Relative cue weighting of the register contrast in Southern Yi

Jianjing Kuang

Aletheia Cui

Yan Lu

About Southern Yi

- Tibeto-Burman
- Spoken in Yunnan, China
- Seven vowels
  - i, ɛ, a, ɿ, ə, o, u
- Three tones
  - Low (21), mid (33), high (55)
- Phonation-based registers (tense vs. lax)
  - Only co-occurs with the low (21) and mid (33) tones
  - Tense vowels are indicated by an underline (e.g., b<u>e</u>21 vs be21)

The Register Contrast

	T21	L21	T33	L33	55
be					
bu					

Phonation as the primary cue for the register contrast

The tense vs. lax contrast is distinct in phonation

- Tense phonation has a greater degree of glottal constriction:
  - Acoustically, tense phonation has less prominent H1 – smaller values for H1*, H1*-H2*, H1*-A1*, H1*-A2*, H1*-A3*
  - Articulatory, tense phonation has greater Contact Quotient (EGG)

Ladefoged and Maddieson, 1985; Shi and Zhou, 2005; Kuang, 2011; Kuang & Keating, 2014

Coarticulated Cues for the Register Contrast

- Constriction in the larynx and pharynx leads to retracted tongue root
  - Laryngoscope studies of Northern Yi show that the tense vowels show a retraction of the tongue root (Esling et al., 2001; Edmondson and Esling, 2006)
  - F1 for the tense vowels tend to be higher (Kuang, 2011) in one dialect
- Stiffness in the vocal folds leads to a higher f0 (Kuang, 2011)
- There are dialect differences.

Sound Change?

- Kuang (2011) noticed that some younger speakers do not produce phonation contrasts for the low vowels, e.g. /be/
- Maybe a sound change is underway?
- Other Yi dialects
  - One has lost phonation distinctions entirely for low vowels

	Tense	Lax
Dialect with no phonation contrast		
Dialect in this study		

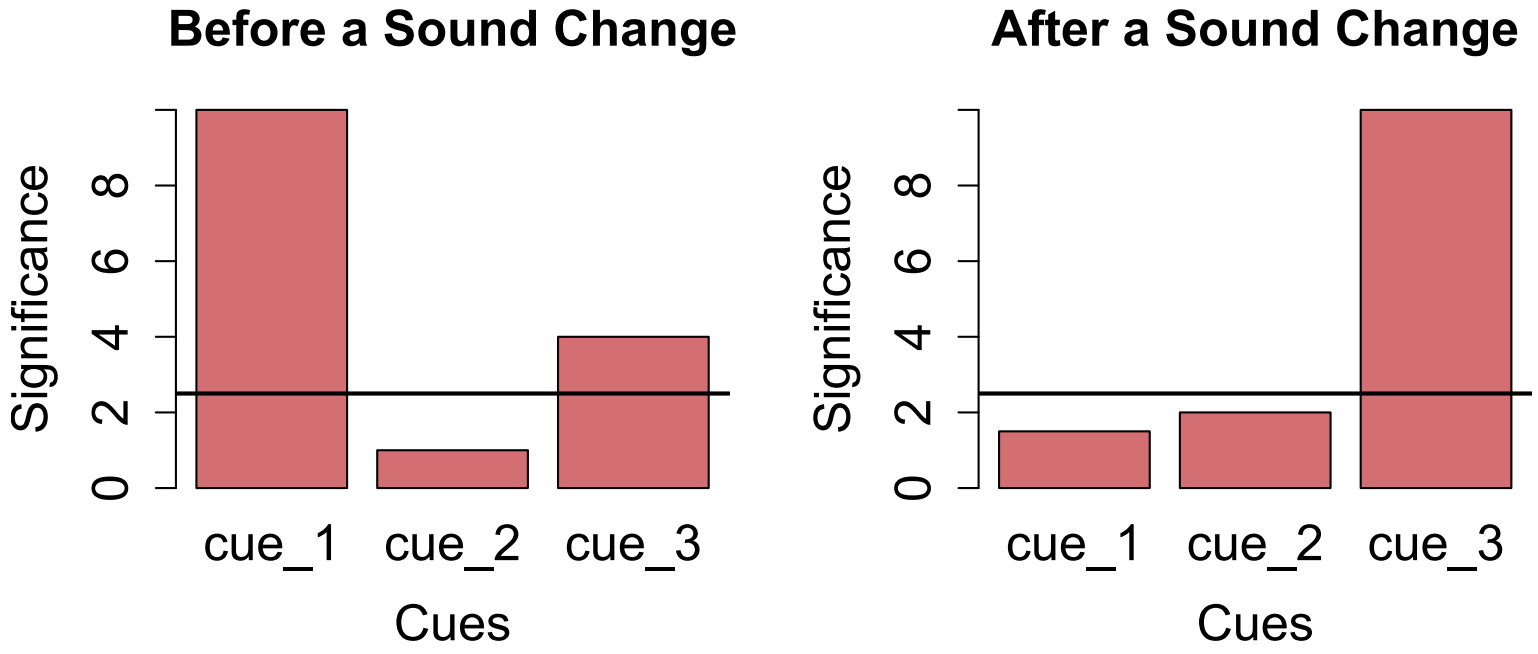
Historical Changes through Cue Shifting

- Tonogenesis
  - Consonant voicing contrast > tonal contrast
  - Vowel pitch is higher after a voiceless consonant (e.g., Ohala, 1973; Hombert et al., 1979; Maddieson, 1984)
    - Increased vocal cord tension to maintain voicelessness, resulting in F0 perturbations on the following vowel (Löfqvist et al, 1989)

- Vowel nasalization:
  - Vowels preceding nasal stops undergo coarticulatory nasalization
  - Results in historical change: VN > Ṽ

Sound Change from the Shifting of the Primary Cue

- Height of bar = importance of cue
- Horizontal line = threshold for significance



Questions

Phonetic variation provides the opportunity for sound change.

- How does this happen?
- The role of production and perception?

Three possibilities:

- Cue shifting starts in production, and then listeners become attuned to the changes in the importance of cues
- Listeners shift their attention to a new cue, and then in turn rely on this cue to mark a phonological contrast in production
- Production and perception at the same time

This Study

- Is Southern Yi undergoing sound change?

  --Low vowels are losing phonation contrasts?

- If so, how is the sound change reflected in production and perception?

→ Relative importance of the coarticulated cues in production and perception with the same group of people

Production and Perception in Sound Change

- Other studies with both production and perception of the same speakers
  - Vowel-to-vowel coarticulation (Grosvald and Corina, 2012)
  - Kammu vowel tonogenesis (Svantesson and House, 2006)
  - Standard South British /u/-fronting (Harrington et al., 2008)

Participants

- 41 native speakers
  - 17 female speakers (age range [26:70], mean 47)
  - 24 male speakers (age range [30:71], mean 45).
- Divided into 3 age groups
  - younger than 40 (10 participants),
  - 40-50 (17 participants)
  - older than 50 (14 participants)

PERCEPTION 1: NATURAL STIMULI

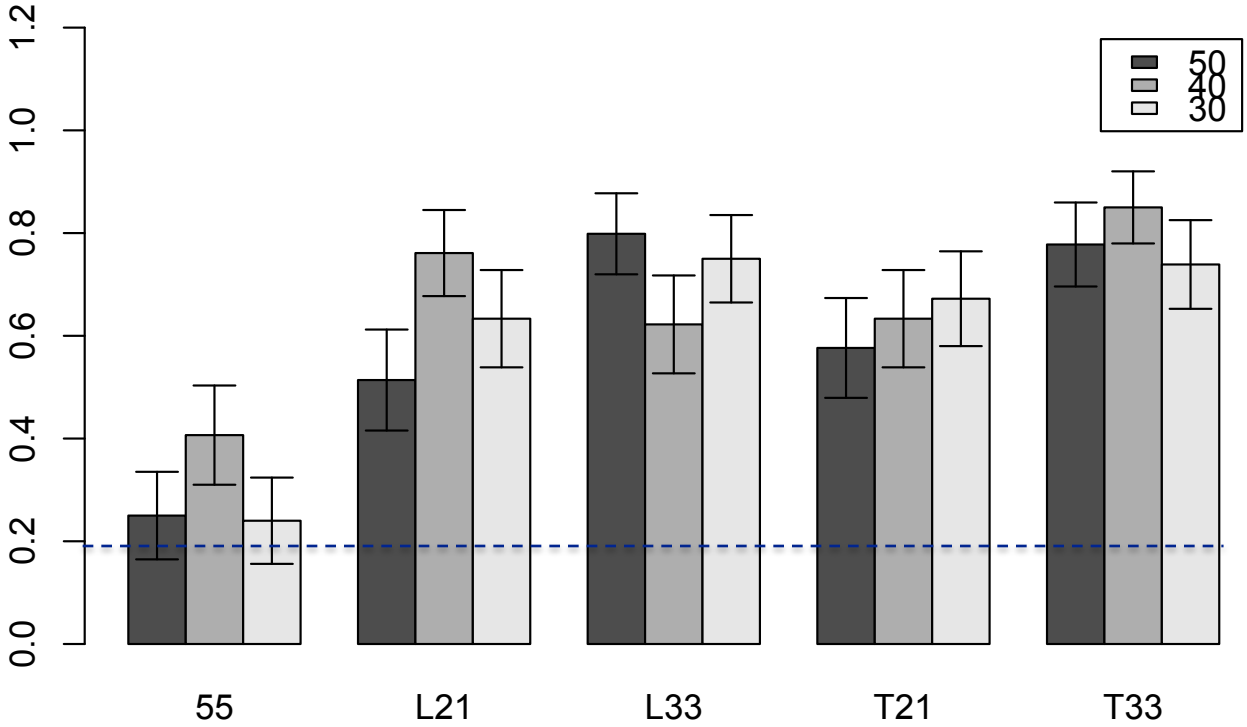
Goals

- Create a baseline for comparison
- Test whether the register contrast is maintained by all age groups

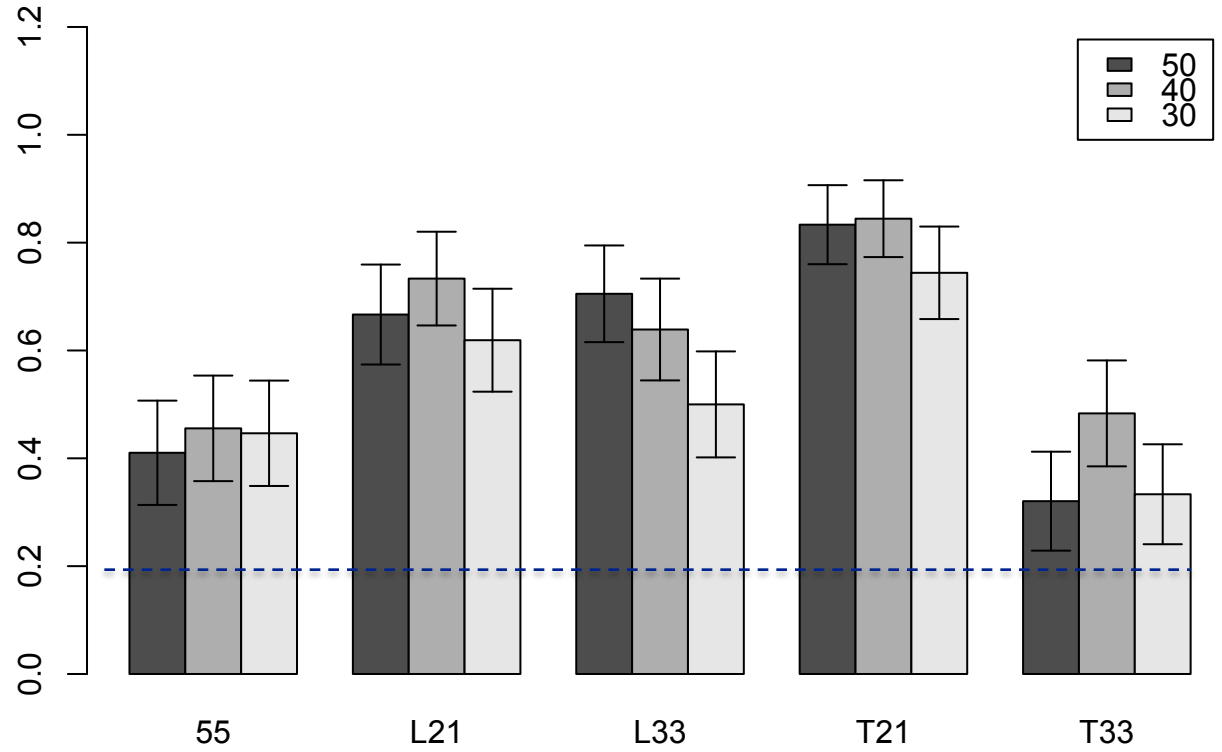
Methods

- Stimuli
  - naturally produced minimal sets *be* (be21, b̲e21, be33, b̲e33, and be55), and *bu* (bu21, b̲u21, bu33, b̲u33 b̲u55) from a previous production experiment from 3 males and 3 females (Kuang 2011)
- Task: forced choice identification task
  - Each stimulus was presented in isolation
  - five possible choices for each stimulus, each corresponding to a word in the minimal set
  - The listeners were asked to select the choice that best corresponded to the stimulus

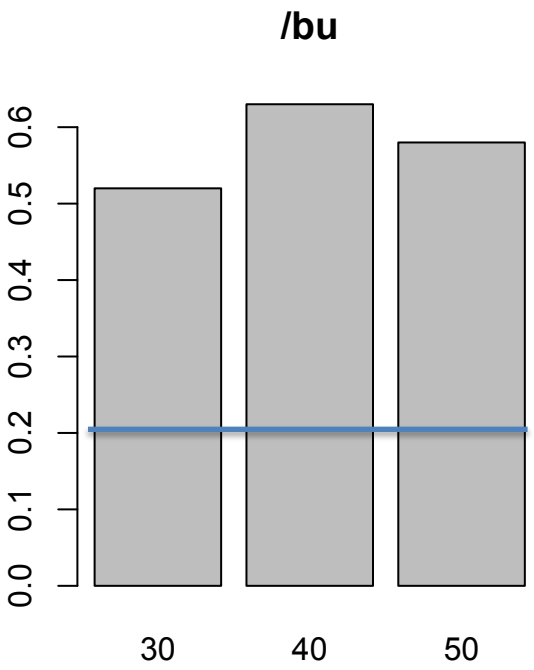
/be/accuracy by age

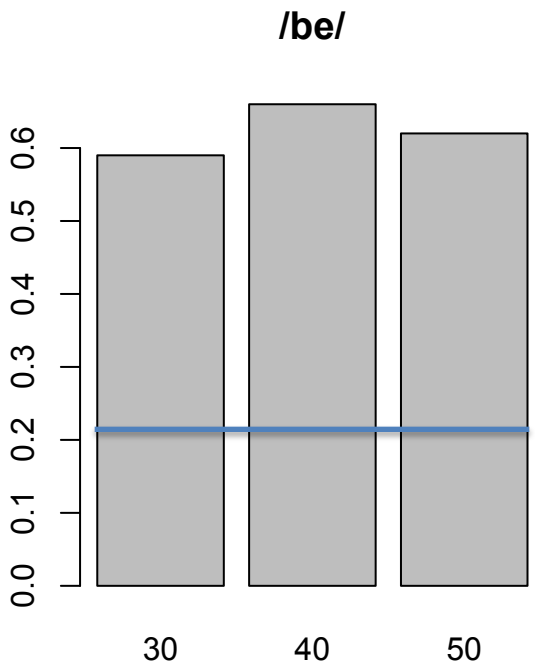


/bu/accuracy by age



Overall accuracy





- Chi-square tests show that there is no age difference in accuracy rate.
- Linguistic contrast is maintained by all age groups

- Whether all age groups use the same cues to make the contrast?

- Whether all age groups use the same cues to make the contrast?

cue weighting in both production and perception

PRODUCTION

Methods

- Speaker were asked to produce minimal sets of /be/ (be21, b̲e21, be33, b̲e33, be55) and /bu/ (bu21, b̲u21, bu33, b̲u33, bu55)
  - Frame: [ŋo33 __ e55 lʏ33 ɣw33] “I say the word ___.”
  - Each utterance was repeated 3 times
- The vowel portion of each target word was extracted
  - acoustic measurements were taken automatically using VoiceSauce
  - Four measures: F1, F2, F0 and H1*-H2*

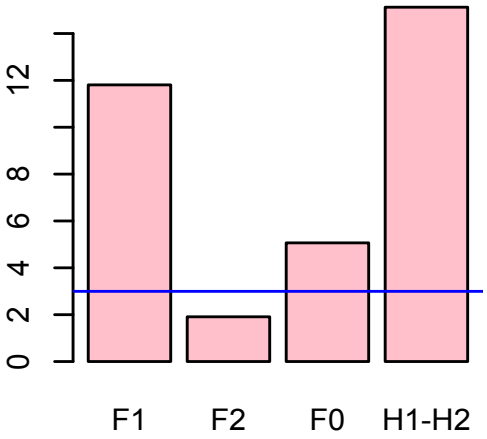
Methods: Data Analysis

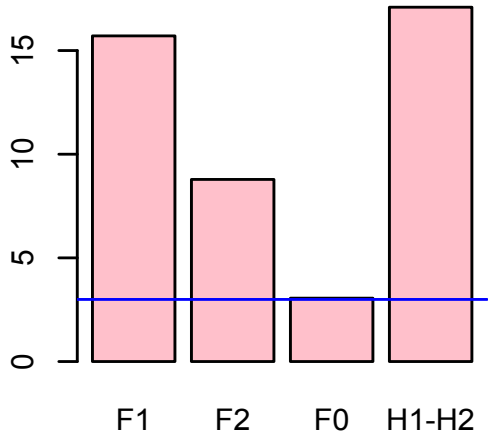
- Stats: logistic regression models

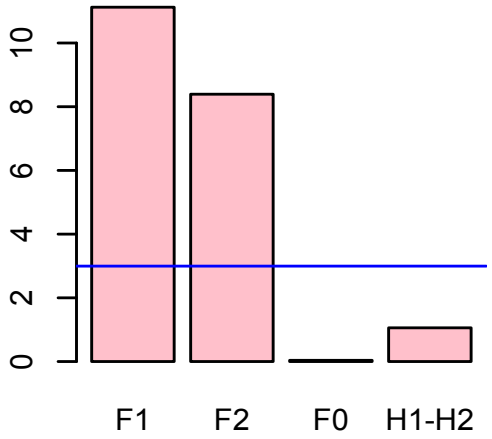
Register ~ F1.z +F2.z + F0.z + H1-H2.z

Plot log(p-value) to indicate relative importance

/be/



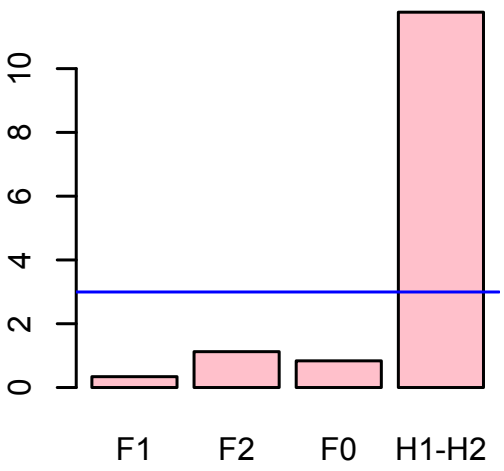


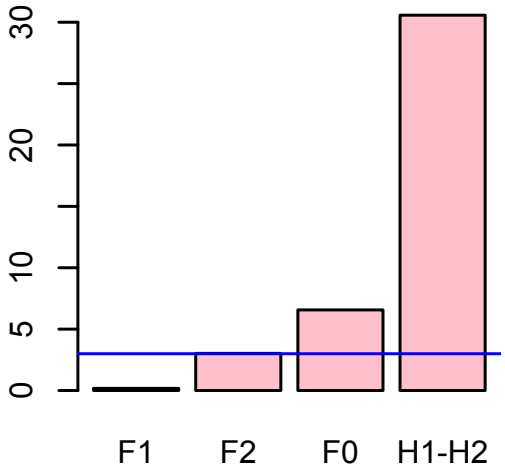


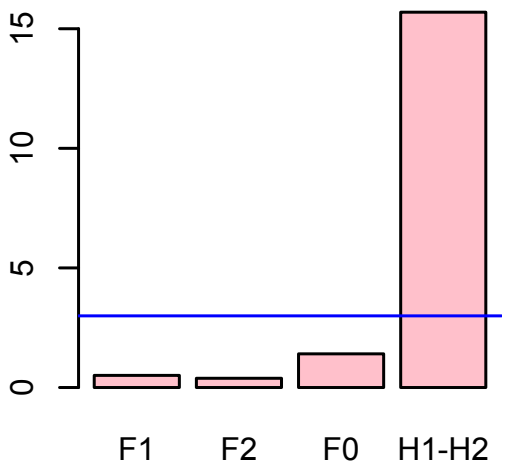
Higher bar = cue is more important

Horizontal line = log(0.05)

/bu/







Higher bar = cue is more important

Horizontal line = log(0.05)

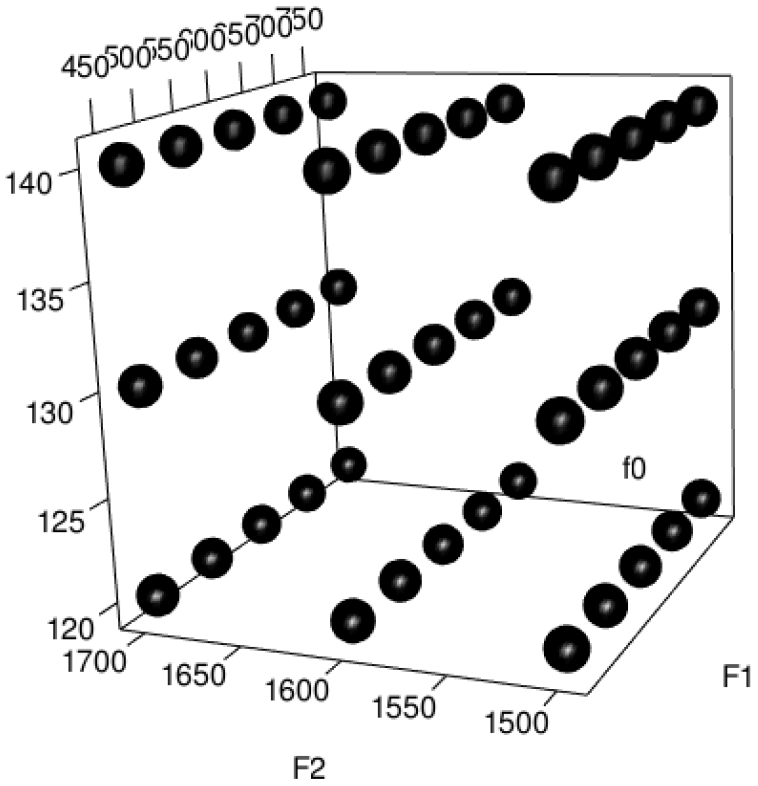
PERCEPTION: RESYNTHESIZED STIMULI

Goals

- Find out age groups differ in the cues they rely on for the register contrast
- Test whether the cues differ for low vowels and high vowels

Methods

- Stimuli
  - resynthesized from naturally produced /be33, b̲e33/ and /bu33, b̲u33/ of two speakers (1F, 1M)
  - F1, F2, and F0 were modified in incremental steps (5 F1 × 3 F2 × 3 F0)
  - Phonation of the original tokens was maintained



Resynthesized Stimuli

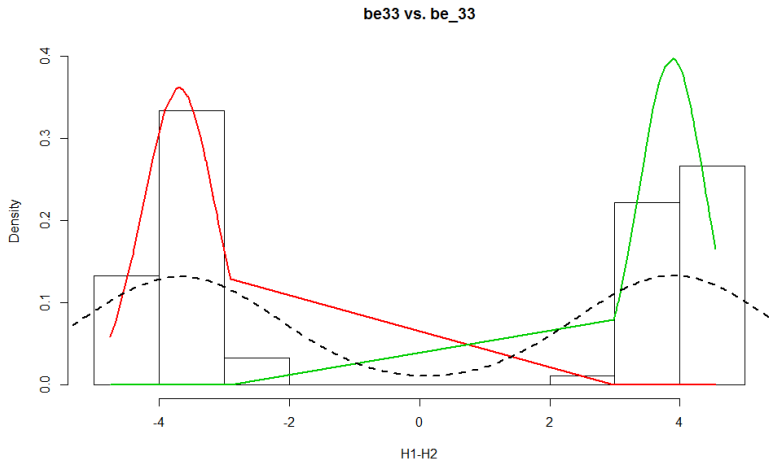
F1	250	325	400	475	550
Tense					
Lax					

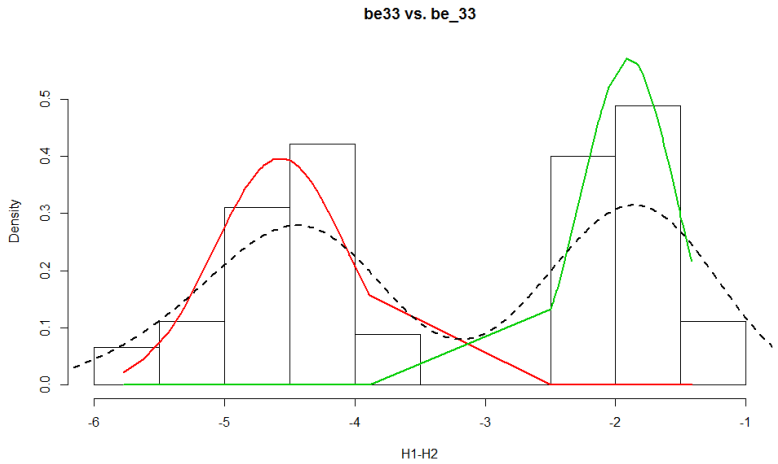
Phonation contrast is successfully maintained

Paired-tests between tense and lax for all four speakers

	Male be	Male bu	Female be	Female bu
H1*	t(44) = 11.2408, p = 1.605e-14	t(44) = 18.1434, p < 2.2e-16	t(44) = 18.5234, p < 2.2e-16	t(44) = 29.8151, p < 2.2e-16
H1*-H2*	t(44) = 98.186, p < 2.2e-16	t(44) = 7.8634, p = 6.341e-10	t(44) = 16.7108, p < 2.2e-16	t(44) = 23.2317, p < 2.2e-16
H1*-A1*	t(44) = 15.565, p < 2.2e-16	t(44) = 5.949, p = 4.01e-07	t(44) = 4.4967, p = 4.98e-05	t(44) = 2.0085, p = 0.05075
H1*-A2*	t(44) = -11.8594, p = 2.695e-15	t(44) = 6.4398, p = 7.606e-08	t(44) = 1.4653, p = 0.1499	t(44) = 10.6477, p = 9.308e-14
H1*-A3*	t(44) = -7.6843, p = 1.15e-09	t(44) = 4.6161, p = 3.388e-05	t(44) = -12.7227, p = 2.424e-16	t(44) = 56.5421, p < 2.2e-16
H2*-H4*	t(44) = -5.8648, p = 5.33e-07	t(44) = -1.0684, p = 0.2912	t(44) = -2.5425, p = 0.0146	t(44) = -3.578, p = 0.0008574
CPP	t(44) = -11.495, p = 7.663e-15	t(44) = -23.4953, p < 2.2e-16	t(44) = -0.3011, p = 0.7648	t(44) = -33.3154, p < 2.2e-16

Phonation contrast is successfully maintained



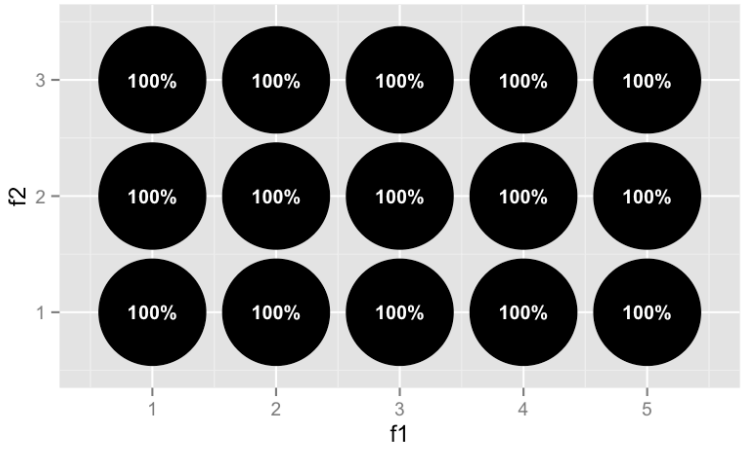


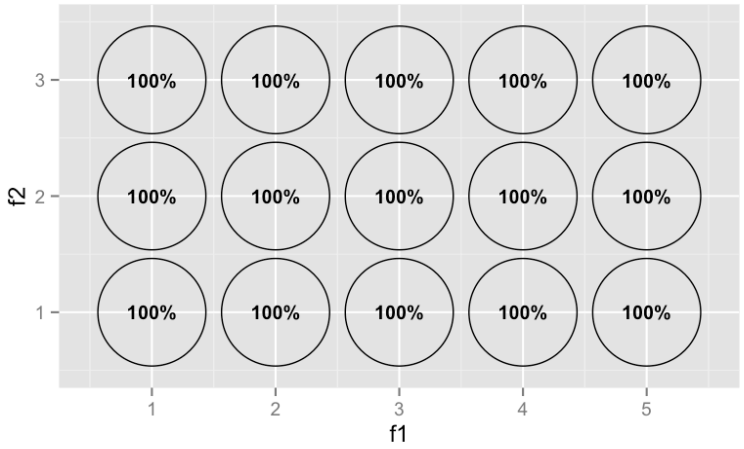
Methods

- Forced choice identification task
  - Each stimulus was presented in isolation
  - the listener was asked to choose between two options whether the token is tense or lax
- Participants: divided into two groups
  - 21 participants heard the resynthesized stimuli of the male speaker
  - 20 heard the resynthesized stimuli of the female speaker.

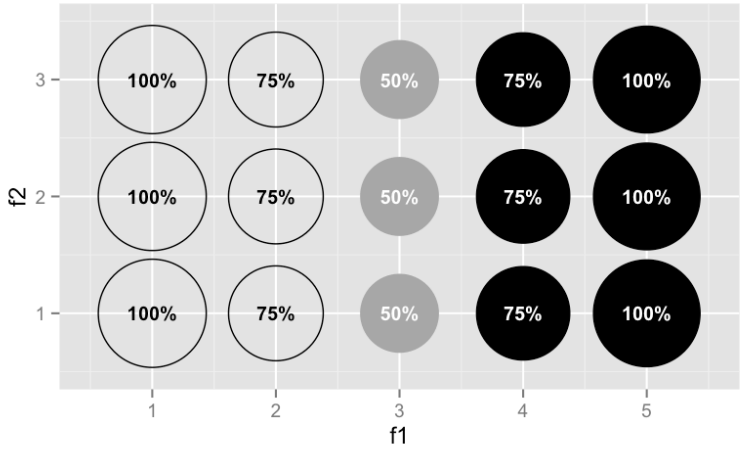
- Predictions:
  - Phonation is the primary cue for /bu/, but maybe not for /be/
  - Older speakers rely more on phonation, younger speakers rely more on vowel quality.

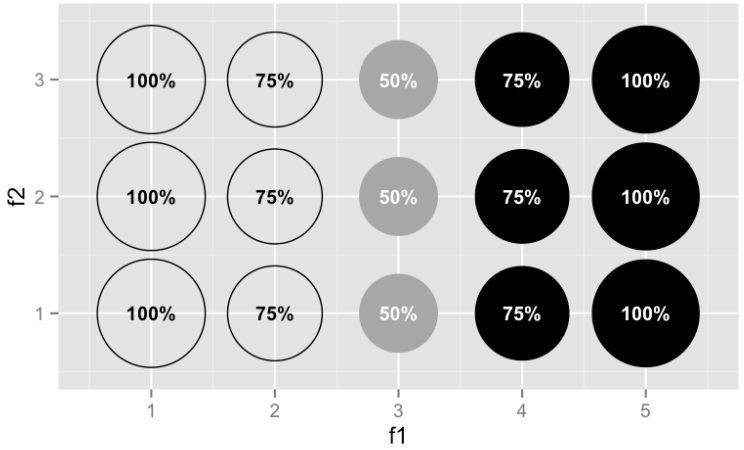
Primarily use phonation



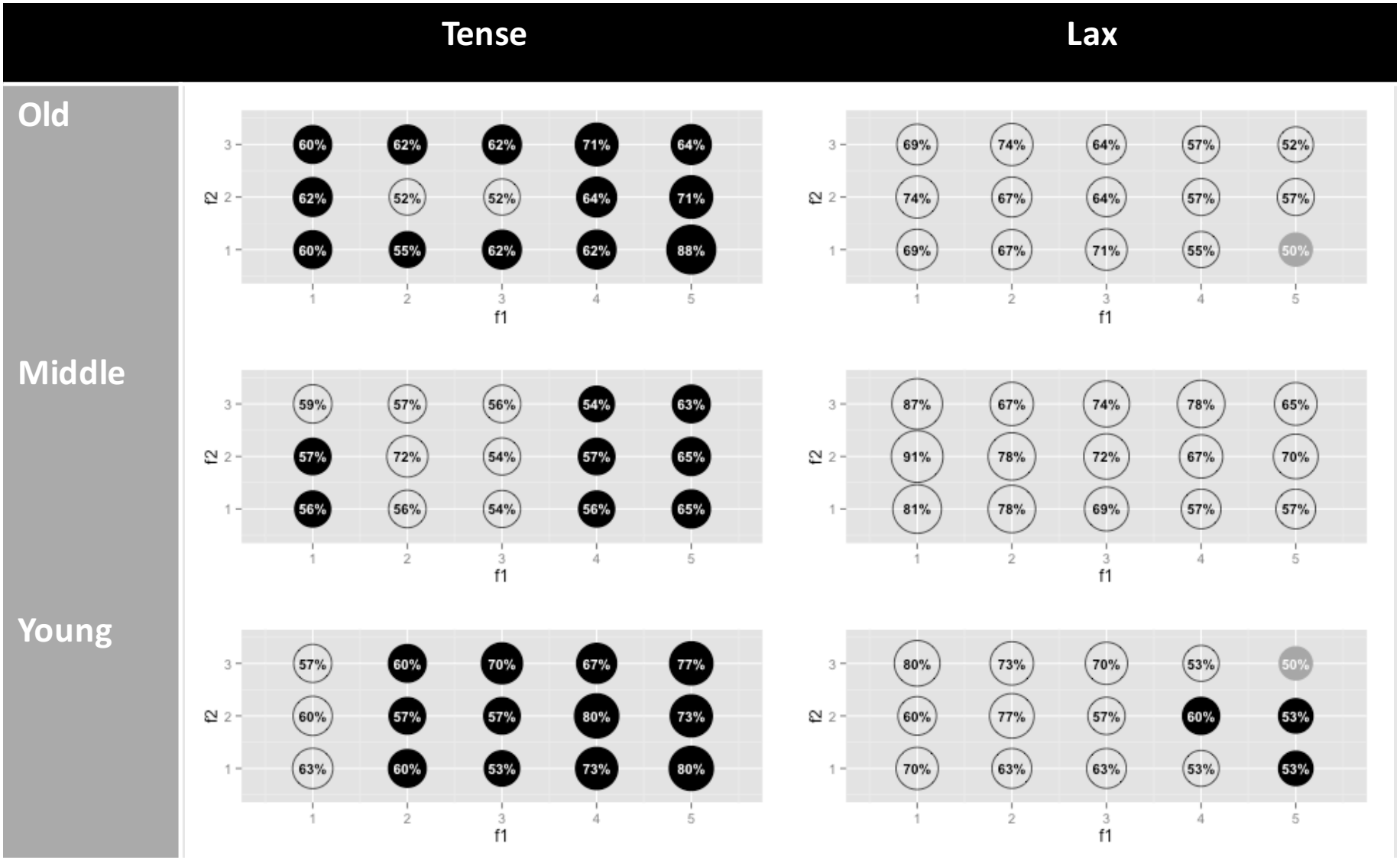


Primarily use F1

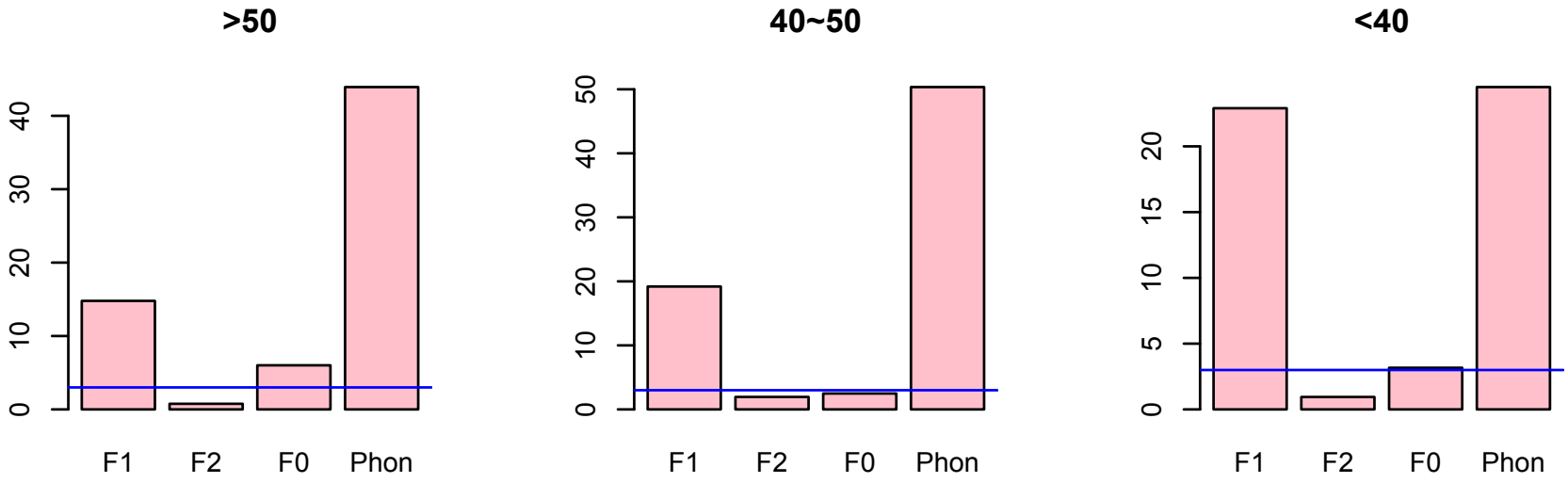




Percentage of response: /bu/



Relative cue weights: /bu/

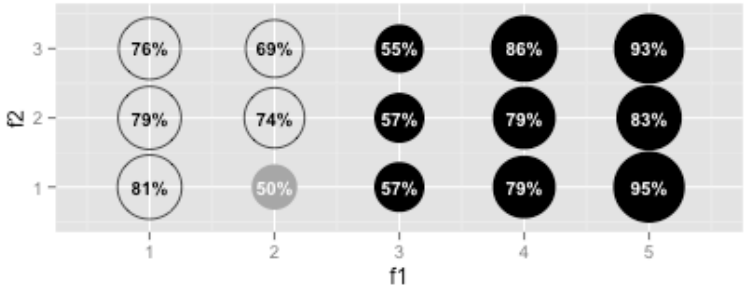
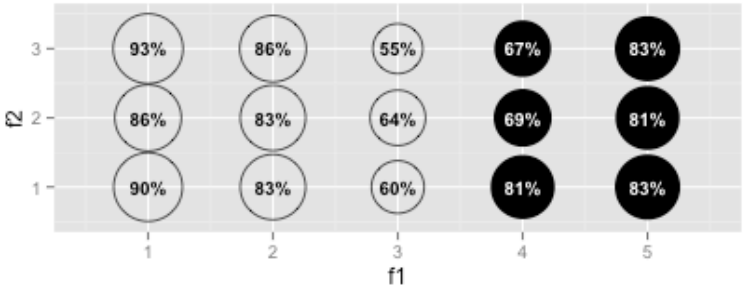
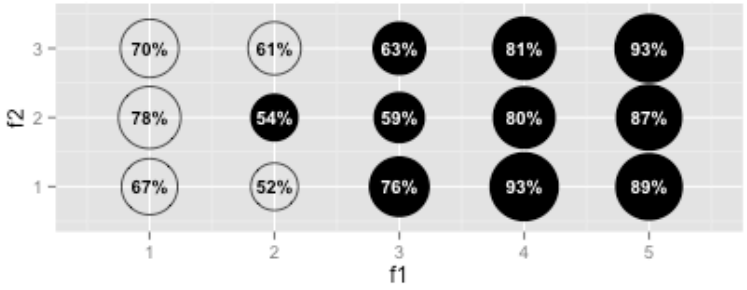
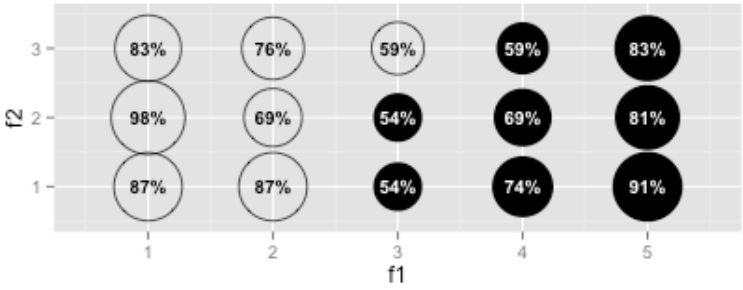
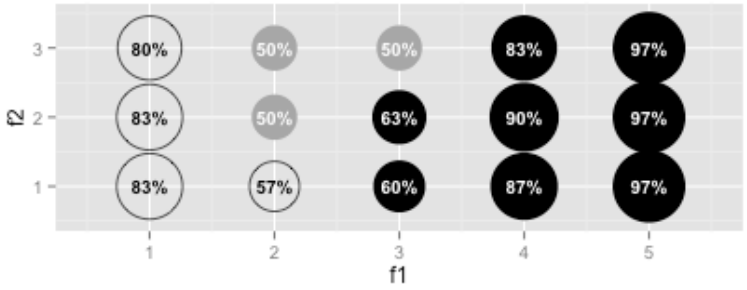
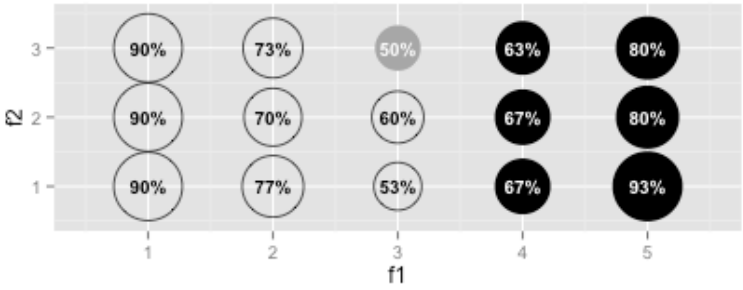


logistic regression model

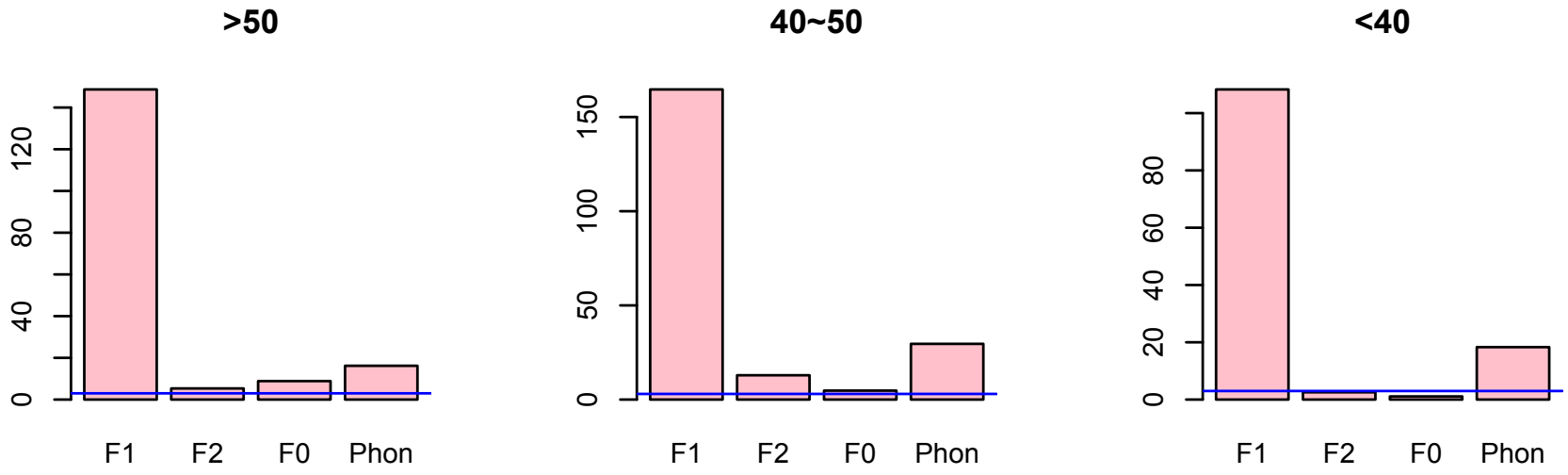
Register ~ F1+F2+F0+ phonation

Horizontal line = log(0.05)

Percentage of response: /be/

	Tense	Lax
Old		
Middle		
Young		

Relative cue weights: /be/



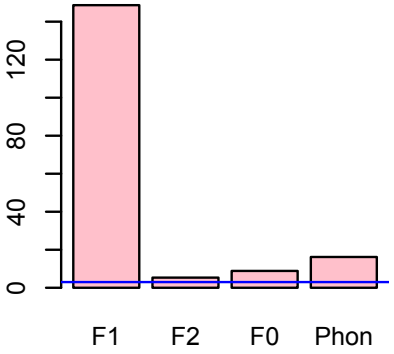
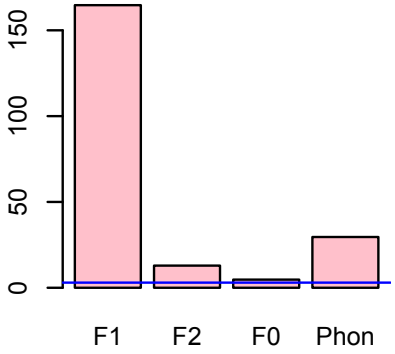
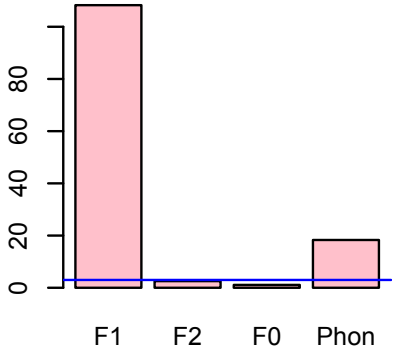
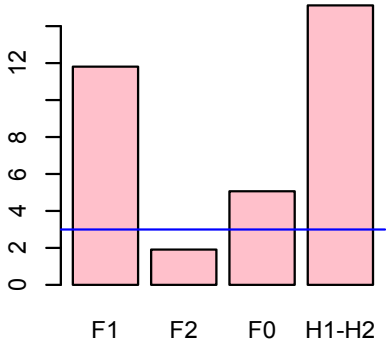
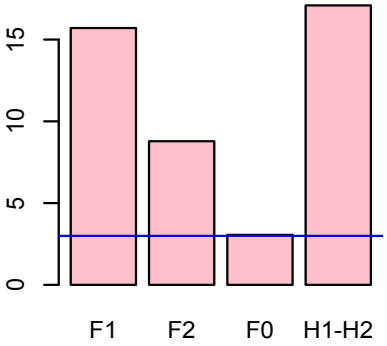
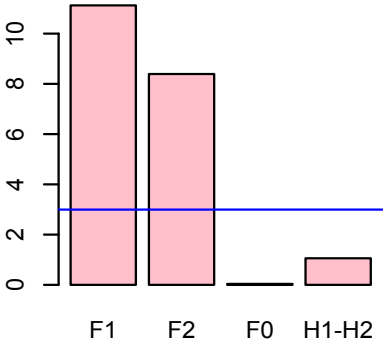
logistic regression model

Register ~ F1+F2+F0+ phonation

Horizontal line = log(0.05)

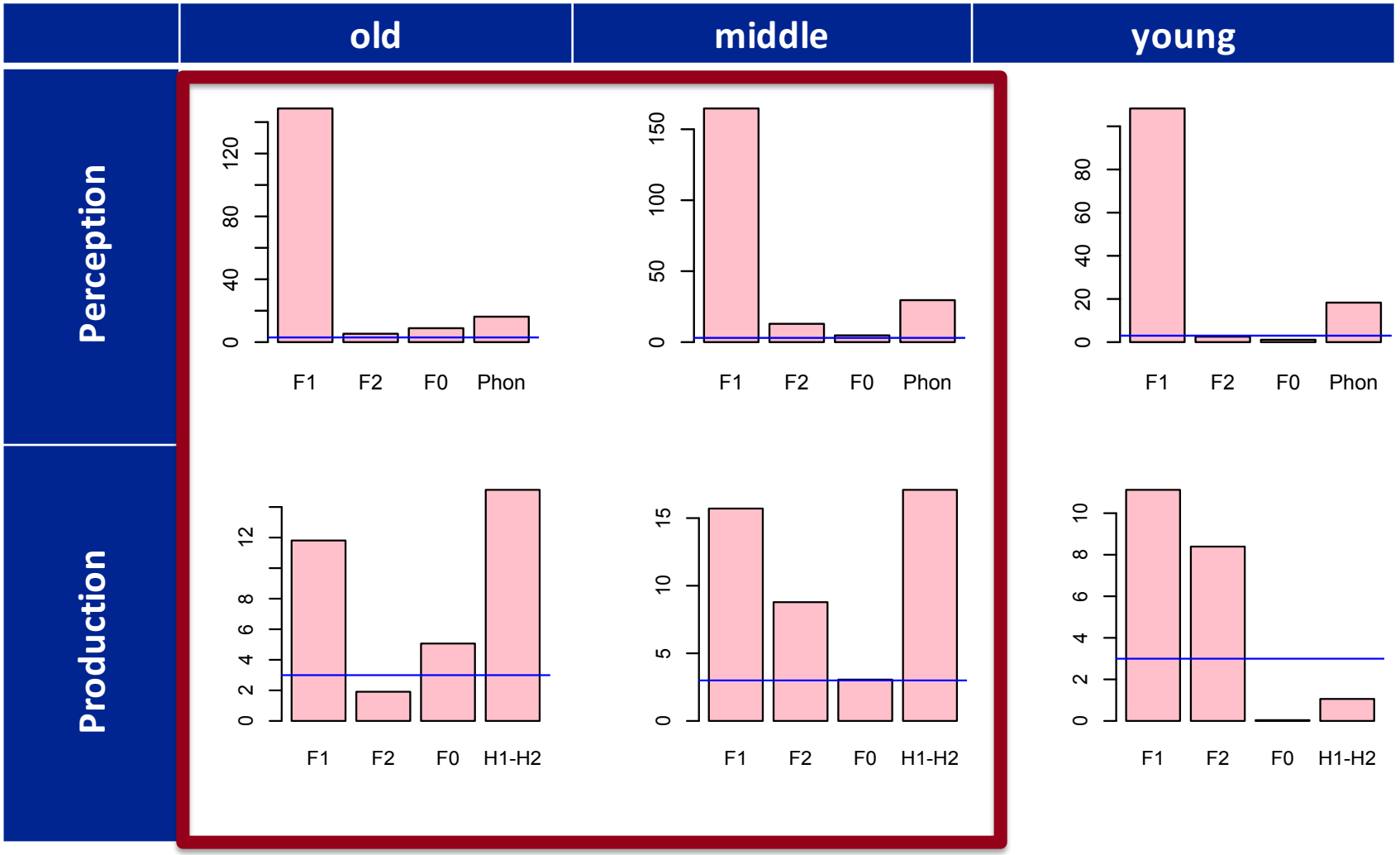
- Mapping between production and perception

/be/

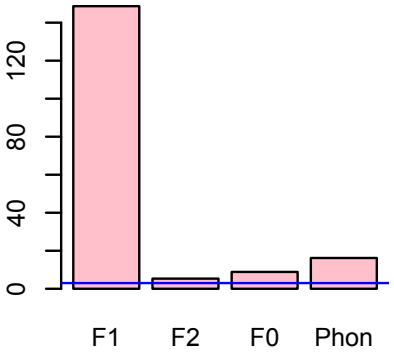
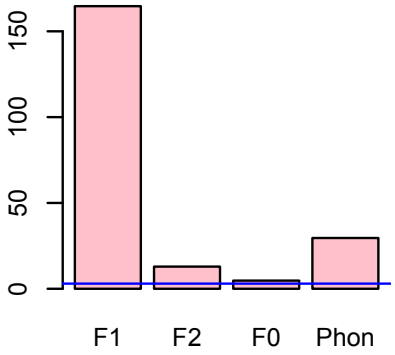
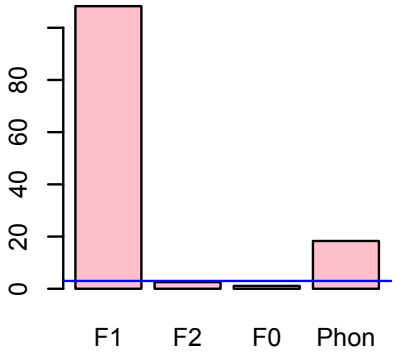
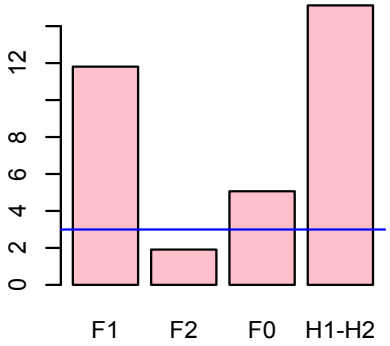
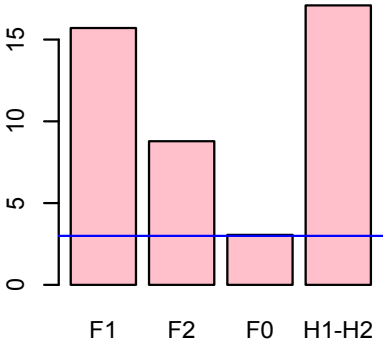
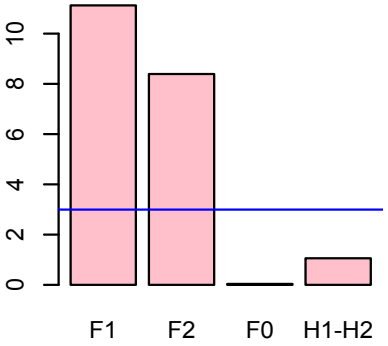
	old	middle	young
Perception			
Production			

/be/

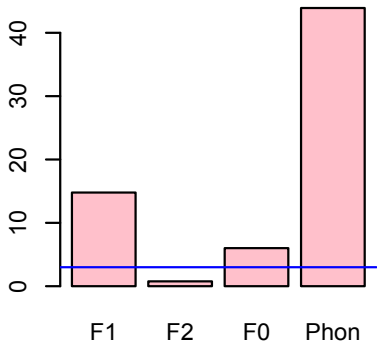
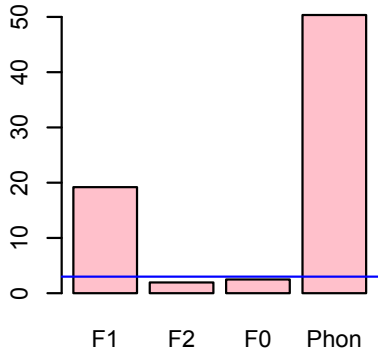
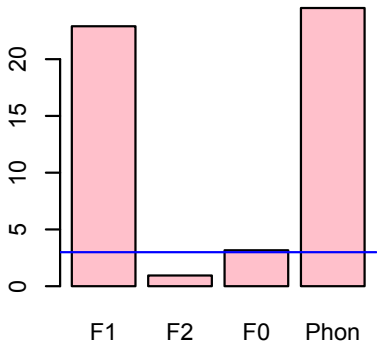
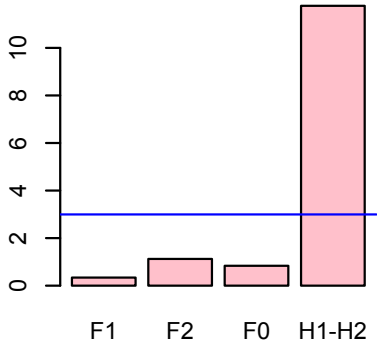
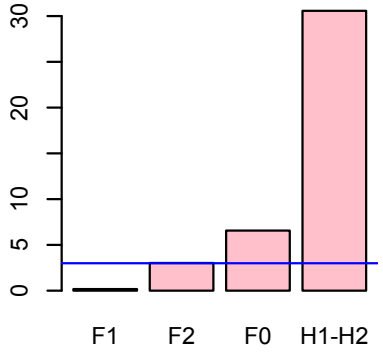
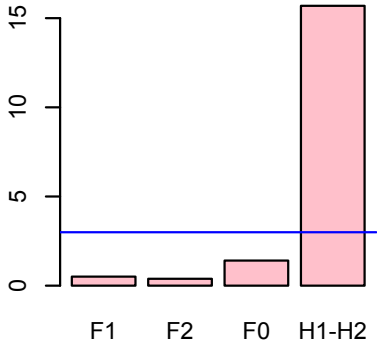
Misaligned!



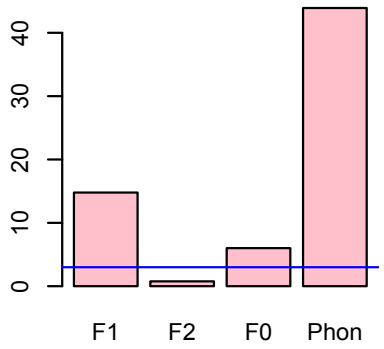
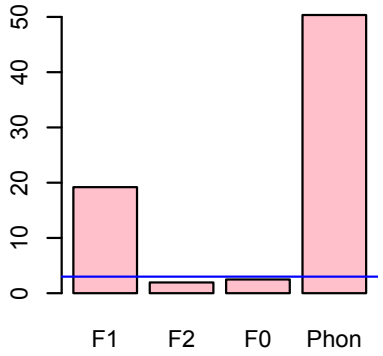
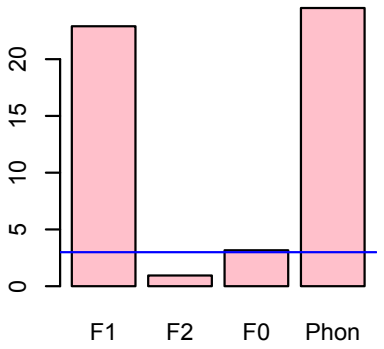
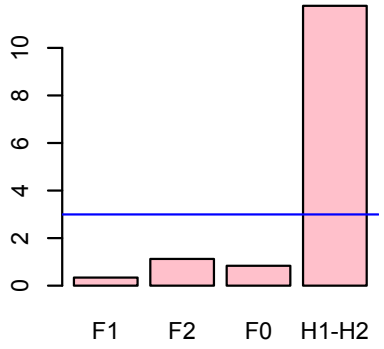
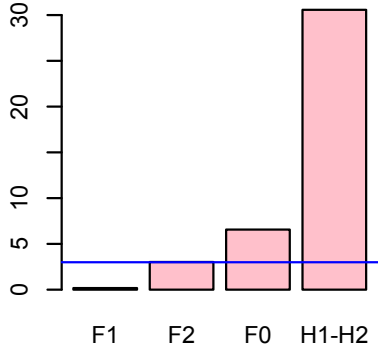
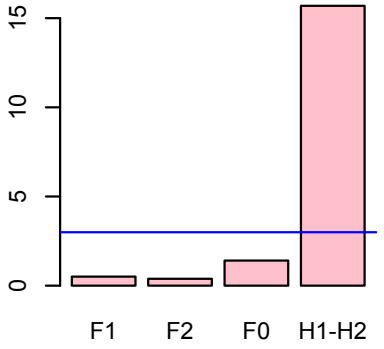
/be/
Cue shifting happens in perception first

	old	middle	young
Perception			
Production			

/bu/

	old	middle	young
Perception			
Production			

/bu/

	old	middle	young
Perception			
Production			

Discussion

- Is sound change underway?
  - Yes. While the register contrast is maintained, formants are overtaking phonation as the primary cues
- How did this happen?
  - Low vowels before high vowels
  - innovation first enters the perception domain, and then production
  - listeners lead the change

Listeners lead the change

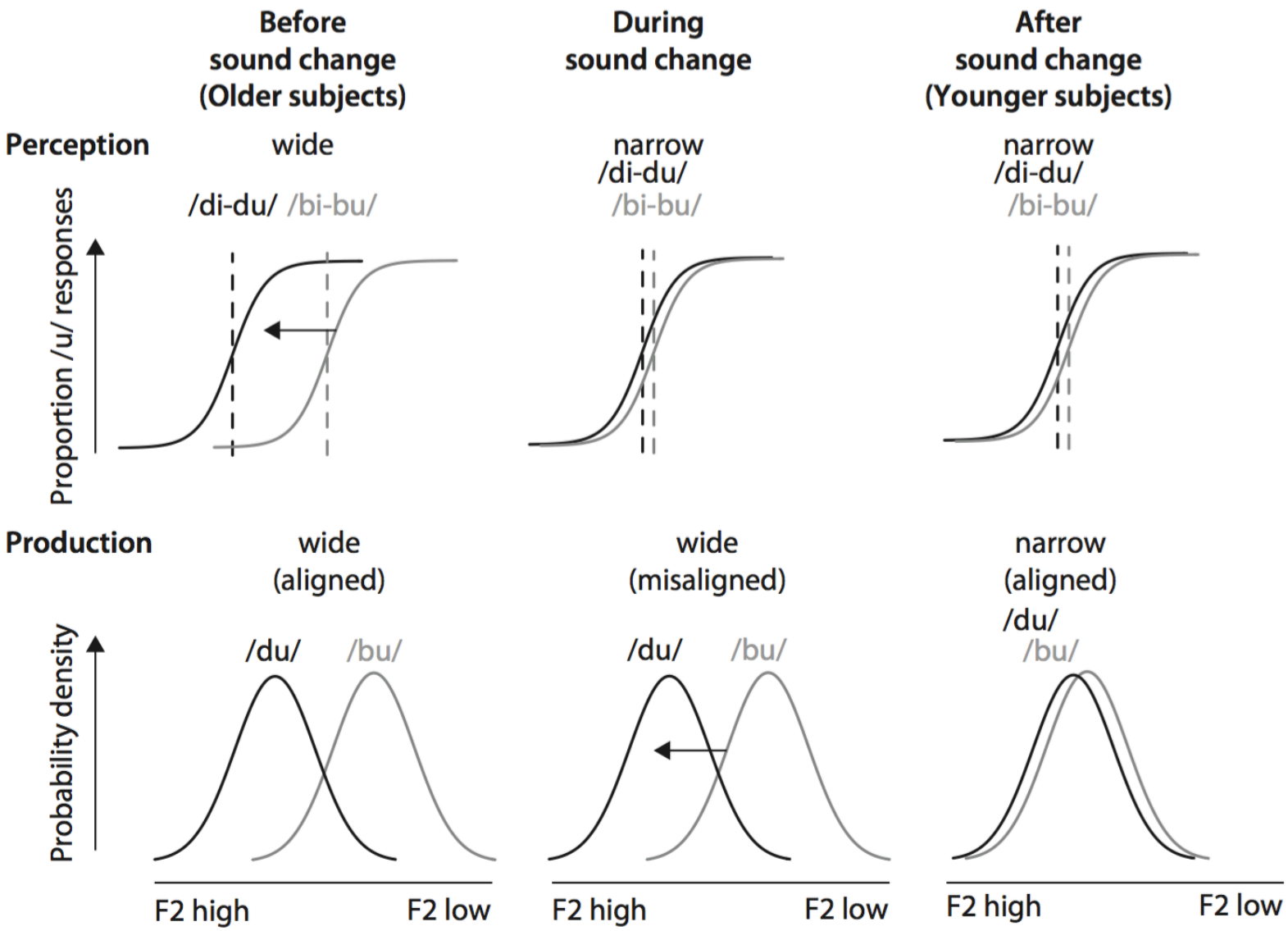
- Ohala (1981, 1993)
  - Variation provides opportunities for change
  - Sound change happens when the listener misinterprets a coarticulated cue as inherent to a segment
- Many-to-many mapping between variable input and flexible perception (Beddor, 2009; 2012)
  - Some listeners may place more weight on coarticulated cues, leading to sound change

Listeners lead the change

- Listeners tend compensate for contextually predictable coarticulation in perception (e.g, Fowler, 2005)
  - Not always accurate (Fowler & Brown, 2000; Beddor et al., 2001; Beddor et al., 2007)
- Age differences in compensation
  - /u/-fronting Standard Southern British (Harrington et al., 2008)
    - Younger speakers compensated less perceptually and produced more fronted /u/

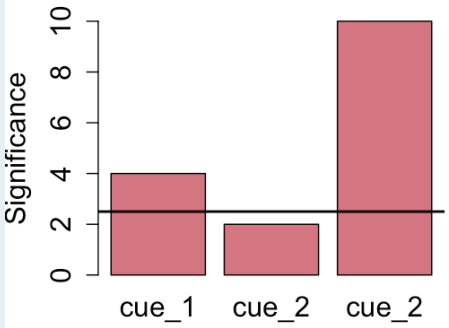
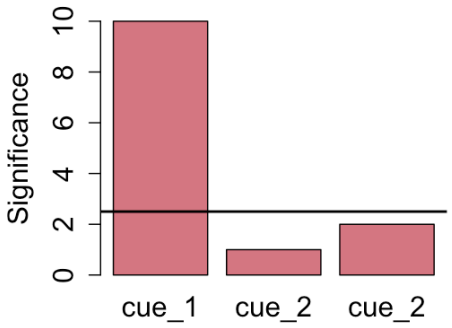
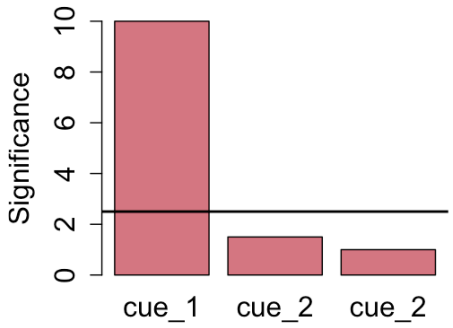
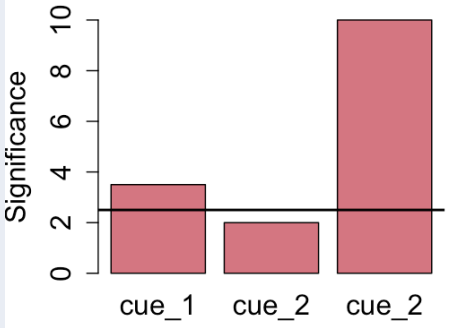
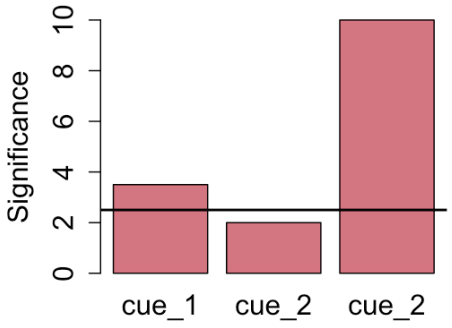
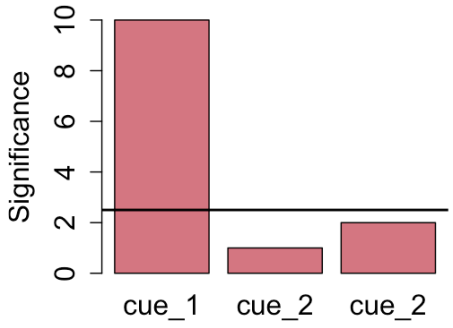
Misalignment of Production and Perception

- Similar case: Standard Southern British (Harrington et al., 2008; Harrington, 2012)
  - /u/-fronting: younger speakers shift in both production and perception; in alignment
  - /ʊ/-fronting: similar production, different perception between age groups
  - /u/-fronting began earlier than /ʊ/-fronting



(Harrington, 2012)

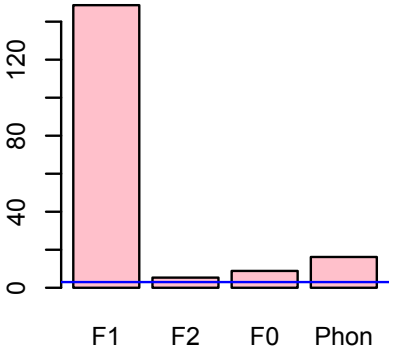
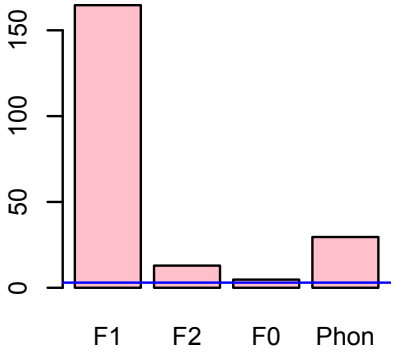
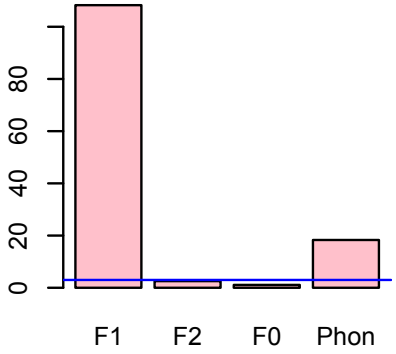
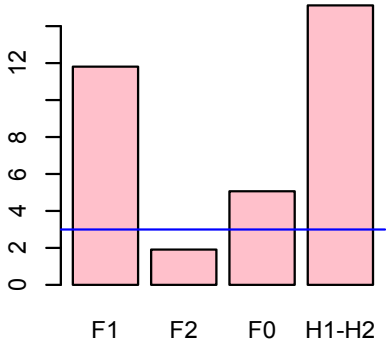
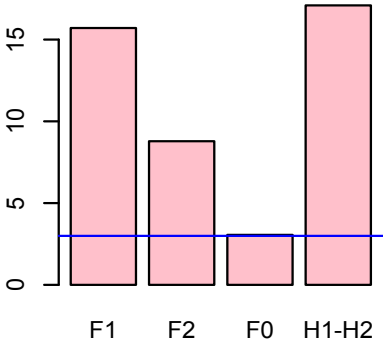
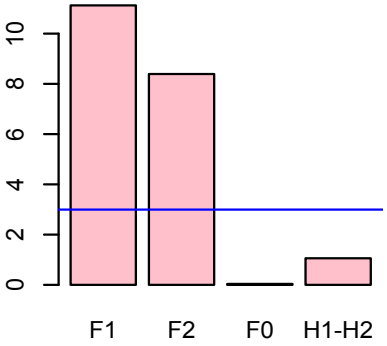
3 Stages of Sound Change

	Before change	During change	After change
Perception			
Production			

Different stages of change

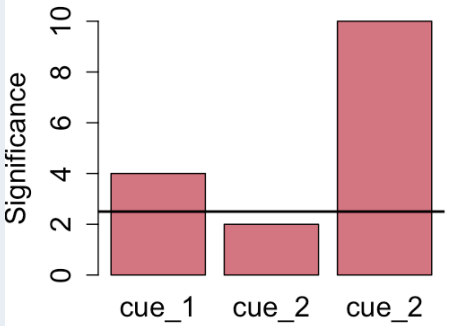
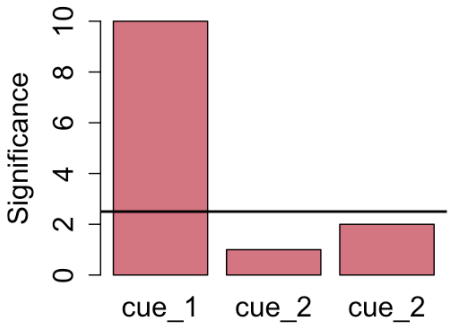
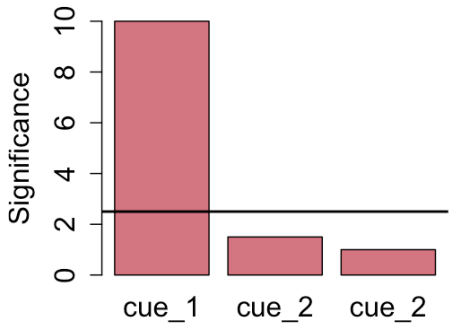
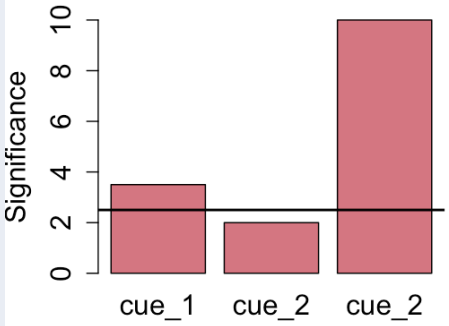
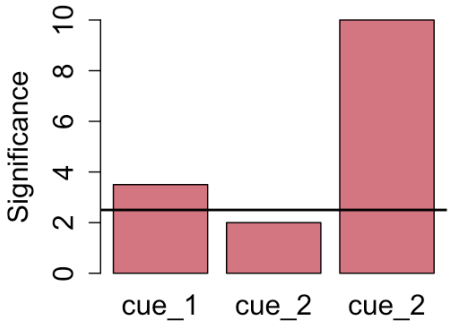
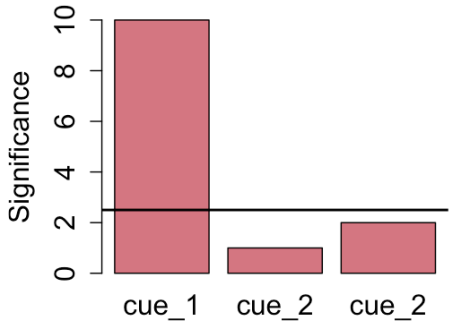
- /be/:
  - Younger speakers: Change complete
    - perception and production both shift from phonation to formants
  - Older speakers: shift started in perception
- /bu/: Earlier stage of change
  - Older speakers:
    - Both production and perception still rely on phonation
  - Younger speakers:
    - Shift started in perception

/be/

	old	middle	young
Perception			
Production			

3 Stages of Sound Change

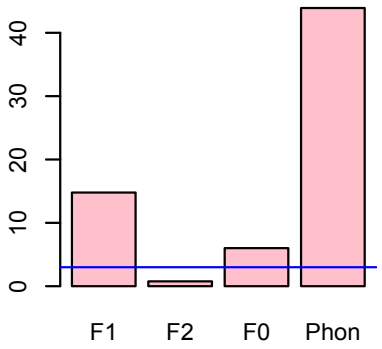
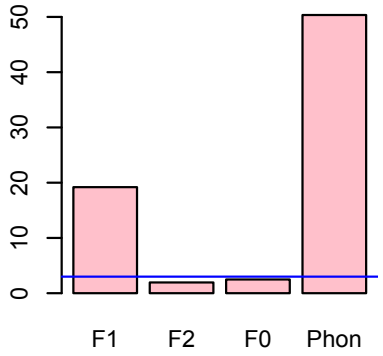
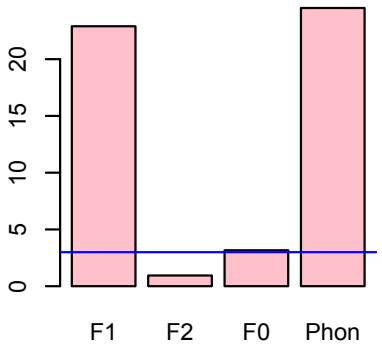
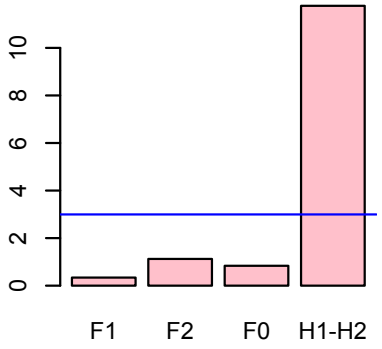
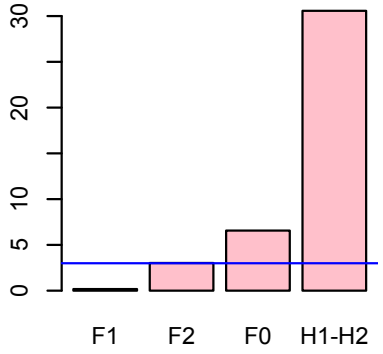
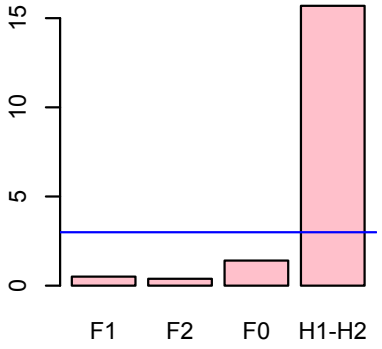
Older **Younger**

	Before change	During change	After change
Perception			
Production			

Different stages of change

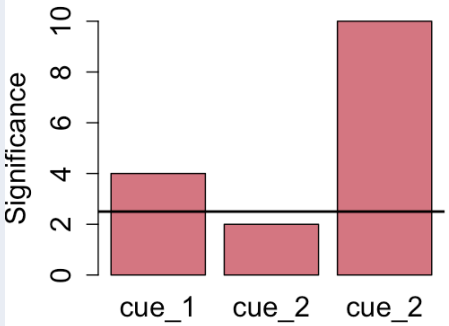
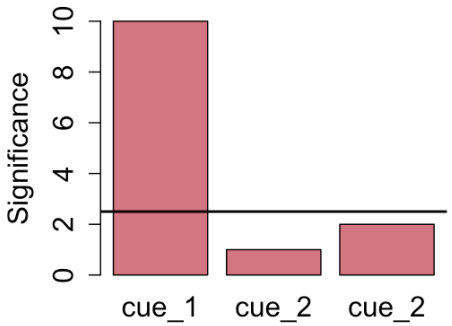
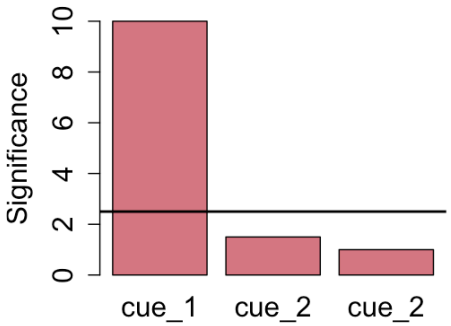
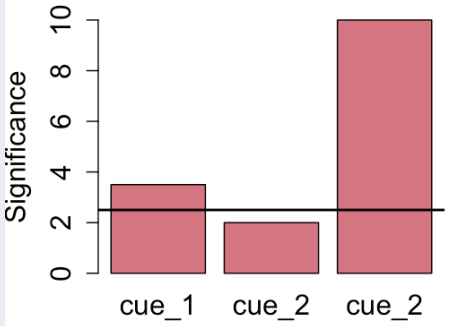
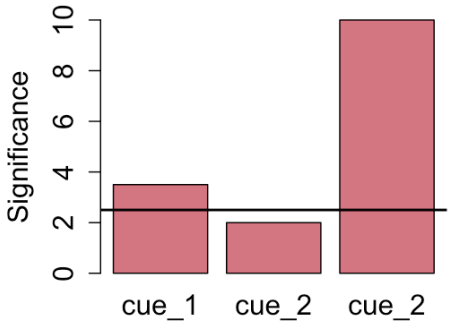
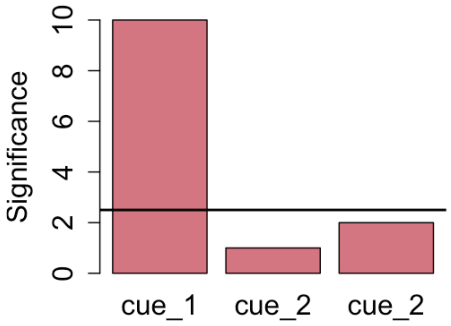
- /be/:
  - Younger speakers: Change complete
    - perception and production both shift from phonation to formants
  - Older speakers: shift started in perception
- /bu/: Earlier stage of change
  - Older speakers:
    - Both production and perception still rely on phonation
  - Younger speakers:
    - Shift started in perception

/bu/

	old	middle	young
Perception			
Production			

3 Stages of Sound Change

Older **Younger**

	Before change	During change	After change
Perception			
Production			

Take home message

- While the register contrast is maintained, formants are overtaking phonation as the primary cues
- Low vowels before high vowels
- Perception before production

Thank you!

Acknowledgments

- This study is supported by a URF award of University of Pennsylvania to Jianjing Kuang
- Thanks to all the Yi friends who kindly participated in our study!

The authors with the hosts