Heinrich Heine Universität Düsseldorf

# Lexical storage and morphological segmentability effects on the production of English derivatives

Simon David Stein
Ingo Plag

DFG Deutsche Forschungsgemeinschaft
FOR2373

## Frequency and duration

## Frequency and duration

### Lexical frequency

How often does a linguistic unit occur in a language?

### Acoustic duration

How long do we pronounce linguistic units?

## Frequency and duration

### Lexical frequency

How often does a linguistic unit occur in a language?

### Acoustic duration

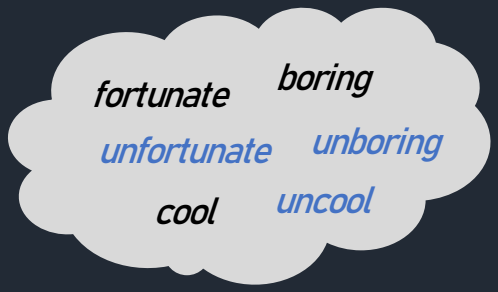
How long do we pronounce linguistic units?

higher → shorter

## Storage in the mental lexicon

## Storage in the mental lexicon

### Whole-word storage



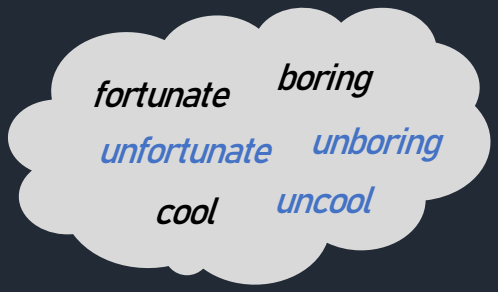

complex words are stored unanalyzed



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

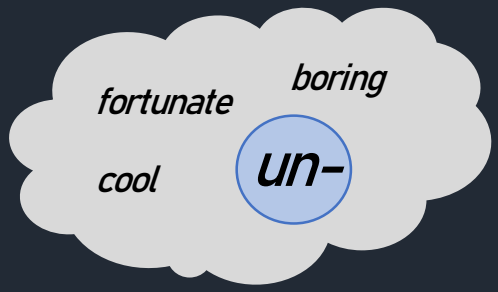
## Storage in the mental lexicon

### Whole-word storage




complex words are stored unanalyzed

### Compositional models



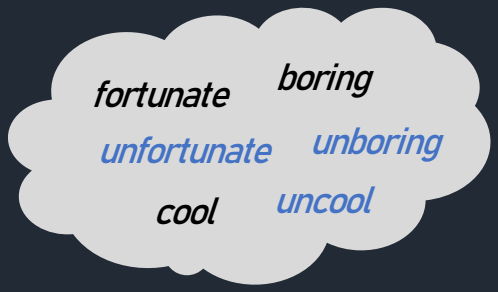

morphemes are stored separately



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

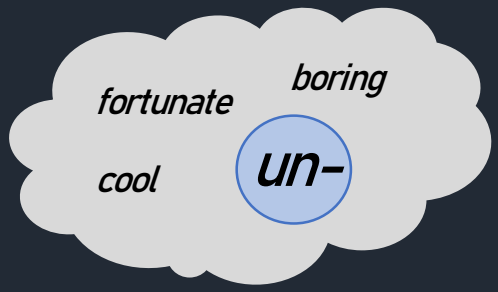
## Storage in the mental lexicon

### Whole-word storage



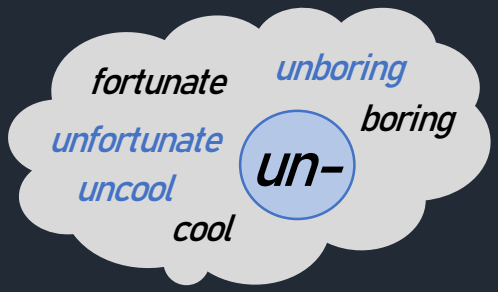

complex words are stored unanalyzed

### Compositional models




morphemes are stored separately

### Dual-route models



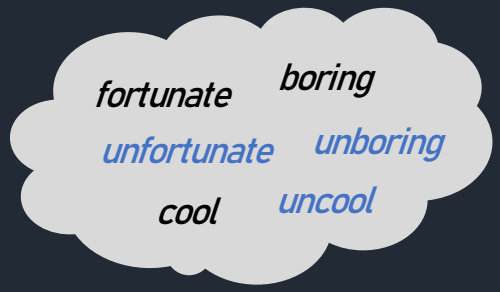

both morphemes and complex words are stored



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

## Storage in the mental lexicon

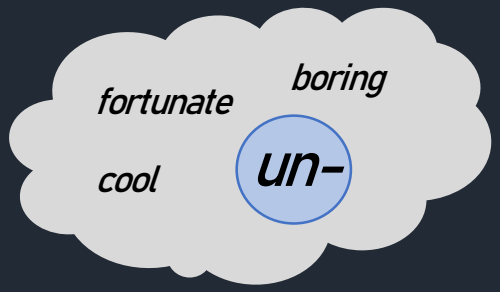
### Whole-word storage




complex words are stored unanalyzed

durations will be shorter the higher the word frequency

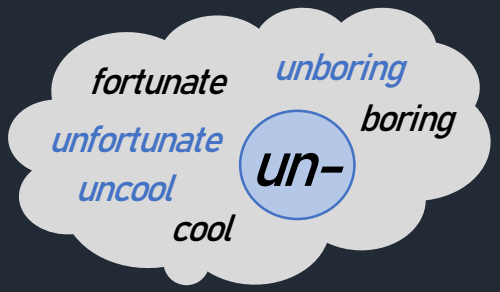
### Compositional models




morphemes are stored separately

durations will be shorter the higher the base frequency

### Dual-route models




both morphemes and complex words are stored

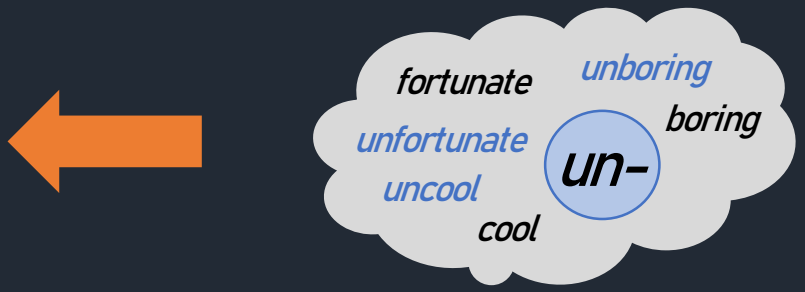
durations will be shorter the lower the relative frequency



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

## Segmentability

Dual-route models

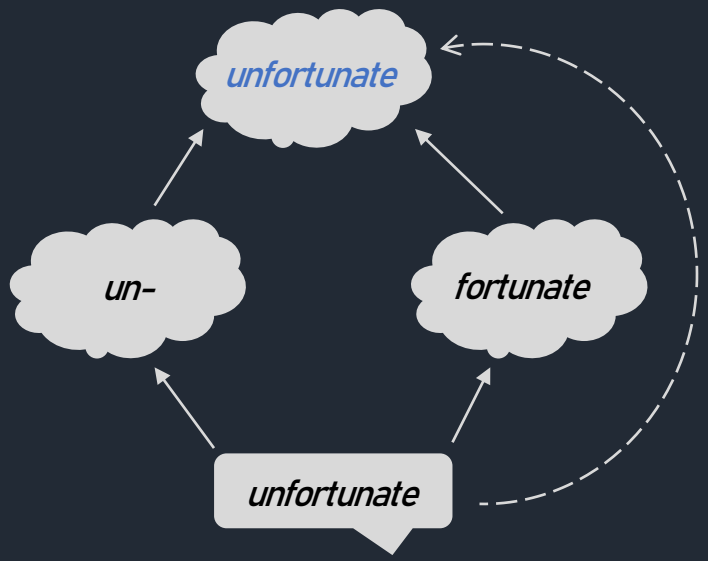



both morphemes and complex words are stored

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

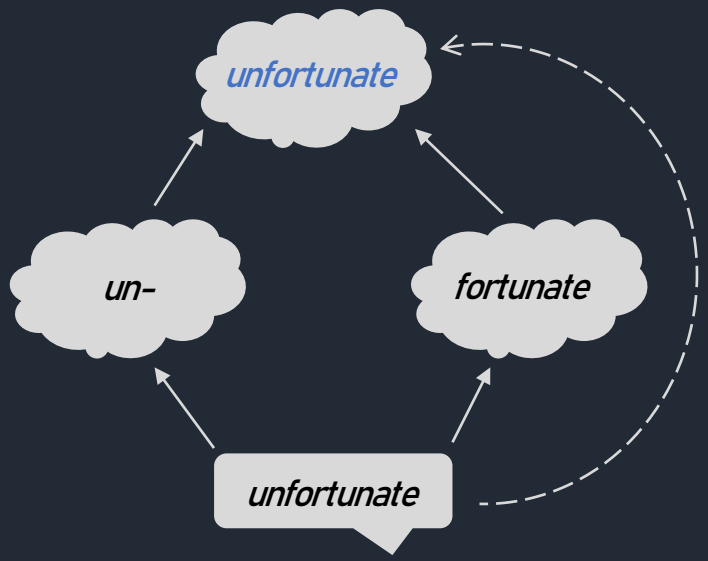

both morphemes and complex words are stored

durations will be shorter the lower the relative frequency



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

## Segmentability




adapted from
Hay 2001: 1045

Dual-route models


both morphemes and complex words are stored

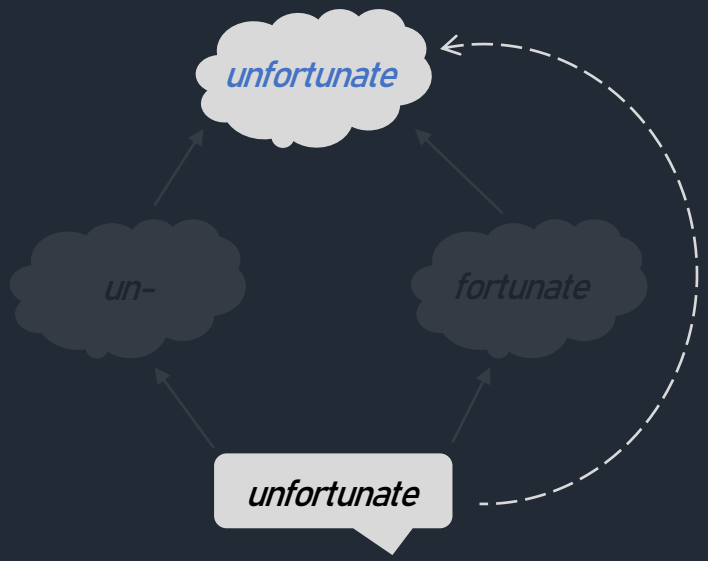
durations will be shorter the lower the relative frequency

| 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

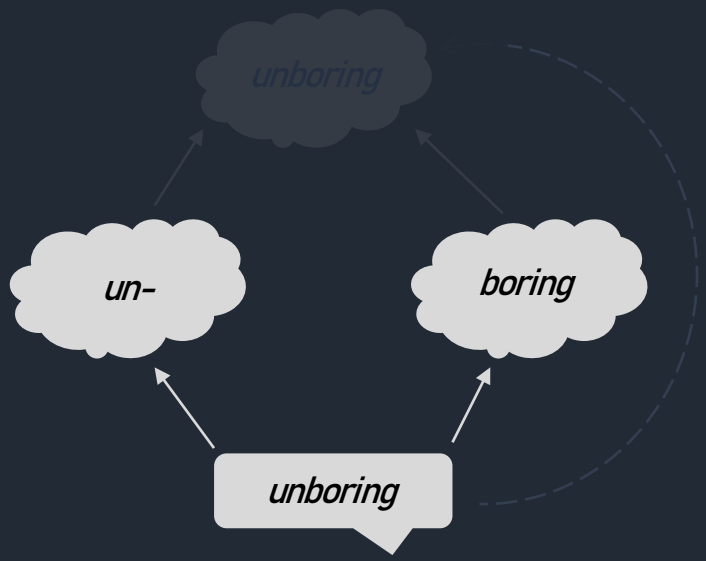
durations will be shorter the lower the relative frequency

| 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

## Segmentability




adapted from
Hay 2001: 1045

Dual-route models


both morphemes and complex words are stored

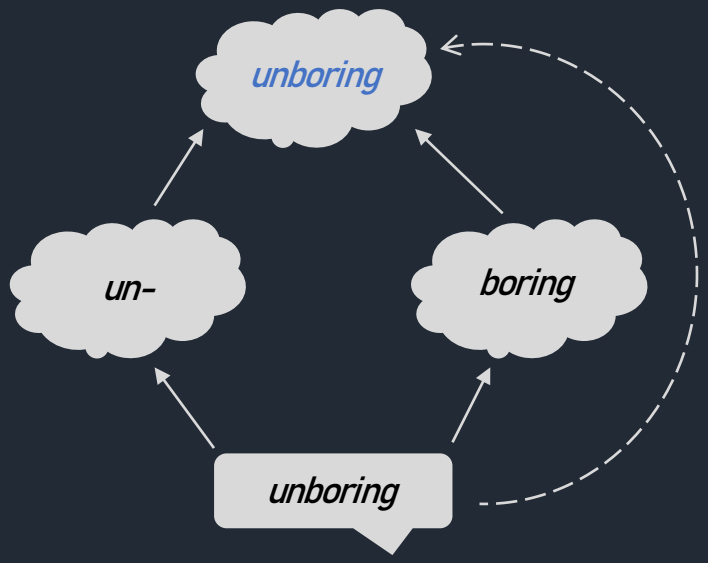
durations will be shorter the lower the relative frequency

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



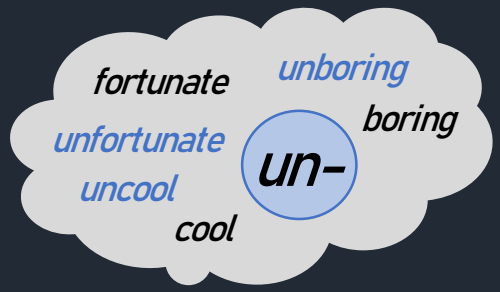
Caselli et al. 2016, examples from Hay 2007: 40–41, frequencies from COCA, Davies 2008

## Segmentability




adapted from
Hay 2001: 1045

### Dual-route models




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

both morphemes and complex words are stored

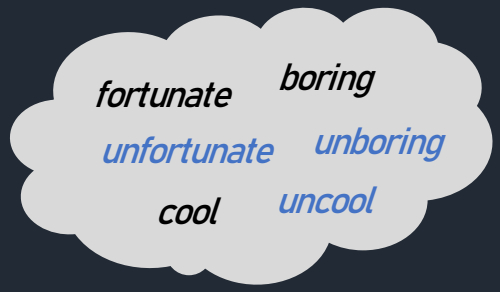
durations will be shorter the lower the relative frequency



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

## Storage in the mental lexicon

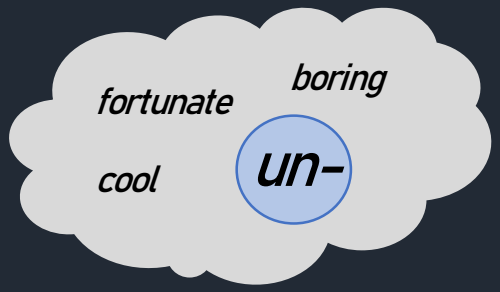
### Whole-word storage




complex words are stored unanalyzed

durations will be shorter the higher the word frequency

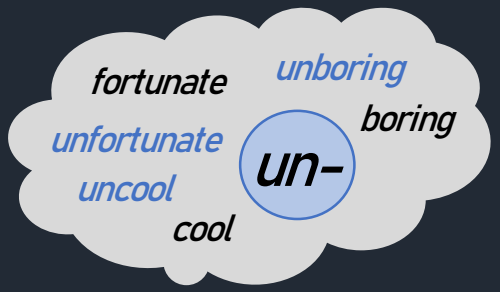
### Compositional models




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

## Previous research

## 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

## 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*



Caselli et al. 2016, Hay 2003, 2007

## 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*



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

## 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

## Present study

### Hypothesis 1

Higher word frequency → shorter duration

## Present study

### Hypothesis 1

Higher word frequency → shorter duration

### Hypothesis 2

Higher base frequency → shorter duration

## Present study

### Hypothesis 1

Higher word frequency → shorter duration

### Hypothesis 2

Higher base frequency → shorter duration

### Hypothesis 3

Higher relative frequency → longer duration

≈ more segmentability

## 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 and measurement

### Data collection

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



Coleman et al. 2012, Boersma and Weenik 2014

## Data and measurement

### Data collection

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

### Affixes N

| | | | |
|---|---|---|---|
| *-ness* | 363 | *pre-* | 123 |
| *-less* | 216 | *dis-* | 689 |
| *-wise* | 289 | *un-* | 960 |
| *-ize* | 476 | *in-* | 342 |
| *-ation* | 3979 | | |



Coleman et al. 2012, Boersma and Weenik 2014

## Data and measurement

### Data collection

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

### Affixes N

| | | | |
|---|---|---|---|
| *-ness* | 363 | *pre-* | 123 |
| *-less* | 216 | *dis-* | 689 |
| *-wise* | 289 | *un-* | 960 |
| *-ize* | 476 | *in-* | 342 |
| *-ation* | 3979 | | |

### Modeling

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



Coleman et al. 2012, Boersma and Weenik 2014, R Core Team 2017

## Data and measurement

### Data collection

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

### Affixes N

| | | | |
|---|---|---|---|
| *-ness* | 363 | *pre-* | 123 |
| *-less* | 216 | *dis-* | 689 |
| *-wise* | 289 | *un-* | 960 |
| *-ize* | 476 | *in-* | 342 |
| *-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



Coleman et al. 2012, Boersma and Weenik 2014, R Core Team 2017

## Data and measurement

### Data collection

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

### Affixes N

| | | | |
|---|---|---|---|
| *-ness* | 363 | *pre-* | 123 |
| *-less* | 216 | *dis-* | 689 |
| *-wise* | 289 | *un-* | 960 |
| *-ize* | 476 | *in-* | 342 |
| *-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



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

## Data and measurement

### Data collection

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

### Affixes N

| | | | |
|---|---|---|---|
| *-ness* | 363 | *pre-* | 123 |
| *-less* | 216 | *dis-* | 689 |
| *-wise* | 289 | *un-* | 960 |
| *-ize* | 476 | *in-* | 342 |
| *-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



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

## Data and measurement

### Data collection

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

### Affixes N

| | | | |
|---|---|---|---|
| *-ness* | 363 | *pre-* | 123 |
| *-less* | 216 | *dis-* | 689 |
| *-wise* | 289 | *un-* | 960 |
| *-ize* | 476 | *in-* | 342 |
| *-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: 81 models



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

## Frequency and segmentability effects

| duration | word | affix | base |
|---|---|---|---|
| affix | -ness | | |
| word frequency | | | |
| base frequency | | | |
| relative frequency | | | |

p < .001 expected direction

## Frequency and segmentability effects

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

p < .001 expected direction
p < .001 unexpected direction

## Frequency and segmentability effects

| duration | word | affix | base | word | affix | base | word | affix | base |
|---|---|---|---|---|---|---|---|---|---|
| affix | -ness | | | -ize | | | -ation | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |

p < .001 expected direction
p < .001 unexpected direction

## Frequency and segmentability effects

| duration | word | affix | base | word | affix | base | word | affix | base |
|---|---|---|---|---|---|---|---|---|---|
| affix | -ness | | | -ize | | | -ation | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |
| affix | -less | | | | | | | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |

p < .001 expected direction
p < .001 unexpected direction

## Frequency and segmentability effects

| duration | word | affix | base | word | affix | base | word | affix | base |
|---|---|---|---|---|---|---|---|---|---|
| affix | -ness | | | -ize | | | -ation | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |
| affix | -less | | | pre- | | | | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |

p < .001 expected direction

p < .001 unexpected direction

## Frequency and segmentability effects

| duration | word | affix | base | word | affix | base | word | affix | base |
|---|---|---|---|---|---|---|---|---|---|
| affix | -ness | | | -ize | | | -ation | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |
| affix | -less | | | pre- | | | -wise | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |

p < .001 expected direction

p < .001 unexpected direction

## Frequency and segmentability effects

| duration | word | affix | base | word | affix | base | word | affix | base |
|---|---|---|---|---|---|---|---|---|---|
| affix | -ness | | | -ize | | | -ation | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |
| affix | -less | | | pre- | | | -wise | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |
| affix | dis- | | | | | | | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |

p < .001 expected direction
p < .001 unexpected direction

## Frequency and segmentability effects

| duration | word | affix | base | word | affix | base | word | affix | base |
|---|---|---|---|---|---|---|---|---|---|
| affix | -ness | | | -ize | | | -ation | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |
| affix | -less | | | pre- | | | -wise | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |
| affix | dis- | | | un- | | | | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |

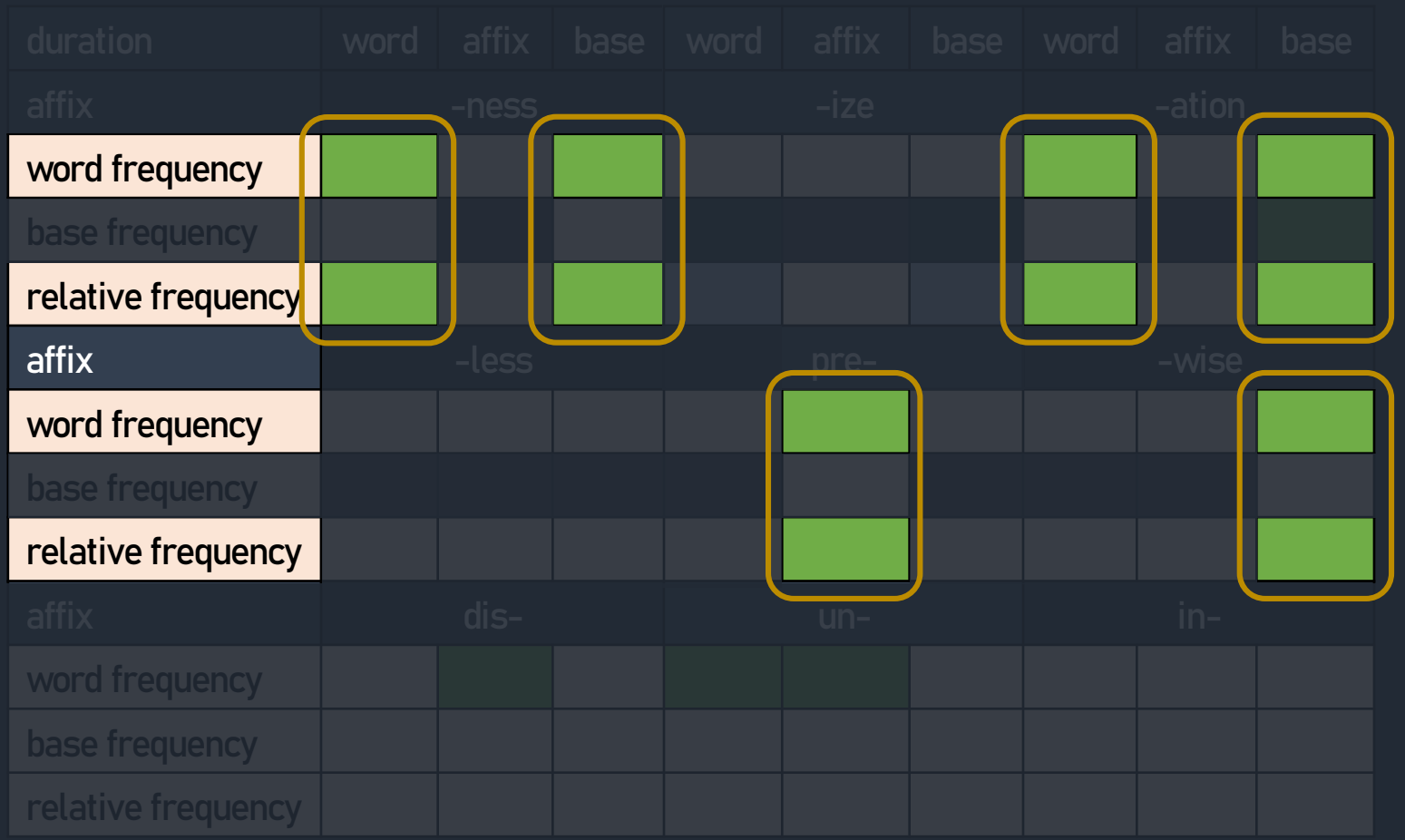
| | | |
|---|---|---|
| | $p < .001$ | expected direction |
| | $p < .001$ | unexpected direction |

## Frequency and segmentability effects

| duration | word | affix | base | word | affix | base | word | affix | base |
|---|---|---|---|---|---|---|---|---|---|
| affix | -ness | | | -ize | | | -ation | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |
| affix | -less | | | pre- | | | -wise | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |
| affix | dis- | | | un- | | | in- | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |

p < .001 expected direction
p < .001 unexpected direction

# Frequency and segmentability effects




p < .001 expected direction

p < .001 unexpected direction

## Frequency and segmentability effects

| duration | word | affix | base | word | affix | base | word | affix | base |
|---|---|---|---|---|---|---|---|---|---|
| affix | -ness | | | -ize | | | -ation | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |
| affix | -less | | | pre- | | | -wise | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |
| affix | dis- | | | un- | | | in- | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |

p < .001 expected direction
p < .001 unexpected direction

## Frequency and segmentability effects

| duration | word | affix | base | word | affix | base | word | affix | base |
|---|---|---|---|---|---|---|---|---|---|
| affix | -ness | | | -ize | | | -ation | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |
| affix | -less | | | pre- | | | -wise | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |
| affix | dis- | | | un- | | | in- | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |

p < .001 expected direction

p < .001 unexpected direction

Are the differences related to ...

## Prefixes vs. suffixes

| duration | word | affix | base | word | affix | base | word | affix | base |
|---|---|---|---|---|---|---|---|---|---|
| affix | -ness | | | -ize | | | -ation | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |
| affix | -less | | | pre- | | | -wise | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |
| affix | dis- | | | un- | | | in- | | |
| 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

| duration | word | affix | base | word | affix | base | word | affix | base |
|---|---|---|---|---|---|---|---|---|---|
| affix | -ness | | | -ize | | | -ation | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |
| affix | -less | | | pre- | | | -wise | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |
| affix | dis- | | | un- | | | in- | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |

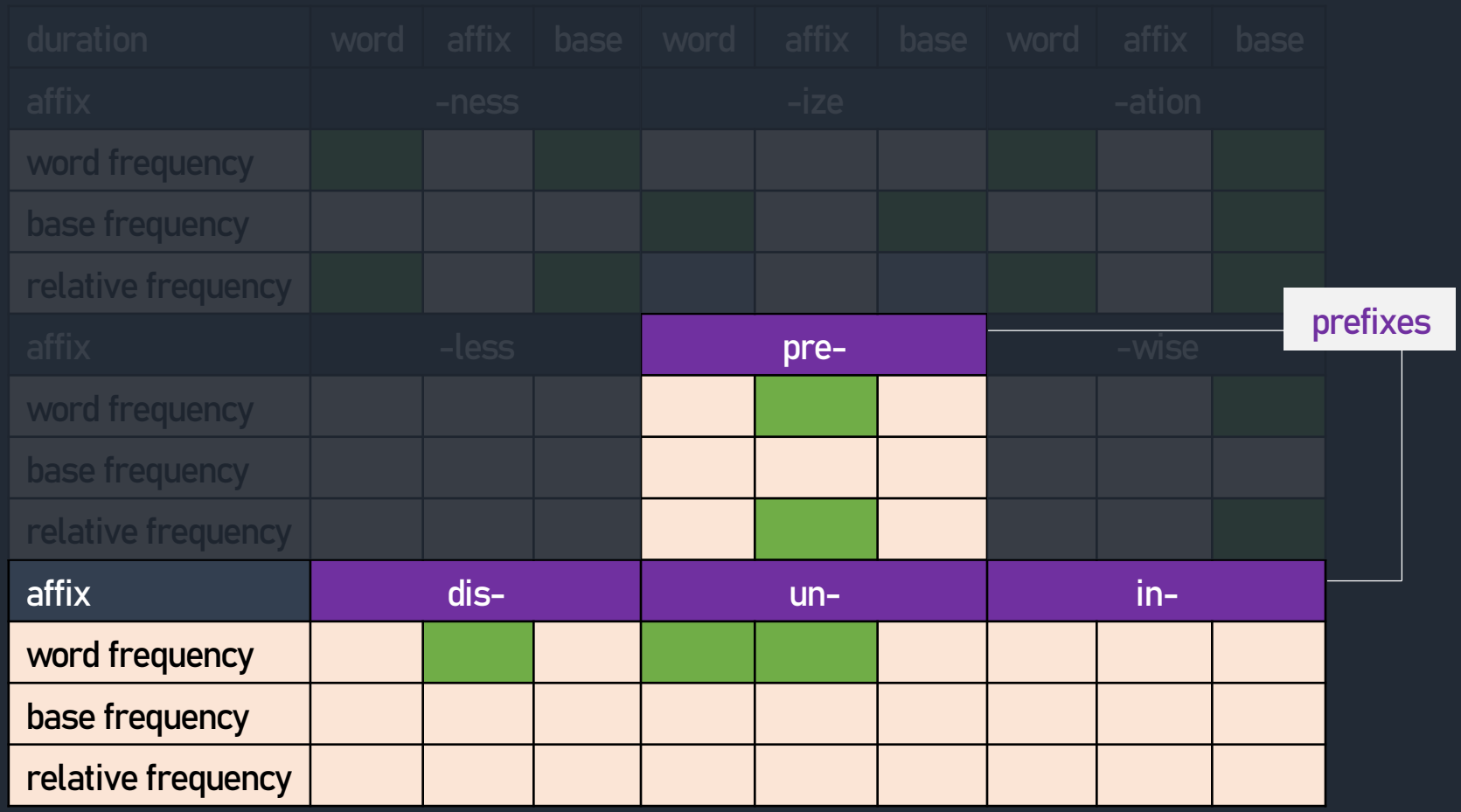
suffixes

p < .001 expected direction

p < .001 unexpected direction

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

## Prefixes vs. suffixes




p < .001 expected direction

p < .001 unexpected direction

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

## Affix length

| duration | word | affix | base | word | affix | base | word | affix | base |
|---|---|---|---|---|---|---|---|---|---|
| affix | -ness | | | -ize | | | -ation | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |
| affix | -less | | | pre- | | | -wise | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |
| affix | dis- | | | un- | | | in- | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |

p < .001 expected direction

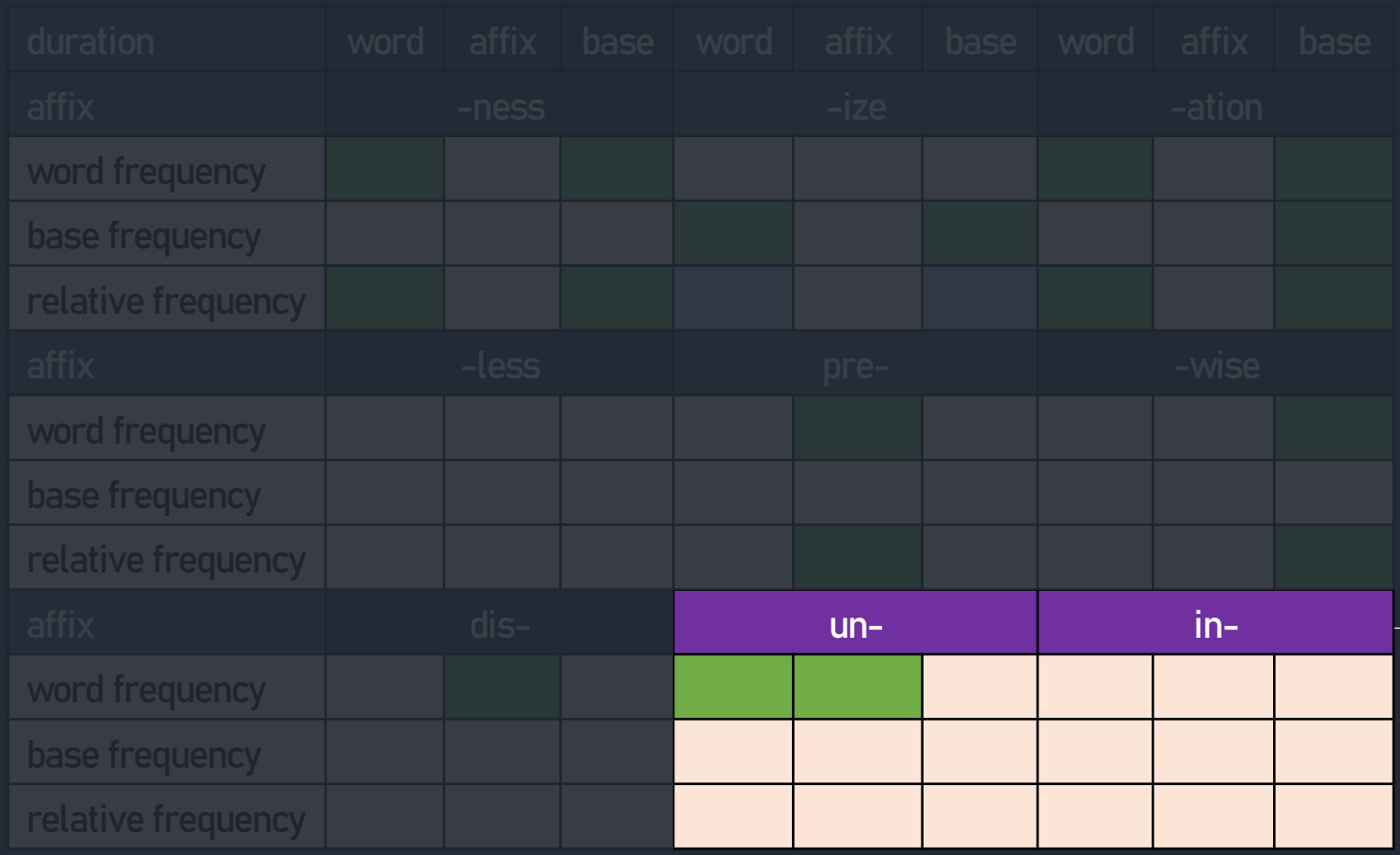
p < .001 unexpected direction

Are the differences related to ...

the type of affix? ✗

the affix length?

## Affix length




around 100–150 ms

p < .001 expected direction

p < .001 unexpected direction

Are the differences related to ...

the type of affix? ✗

the affix length?

## Affix length


p < .001 expected direction

p < .001 unexpected direction

Are the differences related to ...

the type of affix? ✗

the affix length? ✗

## Manual resegmentation

| duration | word | affix | base | word | affix | base | word | affix | base |
|---|---|---|---|---|---|---|---|---|---|
| affix | -ness | | | -ize | | | -ation | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |
| affix | -less | | | pre- | | | -wise | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |
| affix | dis- | | | un- | | | in- | | |
| 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

| duration | word | affix | base | word | affix | base | word | affix | base |
|---|---|---|---|---|---|---|---|---|---|
| affix | -ness | | | -ize | | | -ation | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |
| affix | -less | | | pre- | | | -wise | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |
| affix | dis- | | | un- | | | in- | | |
| 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? ✗ |
| p < .01 | weaker effect | | the segmentation? ✗ |

## Type of prosodic integration

## Type of prosodic integration

### The prosodic hierarchy

- U — Phonological utterance
- IP — Intonation phrase
- ϕ — Phonological phrase
- ω — Prosodic word
- Σ — Foot
- σ — Syllable



Hildebrandt 2015, Raffelsiefen 1999, 2007

## Type of prosodic integration

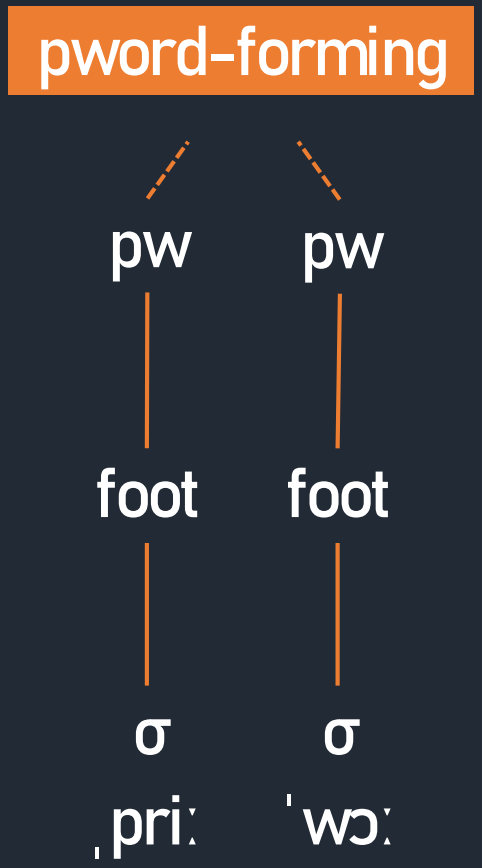
### The prosodic hierarchy

- U — Phonological utterance
- IP — Intonation phrase
- ϕ — Phonological phrase
- ω — Prosodic word
- Σ — Foot
- σ — Syllable



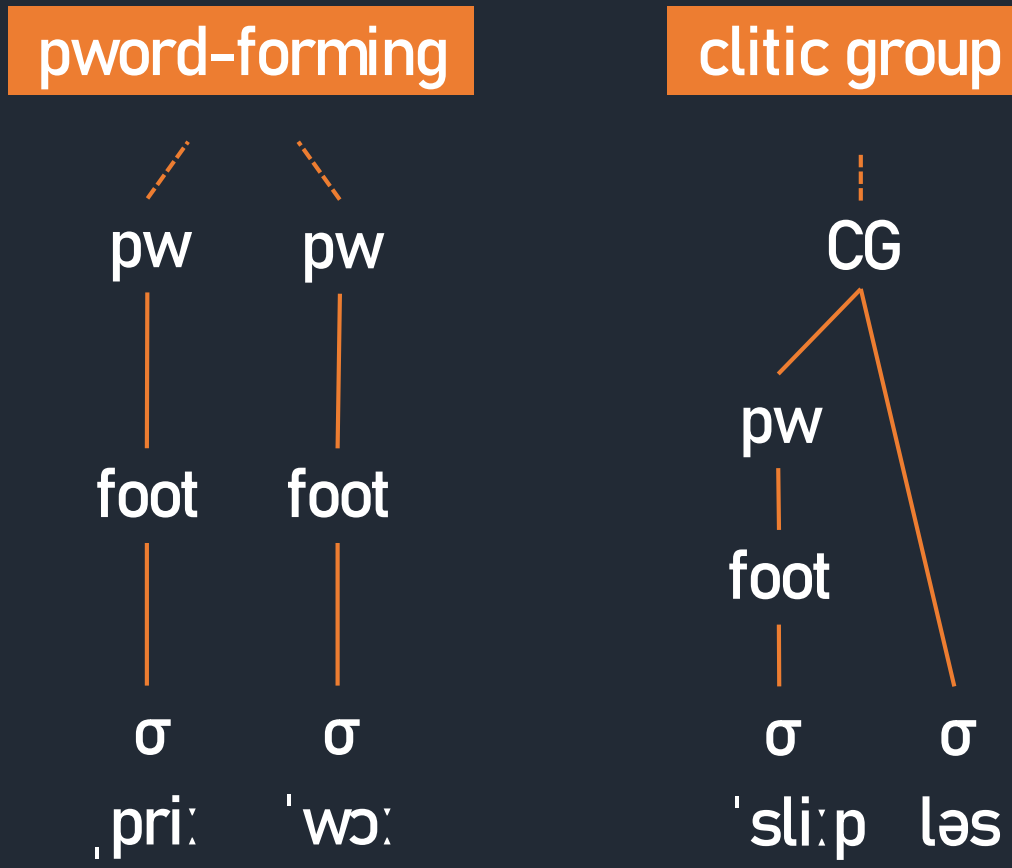
Hildebrandt 2015, Raffelsiefen 1999, 2007

## Type of prosodic integration



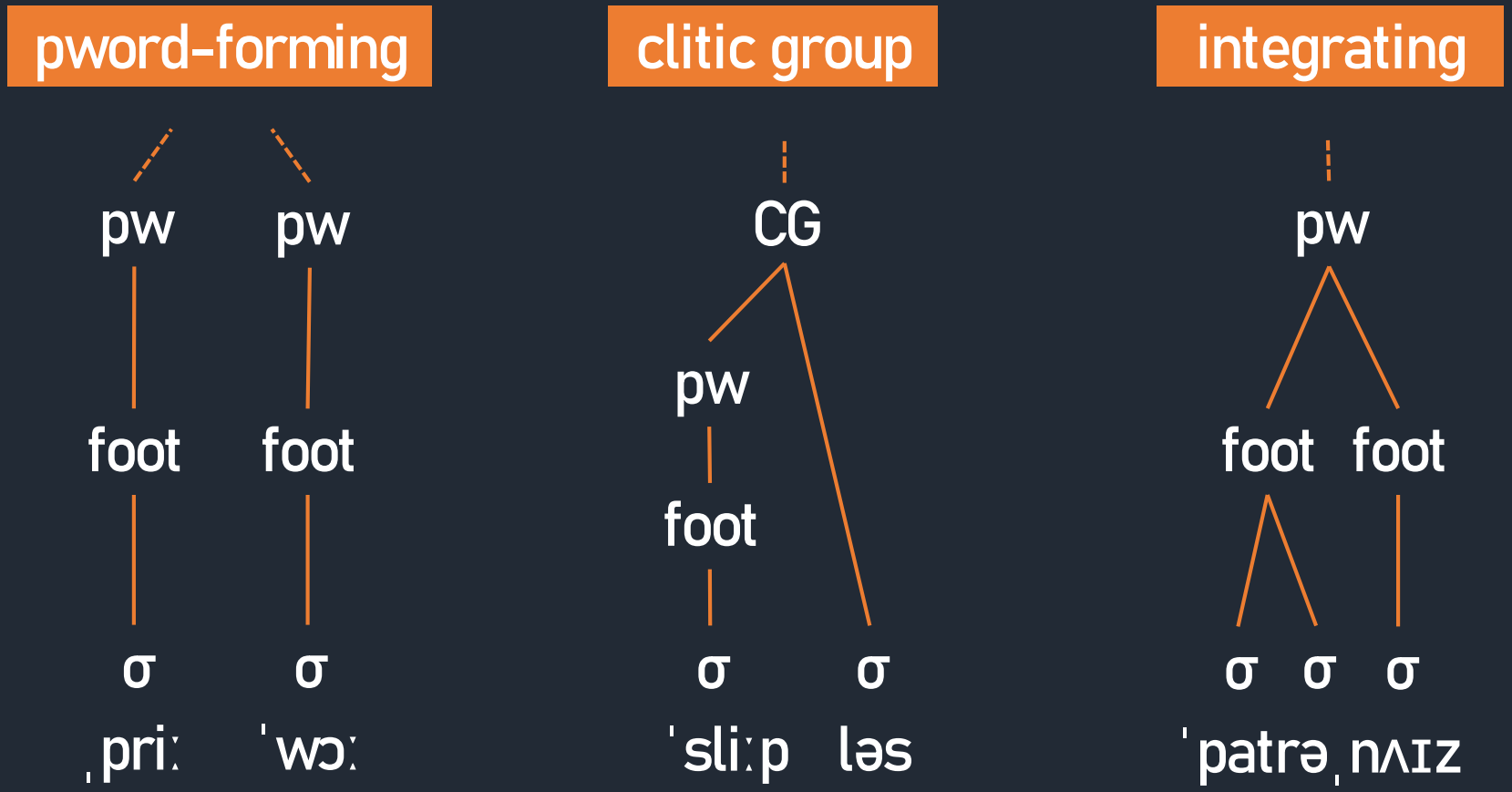

following Raffelsiefen 1999

## Type of prosodic integration




following Raffelsiefen 1999

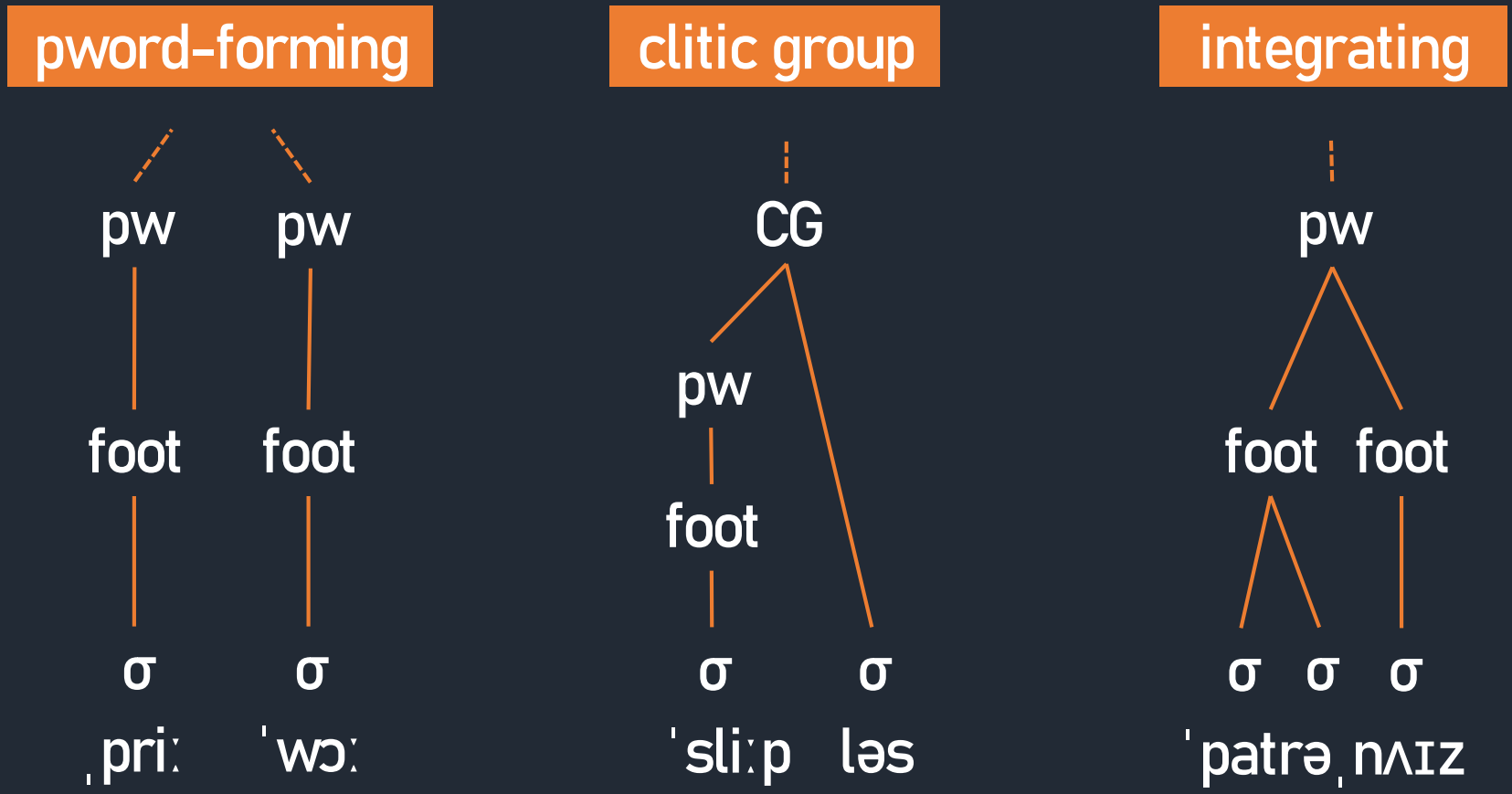
## Type of prosodic integration




following Raffelsiefen 1999

## Type of prosodic integration

more integration




following Raffelsiefen 1999

## Type of prosodic integration

duration longer — duration shorter

### pword-forming


### clitic group


### integrating


following Raffelsiefen 1999

## Type of prosodic integration

duration longer — unclear — duration shorter

### pword-forming


### clitic group


### integrating


following Raffelsiefen 1999

## Type of prosodic integration

duration longer — unclear — duration shorter

### pword-forming


### clitic group


### integrating


following Raffelsiefen 1999

## Type of prosodic integration

duration longer — unclear — duration shorter

### pword-forming


### clitic group


### integrating


following Raffelsiefen 1999

## Type of prosodic integration

| duration | word | affix | base | word | affix | base | word | affix | base |
|---|---|---|---|---|---|---|---|---|---|
| affix | -ness | | | -ize | | | -ation | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |
| affix | -less | | | pre- | | | -wise | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |
| affix | dis- | | | un- | | | in- | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |

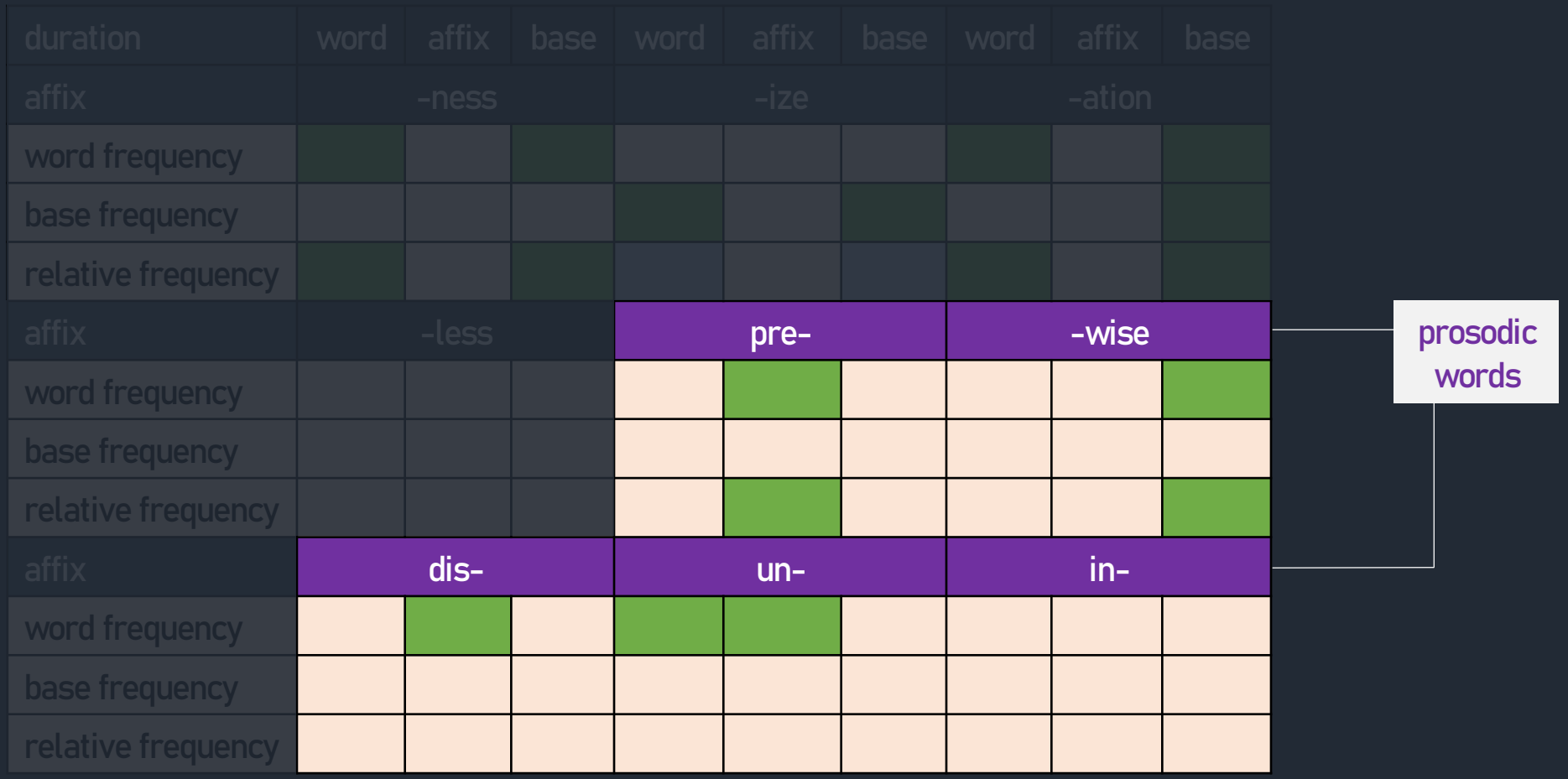
p < .001 expected direction

p < .001 unexpected direction

Are the differences related to ...

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

## Type of prosodic integration




p < .001 expected direction

p < .001 unexpected direction

Are the differences related to ...

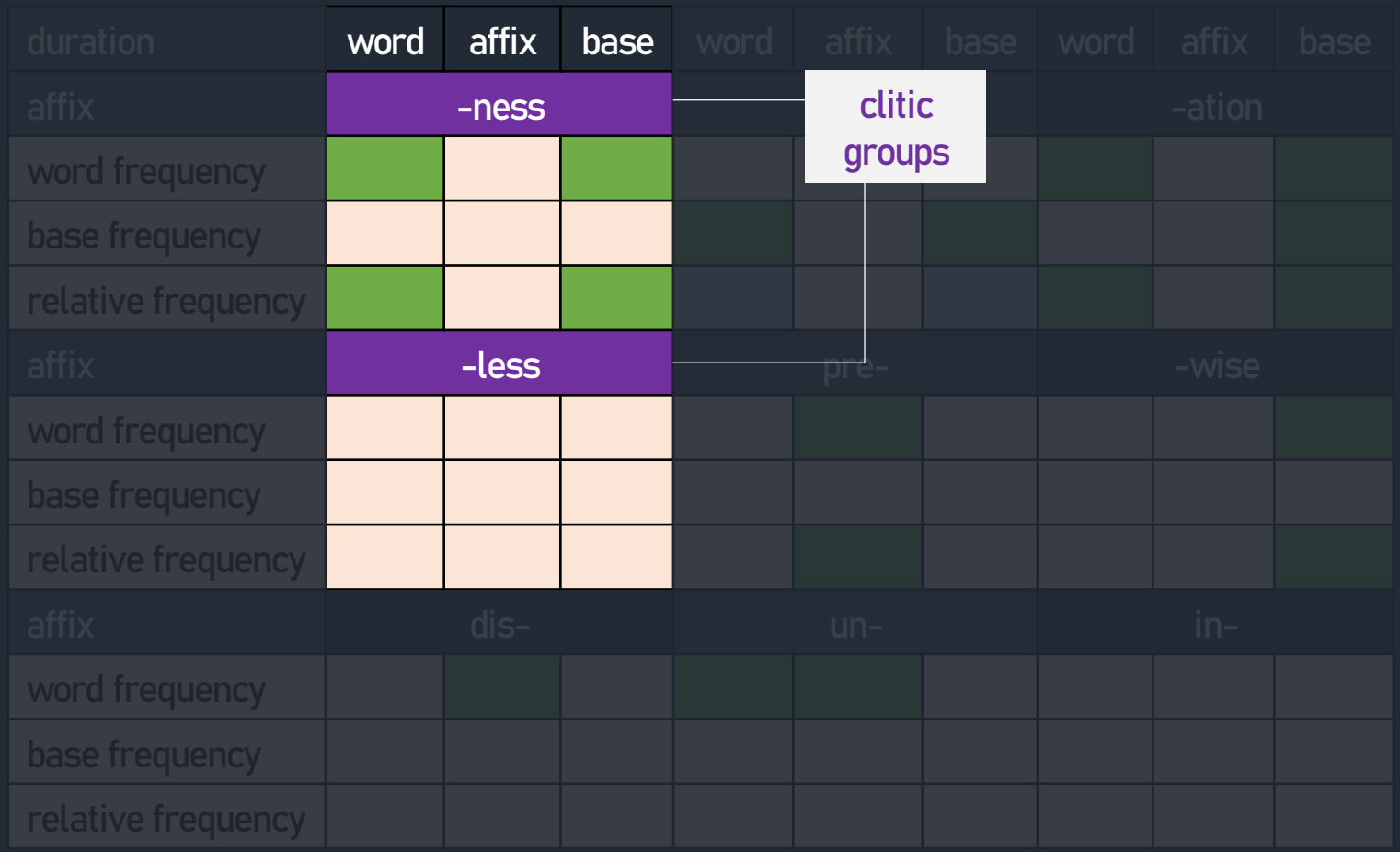
the type of affix? ✗

the affix length? ✗

the segmentation? ✗

prosodic structure?

## Type of prosodic integration




p < .001 expected direction

p < .001 unexpected direction

Are the differences related to ...

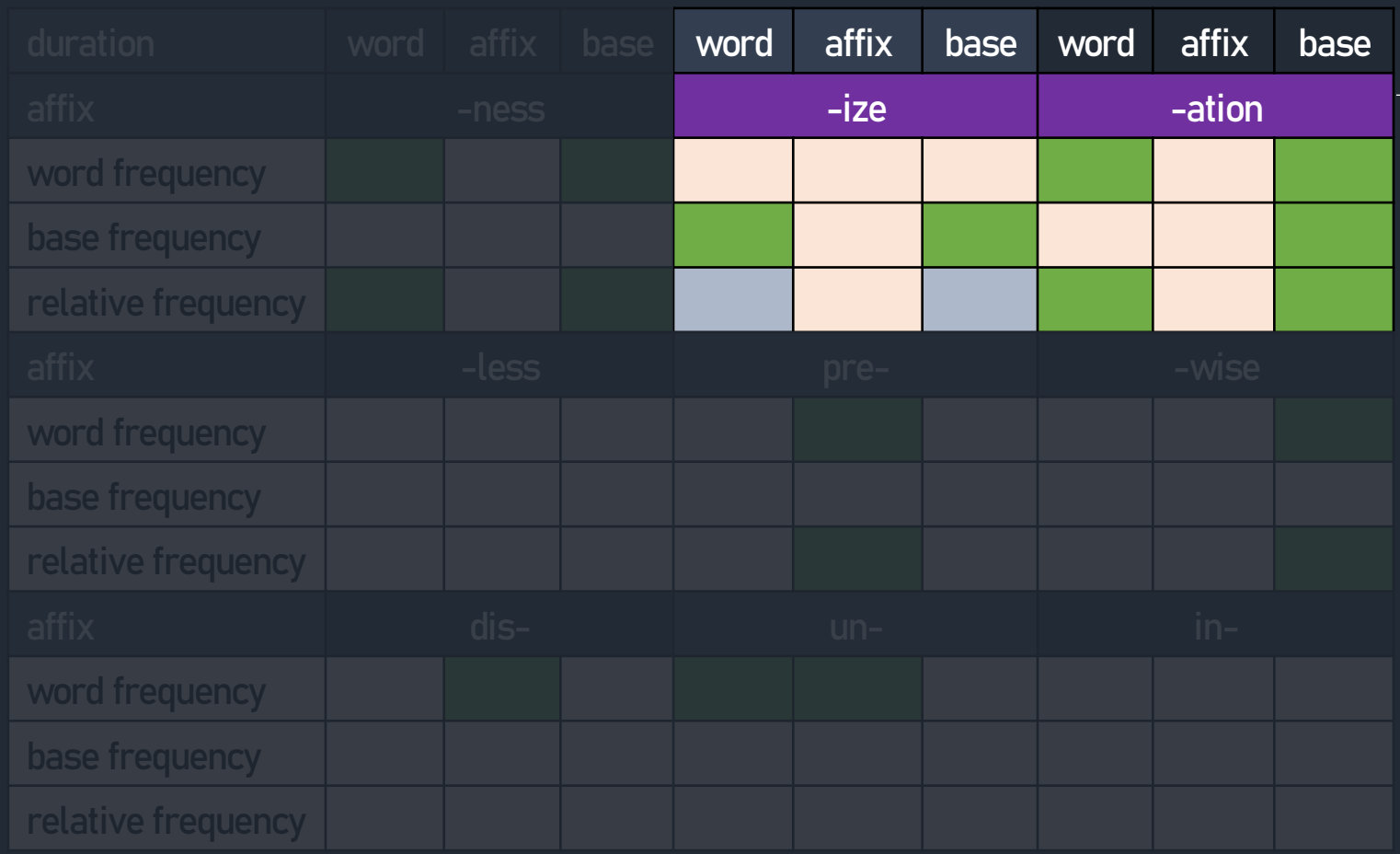
the type of affix? ✗

the affix length? ✗

the segmentation? ✗

prosodic structure?

## Type of prosodic integration




p < .001 expected direction

p < .001 unexpected direction

Are the differences related to ...

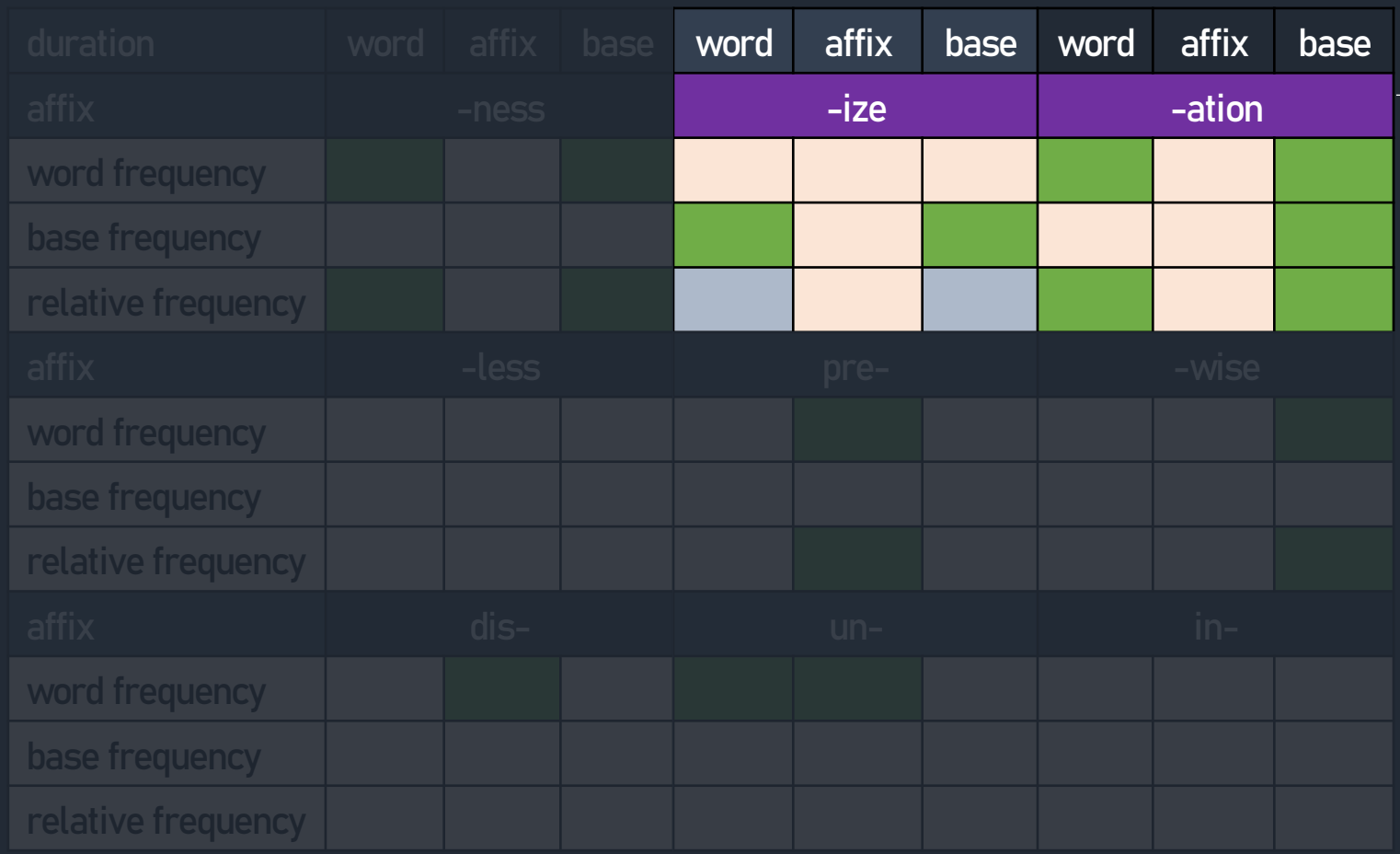
the type of affix? ✗

the affix length? ✗

the segmentation? ✗

prosodic structure?

## Type of prosodic integration




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

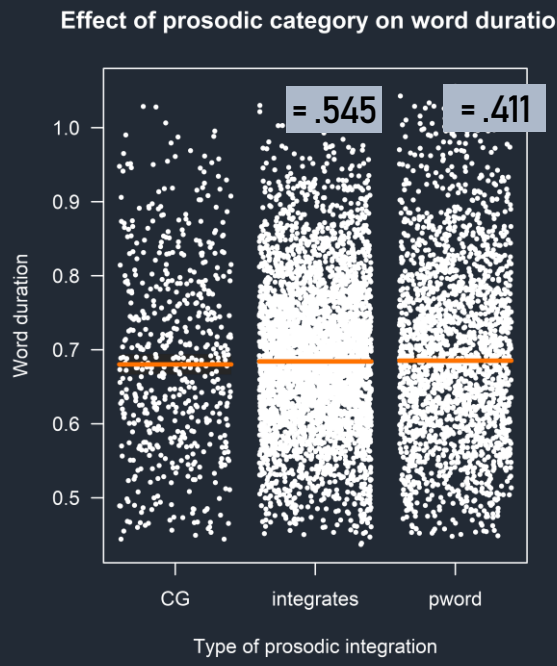
Meta-model including all affixes

- Additional predictor: type of prosodic integration
- Additional covariate: number of timing slots
- N = 7441

## Type of prosodic integration

Meta-model including all affixes

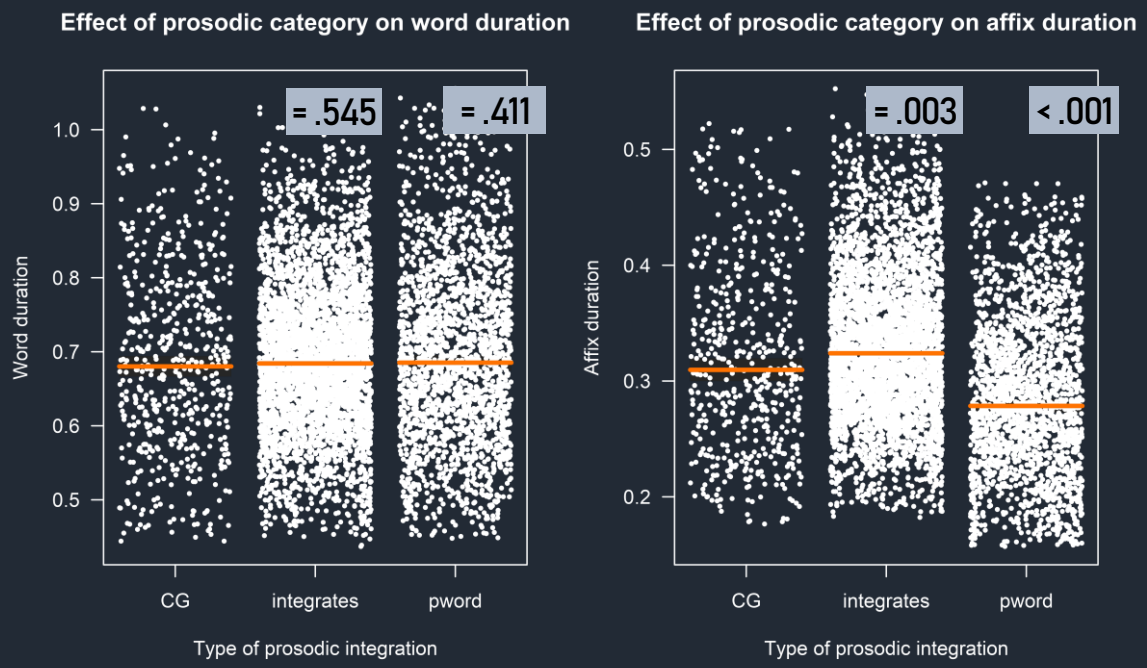
- Additional predictor: type of prosodic integration
- Additional covariate: number of timing slots
- N = 7441



## Type of prosodic integration

### Meta-model including all affixes

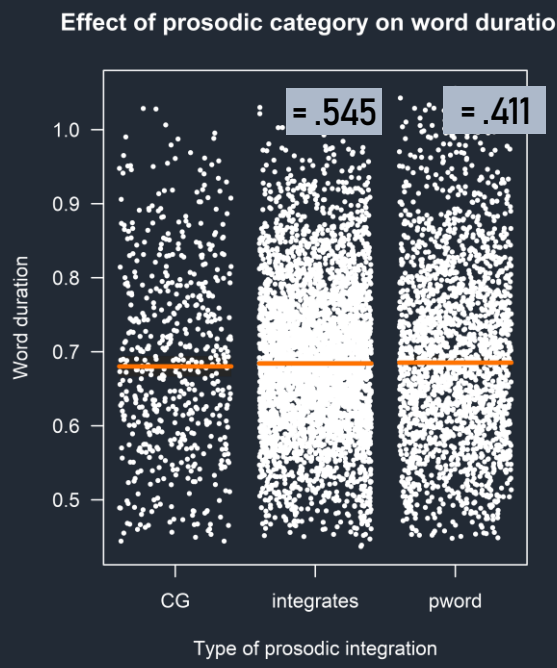
- Additional predictor: type of prosodic integration
- Additional covariate: number of timing slots
- N = 7441

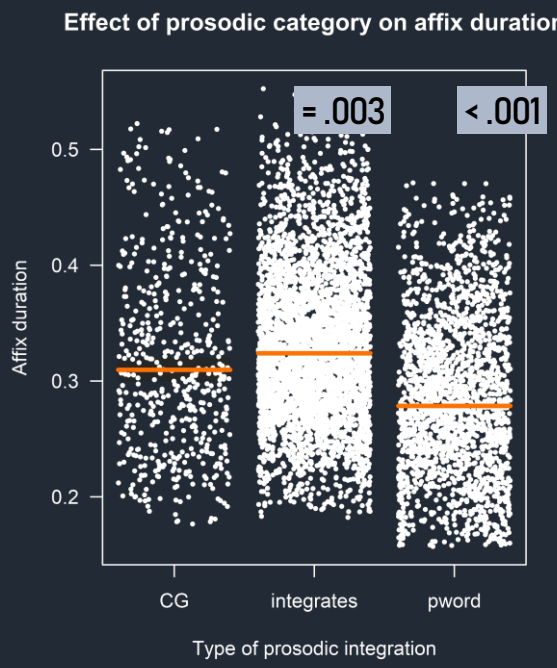


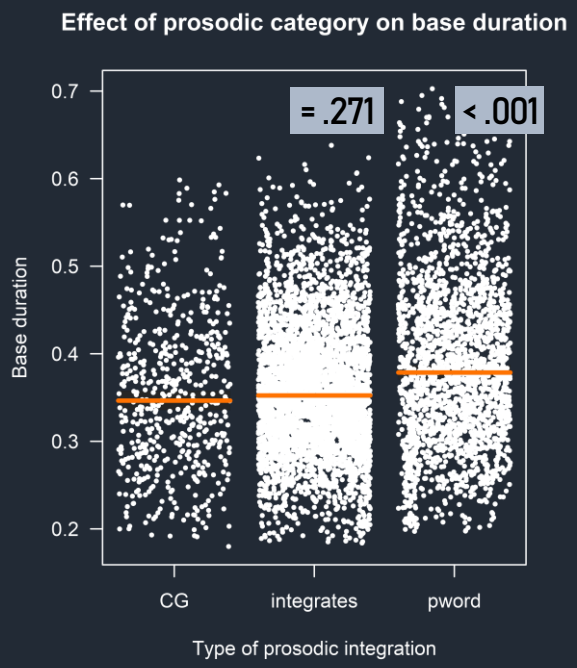
## Type of prosodic integration

### Meta-model including all affixes

- Additional predictor: type of prosodic integration
- Additional covariate: number of timing slots
- N = 7441



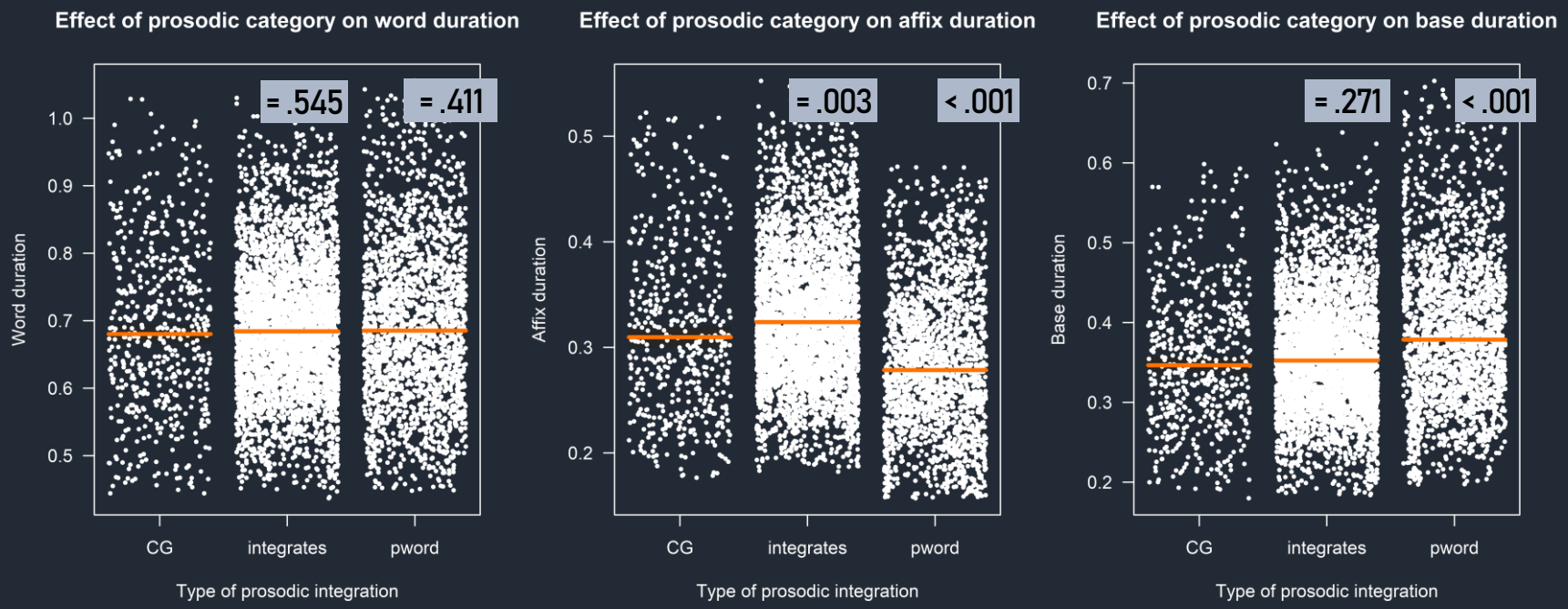



## Type of prosodic integration

### Meta-model including all affixes

- Additional predictor: type of prosodic integration
- Additional covariate: number of timing slots
- N = 7441
- This does not support the predictions of pword integration.



## Summary

## 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.

## 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

## Discussion

Our findings imply that …

- morphological structure can at least partly influence the phonetic output.

## 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) might have to be revised.



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

## 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) might have to 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

Thank you for listening.

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.
- Blazej, Laura J. & Ariel M. Cohen-Goldberg. 2015. Can we hear morphological complexity before words are complex? *Journal of Experimental Psychology*. Human perception and performance 41.1: 50–68.
- Boersma, Paul & David J. M. Weenik. 2014. Praat: Doing phonetics by computer (Version 5.4.04). Computer program. http://www.praat.org/.

Thank you for listening.


- 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.
- Cho, Taehong. 2001. Effects of morpheme boundaries on intergestural timing: Evidence from Korean. *Phonetica* 58: 129–162.
- Cohen-Goldberg, Ariel M. 2013. Towards a theory of multimorphemic word production: The heterogeneity of processing hypothesis. *Language and Cognitive Processes*. DOI: 10.1080/01690965.2012.759241.
- 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/.

Thank you for listening.


- Godfrey, J. E. Holliman & J. McDaniel. 1992. Telephone speech corpus for research and development. *Proceedings of ICASSP-92*, 517–520.
- Gordon, Elizabeth, Margaret Maclagan & Jennifer B. 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*. Basingstoke: Palgrave Macmillan. 82–104.
- Hanique, Iris & Mirjam Ernestus. 2012. The role of morphology in acoustic reduction. *Lingue e Linguaggio* 11: 147–164.
- 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.

Thank you for listening.


- 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*. 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.
- Lee-Kim, Sang-Im, Lisa Davidson & Sangjin Hwang. 2013. Morphological effects on the darkness of English intervocalic /l/. *Laboratory Phonology* 4.2: 475–511.
- 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.

Thank you for listening.


- 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.
- 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.

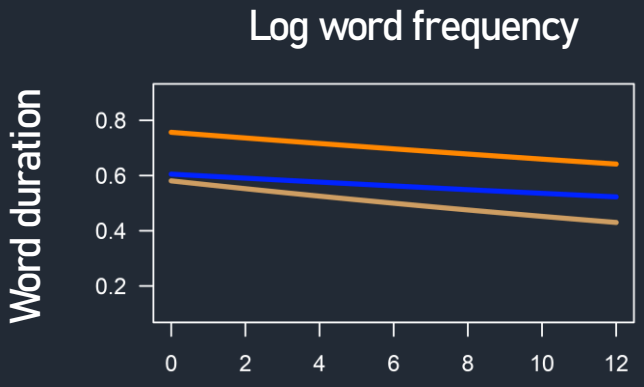
Thank you for listening.

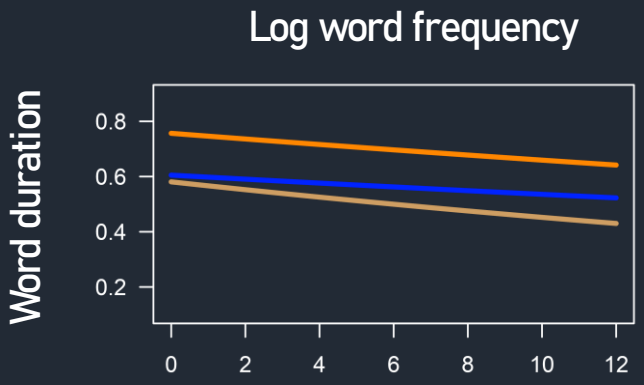

- R Core Team 2017. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing Vienna, Austria. http://www.R-project.org/.
- Seyfarth, Scott, Marc Garellek, Gwendolyn Gillingham, Farrell Ackerman & Robert Malouf. 2017. Acoustic differences in morphologically-distinct homophones. *Language, Cognition and Neuroscience*: 1–18.
- Sugahara, Mariko & Alice Turk. 2009. Durational correlates of English sublexical constituent structure. *Phonology* 26: 477–524.
- 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.

Thank you for listening.


- 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.

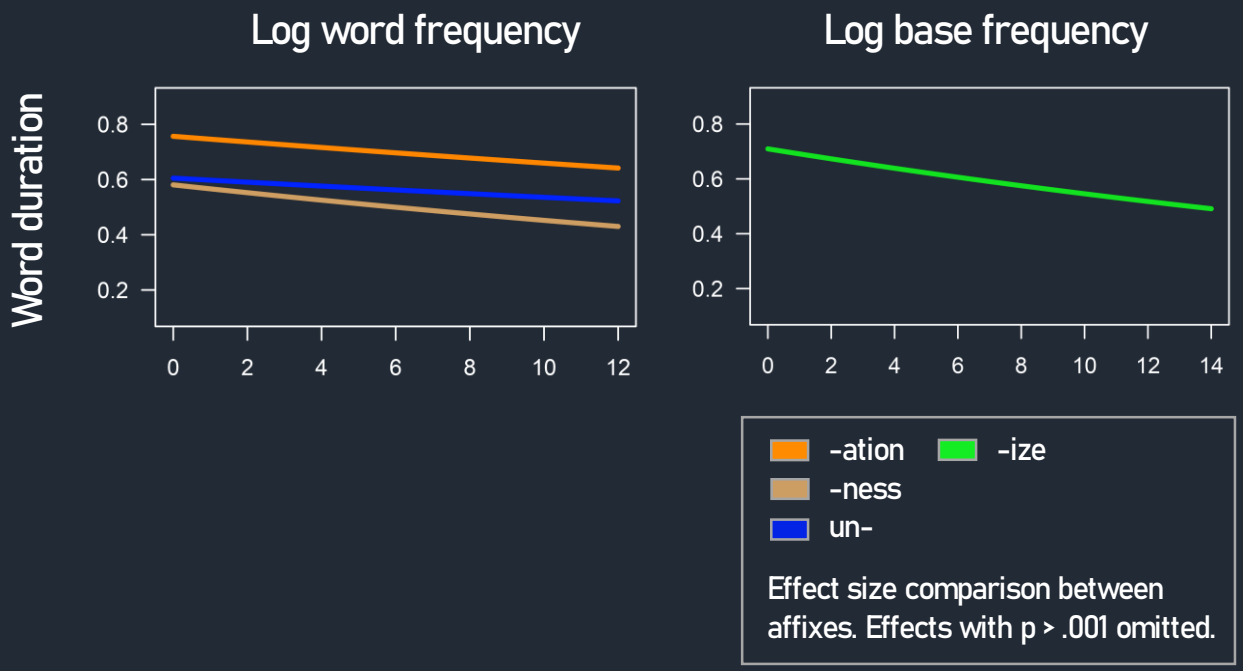
# Appendix

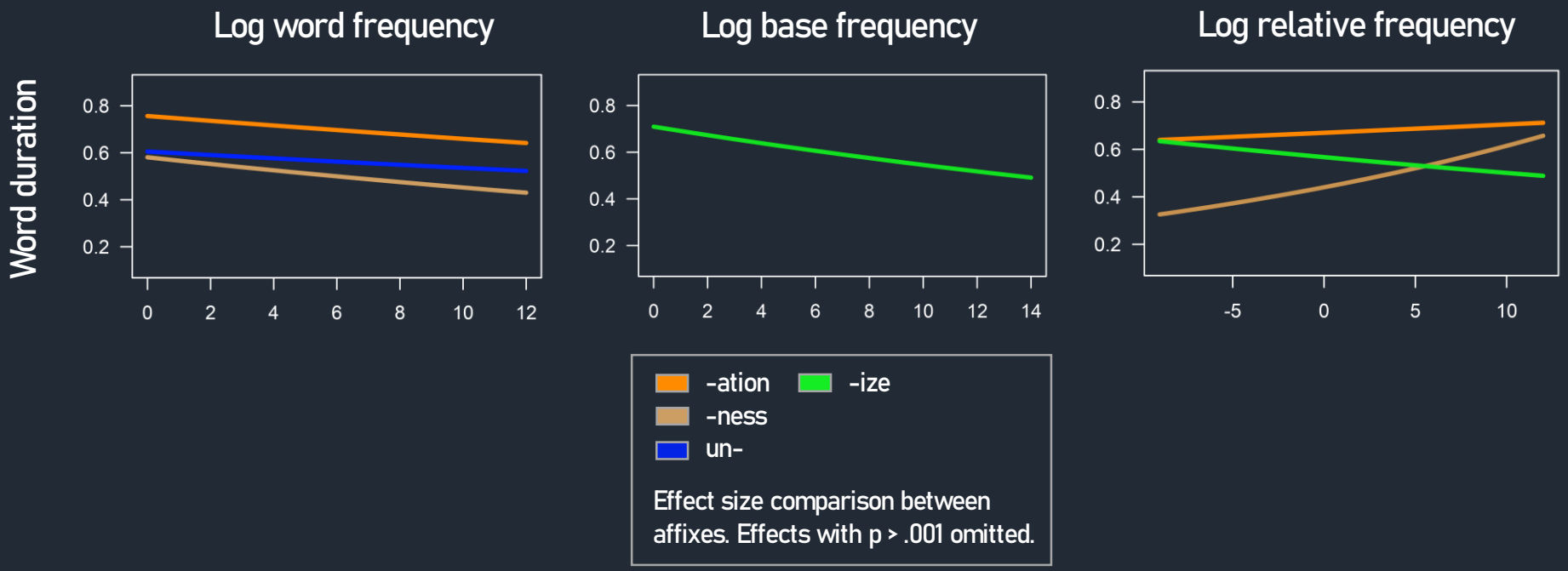

Log word frequency
Word duration
0.8
0.6
0.4
0.2
0
2
4
6
8
10
12

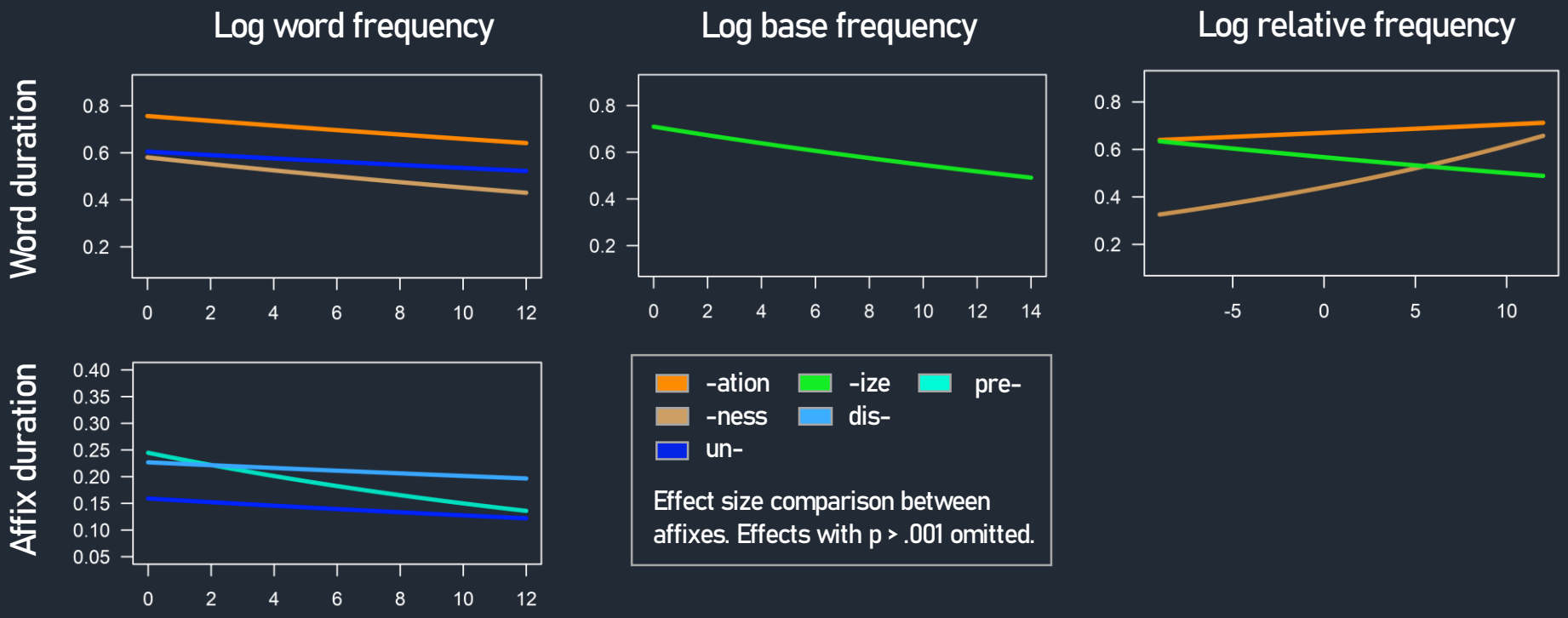


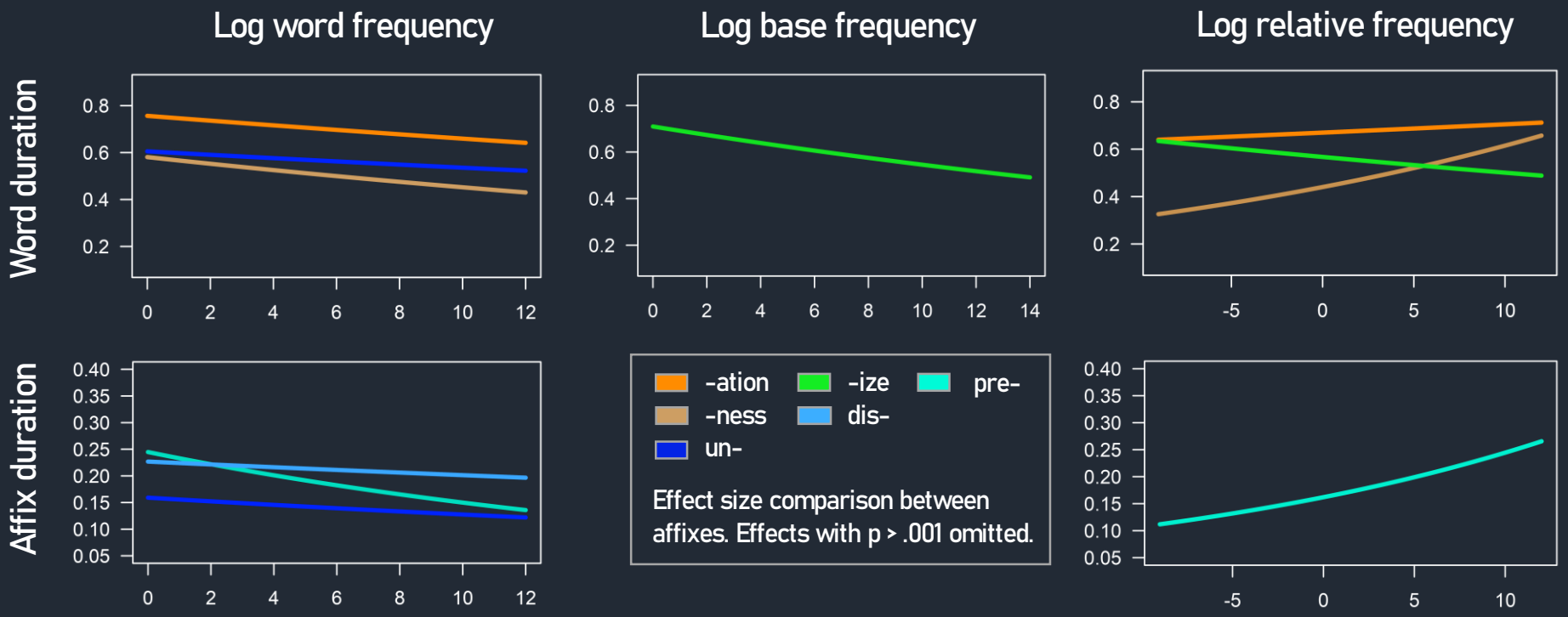

-ation
-ness
un-

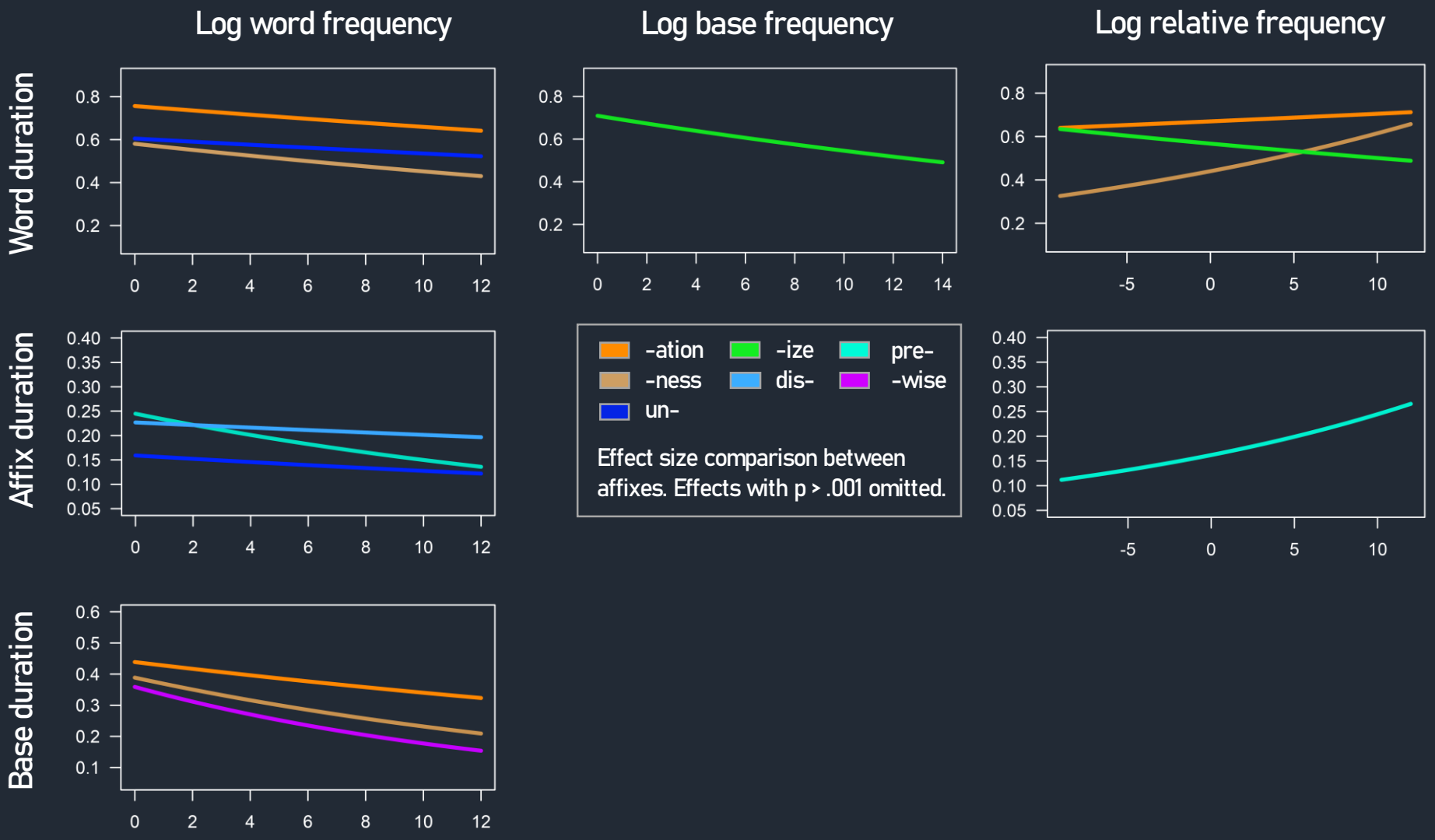
Effect size comparison between affixes. Effects with p > .001 omitted.

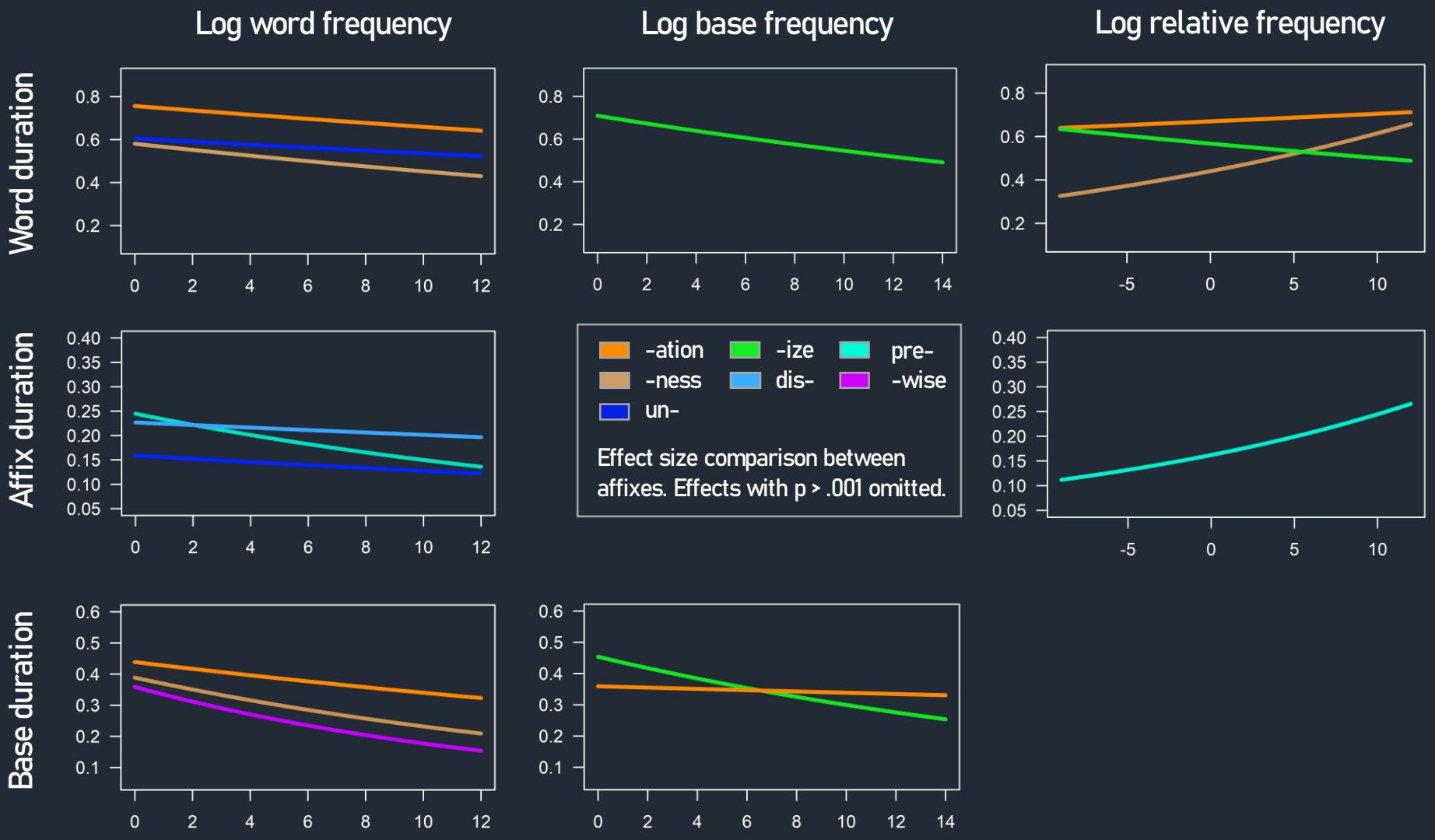

Log word frequency
Log base frequency
Word duration
0.8
0.6
0.4
0.2
0
2
4
6
8
10
12
14
-ation
-ness
un-
-ize
Effect size comparison between affixes. Effects with p > .001 omitted.

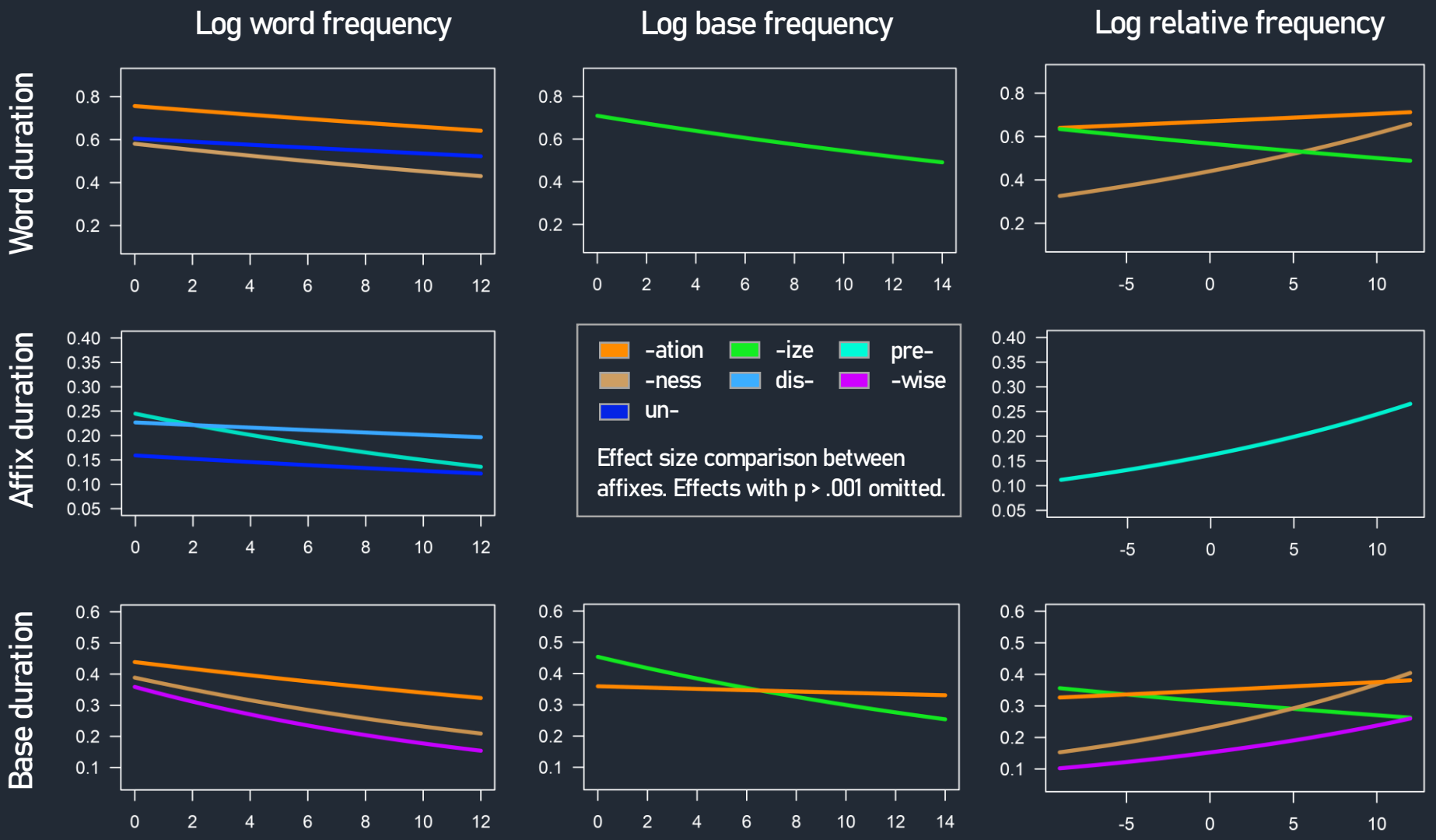

Log word frequency
Log base frequency
Log relative frequency
Word duration
0.8
0.6
0.4
0.2
0 2 4 6 8 10 12
0 2 4 6 8 10 12 14
-5 0 5 10
-ation
-ness
un-
-ize
Effect size comparison between affixes. Effects with p > .001 omitted.


Log word frequency
Log base frequency
Log relative frequency
Word duration
Affix duration
0.8
0.6
0.4
0.2
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0 2 4 6 8 10 12
0 2 4 6 8 10 12 14
-5 0 5 10
-ation
-ness
un-
-ize
dis-
pre-
Effect size comparison between affixes. Effects with p > .001 omitted.


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


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


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


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

# Informativity

| duration | word | affix | base | word | affix | base | word | affix | base |
|---|---|---|---|---|---|---|---|---|---|
| affix | -ness | | | -ize | | | -ation | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |
| affix | -less | | | pre- | | | -wise | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |
| affix | dis- | | | un- | | | in- | | |
| 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? ✗
- affix informativity?

## Informativity

Measured in two ways:

## Informativity

Measured in two ways:

Semantic information load score

## Informativity

Measured in two ways:

Semantic information load score

5-point Likert scales coded for:

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

## Informativity

Measured in two ways:

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

## Informativity

Measured in two ways:

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.

## Informativity

Measured in two ways:

Semantic information load score

Conditional affix probability $C_{aff}$

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.

## Informativity

Measured in two ways:

### 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-* | *...* |

## Informativity

Measured in two ways:

### 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-* | *...* |

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

## Informativity

Measured in two ways:

### 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.

# Informativity: Semantic information load score

| duration | word | affix | base | word | affix | base | word | affix | base |
|---|---|---|---|---|---|---|---|---|---|
| affix | -ness | | | -ize | | | -ation | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |
| affix | -less | | | pre- | | | -wise | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |
| affix | dis- | | | un- | | | in- | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |

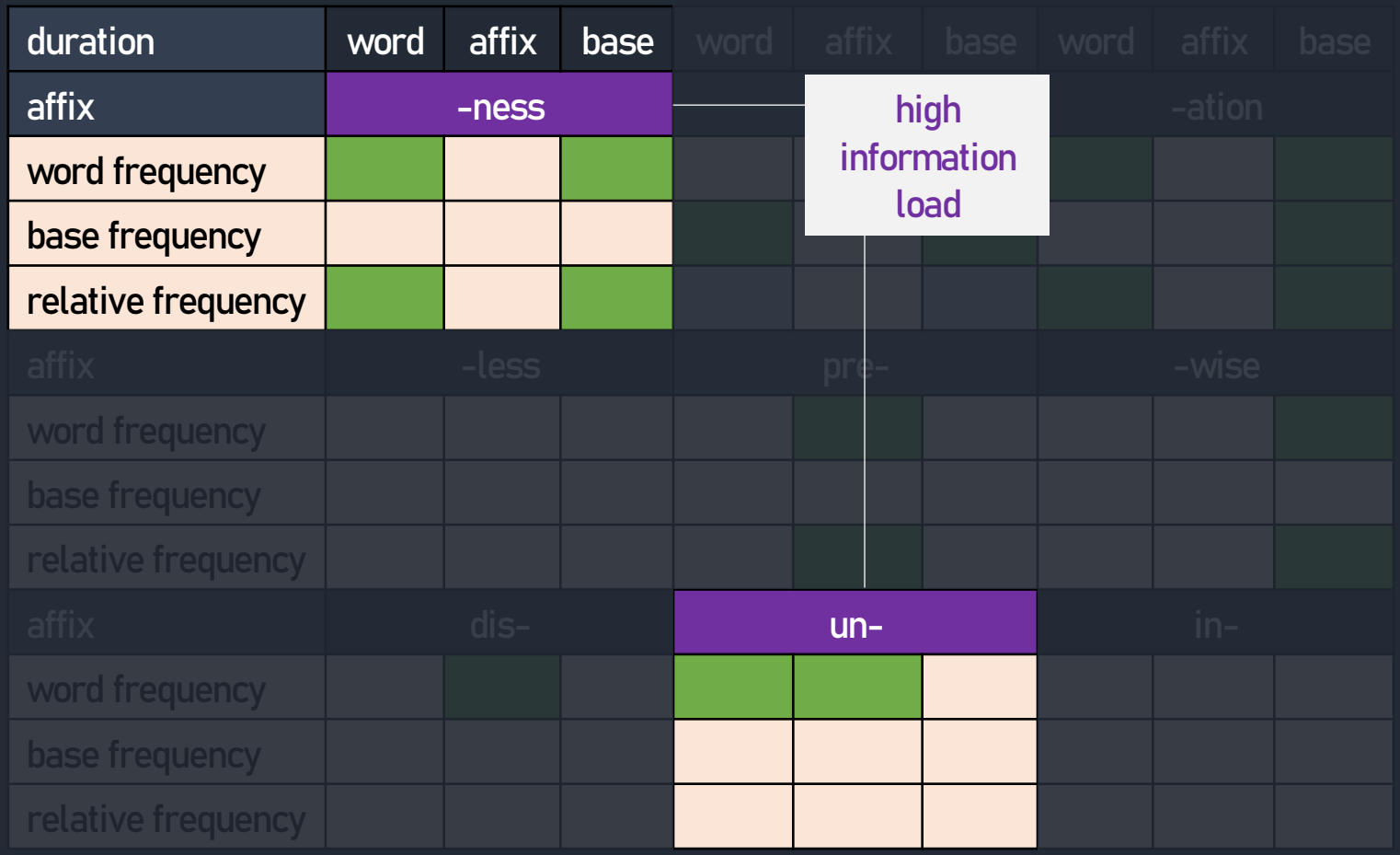
p < .001 expected direction

p < .001 unexpected direction

Are the differences related to ...

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

# Informativity: Semantic information load score




p < .001 expected direction

p < .001 unexpected direction

Are the differences related to ...

- the type of affix? ✗
- the affix length? ✗
- the segmentation? ✗
- prosodic structure? ✗
- affix informativity?

# Informativity: Semantic information load score

| duration | word | affix | base | word | affix | base | word | affix | base |
|---|---|---|---|---|---|---|---|---|---|
| affix | -ness | | | -ize | | | -ation | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |
| affix | -less | | | pre- | | | -wise | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |
| affix | dis- | | | un- | | | in- | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |

low information load

p < .001 expected direction

p < .001 unexpected direction

Are the differences related to ...

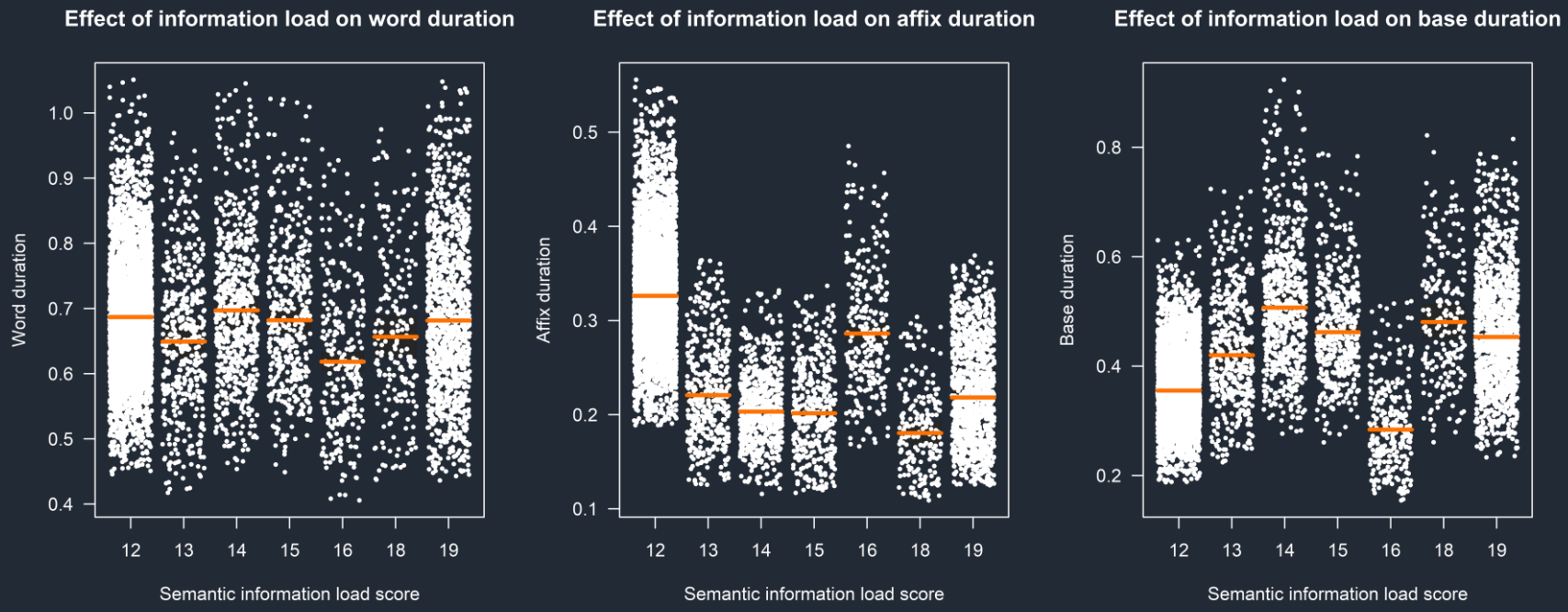
- the type of affix? ✗
- the affix length? ✗
- the segmentation? ✗
- prosodic structure? ✗
- affix informativity?

# Informativity: Semantic information load score

## Meta-model including all affixes

- Additional predictor: semantic information load score
- Additional covariate: number of timing slots
- N = 7441

- This does not support the predictions of semantic information load.



# Informativity: Conditional affix probability

| duration | word | affix | base | word | affix | base | word | affix | base |
|---|---|---|---|---|---|---|---|---|---|
| affix | -ness | | | -ize | | | -ation | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |
| affix | -less | | | pre- | | | -wise | | |
| word frequency | | | | | | | | | |
| base frequency | | | | | | | | | |
| relative frequency | | | | | | | | | |
| affix | dis- | | | un- | | | in- | | |
| 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? ✗
- affix informativity?

# Informativity: Conditional affix probability

| duration | word | affix | base | word | affix | base | word | affix | base |
|---|---|---|---|---|---|---|---|---|---|
| affix | -ness | | | -ize | | | -ation | | |
| affix probability | | | | | | | | | |

| affix | -less | | | pre- | | | -wise | | |
|---|---|---|---|---|---|---|---|---|---|
| affix probability | | | | | | | | | |

| affix | dis- | | | un- | | | in- | | |
|---|---|---|---|---|---|---|---|---|---|
| affix probability | | | | | | | | | |

p < .001 negative correlation

Are the differences related to ...

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

## Results: Updated

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)

## Results: Updated

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)
    - the informativity of the affix (information load, probability).

# Prosodic word diagnostics

# Prosodic word diagnostics

## The prosodic hierarchy

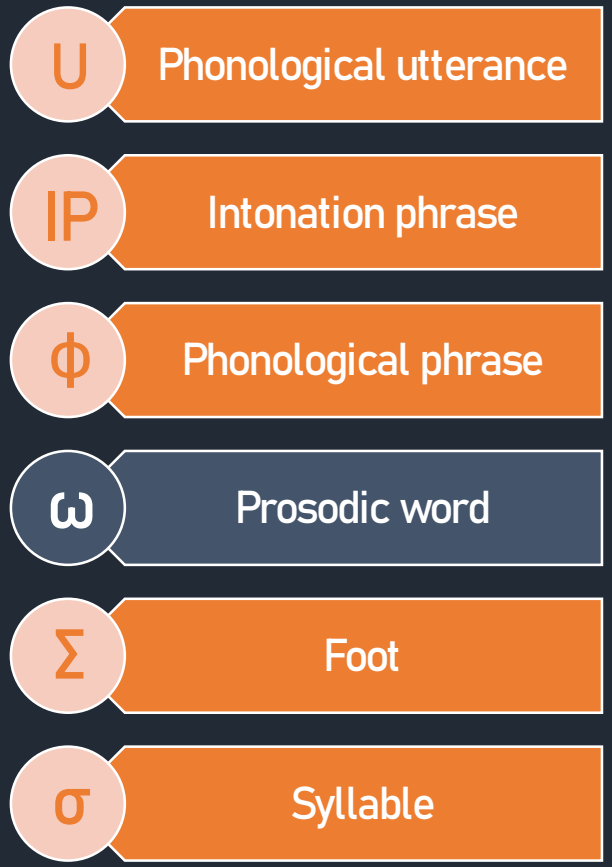
- U — Phonological utterance
- IP — Intonation phrase
- ϕ — Phonological phrase
- ω — Prosodic word
- Σ — Foot
- σ — Syllable



Hildebrandt 2015, Raffelsiefen 1999, 2007

# Prosodic word diagnostics

## The prosodic hierarchy




## Some pword-diagnostics

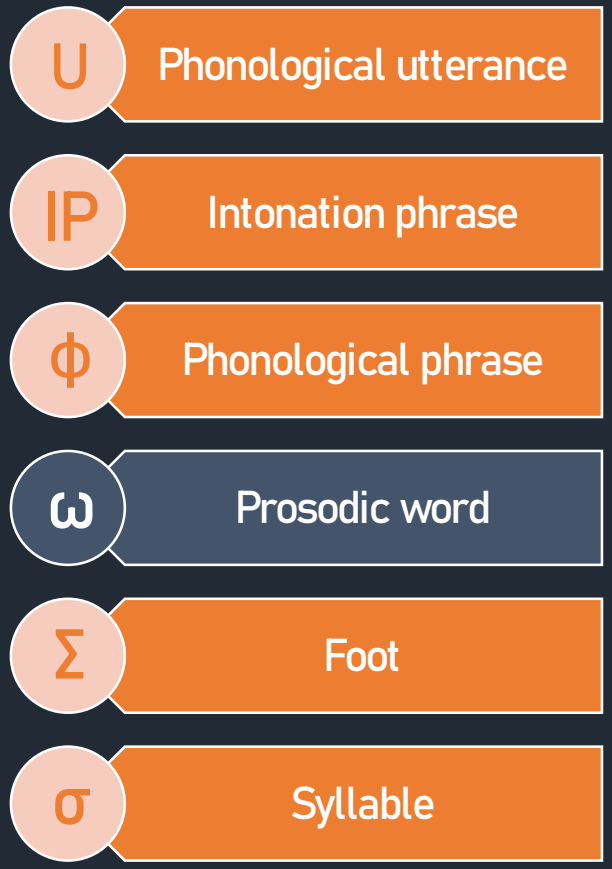
- onset or coda conditions, LOI-violations
- ambisyllabicity
- stress and relative prominence
- trisyllabic laxing, vowel reduction
- minimal word requirements
- compositionality, type of base



Hildebrandt 2015, Raffelsiefen 1999, 2007

# Prosodic word diagnostics

## 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