

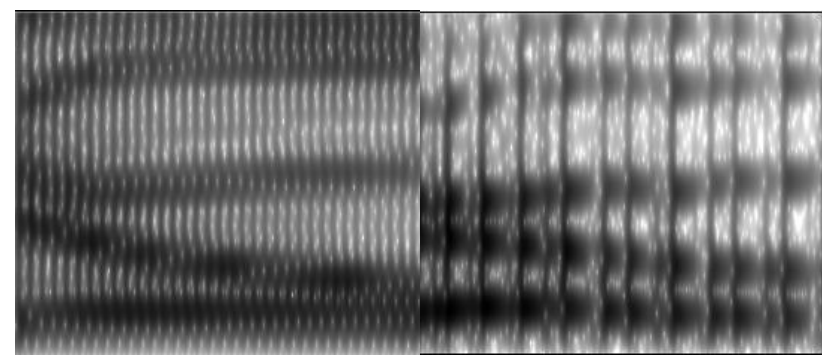
Modelling the acoustic structure of creaky voice: Mediation and factor analysis

1. BACKGROUND

The phonetics of creaky voice

Acoustics (Keating et al. 2015)

- Low f0
- Irregular voicing
- Reduced transglottal airflow



Modal and creaky /ou/ in 'show' produced by speaker F27

Articulation (Wright et al. 2019)

- Increased adductive tension
- Decreased longitudinal tension

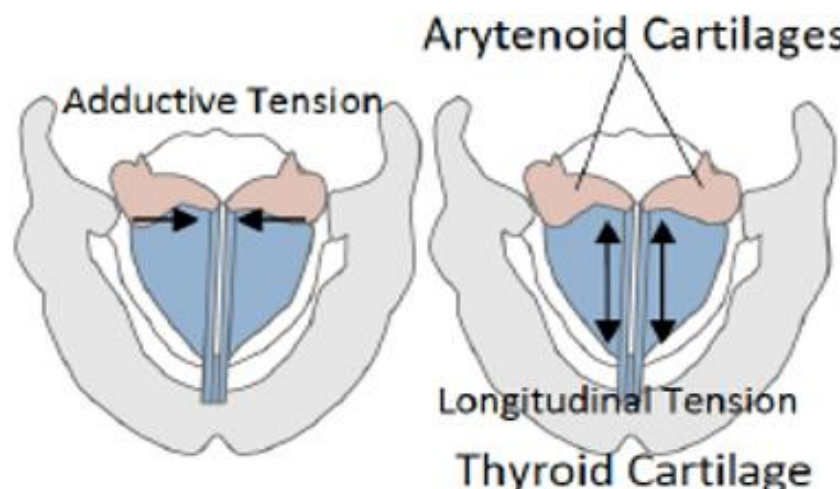
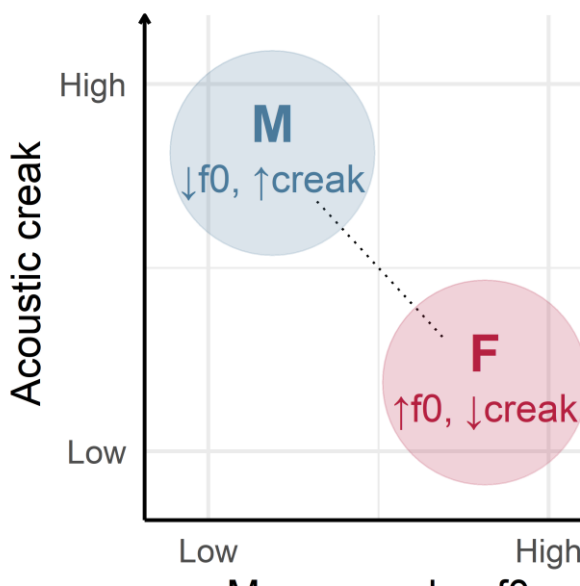


Image from Wright et al. 2019

The multidimensionality of creak

Multiple correlated & co-occurring acoustic cues characterize creak and must be interpreted jointly (Keating et al. 2023).

Creak & f0



Men's voices are both **creakier** and **lower-pitched** than women's (Brown & Sonderegger 2025, Gisladdottir et al. 2023).

... But what is the **causal structure** behind this phonetic pattern (Cohen Priva & Gleason 2020)?

Speakers maintain a **habitual f0 target** (Kearney et al. 2022) shaped by anatomy *and* social practice.

The laryngeal configuration needed to produce that **target** has **downstream** consequences for voice quality.

2. RESEARCH QUESTIONS

- How is the **multidimensional** acoustic space of creaky voice structured?
- To what extent are **gender** effects on creaky voice acoustics **caused** by underlying **f0** variation?

3. METHOD

Data

(Brown & Sonderegger 2025)

~92k vowels

- Spontaneous speech from publicly available podcasts/interviews
- 49 Canadian English-French bilinguals (24M, 25F; born 1937–2000, 21–82 years)

Acoustic cues

- f0, proportion of reliable f0 tracks, H1*–H2*, CPP, HNR < 500 Hz (also H1*–A1/A2/A3*, HNR15/25/35, jitter, shimmer in EFA)

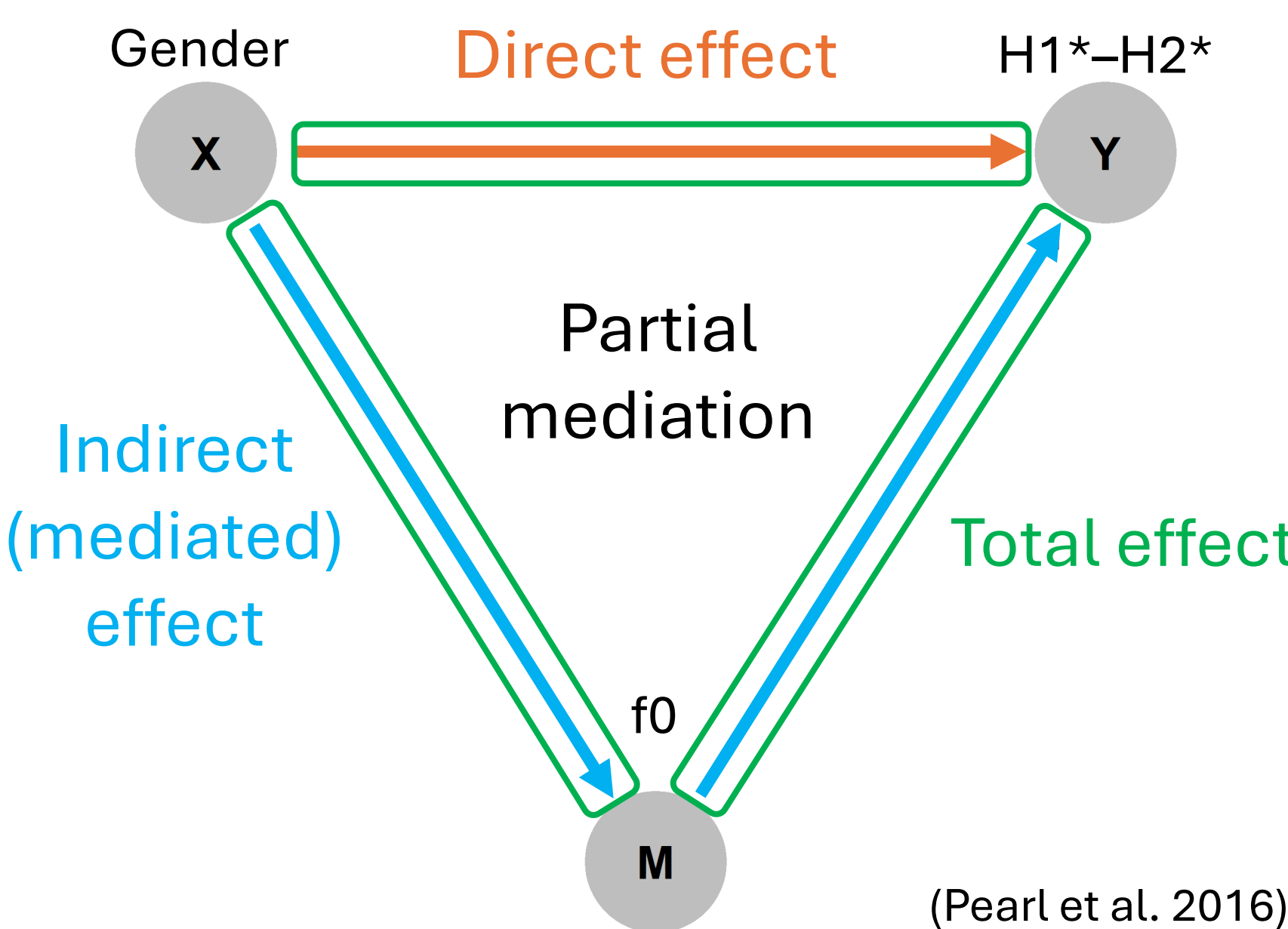
Exploratory factor analysis (EFA)

A dimensionality reduction method that models relationships between variables and latent factors

- Well-suited to variables with different scales and variances (Clopper, 2011)
- Rotation: Varimax oblique

Causal mediation analysis

A way to examine **cause-effect** relationships via observational data



(Pearl et al. 2016)

Implemented with brms (Bürkner 2017)

Total effect of gender on H1*-H2*

(1) $H1^*-H2^* \sim c(\text{Gender})$

Direct effect of gender on H1*-H2*

(2) $H1^*-H2^* \sim c'(\text{Gender}) + b(f0)$

Effect of gender on f0

(3) $f0 \sim a(\text{Gender})$

(Baron & Kenny 1986)

Proportion mediated (MacKinnon 2000):

$$\frac{\text{indirect effect}}{\text{total effect}} = \frac{\text{indirect effect}}{\text{direct} + \text{indirect effect}} = \frac{a \times b}{c' + (a \times b)}$$

ACME (Imai et al. 2010):

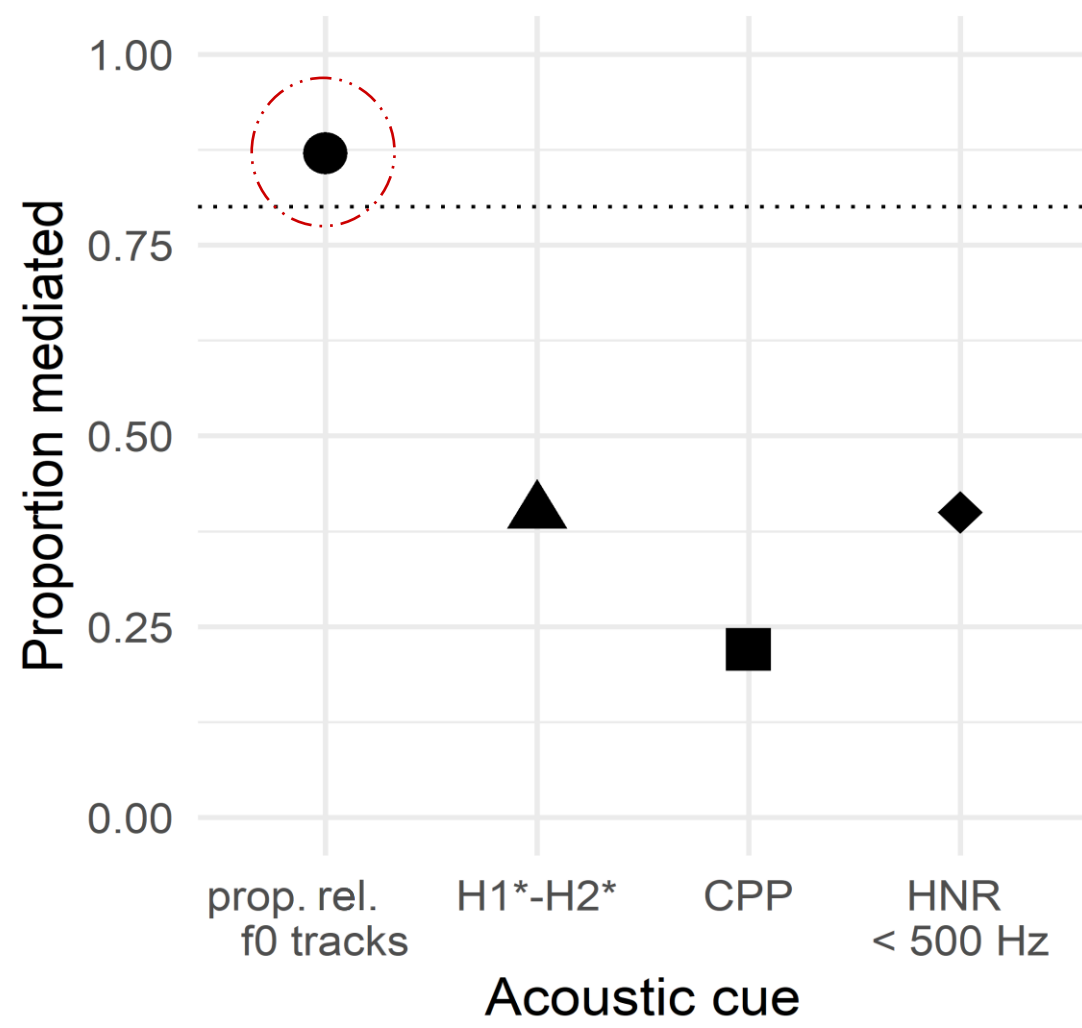
indirect effect + uncertainty

4. RESULTS

Causal mediation analysis

Gender coefficients across models		(1) Y ~ <i>c</i> Gender		(2) Y ~ <i>c'</i> Gender + bf0	(3) f0 ~ <i>a</i> Gender
Prop. reliable f0 tracks	<i>β̂</i>	-0.71*	↓86%	0.10	-72.97*
H1*-H2*		-3.94*	↓33%	-2.63	
CPP		-0.97*	↓30%	-0.68	
HNR < 500 Hz		-3.63*	↓46%	-1.95	

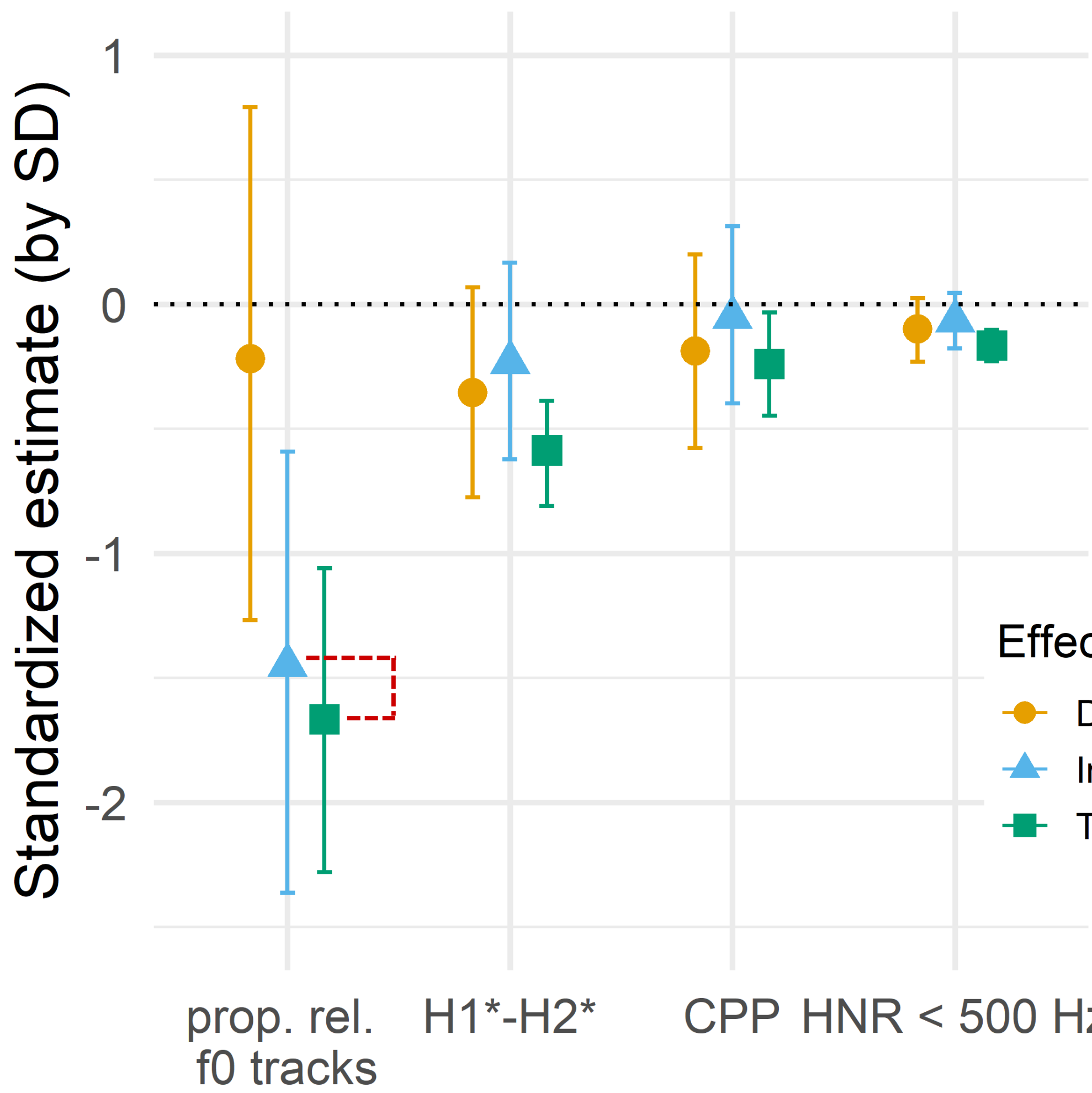
Proportion of gender effect mediated by f0
(from product of coefficients)



Mediation is stronger for **prop. reliable f0 tracks** (than H1*–H2* > HNR05 > CPP)

Note: YOB effects are not reported here. Total effects of YOB were small and inconsistent (sig. only for CPP), making mediation analysis ill-motivated for most cues. When estimated, f0 mediation was correspondingly weaker for YOB than for gender across all cues.

Estimated gender effects
(from posterior draws)

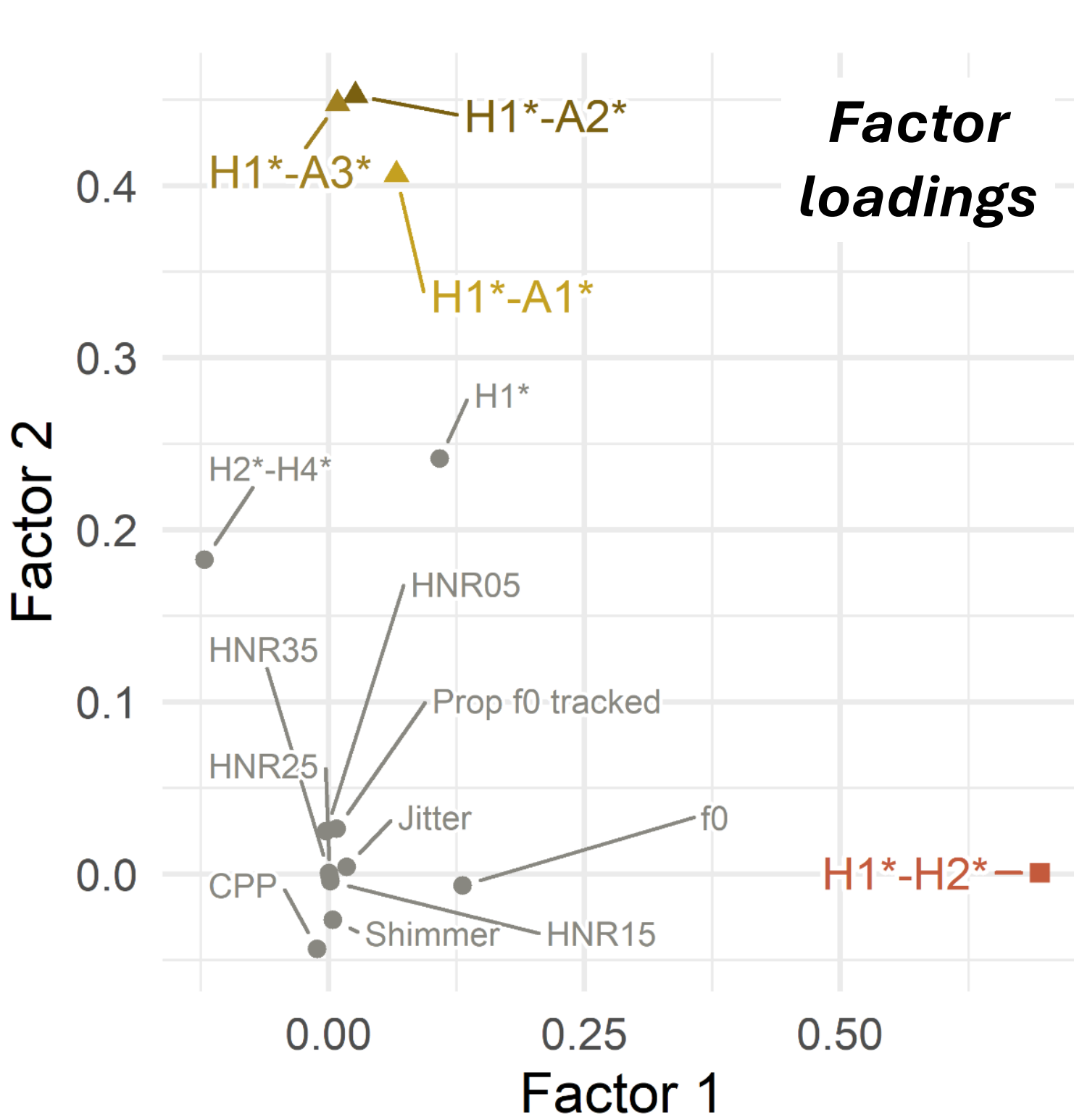


Acoustic cue

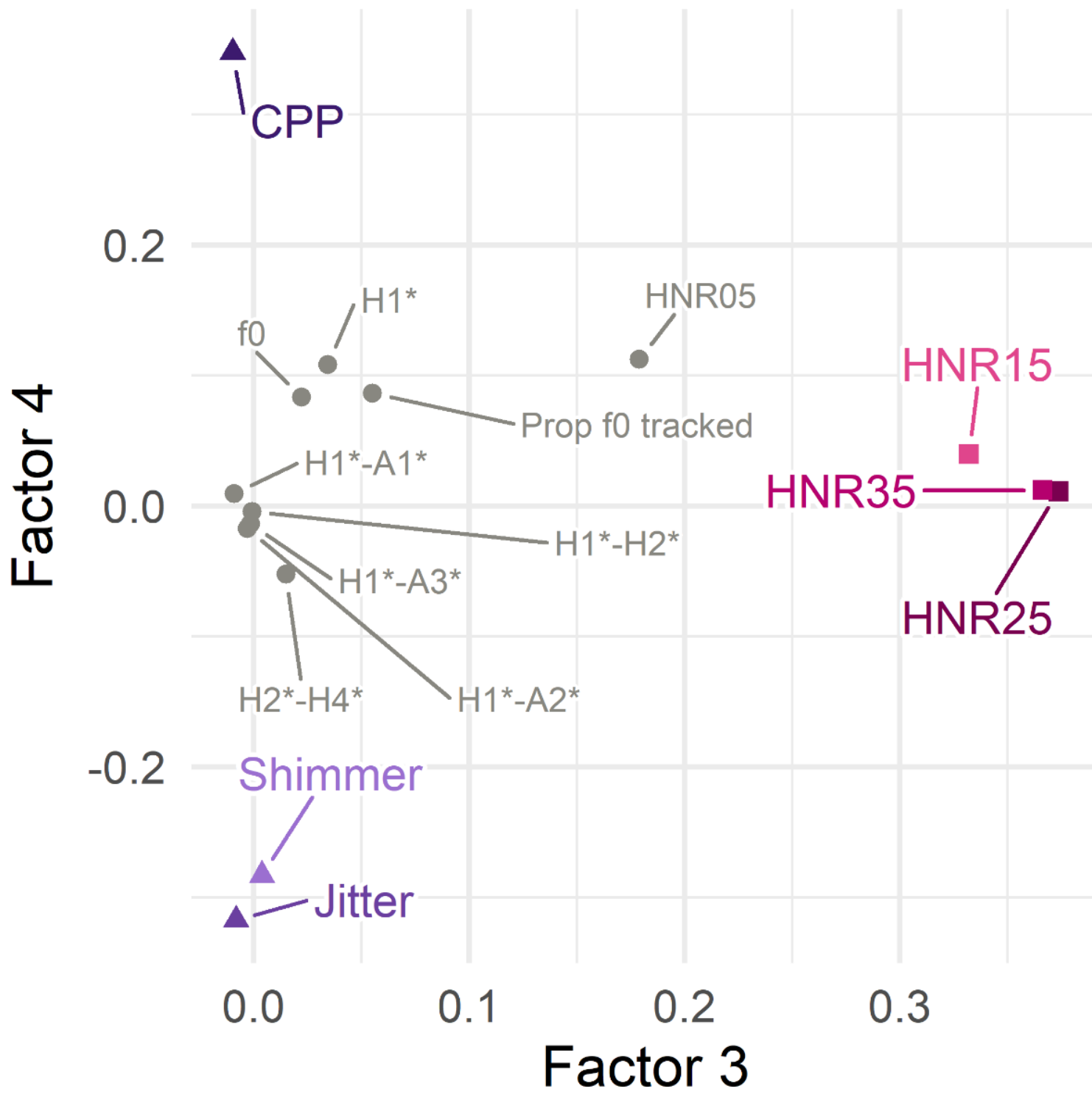
Prop. reliable f0 tracks directly reflects f0 stability → The gender effect is primarily **physiological**.

H1*–H2*, CPP & HNR05 capture additional aspects of voice quality beyond f0 → Gender may operate via other **physiological** or **social** mechanisms.

Exploratory factor analysis (EFA)

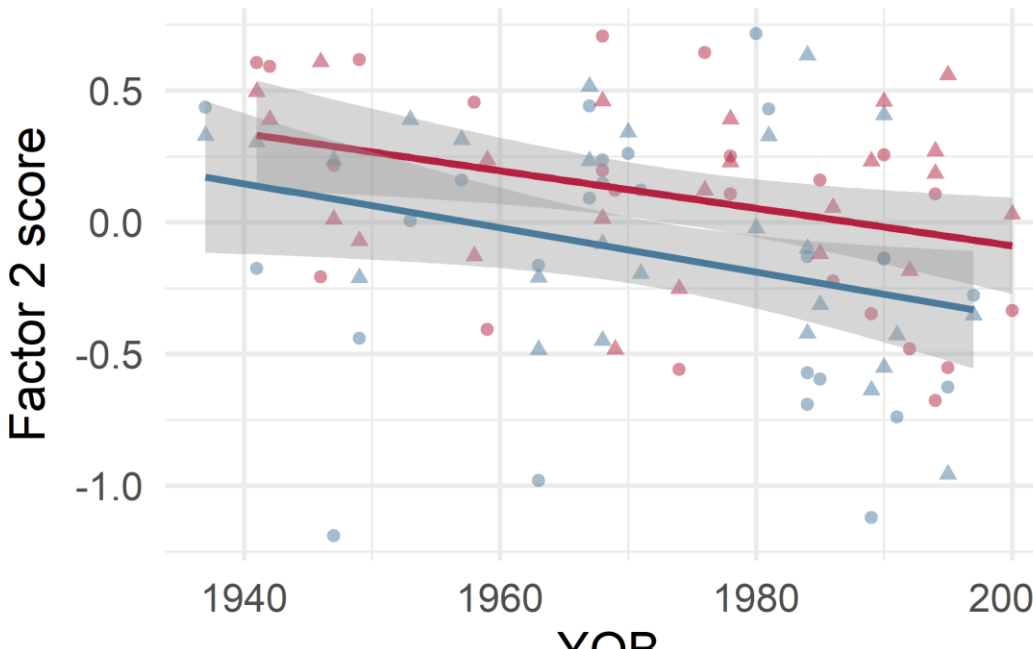
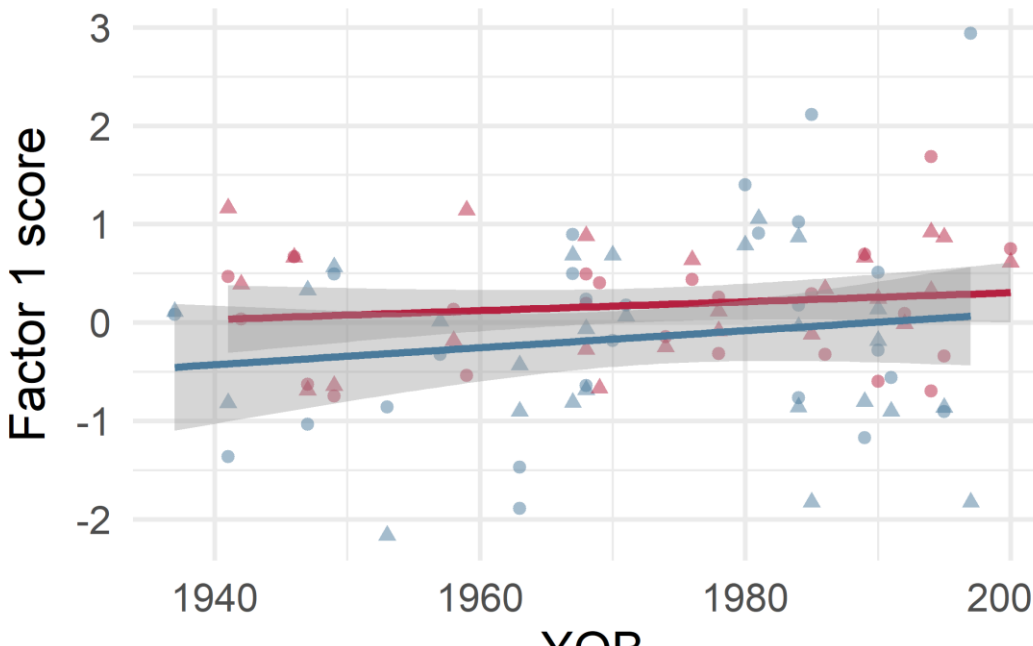


2 interpretable dimensions emerge:
spectral tilt (F1/F2) & **noise** (F3/F4)

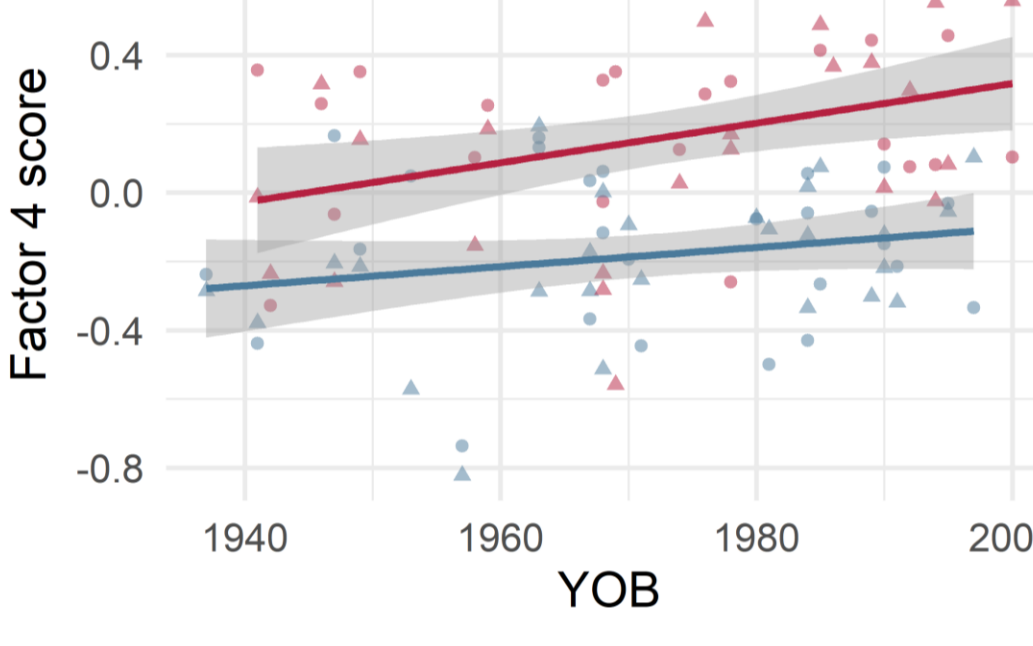
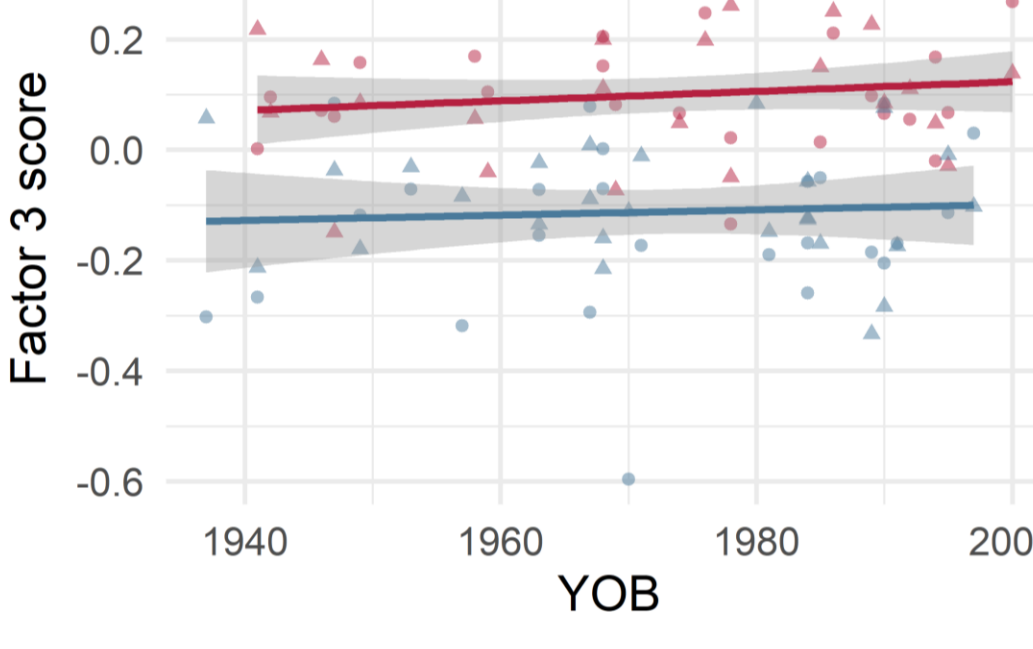


The **gender effect is not uniform** across the acoustic space, consistent with f0-mediation results (different cues, different mediation levels, different gender effect).

Factor scores



Men's voices show slightly more glottal constriction (↓ F1/F2 scores) & clearly more noise (↓ F3/F4 scores) than women's.



Language • EN • FR Gender — F — M

ACKNOWLEDGEMENTS & REFERENCES

Thank you to Morgan Sonderegger for invaluable feedback and guidance throughout this project, and to the causal inference reading group and LING683 for helpful discussion! This work is funded by FRQSC Doctoral Research Scholarship (#373124). Baron & Kenny (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J. Pers. & Soc. Psych.*, 51. Brown & Sonderegger (2025). A sociophonetic study of creaky voice across language, gender and age in Canadian English-French bilinguals. *J. Phonetics*, 172. Bürkner (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *J. Stat. Software*, 80(1). Clopper (2011). Clustering and classification methods. in *The Oxford handbook of laboratory phonology*. Cohen Priva & Gleason (2020). The causal structure of lenition: A case for the causal precedence of durational shortening. *Language*, 96. Gisladdottir et al. (2023). Sequence variants affecting voice pitch in humans. *Sci. Adv.*, 9(23). Imai et al. (2010). A general approach to causal mediation analysis. *Psych. Methods*, 15(4). Keating et al. (2015). Acoustic properties of different kinds of creaky voice. *ICPhS*. Kearney et al. (2022). Quantitatively characterizing reflexive responses to pitch perturbations. *Front. Hum. Neurosci.* 16. Keating et al. (2023). A cross-language acoustic space for vocalic phonation distinctions. *Language*, 99(2). MacKinnon (2000). Contrasts in multiple mediator models. in *Multivariate Applications in Substance Use Research*. Pearl et al. (2016). *Causal inference in statistics: A Primer*. Wright et al. (2019). Voice quality types and uses in North American English. *Anglophonia*, (27).

