The Effect of Musicality on Cue Selection in Pitch Perception by English and Mandarin Speakers

Aletheia Cui

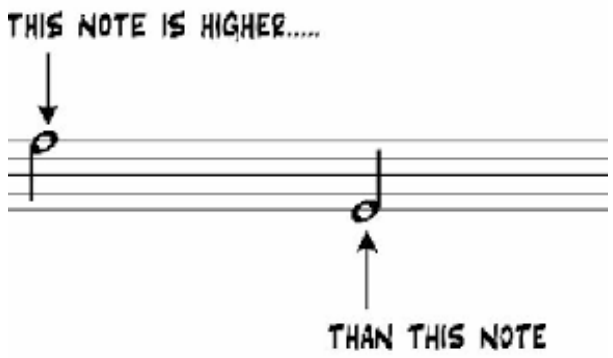
Jianjing Kuang

ICPhS2019

Penn
University of Pennsylvania

Music and speech share similar acoustic cues

- Pitch
- Rhythm
- Spectral shape



Tea = 𝅗𝅥 1 Beat

Coffee = ♫ ½ beat each

Coca cola = 𝅘𝅥𝅯𝅘𝅥𝅯𝅘𝅥𝅯𝅘𝅥𝅯 ¼ beat each

Lemonade = 𝅘𝅥𝅯𝅘𝅥𝅯𝅘𝅥𝅮 ¼ + ¼ + ½

Pineapple = 𝅘𝅥𝅮𝅘𝅥𝅯𝅘𝅥𝅯 ½ + ¼ + ¼

Musicality and language ability

- Musicality: aptitude for music
  - Innate (nature)
  - Learned (nurture)
- Good musicality -> better language abilities
  - Improved reading/verbal skills for L1 (Spychiger 1993; Douglas and Willatts 1994; Costa-Giomi 1999)
  - Better perception and production of L2 (Harrison 1979; Stevenson 1999; Pastuszek-Lipinska 2008)

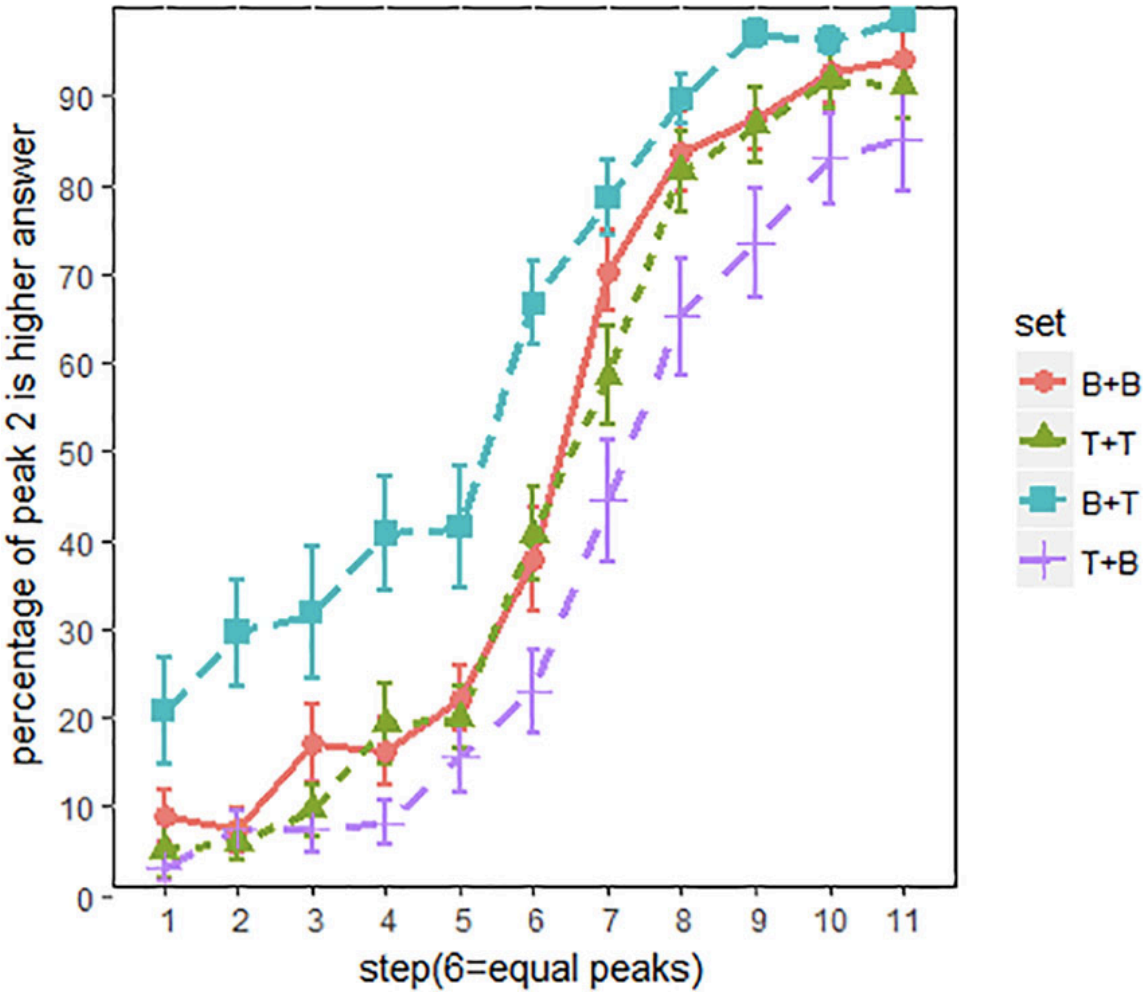
What about people with poor musicality?

- One might expect disadvantages in language ability
  - E.g., Learning tones
- However, congenital amusics acquire tone languages
  - Have normal tone production (Nan et al 2011)
  - Show categorical perception of tones (Huang et al 2015)
- How do people with amusia acquire tones, if they are unable to process pitch?

Pitch in speech is multidimensional

- In addition to f0, **phonation cues** also play a role in pitch perception
- Co-variation between f0 and phonation cues:
  - Spectral tilt: more flat spectrum (tenser phonation) -> higher pitch (Kuang and Liberman, 2015, 2016a)
  - Jitter/periodicity: more irregularity (creakier phonation) -> lower pitch (Kuang and Liberman, 2016b)
  - Gender/vocal tract length information (Kuang and Liberman 2016c)
- People who struggle with f0 can use co-varying cues

Previous findings



Research Questions

- How does musicality affect cue selection in pitch perception?
  - Hypothesis: More musical -> Preference for f0 cues
  - Approach: Correlation between musicality and cue preference
- Is pitch perception strategy affected by language background?
  - Hypothesis: Tone language -> Enhanced pitch perception
  - Approach: Recruit tone (Mandarin) vs non-tone (English) speakers

Methods

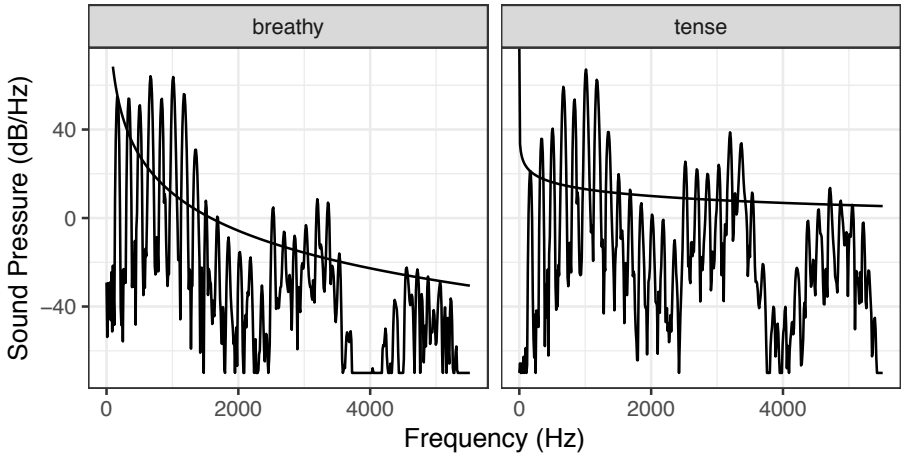
Overview

- Experiment 1: Pitch perception
- Experiment 2: Musicality
- Subjects: Two language groups
  - Non-tone: 71 English speakers (Age mean 19.74, sd 1.60, range 18-25)
  - Tone: 44 Mandarin speakers (Age mean 24.96, sd 6.93, range 18-50)

Experiment 1: Pitch Perception

Spectral manipulation of stimuli

- 4 spectral slope conditions
  - Breathy = natural spectrum
  - Tense = boosted 6 dB/Ocatve up

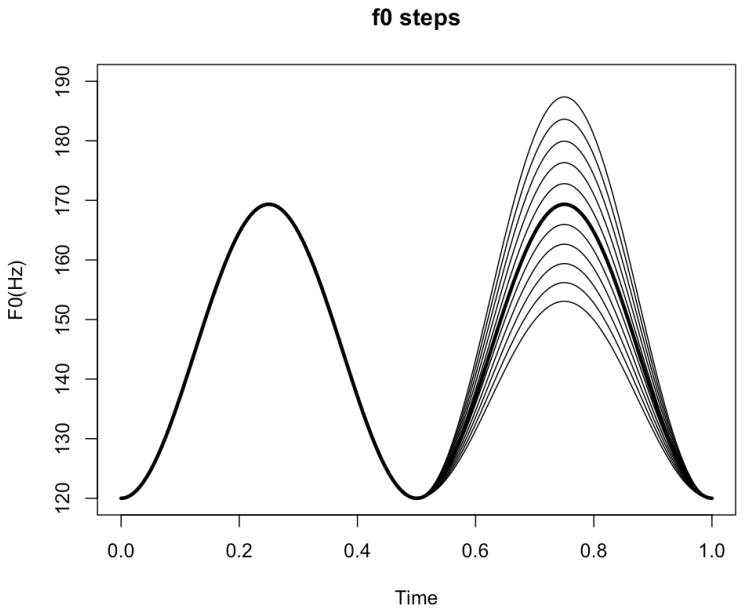


	Peak 1	Peak 2	Example
1.	Breathy	Breathy	
2.	Tense	Tense	
3.	Breathy	Tense	
4.	Tense	Breathy	

Experiment 1: Pitch Perception

f0 manipulation of stimuli

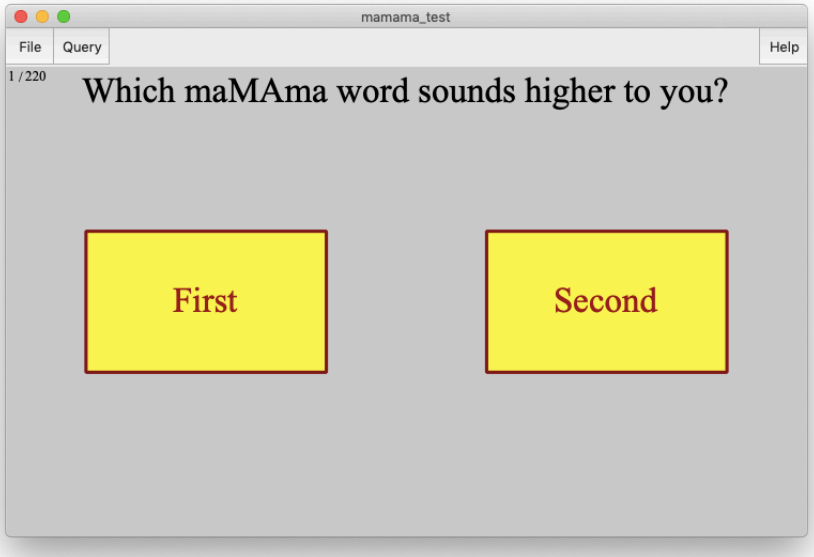
- First peak has constant peak f0
  - 169.34 Hz
- 11 f0 steps on the second peak
  - 0.5 semitones per step
  - equal as peak 1 at step 6



- 44 distinct stimuli in total after manipulations
  - 4 voice quality conditions × 11 f0 steps

Experiment 1: Pitch Perception

Task

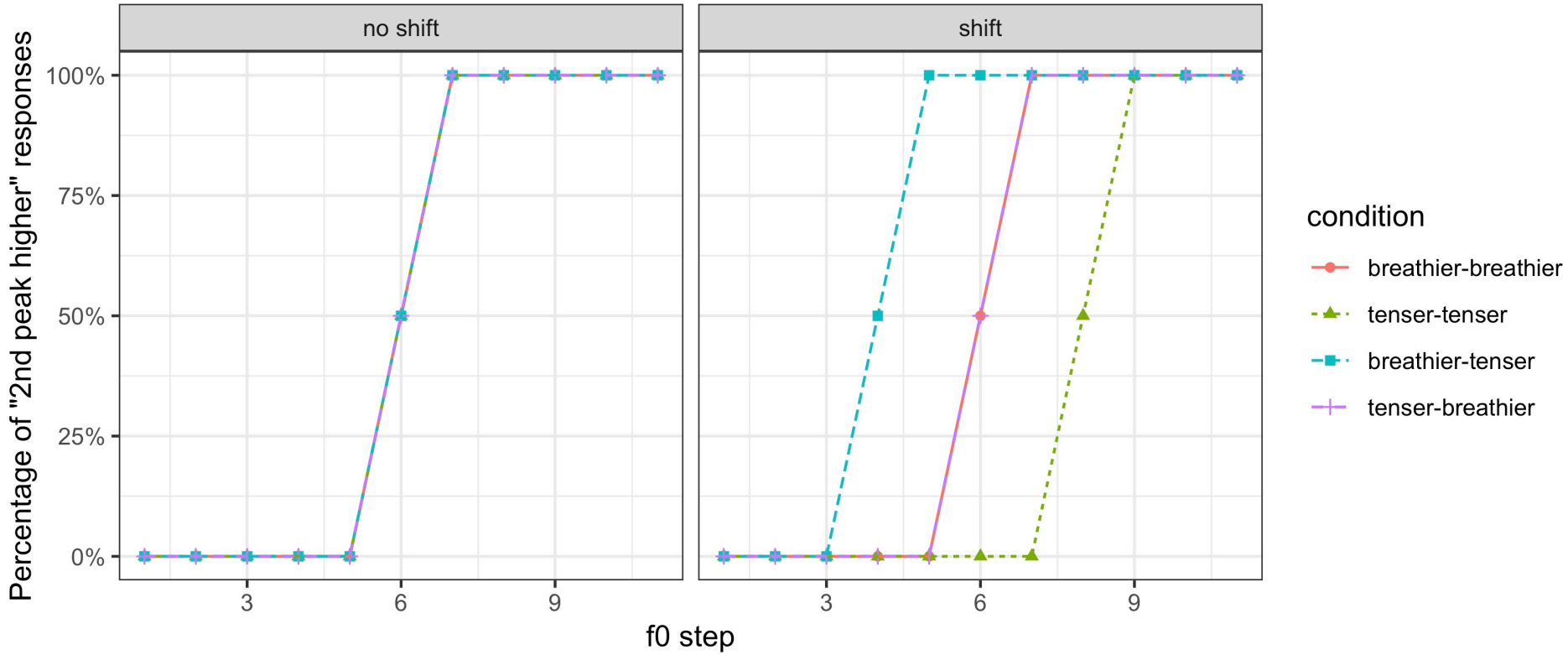


- Forced-choice classification
- Which “mamama” word sounds higher in pitch?
- 5 repetitions
- All stimuli randomized

Experiment 1: Pitch Perception

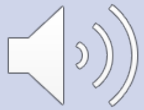
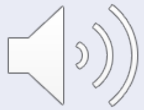
Predictions

Are listeners affected by spectral slope manipulations?



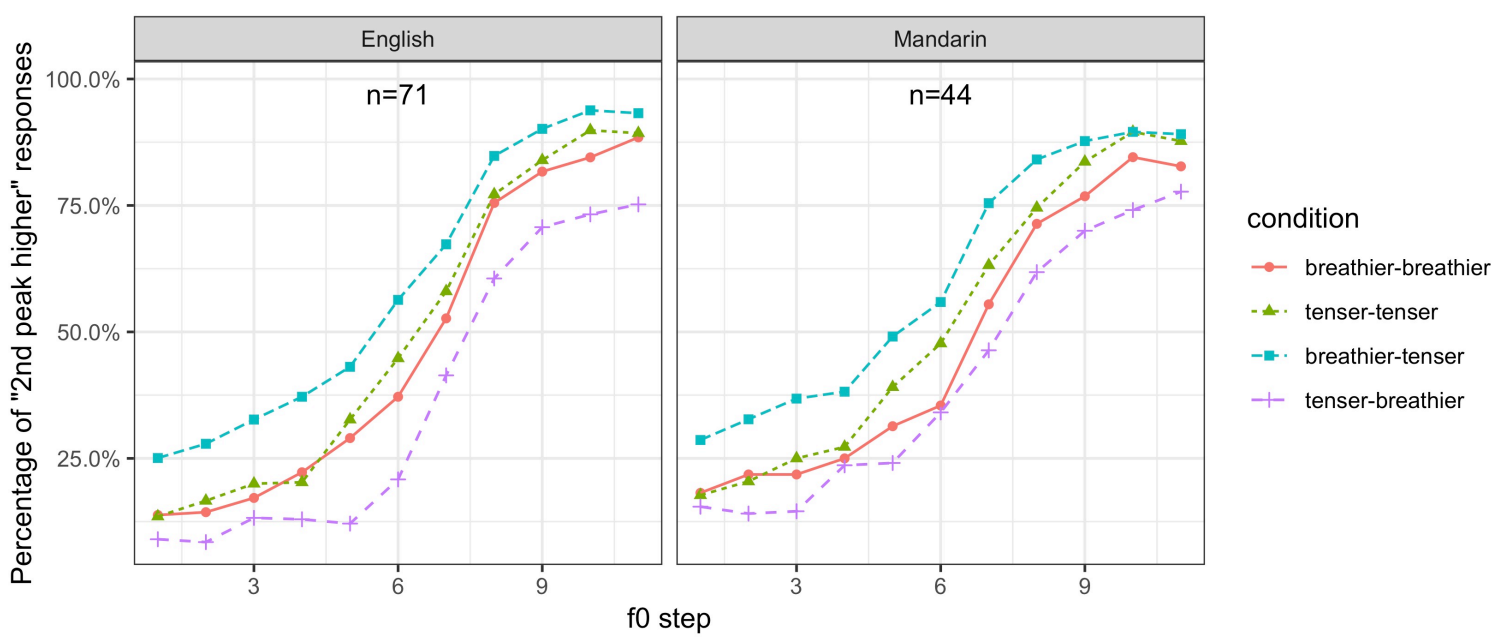
Experiment 2: Musicality

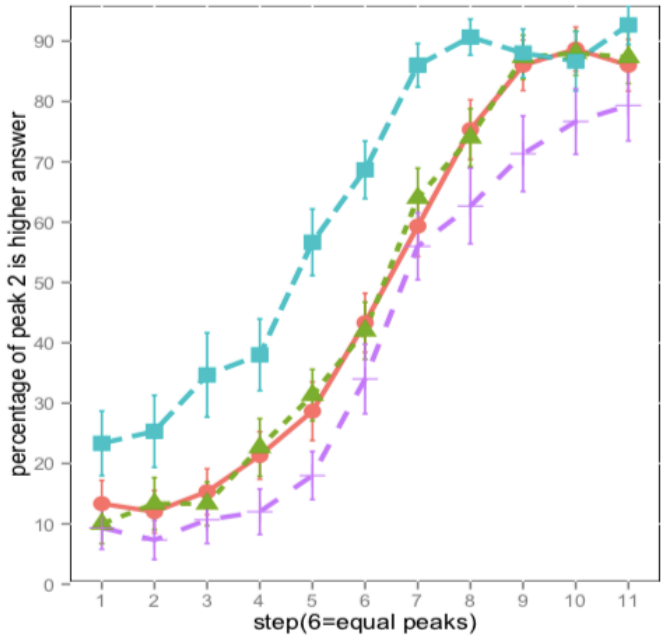
- The Montreal Battery of Evaluation of Musical Abilities (MBEMA) (Peretz et al. 2013)
- Musicality score: Percentage correct across all the tasks

Tasks	Stimuli	Number of stimuli
Melody comparison		20
Rhythm comparison		20
Memory		20

Results

Pitch Perception





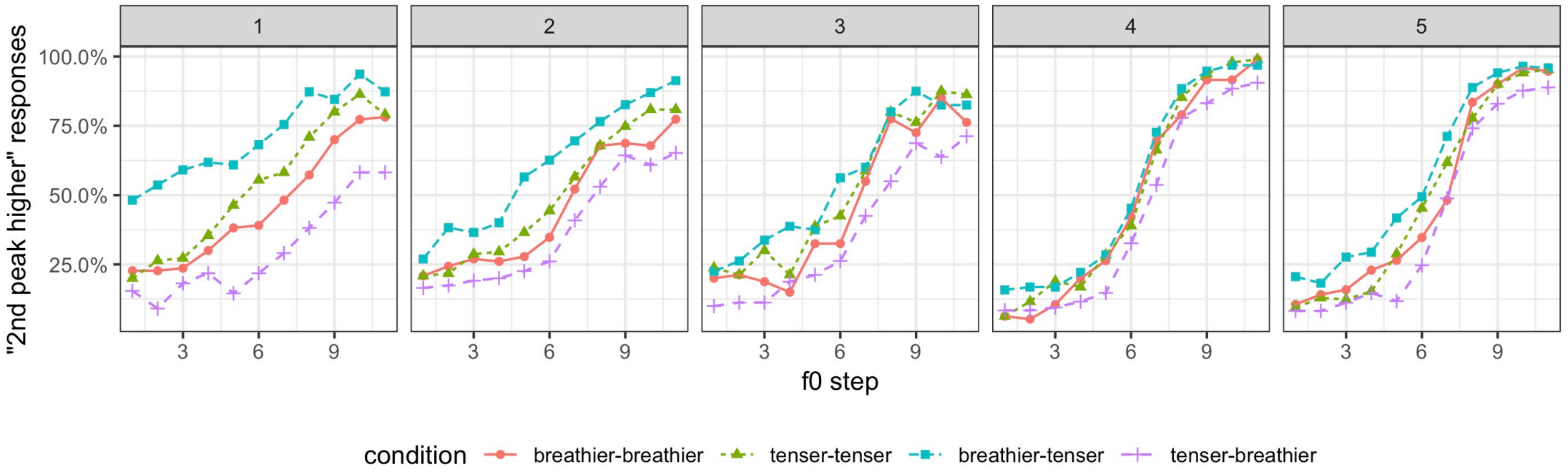
Replicated Kuang & Liberman (2018)

Musicality

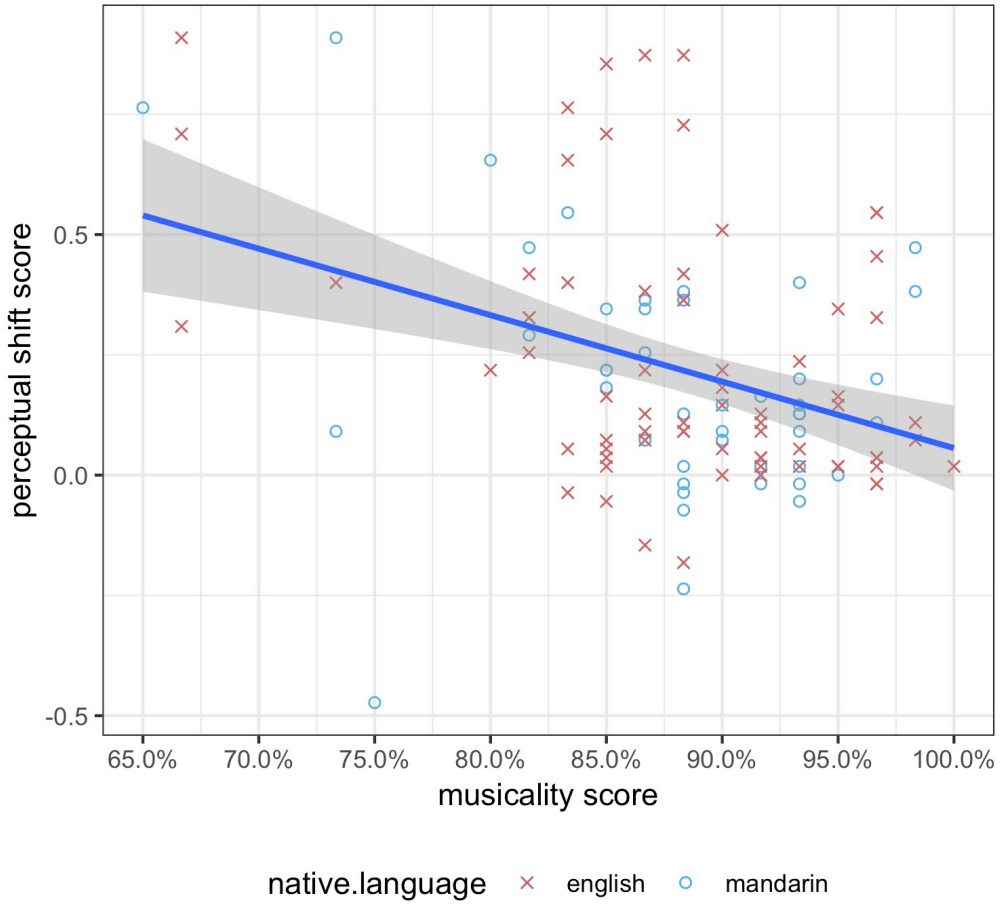
Language	N	melody	rhythm	memory	total
English	70	0.881	0.892	0.881	2.654
Mandarin	44	0.868	0.880	0.895	2.643

T-test on total scores between language groups (p = 0.784)
No significant difference in language, consistent with

Musicality and Pitch Perception by quantiles

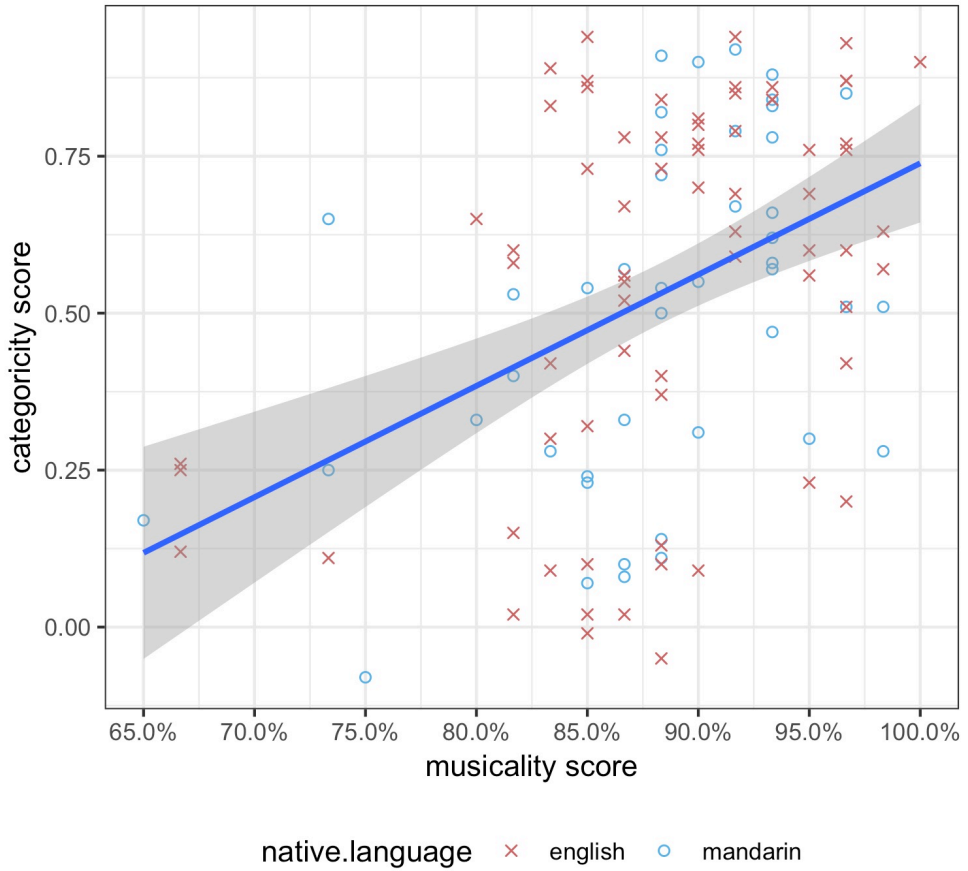


Musicality and Perceptual Shift



r = -0.3695 (t = -4.2091, df = 112, p-value < 0.001)

Musicality and f0 categoricity in perception



r = 0.4314 (t = 5.0616, df = 112, p-value < 0.001)

Discussion

Musicality and pitch perception

- Musicality is a significant predictor of pitch processing strategies
- Better musicality -> enhanced sensitivity to f0
  - More categorical perception along the f0 dimension
  - Less affected by spectral slope differences
- Implications: people who are less sensitive to f0 can rely on spectral cues in pitch perception
  - Amusic people might acquire tonal contrasts through cues co-varying with f0
  - Needs to be validated through studies with amusic population

Language background and pitch perception

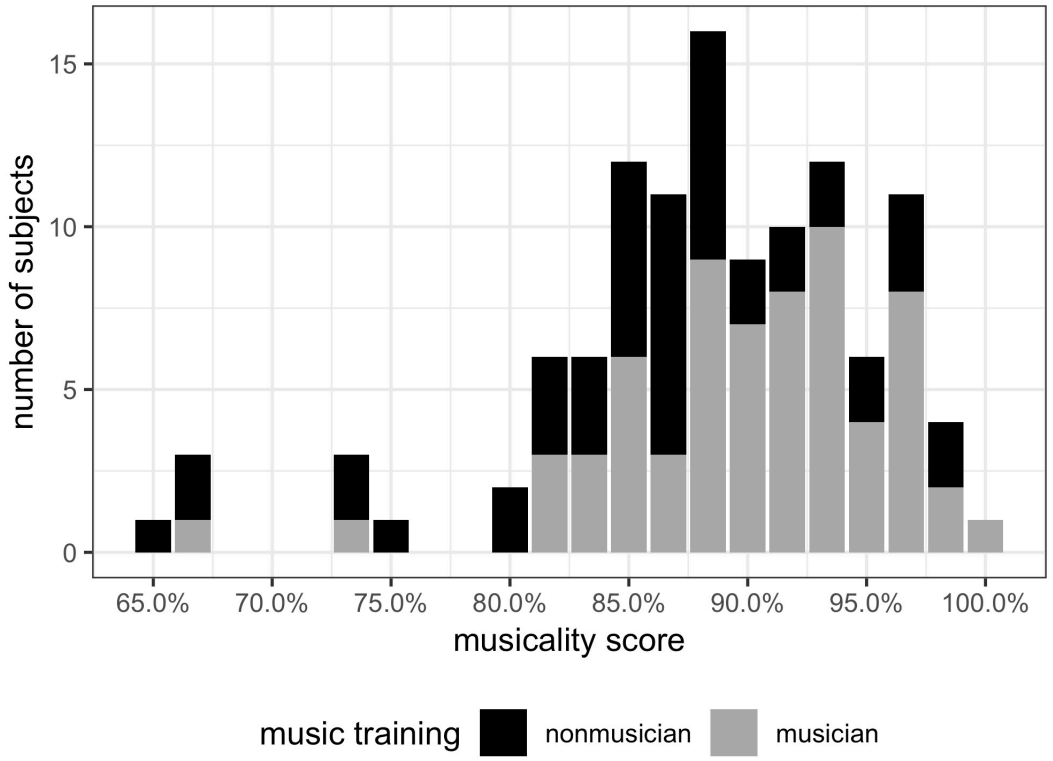
- language experience has little effect on
  - musicality scores
  - general pitch perception (Ngo et al., 2016; Zheng and Samuel, 2018)
- English and Mandarin speakers have similar strategies in pitch perception
  - Higher musicality scores -> more likely to rely on f0 in pitch perception
  - Lower musicality scores -> more likely to rely on spectral slope cue
- Maybe with multiple level tones will show language effects
  - E.g., Cantonese?

Thank you!

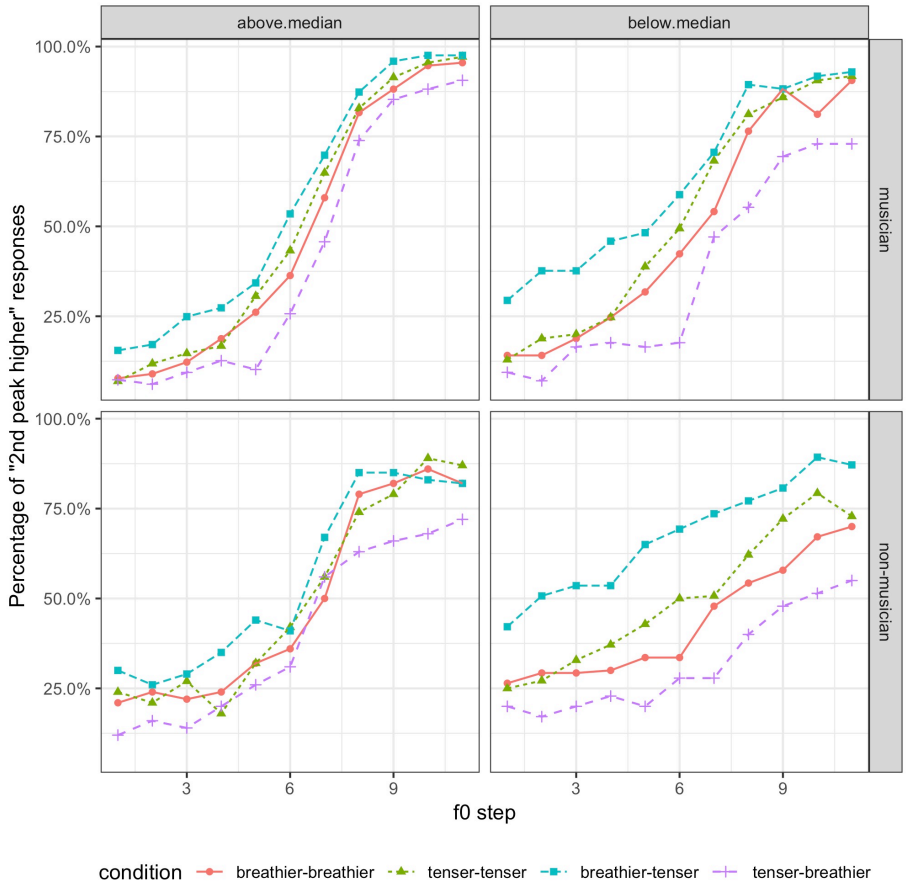
- Main effects of spectral slope for every pair of conditions for English and Mandarin speakers. Means of regression coefficients are followed by 95% highest posterior density intervals in square brackets and p-values.

		BB	TT	BT
English	**TT**	0.11 [0.03, 0.18], p = 0.006		
	BT	0.53 [0.44, 0.62], p < 0.001	0.69 [0.46, 0.89], p < 0.001	
	TB	-0.47 [-0.56, -0.38], p < 0.001	-0.97 [-1.20, -0.64], p < 0.001	-1.12 [-1.28, -0.97], p < 0.001
		BB	**TT**	**BT**
Mandarin	**TT**	0.22 [0.11, 0.34], p < 0.001		
	BT	0.62 [0.49, 0.74], p < 0.001	0.54 [0.37, 0.74], p < 0.001	
	TB	-0.29 [-0.41, -0.18], p < 0.001	-0.71 [-0.94, -0.52], p < 0.001	-1.69 [-2.24, -1.15], p < 0.001

Musicianship vs. musicality



Musicianship vs. musicality



Perceptual shift

- shift_score~musicality*language

	Estimate	Std. Error	t value	Pr(>t)
(Intercept)	1.5439	0.3694	4.18	0.0001 ***
musicality.score	-0.4938	0.1387	-3.56	0.0006 ***
native.language	-0.2630	0.6035	-0.44	0.6639
musicality.score:native.language	0.0819	0.2273	0.36	0.7194

Categoricity

- categoricity_score~musicality*language

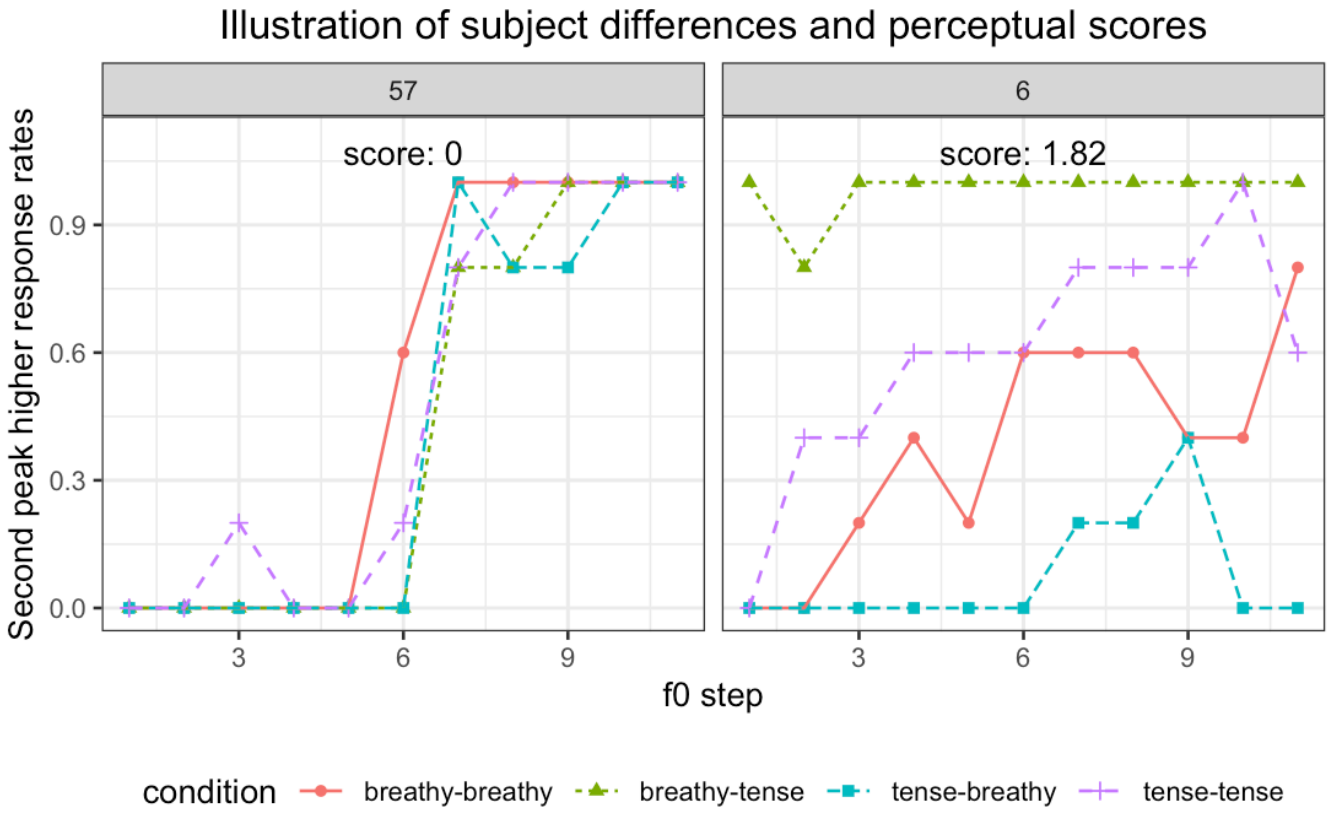
	Estimate	Std. Error	t value	Pr(>t)
(Intercept)	-1.0154	0.3948	-2.57	0.0114 *
musicality.score	0.5904	0.1483	3.98	0.0001 ***
native.language	-0.0294	0.6449	-0.05	0.9638
musicality.score:native.language	-0.0059	0.2428	-0.02	0.9805

Quantifying Perceptual Shift:

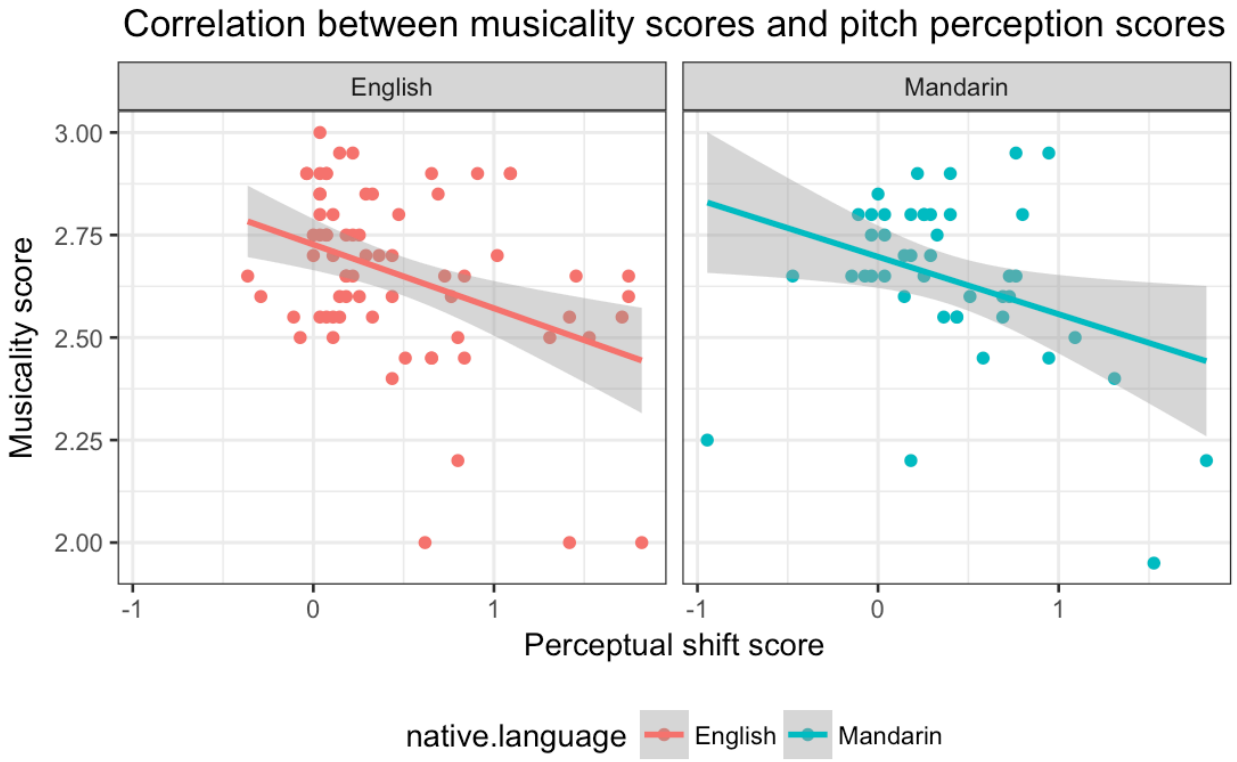
F0 dominant vs. spectral dominant

Shift = $(\overline{BT} - \overline{BB}) + (\overline{BT} - \overline{TT}) + (\overline{BB} - \overline{TB}) + (\overline{TT} - \overline{TB})$

Higher score = more shift



Perceptual Shift and Musicality



English: r = -0.392 (t = -3.51, df = 68, p = 0.00079)
Mandarin: r = -0.340 (t = -2.34, df = 42, p = 0.024)