates" the mixture of effects and null effects.
- Additional analyses suggest that positive relative frequency effects on duration only emerge in the presence of word frequency effects.
- We might need to consider discarding relative frequency as a predictor of morpho-phonetic variation.

Hay 2003, Pluymaekers et al. 2005b, Hay 2007, Schuppler et al. 2012, Zimmerer et al. 2014, Ben Hedia & Plag 2017, Plag & Ben Hedia 2018, Zuraw et al. 2020, Bowden et al. 2010

H_2 rejected

The degree of prosodic word integration does not influence whether higher relative frequency can protect against reduction.

Hay 2003, Pluymaekers et al. 2005b, Hay 2007, Schuppler et al. 2012, Zimmerer et al. 2014, Ben Hedia & Plag 2017, Plag & Ben Hedia 2018, Zuraw et al. 2020

H_2 rejected

The degree of prosodic word integration does not influence whether higher relative frequency can protect against reduction.

- This is indirectly consistent with previous studies, which have found effects of relative frequency on duration for both non-integrating affixes and integrating affixes.

Hay 2003, Pluymaekers et al. 2005b, Hay 2007, Schuppler et al. 2012, Zimmerer et al. 2014, Ben Hedia & Plag 2017, Plag & Ben Hedia 2018, Zuraw et al. 2020

H_2 rejected

The degree of prosodic word integration does not influence whether higher relative frequency can protect against reduction.

- This is indirectly consistent with previous studies, which have found effects of relative frequency on duration for both non-integrating affixes and integrating affixes.
- Previous studies also show that neither an integrating nor a non-integrating affix guarantees a relative frequency effect.

Hay 2003, Pluymaekers et al. 2005b, Hay 2007, Schuppler et al. 2012, Zimmerer et al. 2014, Ben Hedia & Plag 2017, Plag & Ben Hedia 2018, Zuraw et al. 2020

H_3 partial support

The prosodic structure of complex words cannot consistently explain durational variation.

Sproat & Fujimura 1993, Auer 2002, Sugahara & Turk 2009, Bergmann 2018, also see Klatt 1975, Vaissière 1983, Edwards & Beckman 1988, Beckman & Pierrehumbert 1986, Campbell 1990, Wightman et al. 1992

H_3 partial support

The prosodic structure of complex words cannot consistently explain durational variation.

- Some previous studies had suggested that prosodic structure can account for some durational variation.

Sproat & Fujimura 1993, Auer 2002, Sugahara & Turk 2009, Bergmann 2018, also see Klatt 1975, Vaissière 1983, Edwards & Beckman 1988, Beckman & Pierrehumbert 1986, Campbell 1990, Wightman et al. 1992

H_3 partial support

The prosodic structure of complex words cannot consistently explain durational variation.

- Some previous studies had suggested that prosodic structure can account for some durational variation.
- However, there are important differences between these studies and ours (domains, conditions, methodologies, level of prosodic boundaries).

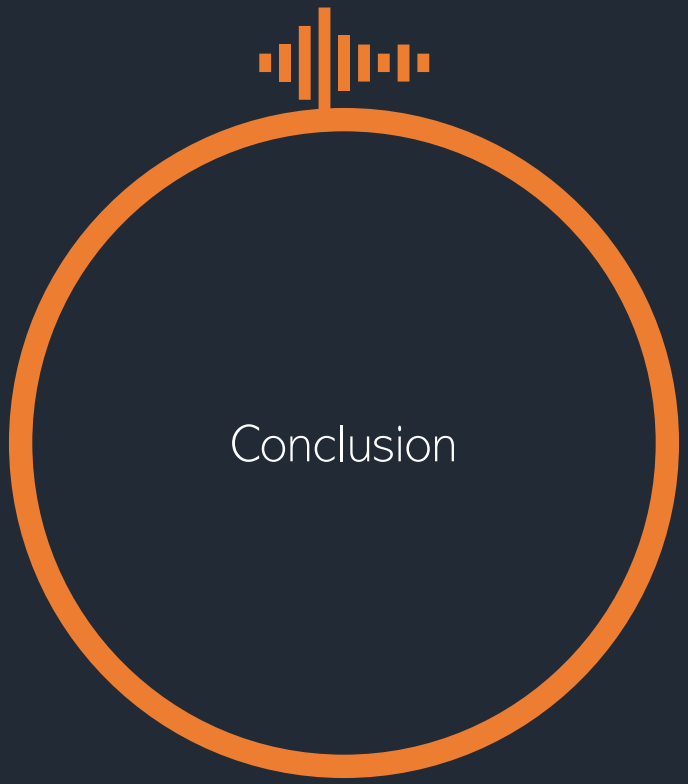
Sproat & Fujimura 1993, Auer 2002, Sugahara & Turk 2009, Bergmann 2018, also see Klatt 1975, Vaissière 1983, Edwards & Beckman 1988, Beckman & Pierrehumbert 1986, Campbell 1990, Wightman et al. 1992

H_3 partial support

The prosodic structure of complex words cannot consistently explain durational variation.

- Some previous studies had suggested that prosodic structure can account for some durational variation.
- However, there are important differences between these studies and ours (domains, conditions, methodologies, level of prosodic boundaries).
- In phonological theory and in models of speech production, it is unclear how the supposed word-internal boundaries translate into articulatory gestures or acoustic properties.

Sproat & Fujimura 1993, Auer 2002, Sugahara & Turk 2009, Bergmann 2018, also see Klatt 1975, Vaissière 1983, Edwards & Beckman 1988, Beckman & Pierrehumbert 1986, Campbell 1990, Wightman et al. 1992



Conclusion

Takeaway

Relative frequency and prosodic word integration do not reliably predict duration, and prosodic word integration is not responsible for the emergence of relative frequency effects.

Takeaway

Relative frequency and prosodic word integration do not reliably predict duration, and prosodic word integration is not responsible for the emergence of relative frequency effects.

What's next?

Takeaway

Relative frequency and prosodic word integration do not reliably predict duration, and prosodic word integration is not responsible for the emergence of relative frequency effects.

What's next?

- We may need to explore other factors for the morphology-phonetics interaction and for processing in the mental lexicon.

Takeaway

Relative frequency and prosodic word integration do not reliably predict duration, and prosodic word integration is not responsible for the emergence of relative frequency effects.

What's next?

- We may need to explore other factors for the morphology-phonetics interaction and for processing in the mental lexicon.
- The morphology-phonology-phonetics interface might be better modeled by non-morphemic, word-based approaches, such as discrimination learning.

Stein & Plag (submitted)


Thank you!

- Arndt-Lappe, Sabine & Mirjam Ernestus. 2020. Morpho-phonological alternations: The role of lexical storage. In Vito Pirrelli, Ingo Plag & Wolfgang U. Dressler (eds.), *Word knowledge and word usage: A cross-disciplinary guide to the mental lexicon* (Trends in Linguistics: Studies and Monographs 337), 191–227. Berlin, Boston: Mouton de Gruyter.
- Auer, Peter. 2002. Die sogenannte Auslautverhärtung in ne[b]lig vs. lie[p]lich: Ein Phantom der deutschen Phonologie? In Michael Bommes, Christina Noack & Doris Tophinke (eds.), *Sprache als Form: Festschrift für Utz Maas zum 60. Geburtstag*, 74–86. Wiesbaden: Westdeutscher Verlag.
- Baayen, R. Harald. 2008. *Analyzing Linguistic Data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.
- Baayen, R. Harald & Petar Milin. 2010. Analyzing reaction times. *International Journal of Psychological Research* 3.2: 12–28. doi: 10.21500/20112084.807.
- Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67.1: 1–48. doi: 10.18637/jss.v067.i01.
- Bauer, Laurie, Rochelle Lieber & Ingo Plag. 2013. *The Oxford reference guide to English morphology*. Oxford: Oxford University Press.
- Beckman, Mary E. & Janet B. Pierrehumbert. 1986. Intonational structure in Japanese and English. *Phonology Yearbook* 3: 255–309. doi: 10.1017/S095267570000066X.
- Bell, Alan, Jason M. Brenier, Michelle Gregory, Cynthia Girand & Daniel Jurafsky. 2009. Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language* 60.1: 92–111.

- Bell, Alan, Daniel Jurafsky, Eric Fosler-Lussier, Cynthia Girand, Michelle Gregory & Daniel Gildea. 2003. Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *The Journal of the Acoustical Society of America* 113.2: 1001–1024. doi: 10.1121/1.1534836.
- Bell, Melanie J., Sonia Ben Hedia & Ingo Plag. 2020. How morphological structure affects phonetic realisation in English compound nouns. *Morphology*. 1–34. doi: 10.1007/s11525-020-09346-6.
- Ben Hedia, Sonia. 2019. *Gemination and degemination in English affixation: Investigating the interplay between morphology, phonology and phonetics* (Studies in Laboratory Phonology 8). Berlin: Language Science Press.
- Ben Hedia, Sonia & Ingo Plag. 2017. Gemination and degemination in English prefixation: Phonetic evidence for morphological organization. *Journal of Phonetics* 62: 34–49.
- Bergmann, Pia. 2018. *Morphologisch komplexe Wörter: Prosodische Struktur und phonetische Realisierung* (Studies in Laboratory Phonology 5). Berlin: Language Science Press. doi: 10.5281/ZENODO.1346245.
- Bermúdez-Otero, Ricardo. 2010. Morphologically conditioned phonetics? Not proven. Paper presented at On Linguistic Interfaces II, 2 December, Belfast.
- Boersma, Paul & David J. M. Weenik. 2001. Praat: Doing phonetics by computer. Version 5.4.04. http://www.praat.org/.
- Bowden, Harriet W., Matthew P. Gelfand, Cristina Sanz & Michael T. Ullman. 2010. Verbal inflectional morphology in L1 and L2 Spanish: A frequency effects study examining storage versus composition. *Language Learning* 60.1: 44–87. doi: 10.1111/j.1467-9922.2009.00551.x.

- Bybee, Joan L. 2000. The phonology of the lexicon: Evidence from lexical diffusion. In Michael Barlow & Suzanne Kemmer (eds.), *Usage-based models of language*, 65–85. Cambridge: Cambridge University Press.
- Campbell, W. Nick. 1990. Evidence for a syllable-based model of speech timing. *Proceedings of the International Conference on Spoken Language Processing*. 9–12.
- Caselli, Naomi K., Michael K. Caselli & Ariel M. Cohen-Goldberg. 2016. Inflected words in production: Evidence for a morphologically rich lexicon. *The Quarterly Journal of Experimental Psychology* 69.3: 432–454.
- Coleman, John, Ladan Baghai-Ravary, John Pybus & Sergio Grau. 2012. *Audio BNC: The audio edition of the Spoken British National Corpus.* University of Oxford. http://www.phon.ox.ac.uk/AudioBNC.
- Davies, Mark. 2008. *The Corpus of Contemporary American English: 450 million words, 1990–present*. http://corpus.byu.edu/coca/.
- Divjak, Dagmar. 2019. *Frequency in language: Memory, attention and learning*. Cambridge: Cambridge University Press. doi: 10.1017/9781316084410.
- Edwards, Jan & Mary E. Beckman. 1988. Articulatory timing and the prosodic interpretation of syllable duration. *Phonetica* 45.2–4: 156–174. doi: 10.1159/000261824.
- Edwards, Jan, Mary E. Beckman & Benjamin Munson. 2004. The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *Journal of Speech, Language, and Hearing Research* 47.2: 421–436. doi: 10.1044/1092-4388(2004/034).
- Fox, John & Sanford Weisberg. 2011. *An R companion to applied regression*, 2nd edn. Thousand Oaks: SAGE Publications.

- Fromont, Robert. 2003–2020. LaBB-Cat. University of Canterbury. labbcat.canterbury.ac.nz/.
- Fromont, Robert & Jennifer Hay. 2012. LaBB-CAT: An annotation store. *Proceedings of the Australasian Language Technology Association Workshop* 10: 113–117.
- Gahl, Susanne. 2008. Thyme and time are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language* 84.3: 474–496.
- Gahl, Susanne, Yao Yao & Keith Johnson. 2012. Why reduce? Phonological neighborhood density and phonetic reduction in spontaneous speech. *Journal of Memory and Language* 66.4: 789–806. doi: 10.1016/j.jml.2011.11.006.
- Gordon, Elizabeth, Margaret Maclagan & Jennifer Hay. 2007. The ONZE corpus. In Joan C. Beal, Karen P. Corrigan & Hermann L. Moisl (eds.), *Creating and digitizing language corpora, Volume 2: Diachronic corpora*, 82–104. Basingstoke: Palgrave Macmillan.
- Guy, Gregory R. 1980. Variation in the group and the individual: The case of final stop deletion. In William Labov (ed.), *Locating language in time and space* (Quantitative Analyses of Linguistic Structure 1), 1–36. New York, London: Academic Press.
- Guy, Gregory R. 1991. Explanation in variable phonology: An exponential model of morphological constraints. *Language Variation and Change* 3.1: 1–22. doi: 10.1017/S0954394500000429.
- Hall, Tracy A. 1999. The phonological word: A review. In Tracy A. Hall & Ursula Kleinhenz (eds.), *Studies of the phonological word* (Current Issues in Linguistic Theory 174), 1–22. Amsterdam, Philadelphia: John Benjamins.
- Hay, Jennifer. 2001. Lexical frequency in morphology: Is everything relative? *Linguistics* 39.6: 1041–1070.

- Hay, Jennifer. 2003. *Causes and consequences of word structure*. New York, London: Routledge.
- Hay, Jennifer. 2007. The phonetics of un. In Judith Munat (ed.), *Lexical creativity, texts and contexts*, 39–57. Amsterdam, Philadelphia: John Benjamins.
- Hildebrandt, Kristine A. 2015. The prosodic word. In John R. Taylor (ed.), *The Oxford handbook of the word*, 221–245. Oxford: Oxford University Press.
- Jurafsky, Daniel. 2003. Probabilistic modeling in psycholinguistics: Linguistic comprehension and production. In Rens Bod, Jennifer Hay & Stefanie Jannedy (eds.), *Probabilistic linguistics*, 39–95. Cambridge, London: MIT Press.
- Jurafsky, Daniel, Alan Bell, Michelle Gregory & William D. Raymond. 2001. Probabilistic relations between words: Evidence from reduction in lexical production. In Joan Bybee & Paul J. Hopper (eds.), *Frequency and the emergence of linguistic structure* (Typological Studies in Language 45), 229–254. Amsterdam: John Benjamins.
- Kiparsky, Paul. 1982. Lexical morphology and phonology. In In-Seok Yang (ed.), *Linguistics in the morning calm: Selected papers from SICOL*, 3–91. Seoul: Hanshin.
- Klatt, Dennis H. 1975. Vowel lengthening is syntactically determined in a connected discourse. *Journal of Phonetics* 3.3: 129–140. doi: 10.1016/S0095-4470(19)31360-9.
- Kunter, Gero. 2016. Coquery: A free corpus query tool. www.coquery.org.
- Kuznetsova, Alexandra, Per B. Brockhoff & Rune H. B. Christensen. 2016. lmerTest: Tests in linear fixed effects models. https://cran.r-project.org/web/packages/lmerTest/index.html.
- Labov, William. 1989. The child as linguistic historian. *Language Variation and Change* 1.1: 85–97. doi: 10.1017/S0954394500000120.

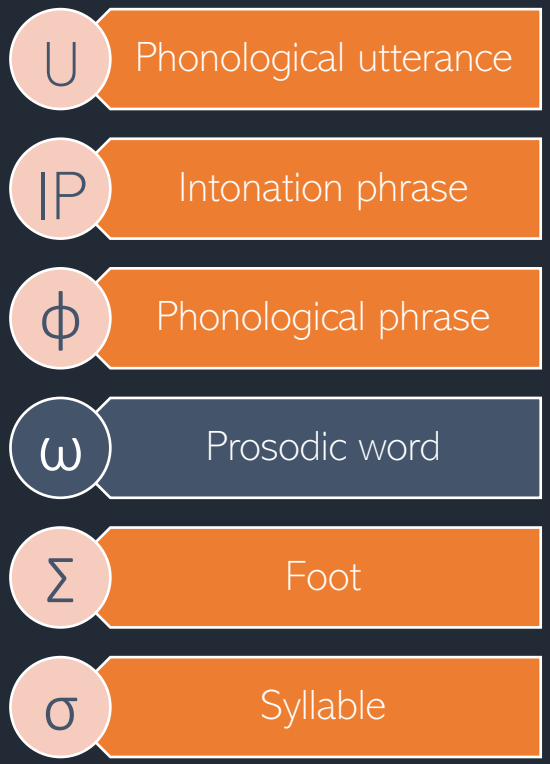
- Ladefoged, Peter & Keith Johnson. 2011. *A course in phonetics*, 6th edn. Boston: Wadsworth Cengage Learning.
- Levelt, William J. M., Ardi Roelofs & Antje S. Meyer. 1999. A theory of lexical access in speech production. *Behavioral and Brain Sciences* 22.1: 1–38.
- Lindblom, Björn. 1963. Spectrographic study of vowel reduction. *The Journal of the Acoustical Society of America* 35.11: 1773–1781. doi: 10.1121/1.1918816.
- Losiewicz, Beth L. 1995. Word frequency effects on the acoustic duration of morphemes. *The Journal of the Acoustical Society of America* 97.5: 3243. doi: 10.1121/1.411745.
- Machač, Pavel & Radek Skarnitzl. 2009. *Principles of phonetic segmentation*. Prague: Epocha Publishing House.
- Munson, Benjamin. 2001. Phonological pattern frequency and speech production in adults and children. *Journal of Speech, Language, and Hearing Research* 44.4: 778–792. doi: 10.1044/1092-4388(2001/061).
- Nespor, Marina & Irene Vogel. 2007. *Prosodic phonology*. Berlin, New York: Walter de Gruyter.
- Nooteboom, Sieb G. 1972. *Production and perception of vowel duration: A study of the durational properties of vowels in Dutch*. Utrecht: University of Utrecht.
- Oxford English Dictionary Online. 2020. Oxford University Press, www.oed.com.
- Plag, Ingo. 2018. *Word-formation in English*, 2nd edn. Cambridge: Cambridge University Press.
- Plag, Ingo & Sonia Ben Hedia. 2018. The phonetics of newly derived words: Testing the effect of morphological segmentability on affix duration. In Sabine Arndt-Lappe, Angelika Braun, Claudine Moulin & Esme Winter-Froemel (eds.), *Expanding the lexicon: Linguistic innovation, morphological productivity, and ludicity*, 93–116. Berlin, New York: Mouton de Gruyter.

- Plag, Ingo, Christiane Dalton-Puffer & R. Harald Baayen. 1999. Morphological productivity across speech and writing. English Language and *Linguistics* 3.2: 209–228. doi: 10.1017/S1360674399000222.
- Plag, Ingo, Julia Homann & Gero Kunter. 2017. Homophony and morphology: The acoustics of word-final S in English. *Journal of Linguistics* 53.1: 181–216.
- Pluymaekers, Mark, Mirjam Ernestus & R. Harald Baayen. 2005a. Articulatory planning is continuous and sensitive to informational redundancy. *Phonetica* 62.2–4: 146–159. doi: 10.1159/000090095.
- Pluymaekers, Mark, Mirjam Ernestus & R. Harald Baayen. 2005b. Lexical frequency and acoustic reduction in spoken Dutch. *The Journal of the Acoustical Society of America* 118.4: 2561–2569.
- Pluymaekers, Mark, Mirjam Ernestus, R. Harald Baayen & Geert Booij. 2010. Morphological effects on fine phonetic detail: The case of Dutch -igheid. In Aditi Lahiri, Cécile Fougeron, Barbara Kühnert, Mariapaola D'Imperio & Nathalie Vallée (eds.), *Laboratory Phonology 10* (Phonology and Phonetics), 511–531. Berlin, New York: Mouton de Gruyter. doi: 10.1515/9783110224917.5.511.
- R Core Team. 2020. R: A language and environment for statistical computing. Version 4.0.1. Vienna: R Foundation for Statistical Computing. http://www.R-project.org/.
- Raffelsiefen, Renate. 1999. Diagnostics for prosodic words revisited: The case of historically prefixed words in English. In Tracy A. Hall & Ursula Kleinhenz (eds.), *Studies of the phonological word* (Current Issues in Linguistic Theory 174), 133–201. Amsterdam, Philadelphia: John Benjamins.
- Raffelsiefen, Renate. 2007. Morphological word structure in English and Swedish: The evidence from prosody. In Geert Booij, Luca Ducceschi, Bernard Fradin, Ernesto Guevara, Angela Ralli & Sergio Scalise (eds.), *Online Proceedings of the Fifth Mediterranean Morphology Meeting* (MMM5): 209–268. Fréjus.

- Schuppler, Barbara, Wim A. van Dommelen, Jacques Koreman & Mirjam Ernestus. 2012. How linguistic and probabilistic properties of a word affect the realization of its final /t/: Studies at the phonemic and sub-phonemic level. *Journal of Phonetics* 40.4: 595–607.
- Sóskuthy, Márton & Jennifer Hay. 2017. Changing word usage predicts changing word durations in New Zealand English. *Cognition* 166: 298–313. doi: 10.1016/j.cognition.2017.05.032.
- Sproat, Richard & Osamu Fujimura. 1993. Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics* 21.3: 291–311. 10.1016/S0095-4470(19)31340-3.
- Stein, Simon David & Ingo Plag. Morpho-phonetic effects in speech production: Modeling the acoustic duration of English derived words with linear discriminative learning. Submitted to *Frontiers in Psychology*.
- Sugahara, Mariko & Alice Turk. 2009. Durational correlates of English sublexical constituent structure. *Phonology* 26: 477–524.
- Torreira, Francisco & Mirjam Ernestus. 2009. Probabilistic effects on French [t] duration. *INTERSPEECH*: 448–451.
- Tucker, Benjamin V. & Mirjam Ernestus. 2016. Why we need to investigate casual speech to truly understand language production, processing and the mental lexicon. *The Mental Lexicon* 11.3: 375–400. doi: 10.1075/ml.11.3.03tuc.
- Turnbull, Rory. 2018. Patterns of probabilistic segment deletion/reduction in English and Japanese. *Linguistics Vanguard* 4.s2. doi: 10.1515/lingvan-2017-0033.
- Vaissière, Jacqueline. 1983. Language-independent prosodic features. In Willem J. M. Levelt, Anne Cutler & D. Robert Ladd (eds.), *Prosody: Models and measurements* (Springer Series in Language and Communication 14), 53–66. Berlin, Heidelberg: Springer. doi: 10.1007/978-3-642-69103-4_5.

- Vitevitch, Michael S. & Paul A. Luce. 2004. A web-based interface to calculate phonotactic probability for words and nonwords in English. *Behavior Research Methods, Instruments, and Computers* 36.3: 481–487.
- Walsh, Liam, Jennifer Hay, Bent Derek, Liz Grant, Jeanette King, Paul Millar, Viktoria Papp & Kevin Watson. 2013. The UC QuakeBox Project: Creation of a community-focused research archive. *New Zealand English Journal* 27: 20–32.
- Wightman, Colin W., Stefanie Shattuck-Hufnagel, Mari Ostendorf & Patti Price. 1992. Segmental durations in the vicinity of prosodic phrase boundaries. *The Journal of the Acoustical Society of America* 91.3: 1707–1717. doi: 10.1121/1.402450.
- Zimmerer, Frank, Mathias Scharinger & Henning Reetz. 2014. Phonological and morphological constraints on German /t/-deletions. *Journal of Phonetics* 45: 64–75. doi: 10.1016/j.wocn.2014.03.006.
- Zuraw, Kie, Isabelle Lin, Meng Yang & Sharon Peperkamp. 2020. Competition between whole-word and decomposed representations of English prefixed words. *Morphology*. 10.1007/s11525-020-09354-6. doi: 10.1007/s11525-020-09354-6.

The prosodic hierarchy



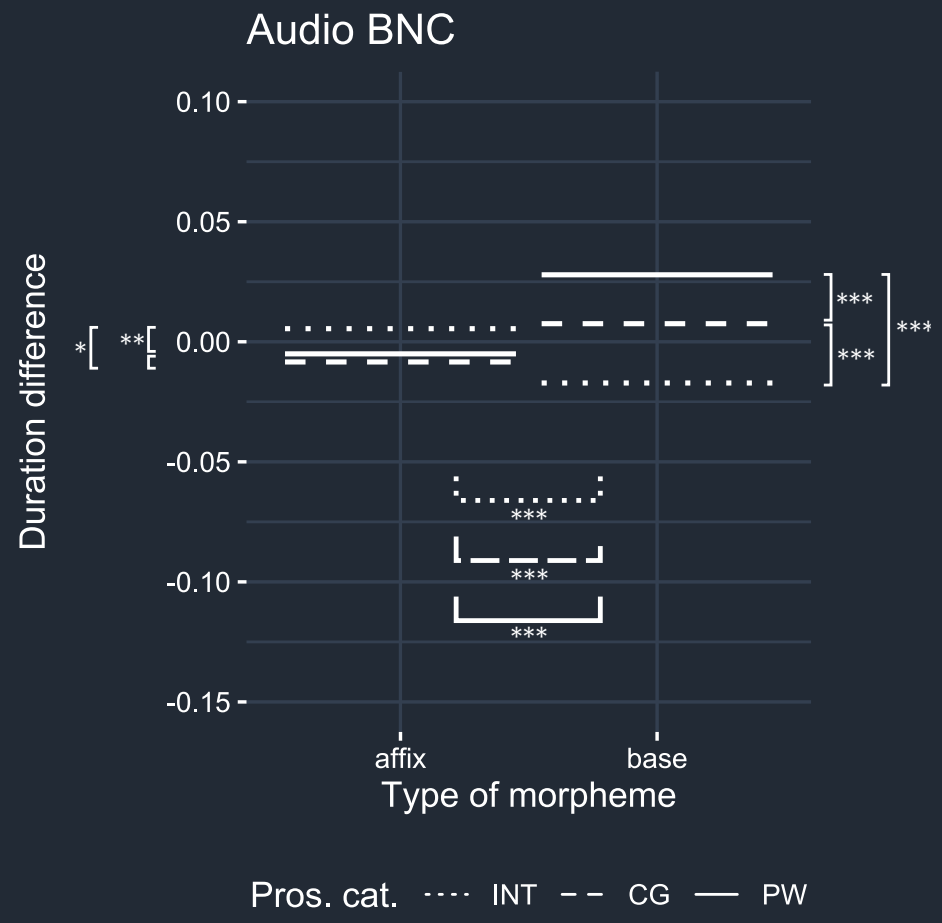
Some pword-diagnostics

- LOI-violations, ambisyllabicity
- stress and relative prominence
- trisyllabic laxing, vowel reduction
- minimal word requirements
- compositionality, type of base

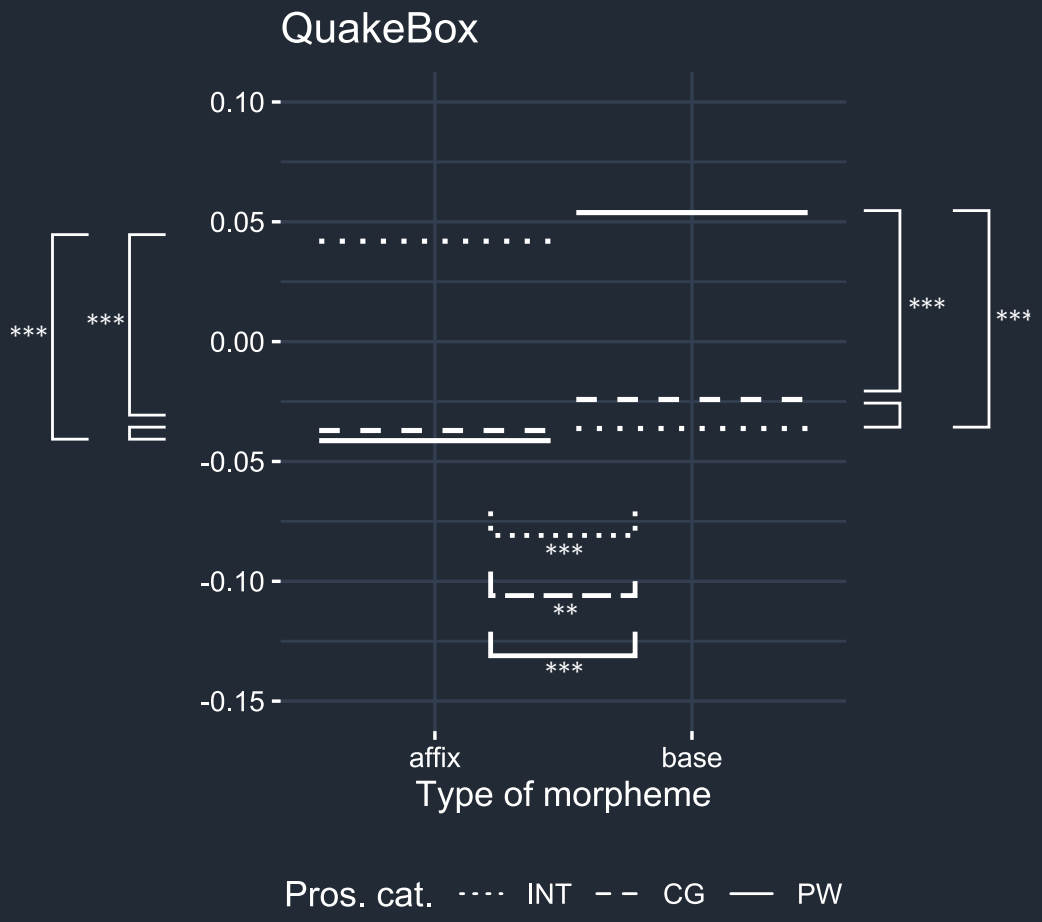
Morpho-prosodic alignment

- A morpheme cannot include multiple pwords, but a pword can include multiple morphemes.

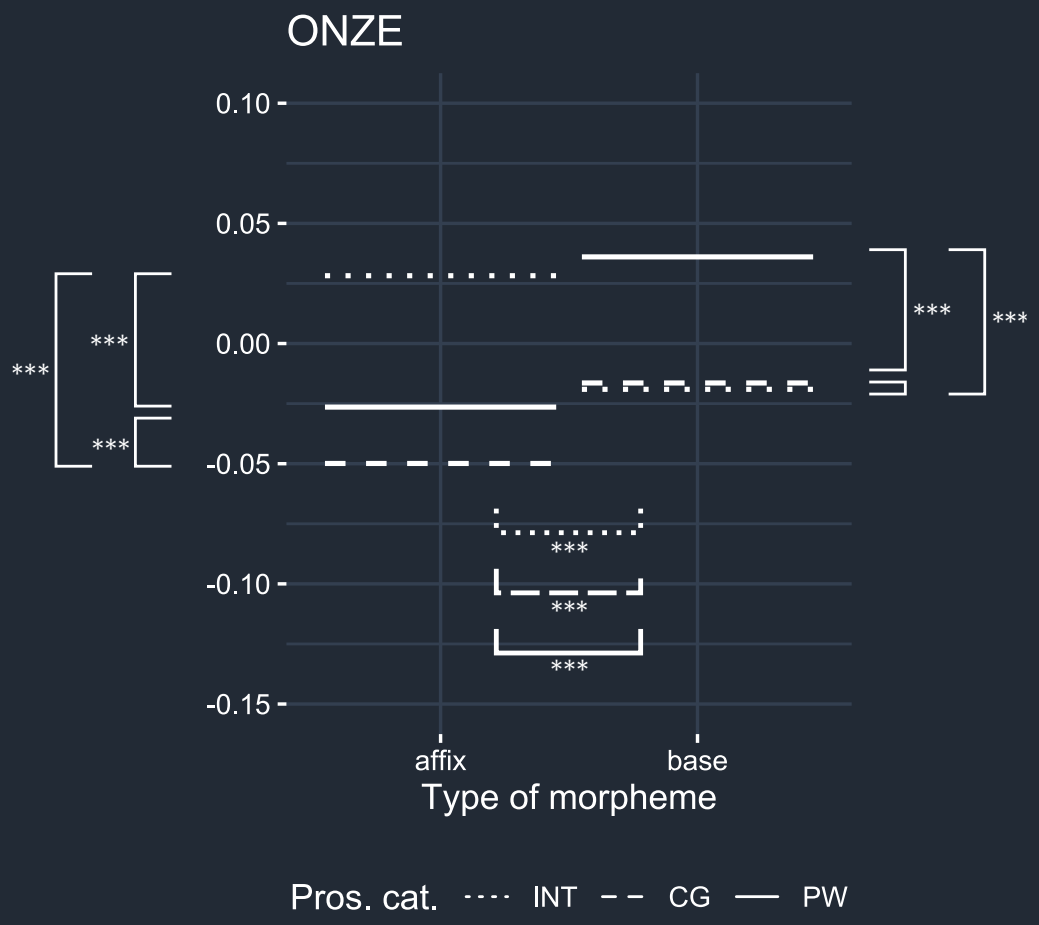
Prosodic category · Type of morpheme



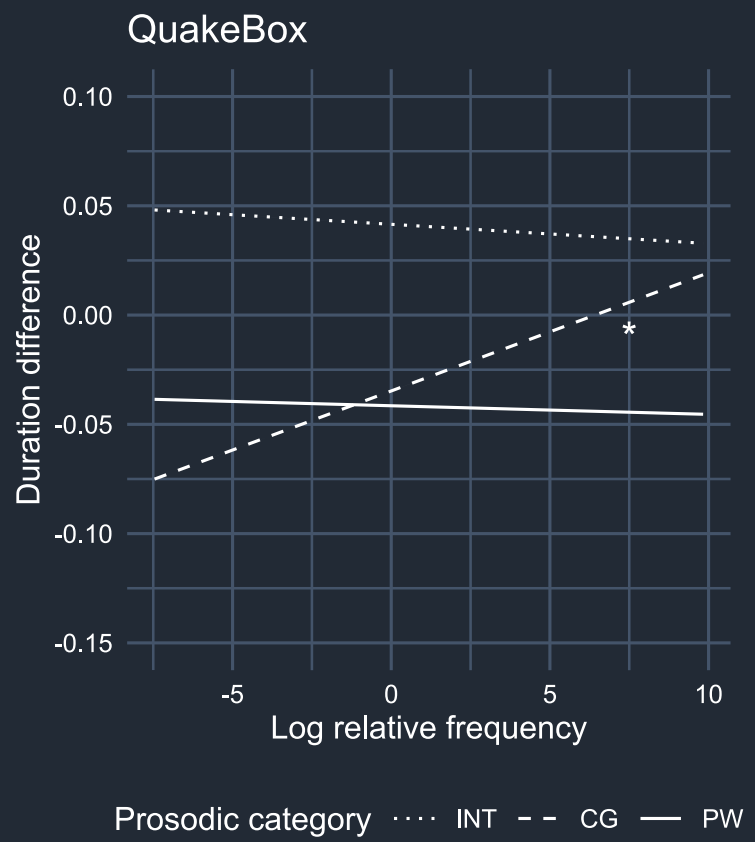
Prosodic category · Type of morpheme



Prosodic category · Type of morpheme



Relative frequency · Prosodic category



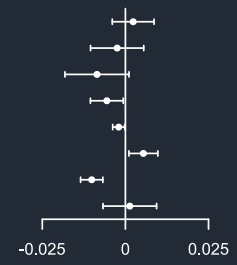
The model is not significantly better than the same model without this interaction.

Category-internal frequency models Audio BNC

word frequency

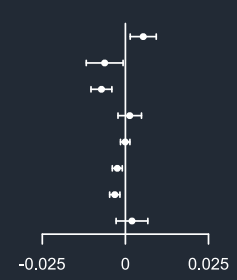
word duration

Affix	β	p
-ness	0.002	–
-less	-0.002	–
pre-	-0.009	–
-ize	-0.006	–
-ation	-0.002	*
dis-	0.005	*
un-	-0.01	***
in-	0.001	–



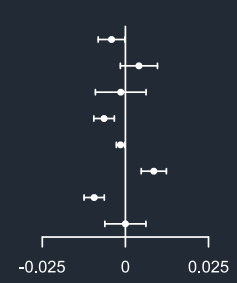
affix duration

Affix	β	p
-ness	0.005	**
-less	-0.006	–
pre-	-0.007	***
-ize	0.001	–
-ation	0	–
dis-	-0.002	**
un-	-0.003	***
in-	0.002	–



base duration

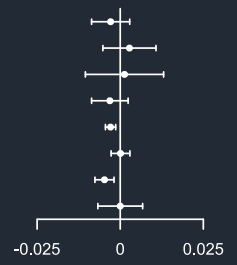
Affix	β	p
-ness	-0.004	–
-less	0.004	–
pre-	-0.001	–
-ize	-0.006	***
-ation	-0.001	*
dis-	0.009	***
un-	-0.009	***
in-	0	–



base frequency

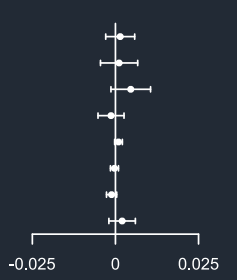
word duration

Affix	β	p
-ness	-0.003	–
-less	0.003	–
pre-	0.001	–
-ize	-0.003	–
-ation	-0.003	***
dis-	0	–
un-	-0.005	**
in-	0	–



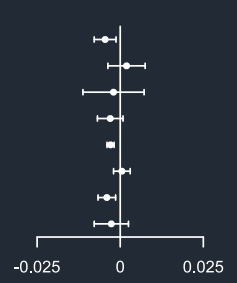
affix duration

Affix	β	p
-ness	0.001	–
-less	0.001	–
pre-	0.005	–
-ize	-0.001	–
-ation	0.001	–
dis-	0	–
un-	-0.001	–
in-	0.002	–



base duration

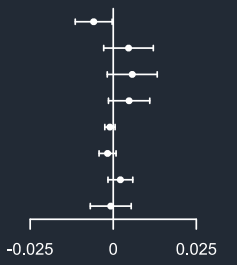
Affix	β	p
-ness	-0.005	**
-less	0.002	–
pre-	-0.002	–
-ize	-0.003	–
-ation	-0.003	***
dis-	0	–
un-	-0.004	**
in-	-0.003	–



relative frequency

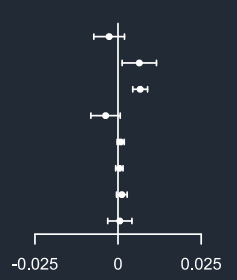
word duration

Affix	β	p
-ness	-0.006	*
-less	0.005	–
pre-	0.006	–
-ize	0.005	–
-ation	-0.001	–
dis-	-0.002	–
un-	0.002	–
in-	-0.001	–



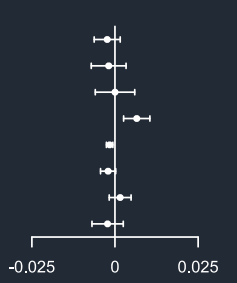
affix duration

Affix	β	p
-ness	-0.003	–
-less	0.006	–
pre-	0.007	***
-ize	-0.004	–
-ation	0.001	–
dis-	0	–
un-	0.001	–
in-	0.001	–



base duration

Affix	β	p
-ness	-0.002	–
-less	-0.002	–
pre-	0	–
-ize	0.007	**
-ation	-0.002	**
dis-	-0.002	–
un-	0.002	–
in-	-0.002	–

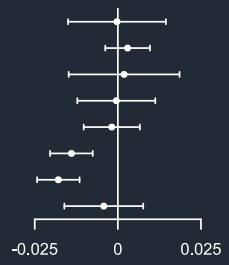


Category-internal frequency models QuakeBox

word frequency

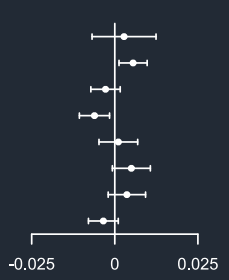
word duration

Affix	β	p
-ness	0	–
-ation	0.003	–
dis-	0.002	–
un-	0	–
-able	-0.002	–
-ity	-0.014	***
-ment	-0.018	***
re-	-0.004	–



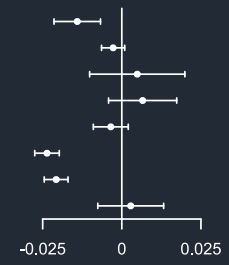
affix duration

Affix	β	p
-ness	0.003	–
-ation	0.006	*
dis-	-0.003	–
un-	-0.006	**
-able	0.001	–
-ity	0.005	–
-ment	0.004	–
re-	-0.003	–



base duration

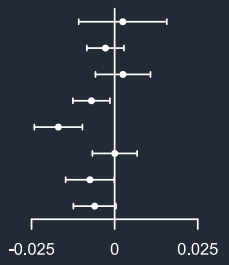
Affix	β	p
-ness	-0.014	***
-ation	-0.003	–
dis-	0.005	–
un-	0.007	–
-able	-0.003	–
-ity	-0.024	***
-ment	-0.021	***
re-	0.003	–



base frequency

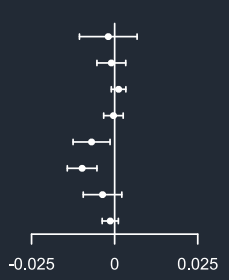
word duration

Affix	β	p
-ness	0.002	–
-ation	-0.003	–
dis-	0.003	–
un-	-0.007	*
-able	-0.017	***
-ity	0	–
-ment	-0.007	*
re-	-0.006	–



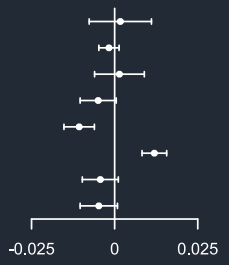
affix duration

Affix	β	p
-ness	-0.002	–
-ation	-0.001	–
dis-	0.001	–
un-	0	–
-able	-0.007	*
-ity	-0.01	***
-ment	-0.004	–
re-	-0.001	–



base duration

Affix	β	p
-ness	0.002	–
-ation	-0.002	–
dis-	0.001	–
un-	-0.005	–
-able	-0.011	***
-ity	0.012	***
-ment	-0.004	–
re-	-0.005	–



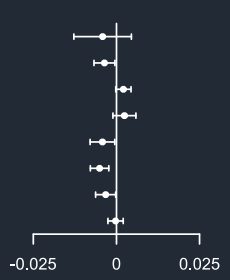
relative frequency

word duration

Affix	β	p
-ness	0.003	–
-ation	-0.003	–
dis-	0.002	–
un-	-0.011	**
-able	-0.007	–
-ity	0.008	***
-ment	0.007	**
re-	-0.004	–

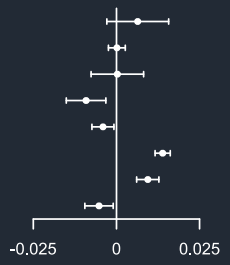
affix duration

Affix	β	p
-ness	-0.004	–
-ation	-0.004	*
dis-	0.002	–
un-	0.002	–
-able	-0.004	*
-ity	-0.005	***
-ment	-0.003	*
re-	0	–



base duration

Affix	β	p
-ness	0.006	–
-ation	0	–
dis-	0	–
un-	-0.009	**
-able	-0.004	*
-ity	0.014	***
-ment	0.009	***
re-	-0.005	*

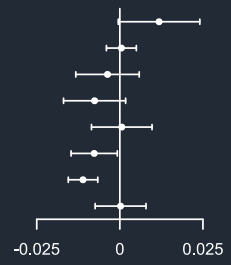


Category-internal frequency models ONZE

word frequency

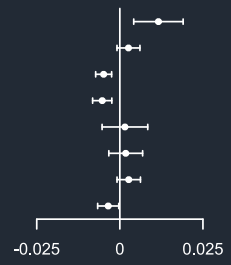
word duration

Affix	β	p
-ness	0.012	–
-ation	0	–
dis-	-0.004	–
un-	-0.008	–
-able	0.001	–
-ity	-0.008	–
-ment	-0.011	***
re-	0	–



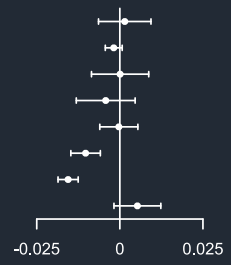
affix duration

Affix	β	p
-ness	0.012	**
-ation	0.003	–
dis-	-0.005	***
un-	-0.005	***
-able	0.002	–
-ity	0.002	–
-ment	0.003	–
re-	-0.003	–



base duration

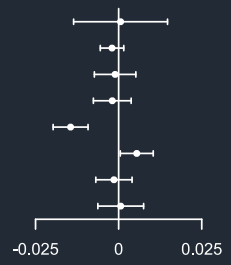
Affix	β	p
-ness	0.001	–
-ation	-0.002	–
dis-	0	–
un-	-0.004	–
-able	0	–
-ity	-0.01	***
-ment	-0.016	***
re-	0.005	–



base frequency

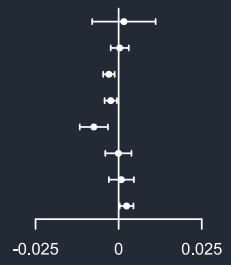
word duration

Affix	β	p
-ness	0.001	–
-ation	-0.002	–
dis-	-0.001	–
un-	-0.002	–
-able	-0.014	***
-ity	0.005	*
-ment	-0.001	–
re-	0.001	–



affix duration

Affix	β	p
-ness	0.002	–
-ation	0	–
dis-	-0.003	***
un-	-0.002	*
-able	-0.007	***
-ity	0	–
-ment	0.001	–
re-	0.002	*



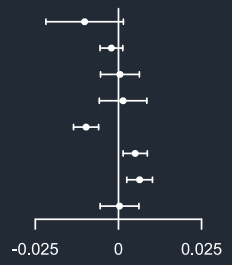
base duration

Affix	β	p
-ness	-0.003	–
-ation	-0.003	**
dis-	0.002	–
un-	-0.001	–
-able	-0.008	***
-ity	0.007	***
-ment	-0.001	–
re-	-0.001	–

relative frequency

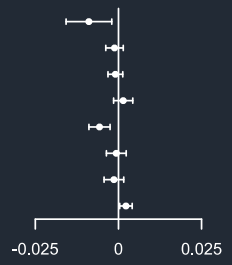
word duration

Affix	β	p
-ness	-0.01	–
-ation	-0.002	–
dis-	0	–
un-	0.001	–
-able	-0.01	***
-ity	0.005	**
-ment	0.006	**
re-	0	–



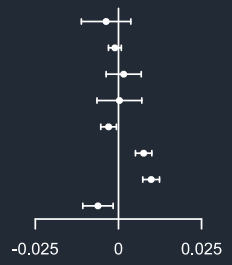
affix duration

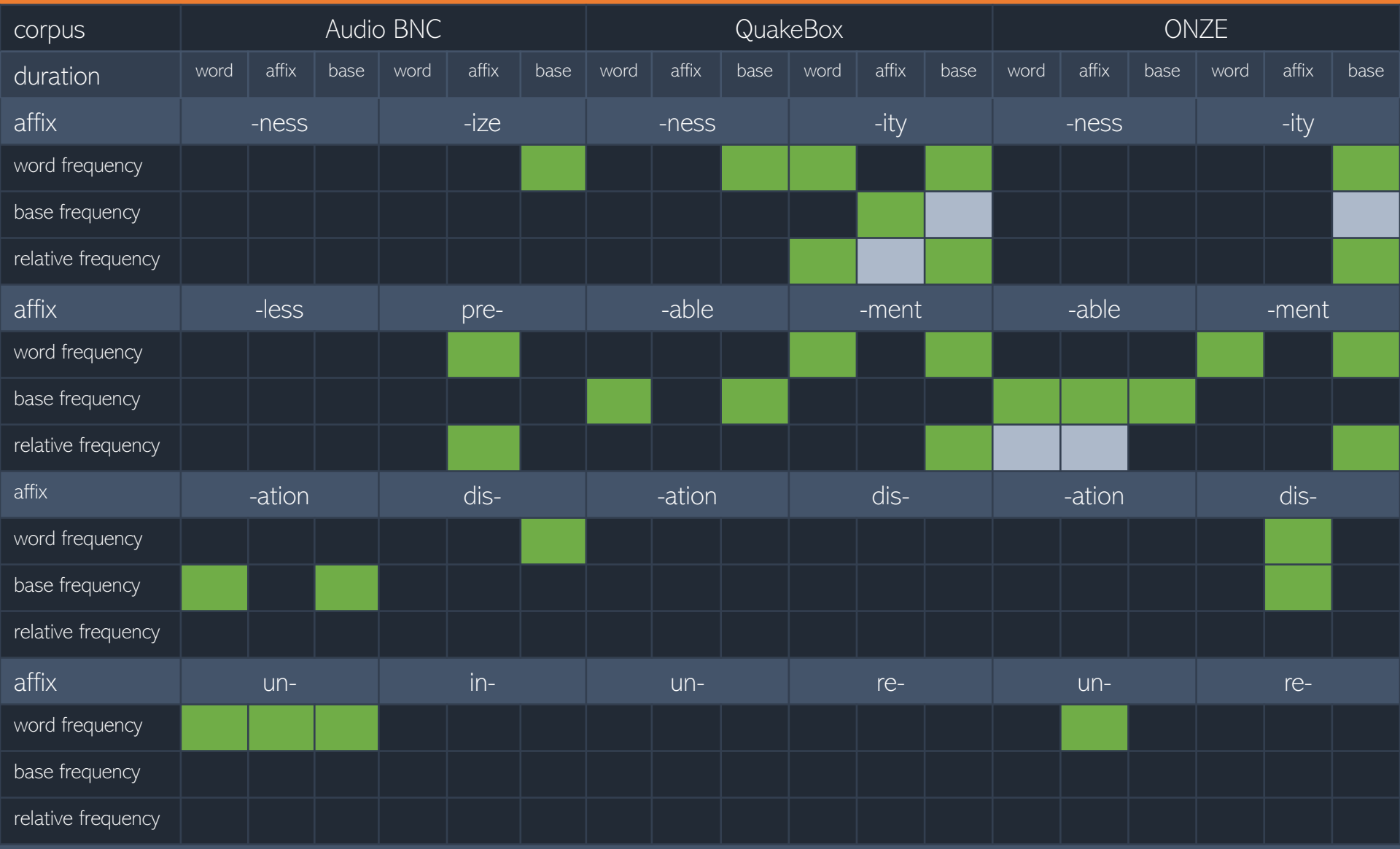
Affix	β	p
-ness	-0.009	*
-ation	-0.001	–
dis-	-0.001	–
un-	0.001	–
-able	-0.006	***
-ity	-0.001	–
-ment	-0.001	–
re-	0.002	*

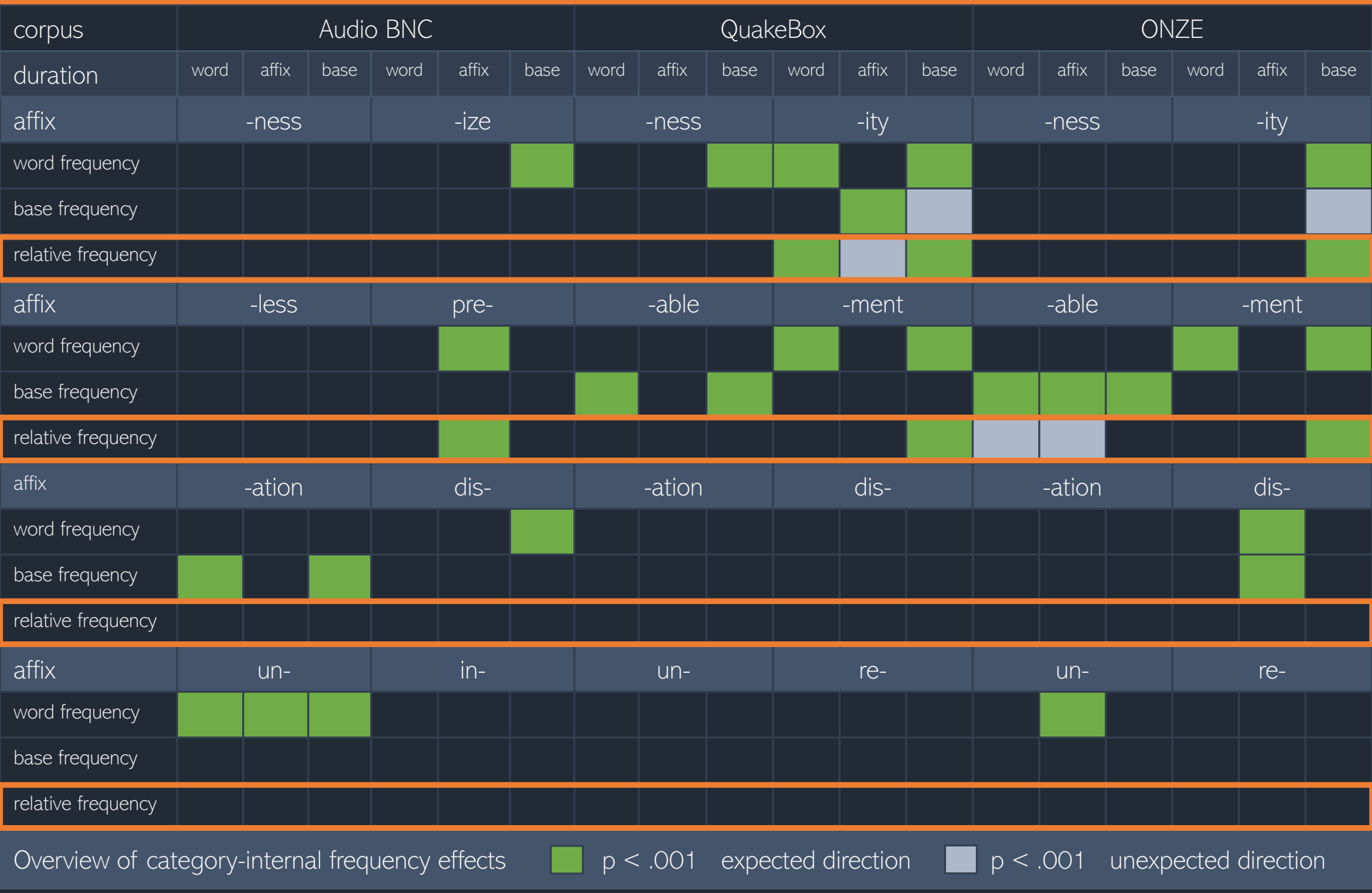


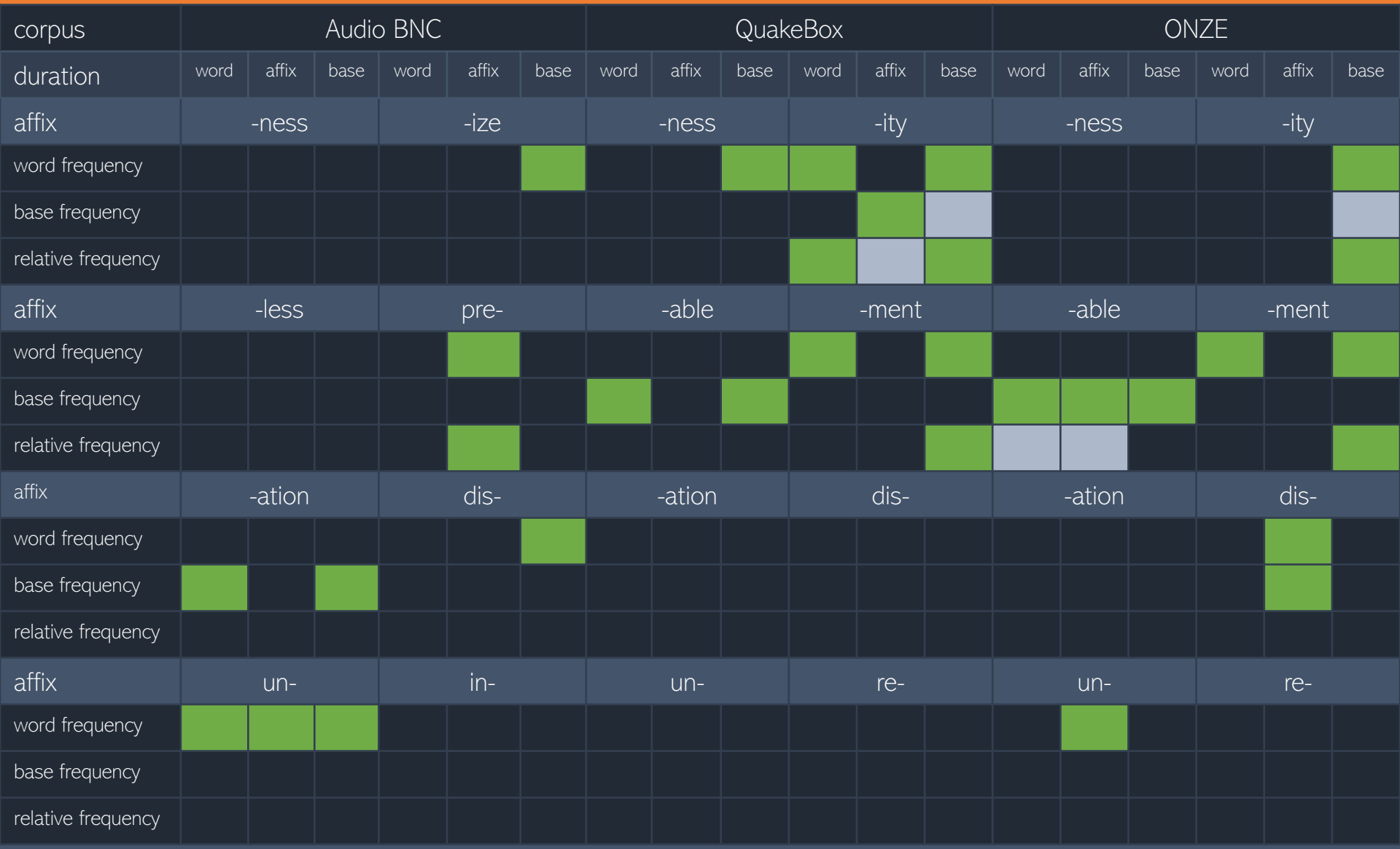
base duration

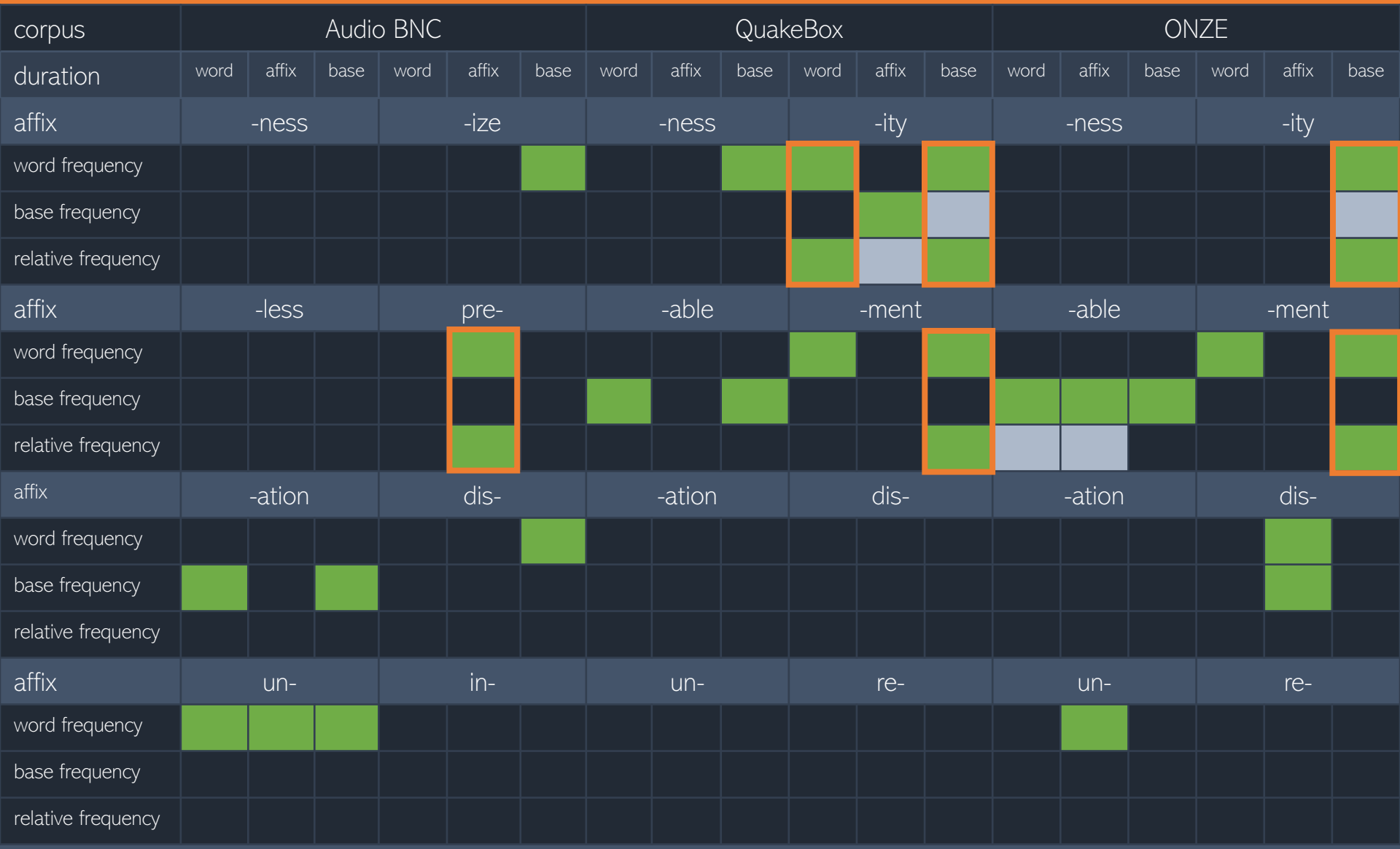
Affix	β	p
-ness	-0.004	–
-ation	-0.001	–
dis-	0.002	–
un-	0	–
-able	-0.003	*
-ity	0.008	***
-ment	0.01	***
re-	-0.006	**

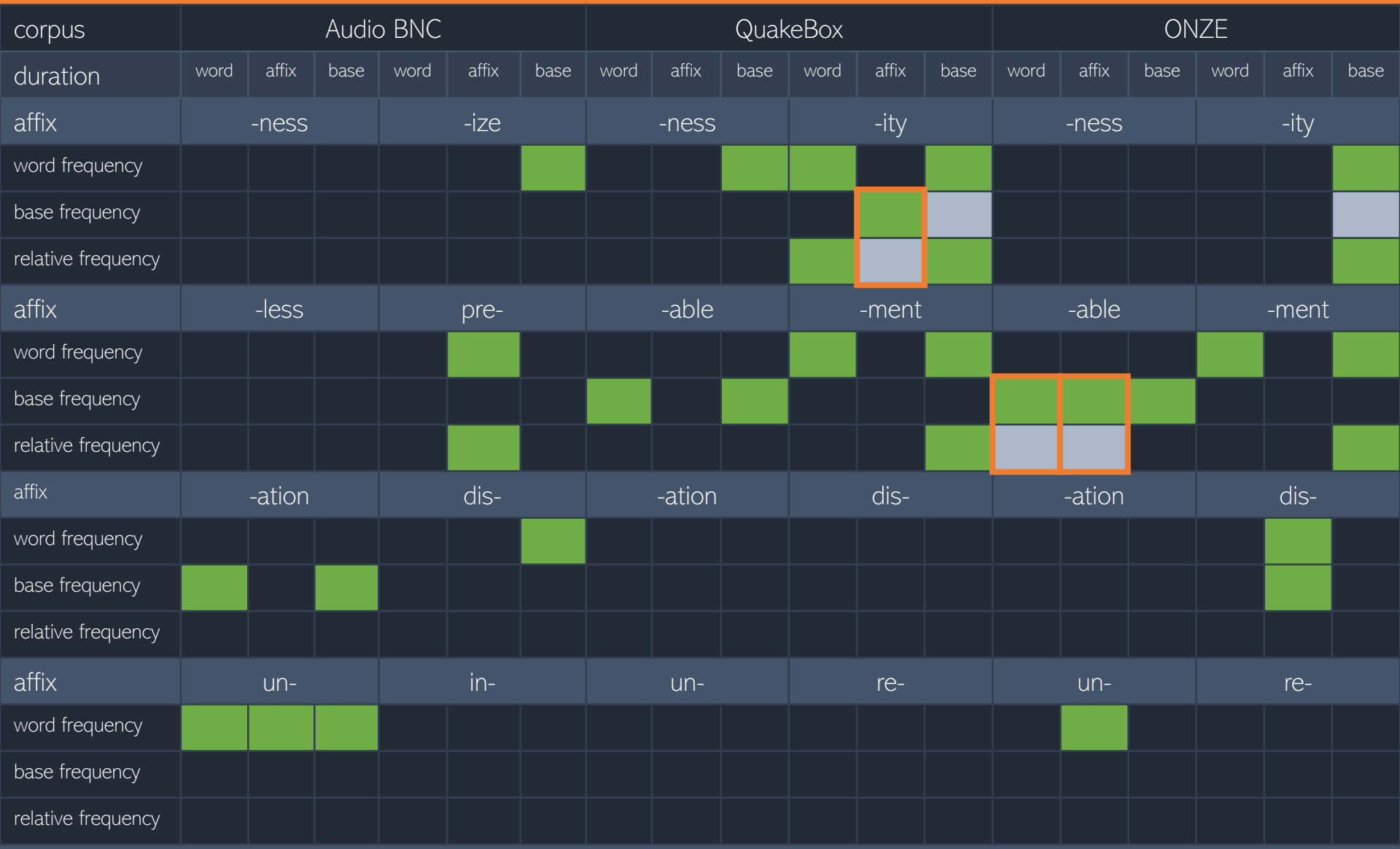


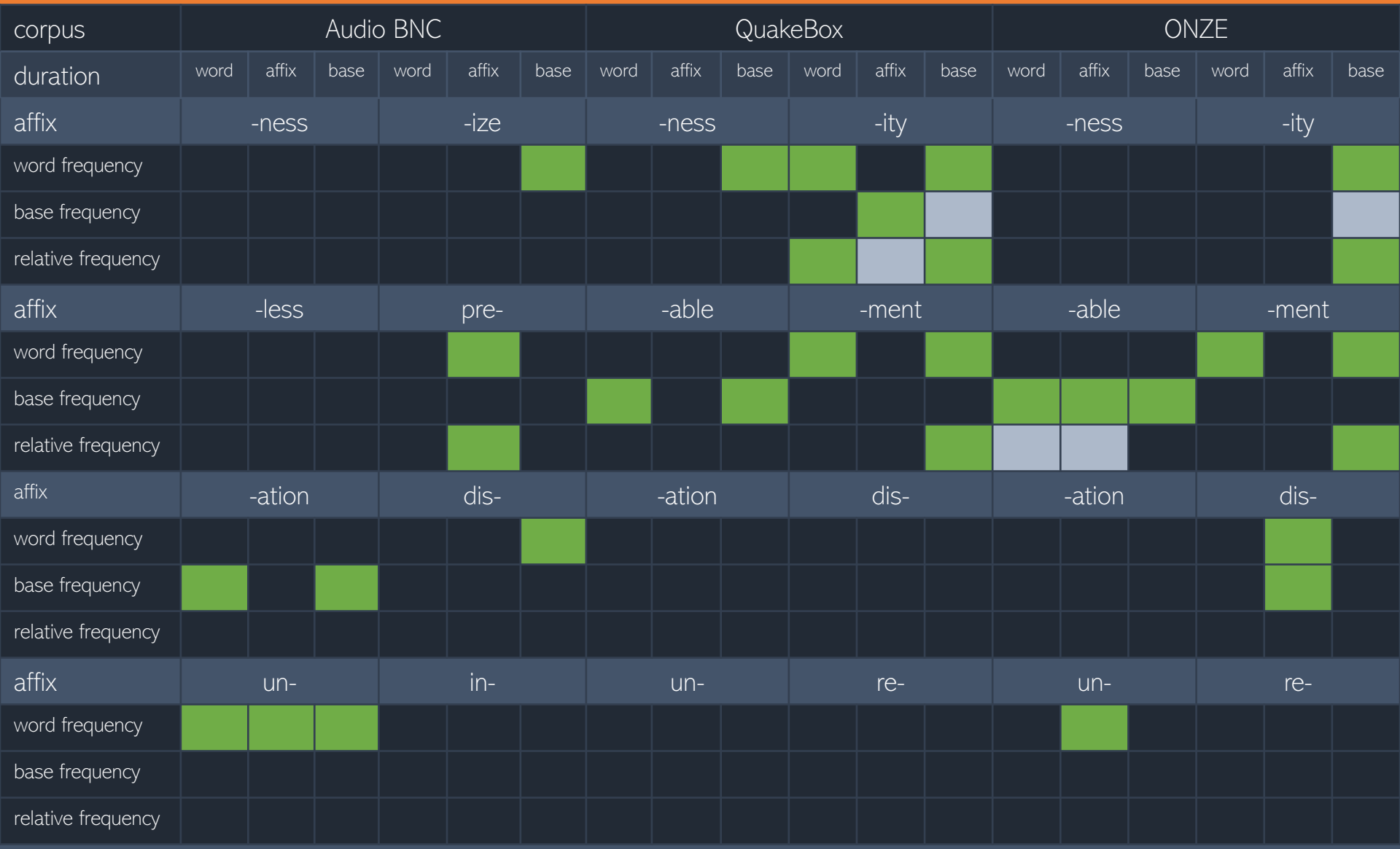

corpus
Audio BNC
QuakeBox
ONZE
duration
word affix base word affix base word affix base word affix base word affix base word affix base
affix
-ness -ize -ness -ity -ness -ity
word frequency
base frequency
relative frequency
affix
-less pre- -able -ment -able -ment
word frequency
base frequency
relative frequency
affix
-ation dis- -ation dis- -ation dis-
word frequency
base frequency
relative frequency
affix
un- in- un- re- un- re-
word frequency
base frequency
relative frequency
Overview of category-internal frequency effects
p < .001 expected direction
p < .001 unexpected direction


corpus
Audio BNC
QuakeBox
ONZE
duration
word affix base word affix base word affix base word affix base word affix base word affix base
affix
-ness -ize -ness -ity -ness -ity
word frequency
base frequency
relative frequency
affix
-less pre- -able -ment -able -ment
word frequency
base frequency
relative frequency
affix
-ation dis- -ation dis- -ation dis-
word frequency
base frequency
relative frequency
affix
un- in- un- re- un- re-
word frequency
base frequency
relative frequency
Overview of category-internal frequency effects
p < .001 expected direction
p < .001 unexpected direction


corpus
Audio BNC
QuakeBox
ONZE
duration
word affix base word affix base word affix base word affix base word affix base word affix base
affix
-ness -ize -ness -ity -ness -ity
word frequency
base frequency
relative frequency
affix
-less pre- -able -ment -able -ment
word frequency
base frequency
relative frequency
affix
-ation dis- -ation dis- -ation dis-
word frequency
base frequency
relative frequency
affix
un- in- un- re- un- re-
word frequency
base frequency
relative frequency
Overview of category-internal frequency effects
p < .001 expected direction
p < .001 unexpected direction


corpus
Audio BNC
QuakeBox
ONZE
duration
word affix base word affix base word affix base word affix base word affix base word affix base
affix
-ness -ize -ness -ity -ness -ity
word frequency
base frequency
relative frequency
affix
-less pre- -able -ment -able -ment
word frequency
base frequency
relative frequency
affix
-ation dis- -ation dis- -ation dis-
word frequency
base frequency
relative frequency
affix
un- in- un- re- un- re-
word frequency
base frequency
relative frequency
Overview of category-internal frequency effects
p < .001 expected direction
p < .001 unexpected direction


corpus
Audio BNC
QuakeBox
ONZE
duration
word
affix
base
affix
-ness
-ize
-ity
-less
pre-
-able
-ment
-ation
dis-
un-
in-
re-
word frequency
base frequency
relative frequency
Overview of category-internal frequency effects
p < .001 expected direction
p < .001 unexpected direction


corpus
Audio BNC
QuakeBox
ONZE
duration
word affix base word affix base word affix base word affix base word affix base word affix base
affix
-ness -ize -ness -ity -ness -ity
word frequency
base frequency
relative frequency
affix
-less pre- -able -ment -able -ment
word frequency
base frequency
relative frequency
affix
-ation dis- -ation dis- -ation dis-
word frequency
base frequency
relative frequency
affix
un- in- un- re- un- re-
word frequency
base frequency
relative frequency
Overview of category-internal frequency effects
p < .001 expected direction
p < .001 unexpected direction

Informativity

Semantic information load score

5-point Likert scales coded for:

- clearness of semantic meaning
- type of base: free vs. bound root
- semantic transparency
- productivity

Affix-specific semantic segmentability hierarchy

H: The higher the semantic information load, the longer the duration.

Conditional affix probability C_{aff}

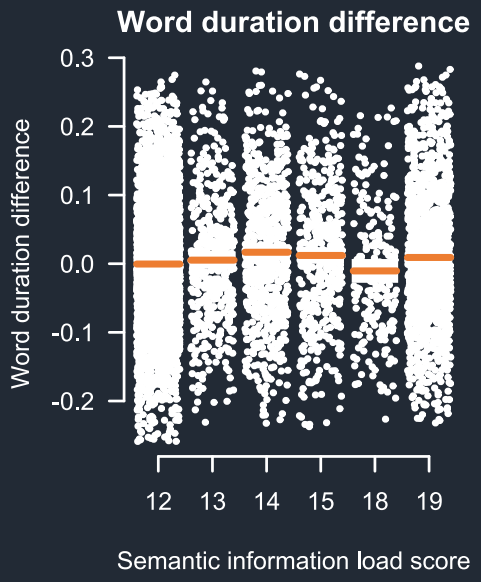
Affix probability given preceding word:

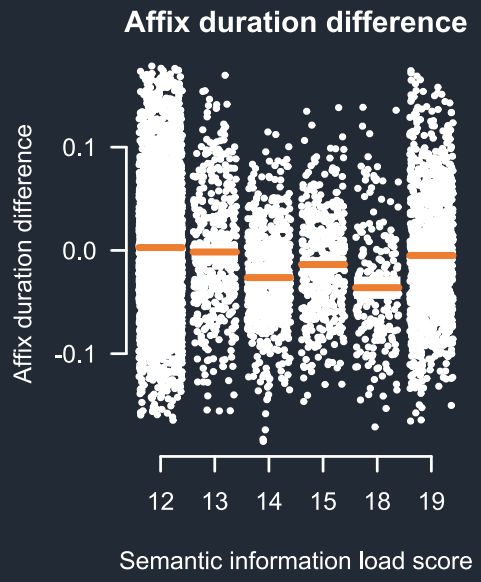
SUFFIX EXAMPLE		PREFIX EXAMPLE		
A	B	A	B	C
random	*ize*	*her*	*pre-*	…

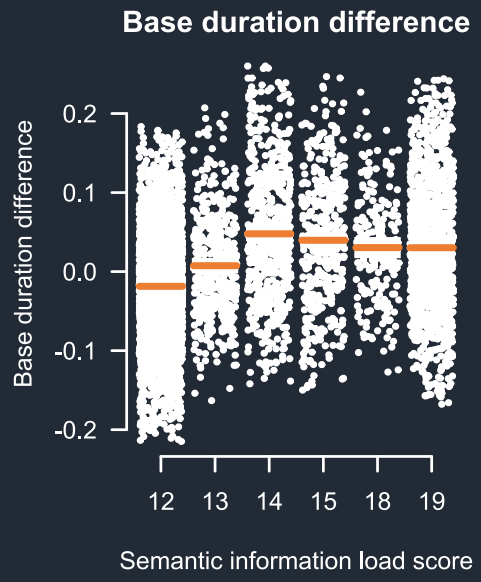
$$C_{aff} = \frac{Freq(AB)}{Freq(A)}$$

H: The higher the conditional affix probability, the shorter the duration.

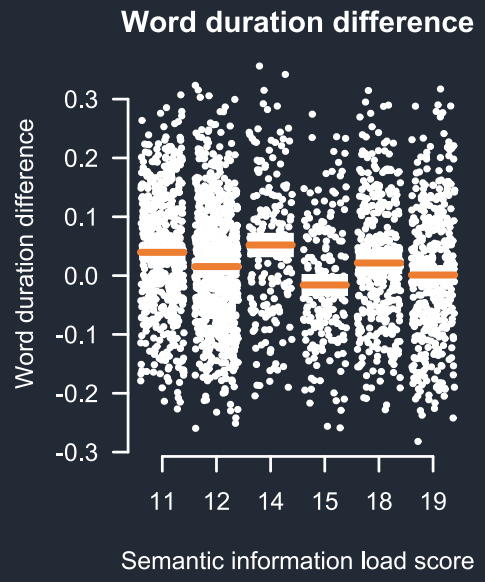
Semantic information load score

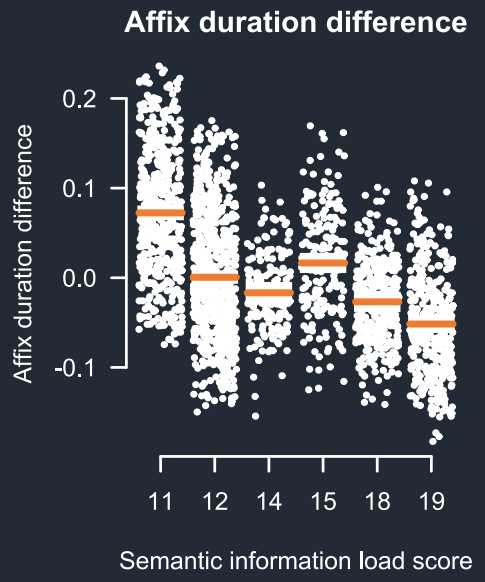


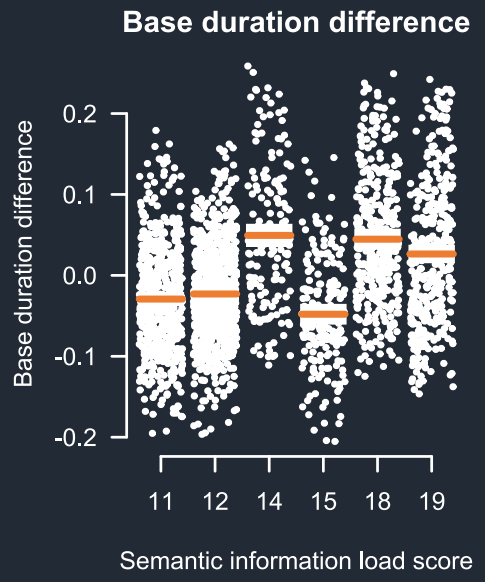




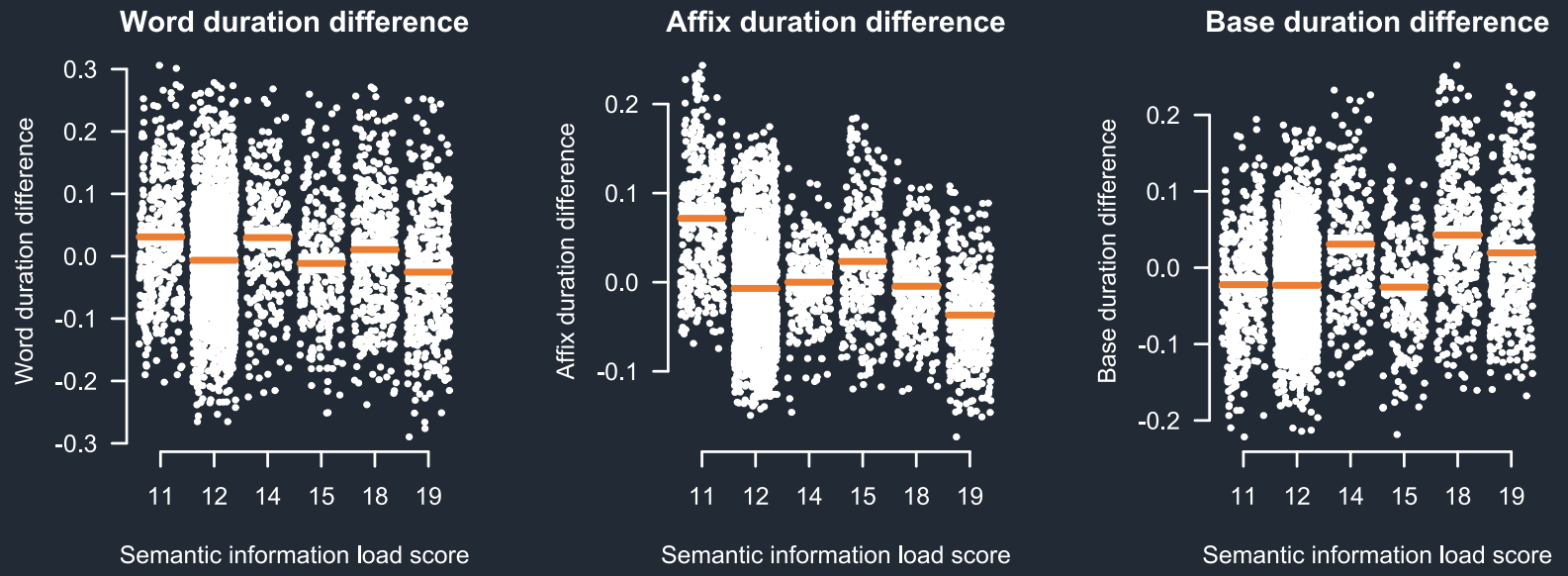
Semantic information load score







Semantic information load score



Conditional affix probability C_{aff}

corpus	Audio BNC						QuakeBox						ONZE					
duration	word	affix	base	word	affix	base	word	affix	base	word	affix	base	word	affix	base	word	affix	base
affix	-ness			-ize			-ness			-ity			-ness			-ity		
C_{aff}																		
affix	-less			pre-			-able			-ment			-able			-ment		
C_{aff}																		
affix	-ation			dis-			-ation			dis-			-ation			dis-		
C_{aff}																		
affix	un-			in-			un-			re-			un-			re-		
C_{aff}																		

$p < .001$ negative effect $p < .001$ positive effect