HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

Lexical storage and morphological segmentability in speech production

New evidence from English derivational affixes

Simon David Stein
Ingo Plag

DFG Deutsche Forschungsgemeinschaft
FOR2373

Frequency and duration

Lexical frequency

How often does a word occur in a language?

Acoustic duration

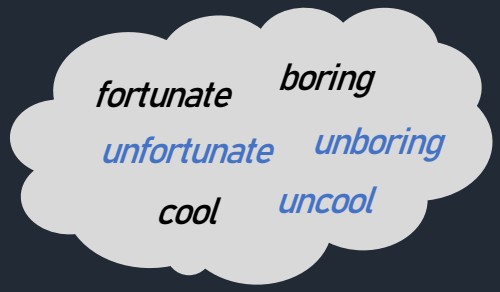
How long do we pronounce linguistic units?

Usual assumption:

The higher the frequency, the shorter the duration of linguistic units such as words, bases, and affixes.

Storage in the mental lexicon

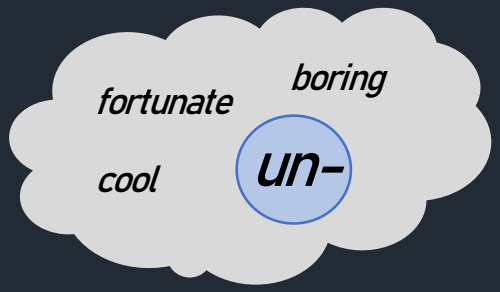
Whole-word storage



complex words are stored unanalyzed

durations will be shorter the higher the word frequency

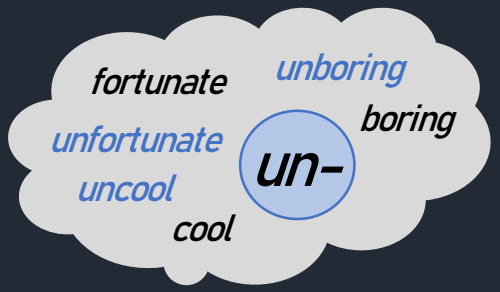
Compositional models



morphemes are stored separately

durations will be shorter the higher the base frequency

Dual-route models

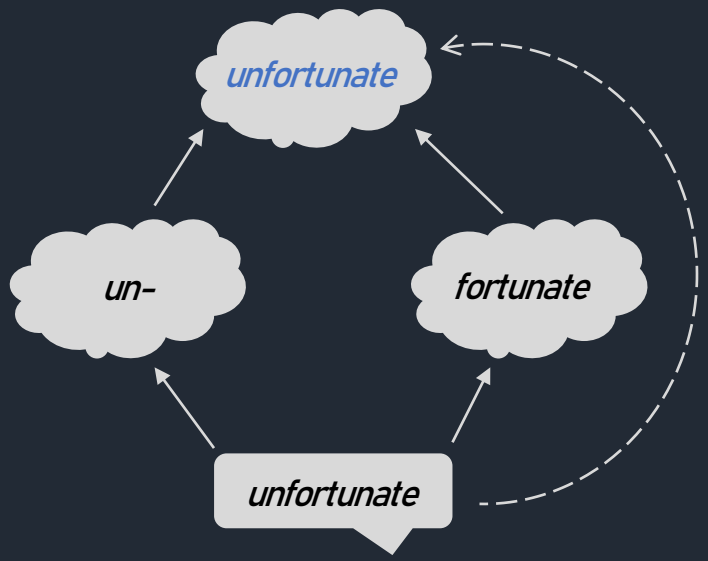


both morphemes and complex words are stored

durations will be shorter the lower the relative frequency

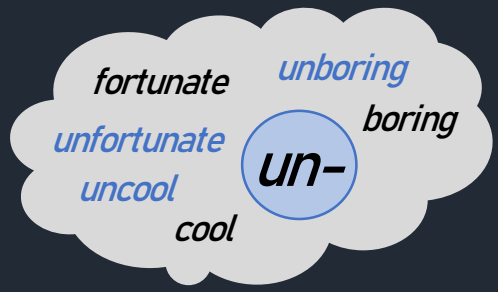
Caselli et al. 2016, examples from Hay 2007: 40–41

Segmentability



adapted from
Hay 2001: 1045

Dual-route models

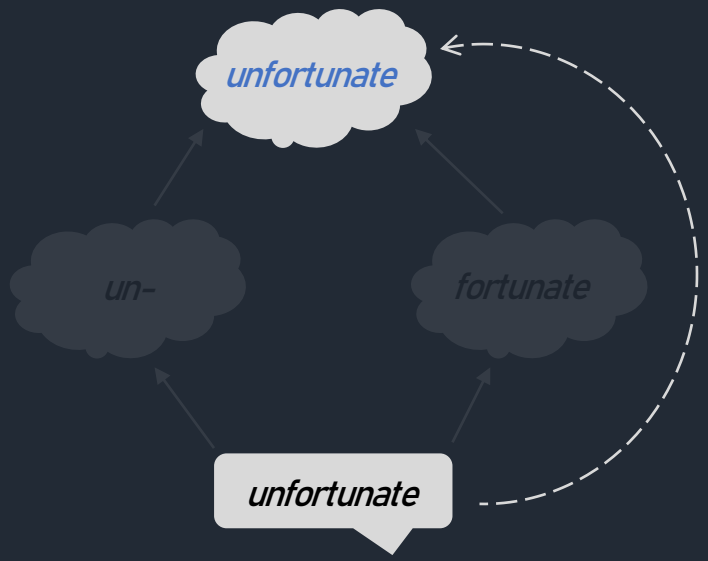


both morphemes and
complex words are stored

durations will be shorter
the lower the
relative frequency

Caselli et al. 2016, examples from Hay 2007: 40–41, frequencies from COCA, Davies 2008

Segmentability



adapted from
Hay 2001: 1045

Dual-route models

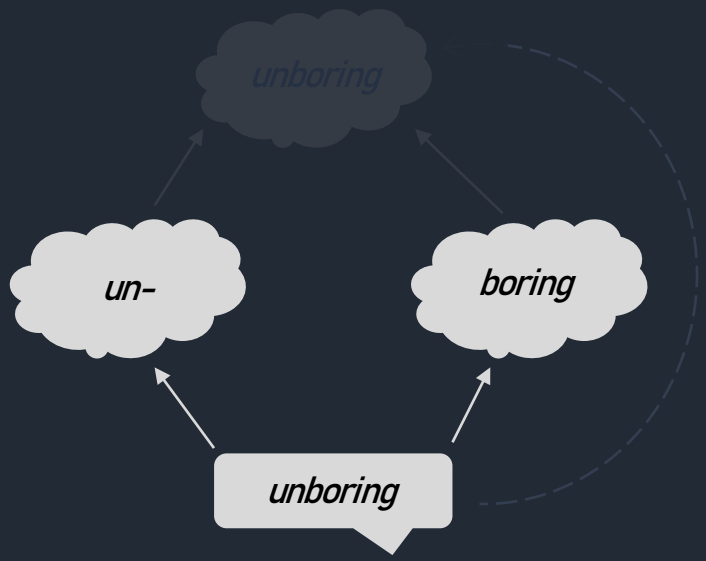
both morphemes and complex words are stored

durations will be shorter the lower the relative frequency

Word	Frequency	Segmentability	Prediction
fortunate	6000	low	shorter duration
unfortunate	6915		
boring	7483	high	longer duration
unboring	4		

Caselli et al. 2016, examples from Hay 2007: 40–41, frequencies from COCA, Davies 2008

Segmentability



adapted from
Hay 2001: 1045

Dual-route models

both morphemes and complex words are stored

durations will be shorter the lower the relative frequency

Word	Frequency	Segmentability	Prediction
fortunate	6000	low	shorter duration
unfortunate	6915		
boring	7483	high	longer duration
unboring	4		

Caselli et al. 2016, examples from Hay 2007: 40–41, frequencies from COCA, Davies 2008

Previous research

Caselli et al. 2016

- inflectional suffixes *-ing*, *-ed*, and *-s*
- evidence for both whole-word storage and composition
  - higher base frequency → shorter word duration
  - higher word frequency → shorter word duration

Hay 2003, 2007

- segmentability effects for *un-* and *-ly*

Plag and Ben Hedia 2018

- segmentability effects for *un-* and *dis-*
- null effects for negative *in-*, locative *in-*, and *-ly*

Contradictory evidence:
Why do the frequency measures sometimes show and sometimes not show effects?

Caselli et al. 2016, Hay 2003, 2007, Plag and Ben Hedia 2018

Present study

Hypothesis 1

Higher word frequency → shorter duration of word, base, and affix

Hypothesis 2

Higher base frequency → shorter duration of word, base, and affix

Hypothesis 3

Higher relative frequency → longer duration of word, base, and affix

≈ more segmentability

Data and measurement

Data collection

- AudioBNC
- Forced Alignment
- Praat textgrids
- manual cleaning of results

Affixes	N
-ness	364
-less	216
pre-	118
-wise	289
-ize	476
-ation	3979

Modeling

- multiple linear regression in R using lm-function
- variable transformations
- trimming of datasets
- backwards exclusion of non-significant variables

Responses

- word duration
- affix duration
- base duration

Predictors

- word frequency
- base frequency
- relative frequency

Covariates

- speech rate
- number of syllables
- biphone probability sum
- bigram frequency

- separate models for durations and frequencies: 54 models

Coleman et al. 2012, Boersma and Weenik 2014, R Core Team 2017, Davies 2008, Vitevitch and Luce 2004

Frequency and segmentability effects

affix	pre-			-ness			-ize		
duration	word	affix	base	word	affix	base	word	affix	base
word frequency									
base frequency									
relative frequency									

affix	-wise			-less			-ation		
duration	word	affix	base	word	affix	base	word	affix	base
word frequency									
base frequency									
relative frequency									

p < .001 expected direction

p < .001 unexpected direction

Are the differences related to ...

Results

Prefixes vs. suffixes

affix	pre-			-ness			-ize		
duration	word	affix	base	word	affix	base	word	affix	base
word frequency									
base frequency									
relative frequency									

affix	-wise			-less			-ation		
duration	word	affix	base	word	affix	base	word	affix	base
word frequency									
base frequency									
relative frequency									

p < .001 expected direction

p < .001 unexpected direction

Are the differences related to ... the type of affix?

Prefixes vs. suffixes

affix	pre-			-ness			-ize		
duration	word	affix	base	word	affix	base	word	affix	base
word frequency									
base frequency									
relative frequency									

affix	-wise			-less			-ation		
duration	word	affix	base	word	affix	base	word	affix	base
word frequency									
base frequency									
relative frequency									

p < .001 expected direction

p < .001 unexpected direction

Are the differences related to ... the type of affix?

Results

Affix length

affix	pre-			-ness			-ize		
duration	word	affix	base	word	affix	base	word	affix	base
word frequency									
base frequency									
relative frequency									

affix	-wise			-less			-ation		
duration	word	affix	base	word	affix	base	word	affix	base
word frequency									
base frequency									
relative frequency									

p < .001 expected direction

p < .001 unexpected direction

Are the differences related to ...

the type of affix? ✗

the affix length?

Results

Affix length

affix	pre-			-ness			-ize		
duration	word	affix	base	word	affix	base	word	affix	base
word frequency									
base frequency									
relative frequency									

affix	-wise			-less			-ation		
duration	word	affix	base	word	affix	base	word	affix	base
word frequency									
base frequency									
relative frequency									

p < .001	expected direction	Are the differences related to …	the type of affix?	✗
p < .001	unexpected direction		the affix length?	

Manual resegmentation

affix	pre-			-ness			-ize		
duration	word	affix	base	word	affix	base	word	affix	base
word frequency									
base frequency									
relative frequency									

affix	-wise			-less			-ation		
duration	word	affix	base	word	affix	base	word	affix	base
word frequency									
base frequency									
relative frequency									

p < .001 expected direction

p < .001 unexpected direction

Are the differences related to ...

the type of affix? ×

the affix length? ×

the segmentation?

Manual resegmentation

affix	pre-			-ness			-ize		
duration	word	affix	base	word	affix	base	word	affix	base
word frequency									
base frequency									
relative frequency									

affix	-wise			-less			-ation		
duration	word	affix	base	word	affix	base	word	affix	base
word frequency									
base frequency									
relative frequency									

p < .001 expected direction

p < .001 unexpected direction

Are the differences related to ...

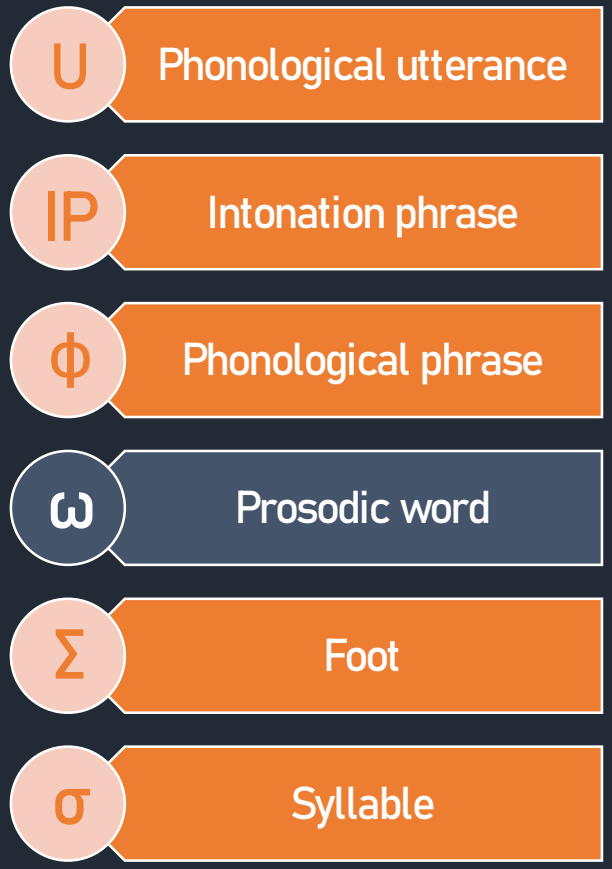
the type of affix? ×

the affix length? ×

the segmentation? ×

Type of prosodic integration

The prosodic hierarchy



Some pword-diagnostics

- onset or coda conditions, 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.

Hildebrandt 2015, Raffelsiefen 1999, 2007

Type of prosodic integration

more integration

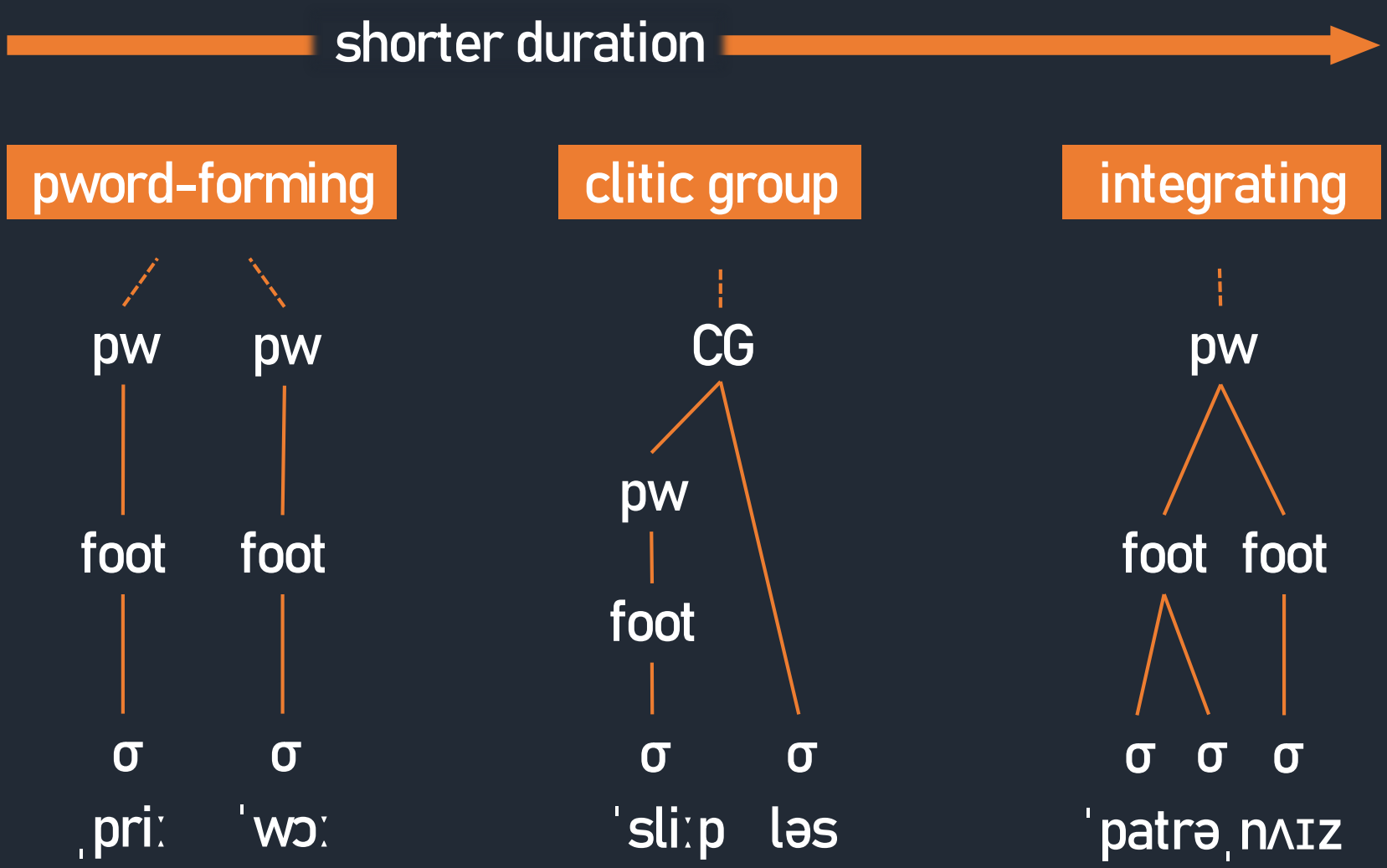
pword-forming

clitic group

integrating

following Raffelsiefen 1999

Type of prosodic integration



following Raffelsiefen 1999

Type of prosodic integration

category	prosodic word			clitic group			integrates		
affix	pre-			-ness			-ize		
duration	word	affix	base	word	affix	base	word	affix	base
word frequency									
base frequency									
relative frequency									

affix	-wise			-less			-ation		
duration	word	affix	base	word	affix	base	word	affix	base
word frequency									
base frequency									
relative frequency									

p < .001 expected direction

p < .001 unexpected direction

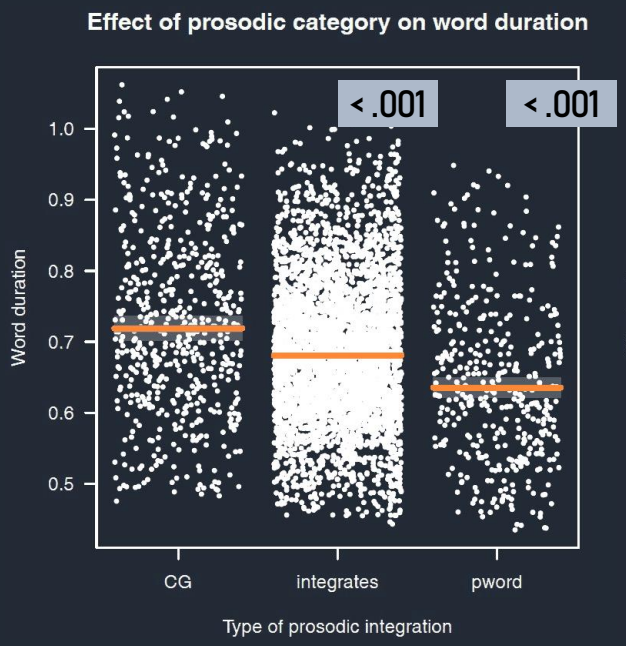
Are the differences related to ...

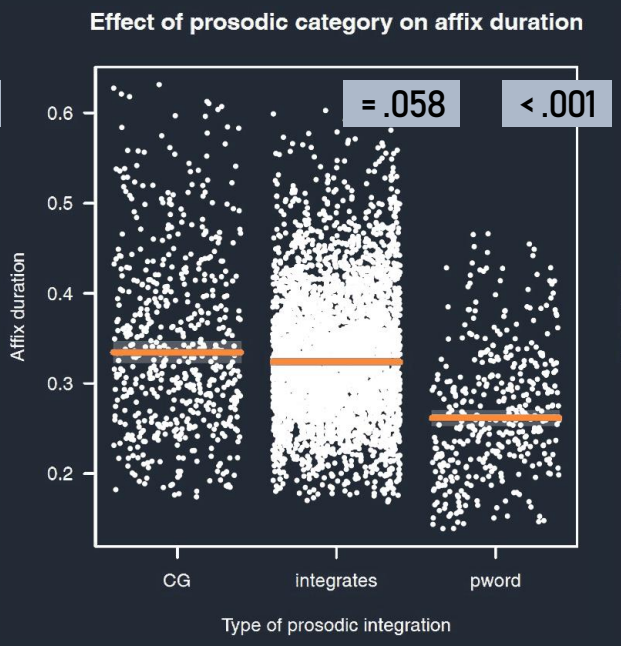
- the type of affix? ×
- the affix length? ×
- the segmentation? ×
- prosodic structure? ×

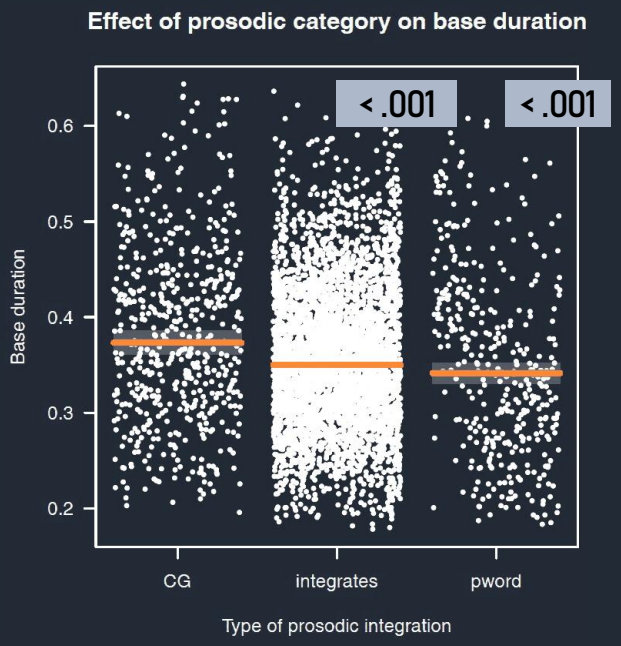
Type of prosodic integration

Meta-model including all affixes

- Additional predictor: type of prosodic integration
- Additional covariate: number of timing slots
- N = 5450
- This does not support the predictions of pword integration.







Summary

In sum, we have a mixed picture.

- Some results are in line with Caselli et al. 2016:
  - All three frequency measures can independently predict duration.
  - This is evidence for both types of storage in the mental lexicon, as well as for segmentability effects.

- However, there are also null effects, which require explanation.
  - So far, we cannot attribute the differences to:
    - the domain of durational measurement (word, affix, base)
    - the type of affix (prefix, suffix)
    - the prosodic category (pword, clitic group, integrating).

Discussion

Our findings imply that …

- morphological structure can at least partly influence the phonetic output.
- models that prohibit post-lexical access of morphological information (e.g. Kiparsky 1982, Levelt et al. 1999, Bermúdez-Otero 2018) should be revised.
- we need to investigate further factors that might cause frequency effects to surface or to not surface.

Kiparsky 1982, Levelt et al. 1999, Bermúdez-Otero 2018

References

Thank you for listening.

- Ben Hedia, Sonia. 2018. *Gemination and Degemination in English Affixation: Investigating the Interplay between Morphology, Phonology and Phonetics*. Ph.D. dissertation: Heinrich-Heine-Universität Düsseldorf.
- Bermúdez-Otero, Ricardo. 2018. Stratal Phonology. In S. J. Hannahs & Anna Bosch (eds.), *Routledge handbook of phonological theory*, 100–143. London: Routledge.
- Boersma, Paul & David J. M. Weenik. 2014. Praat: Doing phonetics by computer (Version 5.4.04). Computer program. http://www.praat.org/.
- Caselli, Naomi K., Michael K. Caselli, and 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.

References

Thank you for listening.

- Coleman, John, Ladan Baghai-Ravary, John Pybus & Sergio Grau. 2012. *Audio BNC: The audio edition of the Spoken British National Corpus*. Phonetics Laboratory, 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/.
- 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.

References

Thank you for listening.

- Hildebrandt, Kristine A. 2015. The prosodic word. In John R Taylor (ed.), *The Oxford Handbook of the Word*. Oxford: Oxford University Press.
- 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.
- 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.
- 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*. Berlin & New York: de Gruyter Mouton.

References

Thank you for listening.

- 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*. 133–201. Amsterdam, Philadelphia: 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)*, Fréjus, 15-18 September 2005, 209–268.
- R Core Team 2017. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing Vienna, Austria. http://www.R-project.org/.

References

Thank you for listening.

› Vitevitch, Michael S., & Luce, Paul A. 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.

Appendix


Log word frequency
Log base frequency
Log relative frequency
Word duration
Affix duration
Base duration
-ation
-ness
-ize
pre-
-wise
Effect size comparison between affixes. Effects with p > .001 omitted.