# Segmental influences on F0:
## cross-linguistic and interspeaker variability of phonetic precursors

*Morgan Sonderegger,*
*Michael McAuliffe, Hye-Young Bang*
McGill University


4th Workshop on Sound Change

# Introduction

- Common in sound change:

phonetic precursors → phonological pattern

F0 perturbations around p/b — lexical tone

What kind of precursor can be a source of change?

- robust
  - Across speakers, languages (e.g. Hombert et al., 1979, Ohala)
- … but variable
  - Individual differences, language-specific phonetics

(e.g. Baker et al., 2011; Labov, 1967; Kingston, 2007; Yu, 2013; this whole conference)

tension

# Introduction

How robust/variable is each phonetic precursor, across languages and individuals?

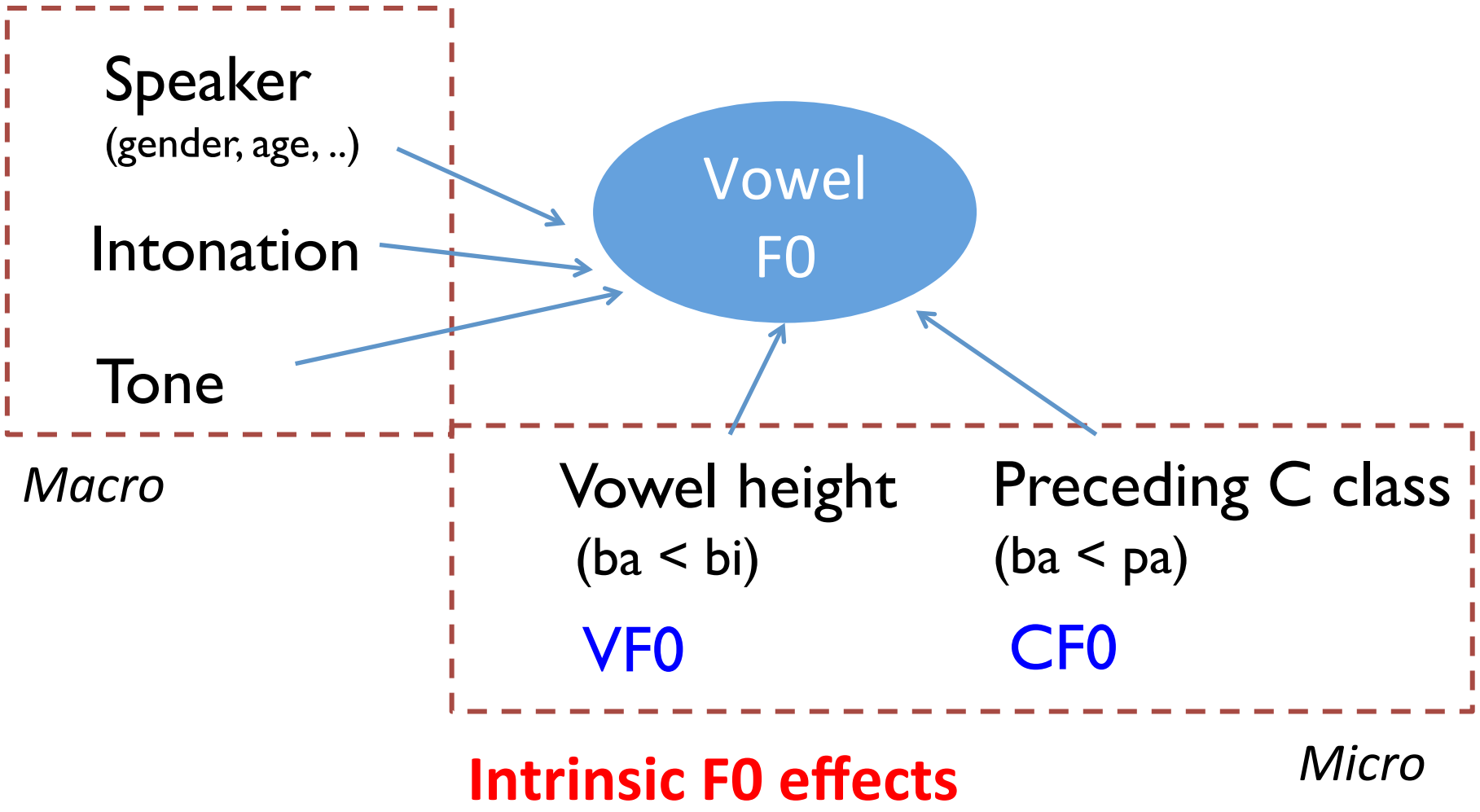
# Introduction

- Methodologically hard
  - big and comparable data: many languages, speakers
  - small effects, big confounds
  - ideally: non-lab speech

- Solution?
  cross-linguistic corpora + automatic analysis + statistical modeling

- Q1: can a "phonetic precursor" be detected in corpus data across languages & speakers?

# Influences on vowel F0




(e.g. Chen, 2011; Connell 2002; Fischer-Jørgenson, 1990; Hanson, 2009; Hoole & Honda, 2011; House & Fairbanks, 1953; Kingston & Diehl, 1994; Kirby & Ladd, 2016; Kingston, 2007; Ladd & Silverman, 1994; Meyer, 1896; Whalen & Levitt, 1995)

# Intrinsic F0

- Huge literature
  - primarily: small n, lab speech
  - focus: mechanism (automatic/controlled)

Across languages:

- CF0
  - "voiced"<"voiceless": most languages
- VF0
  - [-high] < [+high] : (near-)universal
- Effect size: variable
  - Tonal ⇒ smaller effect?

Q1: How much variability in IF0 across 14 languages?

# Intrinsic F0

- Strongly affected by:
  - "Intonation"
  - Gender (VF0)
  - …
- Interspeaker variability:
  - Often noted
- Relationship to sound change:
  - CF0 ⇒ tonogenesis
  - VF0 ⇏ sound change
  - Why?

Q2: How much variability in IF0 across speakers?

# Datasets

| | |
|---|---|
| English | Russian |
| French | Polish |
| German | Spanish |
| Korean | Turkish |

| |
|---|
| Hausa |
| Mandarin |
| Thai |
| Vietnamese |

- Read sentence corpora
  - ~20 hours each
  - Force-aligned

**Montreal Forced Aligner**: trainable for different languages

| |
|---|
| Croatian |
| Swedish |

GlobalPhone (Schulz et al., 2013), Librispeech (Panyatov et al., 2015)

# Datasets

- "Utterance-initial"

  > 150 ms pause or file-initial

C V

obstruent /a/, /i/, /u/

- vowel F0 (Praat)
  - F0 histogram → speaker min, max → re-extract F0
- Other info:
  - Speaker: ID, gender, mean F0
  - Utterance: length (syllables)
  - Surrounding segments
  - Word

Polygot-Speech Corpus Tools

http://speech-corpus-tools.readthedocs.io/

McAuliffe et al. (2017)

# Datasets

- Data cleaning: minimize F0 errors, reduced vowels

- Exclusions:
  - Speakers: Multimodal F0 distribution (non-tonal langs)
  - Vowel tokens:

< 50 msec  < 50% voiced

Extreme values of DV, within-speaker

- Data per language:
  - 1.9-9.5k tokens (~2000)
  - 76-132 speakers (100)

# CF0: Analysis

- One linear mixed effects model / language
- Main terms:

Response: mean F0 in first 50 ms

$C_1$ V X

consonant "voicing"*

overall effect + interspeaker variability

fixed effect by-speaker random slope

* Ex: French p/b, Mandarin p/$p^h$

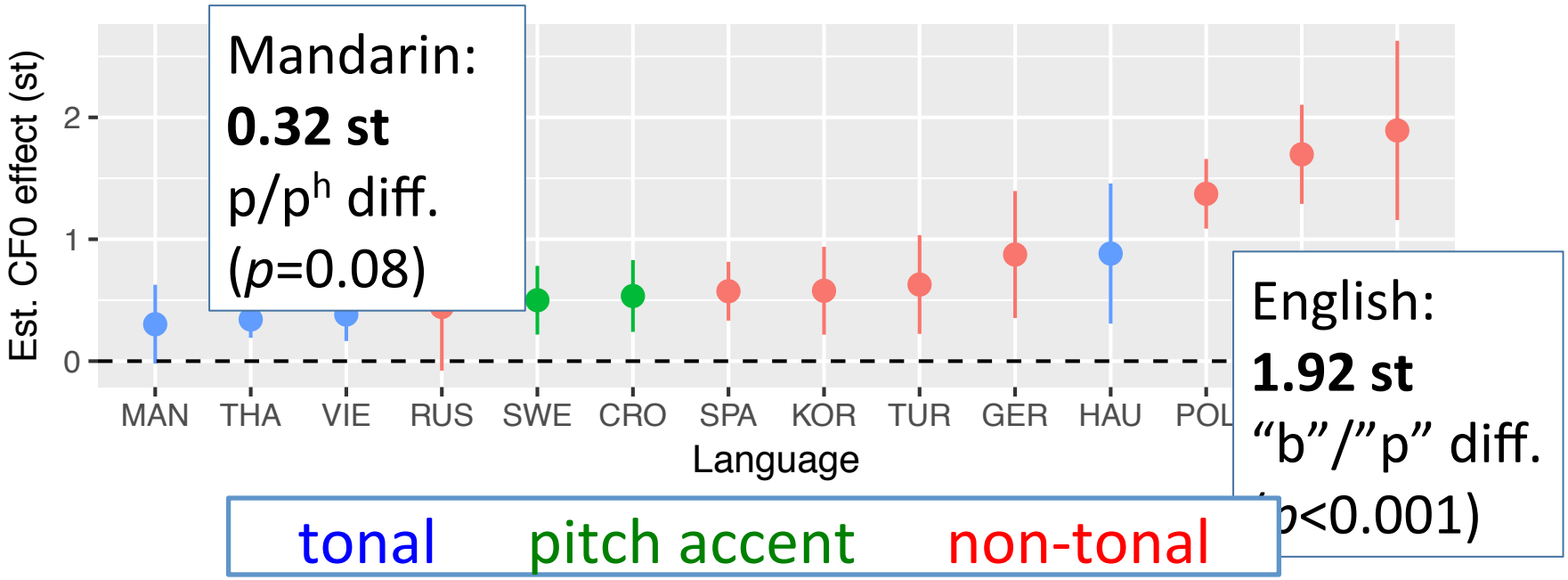
# CF0: analysis

- Other terms
  - "Voicing" interactions: gender
  - Controls:
    - Speaker gender, mean F0
    - Utterance length
    - V identity (incl. height)
    - Speaker, word, preceding/following phone

- Conservative model structure

# CF0: across languages

- "most voiceless" – "most voiced" effect:

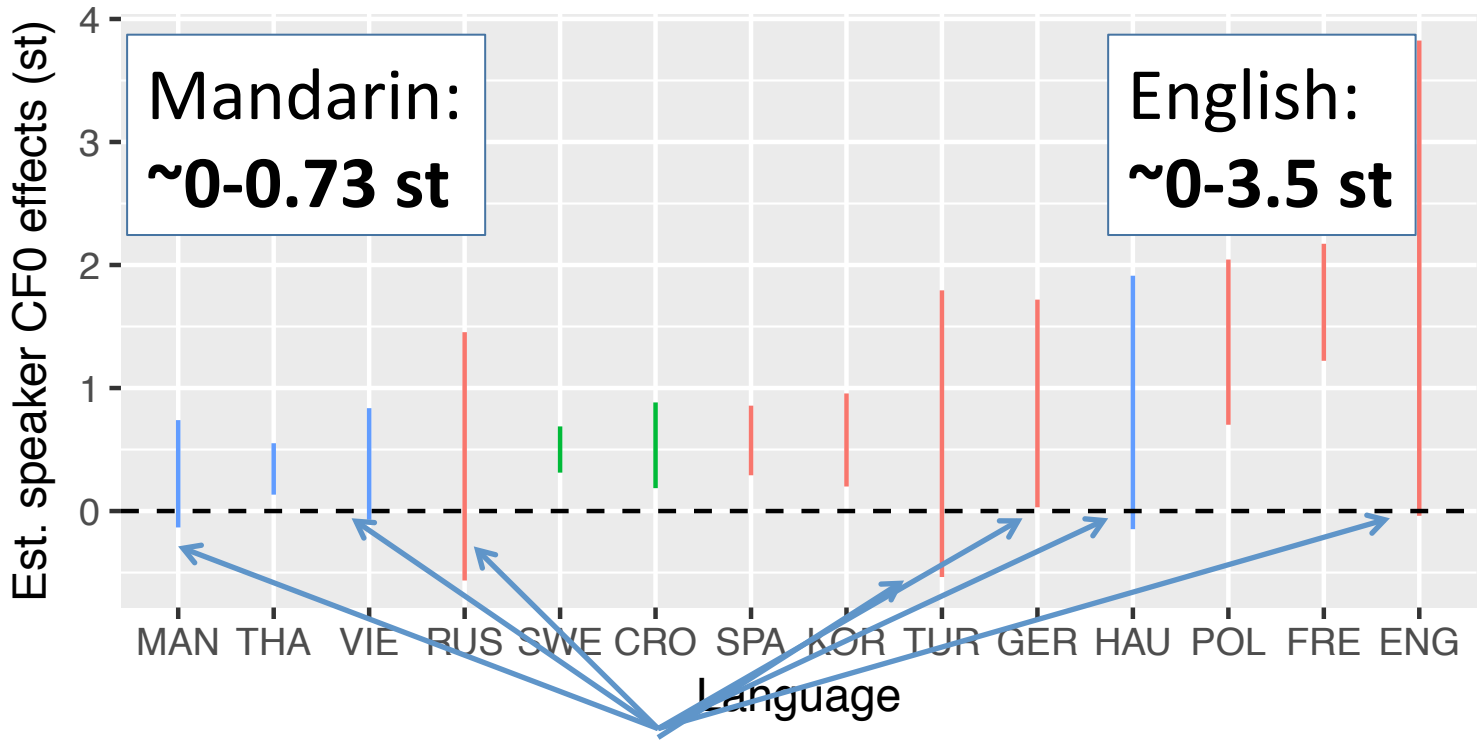



- Robust across languages
- Variable effect size
  - Non-tonal ⇒ larger effect

Average effect across gender, tone, etc.

# CF0: across speakers

- Predicted effects for 95% of individuals:

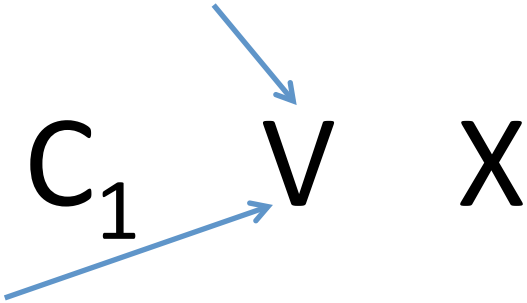



- Common: large interspeaker variability

# VF0: Analysis

- One linear mixed effects model / language
- Main terms:

Response: mean F0




Vowel identity
Height (a vs. i/u)
+ i vs. u

overall effect + interspeaker variability

fixed effect

by-speaker random slope

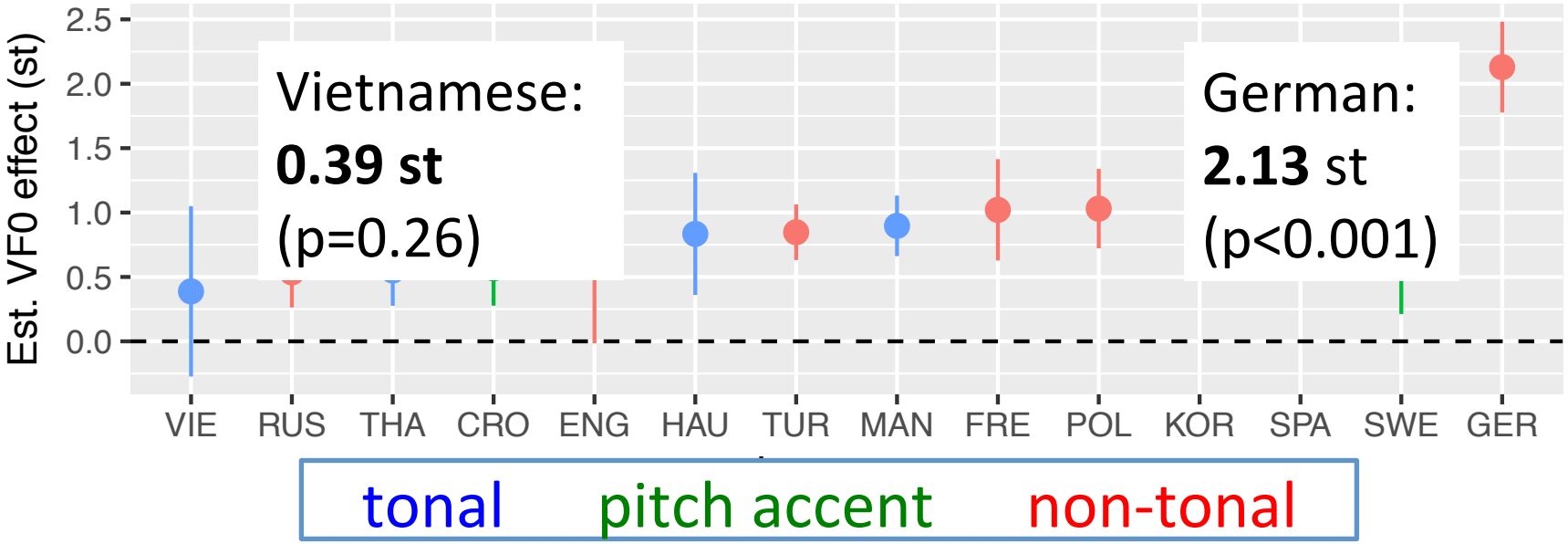
# VF0: analysis

- Other terms
  - V height interactions: gender
  - Controls:
    - Speaker gender, mean F0
    - $C_1$ "voicing"
    - Utterance length
    - V identity
    - Speaker, word, preceding/following phone

- Conservative model structure

# VF0: across languages

- High – low vowel effect:

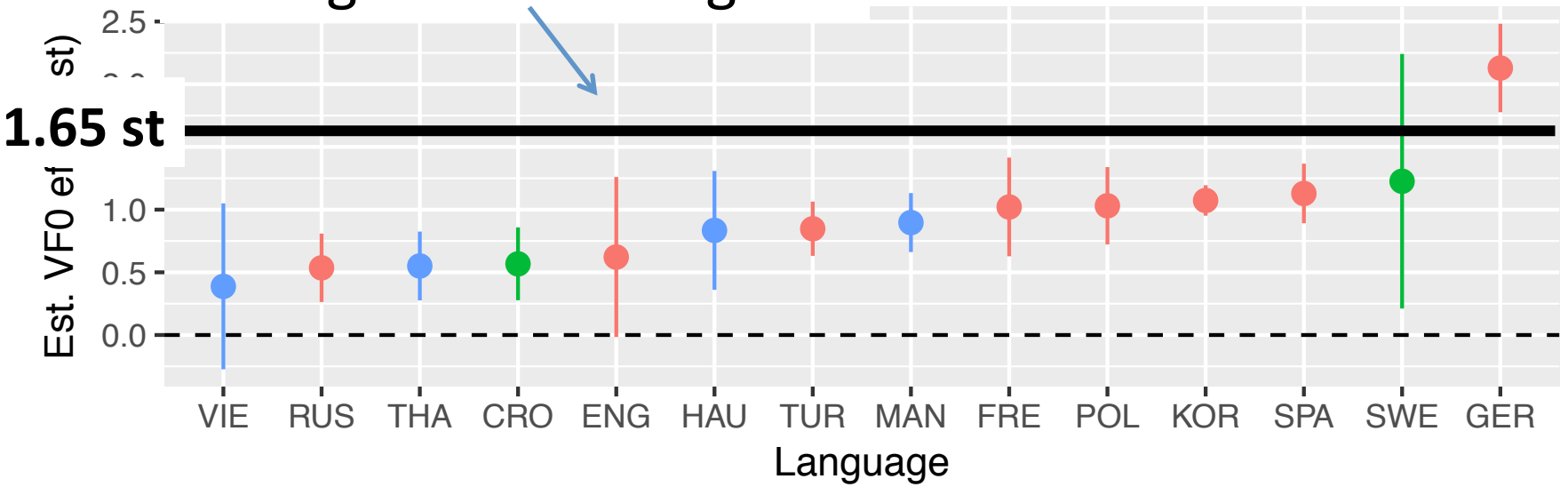



- mostly robust across languages
- variable effect size
  - Non-tonal ⇒ generally larger effect

Average effect across gender, tone, etc.

# VF0: across languages

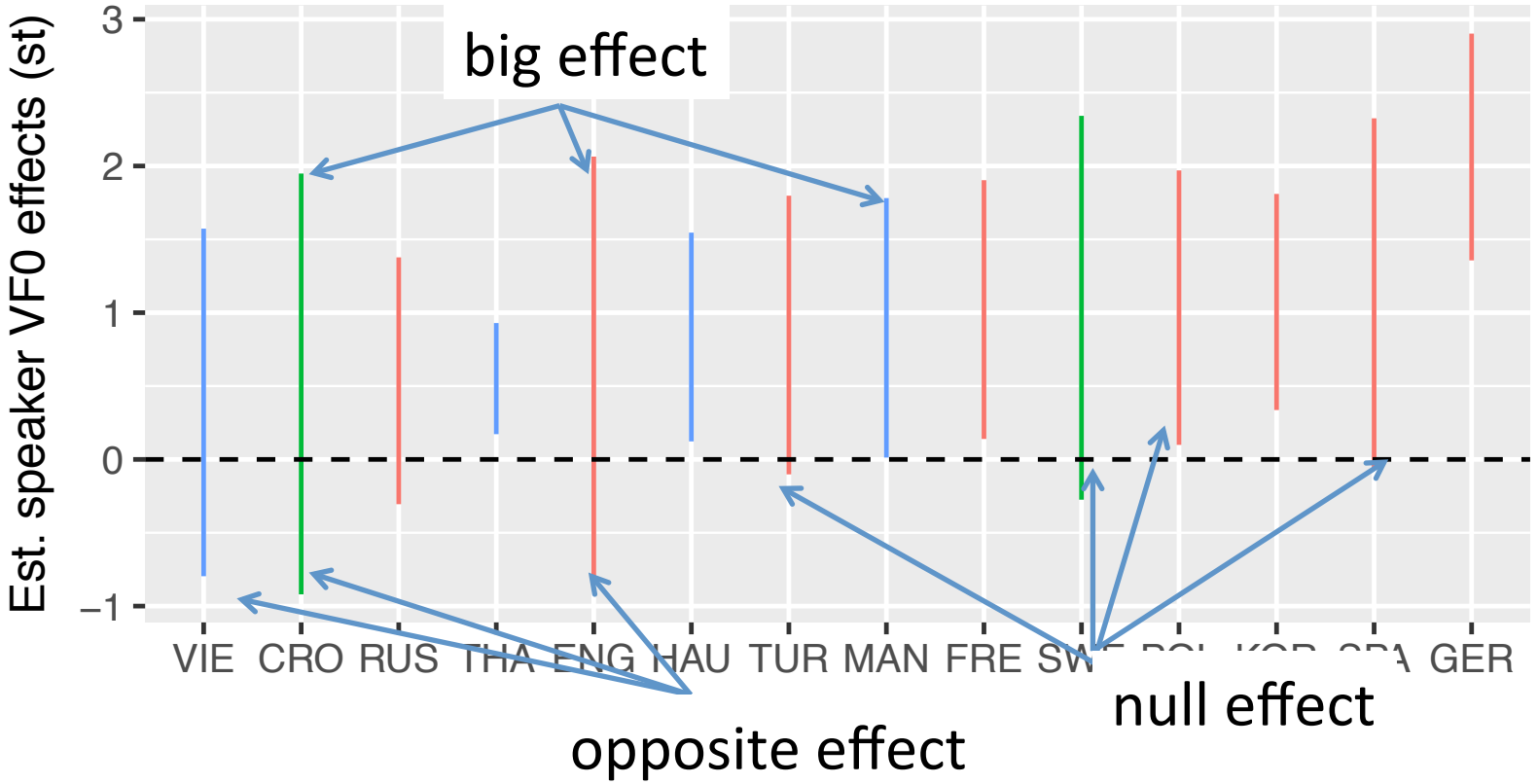
Whalen & Levitt (1995)
average over 31 langs:




- Read vs. lab speech?
- (or artifact of methodology?)

# VF0: across speakers

- Predicted effects for 95% of individuals




- Common: large interspeaker variability

# Discussion

- IF0 effects can be detected using
  - Corpus data
  - Fully automatic analysis
  - Basic statistical controls
  - *n* =~2-4k
- Not obvious!

- Demonstrates feasibility of large-scale studies of phonetic precursors (involving F0)

# Discussion

- Robust group-level IF0 effects across languages
  - same direction
  - "universality" (Whalen & Levitt, 1995)

- Very different effect sizes
  - One reason: tonal/pitch accent language
    ⇒ smaller IF0 more likely
    (hypothesized for VF0: Connell 2002)

- Fits with automatic + controlled mechanism
  (c.f. Hoole & Honda, 2011)

# Discussion

- Large interspeaker variability in IF0 magnitude common, within language
  - ⇒ there are some speakers with null/large effects
  - Still, most speakers show effect in same direction
- Overall: IF0 effects
  - robust across languages
  - variable across speakers
- Both important for sound change
- Related to actuation: why sound changes from IF0 possible, but rare? (Kingston, 2007)

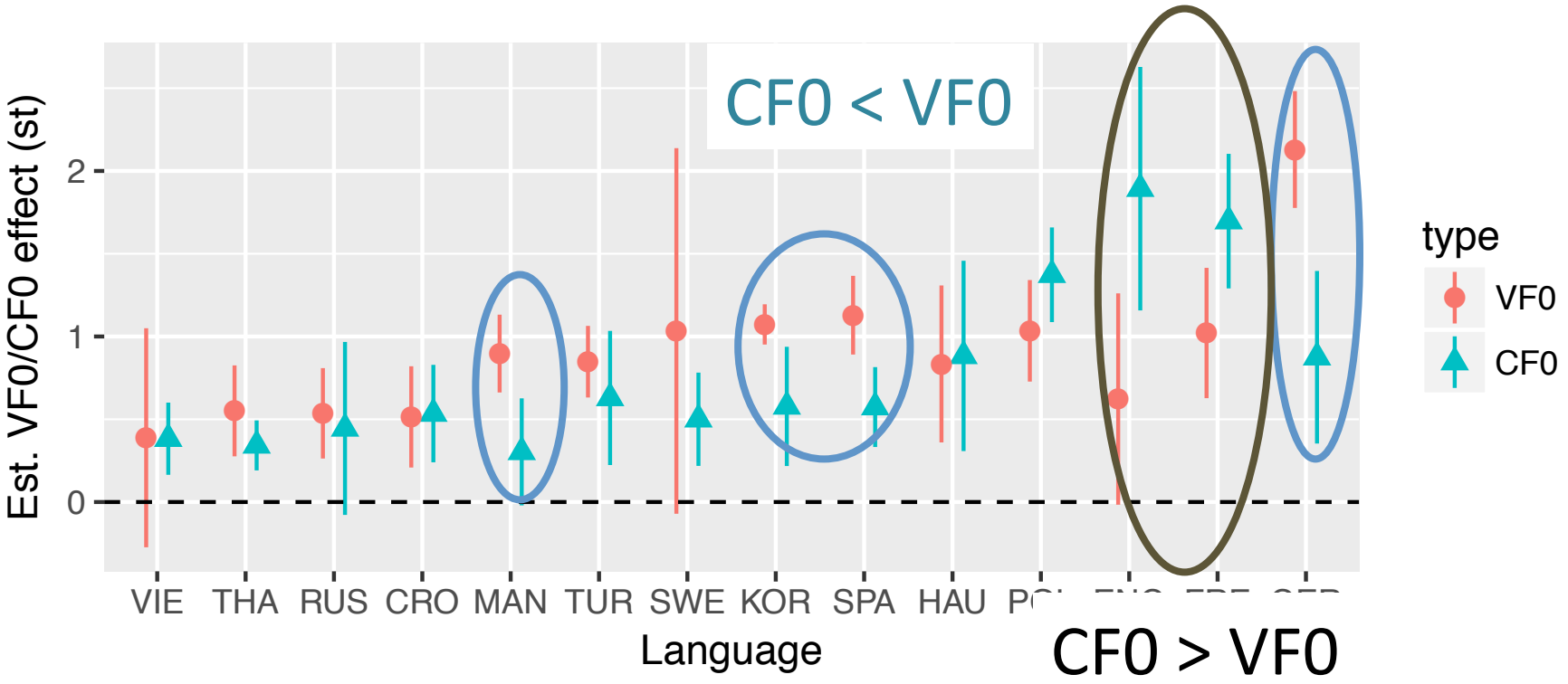
# Extra: VF0 vs. CF0

- Asymmetry between IF0 effects w.r.t. sound change:
  - CF0: many attested changes
  - VF0: ~none

- Why?
  - VF0/CF0 magnitude roughly similar? (Hombert et al., 1979)
  - Perhaps perception is different (Hombert, 1979)
  - VF0 effects show more variability? (Kingston, 2011)

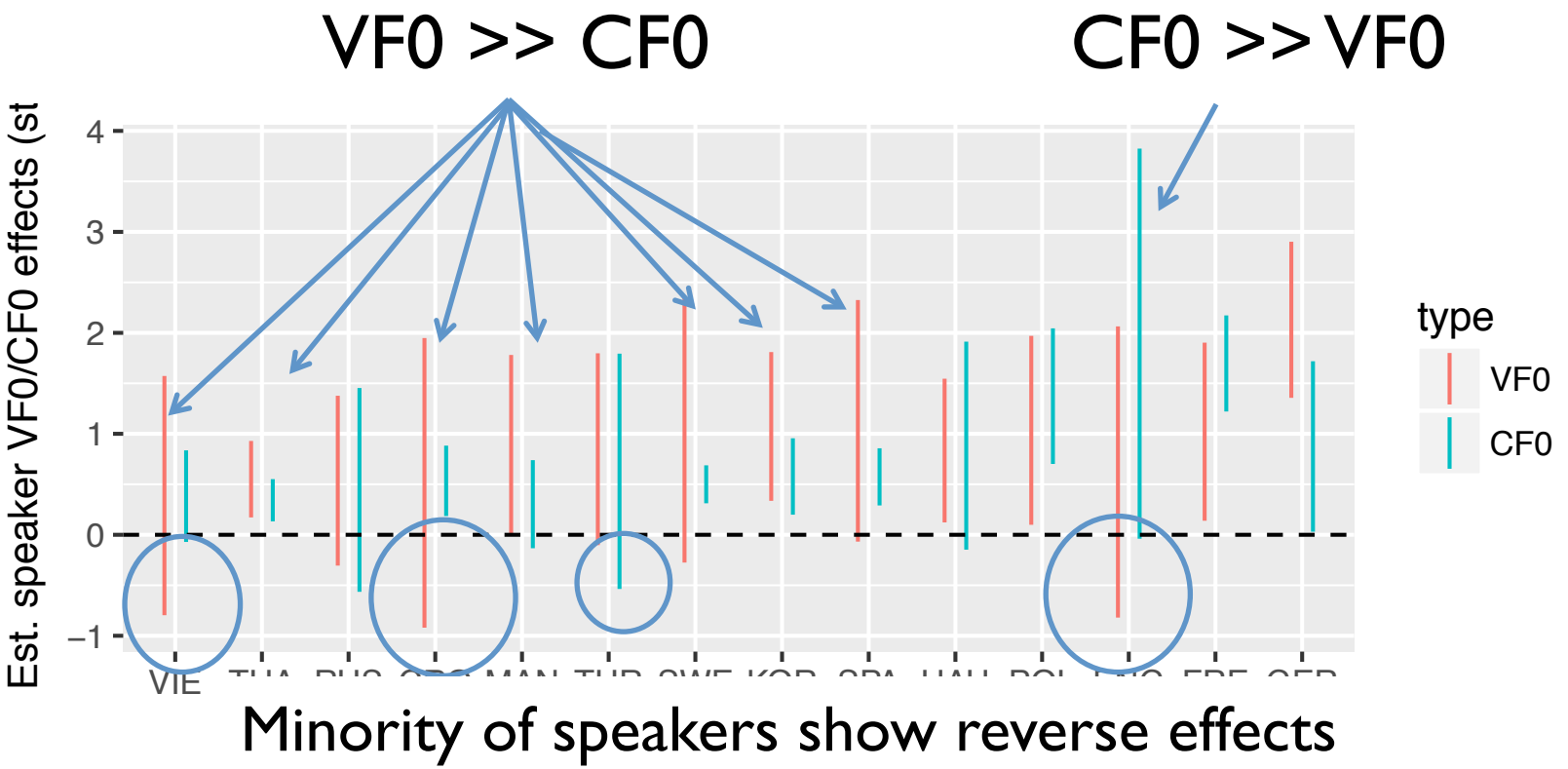
- Q4: Relative magnitude, variability of CF0 & VF0 across languages?

# VF0 vs. CF0: effect size




- No clear pattern
- CF0, VF0 of ~comparable size

# VF0 vs. CF0: speaker variability




- Overall: no obvious pattern
- But: some evidence that VF0 "more variable" than CF0

# Thanks


- RA: Elias Stengel-Eskin
- Comments: James Kirby, Simon King, Montreal Language Modeling Lab members

- Funding:

SSHRC CRSH

Fonds de recherche Société et culture Québec

INNOVATION.CA
CANADA FOUNDATION FOR INNOVATION | FONDATION CANADIENNE POUR L'INNOVATION

# Questions