Linguistic voice quality

Pat Keating

UCLA Linguistics Department

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 (individuals can adjust vibration)

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

Some examples by John Laver - 3 major phonation types

- Laver modal voice
- Laver breathy voice
- Laver creaky voice

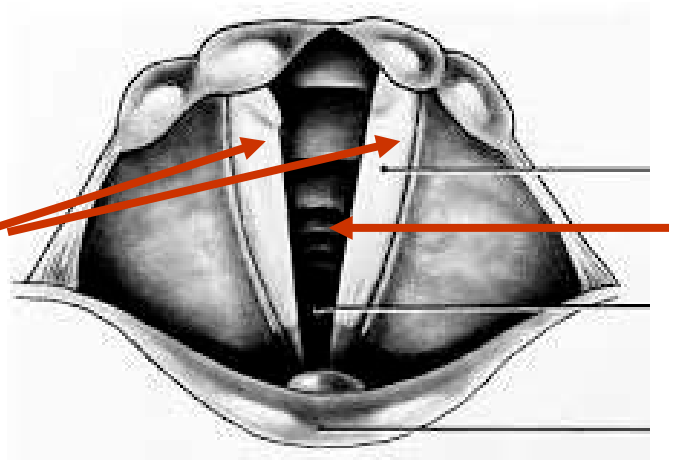
Cassette with Laver 1980, *The Phonetic Description of Voice Quality*

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

Ladefoged’s glottal continuum

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

IPA diacritics: a̤ ã

On the breathy side of modal: lax, slack, or lenis
On the creaky side of modal: tense, stiff, fortis, or pressed

Ladefoged (1971) *Preliminaries to linguistic phonetics*

Phonation contrasts in languages of the world

- Many languages contrast phonations on vowels and/or consonants
- Common especially in SE Asia, the Americas, India

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	--	bə^{33}	--	bə̲^{33}	--

New tools for voice analysis

- For acoustic analysis: VoiceSauce
- For physiological analysis: EggWorks, used with VoiceSauce
- Both = UCLA free software

Shue 2010, Shue et al. 2011, Tehrani 2012

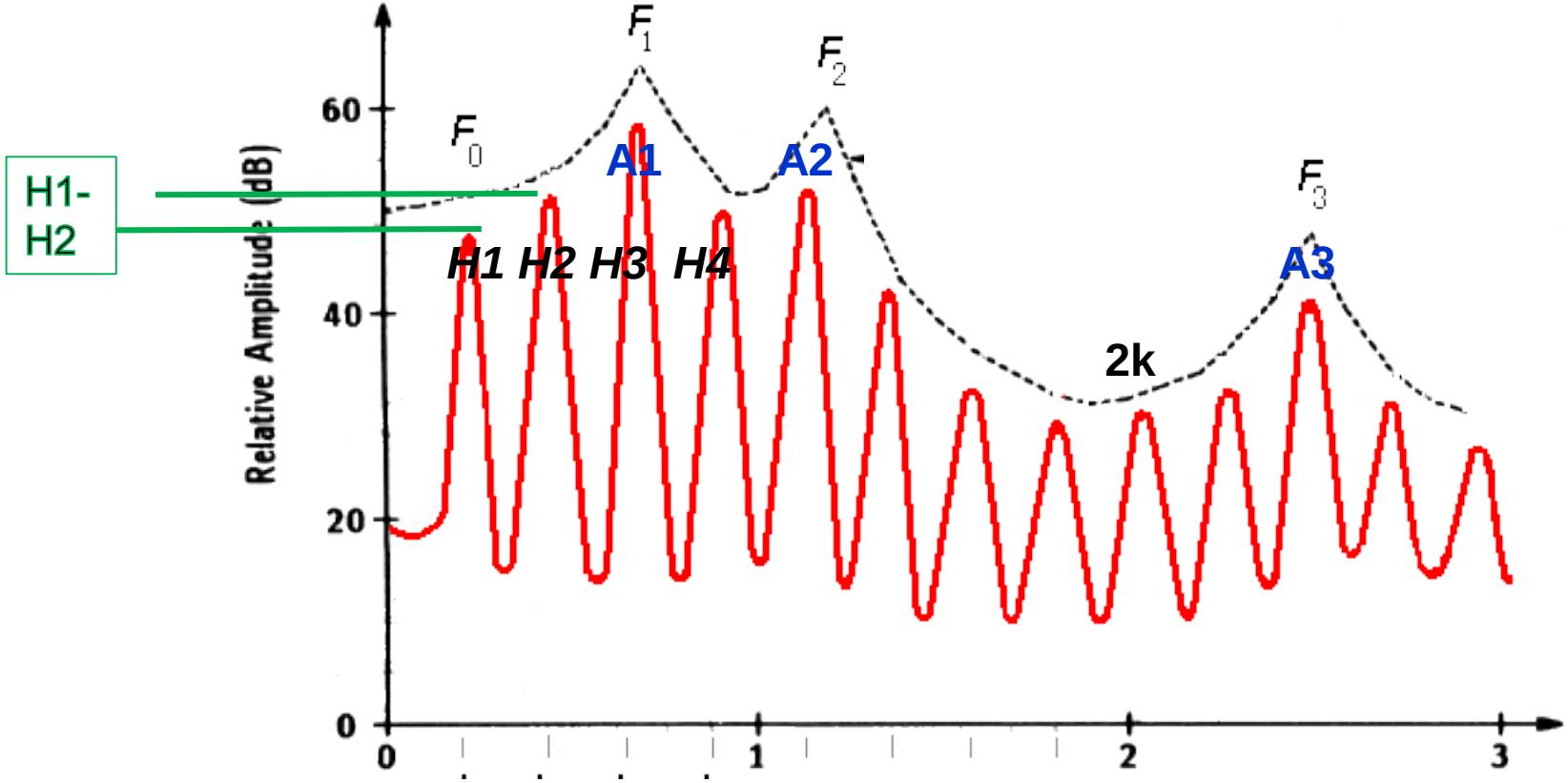
VoiceSauce measures, and those used here

- **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 Harm. Ratio
- **Cepstral Peak Prominence**
- **Harmonic to Noise Ratios (4 freq. bands)**
- Strength of Excitation

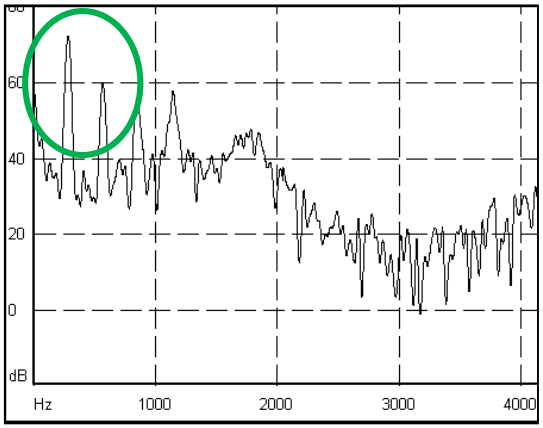
Acoustic measures based on harmonics in spectrum



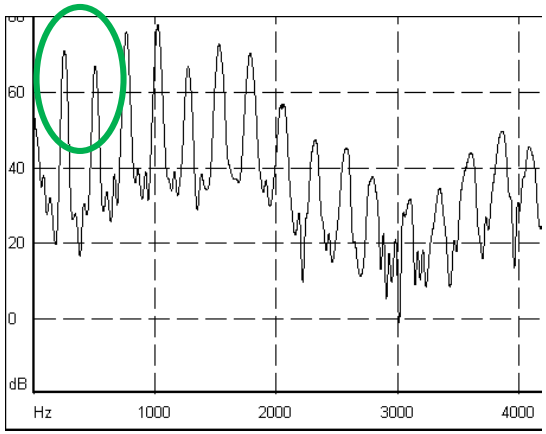
Shue 2010, Kreiman et al. 2007, Kreiman et al 2014

H1-H2 example: Jalapa Mazatec

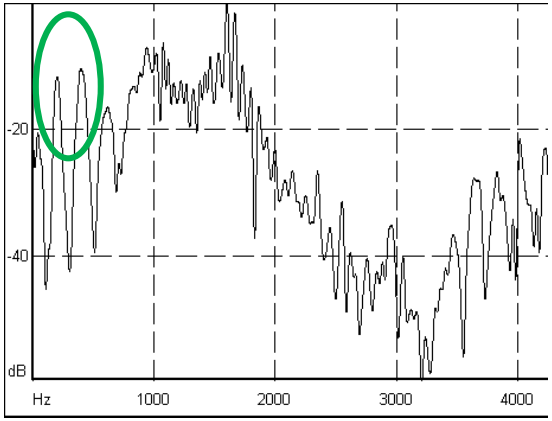
breathy



modal



creaky



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

Kirk et al. 1984, Garellek & Keating 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**

Category means

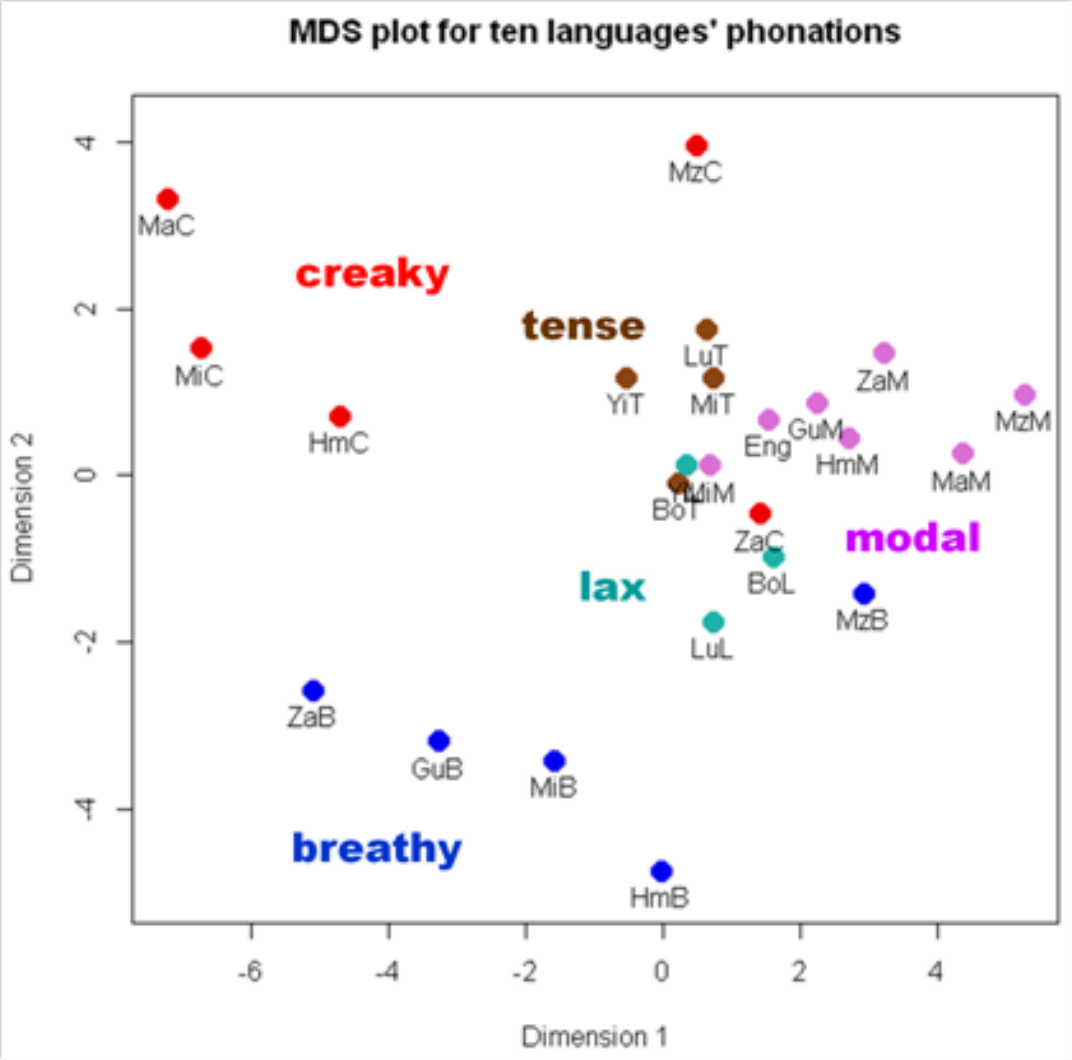
- Each of 17 measures standardized for each speaker
- Mean for each measure for each of 24 phonation categories – across all tokens and speakers
- = 17 x 24 mean measures

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

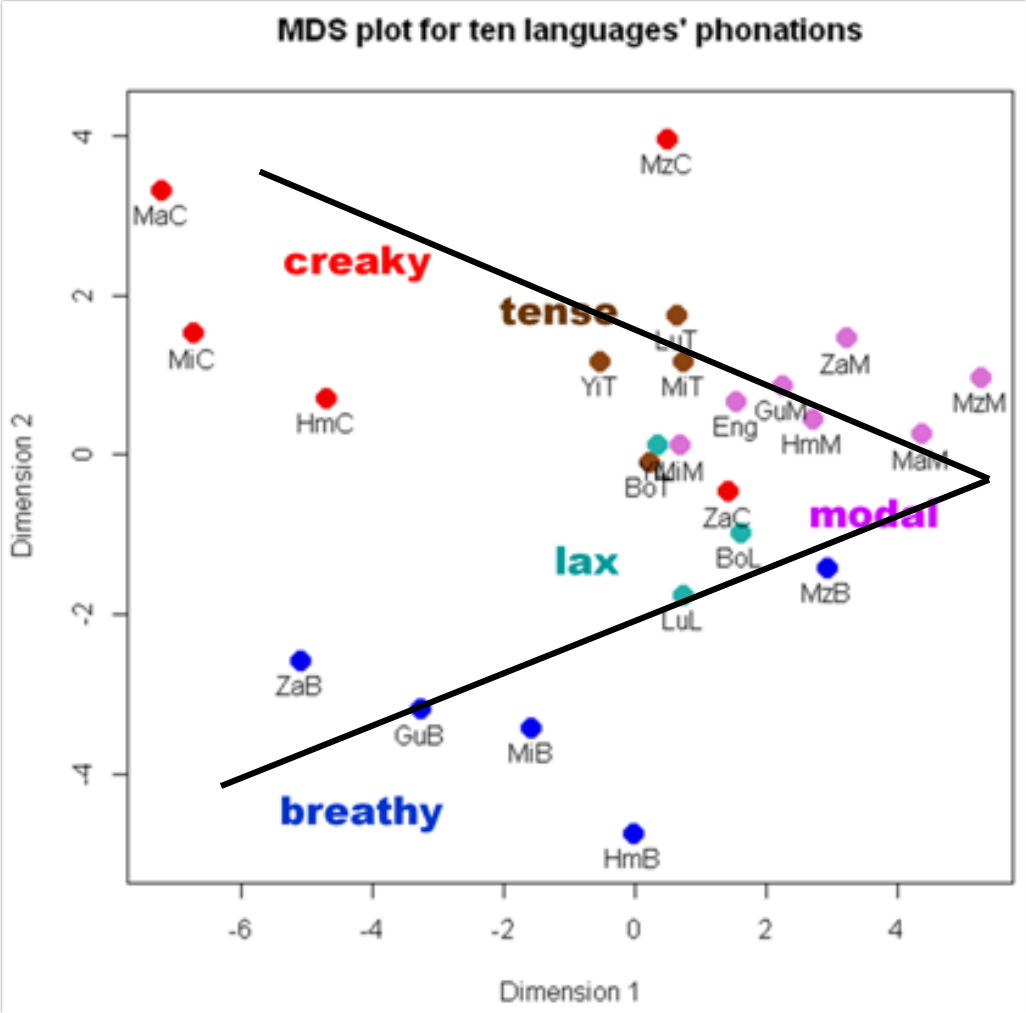
Languages:
Bo
English
Gujarati
Luchun Hani
Hmong
Mandarin
Mazatec
Miao (Black)
Yi (Southern)
Zapotec (Valley)



Keating et al. 2012

2-D acoustic space from MDS

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

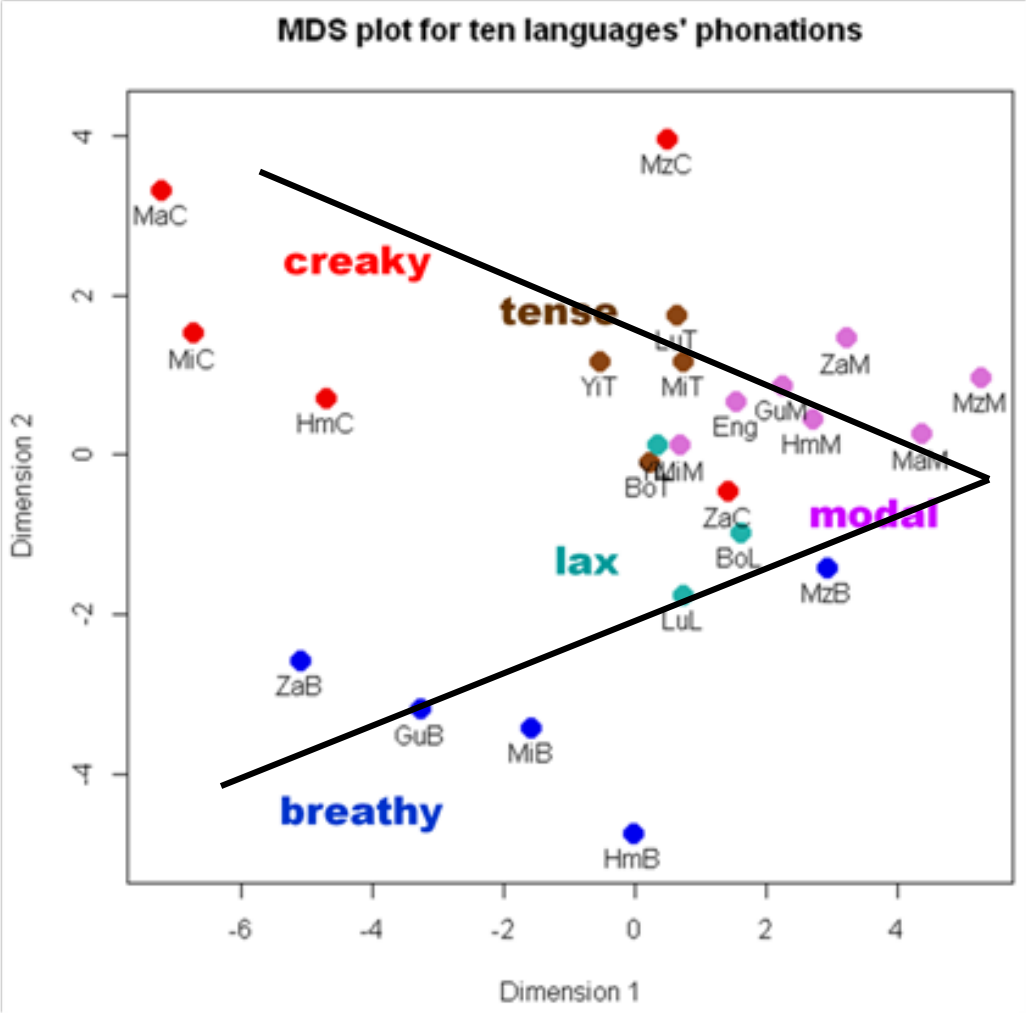


Keating et al. 2012

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

Summary, contrast space

- The acoustic-phonetic space for (vowel) voice quality contrasts is largely 2-D: modal-ness vs. glottal aperture
- Both derived from spectral measures (low and low-mid frequencies)
- 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: Mpi (plays by rows)

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)

James Harris & Peter Ladefoged: 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 creaky voice tone

- It seems that in many languages, the lowest-pitch tone can be produced with creaky voice, or at least laryngealization
- E.g. Mandarin Tone 3 – Kuang (2013) found that 12 speakers produced 60/60 tokens with creak (and 39/60 of Tone 4),
- See also Hockett, 1947; Chao, 1956; Davison, 1991; Belotel-Grenié & Grenié, 1994, 2004

Mandarin example

- 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 example:
Santa Ana del Valle Zapotec

- Modal High and Rising tones
- Breathy and creaky Falling tones
- Triple with Modal-High:
  - Modal: ‘can’ lat
  - Breathy: ‘place’ lat̤
  - Creaky: ‘field’ lat̰s

Esposito 2010

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

Summary, tone/phonation

- In tone languages, pitch and phonation can be independent or correlated, even within one language

- Question: When pitch and phonation are correlated, do listeners use both kinds of information in recognizing tones?

Correlated creaky voice can help in perceiving low tones

- Creaky Tone 3 speeds up judgment, but doesn’t affect accuracy, which is at ceiling
- Creak helps distinguish synthesized Tone 3 from Tone 2
- Cantonese: creaky stimuli perceived more often as low tone (T4)

Belotel-Grenié & Grenié 1997; Yang 211; Yu & Lam 2011

Hmong 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

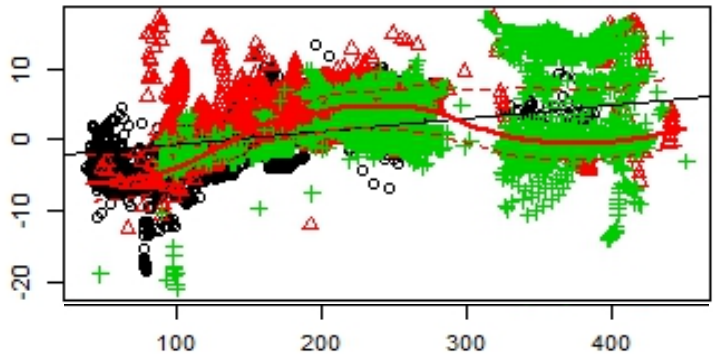
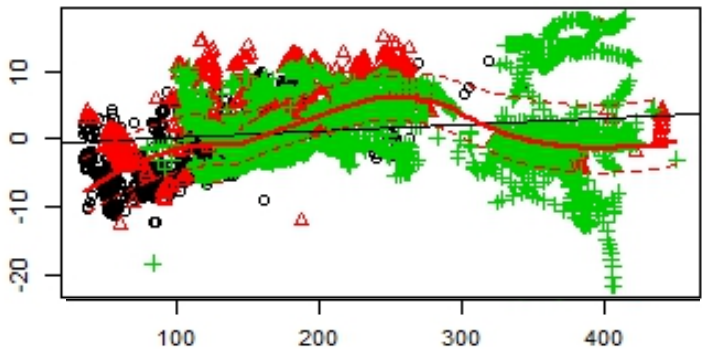
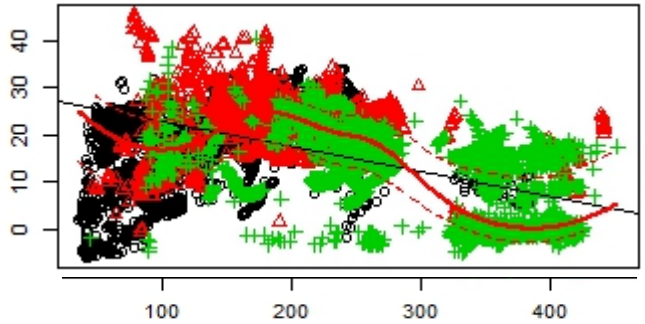
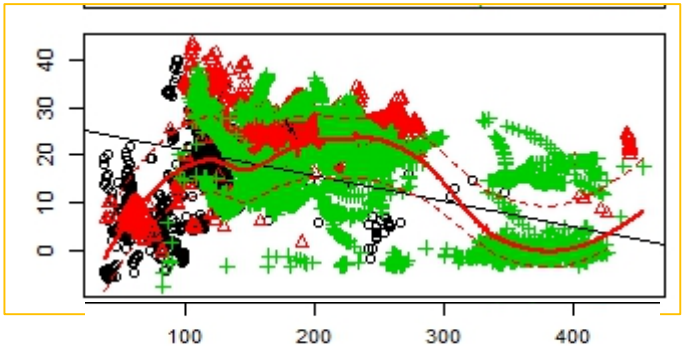
Voice quality in relation to voice pitch

- Generally, phonation varies with pitch
- Speakers vary how their vocal folds vibrate, to help them vibrate faster or slower
- Speakers can thus reach higher and lower pitches than would otherwise be comfortable

Experiment on full F0 range

- Audio recordings of pitch glides up or down by English and Mandarin men and women, on vowel [a]
- On glides down, speakers told either that creak is ok, or creak is not ok
- Examples:
- Measure voice quality as pitch changes within each glide – next slide shows 2 acoustic measures

2 acoustic measures vs. F0

	Women	**Men**
H1-H2		
H1-A1		

Breathier →

Pitch (F0) →

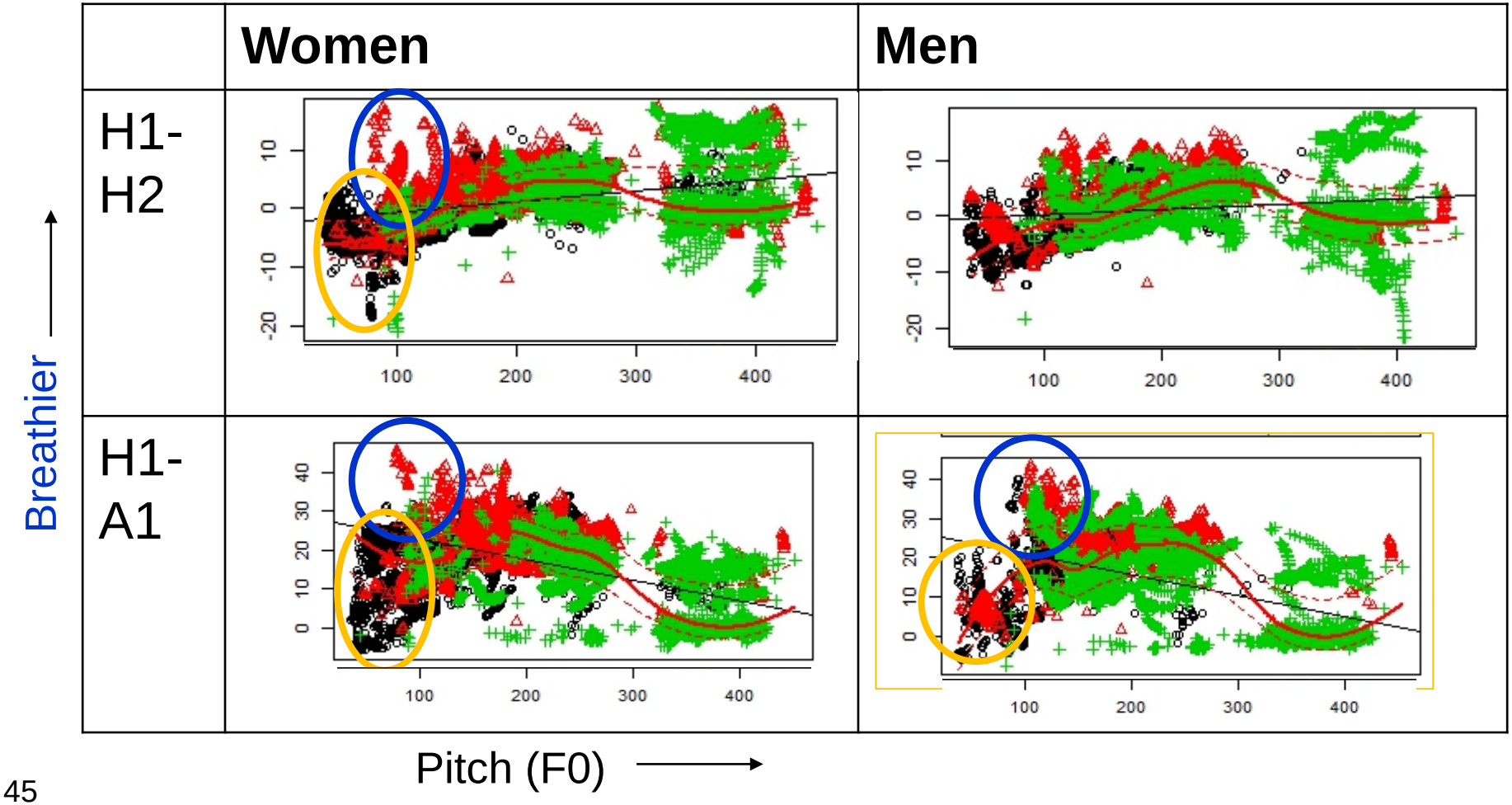
Red △ = falling pitch (don't creak)

Black ○ = falling pitch (creak is ok)

Time runs right-to-left

Green + = rising pitch

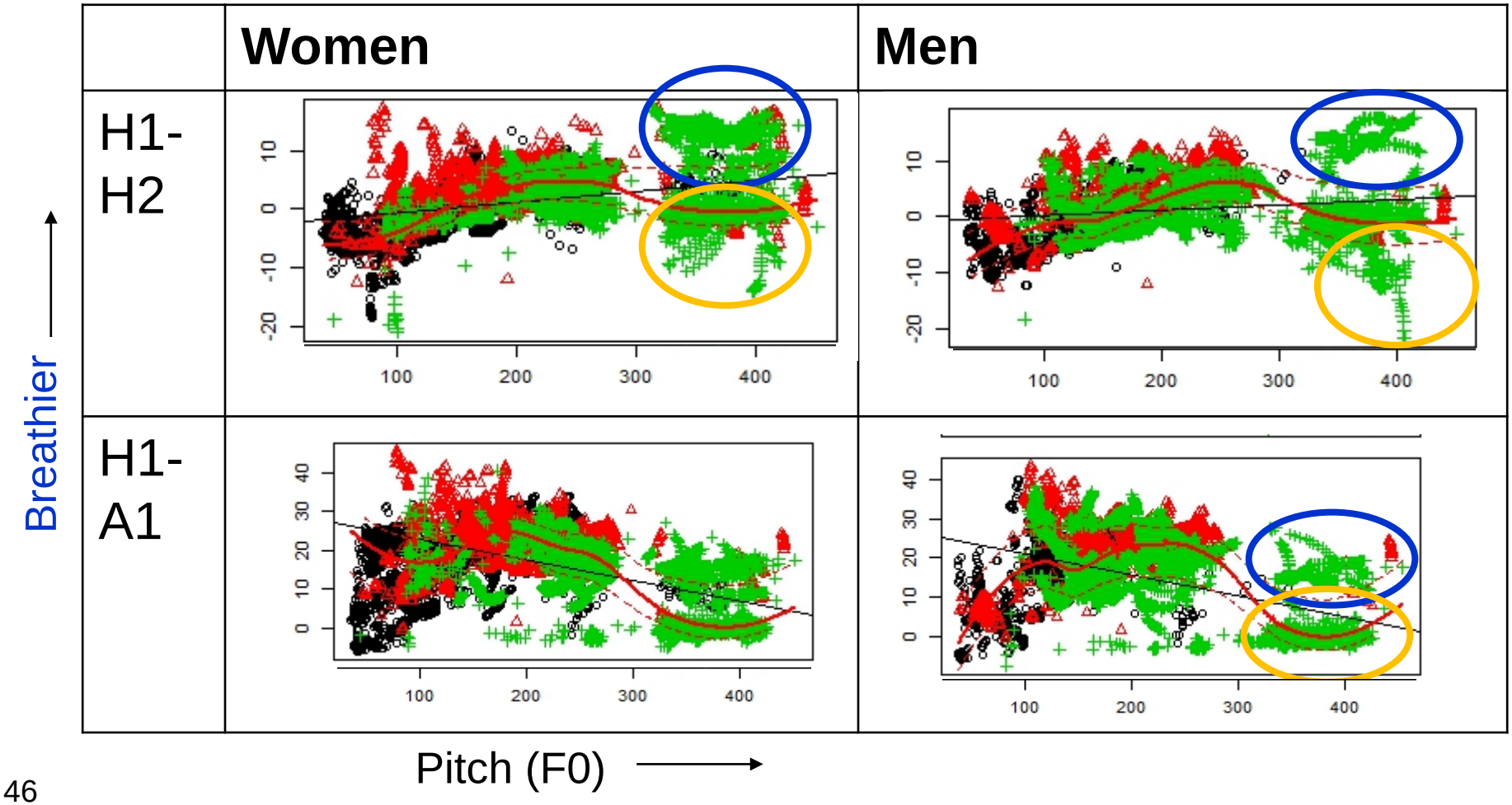
Time runs left-to-right



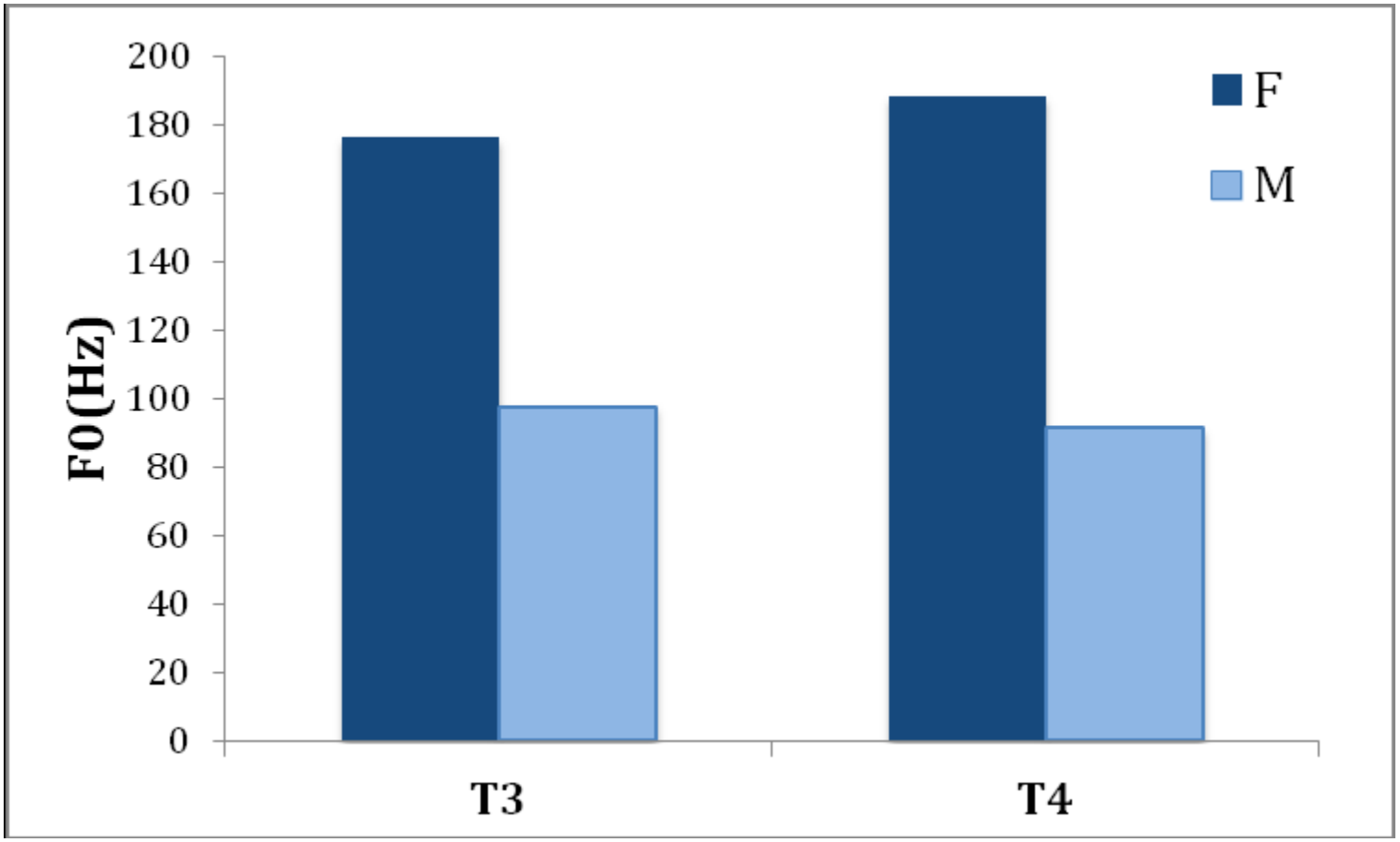
Red △ = falling pitch (don't creak)

Black ○ = falling pitch (creak is ok)

Green + = rising pitch



Mandarin creaky tones:
F0 at break into creak



Kuang 2013

Independent tone and phonation?

- Tone and phonation contrasts can be independent, combining orthogonally within a single language
- In these languages, speakers must largely de-couple pitch and quality, so that any tone can occur with any phonation
- How well can they do this – how *phonetically* independent are these phonological contrasts?

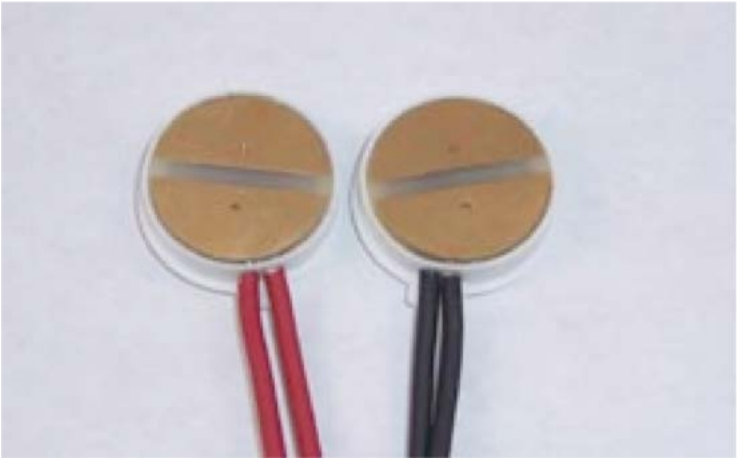
Yi languages: cross-classifying *tense* vs *lax* with tones

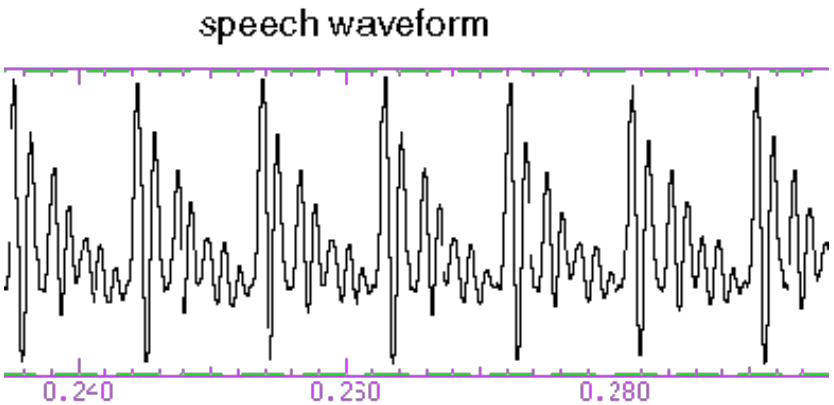
	Low tone	**Mid tone**
Lax phonation	bə ^{21} (mountain)	bə ^{33} (fight)
Tense phonation	bə̱ ^{21} (foot)	bə̱ ^{33} (shoot)

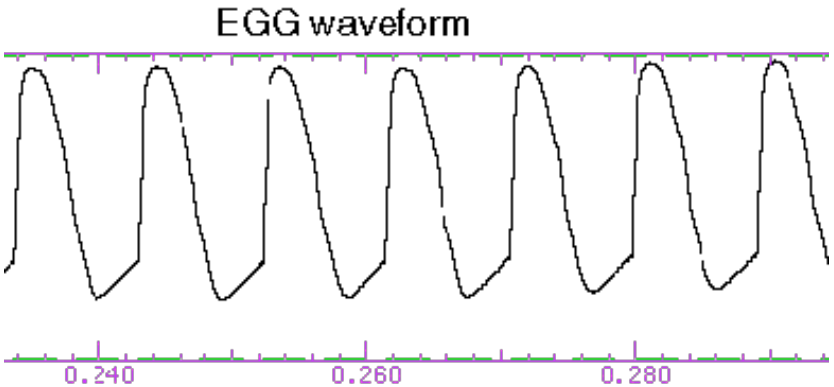
Example from Southern Yi

Kuang 2011, Kuang & Keating 2014

Electroglottography (EGG)







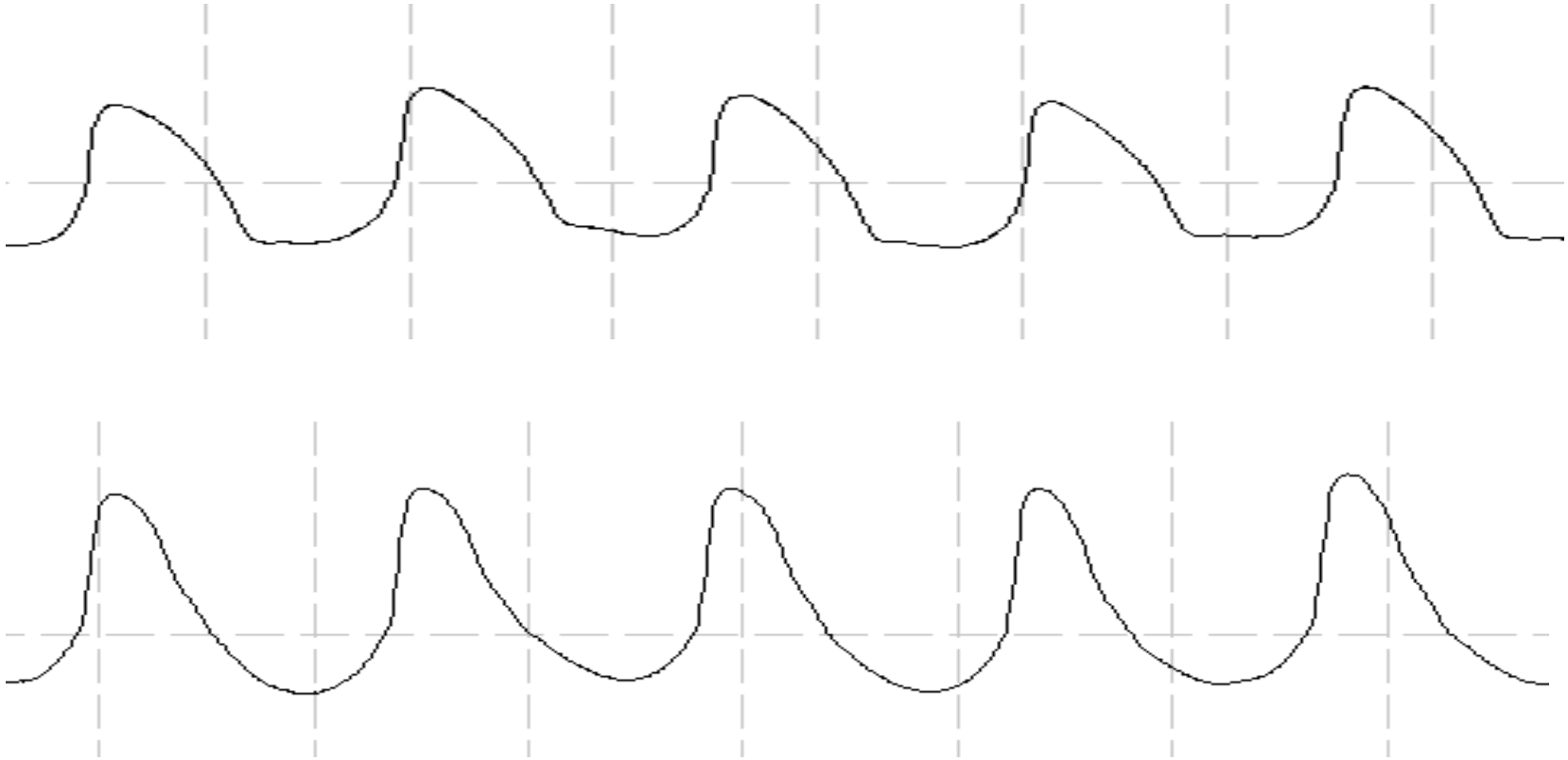
more

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 glottal constriction (higher CQ in tense or creaky voice)

Rothenberg & Mahshie 1988, Herbst & Ternstrom 2006

CQ example: White Hmong

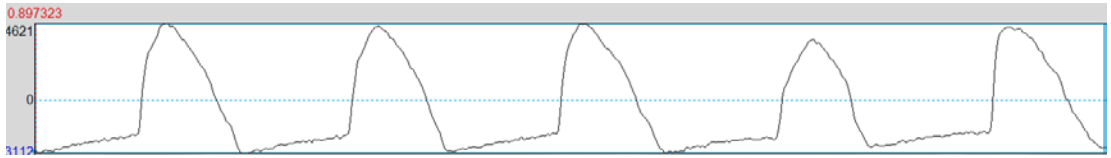
EGG waveforms of 3 phonations

more contact

less contact

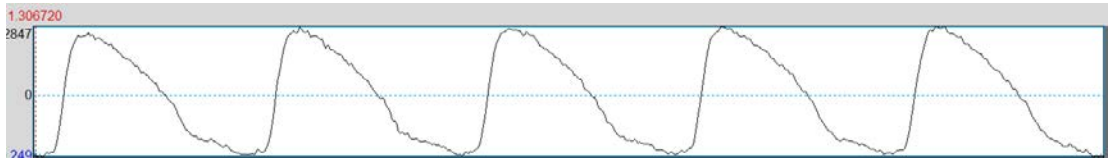
Breathy:

CQ = .41



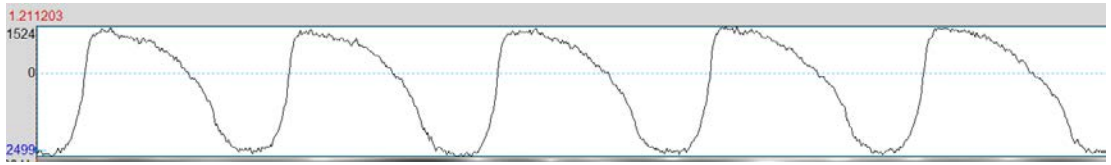
Modal:

CQ = .57



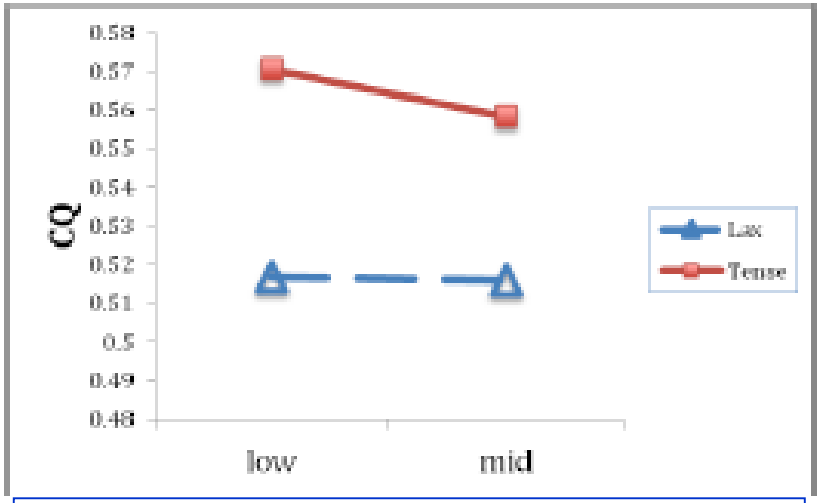
Creaky:

CQ = .65



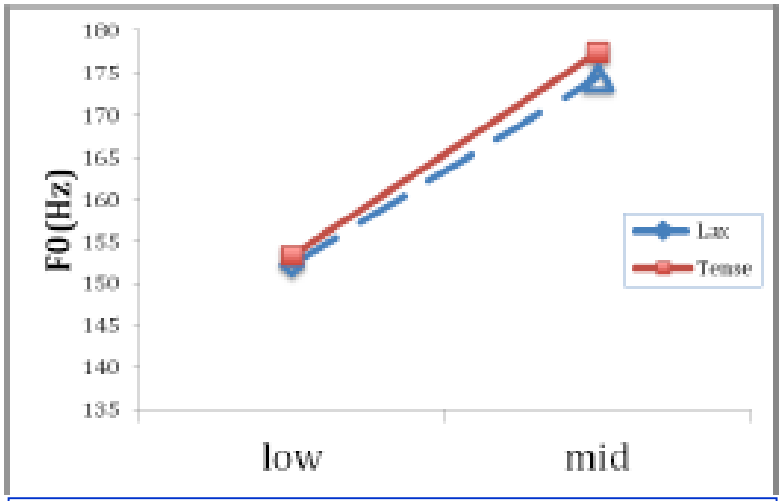
Independence of pitch and phonation in Yi

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

Summary, voice quality and pitch

- Voice quality generally varies with voice pitch, allowing pitch-range expansion
- But this is not necessary – voice quality and pitch can be quite independent in languages that cross-classify tone and phonation contrasts (e.g. Yi languages)

Conclusions

- New tools for analysis of voice quality make large-scale phonetic descriptions possible
- The cross-language phonetic space for phonations is based on the low-frequency harmonic spectrum
- Phonation can be correlated with pitch in tone- and non-tone languages, or independent of pitch in tone languages

Some of my collaborators

Jody Kreiman
UCLA Head&Neck

Abeer Alwan
UCLA Engineering

Jianjing Kuang
U Penn

Marc Garellek
UCSD

Sameer Khan
Reed College

Christina Esposito
Macalester College

Yen-Liang Shue
Dolby Australia

Grace Kuo
Concordia U