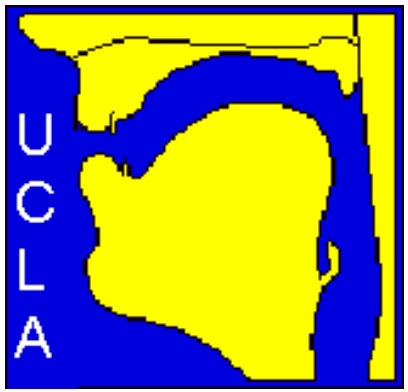


Registers in tonal contrasts

Jianjing Kuang
UCLA Linguistics Department

Contrasting levels?

- IPA (2005) and Y-R Chao: five
- Tonal theories (e.g. Clements 1979 and Yip 1980): Two ~ three
  - four → [+/−upper] register
  - Cantonese:

11 [-upper, L]	33[+upper, L]
22 [-upper, H]	55[+upper, H]

registers

- Tonal registers: [+/- upper] (Yip 1980, 2002)
  - underlying distinctive pitch ranges

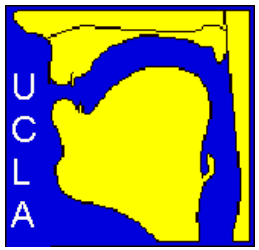
    e.g. Cantonese

    11 [-upper, L] 33[+upper, L]

    22 [-upper, H] 55[+upper, H]
- Phonation-based registers
  - E.g. Wu dialects

These two are historically related

- However, [+/-upper] tonal registers are not perceivable cues:
  - E.g. Cantonese: 22 and 33 are the most confusable tonal pair (Mok and Wong 2011)

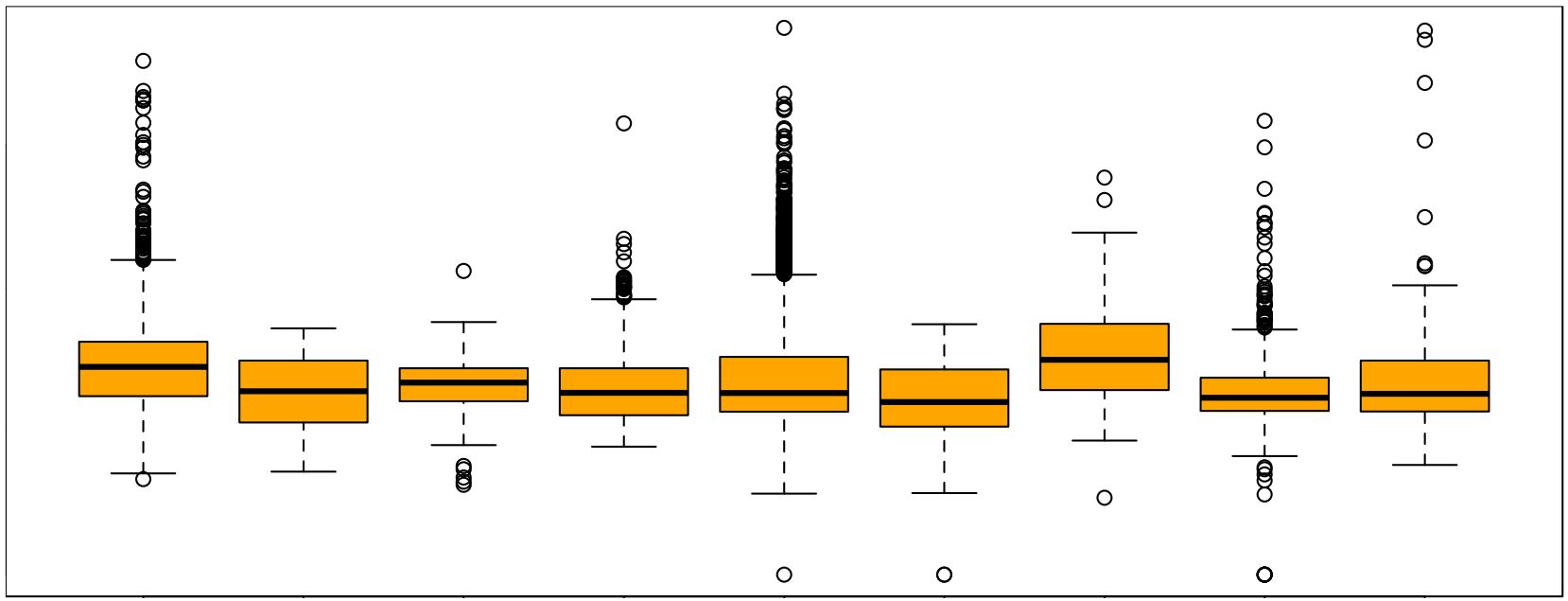


Five-level-tone contrast is very hard to maintain, because…

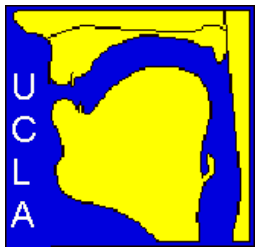
Limitation in production:

- pitch range of normal speech is around 100Hz (Baken and Orlikoff 2000)
- Also see next slide, our data

Pitch range across languages (male speakers)

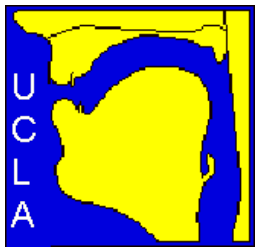


UCLA languages corpus

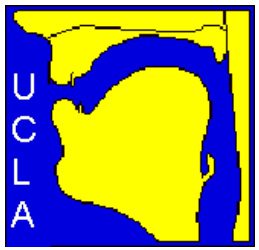


Limitation in perception:

- JND of pitch in lexical tones is about 9Hz (Silverman 2003), but a phonological contrast requires much greater difference
- 20-30Hz difference for a tonal contrast is a small number, e.g. Cantonese 22 and 33 are very confusable and merging (Mok and Wong 2011)

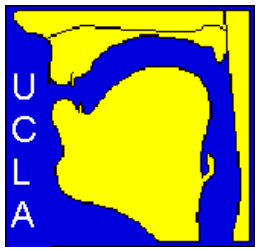


- Even a three-level contrast is very hard to maintain in a 100 Hz range, not to mention a fourth or fifth level.



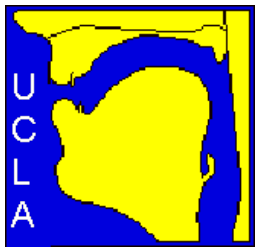
Dispersion

- Seems very dispreferred contrast system (Lindblom and Maddieson 1988, Flemming 2002)
- Violate both goals: maximize perceptual contrasts; minimize articulatory efforts



This talk

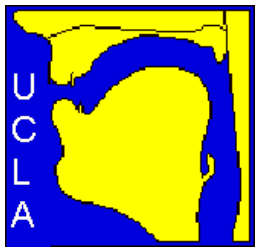
Given normal hearing and speaking ability, how can native speakers produce and hear multiple contrasting level tones?



This talk

Given normal hearing and speaking ability, how can native speakers produce and hear multiple contrasting level tones?

àThe tonal production and perception of a language with five-level-tone contrasts.



This talk

Given normal hearing and speaking ability, how can native speakers produce and hear multiple contrasting level tones?

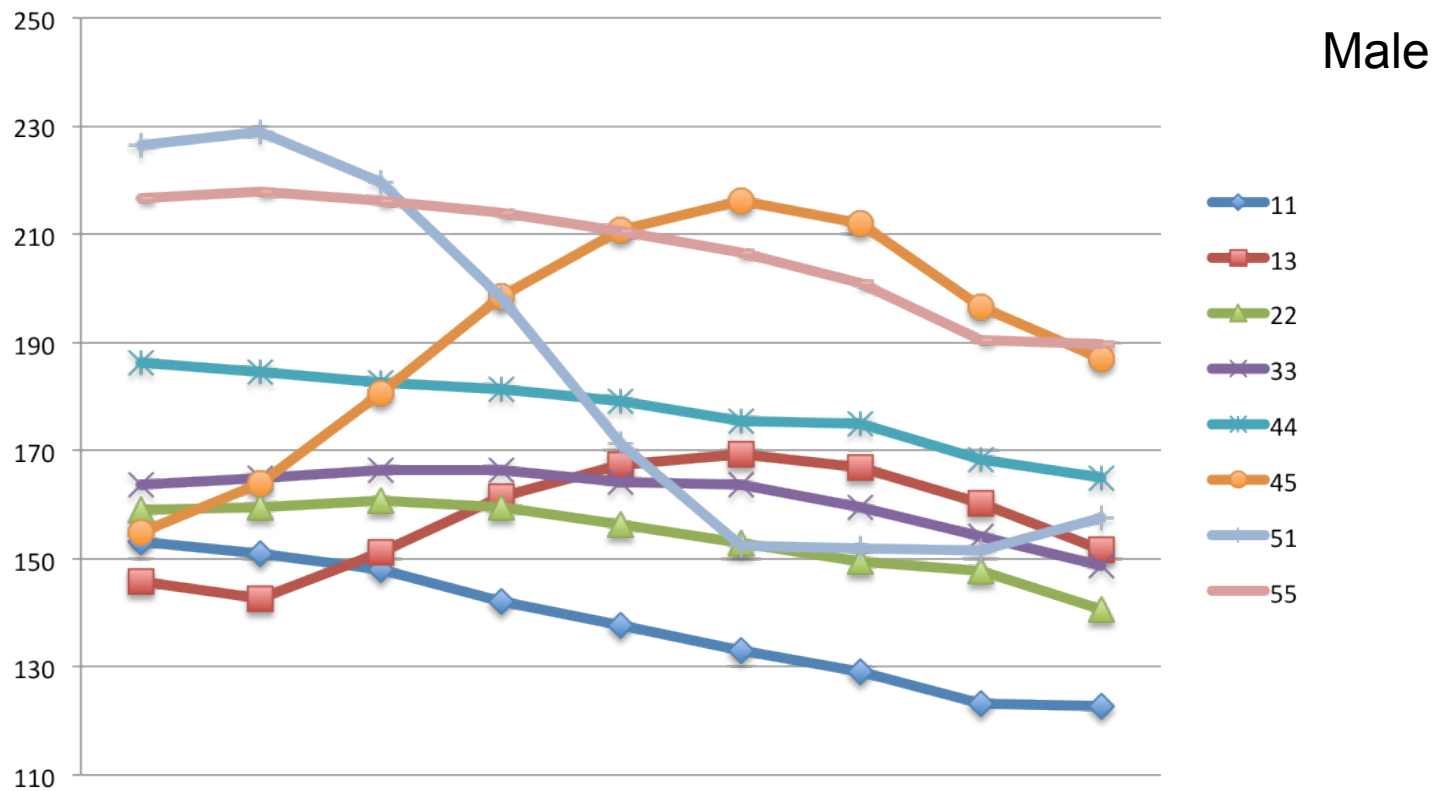
àThe tonal production and perception of a language with five-level-tone contrasts.

Black Miao

- Black Miao dialect, called Qingjiang Miao (Ch'ing Chiang Miao). This dialect is spoken at Shidong Kou (Shih-Tung-K'ou), Taijiang (T'ai-Kung) county of Guizhou (Kweichow) province in China.
- First investigated by Fang-Kuei Li in 1940s

Tonal system

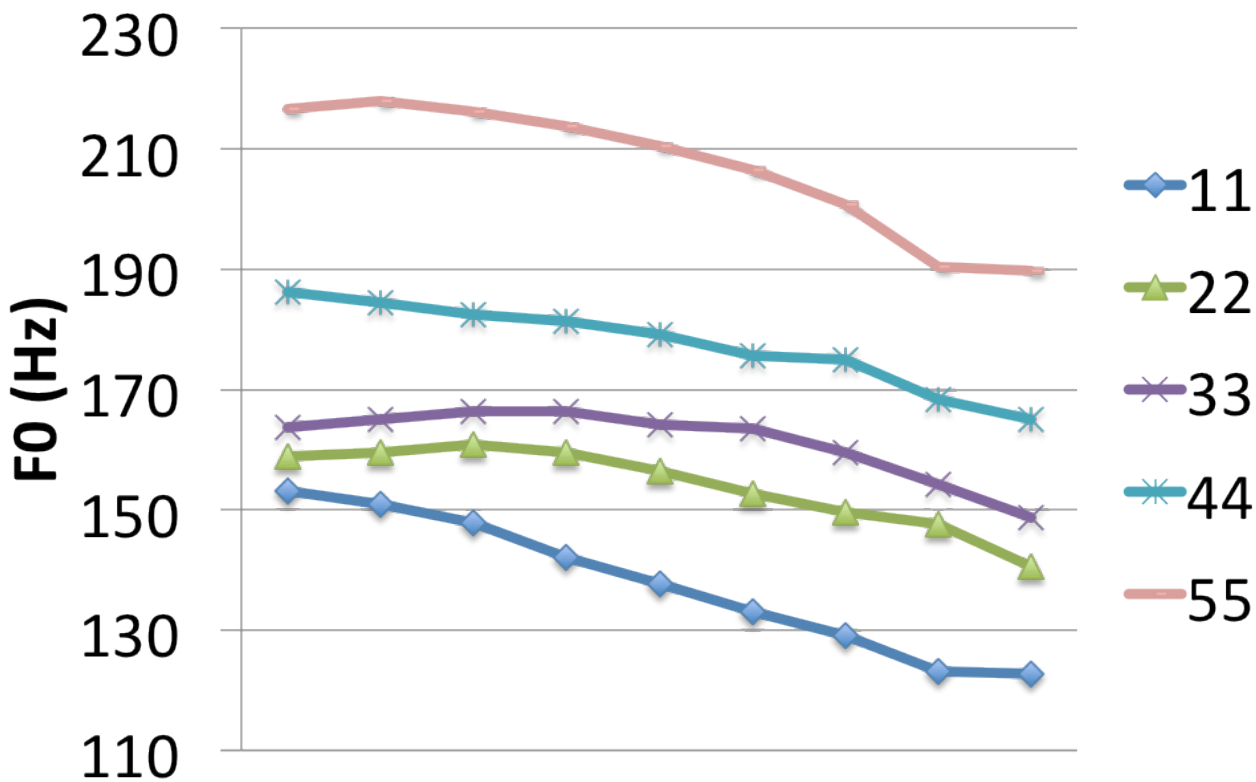
I	II	III	IV	V	VI	VII	VIII
44	51	55	22	45	33	13	11

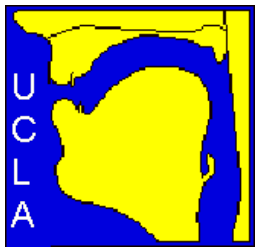


Tonal system

I	II	III	IV	V	VI	VII	VIII
44	51	55	22	45	33	13	11

Male





Registers:

- upper: 44, 55, 45, 13
- lower: 51, 22, 33, 11

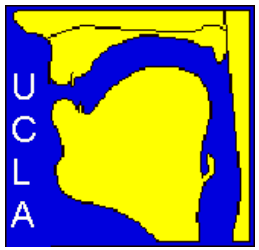
Perception experiment

Stimuli:

- A minimal set of eight real monosyllabic words with [pa]. Produced by a male native speaker.
- /pa55/ "(water) full” 
- /pa44/ "send” 
- /pa33/ "fail” 
- /pa22/ "net” 
- /pa11/ ”pull” 

Procedure

- Familiarity phase: testing words were instructed in proper contexts
- Identification: single audio target; preceded by an audio introduction
- AX discrimination: two audio stimuli (possible pairs among eight tones); measuring RT and accuracy.

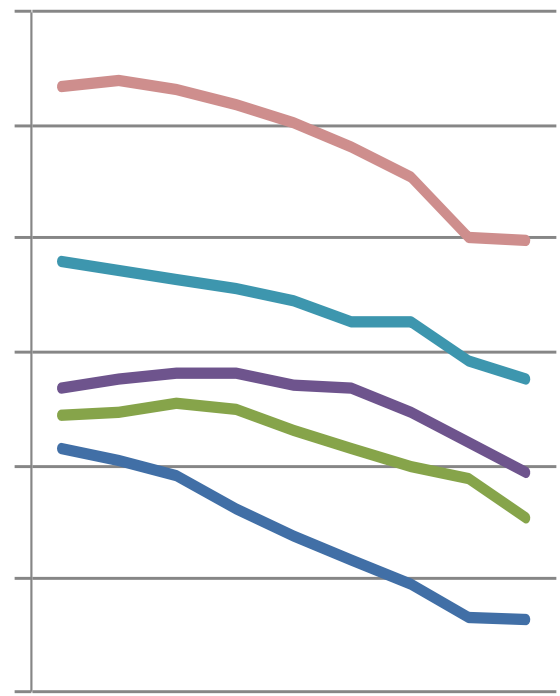


Subjects

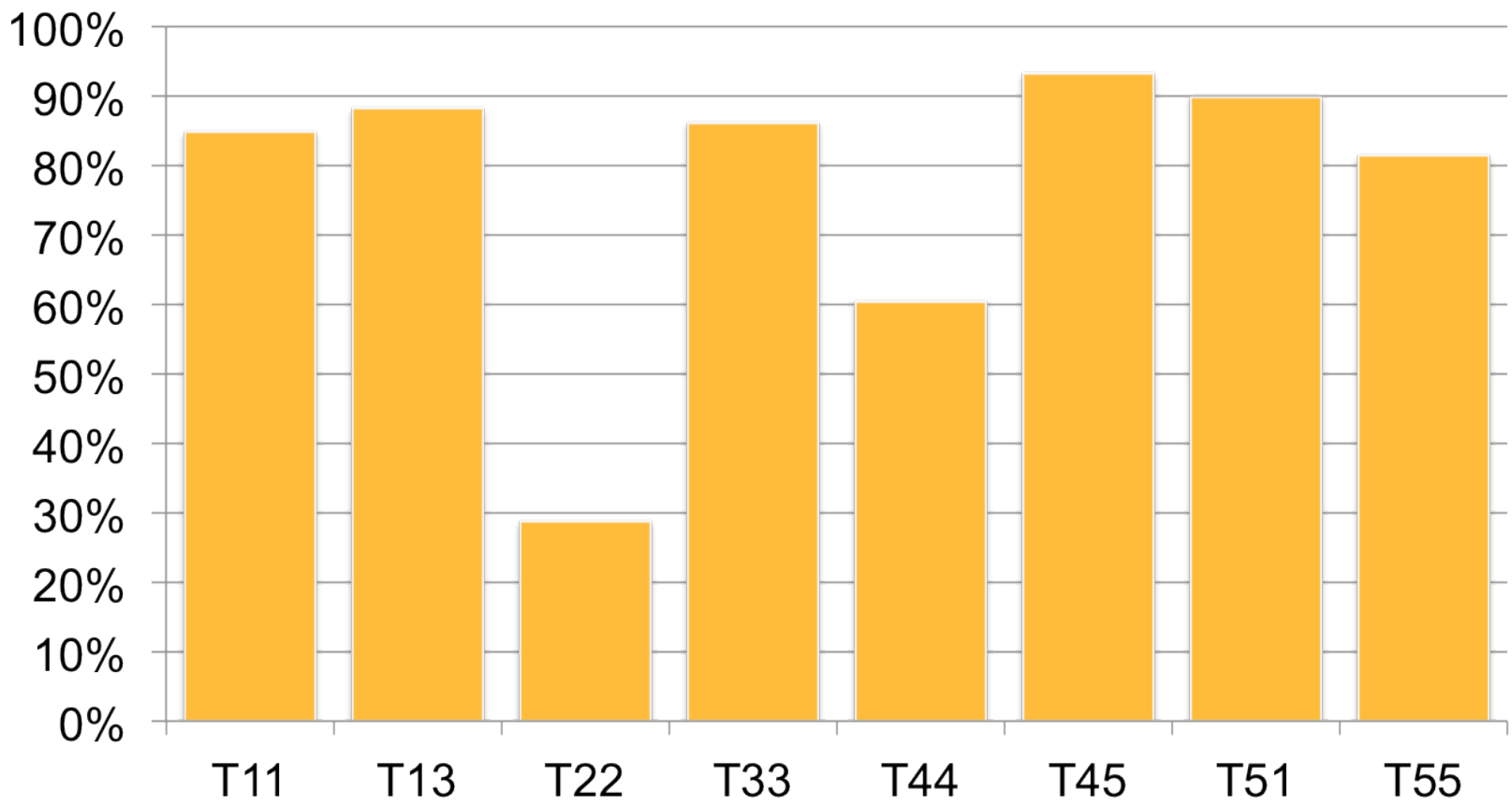
¢A total of 18 subjects, eight males and ten females, participated in this experiment. Four females, who were not native speakers of this particular Black Miao dialect, were excluded from the current analysis, leaving 14 subjects.

Hypotheses

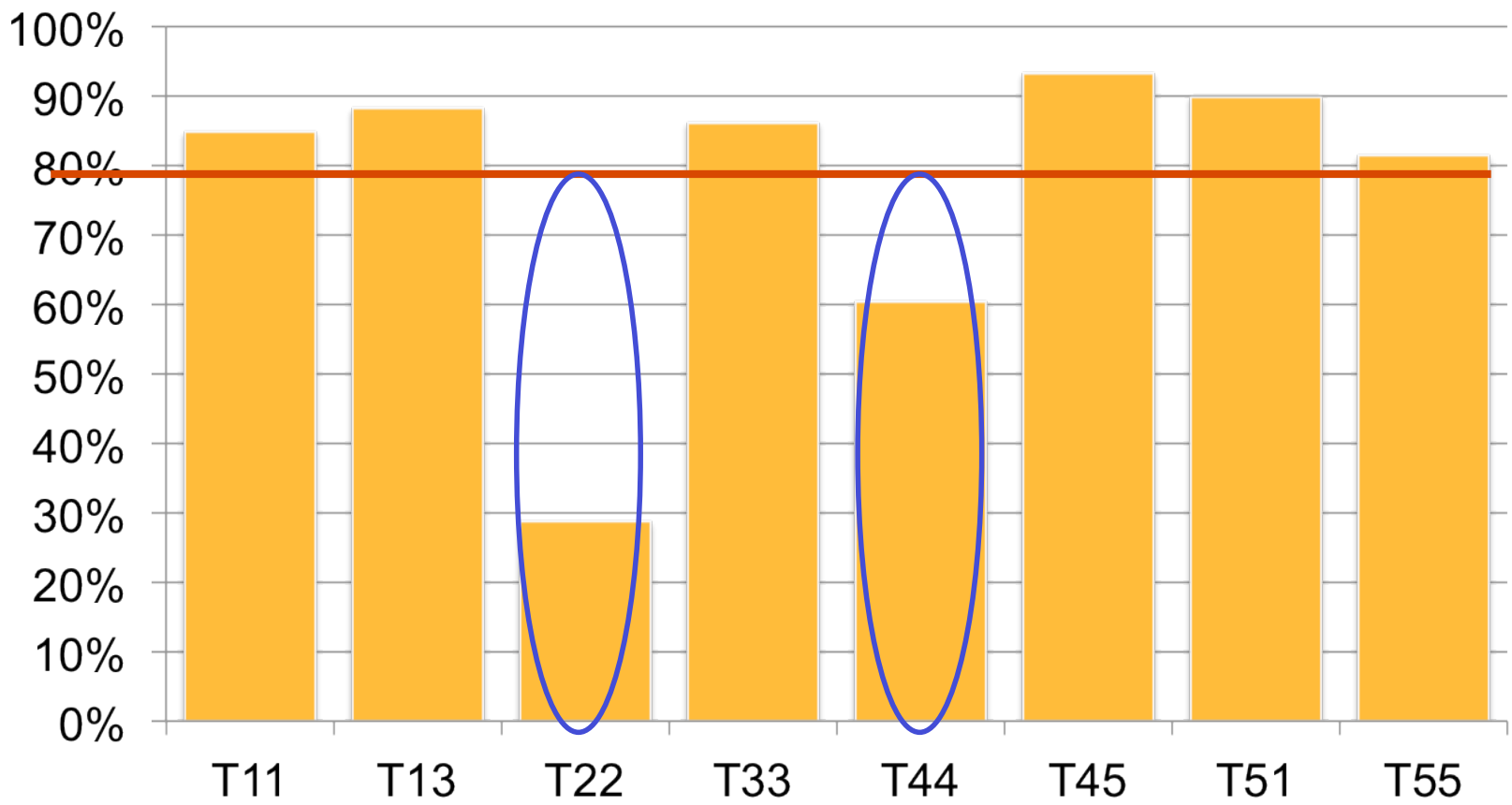
- Tones with adjacent pitch values are in trouble
- If F0 is the only cue, accuracy for 33 should be the worst.



ID Accuracy



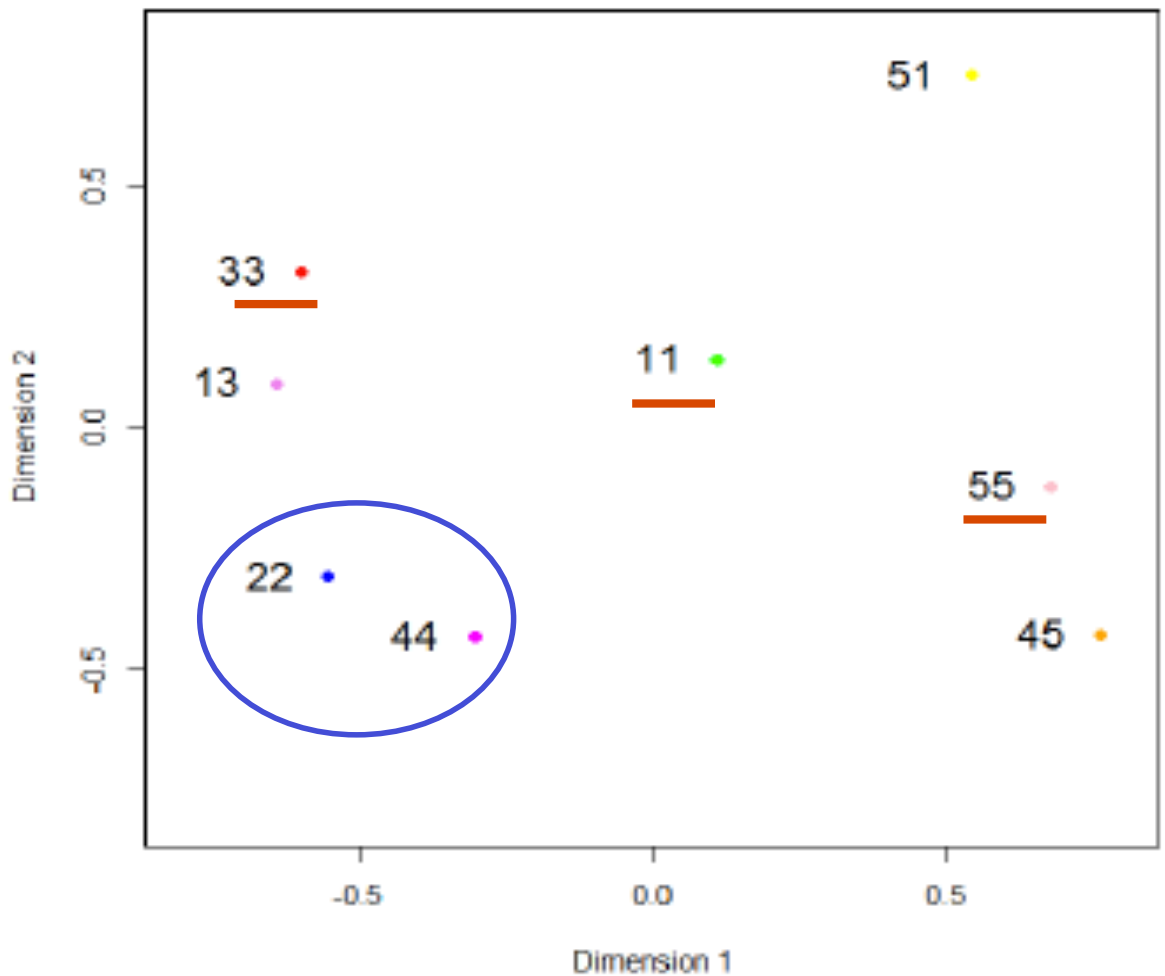
ID Accuracy



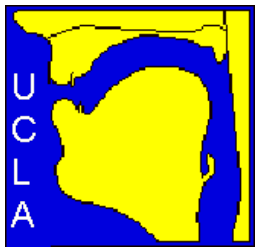
Dissimilarity matrix for all listeners.

	T11	T13	T22	T33	T44	T45	T51	T55
T11	0.05							
T13	0.94	0.00						
T22	0.93	0.88	0.03					
T33	0.97	0.78	0.95	0.05				
T44	0.98	1.00	0.70	0.98	0.03			
T45	0.94	1.00	1.00	1.00	1.00	0.00		
T51	0.94	1.00	1.00	1.00	1.00	1.00	0.00	
T55	0.95	1.00	1.00	1.00	0.92	0.88	0.88	0.00

MDS perceptual space

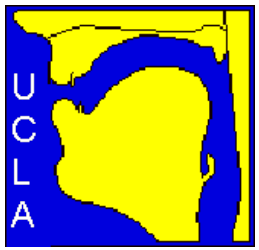


Followed Johnson (2003)



Production experiment

- A wordlist of minimal monosyllabic sets for the eight tones was created based on Li's transcriptions (Kwan1966). 23 minimal or near-minimal sets were confirmed by the speakers.
- Simultaneous EGG and audio recordings were then collected from 15 native speakers (ten males and five females).



Measures

- Pitch related:

mean F0, F0 change, offset, onset

- Duration
- Voice quality related (* corrected):

H1*, H2*, H4*, A1*, A2*, A3*;

H1*-H2*/A1*/A2*/A3*;H2*-H4*

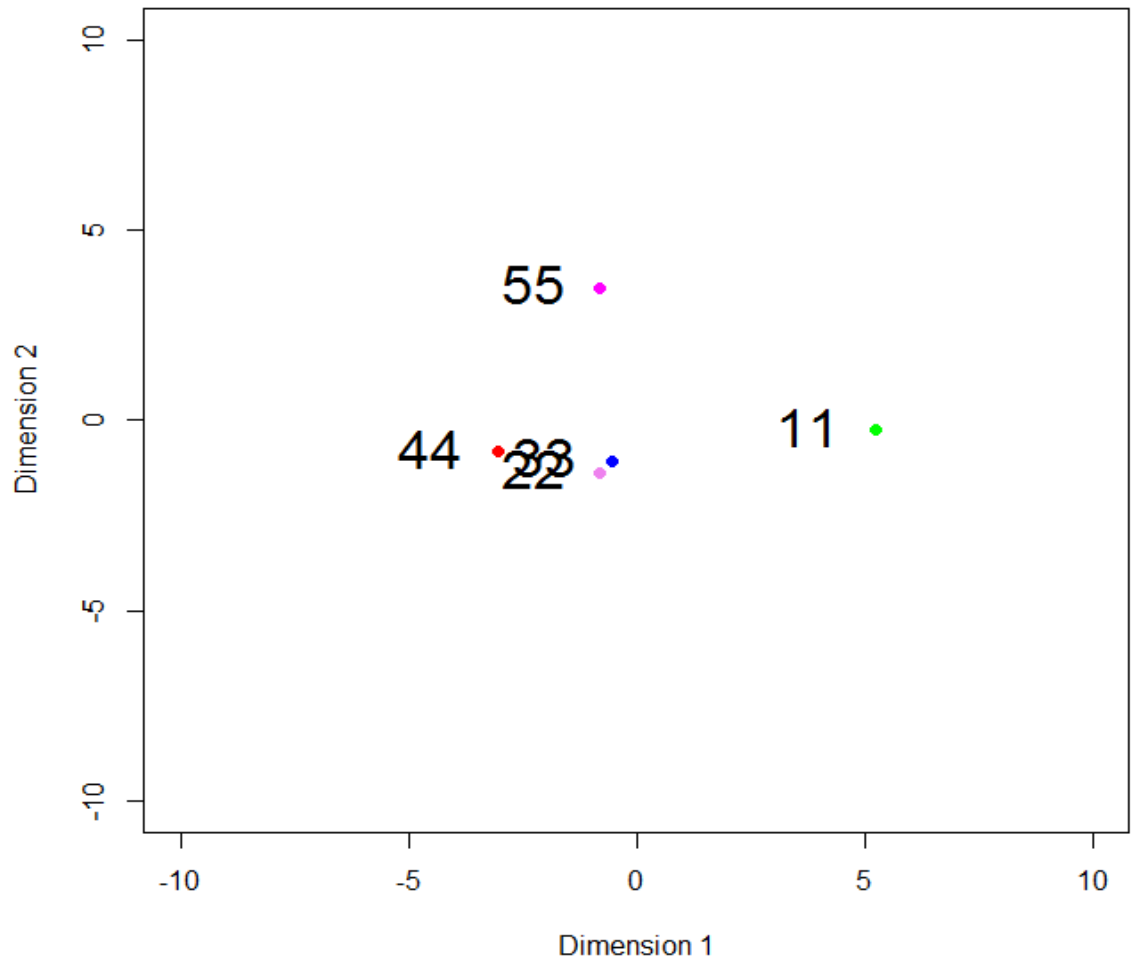
- EGG

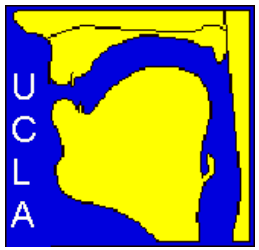
CQ (contact quotient), SQ (skew quotient), PIC, PDC

MDS Production space

Pitch measures only

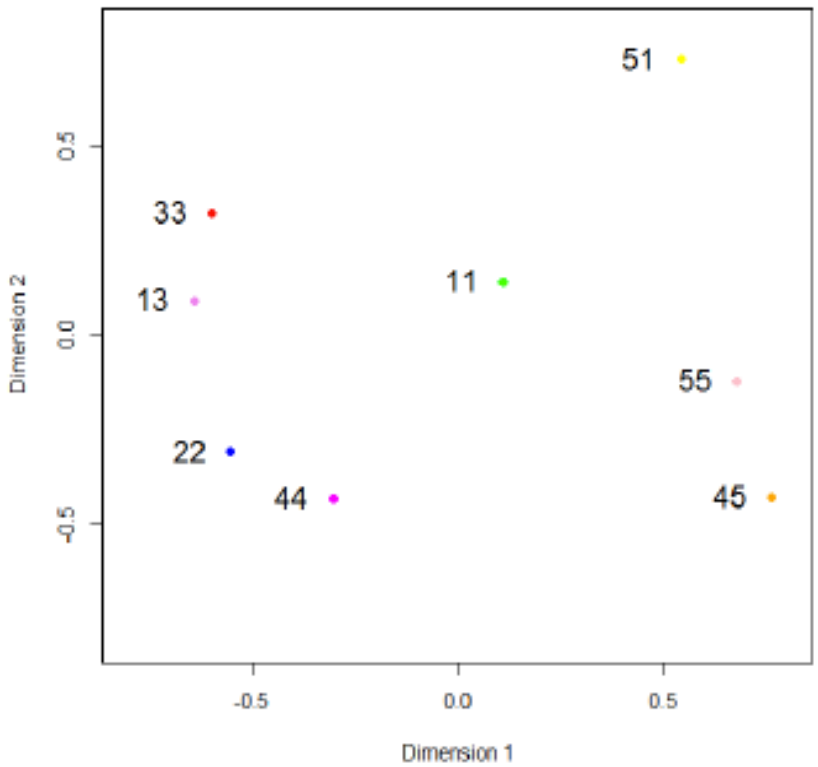
- Mean F0,
- F0 change
- onset
- offset
- duration



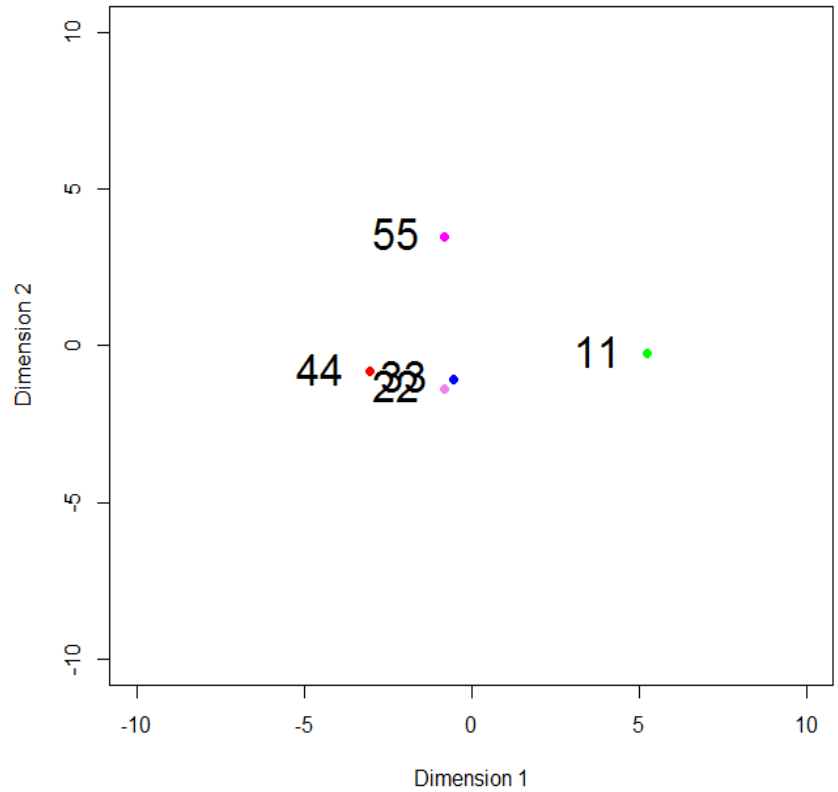


R=0.17

Perceptual space



Production space



MDS Production space

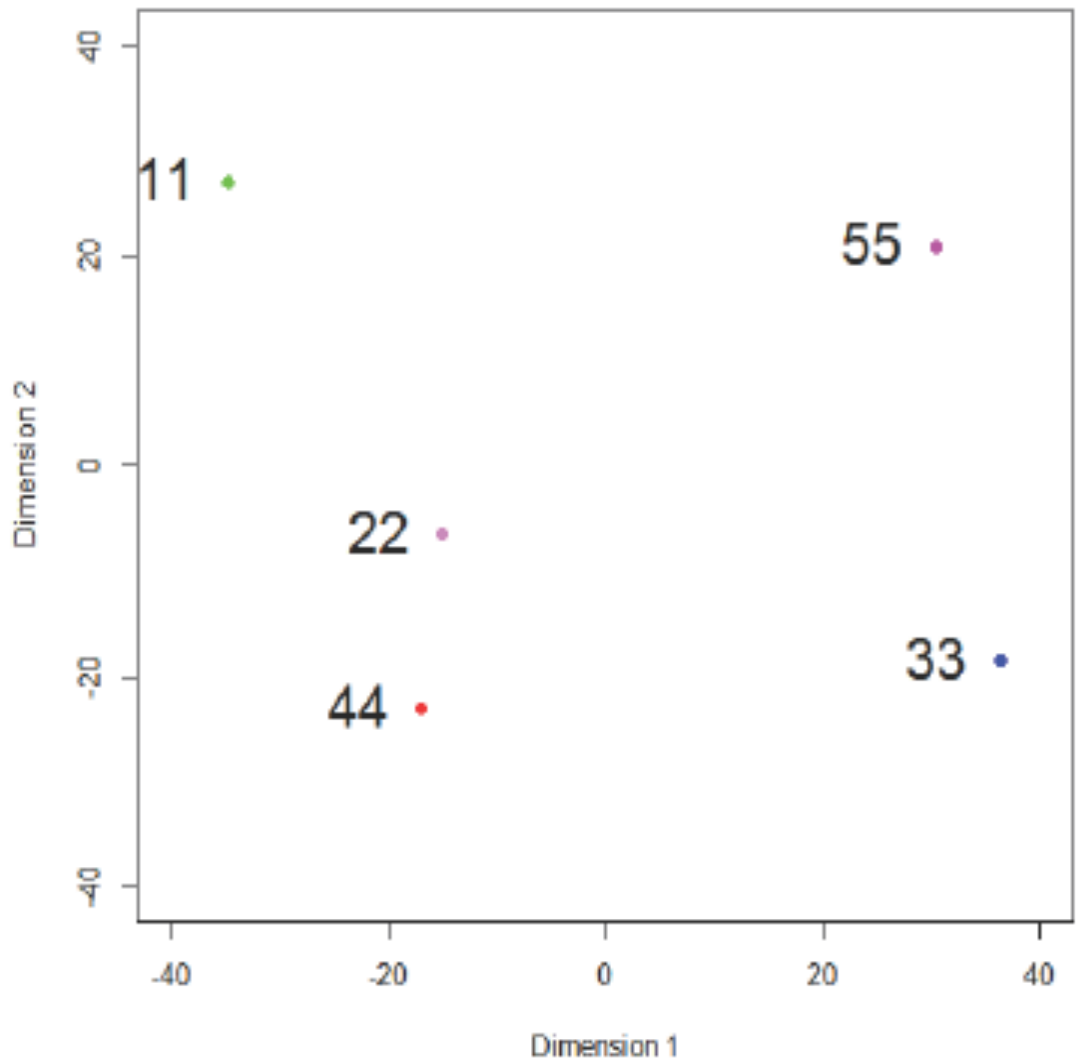
Pitch + voice quality

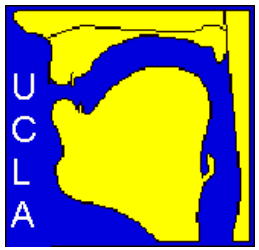
H1*,H2*,H4*

H1*-H2*

H1*-A1*/A2*/A3*

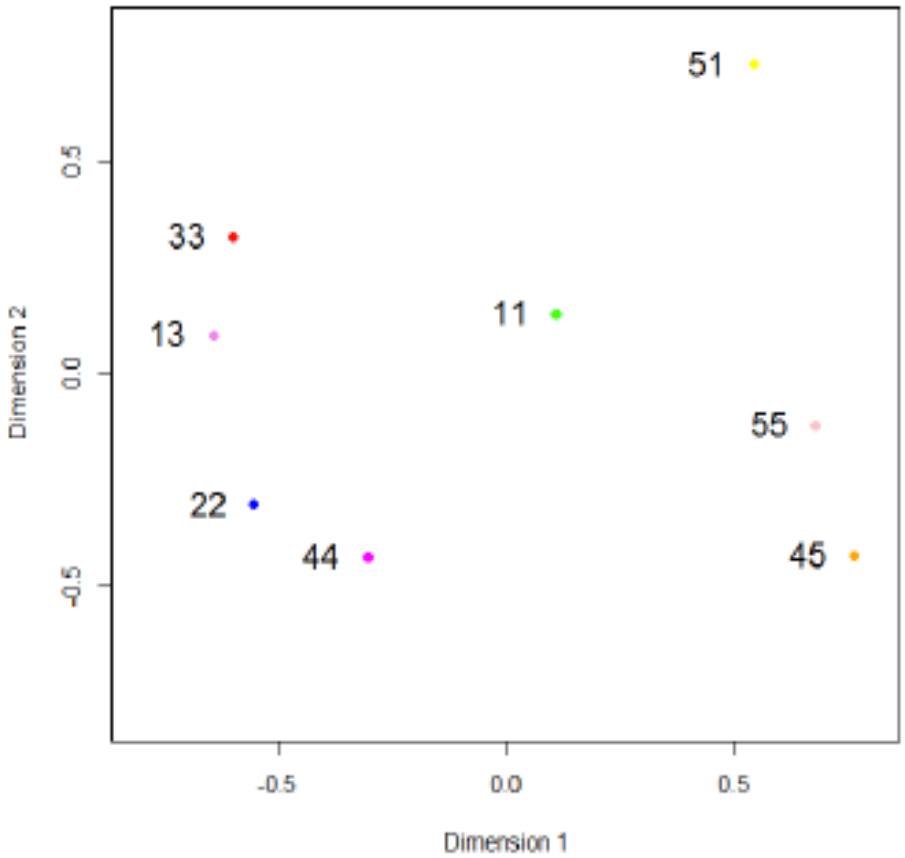
H2*-H4*



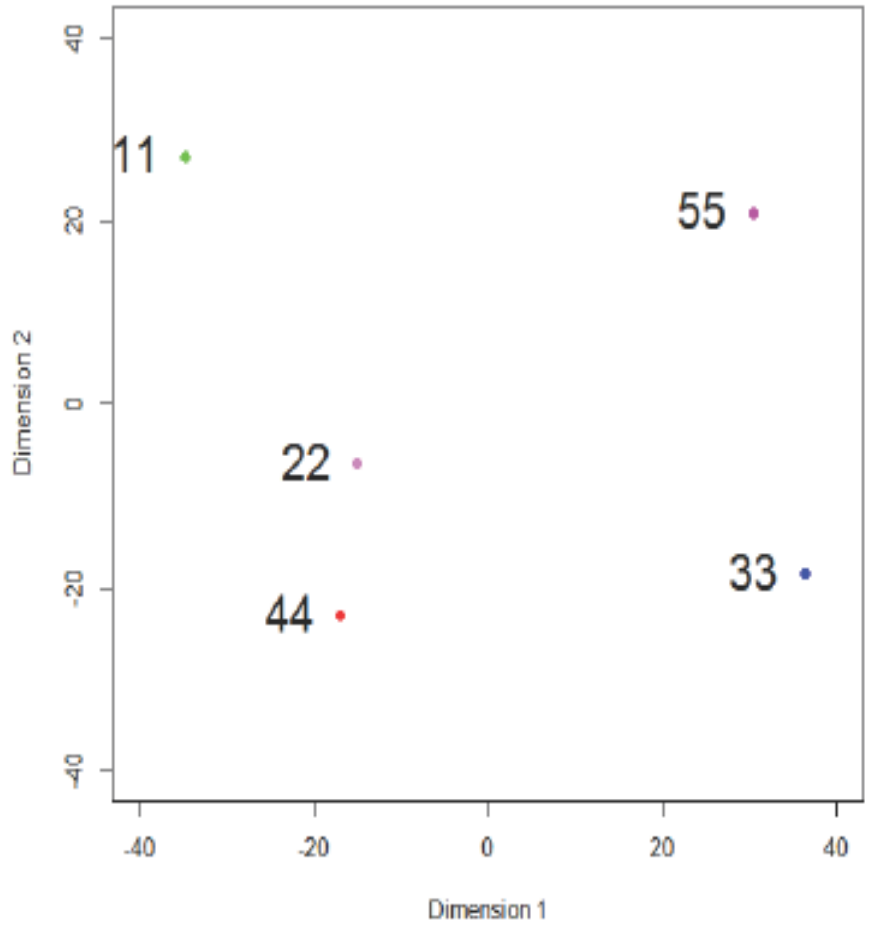


R=0.76

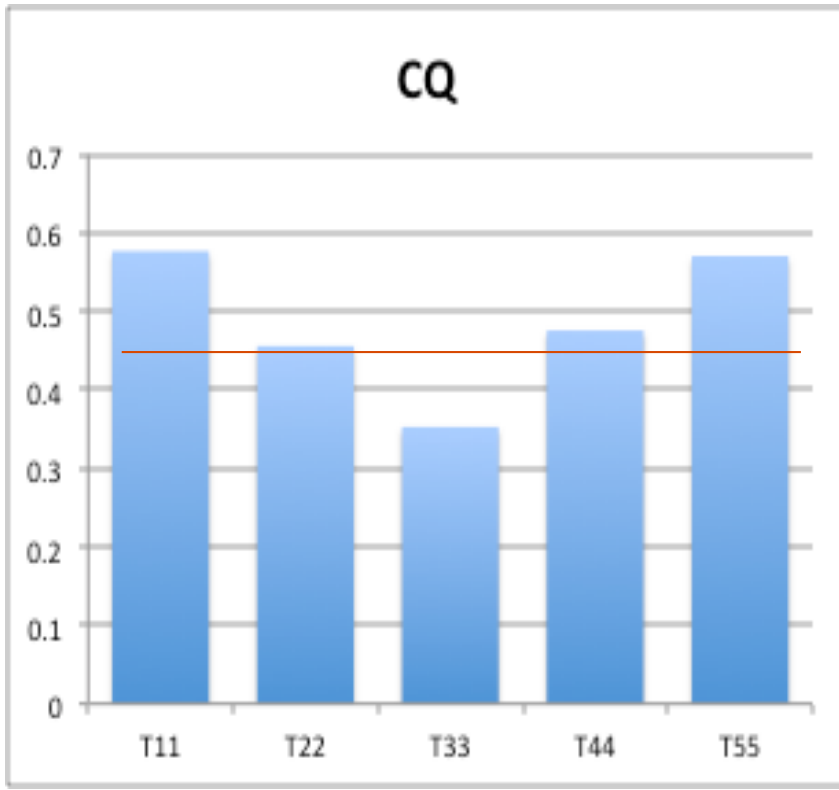
Perceptual space

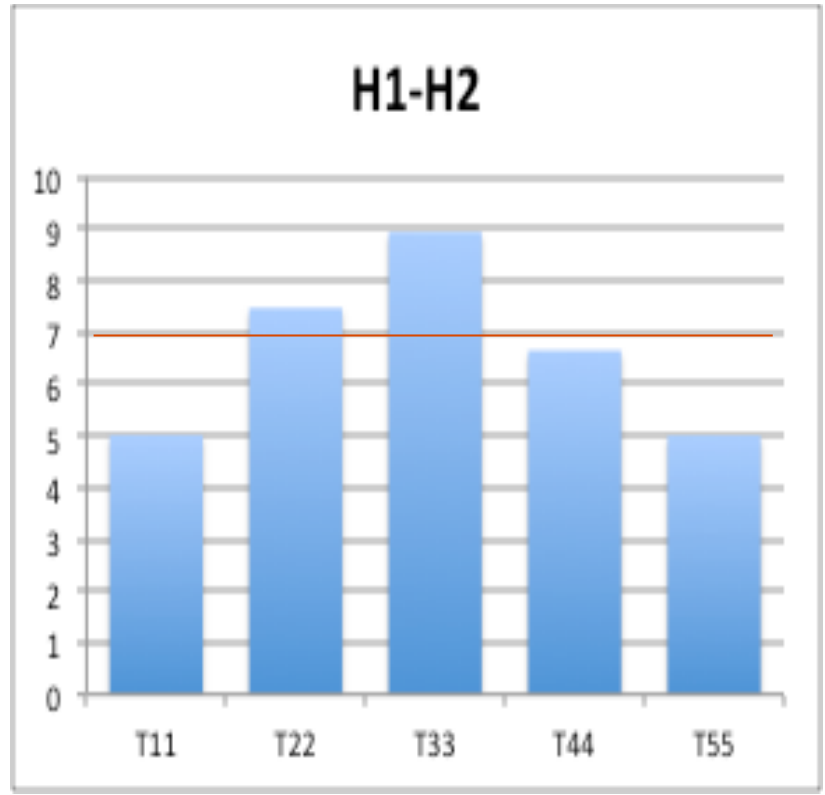


Production space

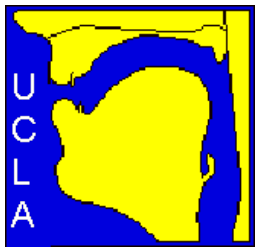


Voice qualities in tones



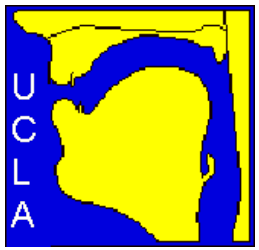


11, 33 and 55 benefit from phonation cues

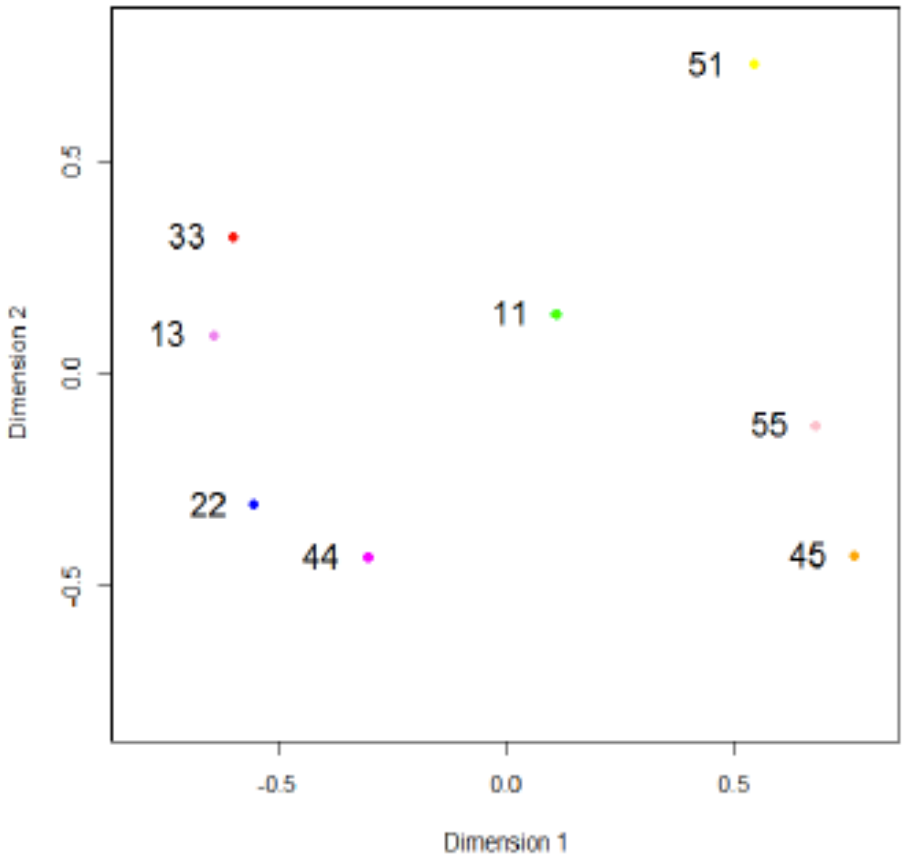


Discussions

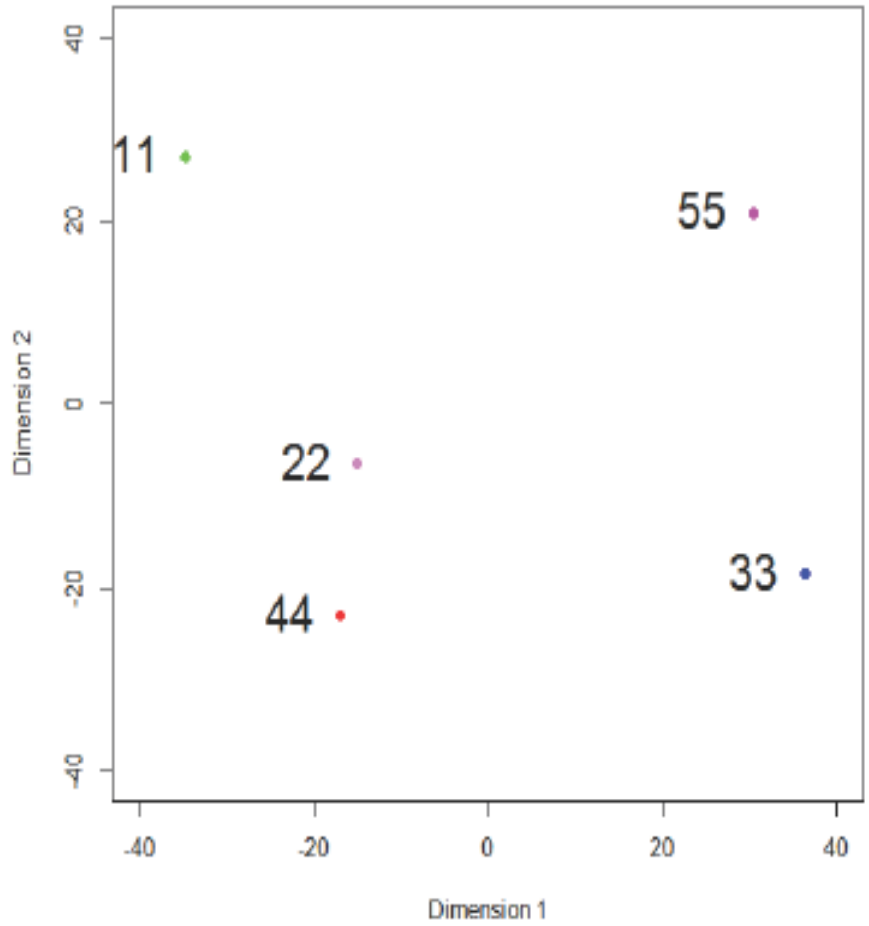
- 55 and 11 can benefit from both pitch cues and phonation cues
- For the mid-range tones that have very similar pitch cues, 33 is distinctive from 22 and 44 primarily by the phonation cue.
- 22 vs. 44, the tonal contrast with only a pitch difference is the most confusable



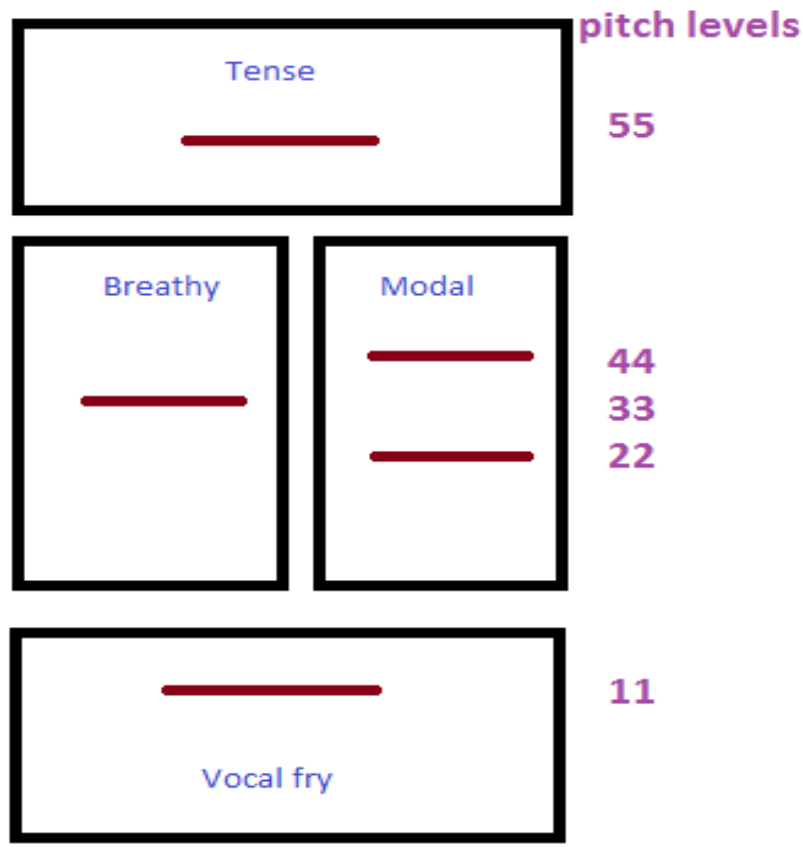
Perceptual space

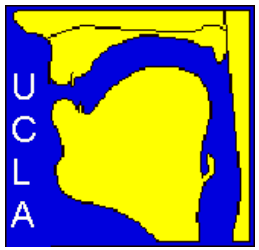


Production space



Tonal registers





Contributions of non-modal phonations

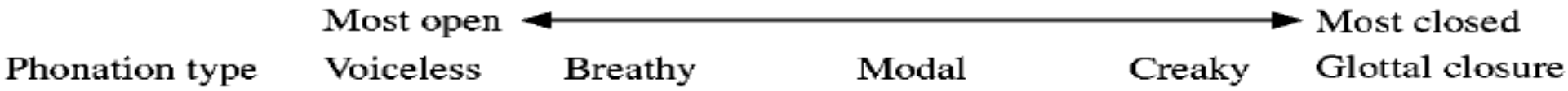
- 55 and 11: enhance the tonal contrasts
- 33: distinctive cue from the other mid tones

Two types of non-modal phonations

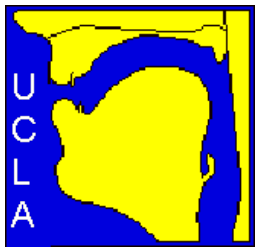
- Pitch-driven type: e.g. vocal fry, falsetto and tense -- part of the fundamental frequency scale
- Pitch-range production is related to certain phonation types: (Hollien 1974, Titze 1988, Baken&Orlikoff 2000)
  - The lowest pitch range (i.e. < 70Hz) – vocal fry
  - The highest pitch range (i.e. > 175Hz for males, 275Hz for females) – falsetto/tense

Two types of non-modal phonations

- Pitch independent type: create an independent dimension for tonal contrasts
  - 33 vs. 22 and 44 (similar pitch but different registers are not confuable)
  - Also Green Mong (Andruski 2006) and White Hmong (Garelleck et al. 2012)
- breathiness contrast – relative on a phonation continuum

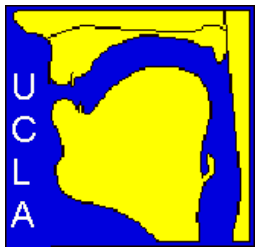


Ladefoged's model of continuous breathiness



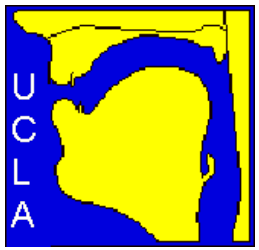
Concluding remarks

- Dispersion of the five level tones is fine:
  - Pitch-driven phonations help to produce extreme F0 targets, and thus enhance the perceptual differences for the highest and lowest tones
  - Pitch-independent phonations create an independent dimension for tonal contrasts so that tones with similar pitches are very well distinguished



associations

- Pitch register – pitch-driven non-modal phonations – allophonic
  - Realization of extreme pitch targets (highest, lowest)
  - E.g. vocal fry: Mandarin, Cantonese
    falsetto: Gaoba Dong, PPhN Thai
- Phonation-based register – pitch independent non-modal phonations – phonemic
  - E.g. Yi, Mazatec



Thank you! ☺

put vot

Acknowledgments

- Advising committee: Patricia Keating, Jody Kreiman, Sun-Ah Jun, Bruce Hayes and Megha Sundra for invaluable comments
- Colleagues in UCLA phonetics lab for all the discussions
- Professor Jiangping Kong of Peking University for the equipment support
- Professor Defu Shi of Minzu University of China for introduction of Black Miao letters
- Lindy Mark for the help of getting Prof Li's word list
- Black Miao friends who kindly participated in this study