A corpus-based study of positional variation in Seoul Korean vowels

Yoonjung Kang
October 12, 2013
Japanese/Korean Linguistics Conference 23
MIT

"A corpus-based study of Korean dialects: microvariation and language universals"

- Dr. Tae-jin Yoon (Cheongju University)
- Dr. Sungwoo Han (Inha University)
- The National Institute of the Korean Language (NIKL)
- The Archive and Research Center for Korean Recordings
- UTSC Library Digital Scholarship Unit
- *The Social Science and Humanities Research Council of Canada—Partnership Development Grant* (# 890-2012-25)

Social Sciences and Humanities Research Council of Canada

Conseil de recherches en sciences humaines du Canada

Canada

Structure of the talk

Part I. Overview & context

Part II. Positional asymmetry in Korean vowels

Part III. Methods

Part IV. Age and gender-based variation

Part V. Positional variation

Part I: Overview and Context

Privileged positions in phonology

- Stressed vs. unstressed syllable
- Word-initial vs. non-initial syllable
- Root vs. affix
- Onset vs. coda

(Beckman 1998, Crosswhite 2004, Smith 2005, Barnes 2006, Walker 2011, etc.)

Vowels in privileged positions

- Resist deletion
- Preserve contrast lost in other positions
- Controls harmony

Source of position effect

- Phonetic prominence
  - **Stressed syllable (vs. Unstressed syllable)**
  - Long vowel
- Psycholinguistic prominence
  - **Word-initial syllable**
  - Root

(Smith 2005)

Stressed vowels

- Phonetics of stressed vowels
  - Stressed vowels are longer; More likely to reach the articulatory target
  - Unstressed vowels are shorter; non-high vowels are raised, not reaching the articulatory target
  - Reduction of vowel contrast in unstressed vowels

(Lindblom 1963, Flemming 2004, 2005, Barnes 2006, Giavazzi 2010)

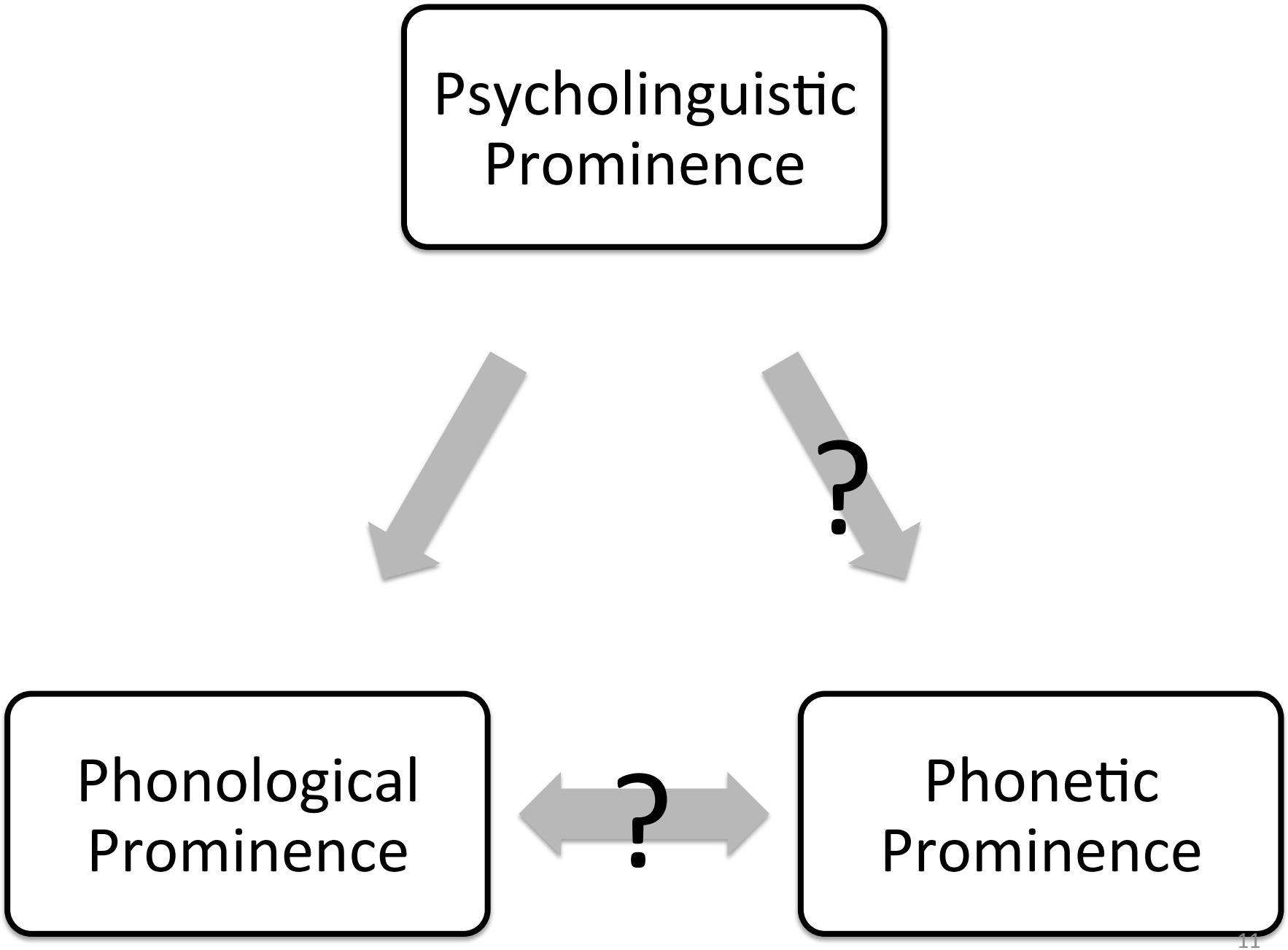
Initial syllables

- Psycholinguistic grounding
  - Important in lexical access and retrieval
  - More important to maintain the contrast in initial than non-initial position

(Smith 2005)

Phonetics of vowels in initial syllables

- Domain-initial strengthening: duration and articulatory strength (Fougeron 1999, Cho and Jun 2001, etc.)
- Effect limited mainly to segments in absolute initial position
- Not as widely studied
  - Turkish: Vowels in Initial syllables are longer than those in non-initial syllables (Barnes 2006)
  - Creek: Vowels in initial syllables are more peripheral than vowels in final syllables (Johnson and Martin 2003)


Psycholinguistic Prominence
?
Phonological Prominence
?
Phonetic Prominence

Part II: Positional asymmetry in Korean vowels

Type I: resistance to contrast reduction

- Vowel length contrast only in initial position
  - /nun/ ‘eye’ /nu:n/ ‘snow’
  - /k’amak-nun/ ‘illiterate’ /hampak-nun/ ‘large snow flake’
  - /mal/ ‘horse’ /ma:l/ ‘speech’
  - /colaŋ-mal/ ‘pony’ /pa:n-mal/ ‘talk down’

Type I: resistance to contrast reduction

- Two-step merger of /ɔ/
  - /ɔ/ merges with /ɨ/ in **non-initial syllable.** (15th -16th century)

    /hanɔ̲l/ > /hanɨ̲l/ 'sky'

    /nakɔ̲naj/ > /nakɨ̲ne/ 'wanderer'

    /talɔ̲-/ > /talɨ̲-/ 'different'
  - /ɔ/ merges with /a/ in **initial syllable**. (18th century)

    /pɔ̲lam/ > /pa̲lam/ 'wind'

    /tɔ̲l/ > /ta̲l/ 'moon'

    /hɔ̲-/ > /ha̲-/ 'do'
- Older speakers of Jeju dialect still retain /ɔ/, only in **word-initial** position.

  /nɔ̲mɔ̲l/ > /nɔ̲mʌ̲l/ 'vegetables'

(K.M. Lee 1972)

Type I: resistance to contrast reduction

- Merger of /ɛ/ with /e/ in central, southern, & Jeju dialects
  - /ɛ/ raises and merges with /e/.
  - The contrast is merged in **non-initial** syllables and is retained in **initial** syllables.

  (Chung 2002, Dialect Dictionary 2001)

Type I: resistance to contrast reduction

- Raising of **non-initial** /o/ to /u/ (19th century~)
  - Lexically conditioned

    /namo/ > /namu/ 'tree'

    /ʌlkol/ > /ʌlkul/ 'face'

    /ocom/ > /ocum/ 'urine'

  Cf. /kalo/ > */kalu/ 'width'

  - Lexical strata

    /kacok/ > /kacuk/ 'leather' (Native)

    /kacok/ > */kacuk/ 'family' (Sino-Korean)

  (Chae 1995)

Type II: controls harmony

- Vowel harmony in ideophones

  “dark”: ʌluk tʌluk (e,ʌ,u)

  D D D D

  “light”: ɑlok tɑlok (ɛ,ɑ,o)

  L L L L

- High vowels (i, ɨ):
  - “dark” in **initial syllable**

    k’ɨcʌk, *k’ɨcɑk (“light”: k’ɛcɑk)

    D D *__D__ L L L

  - “neutral” in **non-initial syllable**

    tɑlkɨlɑk, tʌkɨkʌk

    L **N** L D **N** D

Type II: controls harmony

- Vowel harmony in verbal suffixes
  - "dark" vowels (i,ɨ,u,e,ɛ,ʌ) take /-ʌ/: /mʌk-ʌ/ 'eat'
    D D
  - "light" vowels (o,ɑ) take /-ɑ/: /cop-ɑ/ 'narrow'
    L L
- Variation in p-irregular verbs
  - **Monosyllabic** p-irregular stems: /-ɑ/ is obligatory
    /top-a/ [tow-ɑ], *[tow-ʌ] 'help'
    L L ***L D**
  - **Polysyllabic** p-irregular stems: /-ɑ/ ~ /-ʌ/
    /sɛlop-a/ [sɛlow-ɑ] ~ [sɛlow-ʌ] 'new'
    L L **L D** (J.K. Kim 2000)

Questions

- Are vowels in initial syllable phonetically special (in Korean)?
  - Longer duration? (See Tae-Jin Yoon's presentation later today)
  - Vowel quality difference? How?
  - If vowel quality is different, is it merely an epiphenomenon of longer duration?
  - Is the positional difference due to root-affix asymmetry?

Part III: Methods

The Reading-Style Speech Corpus of Standard Korean

- The National Institute of the Korean Language (2002)
- 60 male, 60 female, Seoul-Kyeoungki area
- Age/Year of Birth distribution

Age	**20s**	**30s**	**40s**	**50s**	**60~70s**
Male	20	20 (19)		20	20
Female	20		20	20	20 (19)

Year of Birth	**1930s**	**1940s**	**1950s**	**1960s**	**1970s**	**1980s**
Male	4	12	4	8	27 (26)	5
Female	2	10 (9)	25	3	11	9

	Text	# of sentences	Duration (mv01)
1.	요람기	51	321.92 sec (5.37 min)
2.	소나기	87	240.12 sec (4.00 min)
3.	아름다움에 대하여	69	311.92 sec (5.20 min)
4.	방망이를 깎던 노인	62	275.50 sec (4.59 min)
5.	문학의 세계와 삶의 세계	47	321.01 sec (5.35 min)
6.	독서와 인생	54	294.97 sec (4.92 min)
7.	수난 이대	62	317.31 sec (5.29 min)
8.	메밀꽃 필 무렵	94	425.98 sec (7.10 min)
9.	**토끼와 자라**	**60**	**185.27 sec (3.09 min)**
10.	**선녀와 나무꾼**	**73**	**295.03 sec (4.92 min)**
11.	**호랑이와 곶감**	**42**	**159.92 sec (2.67 min)**
12.	**해님 달님**	**28**	**124.10 sec (2.07 min)**
13.	**그리운 시내가**	**36**	**276.49 sec (4.61 min)**
14.	**광화문 지하도 아주머니**	**27**	**148.40 sec (2.47 min)**
15.	**막 지은 밥**	**17**	**106.36 sec (1.77 min)**
16.	**눈 오던 날**	**35**	**147.77 sec (2.46 min)**
17.	**숭늉의 지혜**	**19**	**95.00 sec (1.58 min)**
18.	**까만 눈동자들 앞에서**	**27**	**205.92 sec (3.43 min)**
19.	**내 고향 개울 물에서**	**40**	**163.34 sec (2.72 min)**
	Total	930	4416.40 sec (73.61 min)

From Yoon (2013)

Seoul Korean vowels

i	ɨ	u
e	ʌ	o
(ɛ)	ɑ	

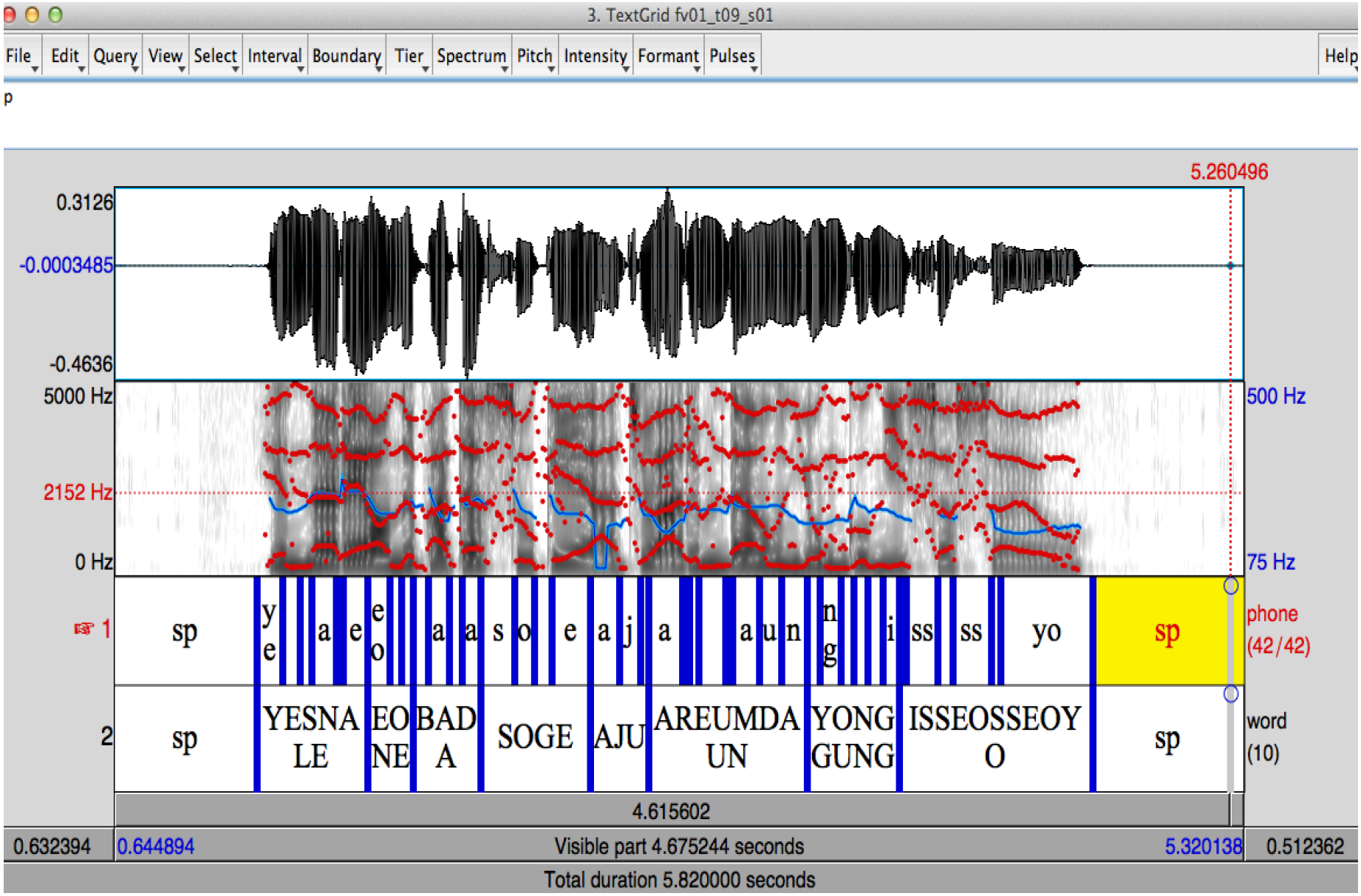
of monophthongal vowels

- 9738 tokens per speaker

ɑ (ㅏ)	ɛ (ㅐ)	e (ㅔ)	ʌ (ㅓ)	ɨ (ㅡ)	i (ㅣ)	o (ㅗ)	u (ㅜ)
2540	421	427	1224	1555	1683	1165	724

Korean phone-aligner

- Developed by Tae-Jin Yoon (cf. Yoon and Kang 2012, Yoon 2013)
- Given a sound file and its Korean transcription, the aligner provides a segmentation of the file into component phones.


3. TextGrid fv01_t09_s01
File Edit Query View Select Interval Boundary Tier Spectrum Pitch Intensity Formant Pulses Help
p
5.260496
0.3126
-0.0003485
-0.4636
5000 Hz
2152 Hz
0 Hz
500 Hz
75 Hz
1
sp
y e a e e o a a s o e a j a a u n n g i ss ss yo
sp
phone (42/42)
2
sp
YESNA LE
EO NE
BAD A
SOGE
AJU
AREUMDA UN
YONG GUNG
ISSEOSSEOY O
sp
word (10)
4.615602
0.632394
0.644894
Visible part 4.675244 seconds
5.320138
0.512362
Total duration 5.820000 seconds

Formant measurements

- Linear Predictive Coding (LPC) method, as implemented in *Praat*'s "To Formant (burg)" function.
- To minimize the influence of local formant tracking errors, the average of the measurements from mid 20% of the vowel duration was used.

LPC formant ceilings

- In the LPC method, users define the number of formants to be tracked and the frequency range where those formants should be searched for.
- Adult male [ə]
  - 500Hz, 1500Hz, 2500Hz, 3500Hz, 4500Hz…
  - 5 formants under **5000** Hz
- Adult female [ə]
  - 583Hz, 1750Hz, 2916Hz, 4083Hz, 5250Hz…
  - 5 formants under **5500** Hz

→ *Praat* recommended default settings

But, there are vowel-specific and speaker-specific variations.

Optimizing formant ceilings

- Compare formant measurements based on many different formant ceiling settings
- Choose the setting that minimizes variance for a given vowel for a given speaker.
  - **Escudero et al. (2009, *JASA*: Portuguese dialect comparison)**
  - Yao et al. (2010, *UCBerkley lab report*: Buckeye Corpus),
  - Labov et al. (2013, *Language*: Philadelphia Language Variation and Change Corpus)

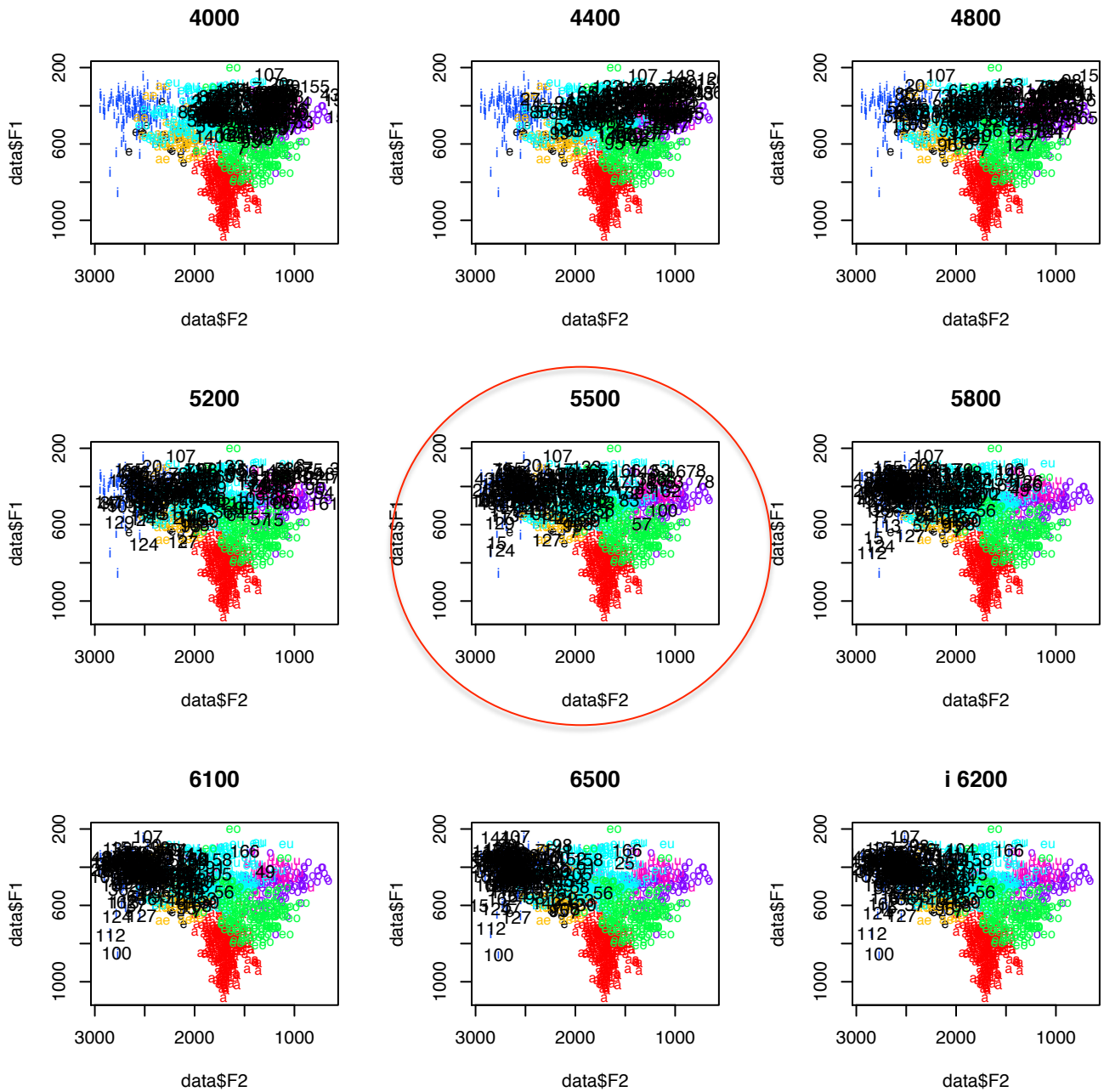
Optimal formant ceilings

- At every 100 Hz interval in 4000~6500 Hz

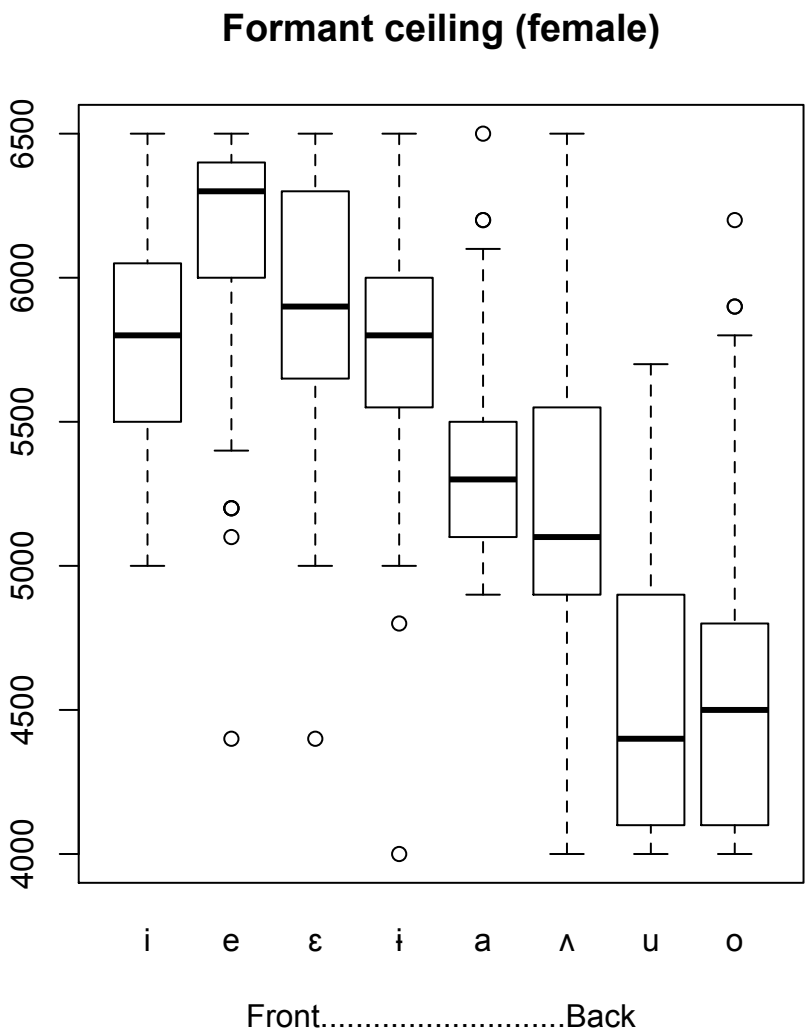
➔ 26 formant measurements per token

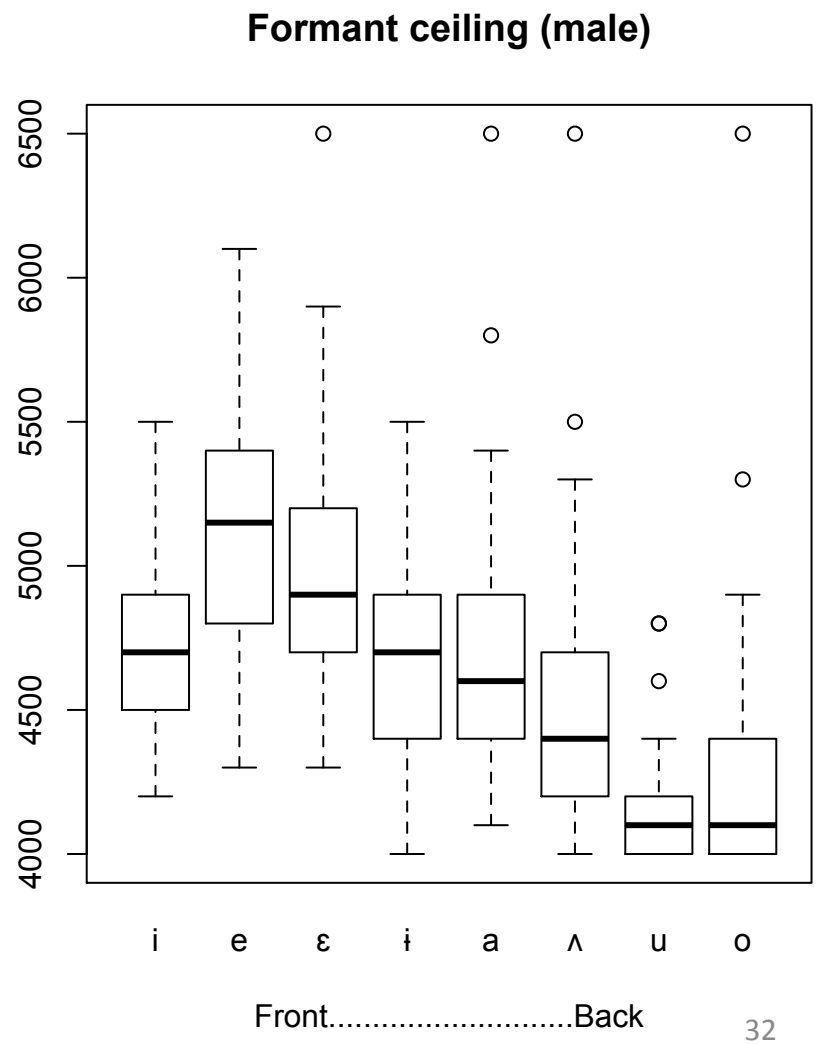
- For each vowel for each speaker, pick the formant ceiling that minimizes the variance (=minimizes the measurement errors).
  - variance(20*log(F1)) + variance(20*log(F2))

fv04: /i/
Ceiling: 6200 Hz



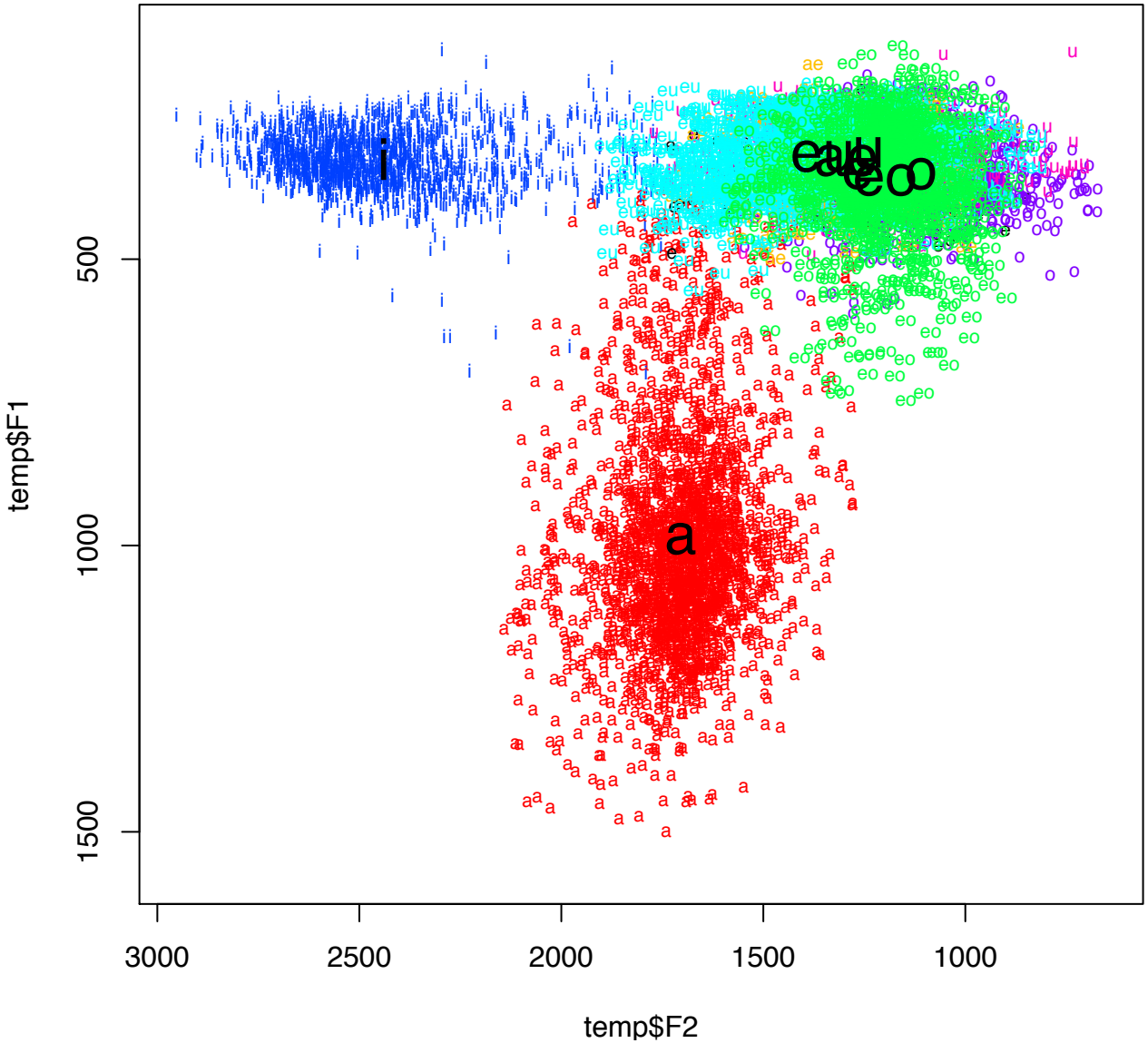
Distribution of optimal ceilings





Bad majority rules

fx07



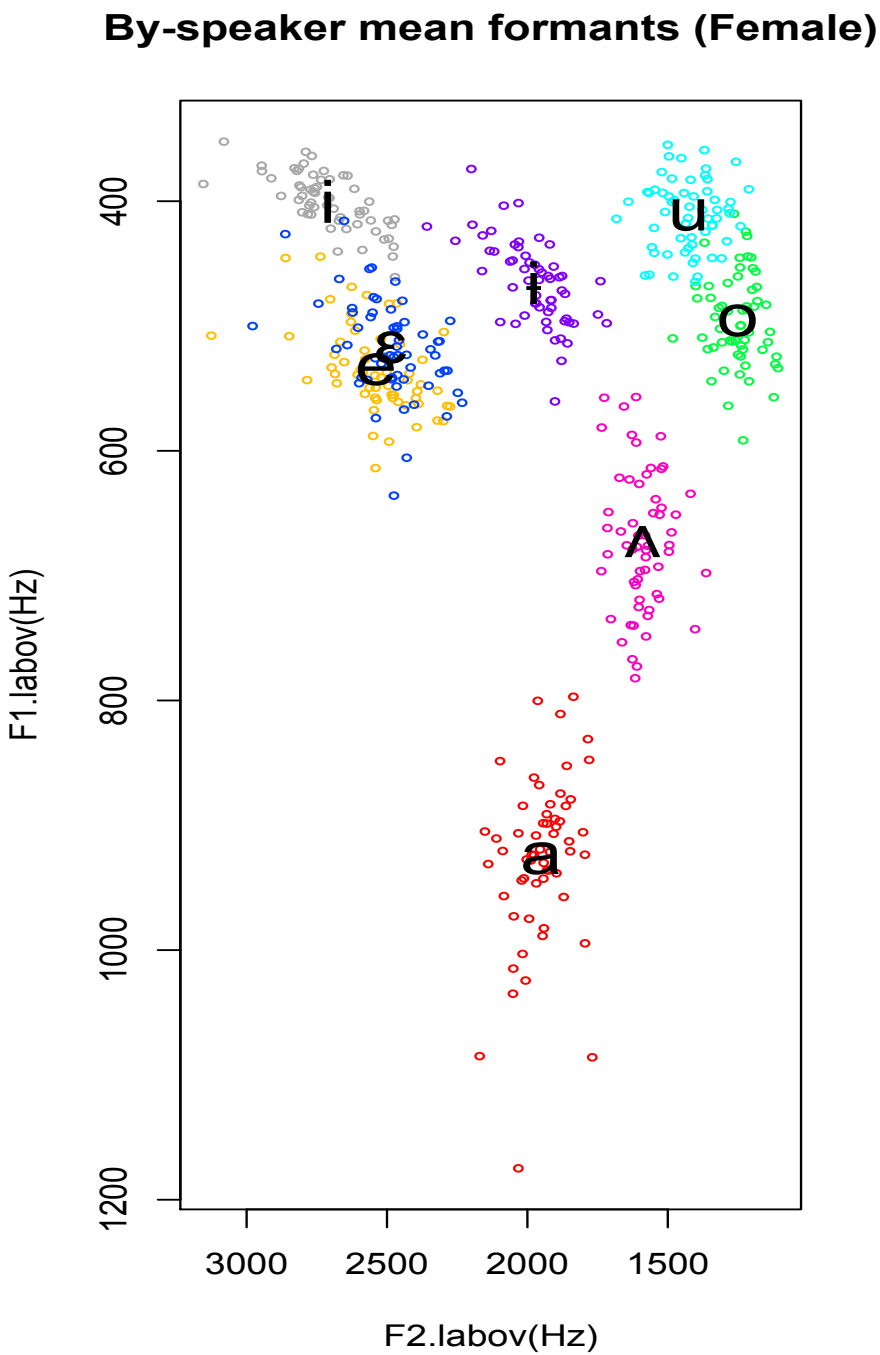
Gender and Vowel-specific ceilings

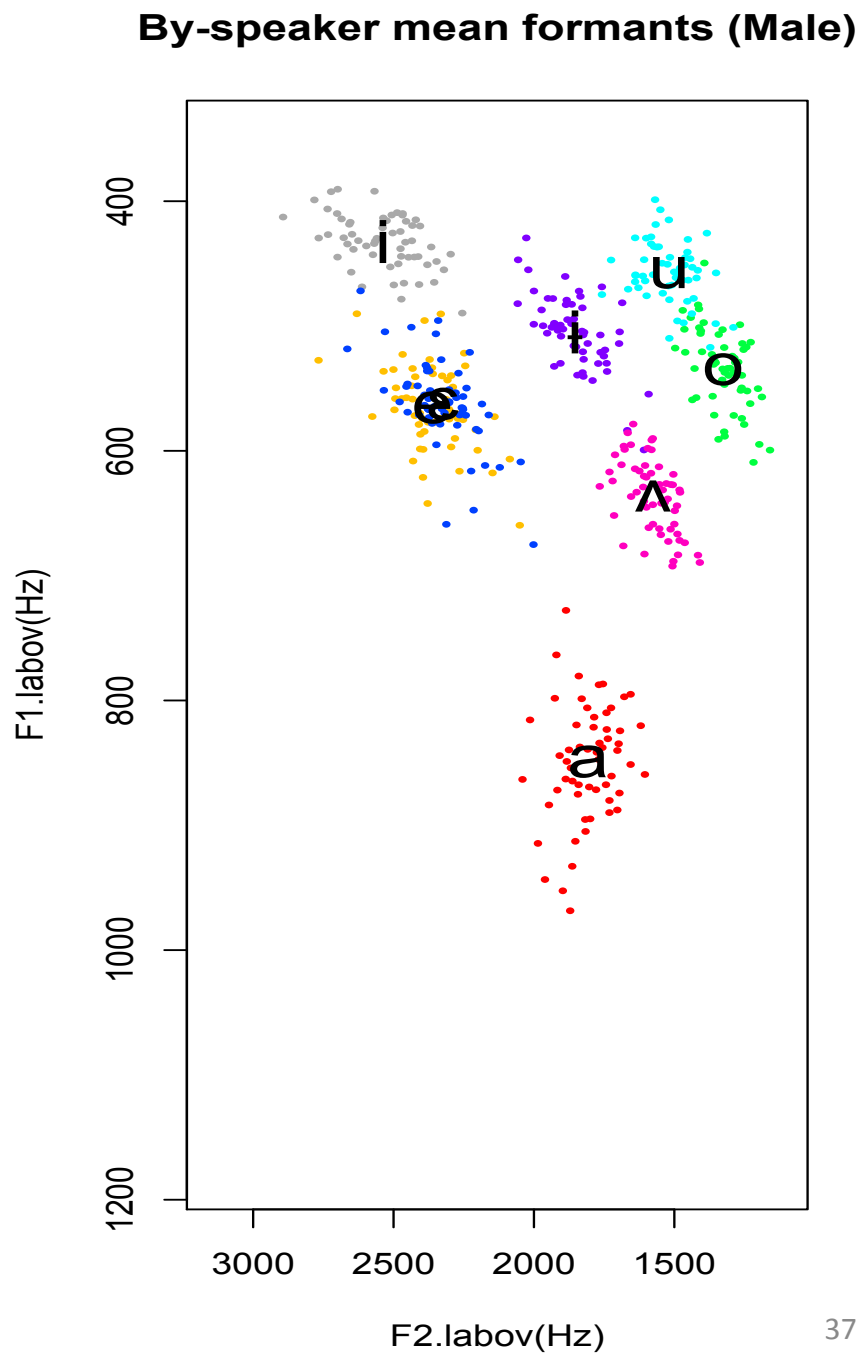
- For each vowel for each gender, the mean ceiling values are used as the optimal ceiling.
- This method
  - provides improvement in the accuracy of formant measurements compared to the fixed ceiling method but
  - avoids the “bad majority” problems.

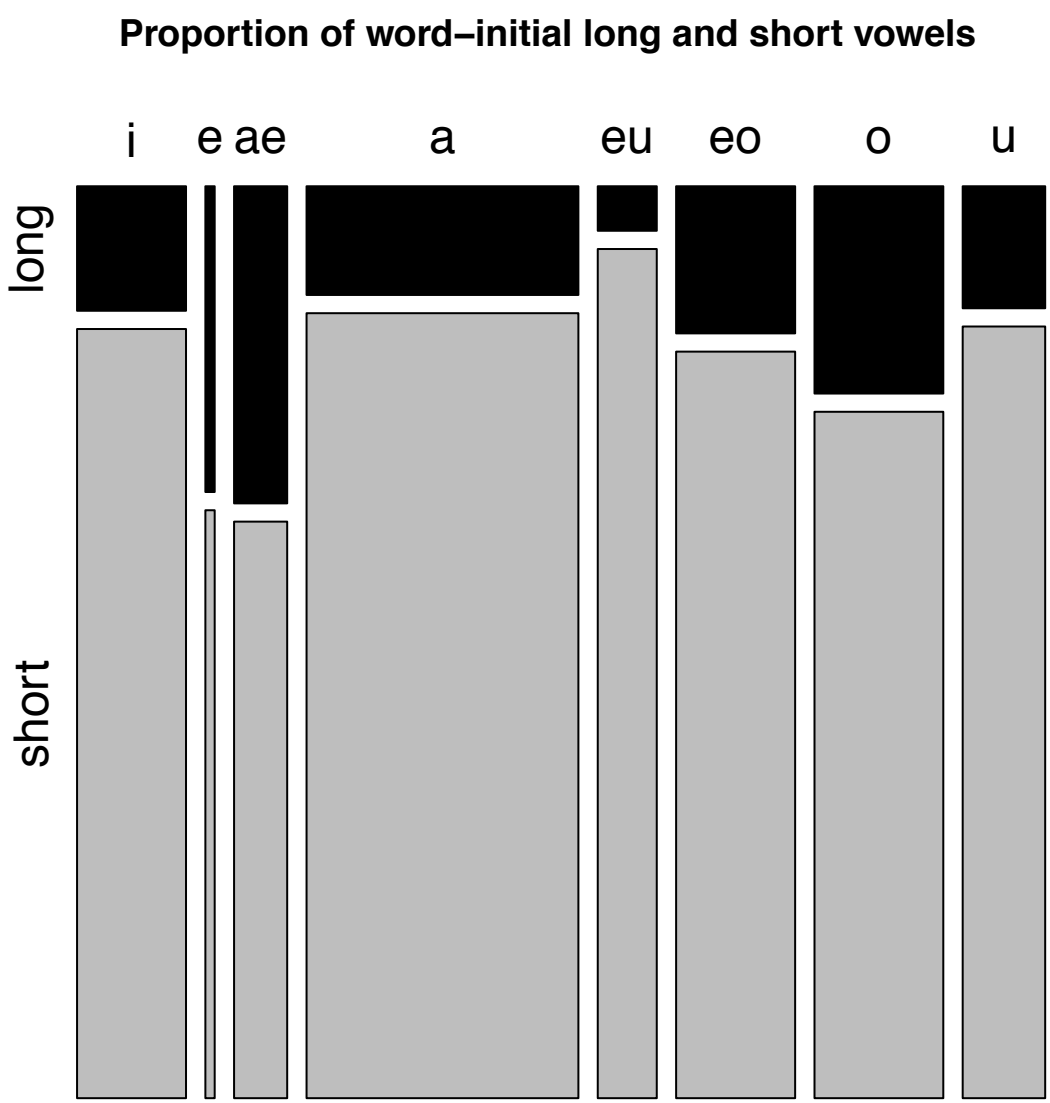
Normalization

- Factor out inter-speaker variation due to physiological difference
- Lobanov
  - By-speaker z-score transformations for each F1 and F2
- Labov
  - Recalibration of formants (Hz) based on the by-speaker grand mean of F1 and F2.
- Below we report results that are consistent through both normalization methods.

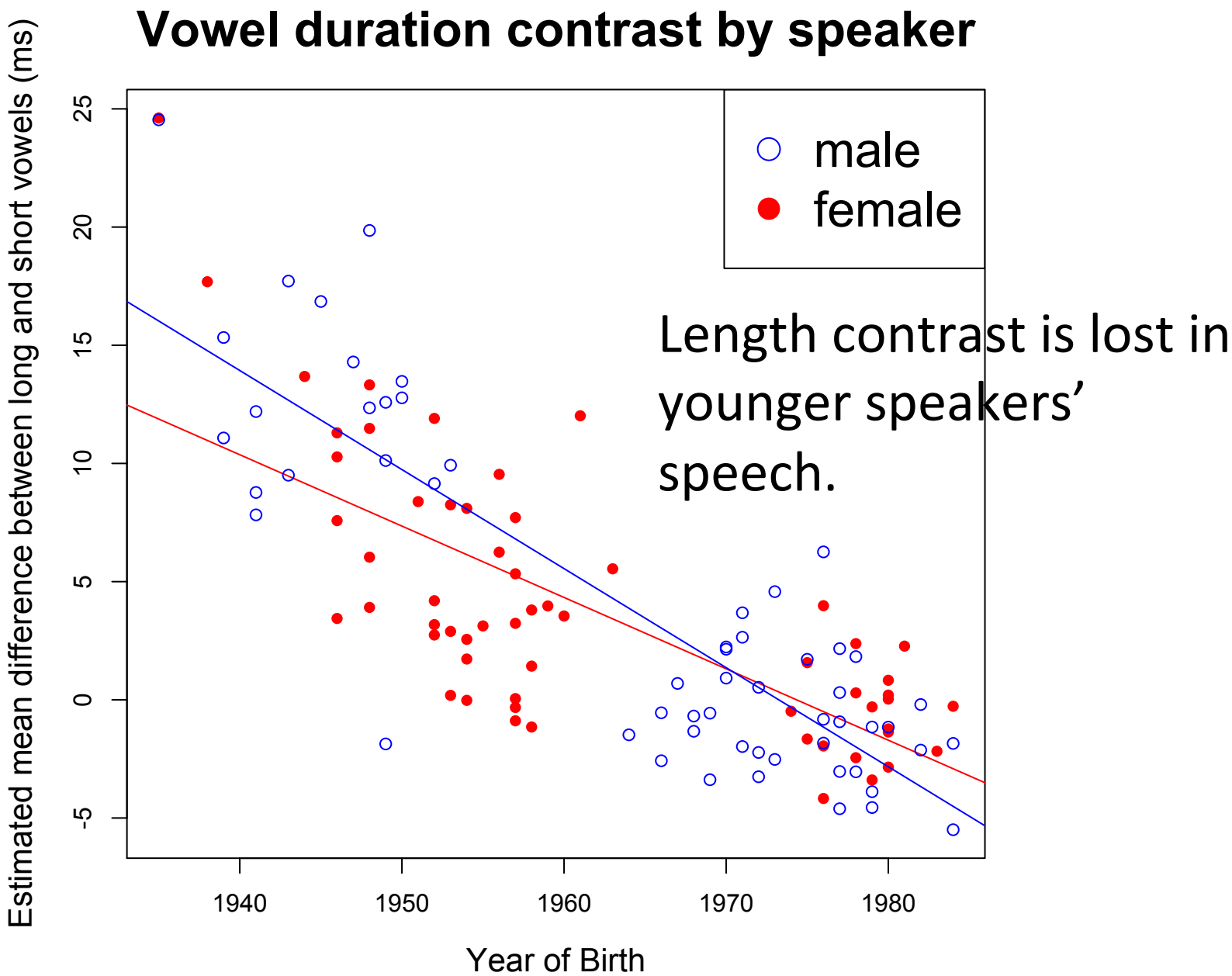
Part IV: Age and gender-based variation

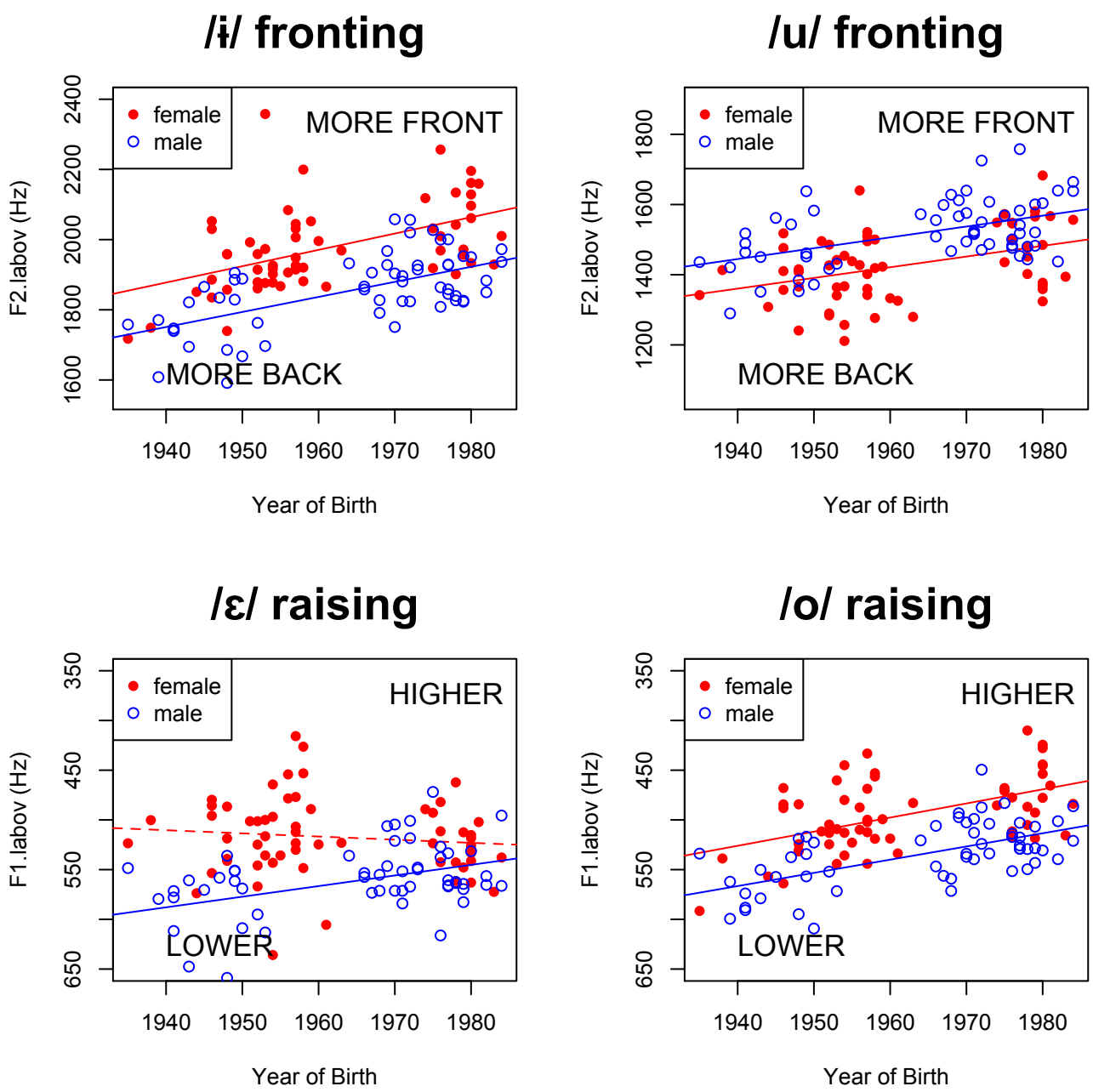

By-speaker mean formants (Female)
F1.labov(Hz)
F2.labov(Hz)
400
600
800
1000
1200
3000
2500
2000
1500
i
ɨ
u
o
ɛ
æ
ʌ
a


By-speaker mean formants (Male)
F1.labov(Hz)
F2.labov(Hz)
400
600
800
1000
1200
3000
2500
2000
1500
i
ɨ
u
o
ɛ
æ
ʌ
a



Based on length specification in *Great Standard Korean dictionary (*http://stdweb2.korean.go.kr/main.jsp)


Vowel duration contrast by speaker
Estimated mean difference between long and short vowels (ms)
25
20
15
10
5
0
-5
1940
1950
1960
1970
1980
Year of Birth
male
female
Length contrast is lost in younger speakers' speech.


/ɨ/ fronting
female
male
MORE FRONT
MORE BACK
F2.labov (Hz)
1600
1800
2000
2200
2400
1940
1950
1960
1970
1980
Year of Birth
/u/ fronting
female
male
MORE FRONT
MORE BACK
F2.labov (Hz)
1200
1400
1600
1800
1940
1950
1960
1970
1980
Year of Birth
/ɛ/ raising
female
male
HIGHER
LOWER
F1.labov (Hz)
350
450
550
650
1940
1950
1960
1970
1980
Year of Birth
/o/ raising
female
male
HIGHER
LOWER
F1.labov (Hz)
350
450
550
650
1940
1950
1960
1970
1980
Year of Birth

i ← ɨ ← u

↑

o

e

↑

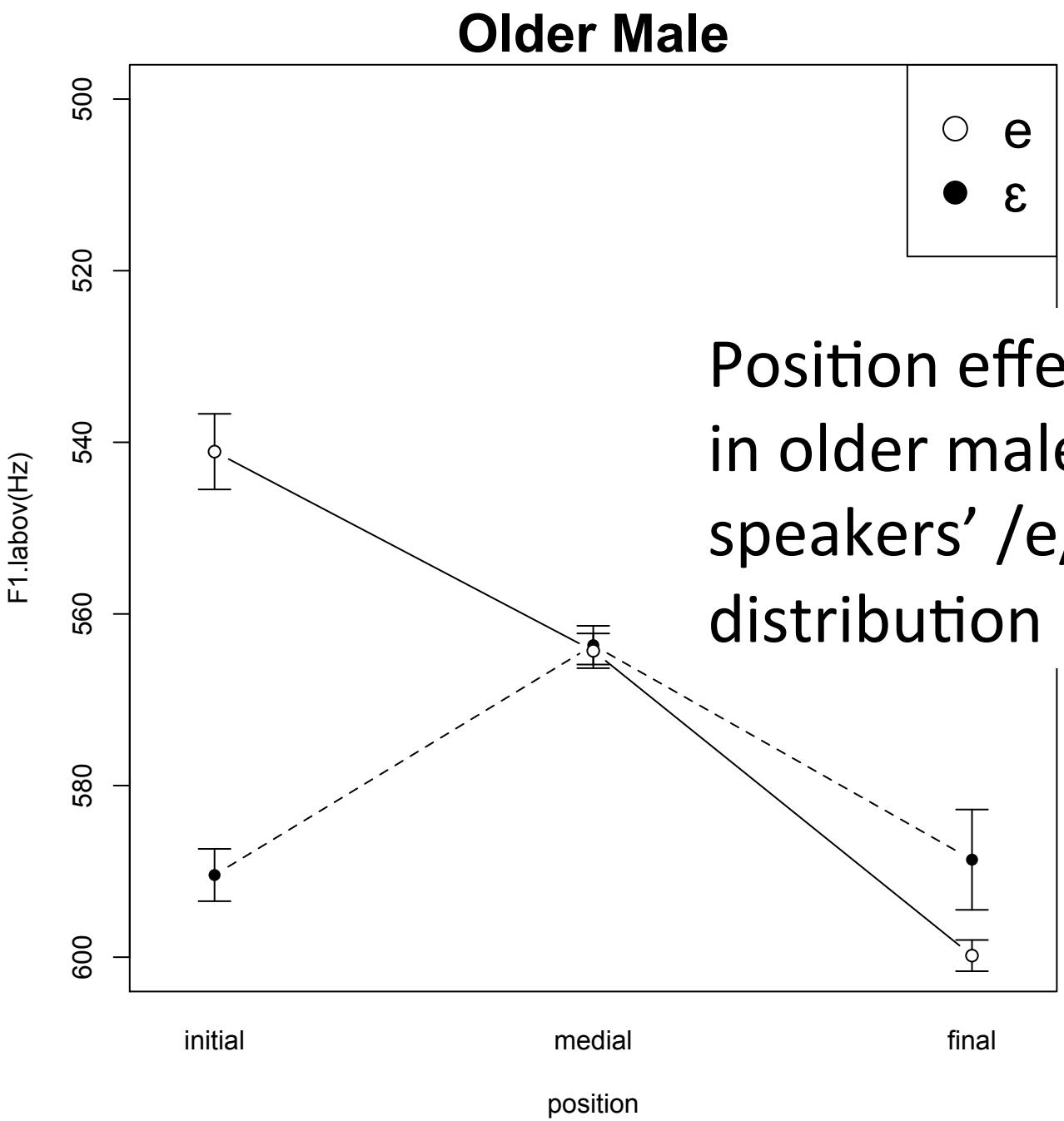
ɛ

ʌ

No loss of contrast: chain shift-like movement

/ɛ/ > /e/ raising & merger is complete for females and nearing the end for males.

ɑ

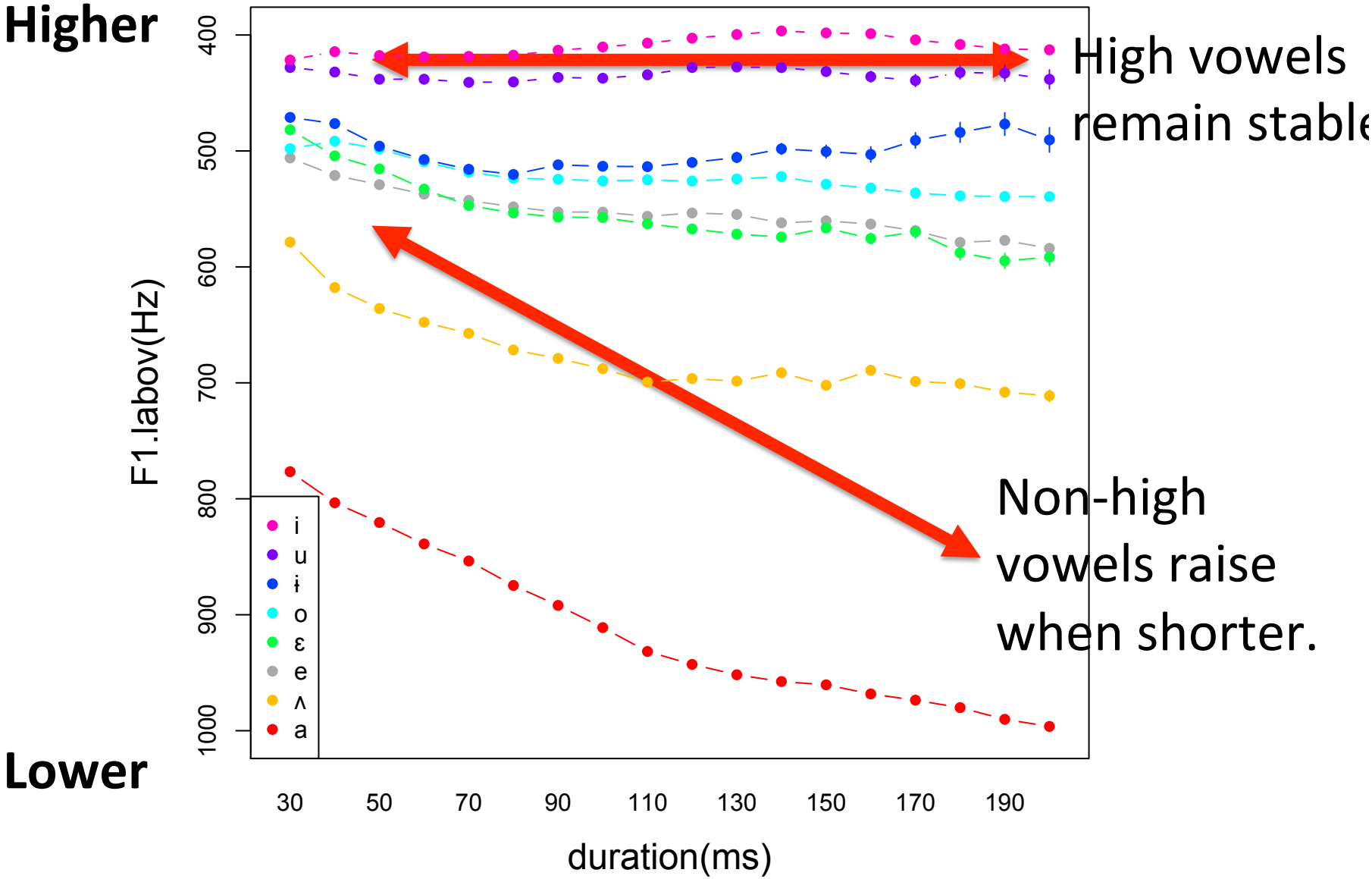


Position effect visible in older male speakers' /e/ ~ /ε/ distribution

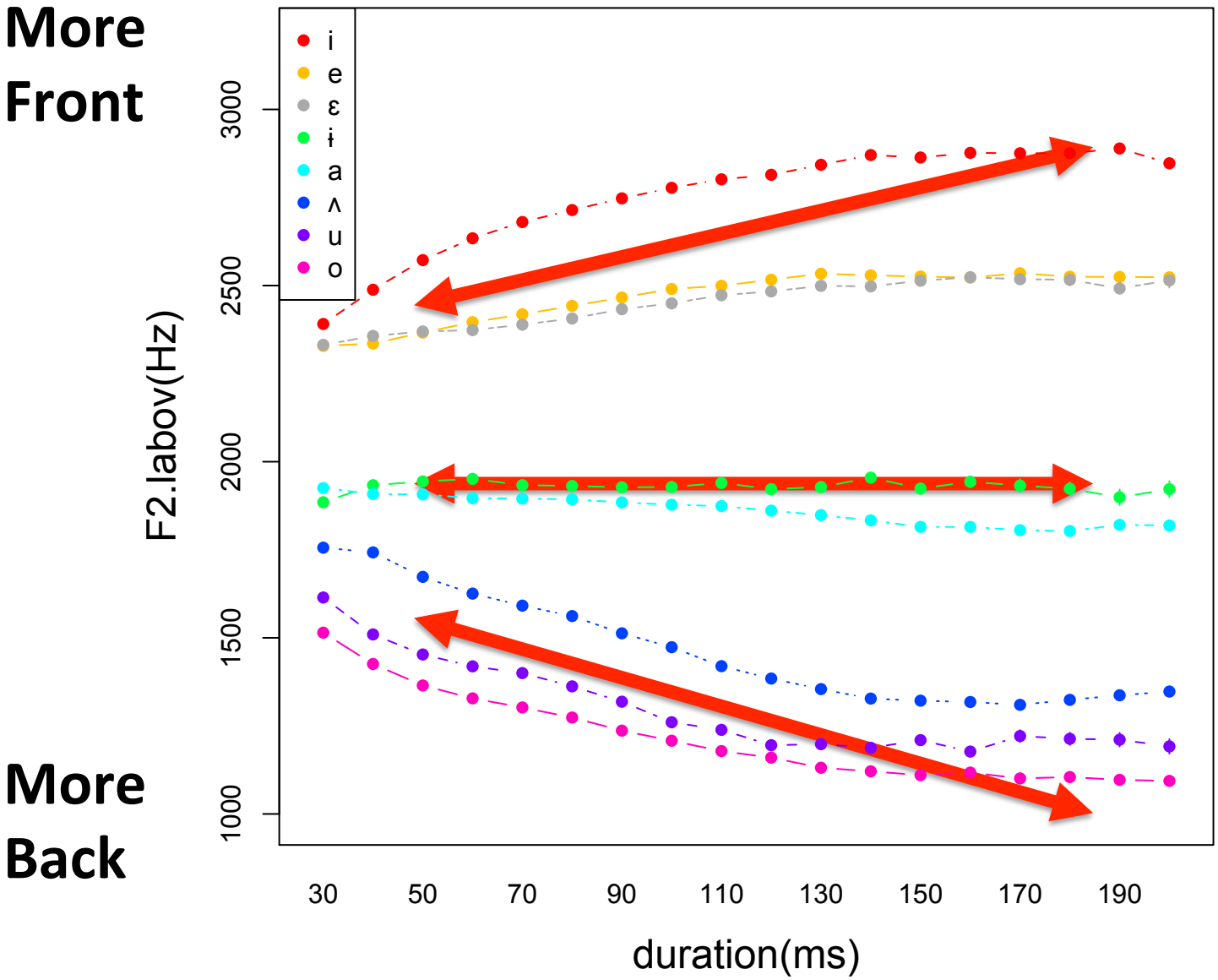
Part IV: Positional asymmetry

(1) Duration and vowel quality

Short vowel raising



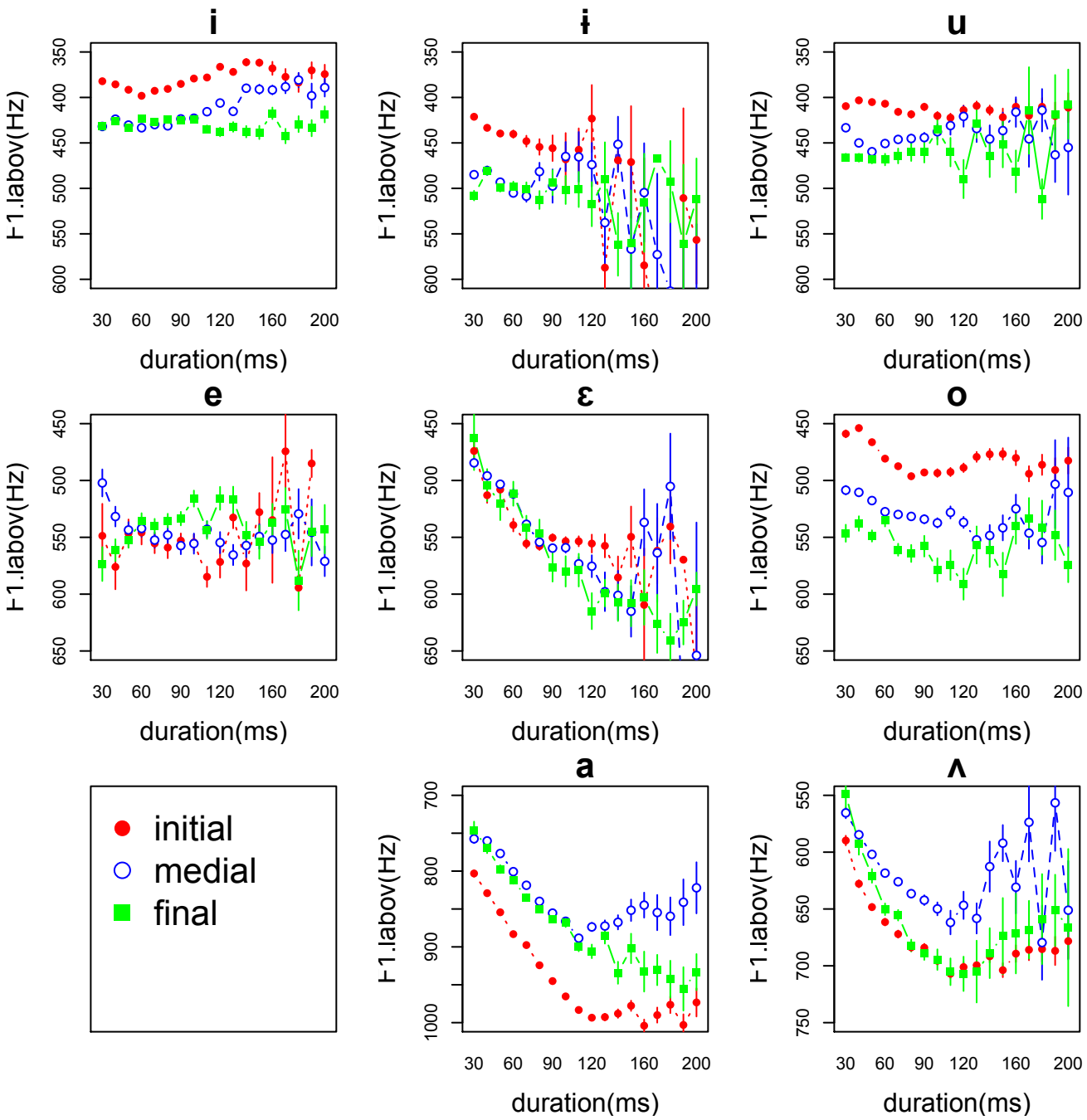
Short vowels centralize



(2) Position and vowel quality

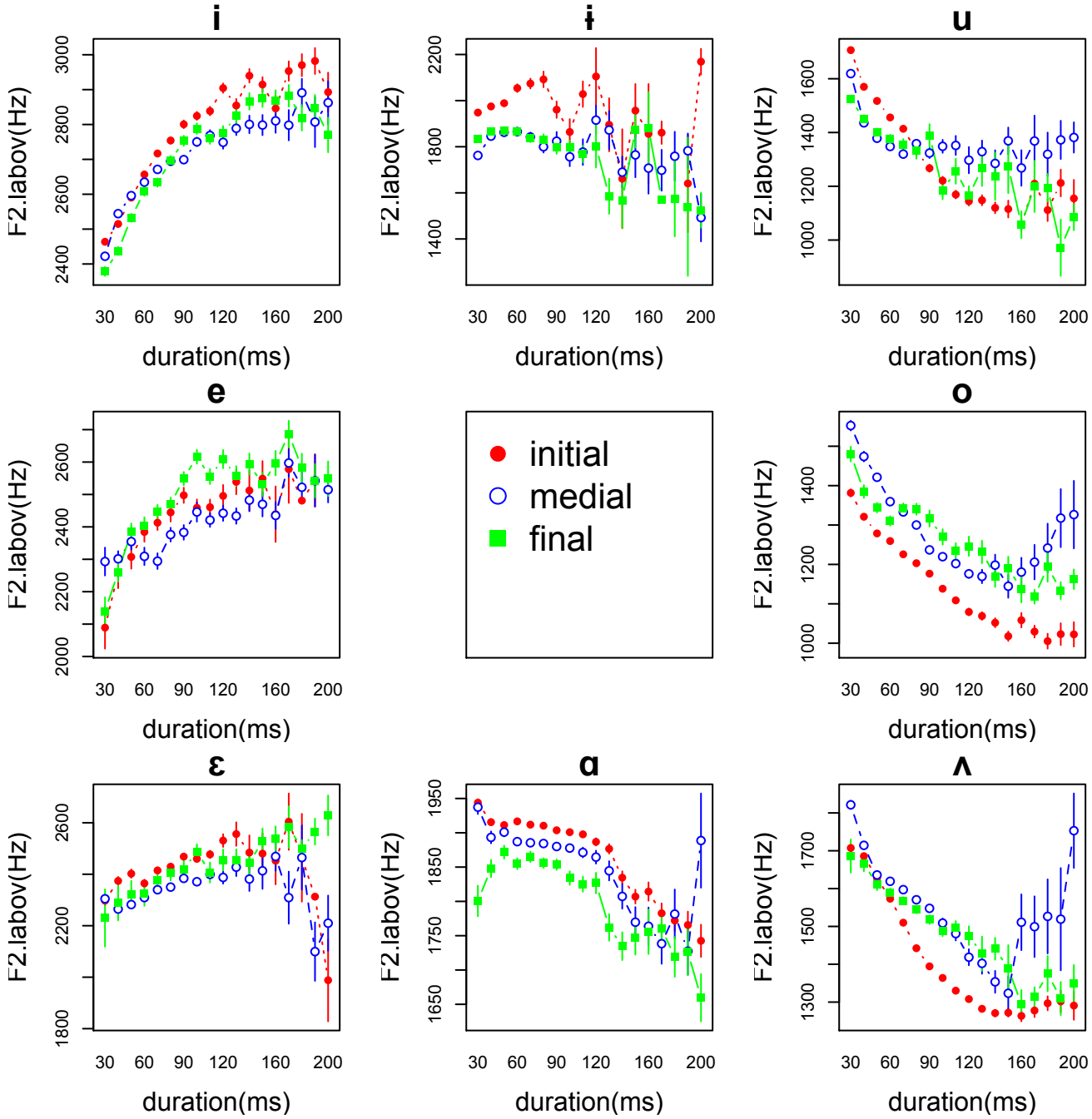
Position and Vowel quality

- How does position affect vowel quality?
  - To avoid the vowel length contrast confound, only "short" vowels are included
  - To avoid affix-root asymmetry counfound, only roots (nouns, verbs/adjectives, adverbs) are included. (cf. Morphological analysis based on R library (koNLP))


i
ɨ
u
e
ɛ
o
a
ʌ
F1.labov(Hz)
duration(ms)
initial
medial
final

Position and F1

- Vowels in initial position have more peripheral vowel targets; low vowels are lower and high vowels are higher.
- /o/ patterns like a high vowels. Maybe related to being part of the raising.


i
ɨ
u
e
o
ɛ
ɑ
ʌ
F2.labov(Hz)
duration(ms)
initial
medial
final

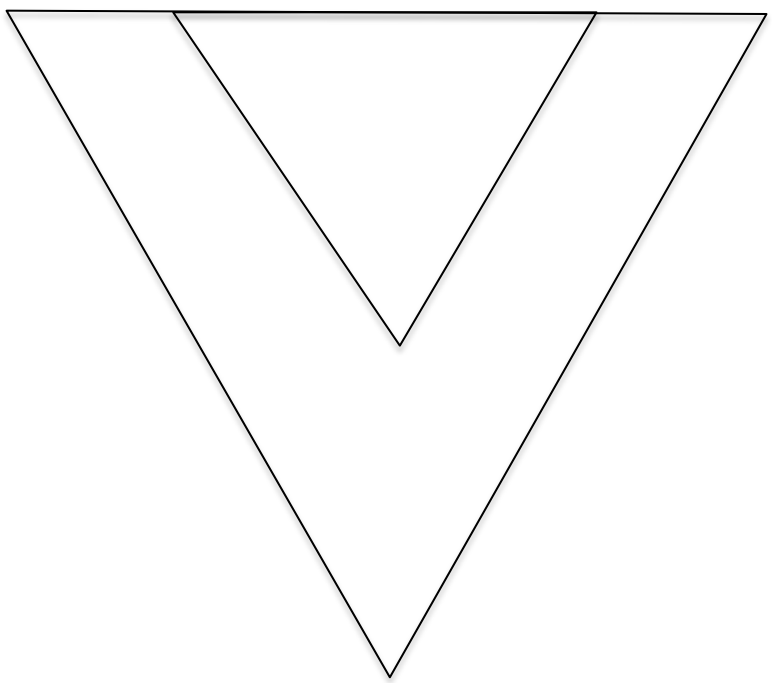
Position and F2

- F2: Vowels in initial position have more peripheral vowel targets; back vowels are more back and front vowels are more front (less clear)
- Effect seems less pronounced.
- /u/ patterns ambiguous; may be related to the fact that the vowel is in the move as part of the chain shift.
- /ɨ/ patterns like a front vowel; may be related to the fact that the vowel is in the move as part of the chain shift.

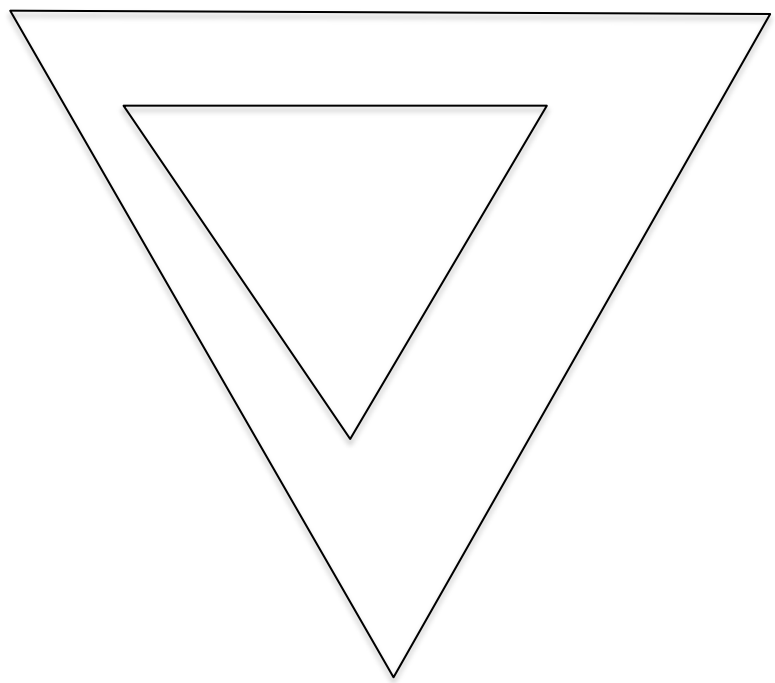
Positional vs. durational reduction

- The patterns of reduction are different.

Duration



Positional



Summary

- We identified a pattern of positional effect on vowel quality, independent of duration and root-affix asymmetry.
  - Duration-based quality difference is due to articulatory undershoot (vowel raising in short vowels),
  - position-based quality difference is speaker-controlled: more peripheral vowel targets for prominent position and the targets interact with the direction of sound change.

Selected references

Barnes, J. 2006. Strength and Weakness at the interface: Positional neutralization in Phonetics and phonology.

Crosswhite, K. 2004. Vowel reduction. In Hayes et al. ed. *Phonetically-based phonology.*

Flemming, E. 2005. A phonetically-based model of phonological vowel reduction.

Johnson, Keith, and Jack B. *Martin*. 2001. Acoustic *vowel* reduction in *Creek*: Effects of distinctive length and position in the word. Phonetica 58:81-102.

Kim, J. K. 2000. *Quantity sensitivity and feature sensitivity of vowels: A constraint-based approach to Korean vowel phonology*. PhD dissertation, Indiana University.

Lee, K.-M. 1972. Kwuke umwunsa yenkwu (A study of the history of Korean phonemes). Seoul: Hankwuk Mwunhwa Yenkwuso.

Smith, J. 2002. Phonological Augmentation in Prominent Positions. PhD dissertation, University of Massachusetts, Amherst.

Walker, R. 2011. *Vowel patterns in language.* Cambridge University Press.

Yoon, T.J. and Y.J. Kang 2012. A forced-alignment-based study of declarative sentence-ending 'da' in Korean. *The Proceedings of 2012 Speech Prosody.*