Linguistic voice quality

Pat Keating

UCLA Linguistics Department

UCLA Voice Projects

- Voice Science Consortium at UCLA:
  - Phonetics Lab, Linguistics Department
  - Abeer Alwan, Speech Processing and Auditory Perception Lab, Electrical Engineering Department
  - Jody Kreiman, Voice Perception Lab, Department Head & Neck Surgery, Geffen Medical School
- "Production and perception of linguistic voice quality"
- "A new voice source model"
- "Variance and invariance in voice quality"

Phonation

- **Phonation**: sound production in the larynx, usually by vocal fold vibration (**voice**, or **voicing**)
- How fast the folds vibrate determines voice **pitch**; how they move determines voice **quality**
- These vary *across* speakers (people's voices sound different) and *within* speakers (people can adjust vibration)

Ladefoged gif: http://www.linguistics.ucla.edu/faciliti/demos/vocalfolds/vocalfolds.htm

Some voices by Laver

- Laver modal voice
  - (modal: typical quality)
- Laver breathy voice
  - (breathy: with noise)
- Laver creaky voice
  - (creaky: with constriction and tension)

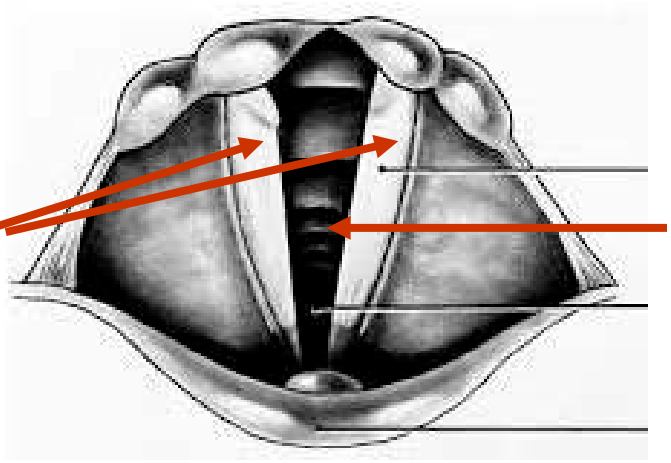
Laver (1980) *The phonetic description of voice quality*

Phonation types and glottal opening

How large is the glottal opening?

back

2 vocal folds



glottis (space between)

front

http://pixgood.com/glottis-and-epiglottis.html

A continuum of glottal openings

- from most open (spread as for voiceless sounds)
- to most closed (glottal stop)
- many intermediate openings that allow voicing, of different qualities

Ladefoged's continuum

	Most open ←				→ Most closed
Phonation type	Voiceless	Breathy	Modal	Creaky	Glottal closure

On the breathy side of modal: lax, slack, or lenis
On the creaky side of modal: tense, stiff, fortis, or pressed
(together with creaky: "laryngealized")

Ladefoged (1971) *Preliminaries to linguistic phonetics*

Phonation contrasts in languages of the world

- Many of the world's languages use phonation contrastively on vowels and/or consonants
- Common especially in SE Asia, India, the Americas

UCLA Linguistic Voice Quality project

- How do phonation types (on vowels) differ within and across languages?
- This talk:
  - Cross-language comparison of vowel phonation acoustics: What is the overall phonetic space for vowel voice quality?
  - Phonation in tone languages: How do pitch and phonation interact?

Our project: 10 languages
from four language families

Sino-Tibetan

- ***Yi** (Southern: Xinping & Jiangcheng)
  - lax vs. tense
  - crossed with L, M lexical tones
- ***Bo** (Shizong & Xingfucun) – like Yi
- ***Hani** (Luchun) – like Yi
- ***Mandarin** (Beijing) – creaky tone3

Indo-European

- ***Gujarati** (Standard Mumbai)
  - breathy vs. modal, no tones
- **English** (Californian)
  - no contrasts

*8 languages with electroglottography

Hmong-Mien

- ***Hmong** (White Hmong)
  - modal vs breathy H-falling tone, creaky L tone, others modal
- ***Black Miao** (Shidong Kou)
  - modal vs breathy M tone, creaky L tone, pressed H tone

Oto-Manguean

- **Mazatec** (Jalapa de Diaz)
  - breathy vs. modal vs. laryngealized (creaky)
  - fully crossed with lexical tones
- ***Valley Zapotec** (Santiago Matatlán and San Juan Guelavia combined)
  - Modal H tone, creaky H-falling tone, breathy L-falling tone

Keating et al. 2011

Black Miao fieldwork in Guizhou (Jianjing Kuang)

Kuang 2013

Yi-languages fieldwork in Yunnan (Jianjing Kuang)

Kuang 2011, 2013

Hmong fieldwork in Minnesota

(Christina Esposito)

Esposito 2012

In Los Angeles

- Mandarin and English students at UCLA
- Gujarati students at USC
- Zapotec speakers in Koreatown

- Mazatec recordings from online UCLA Phonetic Archive

Keating & Kuo 2012, Khan 2012, Garellek & Keating 2011

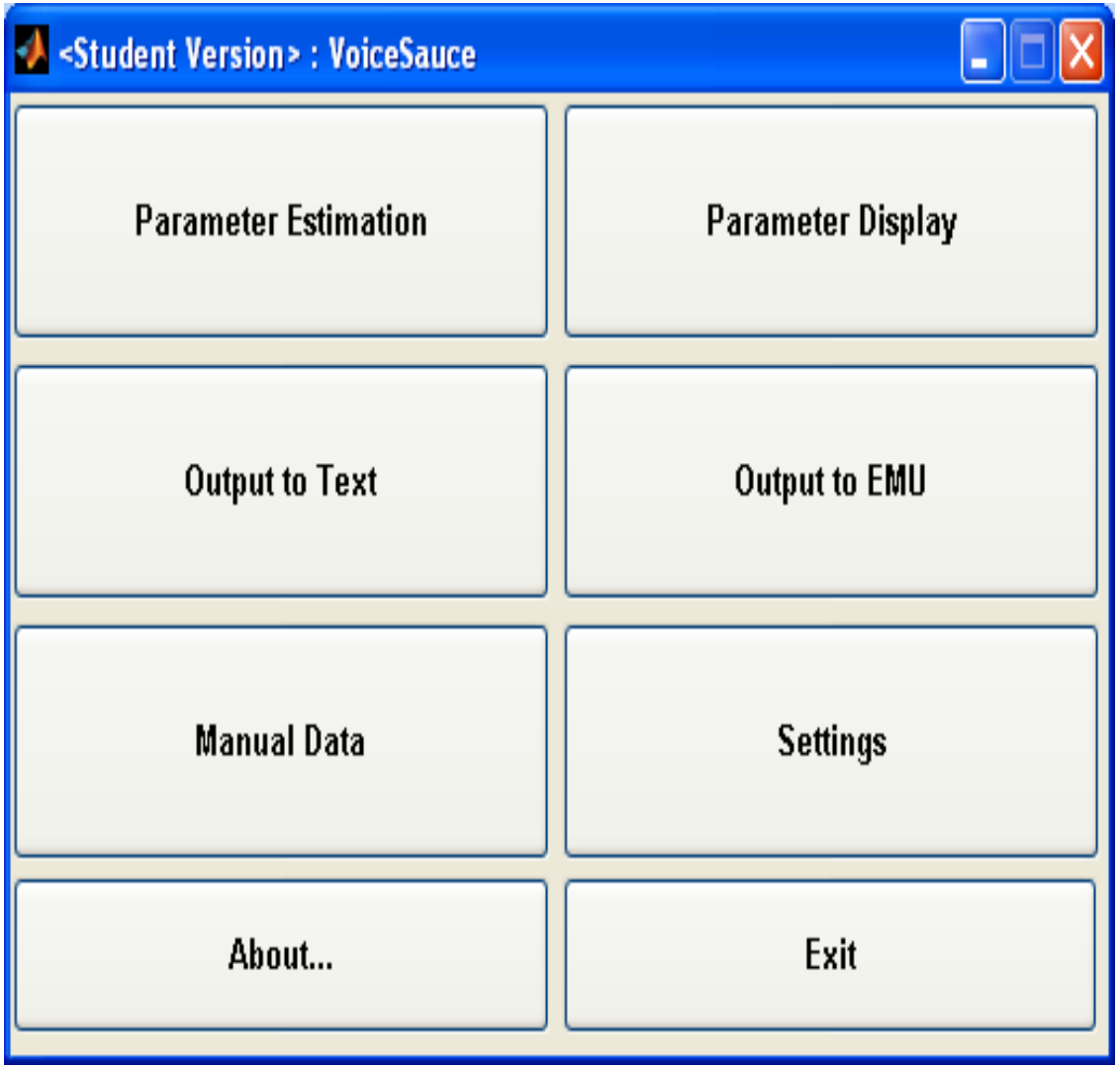
Sample tokens 4 languages

(1 female speaker each language)

	Breathy	Lax	Modal	Tense	Creaky
Gujarati	ba̤ɾ	--	baɾ	--	--
Hmong	pɔ̤^{42}	--	pɔ^{52} pɔ^{22}	--	pɔ̰^{21}
Mazatec	ba̤^{34}	--	ba^{32}	--	ba̰^{3}
S. Yi	--	be^{33}	--	be̱^{33}	--

UCLA analysis tools

- **EggWorks** – for PC
- **VoiceSauce** - for Matlab, or PC
- Free by downloading



Shue 2010, Shue et al. 2011

Acoustic space for phonation across languages

Many acoustic measures of 24 phonation categories from 10 languages, all speakers

24 categories (for non-high, oral vowels after unaspirated consonants, at mid-vowel)

- **Bo**
  - **Lax, Tense**
- **English**
  - **Modal**
- **Gujarati**
  - **Breathy, Modal**
- **Hani (Luchun)**
  - **Lax, Tense**
- **Hmong**
  - **Breathy, Modal, Creaky**
- **Mandarin**
  - **Modal, Creaky**
- **Mazatec**
  - **Breathy, Modal, Creaky**
- **Miao (Black)**
  - **Breathy, Modal, Tense, Creaky**
- **Yi (Southern)**
  - **Lax, Tense**
- **Zapotec (Valley)**
  - **Breathy, Modal, Creaky**

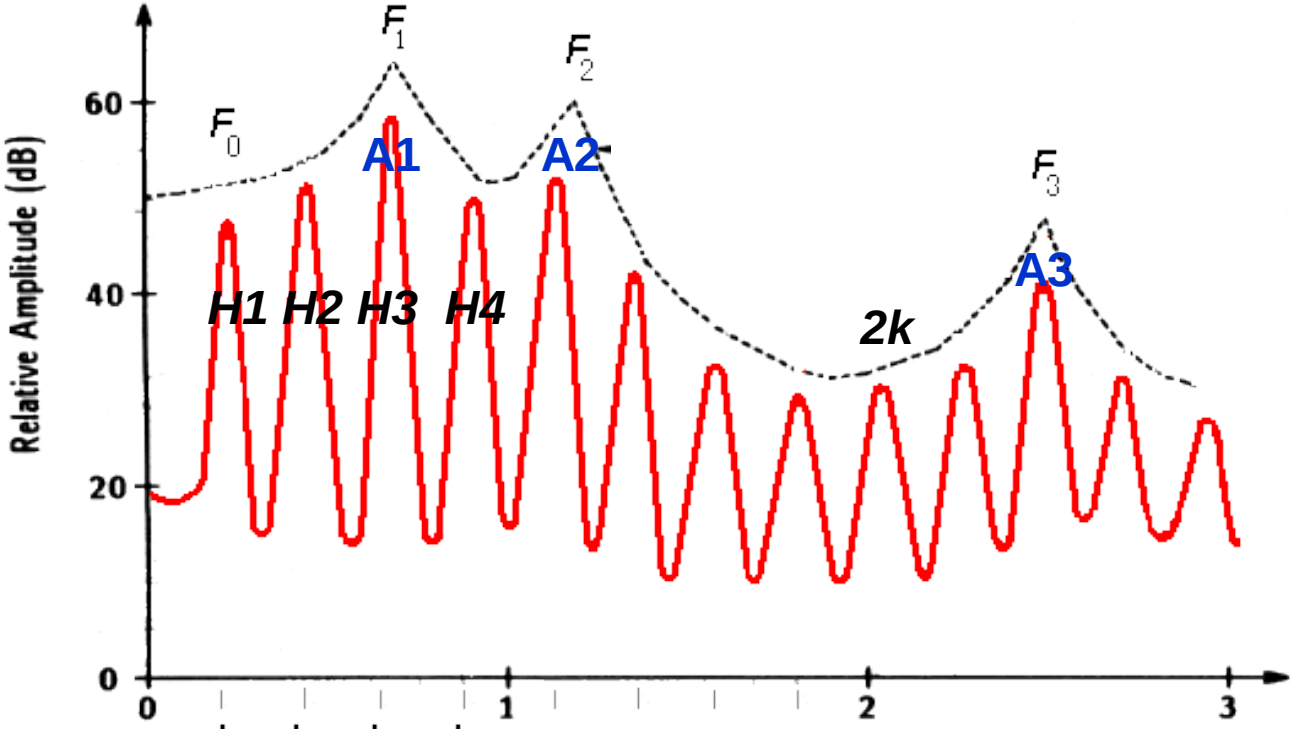
VoiceSauce measures, and those used here

- **F0 from STRAIGHT**, Snack, or Praat
- **H1*, H2*, H4***
- H2kHz*, H5kHz
- F1-F4 and B1-B4 from Snack or Praat
- **A1*, A2*, A3***

- All * harmonic measures come both corrected (*) and uncorrected for formants

- **H1*-H2***
- **H1*-A1***
- **H1*-A2***
- **H1*-A3***
- **H2*-H4***
- H4*-H2k*, H2k*-H5k
- Energy
- Subharmonic to Harmonic Ratio
- **Cepstral Peak Promin.**
- **Harmonic to Noise Ratios (4 freq. bands)**

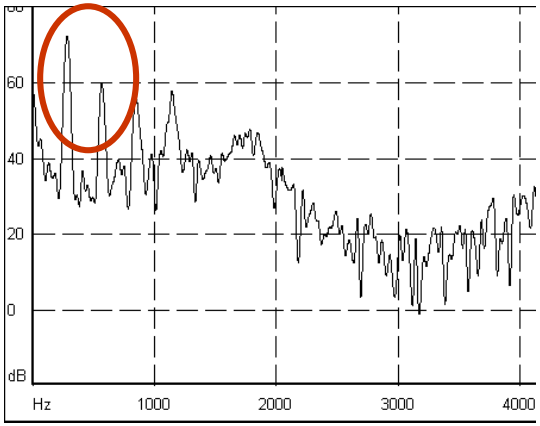
Harmonics in spectrum



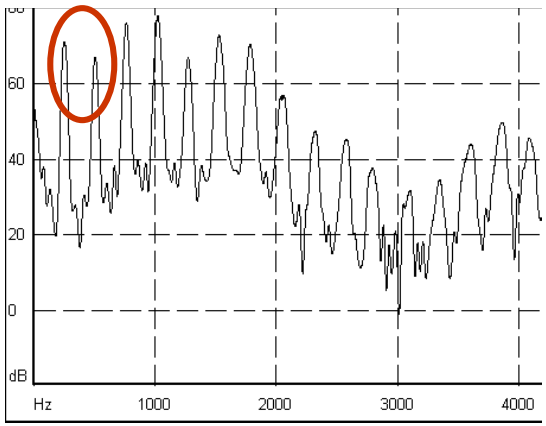
Kreiman et al. 2007, Kreiman et al 2014

H1-H2: Mazatec example

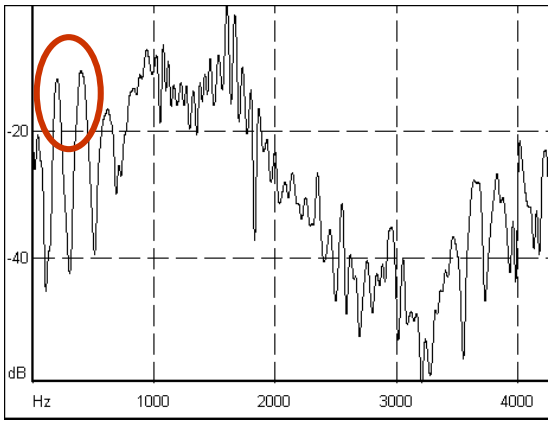
breathy



modal



creaky



Breathy	Modal	Creaky
ba̤^{34}	ba^{32}	ba̰^{3}

Kirk et al. 1984, Garellek & Keating 2011

Category means

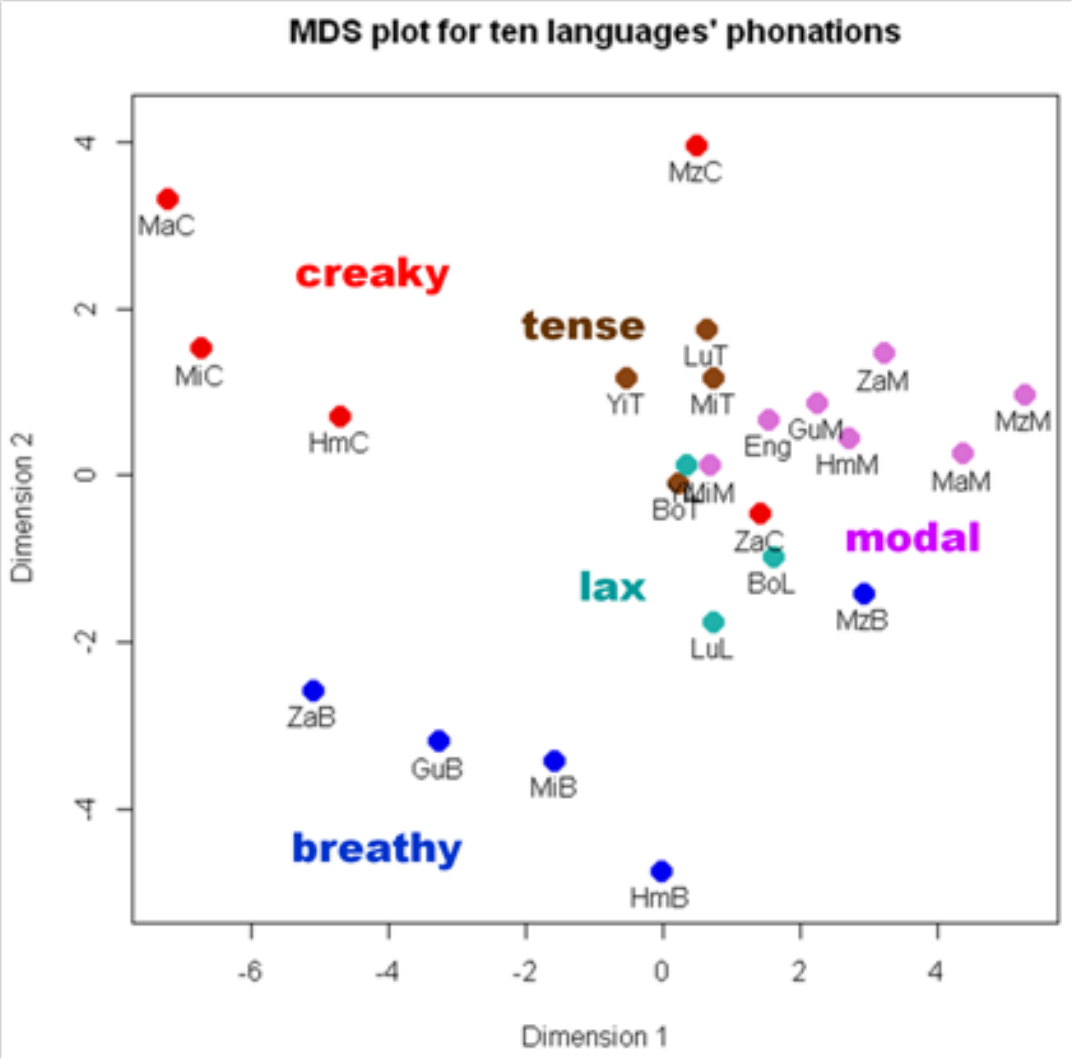
- Each of 17 measures standardized for each speaker
- Mean for each measure for each of 24 phonation categories – across all tokens and speakers
- = 17 x 24 mean measures

Multi-Dimensional Scaling

- MDS is a reduction of high-dimensional data to a low-dimension map of distances that can be visualized.
- Usually used with ***perception*** data, but here applied to ***acoustic*** data. Each acoustic measure is a dimension, and each category mean has a multi-dimensional physical acoustic distance from all other category means.
- Can test for strength of contribution of measures to dimensions.

2-D acoustic space from MDS

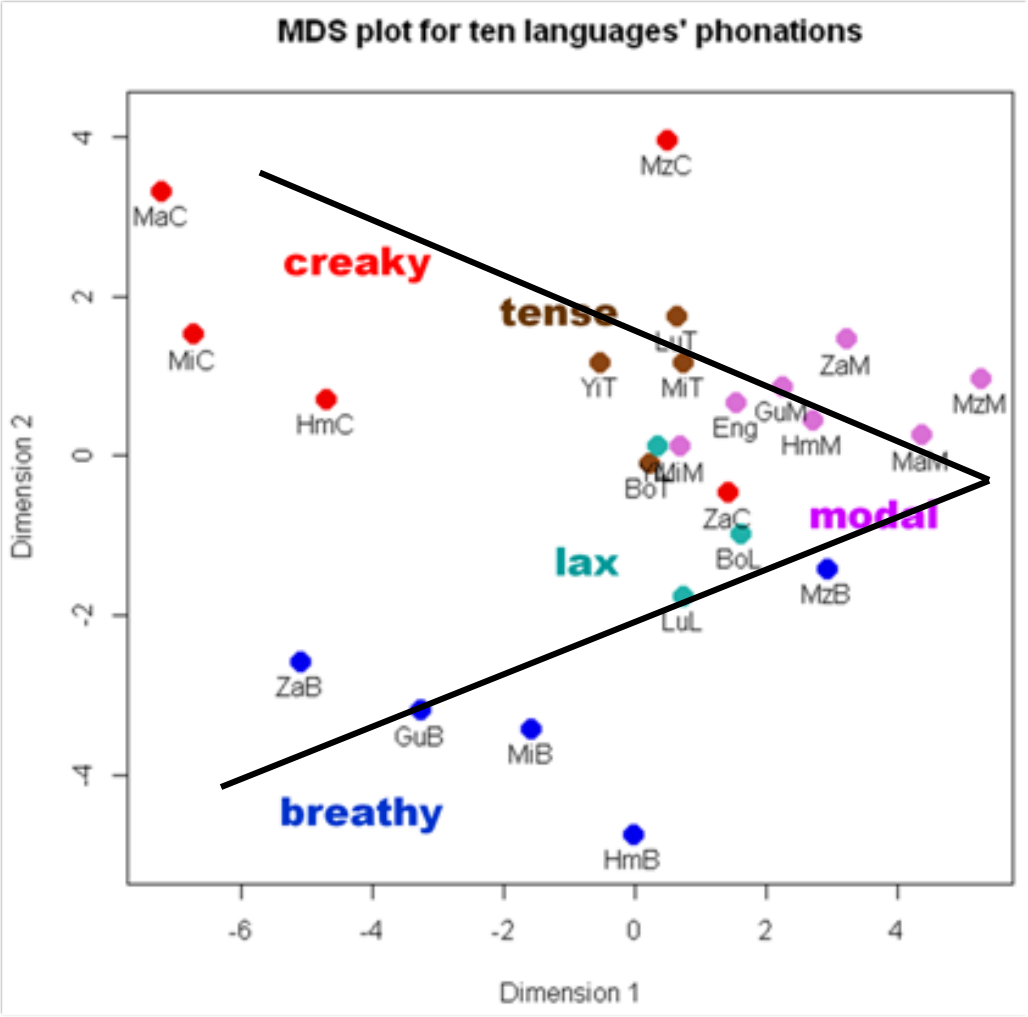
Bo
English
Gujarati
Luchun Hani
Hmong
Mandarin
Mazatec
Miao (Black)
Yi (Southern)
Zapotec (Valley)



Keating et al. 2012

2-D acoustic space from MDS

Bo
English
Gujarati
Luchun Hani
Hmong
Mandarin
Mazatec
Miao (Black)
Yi (Southern)
Zapotec (Valley)

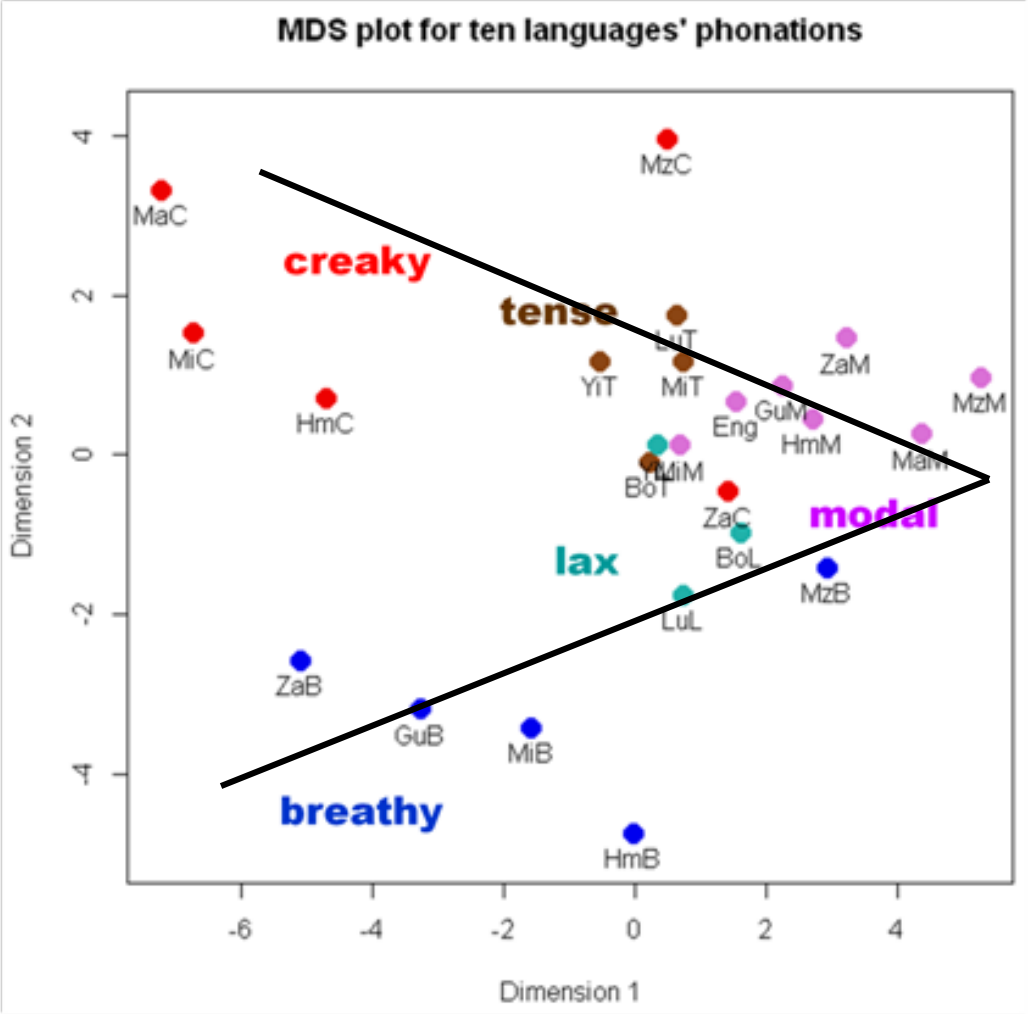


Keating et al. 2012

2-D acoustic space from MDS

Bo
English
Gujarati
Luchun Hani
Hmong
Mandarin
Mazatec
Miao (Black)
Yi (Southern)
Zapotec (Valley)

~ H1*, H1*-H2*, H1*-A1*



~ H1*-A1*, A3*, H2*

Keating et al. 2012

Local summary

- The acoustic-phonetic space for (vowel) voice quality across languages is largely 2-D: modal-ness vs. glottal aperture
- Each phonation type tends to occupy one area of the space, in a V-shaped array
- But languages do differ in exactly how they use the space for contrasts

Pitch and phonation in tone languages

- Pure phonation contrast
- Correlated pitch and phonation
- Mixed system

Relation of phonation to lexical tone in languages

- Some languages with phonation contrasts do not have lexical tone (pitch) contrasts
- Some languages have both, cross-classifying: different tones and phonations co-occur in all possible combinations
- Some languages use phonation as part of the tonal system: certain tones have their own correlated phonations
- Mixed tone systems combine contrast and correlation

Thongkum 1988, Silverman 1997, Thurgood 2015

Non-tonal example: Gujarati modal vs. breathy voice

Orthography	*Dictionary transcription*	IPA	Gloss
કાન	*kan*	kan	ear
કૃષ્ણ	*(not listed)*	kä̤n	Krishna
બાર	*bar*	bar	twelve
બહાર	*bə.har*	bä̤r	outside
બાણ	*baɳ*	baɳ	arrow
બહાનું	*bə.ha.nũ*	bä̤nũ	excuse
માલિક	*ma.lik*	malik	boss, god
મહારાજ	*mə.ha.raɟ*	mä̤raɟ	priest, emperor

Khan 2012

Cross-classifying example: Ladefoged's Mpi set

TONE (PITCH)	REGULAR VOICE	ENGLISH	TENSE VOICE	ENGLISH
Low rising	**si**	'to be putrid'	**si**	'to be dried up'
Low level	**si**	'blood'	**si**	'seven'
Mid rising	**si**	'to roll rope'	**si**	'to smoke'
Mid level	**si**	(a color)	**si**	(classifier)
High falling	**si**	'to die'	**si**	(name)
High level	**si**	'four'	**si**	(name)

http://www.phonetics.ucla.edu/vowels/chapter12/mpi.html

Cross-classifying example: Mazatec

	Creaky	Modal	Breathy
Low	F4 F6	F4	F4 F6
Mid	F4 F6	F6	F6
High	F4	F4	F6

Garellek & Keating 2011

Correlated example: Mandarin Chinese

- Female speaker
- Minimal tone set:
  - Tone 1: High 師
  - Tone 2: Rising 十
  - Tone 3: Low 使 (creaky at the end)
  - Tone 4: Falling 示 (creaky at the end)
- 3 times each

Keating & Kuo 2012

Mixed system:
Santa Ana del Valle Zapotec

- Modal High and Rising tones
- Breathy and creaky Falling tones
- Triple with Modal-High:
  - Modal: ‘can’ **lat**
  - Breathy: ‘place’ **la̤t**
  - Creaky: ‘field’ **la̰ts**

Esposito 2010

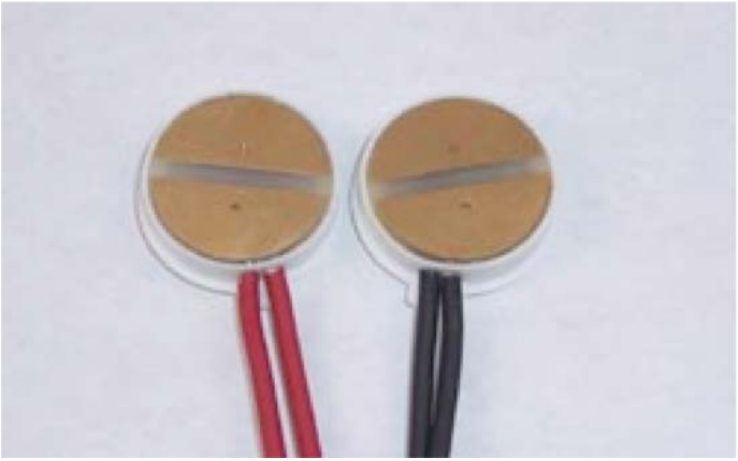
Yi languages: cross-classifying *tense* vs *lax*

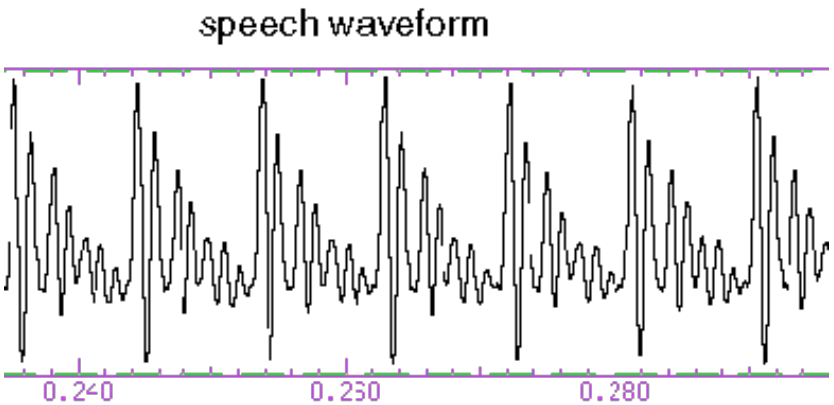
	Low tone	Mid tone
Lax phonation	be ^{21} (*mountain*)	be ^{33} (*fight*)
Tense phonation	be̲ ^{21} (*foot*)	be̲ ^{33} (*shoot*)

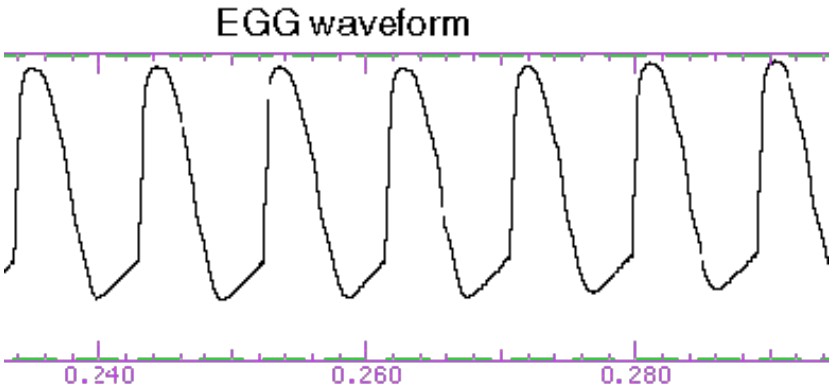
Example from Southern Yi

Kuang 2011, Kuang & Keating 2014

Electroglottography (EGG)







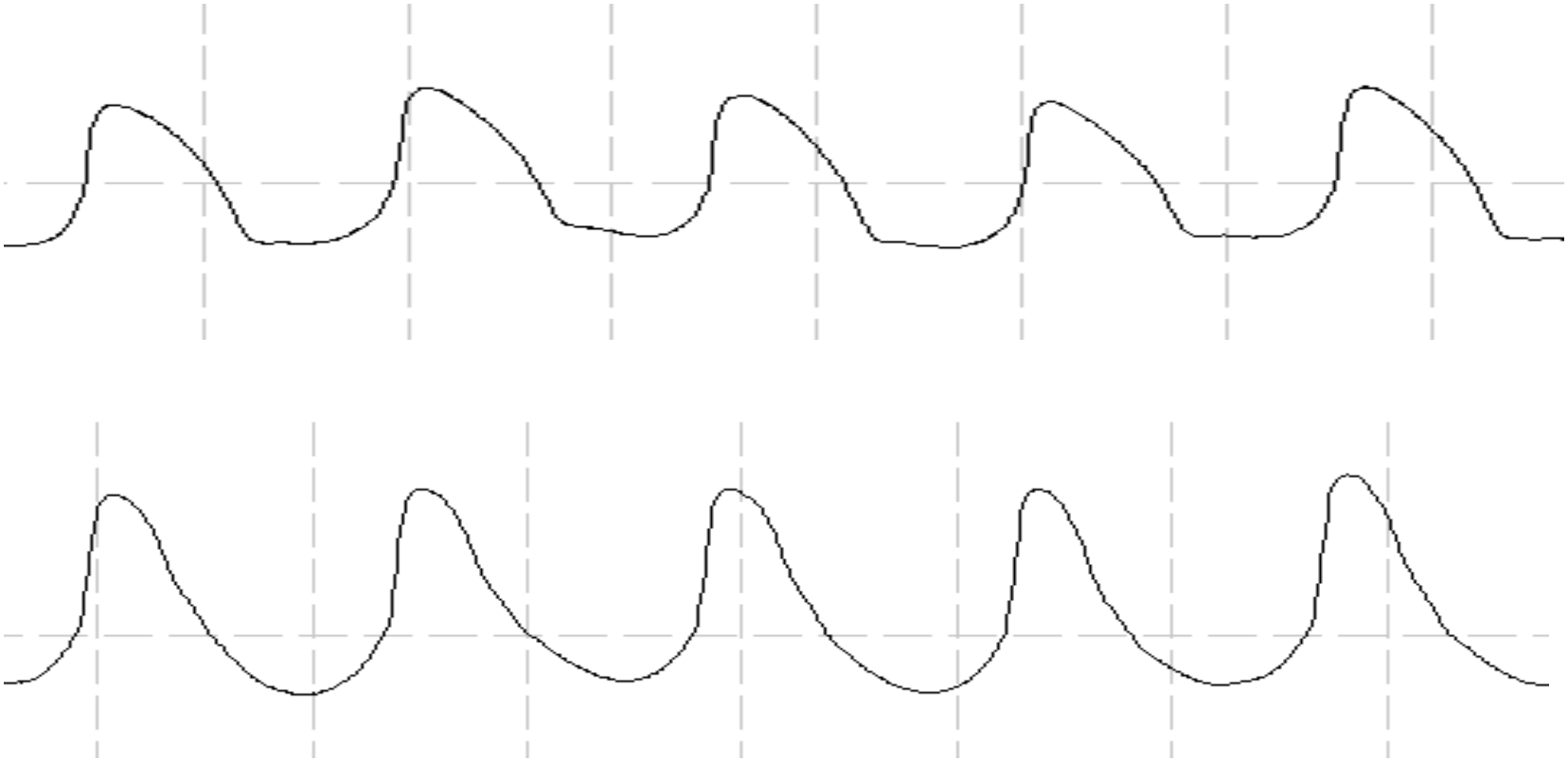
more

less

contact

Fabre 1957, Fourcin 1974; Esling 1984

Sample Yi EGG cycles: tense (top) and lax (bottom)



Kuang & Keating 2014

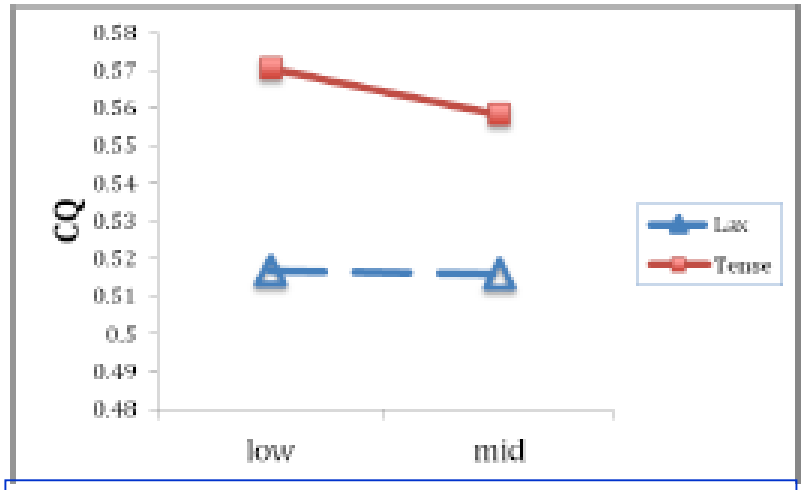
EGG measure: Contact Quotient (CQ)

- A measure of relative (proportional) amount of greater vs lesser vocal fold contact
- High CQ = overall more constriction (should be higher in tense voice)

Rothenberg & Mahshie 1988, Herbst & Ternstrom 2006

Contact Quotient in Yi

No tone effect on CQ

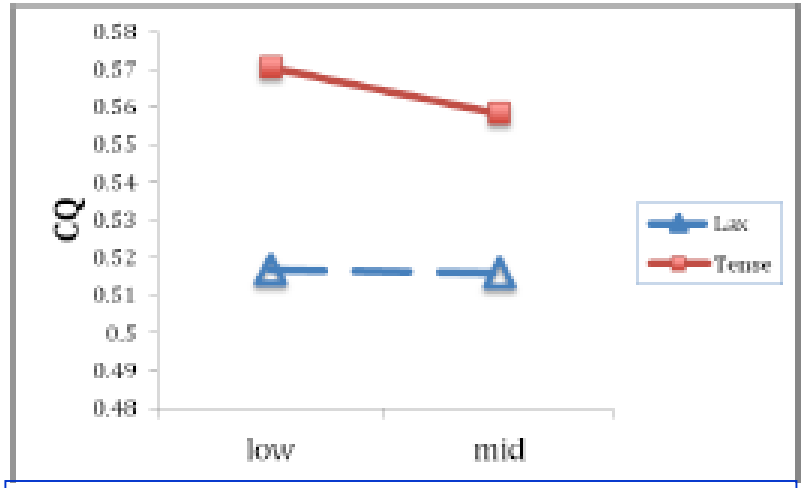


CQ is greater for tense (red) than for lax (blue) phonation, as expected, but tones have same CQ

Kuang & Keating 2012, Kuang 2013

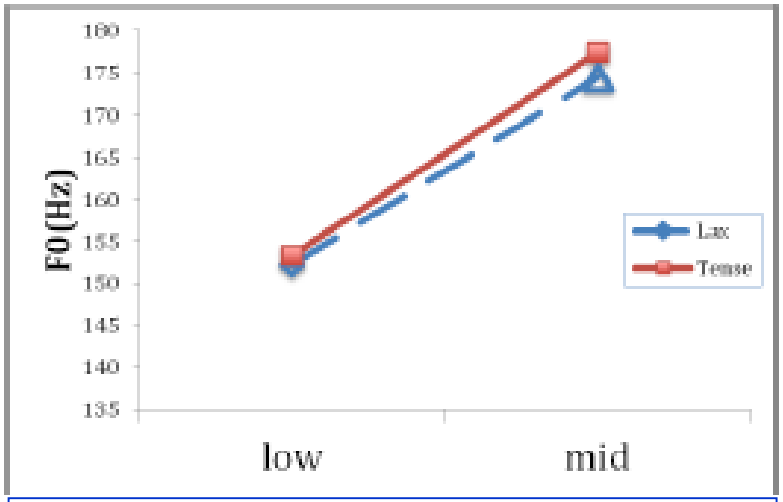
Independence of pitch and phonation in Yi

No tone effect on CQ



CQ is greater for tense (red) than for lax (blue) phonation, as expected, but tones have same CQ

No phonation effect on F0



F0 is greater for mid (right) than for low (left) tone, as expected, but phonations have same F0

Kuang & Keating 2012, Kuang 2013

Mandarin: creaky voice, tones, pitch

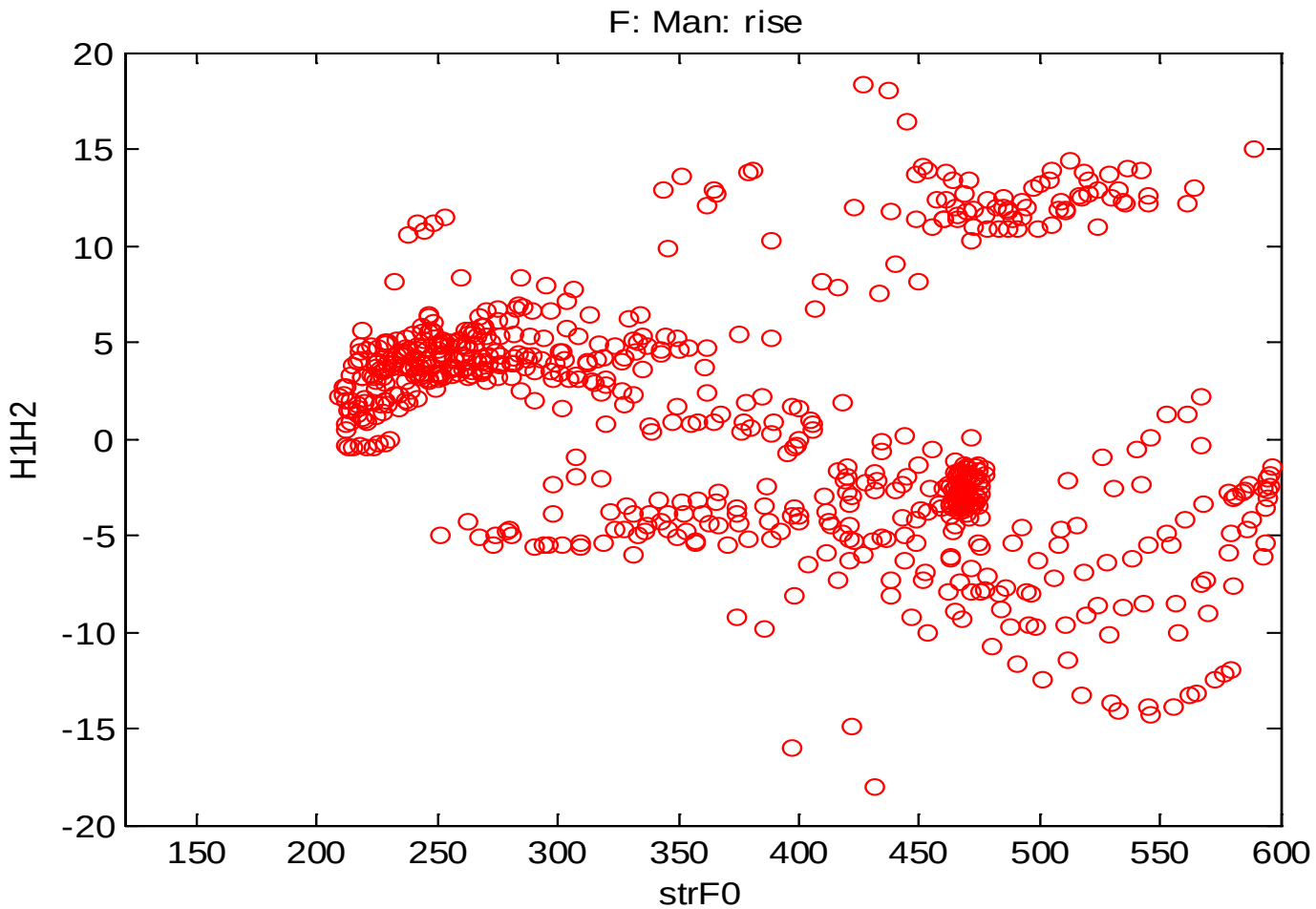
- It seems that in many languages, the lowest-pitch tone(s) can be produced with creaky voice, or at least constriction
- More generally, *voice quality will tend to co-vary with voice pitch* – the vocal folds adjust to reach lower vs higher pitches, and those adjustments affect phonation

Voice quality in rising and falling pitch glides

- Starting from a self-chosen comfortable pitch, speakers let their pitch sweep up or down
- For falls, they either let their voices creak, or did not (2 separate conditions)
- Speakers included 23 Mandarin-speaking UCLA students
- Acoustic measures of voice quality ~ F0

Keating & Kuo 2012

H1*-H2* ~ F0, rising pitch, 10 Mandarin women

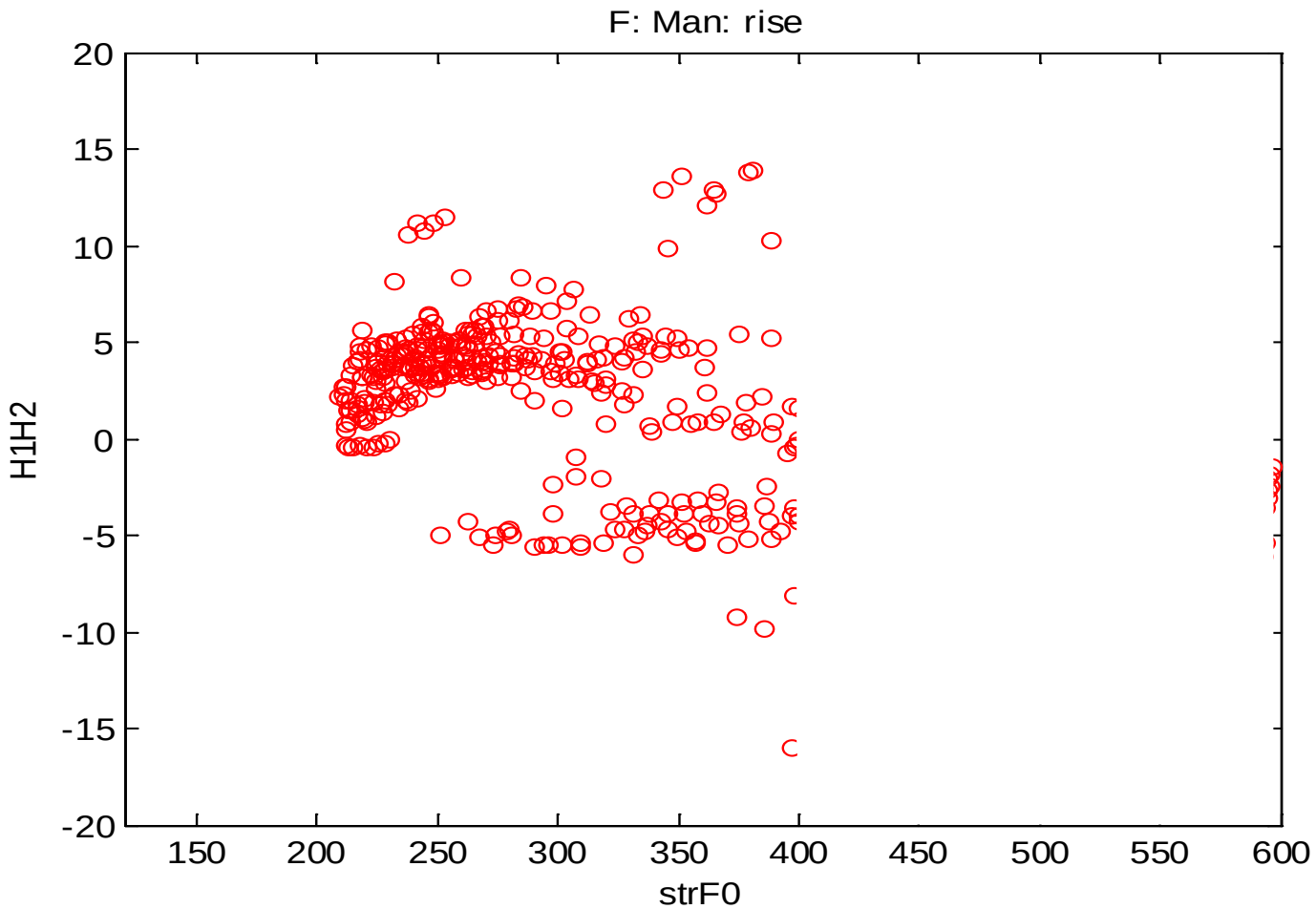


laxer

tenser

Keating & Shue 2009

H1*-H2* ~ F0, rising pitch, 10 Mandarin women



laxer

tenser

Combining rises and falls

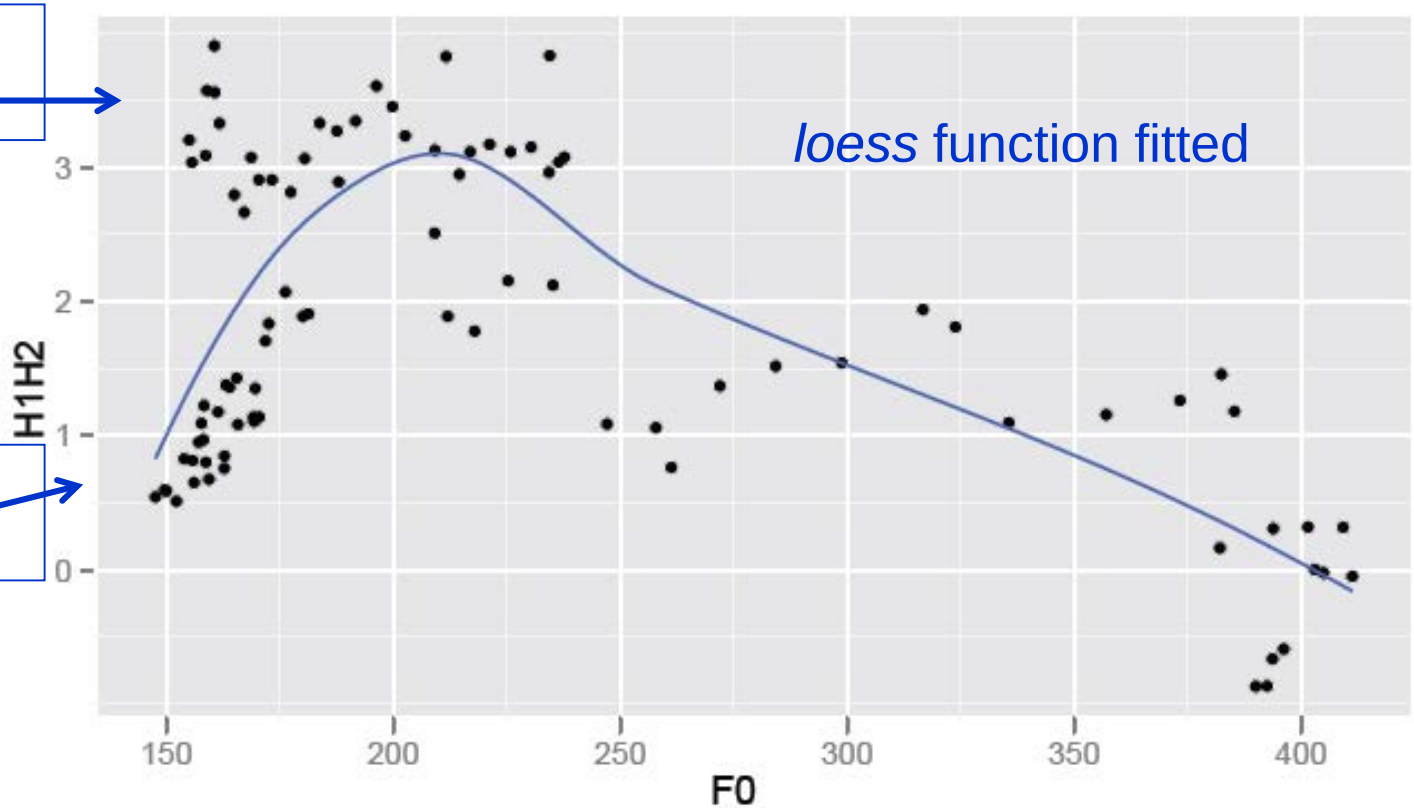
- The 3 glide types were kept separate: rise, fall with creak, fall without creak; speakers were again divided by language and sex
- Within each glide type, mean H1*-H2* was calculated across speakers (here, Mandarin women); average values within 30 time-intervals are plotted and fitted with a *loess* smoother

H1*-H2* ~ F0 (Mand. females)

Speakers averaged for each time interval within each glide type

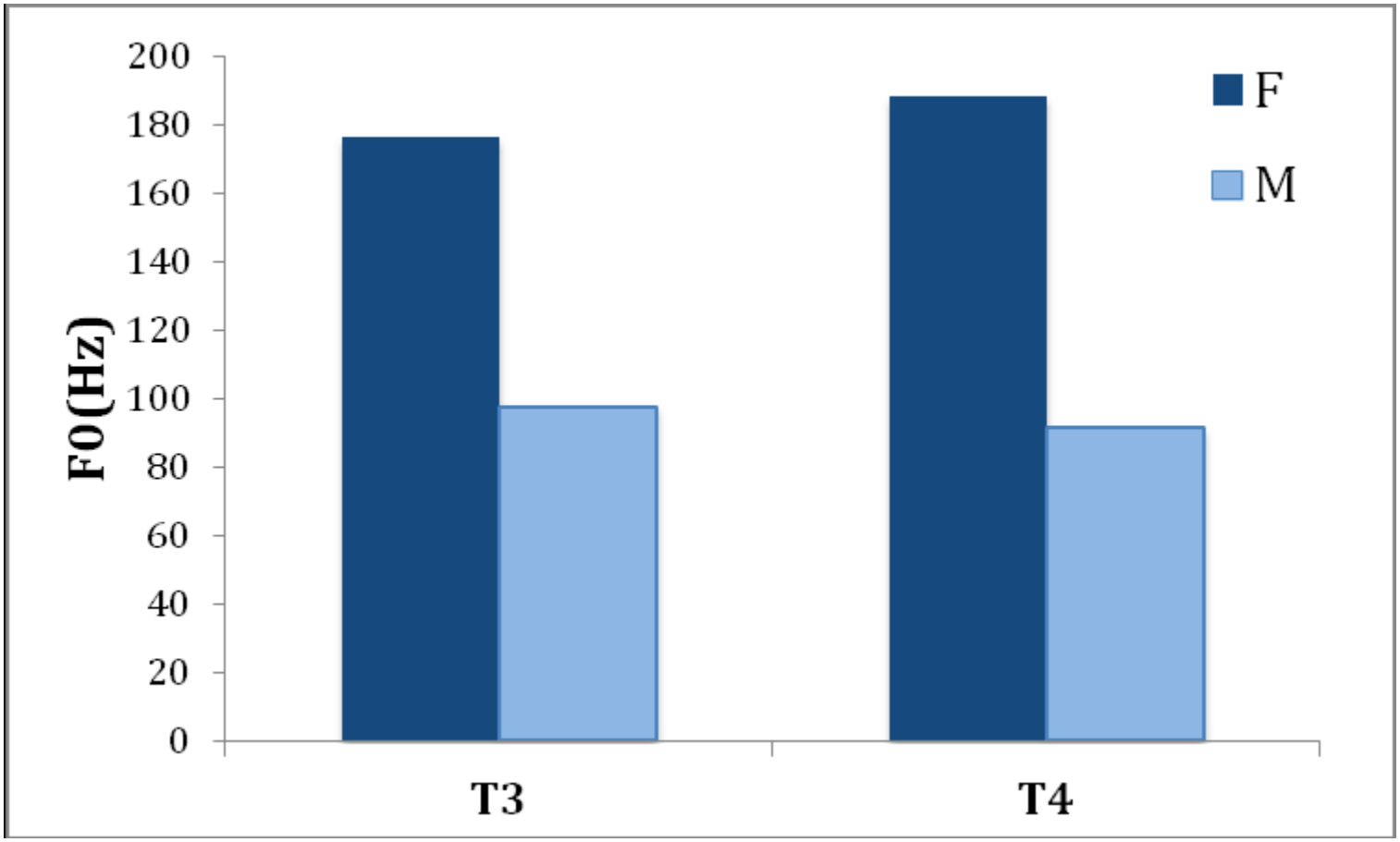
Low F0 is breathier

Low F0 is creakier



Kuang 2013, 2015

Mandarin creaky tones:
F0 at break into creak



Kuang 2013

White Hmong tones

Tone	Orthographic tone symbol	Example (IPA)	Example in Hmong orthography with English meaning
High-rising (45)	-b	[pɔ˧˥]	*pob* "ball"
Mid (33)	∅	[pɔ˧]	*po* "spleen"
Low (22)	-s	[pɔ˨]	*pos* "thorn"
High-falling (52)	-j	[pɔ˥˩]	*poj* "female"
Mid-rising (24)	-v	[pɔ˧˦]	*pov* "to throw"
Low-falling creaky (21)	-m	[pɔ̰˨˩]	*pom* "to see"
Mid-to high-falling breathy (52 or 42)	-g	[pɔ˥˩ˀ]	*pog* "grandmother"

Ratliff 1992, Esposito 2012

Perception experiment

- White Hmong minimal set
- Breathy-, creaky-, and modal-tone tokens had their F0 and duration modified by PSOLA re-synthesis
  - originally-breathy words now shortened and/or with lowered/falling F0
  - originally-creaky words now lengthened and/or with raised F0
  - originally-modal words now with varying duration and/or F0
- Original phonation was never modified

Garellek et al. 2013

Stimulus examples

Natural and manipulated tokens of:

- original breathy
- original modal
- original creaky

15 White Hmong listeners identified words

Results

- They heard breathy-tone words only for stimuli made from an original breathy token;

  F0 did not matter

- In contrast, they heard more creaky-tone words when stimulus F0 was low-falling and duration was short – even if originally modal;

  phonation did not matter

So, 2 different outcomes

- Breathy tone is heard when the stimulus is breathy, regardless of F0:

  phonation is criterial for the breathy tone

- “Creaky” tone is heard when the stimulus is low-pitched/short, regardless of modal/creaky phonation:

  phonation is NOT criterial for the “creaky” tone (it’s primarily a pitch contrast)

Comparison of tones

	Breathy	Modal	Creaky
Hmong	pɔ̤^{42}	pɔ^{52} pɔ^{22}	pɔ̰^{21}

phonation contrast

pitch (and duration) contrast

Local summary

- Some languages have phonetically independent tone and phonation contrasts (Mpi; Yi languages)
- Some languages use phonations to help expand the pitch range for tones (Mandarin; also Black Miao)
- When phonation is correlated with pitch, phonation may not be perceptually important (White Hmong)

Conclusions

- New tools for analysis of voice quality make large-scale phonetic descriptions possible
- Phonation contrasts are organized in a 2-D acoustic space, overall similarly across languages, but with language-specific differences
- Phonation can be correlated with pitch in tone- and non-tone languages, or independent of pitch in tone languages
- Languages can differ in their relative weighting of pitch and phonation in tone perception

Acknowledgments

- NSF grant BCS-0720304, and co-PIs Abeer Alwan, Jody Kreiman
- Y.-L. Shue for VoiceSauce
- H. Tehrani for EggWorks

Thanks to my collaborators
on these and other voice projects

Jody Kreiman
UCLA

Abeer Alwan
UCLA

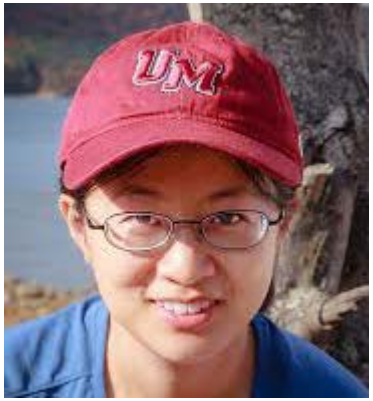
Christina Esposito
Macalester College

Jason Bishop
CUNY/CSI

Sameer Khan
Reed College

Jianjing Kuang
U Penn

Kristine Yu
U Mass



Marc Garellek
UCSD