

# SPADE

SPeech Across Dialects of English

# Exploring the anatomy of articulation rate in spontaneous English speech

Relationships between utterance length effects and  
social factors

James Tanner, Morgan Sonderegger, Jane Stuart-Smith, Tyler Kendall,  
Jeff Mielke, Erik Thomas, Robin Dodsworth, SPADE Data Consortium

INTERSPEECH 2024, 1-5th September



University  
of Glasgow



McGill

UNIVERSITY OF  
OREGON

NC STATE  
UNIVERSITY

# Introduction

- speech rate
  - measure of speech timing – the tempo at which linguistic events unfold
- strong social stereotypes about differences in speech rate
  - *"People from X speak slowly, people from Y are always rushing their speech"* etc

# Introduction

- strong social stereotypes about differences in speech rate

 **Just Tracy**  @TracyOnFire2020 · May 5, 2022

Replying to @\_bee\_kay\_

Well I live in the south and that's not how Southerners sound. Southerners talk very slowly, often so **slow** that if they were talking any slower they'd be talking backwards. I have no idea what tribe this cheetah comes from.

3 1 29

 **Jamey\_Giddens** @Jamey\_Giddens · Dec 13, 2013

I am from the South. It is regionally impossible for us to talk as fast as these folks are talking. Our brains would explode. #Scandal

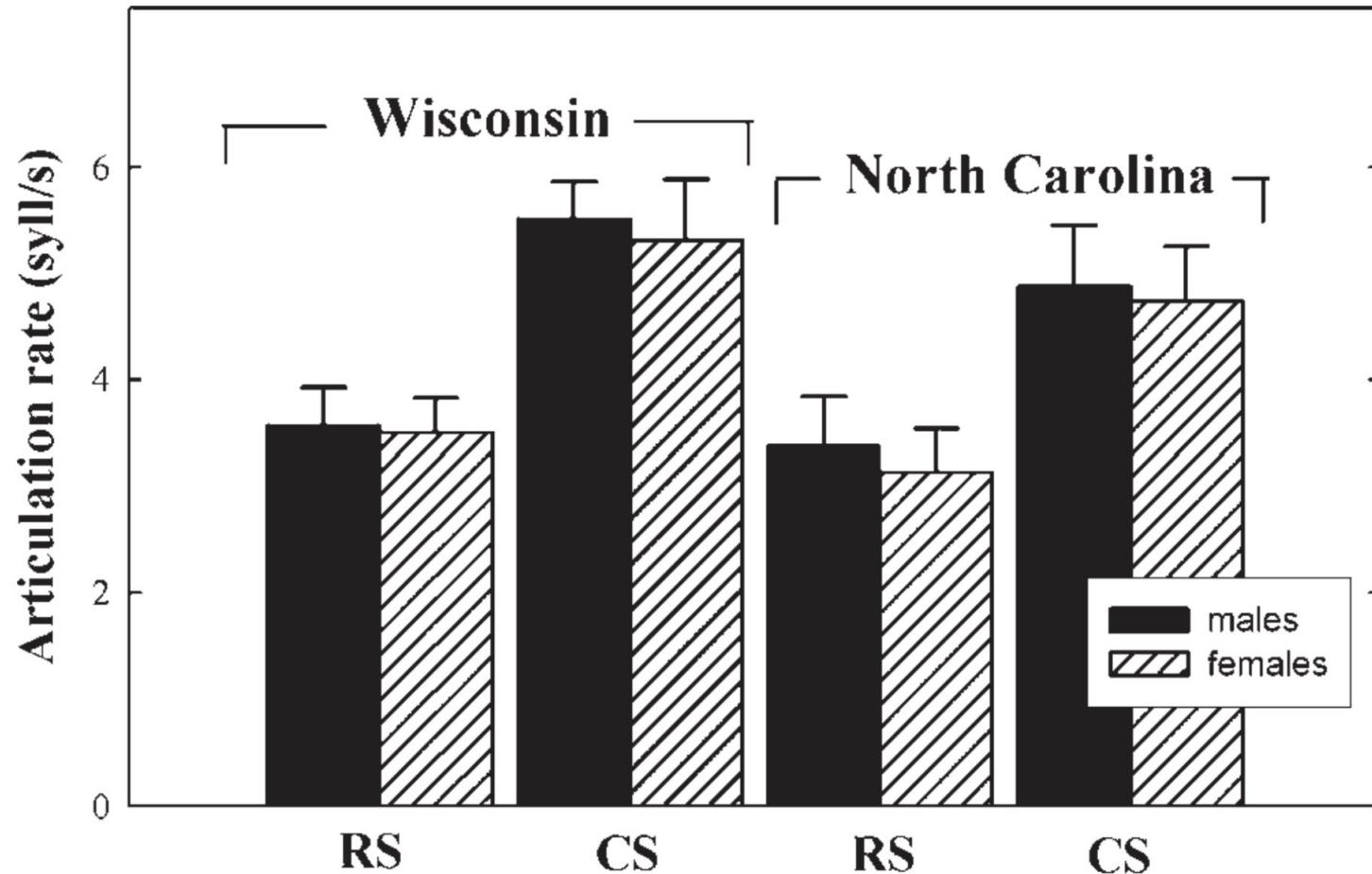
6 8 11

and I will be there, nose pressed against the screen, ears straining to understand fast talking northern english accents



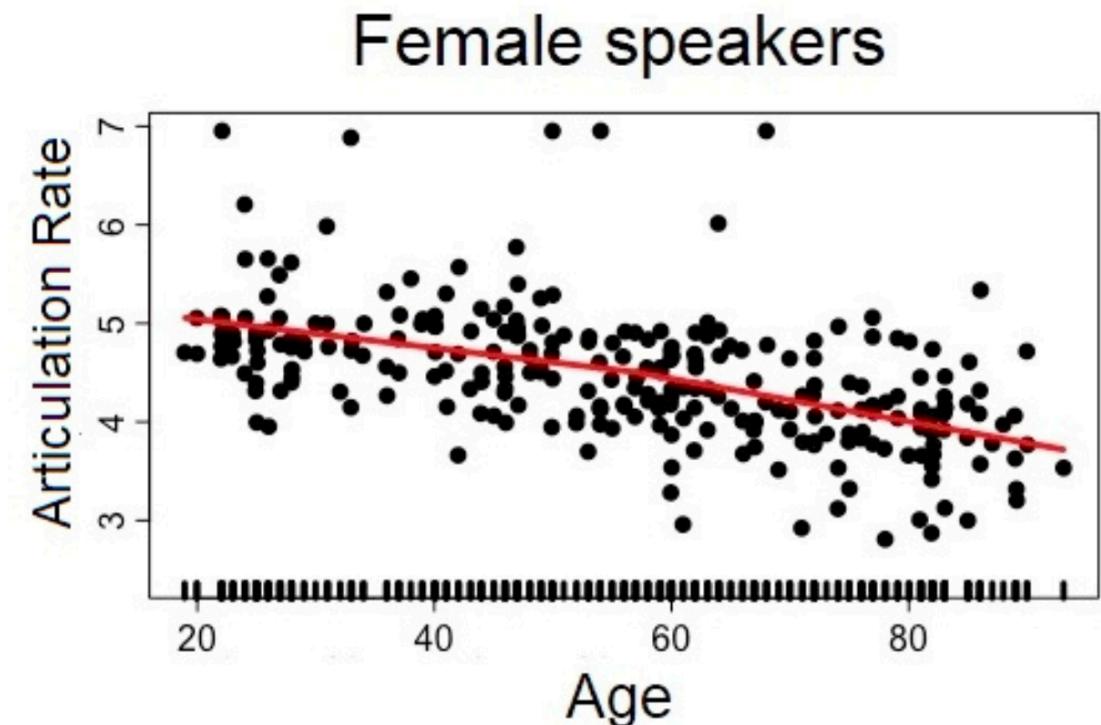
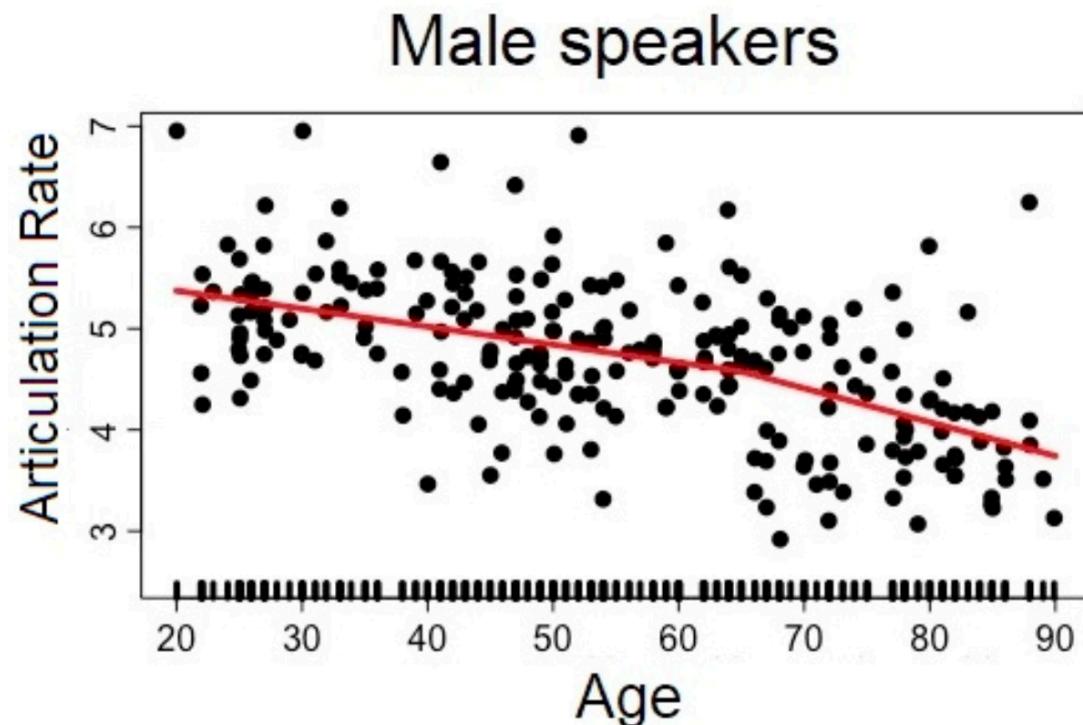
9:20 PM · Nov 20, 2021

# Speech rate by dialect



Some differences between dialects observed, but inconsistent in size of differences

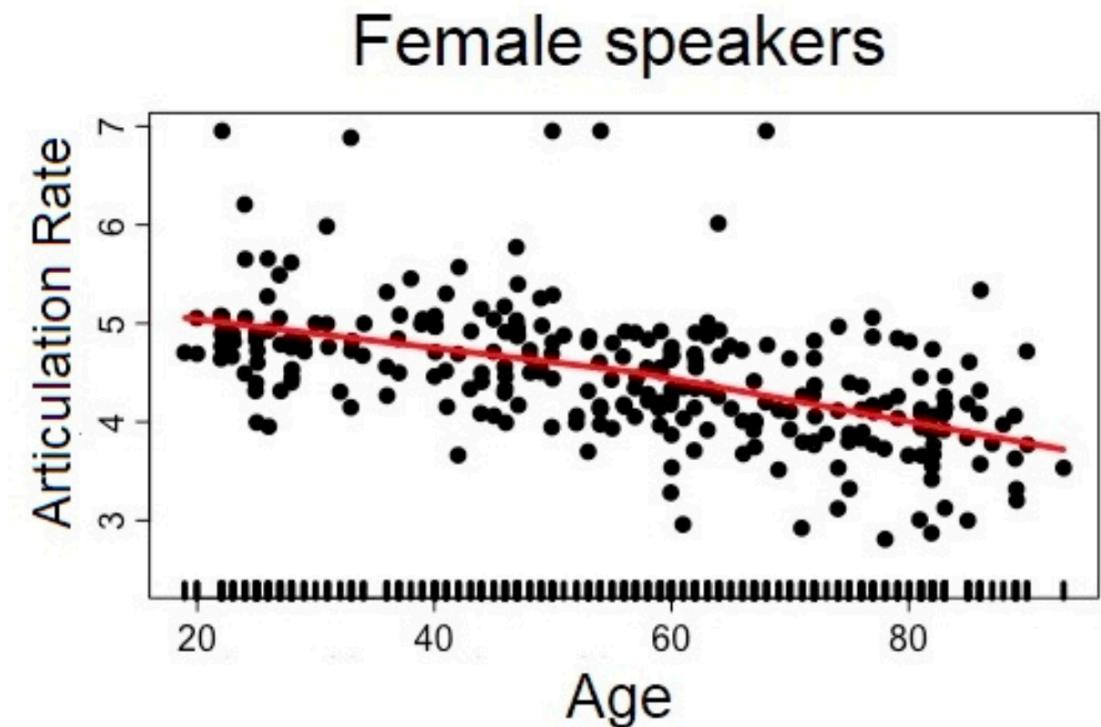
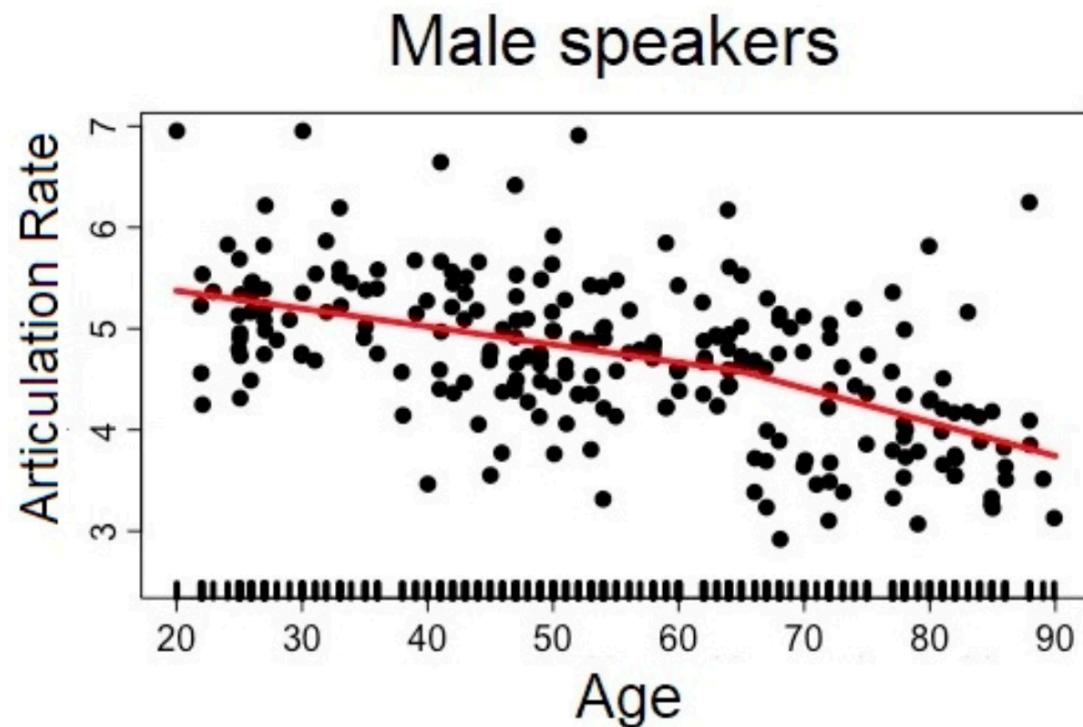
# Speech rate by age and gender



Age effect is inconsistent, but if found, then older speakers are slower than younger speakers

Ramig (1983), Byrd (1994), Yuan et al. (2006), Quené (2008),  
Jacewicz et al. (2009), Fougeron et al (2021)

# Speech rate by age and gender

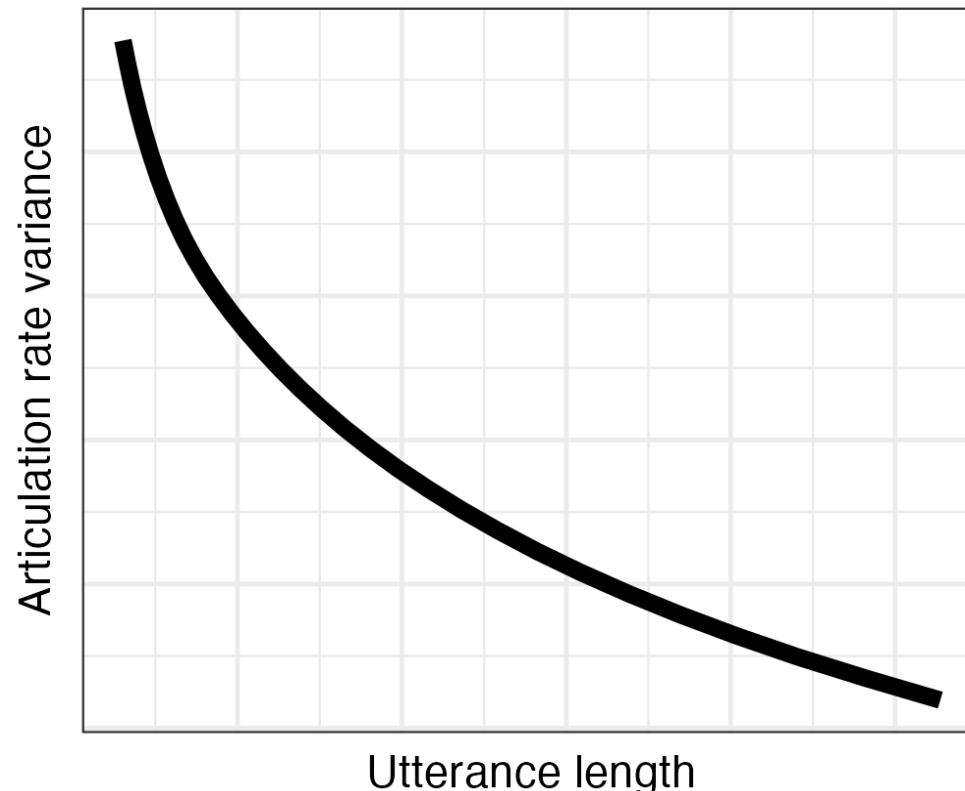
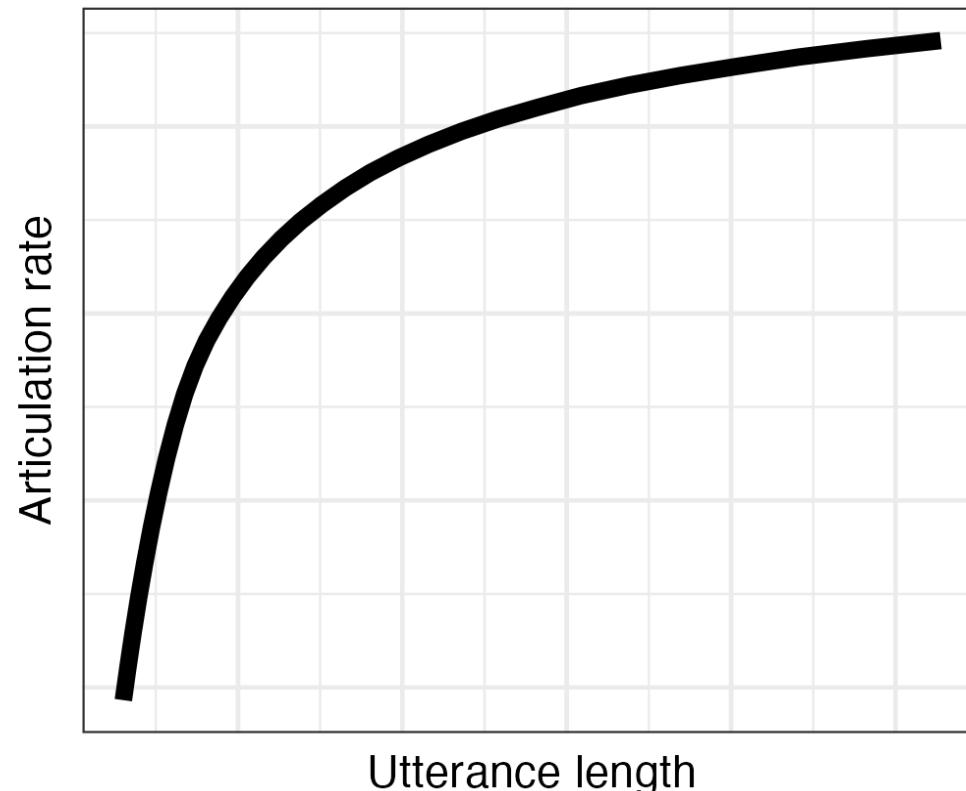


Gender effect is inconsistent, but if found, then female speakers are slower than male speakers

Syrdal (1996), Jacewicz et al. (2009), Kendall (2013), Fougeron et al. (2021) 6

# Speech Rate is contingent on Utterance Length

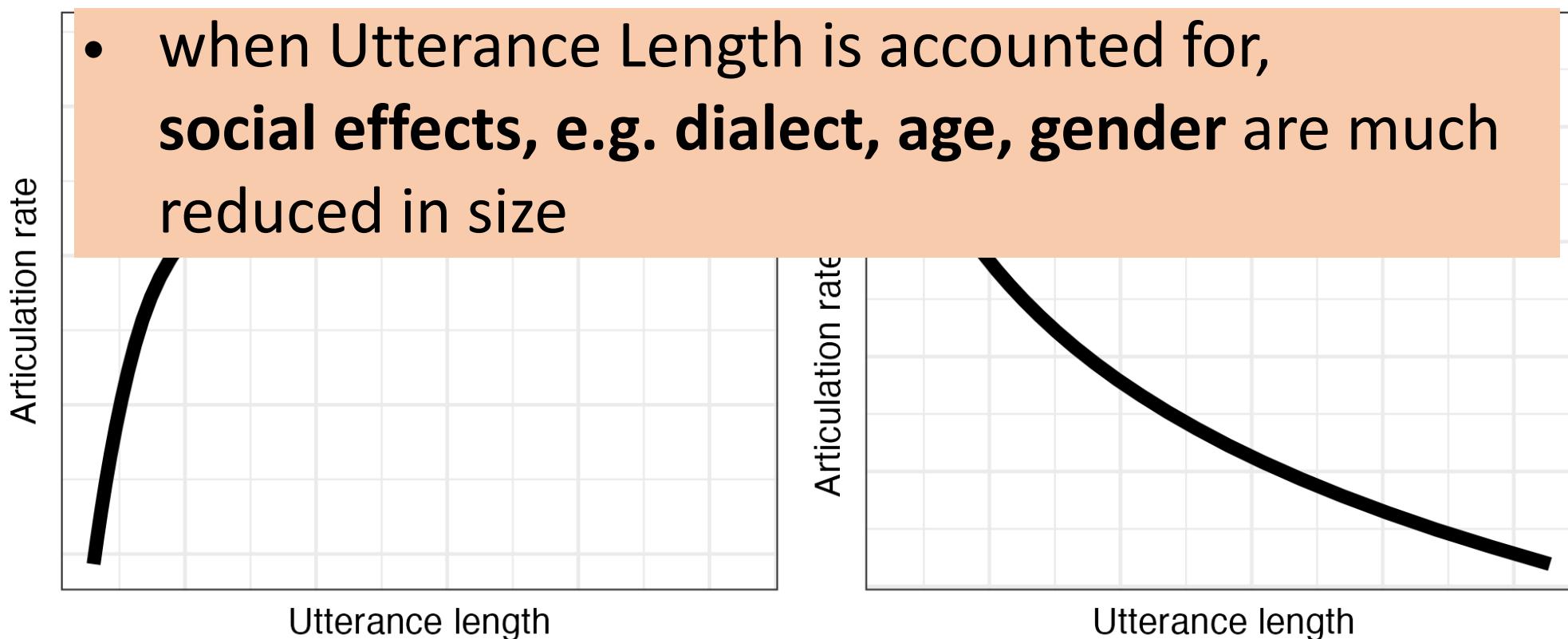
- **Utterance length (UL):** speech rate is *faster* and *less variable* for longer utterances



Altmann (1980), Quéne (2008), Kendall (2013), Jacewicz et al. (2010), White & Turk (2010), Bishop & Kim (2018)

# Speech Rate is contingent on Utterance Length

- **Utterance length (UL):** speech rate is *faster* and *less variable* for longer utterances



Altmann (1980), Quéne (2008), Kendall (2013), Jacewicz et al. (2010), White & Turk (2010), Bishop & Kim (2018)

# Speech Rate is contingent on Utterance Length

- **Utterance length (UL):** speech rate is *faster* and *less variable* for longer utterances

- when Utterance Length is accounted for, **social effects, e.g. dialect, age, gender** are much reduced in size
  - by Menzerath's Law, longer utterances result in shorter syllable durations, and so higher speech rate
- speech rate variation may largely arise from constraints on **speech production planning**

# Research questions for this study

- previous work on speech rate has considered only handful of dialects at a time, and most have not looked at Utterance Length
  - hence, still unclear the extent to which Utterance Length effects are driven by speech planning constraints vs learned social and/or stylistic variation
- To address this, we look at speech rate, alongside Utterance Length, in more, and a more diverse set of, dialects
1. How robust is the Utterance Length effect across dialects, and speakers?
  2. How robust are social effects on speech rate, after controlling for Utterance Length?

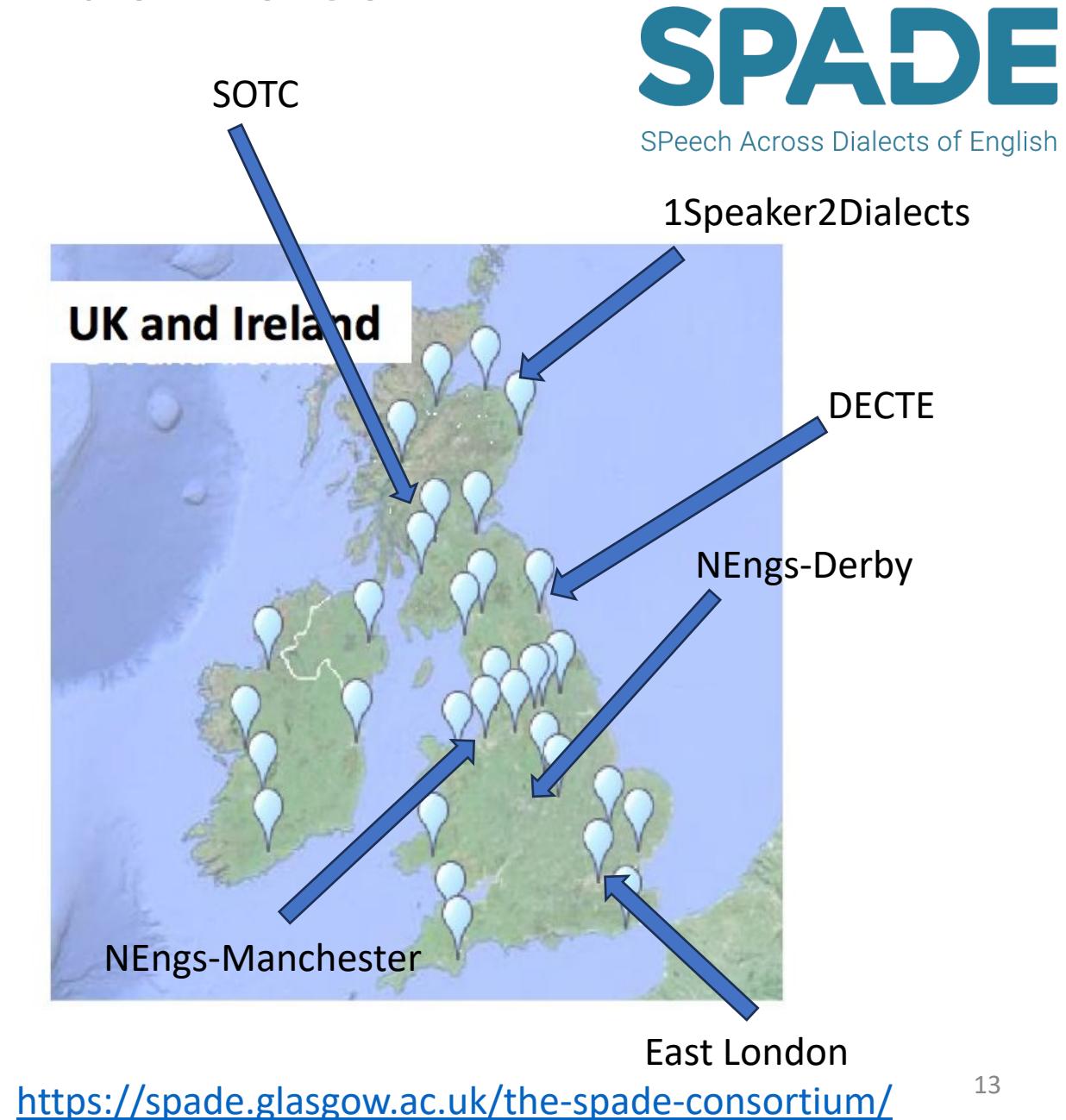
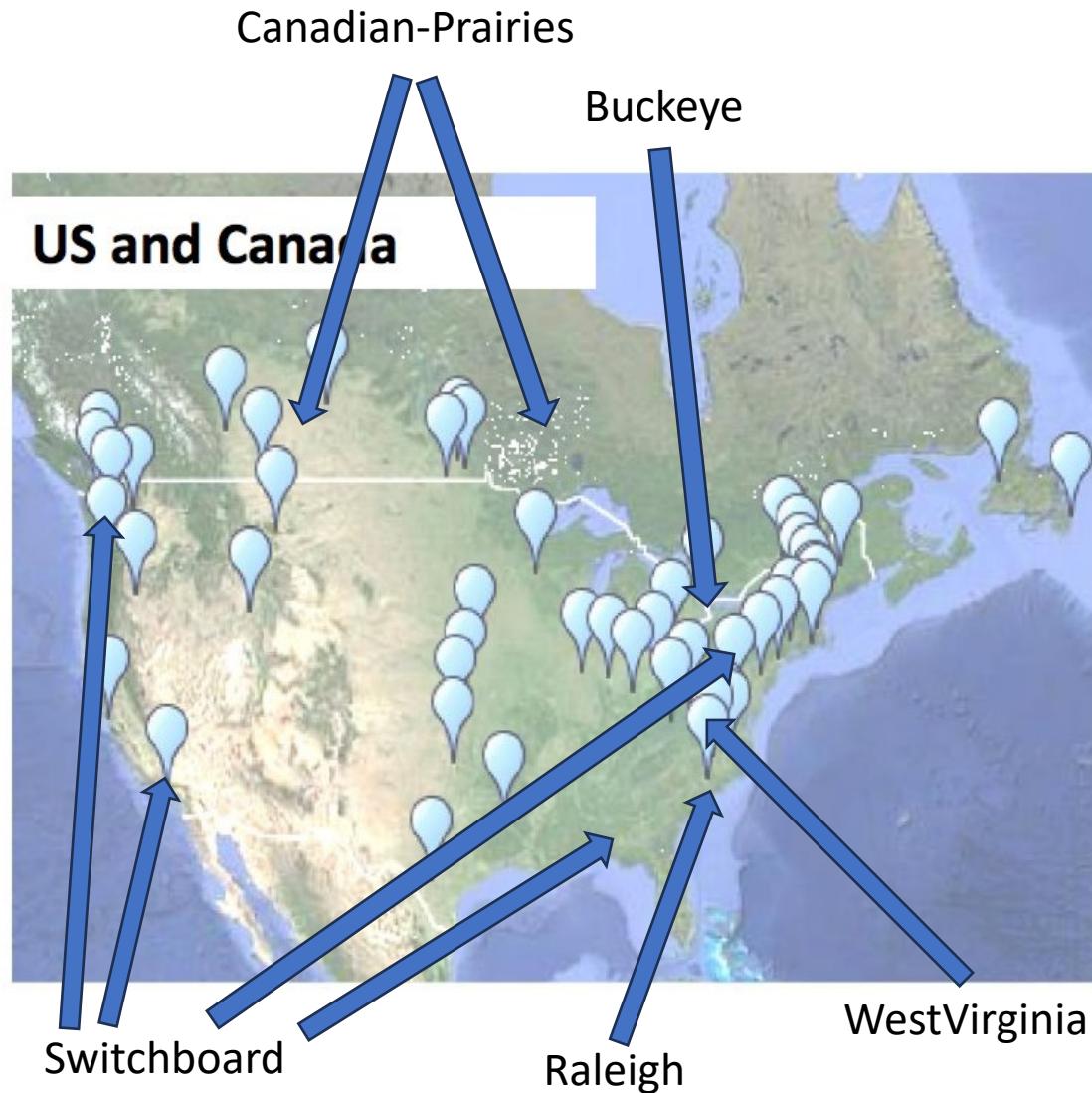
# Methods



# Data

- explores the relationship between Utterance Length and social effects (age, gender) on speech rate across a number of different dialects/ varieties of British Isles and North American English
- Here: varieties are represented by different speech corpora
  - each corpus treated as a different 'instance' of speech, that vary in several ways (dialect, recording period/context, age, style)

# 13 corpora of N American/British Isles English



# Data

- 116k utterances (150ms pause delimited)
  - Exclude utterances shorter than 3 syllables
- 1092 speakers (510 female-labelled)
- **Speech rate measure:** syllables per second within utterance, excluding pauses ('articulation rate')

Code & data: <https://osf.io/j9vny/>

# Model

- Bayesian distributional mixed-effects model (Stan/brms)
  - Model both *mean* ( $\mu$ ) and *variance* ( $\sigma$ ) of (log) articulation rate
- Fixed effects:
  - Utterance length (non-linear): speaker **mean & deviance** from mean
  - Gender
  - Age (z-scored)
- Random effects:
  - By corpora: same as fixed
  - By speaker: UL deviance

# Model

- Bayesian distributional mixed-effects model (Stan/brms)
  - Model both *mean* ( $\mu$ ) and *variance* ( $\sigma$ ) of (log) articulation rate
- Fixed effects:
  - Utterance length (non-linear): speaker **mean & deviance** from mean
  - Gender
  - Age (z-scored)
- Random effects:
  - By corpora: same as fixed
  - By speaker: UL deviance



Longer vs shorter utterances  
"on average" for a given speaker

# Model

- Bayesian distributional mixed-effects model (Stan/brms)
  - Model both *mean* ( $\mu$ ) and *variance* ( $\sigma$ ) of (log) articulation rate
- Fixed effects:
  - Utterance length (non-linear): speaker **mean & deviance** from mean
  - Gender
  - Age (z-scored)
- Random effects:
  - By corpora: same as fixed
  - By speaker: UL deviance

Length of the  
utterance *relative to that*  
speaker's mean

# Model

- Bayesian distributional mixed-effects model (Stan/brms)
  - Model both *mean* ( $\mu$ ) and *variance* ( $\sigma$ ) of (log) articulation rate
- Fixed effects:
  - Utterance length (non-linear): speaker **mean & deviance** from mean
  - Gender
  - Age (z-scored)
- Random effects:
  - By corpora: same as fixed
  - By speaker: UL deviance

*RQ1: How robust is the  
Utterance Length effect across  
dialects, and speakers?*

# Model

- Bayesian distributional mixed-effects model (Stan/brms)
  - Model both *mean* ( $\mu$ ) and *variance* ( $\sigma$ ) of (log) articulation rate
- Fixed effects:
  - Utterance length (non-linear): speaker **mean & deviance** from mean
  - Gender
  - Age (z-scored)
- Random effects:
  - By corpora: same as fixed
  - By speaker: UL deviance

*RQ2: How robust are social effects on speech rate, after controlling for Utterance Length?*



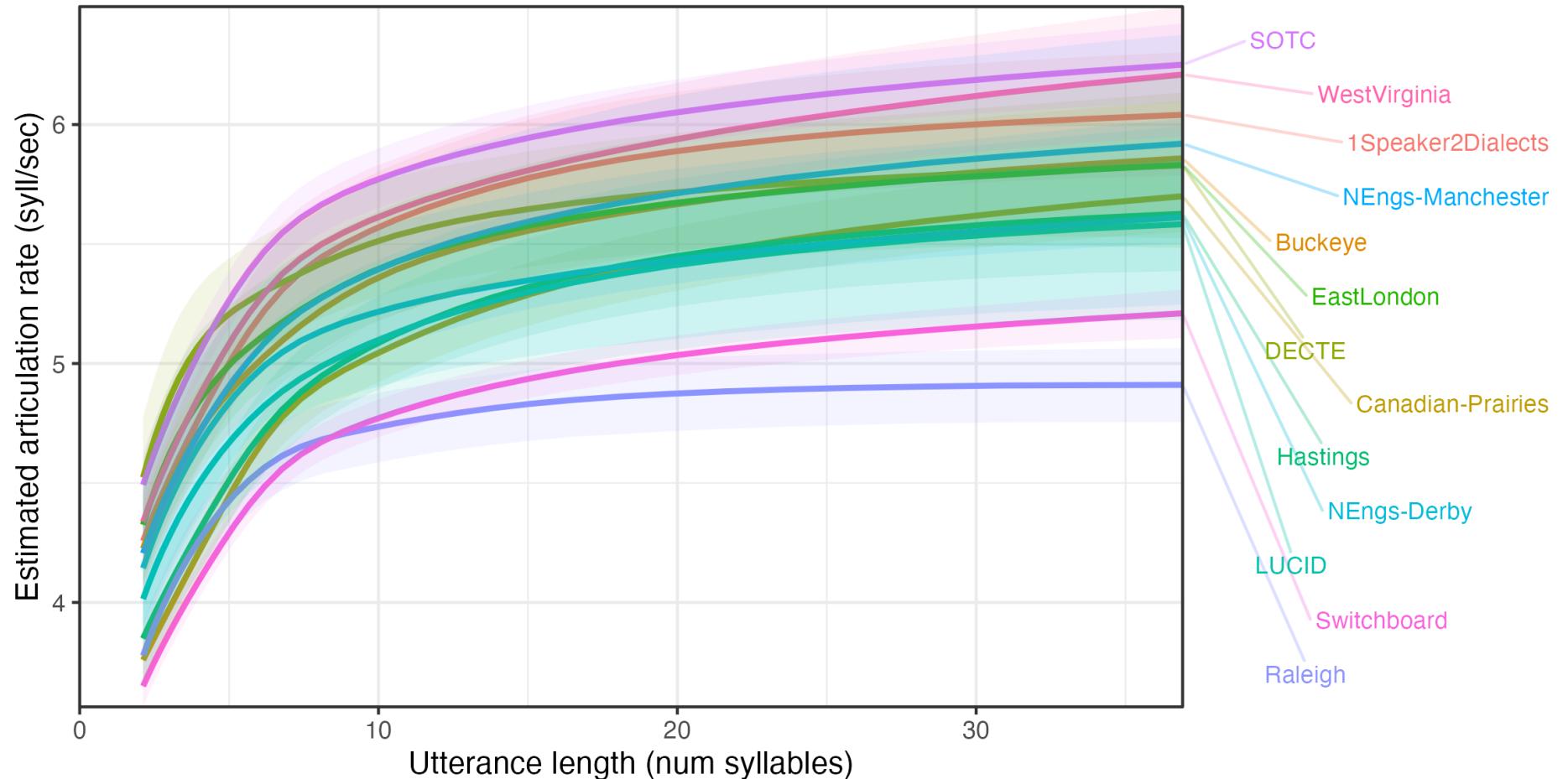
# Results



# Q1: UL effect on rate across corpora

## Mean rate

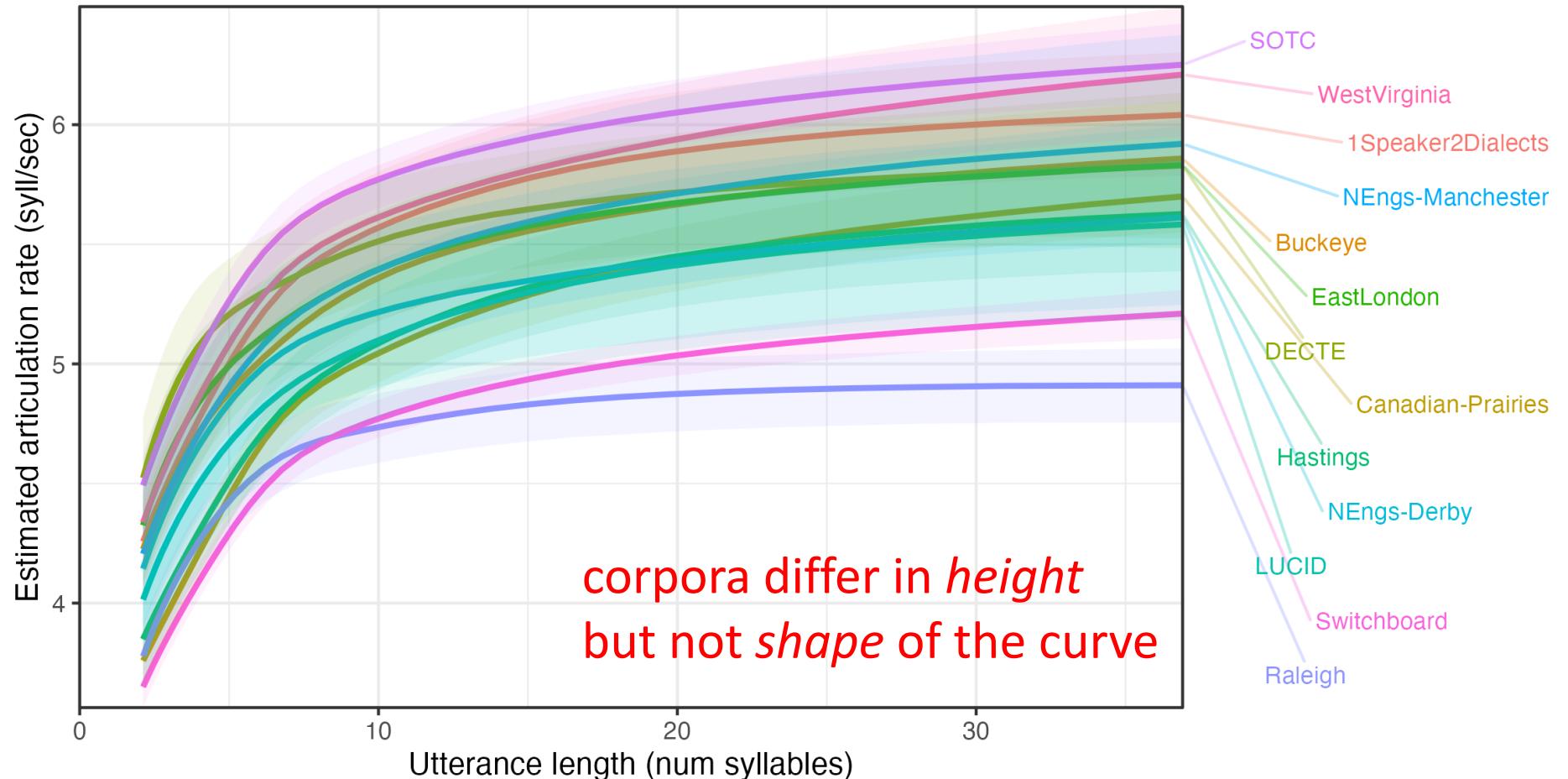
- Longest utterances are approx 33% faster than the shortest utterances
- UL effect similar across corpora
- corpora themselves differ in average rate



# Q1: UL effect on rate across corpora

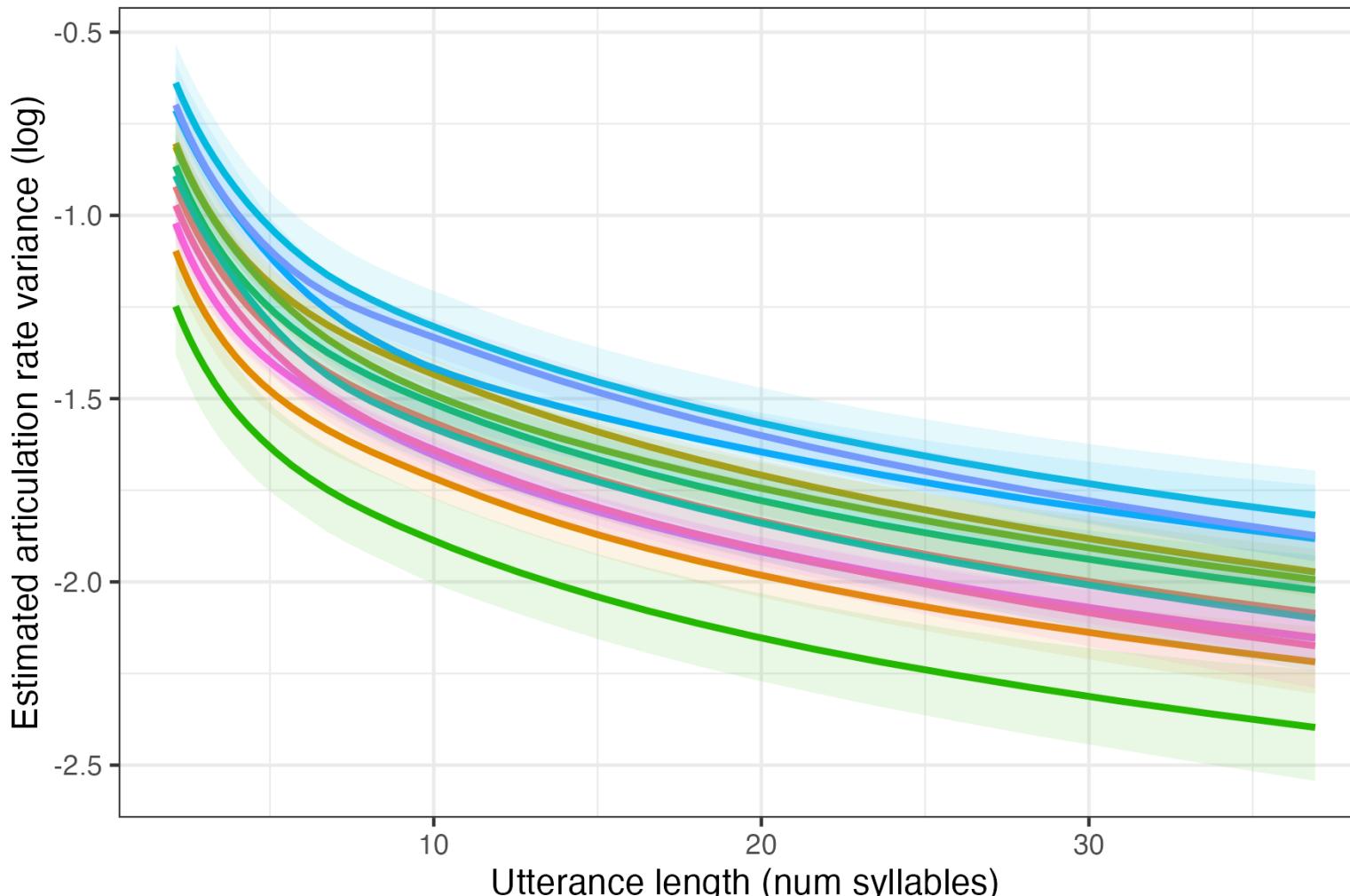
## Mean rate

- Longest utterances are approx 33% faster than the shortest utterances
- UL effect similar across corpora
- corpora themselves differ in average rate



# Q1: UL effect on rate across corpora

## Rate variance



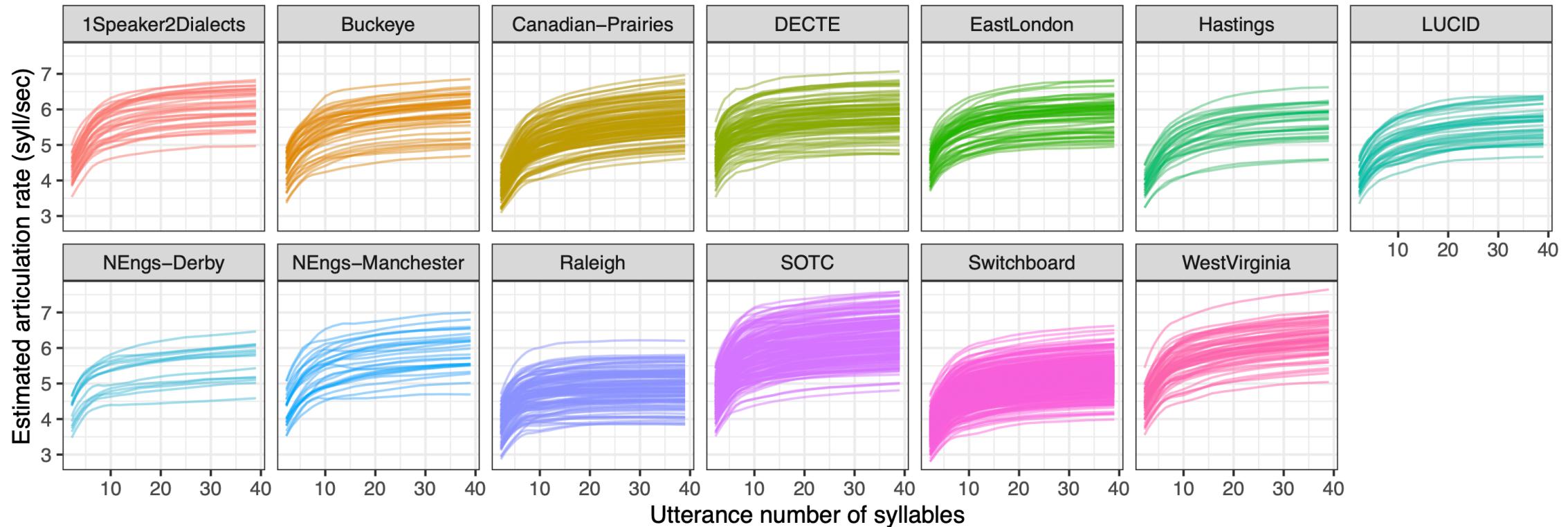
Less *variability* in articulation rate for longer utterances

Longest utterances are 1/3 less-variable than shortest

Corpora vary little in the UL effect on rate *variance*

# Q1: UL effect on rate across corpora & speakers

## Individual speakers

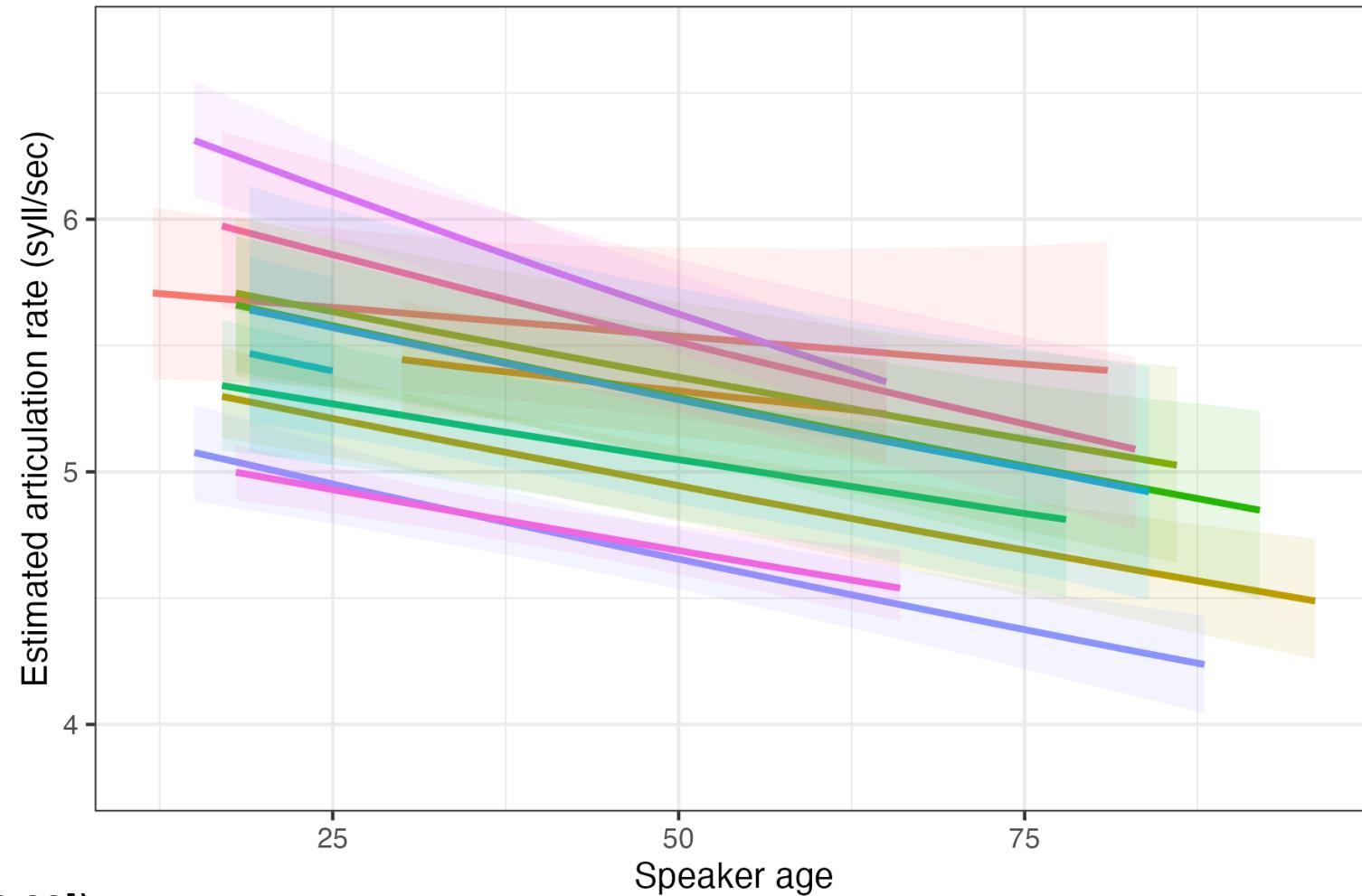


- Utterance length effect *similar* across speakers
- Individual speakers vary a lot from their corpus average

## Q2: social effects across corpora on speech rate

### Age

- Age has negative effect on articulation rate = older speakers have slower rates (on average)
- Age effect itself differs little by corpora
- Age effect is smaller than UL effect (15% vs 33%)

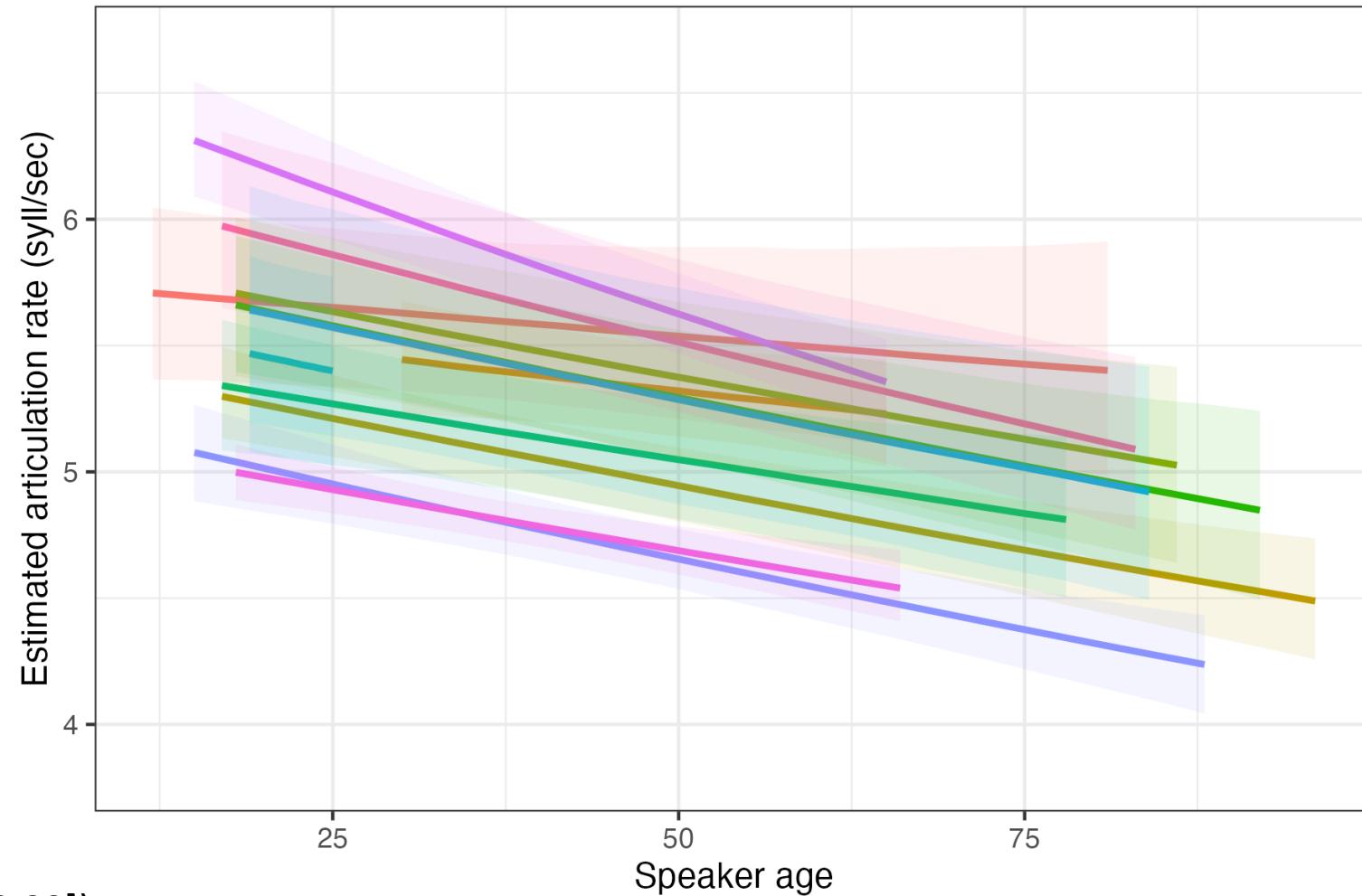


$$(\beta_{\text{age}} = -0.04, 95\% \text{ CrI} = [-0.05, -0.02])$$

## Q2: social effects across corpora on speech rate

### Age

- Age has negative effect on articulation rate = older speakers have slower rates (on average)
- Age effect itself differs little by corpora
- Age effect is smaller than UL effect (15% vs 33%)

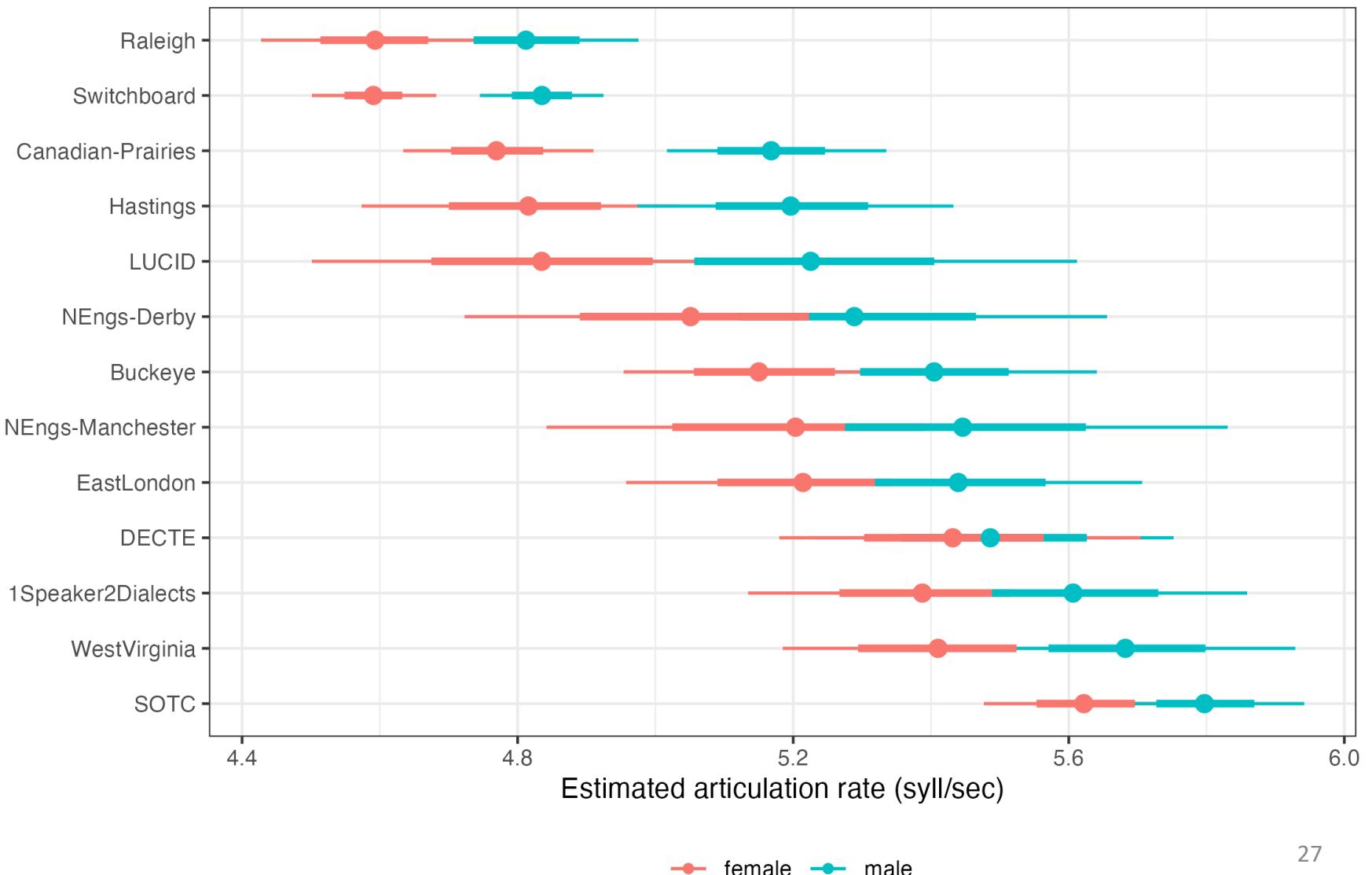


$$(\beta_{\text{age}} = -0.04, 95\% \text{ CrI} = [-0.05, -0.02])$$

# Q2: social effects across corpora on speech rate

## Gender

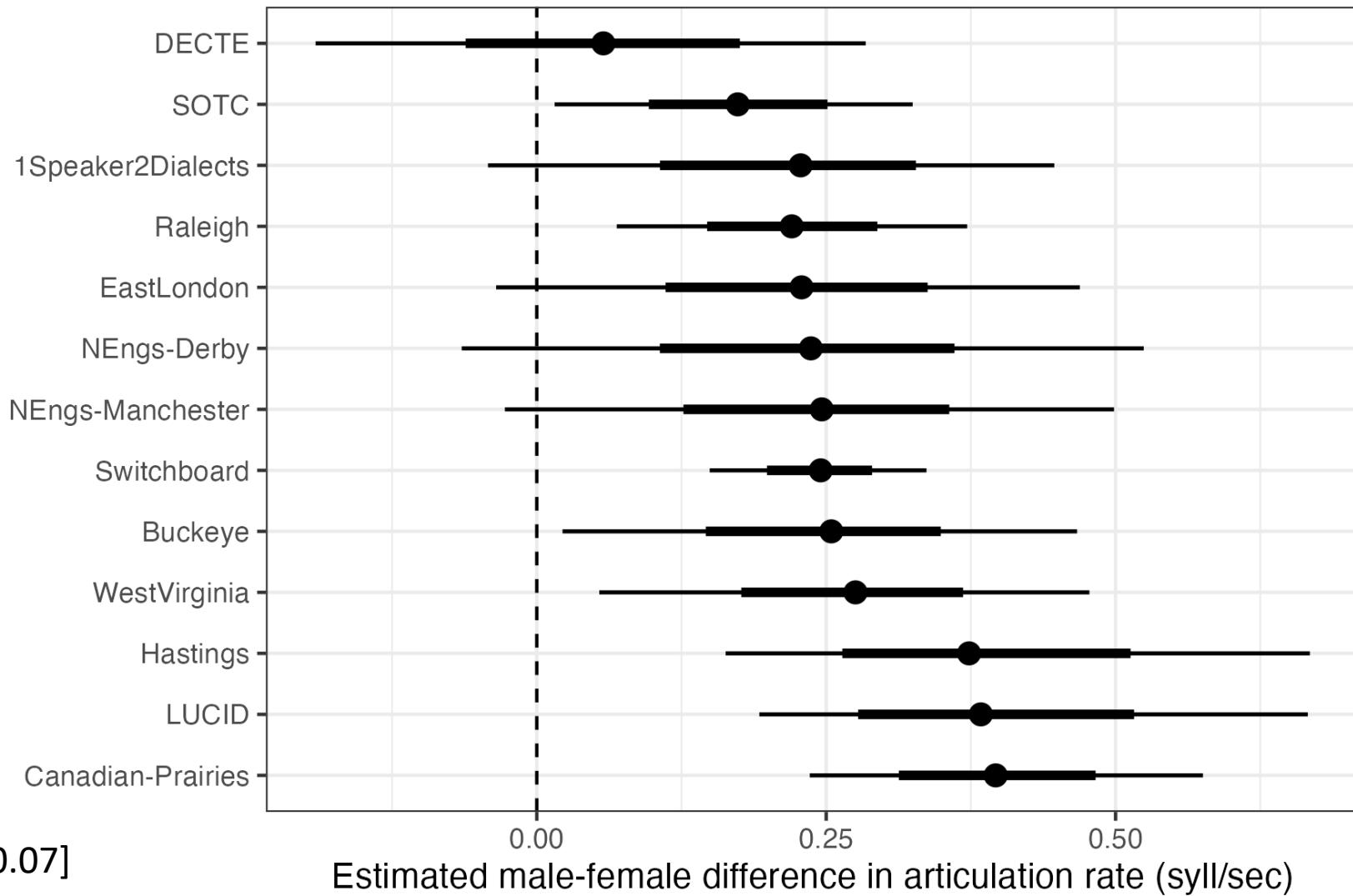
- Female speakers speak more slowly than males
- Corpora differ in *size* but *not direction* of the gender effect
- Gender effect *much* smaller than UL effect



# Q2: social effects across corpora on speech rate

## Gender

- Female speakers speak more slowly than males
- Corpora differ in *size* but *not direction* of the gender effect
- Gender effect *much* smaller than UL effect

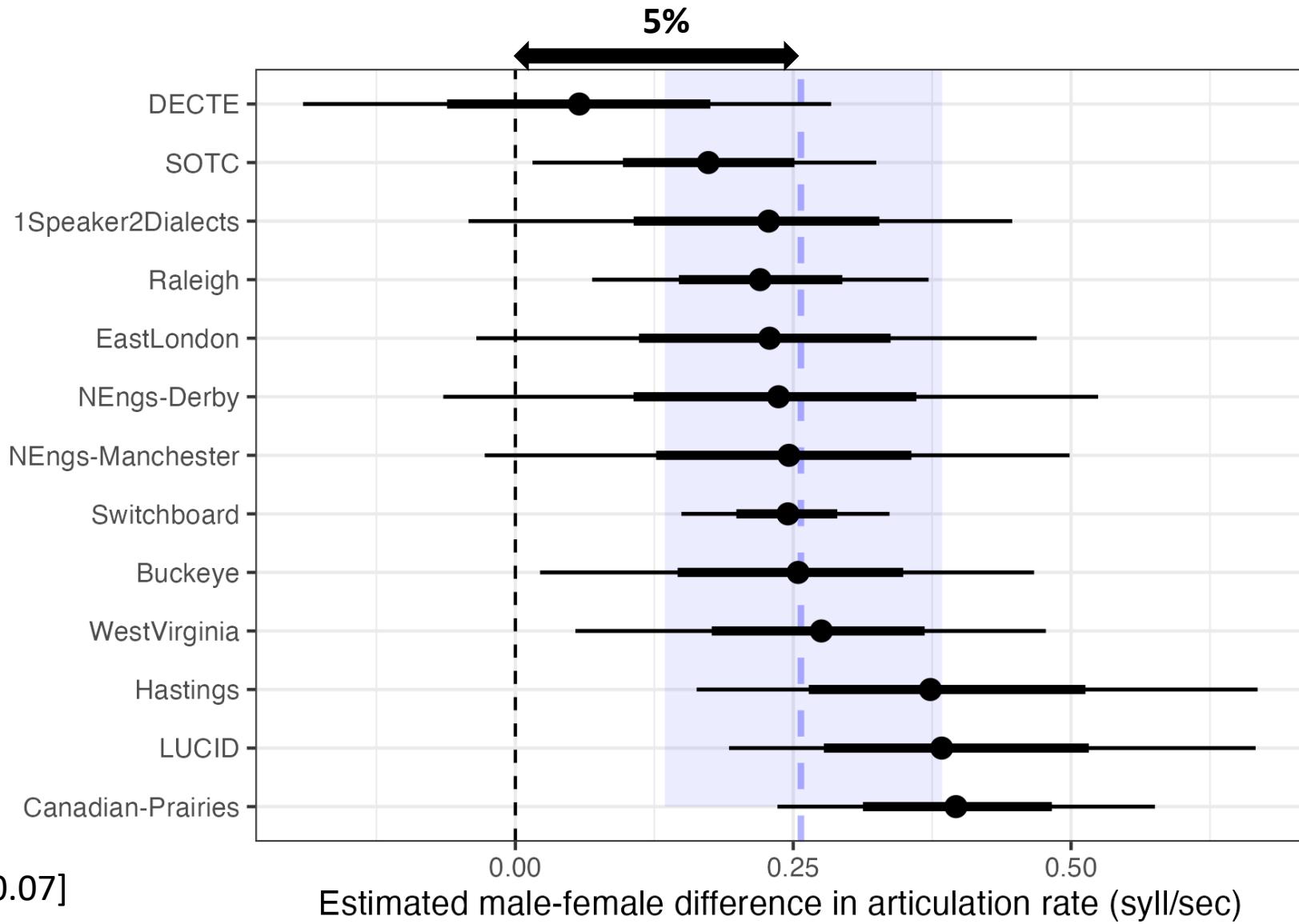


Corpora with effect sizes including 0 (DECTE, 1S2D, EastLondon, NEngs) also those with fewest observations

# Q2: social effects across corpora on speech rate

## Gender

- Female speakers speak more slowly than males
- Corpora differ in *size* but *not direction* of the gender effect
- Gender effect *much smaller than* UL effect



Corpora with effect sizes including 0 (DECTE, 1S2D, EastLondon, NEngs) also those with fewest observations

# Discussion



# Summary

1. How robust is the Utterance Length effect on speech rate across dialects and speakers?
  - Utterance length has large consistent effect across corpora and speakers

# Summary

1. How robust is the Utterance Length effect on speech rate across dialects and speakers?
  - Utterance length has large consistent effect across corpora and speakers
2. How robust are social effects on speech rate, after controlling for Utterance Length?
  - size of age and gender effects present and similar across corpora
  - social effects, while present, much smaller in magnitude
  - Utterance Length by far the largest predictor of articulation rate variation

# Scaling up the study of speech rate

examining speech rate across more, and more diverse varieties...

- confirms the key effect of utterance length on speech rate (longer utterances are faster and less variable)
- demonstrates the *consistency* in the nature of the UL effect across corpora & speakers

# What drives the Utterance Length effect?

- robust Utterance Length effect across many corpora is consistent with Menzerath's Law operating on speech rate
  - As utterance length increases, syllable duration decreases
- this suggests speech rate variation is *primarily* driven by articulatory and/or perceptual constraints
- to what extent is Utterance Length contingent on and/or confounded with, other kinds of linguistic organization of timing, e.g. prosodic structure?

# Speech rate and social factors

- social effects are small, but consistently present across corpora
- age: older speakers are always slightly slower
  - this cannot be accounted for by older speakers showing shorter utterances
  - is this the result of physiological aging?
- gender: females are always slightly slower
  - again, females don't show shorter utterances
  - to what extent is the gender effect a result of gendered speaking styles?
- So where do social stereotypes around speech rate actually come from?
  - articulation rate is only *one* of the measures capturing speech rate – pause frequency and duration likely important in impressions of rate differences

# Thank you!

Code & data: <https://osf.io/j9vny/>



# SPADE

SPeech Across Dialects of English

# Investigators



THE UNIVERSITY  
of EDINBURGH



University  
of Glasgow



McGill



NC STATE  
UNIVERSITY



UNIVERSITY OF  
OREGON



<http://spade.glasgow.ac.uk/>

# SPADE

SPeech Across Dialects of English



James Tanner  
Postdoc



Rachel Macdonald  
Project Manager



Michael McAuliffe  
Software Development



Arts & Humanities  
Research Council



SSHRC CRSH

Social Sciences and Humanities Research Council of Canada  
Conseil de recherches en sciences humaines du Canada



# Extra slides

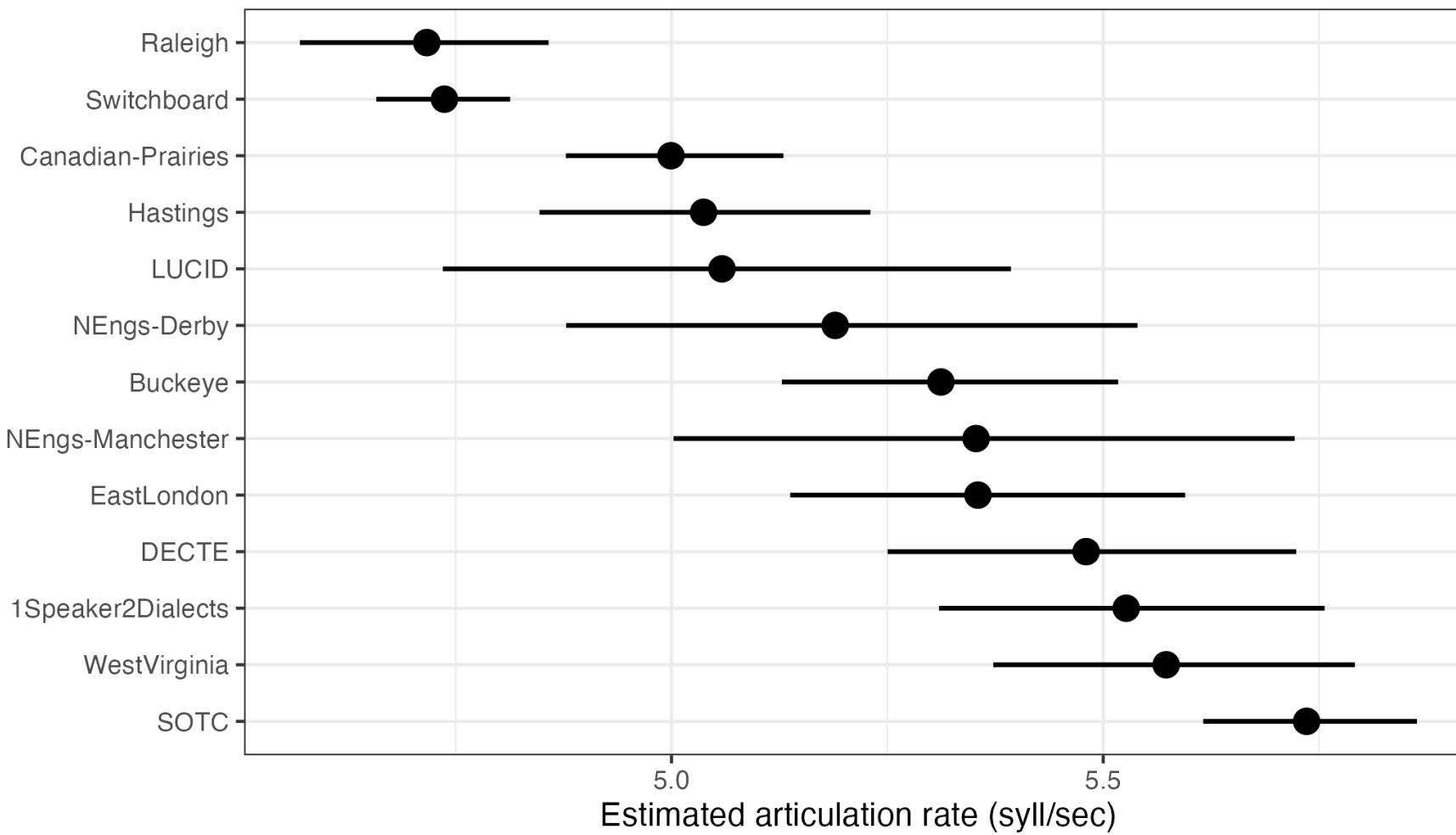


# Data table

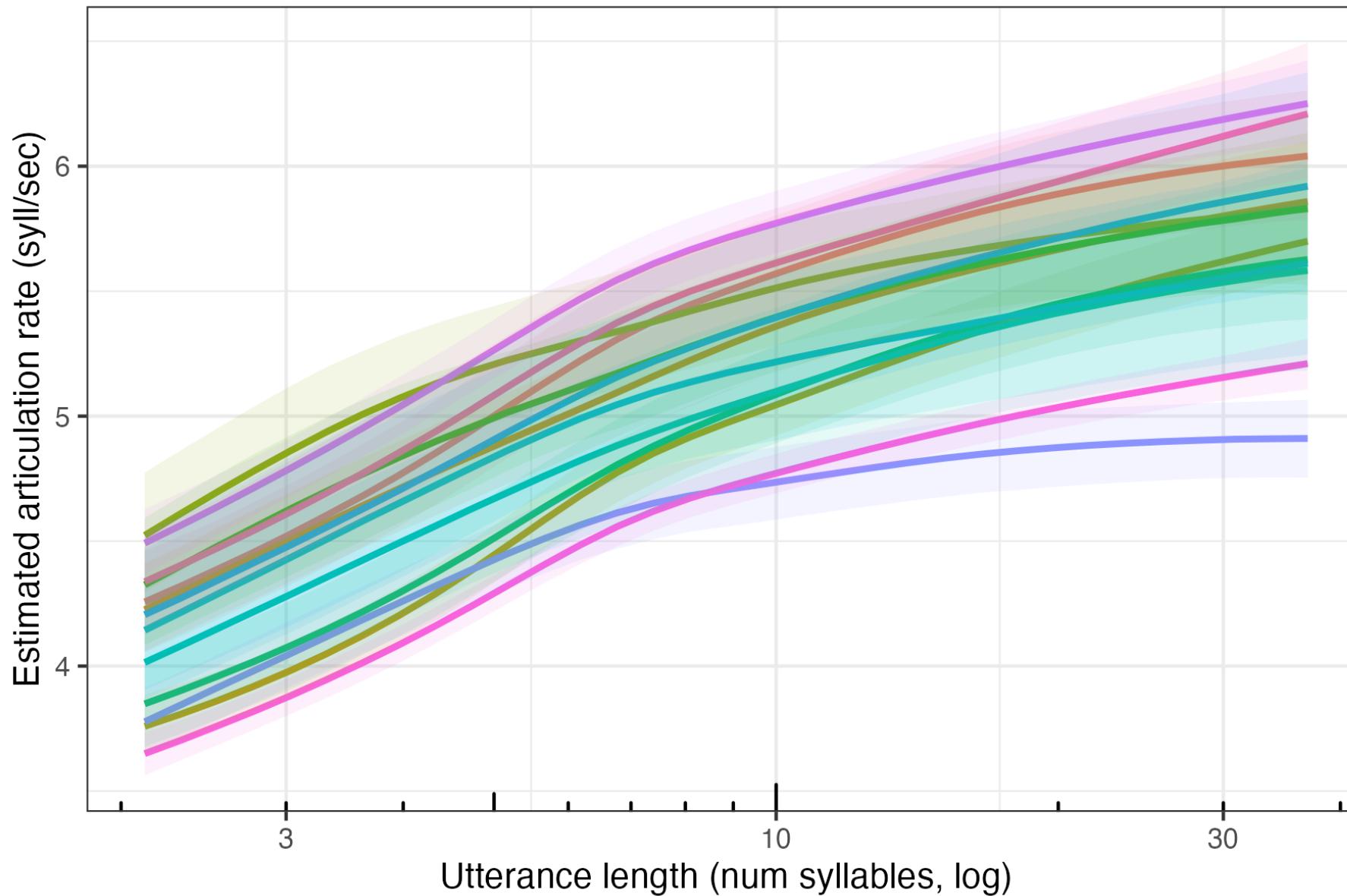
Table 1: *Summary of data used in final analysis: region, speech style/context, number of speakers (female), mean age (standard deviation), and number of utterances (N) by corpus.*

| Corpus                 | Region                      | Style                        | Speakers (F) | Age (SD) | N     |
|------------------------|-----------------------------|------------------------------|--------------|----------|-------|
| 1Speaker2Dialects [24] | NE Scotland (UK)            | Sociolinguistic interviews   | 31(14)       | 42 (27)  | 12640 |
| Buckeye [25]           | Ohio (US)                   | Sociolinguistic interviews   | 40(20)       | 48 (17)  | 7896  |
| Canadian Prairies [26] | Alberta & Manitoba (Canada) | Sociolinguistic interviews   | 108(58)      | 41 (20)  | 18379 |
| DECTE [27]             | Newcastle (UK)              | Interviews, conversation     | 82(46)       | 37 (22)  | 2928  |
| East London            | London (UK)                 | Sociolinguistic interviews   | 57(24)       | 35 (27)  | 327   |
| Hastings [28]          | SE England (UK)             | Sociolinguistic interviews   | 31(11)       | 49 (25)  | 6618  |
| LUCID [29]             | London (UK)                 | Structured conversation task | 40(20)       | 23 (0)   | 2919  |
| NEngs-Derby [30]       | Derby (UK)                  | Sociolinguistic interviews   | 14(7)        | 22 (2)   | 1819  |
| NEngs-Manchester [30]  | Manchester (UK)             | Sociolinguistic interviews   | 27(14)       | 44 (24)  | 2338  |
| Raleigh [31]           | North Carolina (US)         | Spontaneous conversation     | 100(50)      | 51 (19)  | 18203 |
| SOTC [32]              | Glasgow (UK)                | Sociolinguistic interviews   | 162(63)      | 44 (21)  | 18536 |
| Switchboard [33]       | Multiple locations (US)     | Telephone conversations      | 339(152)     | 36 (11)  | 18855 |
| West Virginia [34]     | West Virginia (US)          | Sociolinguistic interviews   | 61(31)       | 39 (22)  | 4562  |

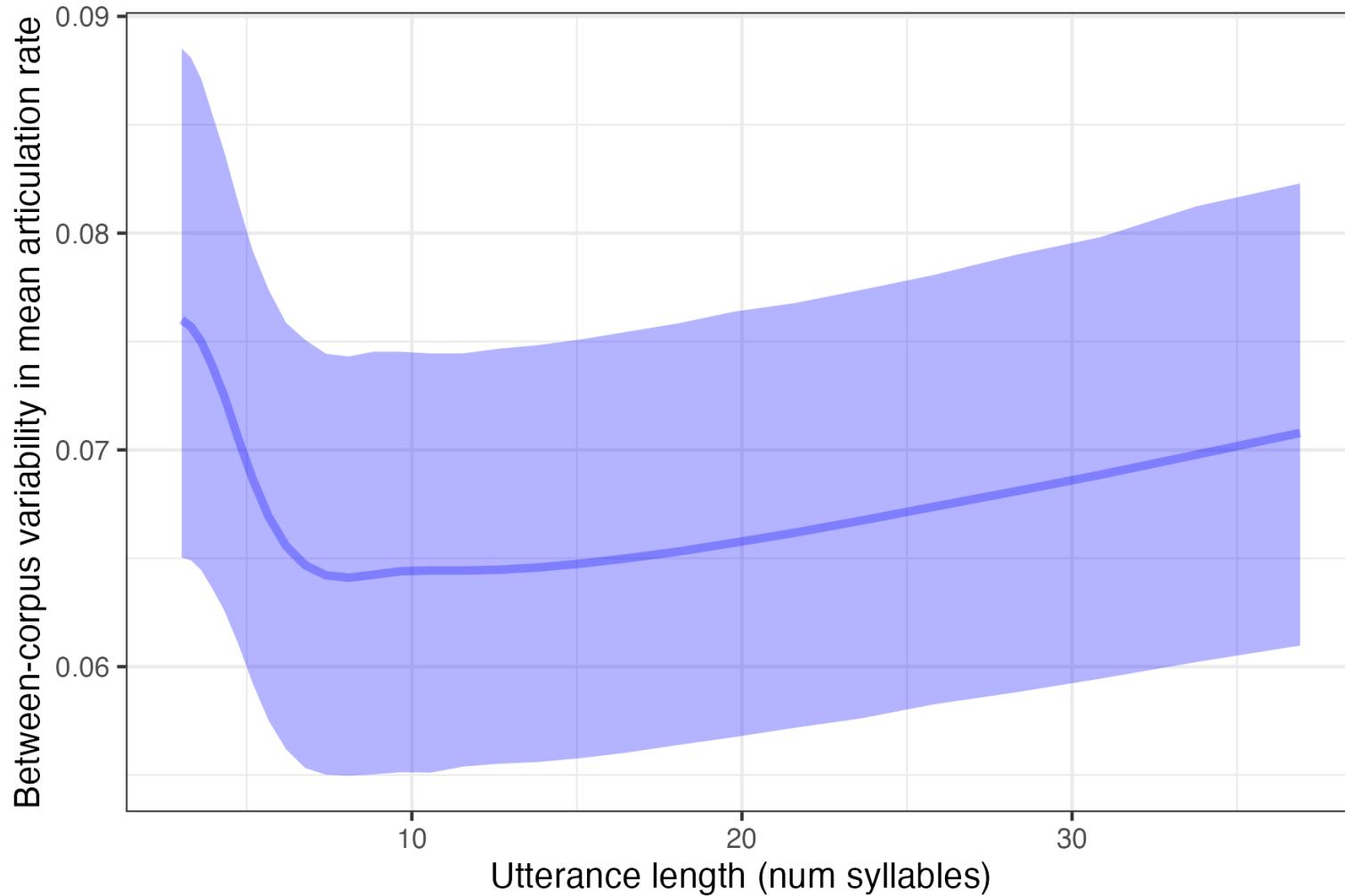
# Estimated articulation rate across corpora



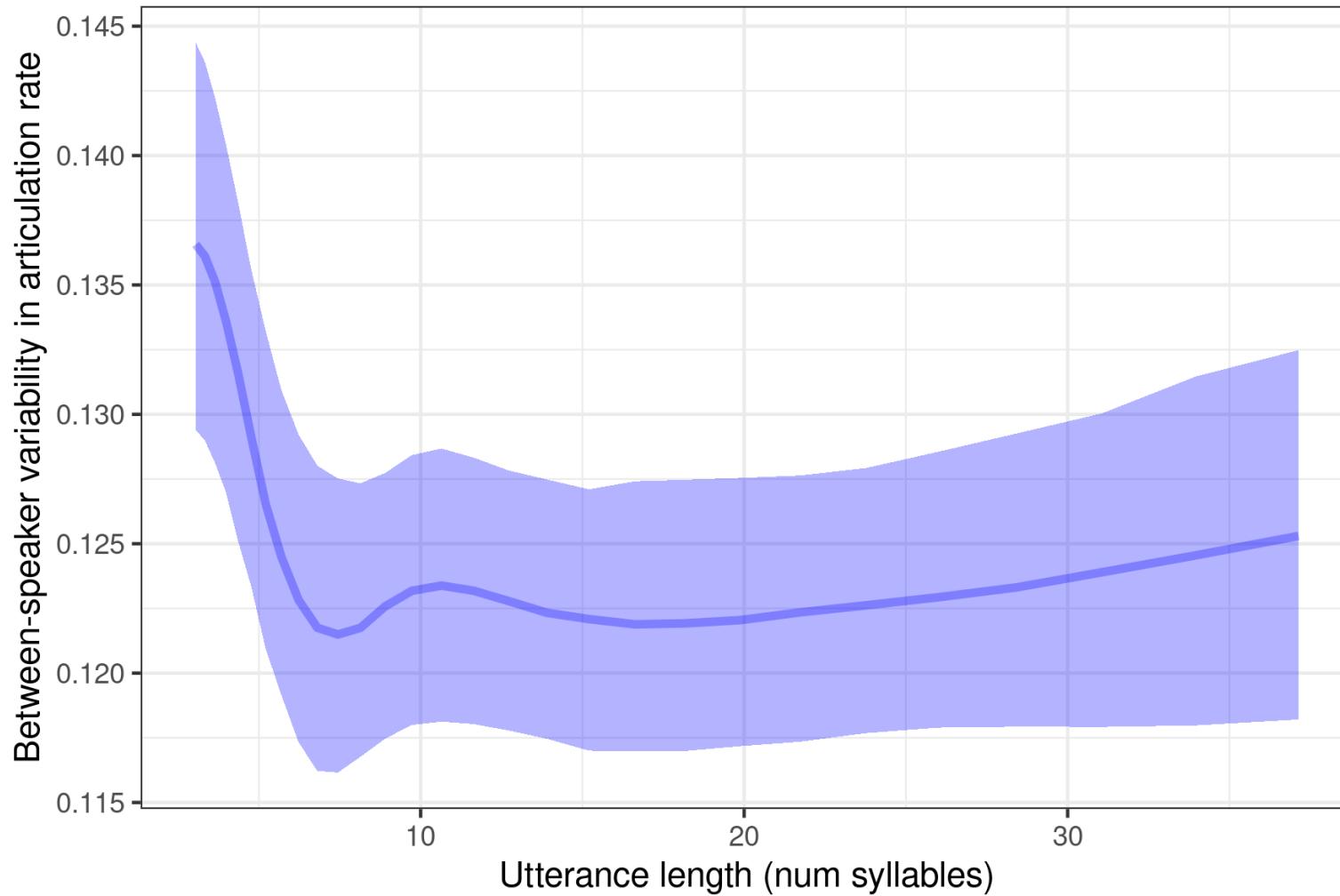
# UL effect across corpora (log)



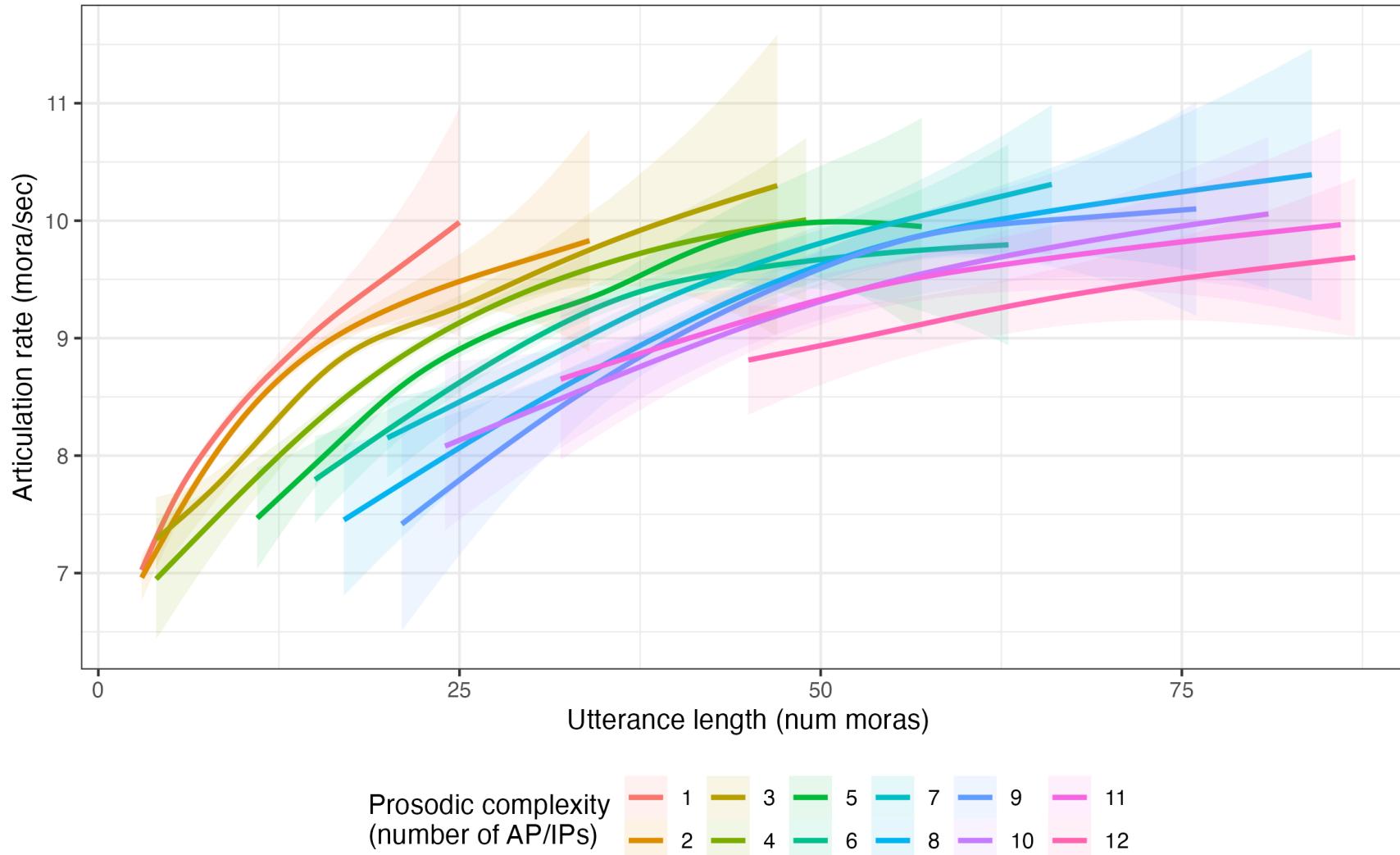
# Between-corpus variability in *mean* rate



# Between-speaker variability in *mean* rate



# Japanese articulation rate by prosodic complexity of the utterance



# Introduction

- speech rate
  - measure of speech timing – the tempo at which linguistic events unfold

 **Just Tracy**  @TracyOnFire2020 · May 5, 2022

Replying to @\_bee\_kay\_

Well I live in the south and that's not how Southerners sound. Southerners talk very slowly, often so **slow** that if they were talking any slower they'd be talking backwards. I have no idea what tribe this cheetah comes from.

3 1 29

 **Jamey\_Giddens** @Jamey\_Giddens · Dec 13, 2013

I am from the South. It is regionally impossible for us to talk as fast as these folks are talking. Our brains would explode. **#Scandal**

6 8 11

and I will be there, nose pressed against the screen, ears straining to understand fast talking northern english accents

