Acoustic Correlates of Performed Sexual Orientation in Thai Media

NWAV-AP 8
Siyu Liang

UNIVERSITY of WASHINGTON

Background

- **Speech encodes social meaning**
  - Speakers index gender, sexual orientation, etc. from acoustic signals such as pitch & phonation (e.g. Gaudio 1994; Pierrehumbert et al. 2004; Zimman 2013).
- **Mixed findings for sexual orientation & acoustics**
  - Often contradictory findings across research settings (Suire et al. 2020; Holmes et al. 2024).
- **Most data are English/non-tonal**
  - Tonal languages exert extra constraints on f_0 and voice quality (Keating et al. 2023).

Motivations

- **Tonal language**
  - Thai has 5 lexical tones → pitch already utilized lexically, so any stylistic f_0 shift would be interesting and non-trivial (Osatananda & Gadavanij 2019).
- **Data**
  - TV dramas cast the same actors as gay and straight characters, holding anatomy constant.
  - Scripted yet near-natural dialogue yields comparable gay/straight speech tokens without reading-style artifacts.
- **Goal**
  - Identify which acoustic cues Thai actors manipulate to signal orientation, extending the sociophonetic landscape beyond the Anglophone focus.

Research Questions & Hypotheses

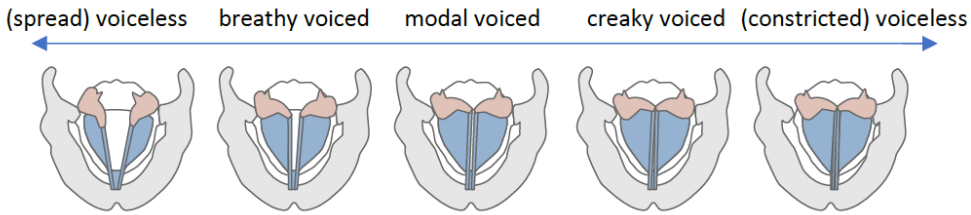
- RQ1: Do Thai actors systematically adjust pitch when portraying gay vs. straight roles?
- RQ2: Do they adjust voice quality cues, specifically breathiness?
- H1: Mean $f_0 \uparrow$ in gay-role speech (Gaudio 1994; Barbuio & Paulino 2021).
- H2: Gay-role speech $\rightarrow$ breathier phonation (Podesva 2007; Becker et al. 2022).

Mean *f*o

- **Stereotype**
  - Gay men sound "higher-pitched," lesbians "lower-pitched" (Zwicky 1997; Lakoff 1973).
- **Relevant findings**
  - Mixed: no universal effect (Gaudio 1994; Munson et al. 2006; Rendall et al. 2008).
  - Small but significant f_0 raise in some corpora/languages (Suire et al. 2020 – French; Barbuio & Paulino 2021 – Portuguese).
  - Some large-scale work shows lower mean f_0 for gay men (Holmes et al. 2024).
- **Interpretation**
  - f_0 is one cue in a constellation; its social meaning depends on listener expectations, speaker style, language norms (Vaughn 2019; Eckert 2008).

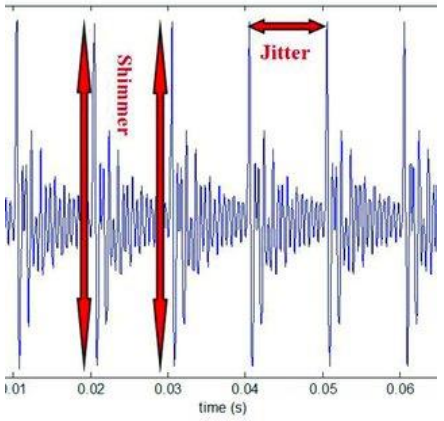
Voice Quality

- **What it is**
  - Long-term settings of laryngeal + vocal-tract settings that shape a voice's "timbre" beyond pitch and loudness (Laver 1968; Klatt & Klatt 1990).
  - A continuum from breathy to creaky (Wright et al. 2019).
- **Sociophonetic use**
  - Features such as breathiness and creak routinely index gender, sexuality, stance, and affect (Podesva 2007; Yuasa 2010).
- **Cross-linguistic use**
  - Languages differ in baseline phonation settings (Keating et al. 2023); tonal languages often tie voice quality to tone targets (Kuang 2013).



Wright et al. 2019

Jitter & Shimmer



Teixeira & Gonçalves 2014

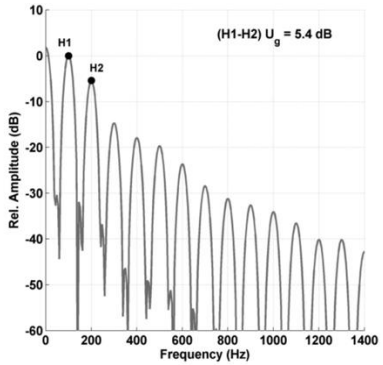
- **Definitions**
  - Periodic-based measurements of voice quality.
  - Jitter = cycle-to-cycle f_0 variability.
  - Shimmer = cycle-to-cycle amplitude variability.
  - Higher jitter & shimmer: breathier/creakier voice quality.
- **Relevant findings**
  - Higher jitter/shimmer correlate with creaky voice, typical of "straight" styles in English (Zimman 2013; Becker et al. 2022).
  - Lesbian voices showed higher jitter, interpreted as a creakier quality (Holmes et al. 2024).
- **Interpretation**
  - Jitter/Shimmer is useful acoustic cues for studying indexing of sexual orientation.
  - Could be used to distinguish creakiness/breathiness vs. modal voice, but not creakiness vs breathiness.

Harmonic–Noise Ratio (HNR)

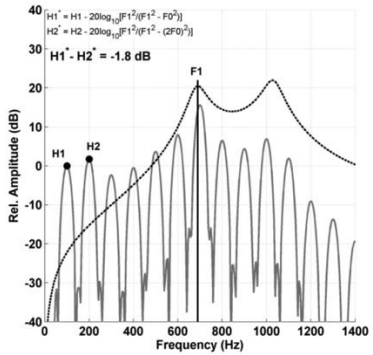
- **What it measures**
  - Ratio of periodic (harmonic) energy to noise.
  - Lower HNR = breathier/creakier voice (Hillenbrand et al. 1994).
- **Relevant Studies**
  - Gay French men had higher HNR than straight men (Suire et al. 2020).
  - Gay men showed lower HNR than straight men; lesbian women had similar trend (Holmes et al. 2024).
- **Interpretation**
  - HNR is useful in indexing of sexual orientation.
  - Similar to jitter/shimmer, it could be used to distinguish creakiness/breathiness vs. modal voice, but not creakiness vs breathiness.

Spectral Tilt (H1–H2)

- **What it measures**
  - Difference in the amplitudes of first two harmonics; larger (more positive) = breathier, smaller = tighter/creakier (Chai & Garellek 2022).
- **Link to orientation**
  - Higher H1*–H2* → more breathy registers used to index flamboyance/femininity (Podesva 2007 & Becker et al. 2022).
- **Caveat**
  - Need formant-correction: H1*–H2* (Iseli et al. 2007).
- **Interpretation**
  - A useful metric for breathiness vs creakiness.



(a)



(b)

Samlan & Story 2011

Thai

- **Central Thai:**
  - **~50 M speakers; 5 lexical tones—H, M, L, Falling, Rising (Tingsabadh & Abramson 1993).**
- **No contrastive phonation**
  - **Breathy/creaky not phonemic.**
  - **However, phonation can co-vary with tone targets (Kuang 2013; Keating et al. 2023).**
- **Glottal/nasal influence**
  - **Vowels after /h, ʔ/ often nasalize & shift voice quality—"rhinoglottophilia" (Cooke 1989; Johnson 2019).**
- **Note**
  - **Thai lets us test how orientation-linked cues fit into a tonal system where f_0 already carries lexical load.**

Z-score (Lobanov) Normalization

- **Why normalize**
  - Removes anatomical differences (vocal-tract length, habitual voice) so role/speaker effects aren't confounded (Adank et al. 2004).
- **For each speaker**
  - $z = \frac{x-\mu}{\sigma}$
  - Where x = token value, μ = speaker mean, σ = speaker SD.
- **Benefits**
  - Comparable effect sizes across speakers & measures.
  - Recommended for sociophonetic mixed-effects models (Johnson 2019; Munson & Babel 2019).
- **Limitations**
  - Loses absolute physiological info; interpret in SD units.

Data

Petch

Ter

- 2 male Thai actors
  - Petch Paopetch Charoensook (straight).
  - Ter Ratthanant Janyajirawong (gay).
- Roles portrayed in TV shows
  - Gay (*Diary of Tootsies* 2016).
  - Straight (*Social Syndrome* 2018).
- 105 vowel tokens (Table 1)
  - No background music/noise.
  - In open syllable content words at non-reducing prosodic positions, not around /h, ʔ/ (Johnson et al. 2019).
  - M. tone.

Diary of Tootsies

Social Syndrome

Table 1. Number of tokens by speaker and role.

Speaker	Role	Tokens
Petch	Gay	19
Petch	Straight	30
Ter	Gay	39
Ter	Straight	17

Data Processing & Analysis

- **Segmentation**
  - Manually using periodic voicing as cue in Praat (Boersma 2007).
- **Measurements**
  - 30 ms at midpoint using Parselmouth (Jadoul et al. 2018).
  - Mean f_0, jitter, shimmer, HNR, H1*–H2*.
- **Normalization**
  - Lobanov *z*-score.
- **Stats**
  - Linear mixed-effects using statsmodels in Python.
  - Model: *z*-Measure ~ Role + Speaker + (1 | Speaker)
    - Role: baseline = [G]ay, contrast level = [S]traight
    - Speaker: baseline = [P]etch, contrast level = [T]er
  - In plain terms: given each actor's own performative gay speech baseline, does portraying a straight character reliably effect pitch, jitter, etc.?

Results

- **Speaker effects**
  - **No significant differences once role is accounted for ($p > .80$).**
- **Role effects: relative to performed gay speech, straight speech is**
  - **Lower pitched: mean f_0 ↓ 1.008 σ.**
  - **Less breathy: H1*–H2* ↓ 0.871 σ.**
  - **Maybe creakier: Jitter & Shimmer ↑, HNR ↓.**

Table 2. Summary of fixed effects from the mixed-effects models (reference levels: *Speaker* = Petch, *Role* = Gay).

Dependent Variable	Predictor	Estimate	Std. Err	z	p-value
f_0	(Intercept)	0.612	0.902	0.679	0.497
	speaker[T]	-0.306	1.267	-0.242	0.809
	role[S]	**-1.008**	**0.183**	**-5.508**	**<0.001**
Jitter	(Intercept)	-0.419	0.973	-0.431	0.666
	speaker[T]	0.214	1.366	0.157	0.875
	role[S]	**0.676**	**0.197**	**3.424**	**<0.001**
Shimmer	(Intercept)	-0.378	0.977	-0.387	0.699
	speaker[T]	0.191	1.372	0.139	0.889
	role[S]	**0.618**	**0.198**	**3.119**	**0.002**
HNR	(Intercept)	0.531	0.936	0.568	0.570
	speaker[T]	-0.267	1.314	-0.203	0.839
	role[S]	**-0.871**	**0.190**	**-4.592**	**<0.001**
H1*-H2*	(Intercept)	0.548	0.933	0.587	0.557
	speaker[T]	-0.283	1.310	-0.216	0.829
	role[S]	**-0.872**	**0.189**	**-4.612**	**<0.001**

Interpretation

- **H1 confirmed.**
  - **A full *σ* pitch increase in performed gay speech.**
  - **In line with some cross-language findings of a higher-pitched "gay voice" (Gaudio 1994; Suire et al. 2020).**
- **H2 confirmed.**
  - **Gay roles are breathier (↑ H1*–H2*).**
  - **However, straight role speech might not be just modal but also shift towards creak (↑ Jitter/Shimmer, ↓ HNR).**

Discussion

- **Performative, not anatomical**
  - No baseline speaker effect → both actors execute the same role shift, showing orientation cues are flexible resources (Daniele et al. 2020).
- **Tone compatibility**
  - Sociophonetic variation of mean-f_0 and laryngeal manipulations occur even in the presence of Thai lexical tones (Cheng 2020; Ordin & Mennen 2017).
- **Indexical bundle**
  - Results underscore Eckert's (2008) indexical field: pitch, voice quality, and maybe other acoustic correlates work together; reading any single cue in isolation risks misinterpretation.

Summary

- **“Gay voice” is multi-cue & context-sensitive.**
- **Mean f_0 & phonation leveraged in Thai, a tonal language.**
- **Media portrayals reinforce enregistered gay style (Bell 1984; Eckert 2008).**

Questions

Bibliography

> **See paper.**