The individual factor of social voice perception

Influence of bottom-up acoustics and top-down personality traits

Claudia Roswandowitz

Department of Computational Linguistics, Phonetics and Speech Sciences Group

University Zurich

claudia.roswandowitz@uzh.ch

SOCIAL PERCEPTION BY VOICE

The speaker´s voice is an important source of social information, which is essential for navigating complex social environment.

Listeners are experts in forming representations about the speakers´ identity, personality and social traits, even from brief utterances *(McAleer et al., 2014, 2017, Fernàndez Gallardo & Weiss, 2017, Lavan 2023)*.

University of
Zurich[UZH]

SOCIAL PERCEPTION BY VOICE - Mental representation of vocal information

Multi-dimensional acoustic voice space (*Latinus, Belin 2011, Lee et al., 2019,* reviews: *Maguinness, Roswandowitz, et al., 2018, Lavan, McGettigan 2023*).

Typical representation aka average voice at its center.

Acoustic distance to the typical representation modulates the perception of the speaker´s voice (*Bruckert et al., 2010, Latinus et al., 2013*).

Acoustic voice space
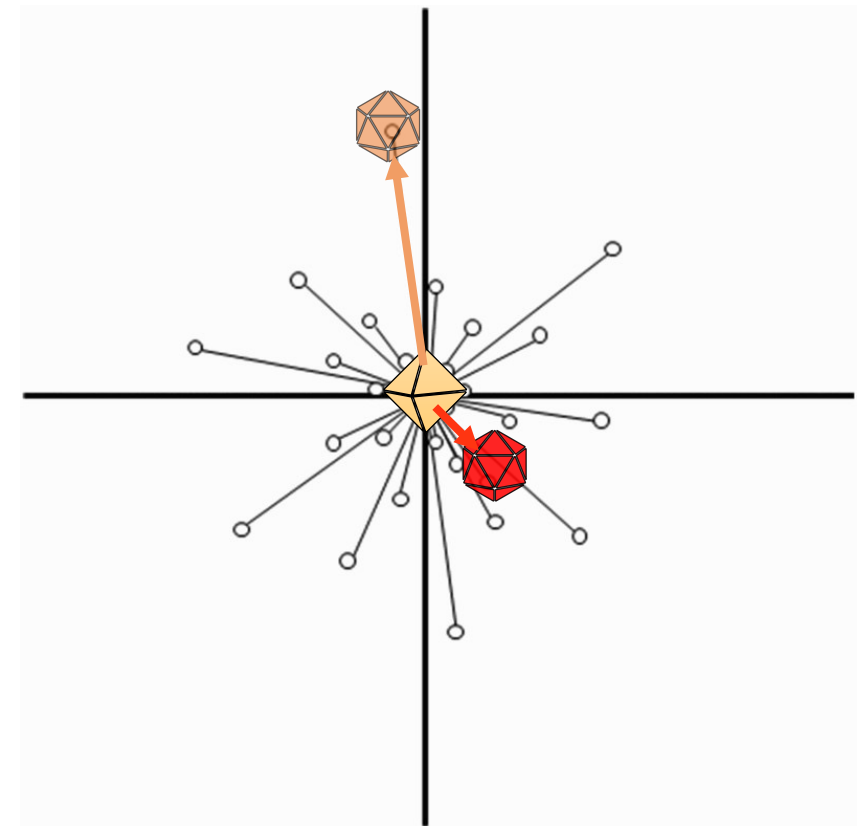
SOCIAL PERCEPTION BY VOICE – Individual differences

Everyone is unique regarding social and cultural experiences, personality, …

… social judgments from voice (and face) differ to a certain degree between listeners *(Lavan & Sutherland 2023, preprint)*.

The listener's personality or taste modulate social voice perception (*Gallardo et al., 2016, Lavan & Sutherland 2023, preprint*).

SPEAKER (Target)   ⬅   LISTENER (Perceiver)

SOCIAL PERCEPTION BY VOICE – Dyadic process

Social interaction is a dyadic process, involving a speaker and a listener - interacting and influencing each other.
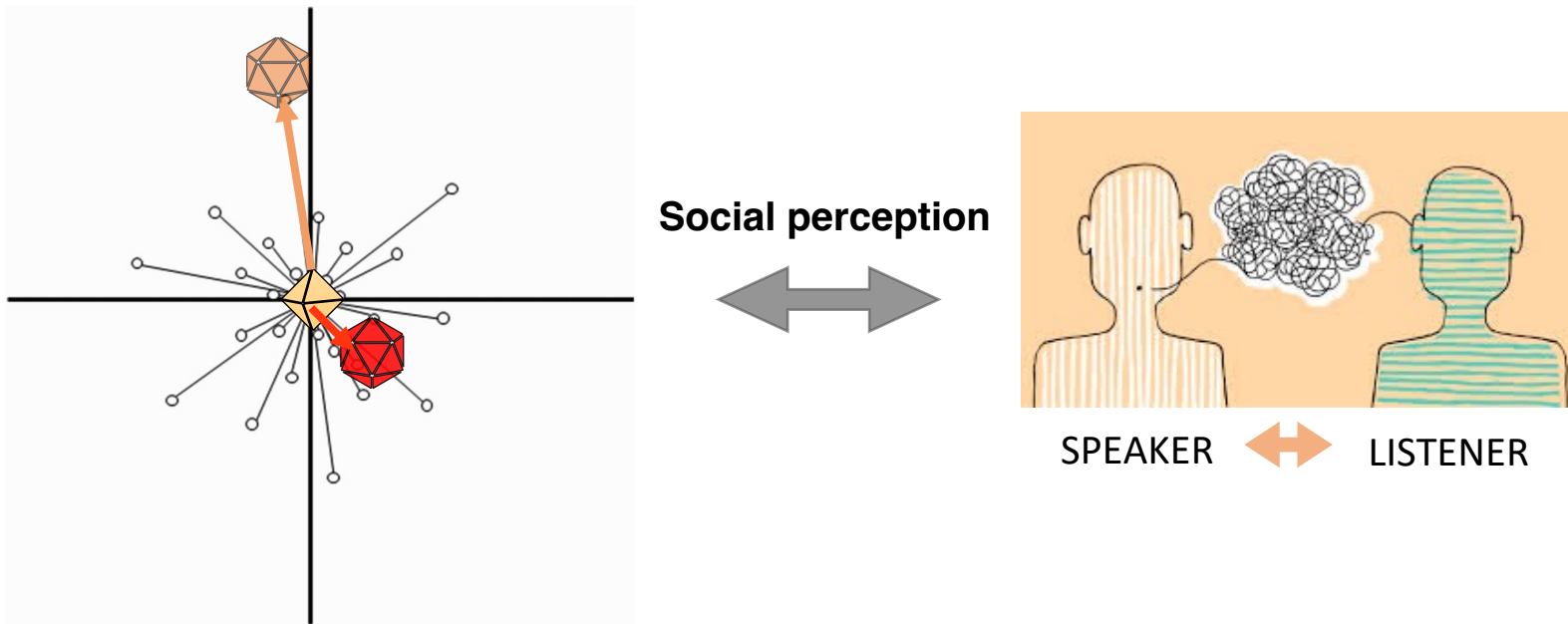
Similarity attraction paradigm (*Byrne & Nelson 1965*).

First evidence from *Gallardo et al., 2016*: people sharing personality traits rate each other's voice more likable.


SPEAKER (Target) ⬌ LISTENER (Perceiver)

Research aim

Acoustic distance between speaker's voice
and typical representation (aka average voice).

Social perception

Distance in personality between
the speaker and the listener.

SPEAKER ⬌ LISTENER

Hypothesis

As the acoustic distance between a speaker's voice and the typical representation increases, the similarity in personality between the speaker and listener becomes relevant for social perception.

Social perception
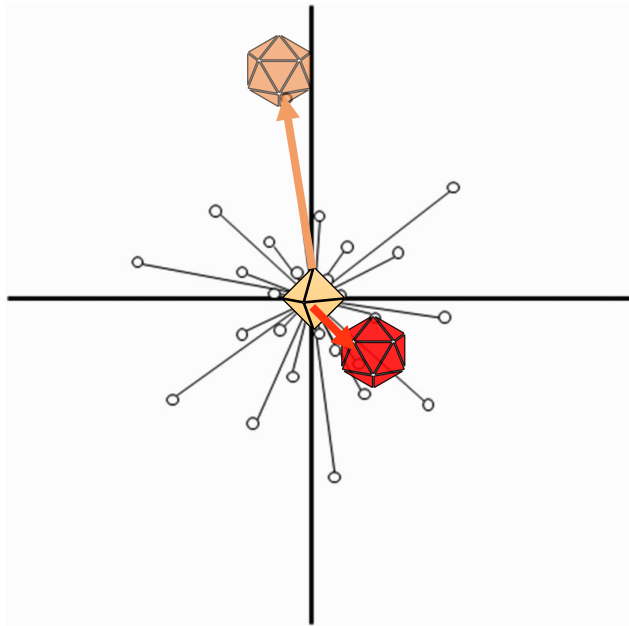
SPEAKER ⟷ LISTENER

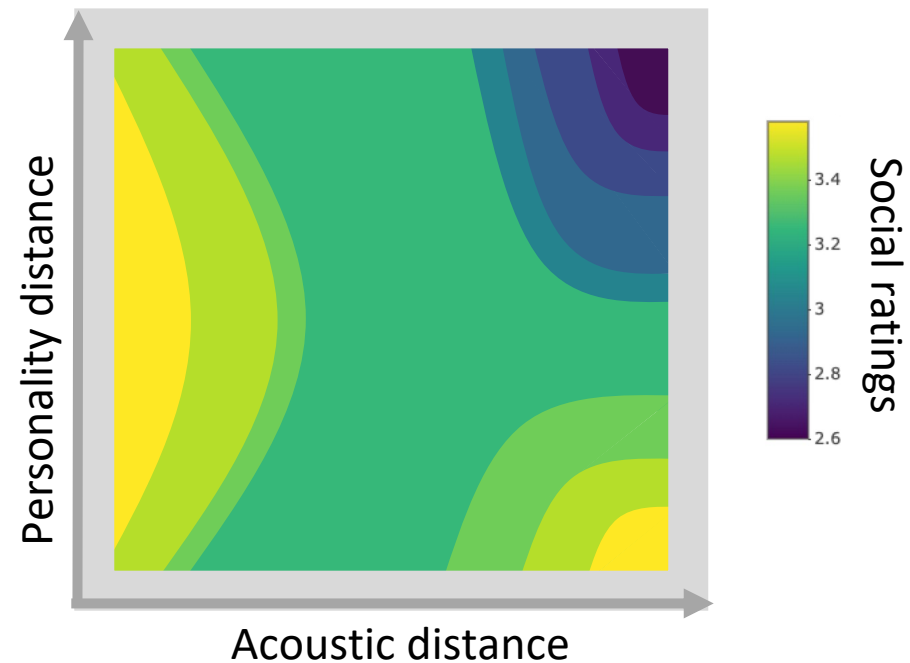With increasing acoustic distance, the acoustic signal becomes less informative.

When personalities overlap, social ratings become more positive.

Hypothesis

As the acoustic distance between a speaker's voice and the typical representation increases, the similarity in personality between the speaker and listener becomes relevant for social perception.

Stimulus Material

Auditory stimulus set

- 30 natural female voices (Nautilus Speaker Characterisation Corpus, *Gallardo & Weiss, 2018*), BFI personality scores available

- *Hallo* words extracted from semi-scripted dialogues (uttered at the dialogue beginning)

- Averaged voice (20 voice composite) generated with the STRAIGHT algorithm (Kawahara, Matsui 2003)

Natural voice   Average voice

Study Design

Web-based experiment (Gorilla)

- Part I: Personality assessment (BFI-10) – 45 questions

- Part II: Voice rating on attractiveness, trustworthiness, likability, dominance (continuous slider from 1 to 5)

    - 30 natural female voices

Participants

- N = 110 (range of age 18-35, M = 22.12, 75 female and 35 male participants)

Distance computations

Acoustic distance between each natural voice and the average voice

- 4 vocal parameters (fundamental pitch, jitter, harmonicity, and formant dispersion)

  == acoustic coordinate system for each voice

- Euclidean distance between the acoustic coordinate system of each natural and the average voice (30 acoustic ED values)
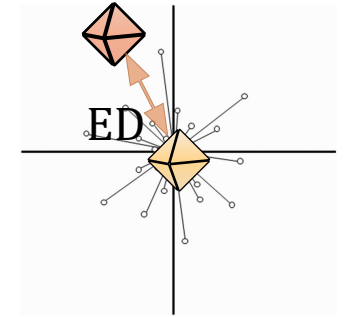
SPEAKER   LISTENER

Personality distance between each speaker-listener pair

- 5 personality scores (Extraversion, Agreeableness, Conscientiousness, Neuroticism, Openness)

  == personality coordinate system for each speaker and listener

- Euclidean distance between the personality coordinate system of each speaker-listener pair (3300 personality ED values)
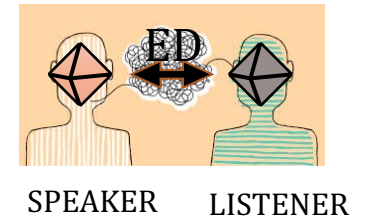
Statistics

Linear mixed effects models:

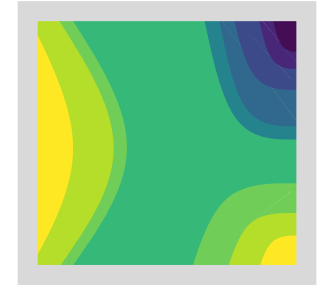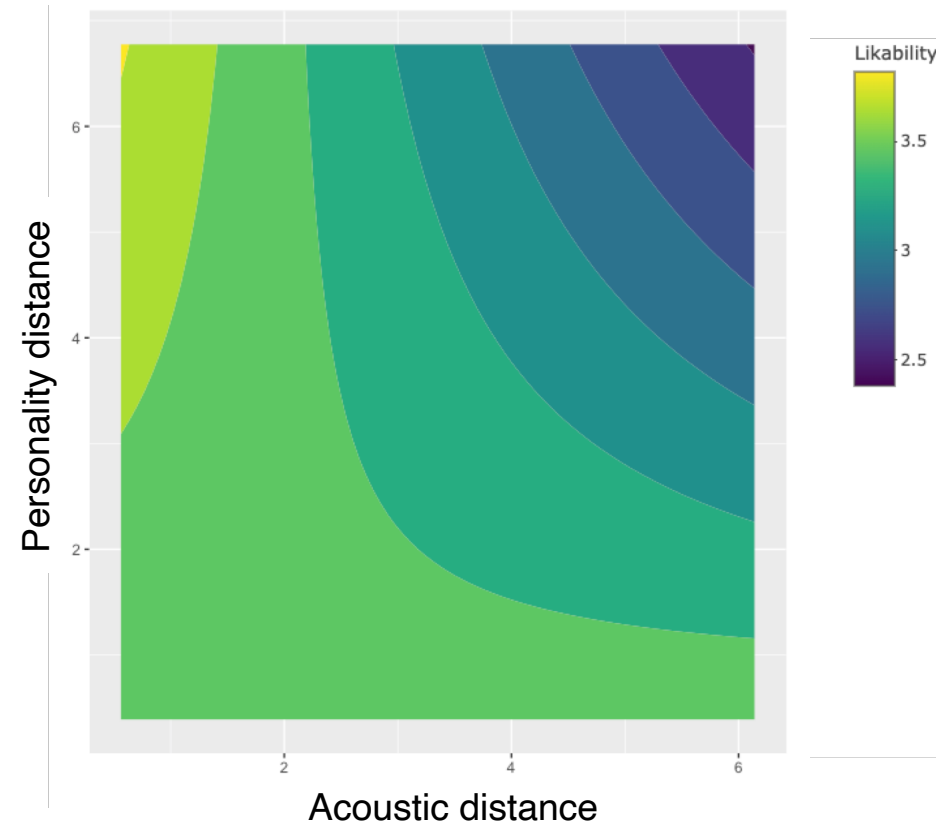Social ratings ~ acoustic distance x personality distance + (1|Participants)

Attractiveness ratings ~ acoustic distance x personality distance

	Attractiveness		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	2.47	2.21 – 2.72	**<0.001**
ED_Acoustic	0.11	0.03 – 0.19	**0.007**
ED_Personality	0.13	0.05 – 0.21	**0.001**
ED_Acoustic * ED_Personality	-0.04	-0.07 – -0.01	**0.003**
Random Effects			
σ^2	0.84		
$\tau_{00\ Participant\_ID}$	0.29		
ICC	0.25		
$N_{Participant\_ID}$	110		
Observations	3300		
Marginal R^2 / Conditional R^2	0.003 / 0.256		
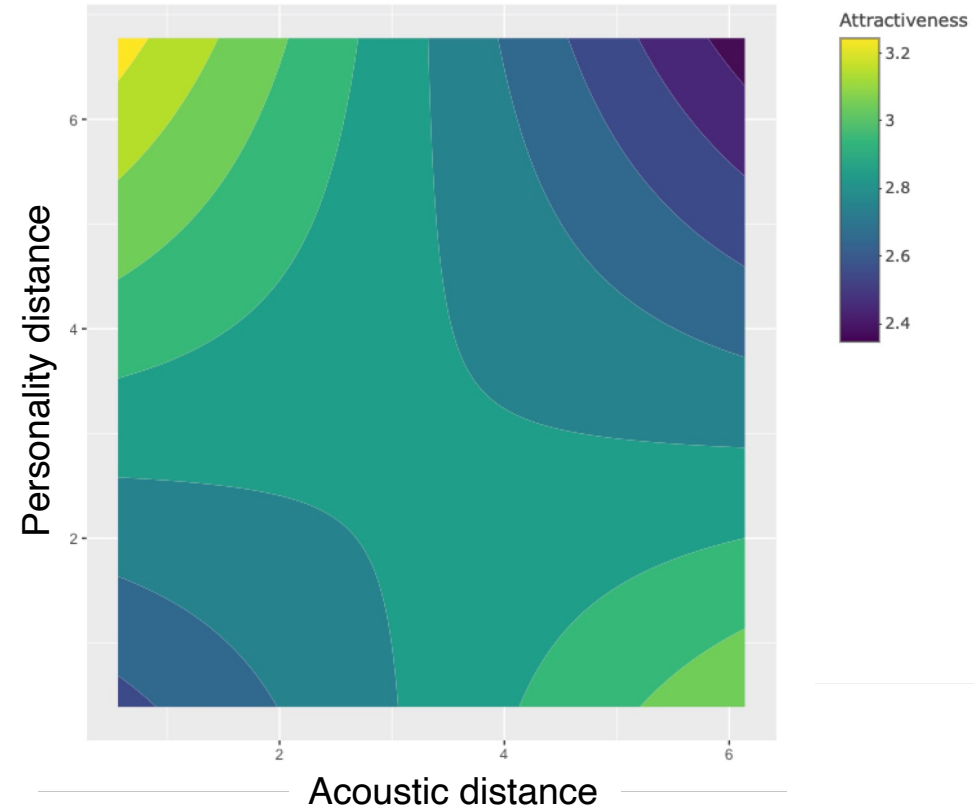
Likability ratings ~ acoustic distance x personality distance

Predictors	Likability		
	Estimates	CI	p
(Intercept)	3.40	3.16 – 3.64	**<0.001**
ED_Acoustic	0.03	-0.04 – 0.11	0.370
ED_Personality	0.08	0.01 – 0.16	**0.027**
ED_Acoustic * ED_Personality	-0.04	-0.07 – -0.02	**0.001**
Random Effects			
σ^2	0.78		
$\tau_{00\ Participant\_ID}$	0.19		
ICC	0.20		
$N_{Participant\_ID}$	110		
Observations	3300		
Marginal R^2 / Conditional R^2	0.015 / 0.210		

Trustworthiness ratings ~ acoustic distance x personality distance

Predictors	Estimates	CI	p
Trustworthiness			
(Intercept)	3.07	2.85 – 3.30	**<0.001**
ED_Acoustic	0.08	0.01 – 0.15	**0.022**
ED_Personality	0.09	0.02 – 0.16	**0.010**
ED_Acoustic * ED_Personality	-0.04	-0.06 – -0.01	**0.002**

Random Effects

σ^2	0.66
$\tau_{00\ Participant\_ID}$	0.20
ICC	0.24
$N_{Participant\_ID}$	110
Observations	3300
Marginal R^2 / Conditional R^2	0.004 / 0.239
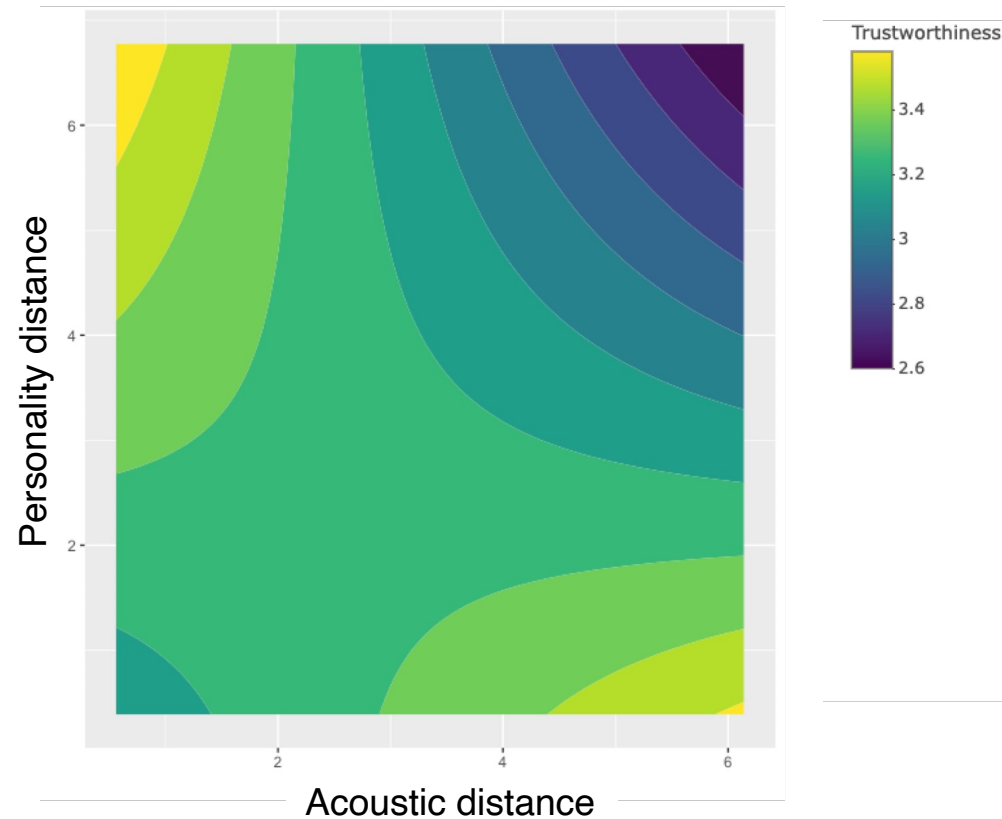
Dominance ratings ~ acoustic distance x personality distance
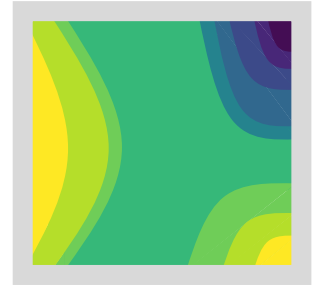
Predictors	Estimates	CI	p
		Dominance	
(Intercept)	2.61	2.39 – 2.84	**<0.001**
ED_Acoustic	-0.05	-0.12 – 0.02	0.183
ED_Personality	-0.03	-0.09 – 0.04	0.435
ED_Acoustic * ED_Personality	0.01	-0.02 – 0.03	0.627
Random Effects			
σ^2	0.64		
τ_{00} Participant_ID	0.25		
ICC	0.28		
N Participant_ID	110		
Observations	3300		
Marginal R^2 / Conditional R^2	0.002 / 0.277		

Summary

H: As the acoustic distance between a speaker's voice and the typical representation increases, the similarity in personality between the speaker and listener becomes relevant for social perception.
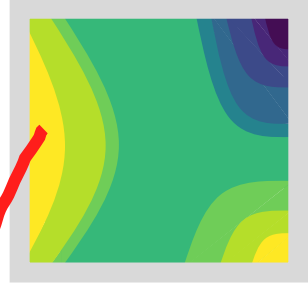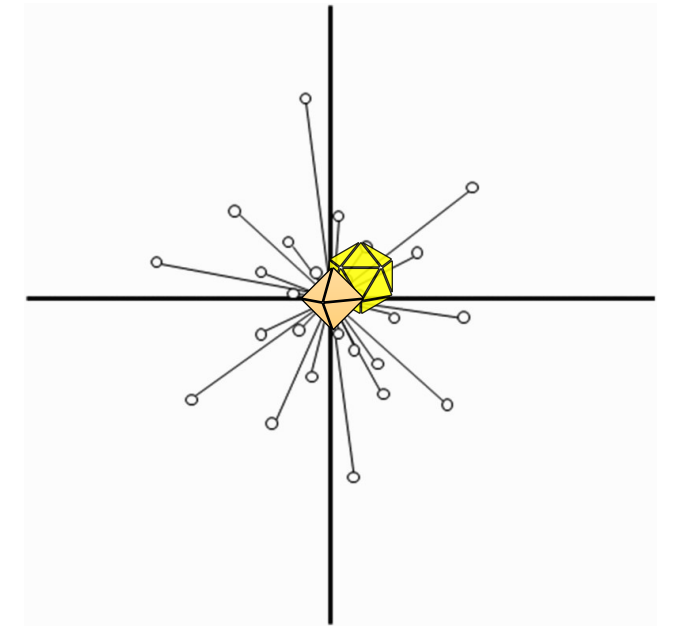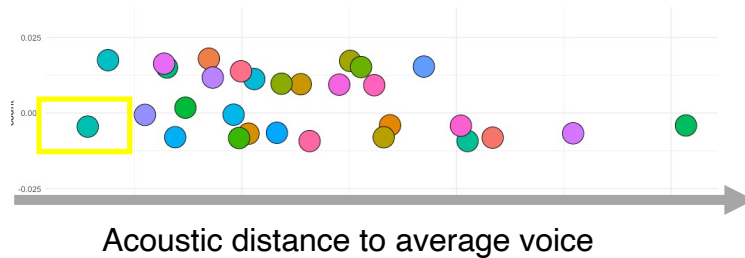
- Large acoustic distance: personality distance clearly modulates social ratings, with higher social ratings when personality distance is smaller (expected)

- Small acoustic distance: personality distance marginally modulates social ratings, and if so higher social ratings when personality distance is larger (unexpected)

Are effects generalisable?

Acoustic distance between speakers' voices and natural reference voice

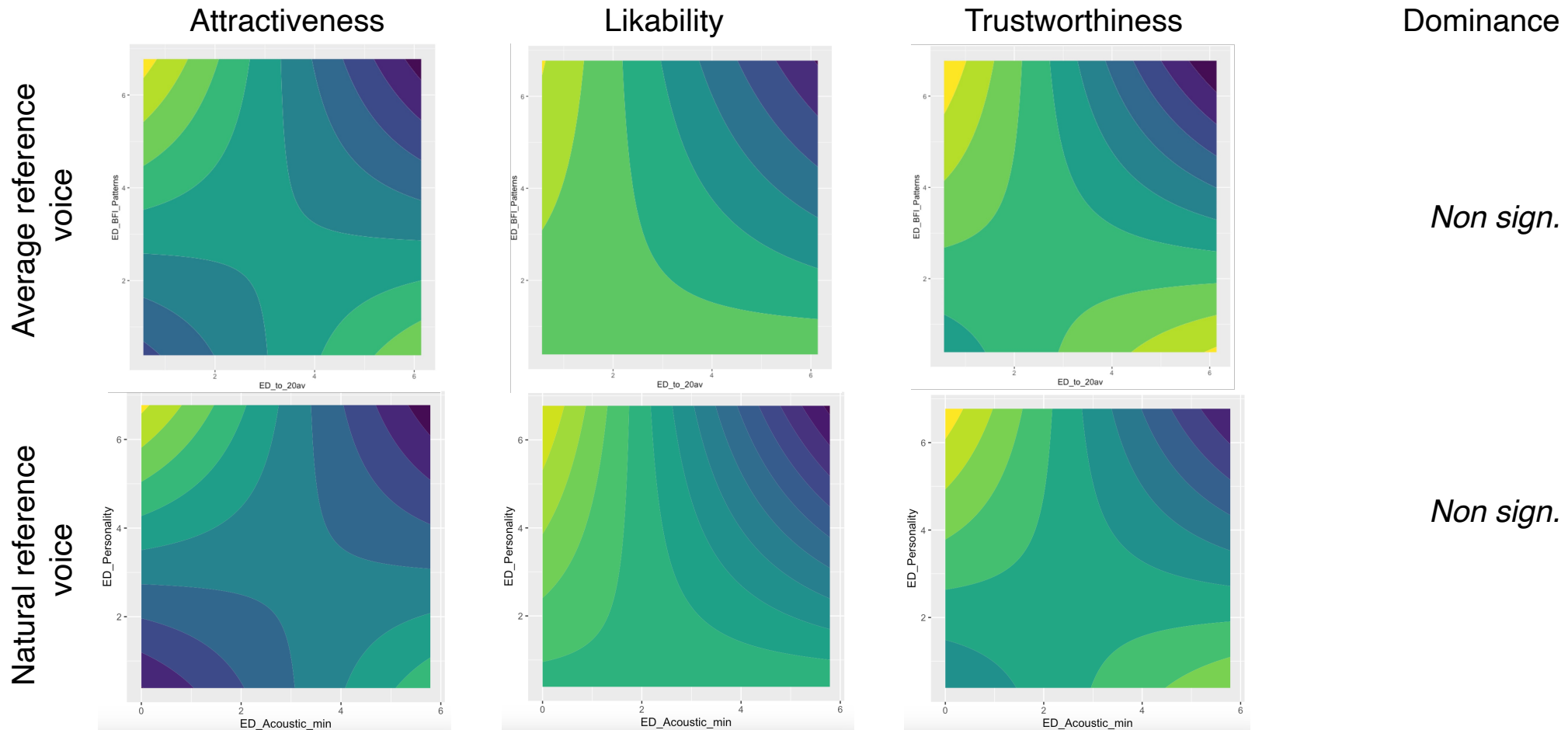- Natural voice with minimum distance to the average voice

Acoustic distance to average voice

Hypothesis

- Minimum distance voice = effect should replicate

Are effects generalisable?

| Attractiveness | Likability | Trustworthiness | Dominance |

Average reference voice — *Non sign.*

Natural reference voice — *Non sign.*

General Conclusion

Our findings suggest that not only the acoustic information of the speaker, but also the similarity between the speaker's and listener's personality influences how we socially perceive our communication partner.

Emphasises the individual and dyadic character of social voice perception.

Our statistical models leave some variability unexplained - likely that also other top-down information form social perceptions, e.g. individual social learning/stereotype activation, cognitive state (motivation, arousal).

SPEAKER (Target) ⬌ LISTENER (Perceiver)

Outlook

Replicate findings in natural and dyadic vocal communication situations

University of Zurich UZH

Thank you very much for your attention!

And thanks to my collaborators:

Selma Bruggisser
Department of Psychology
University Zurich, Switzerland

Prof. Alexis Hervais-Adelman
Department of Basic Neuroscience
Geneva University, Switzerland

Prof. Volker Dellwo
Department of Computational Linguistics,
University Zurich, Switzerland