

# How do we feel about /x/?

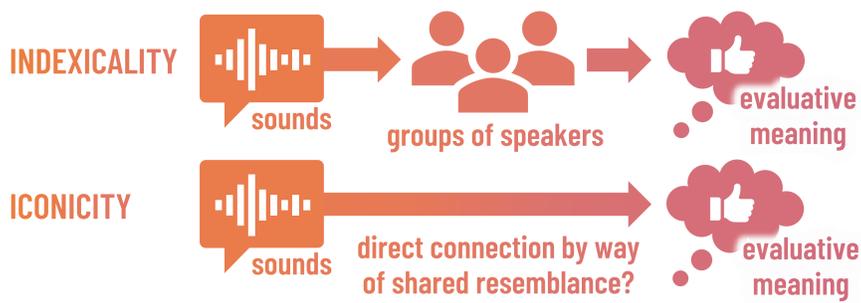
The role of iconicity and indexicality for language attitudes

French sounds beautiful and romantic.

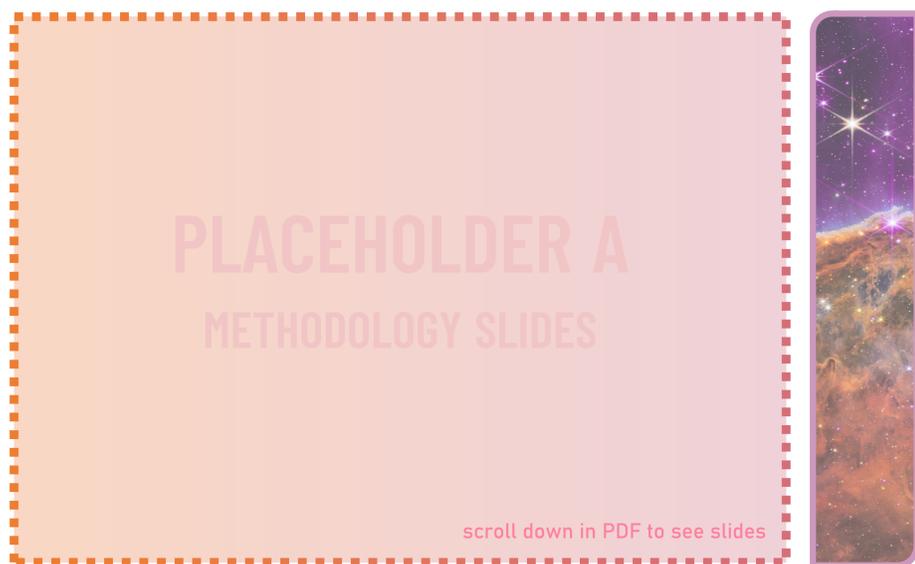
German sounds harsh and aggressive.

language attitudes <sup>1-28</sup>

But why? <sup>29-33</sup> Indexicality <sup>34-37</sup> versus iconicity <sup>38-51</sup>



Methodology <sup>52-54</sup> Creation of new languages

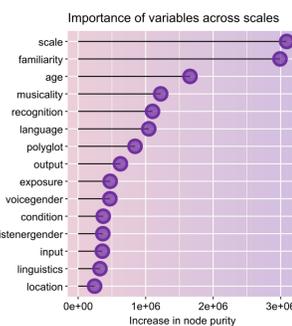
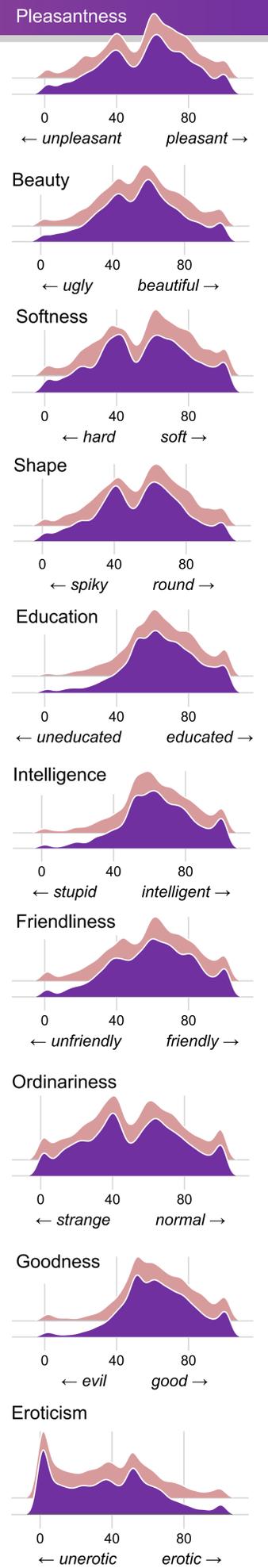
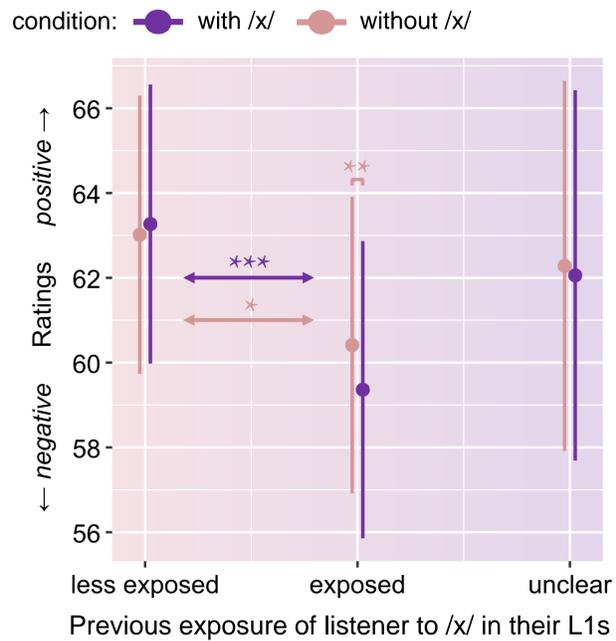


- 1 software development new pseudotext generator called SSPG
- 2 text generation with and without /x/, other properties the same
- 3 speech synthesis with different languages, engines, voices <sup>55</sup>
- 4 rating study semantic differential scale sliders, 501 listeners <sup>56-57</sup>
- 5 modeling mixed models with lme4 <sup>58-59</sup>, randomForest <sup>60</sup> in R <sup>61</sup>

In short You rate /x/ worse only if you're exposed to it, plus:

- 🏠 Language sounds worse to you if you perceive it as being less familiar. <sup>62-64, cf. 30, 32, 33</sup>
- ♂️ You rate language worse if you're male. Gender of voice also affects ratings. <sup>20, 65-70</sup>
- 🌐 You rate language worse if you feel like it's similar to a specific language group, such as those associated with "harshness", like 'Middle East' or 'Germanic'. <sup>71-72, cf. 30, 32</sup>

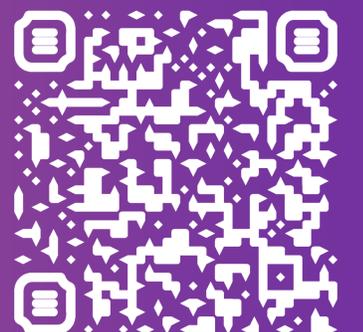
Results All semantic scales



/x/ doesn't sound as terrible as you'd think—unless sociocultural factors make it sound bad.

This supports the idea of iconicity being embedded into and constrained by a deeply indexical system. <sup>35, 73-81</sup>

REFERENCES, DATA, SCRIPTS



## 2 Text generation with the Sonority-Sensitive Pseudotext Generator SSPG

Speaker	Control			Stimulus		
	1	2	3	1	2	3
Mean sonority 1-17	10.91	10.94	10.92	10.94	10.95	10.92
Consonants %	47.62	46.9	47.26	47.33	47.44	47.13
Obstruents %	25.32	23.97	23.73	24.52	26.07	26.43
Vowels %	52.4	53.1	52.7	52.7	52.6	52.9
Voicing %	79.44	80.37	79.51	79.53	78.21	78.28
Syllable structures	CV, V					
Syllable weights 0-1	0.9, 0.1	0.9, 0.1	0.9, 0.1	0.9, 0.1	0.9, 0.1	0.9, 0.1
Number of words in text	100	100	100	100	100	100
Max number of syllables	4	4	4	4	4	4
Sound inventory	base	base	base	base, x	base, x	base, x

Overview of properties of convariety texts in both conditions, stimulus and control, after verification. The sound inventory base refers to the baseline inventory /p, t, k, b, m, n, s, l, w, j, i, u, a, e, o/. The crucial difference between control and stimulus conditions is the /x/ added to the baseline inventory.

## 2 Text generation with the Sonority-Sensitive Pseudotext Generator SSPG

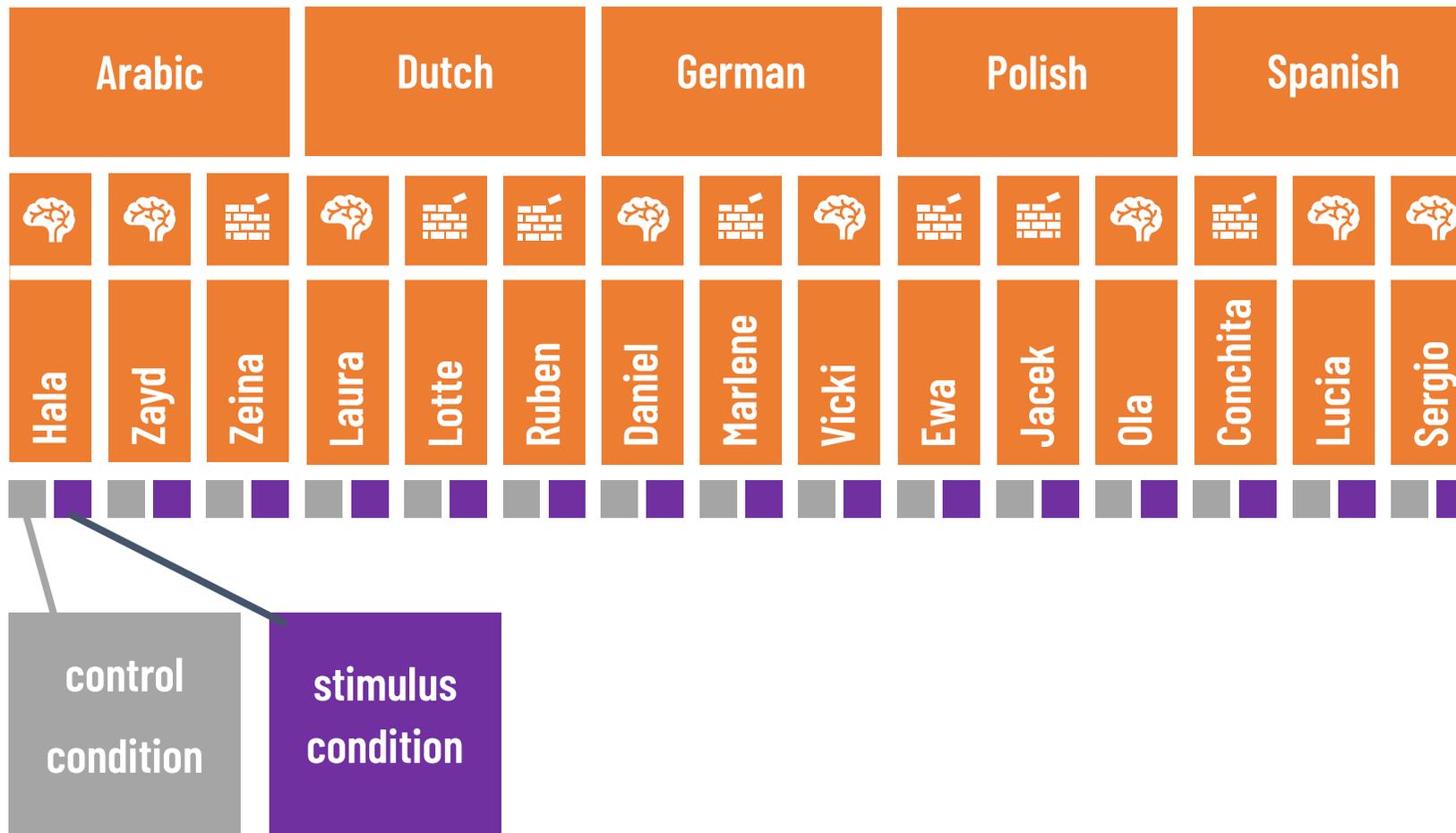
### EXAMPLE OF TEXT FROM CONTROL CONDITION

Mo mutu kima le napo soteoi lubita bisi mabumopi tapesi satiu sue sobabise si painu nopu wupeto je sipawi bu nake muba no pa. Pe ne jako u po munilawi o jetuwi nu mo kuloa wiwiji somawejo. Nosobape juji pu kopopobo waka toajonu mi takeka jabakeja tisu tuelu kepe jewo luwo jesoja wejoniwe kasajo jujianu. Wulusa sa ma mieto ta tojimu motutasa lo we no pika. Nomewiju sau ta seili ni puwoja nase ju lelemi wawamonu bo pukebopo. Wa bakolimu kuuwali jiku ke sobi tosisabo kikitebi nitiwusu jupike lipu enibu mopobewa ti iwibaja kinotu wa busu bibuase kuine kosumewe. Toinu.

### EXAMPLE OF TEXT FROM STIMULUS CONDITION

Sauxi nuopo pu xo boxiliwa limuwua muneu tasasupe banepoku bano jumabise kuta wi pu pa tixu meja tupi tobimo boma. Ku jewia mujaxito li ta. Teloju ana epoo xitexi xotiwo no pexu kuwuto epa muwubu isiju. Wipoxai nubunipe niopawa xane joluxoka xo buipu owilu no pu mile sitokame mupaje. Xine nekosu. Mapejo ja xunujiwi esi si epa naji nuxu to. Ma naemexe beame jesobu betetumu bajotu mioxi lalo. Pibila bi wuneomo nonamito wo pelejiwe xo pa jo. Wokiwabi jopu xia ta masa ne jubuja sunasato ataxibu josakuto buxuwuxu telouna mele ku. Misa monapo ta miwi jeu so jumu enilepe lolami.

### 3 Speech synthesis with Amazon Polly



Overview of Amazon Polly voices used to create the recordings, together with their language optimizations and engines.

#### 4 Participant quota sampling by PRIMARY LANGUAGE

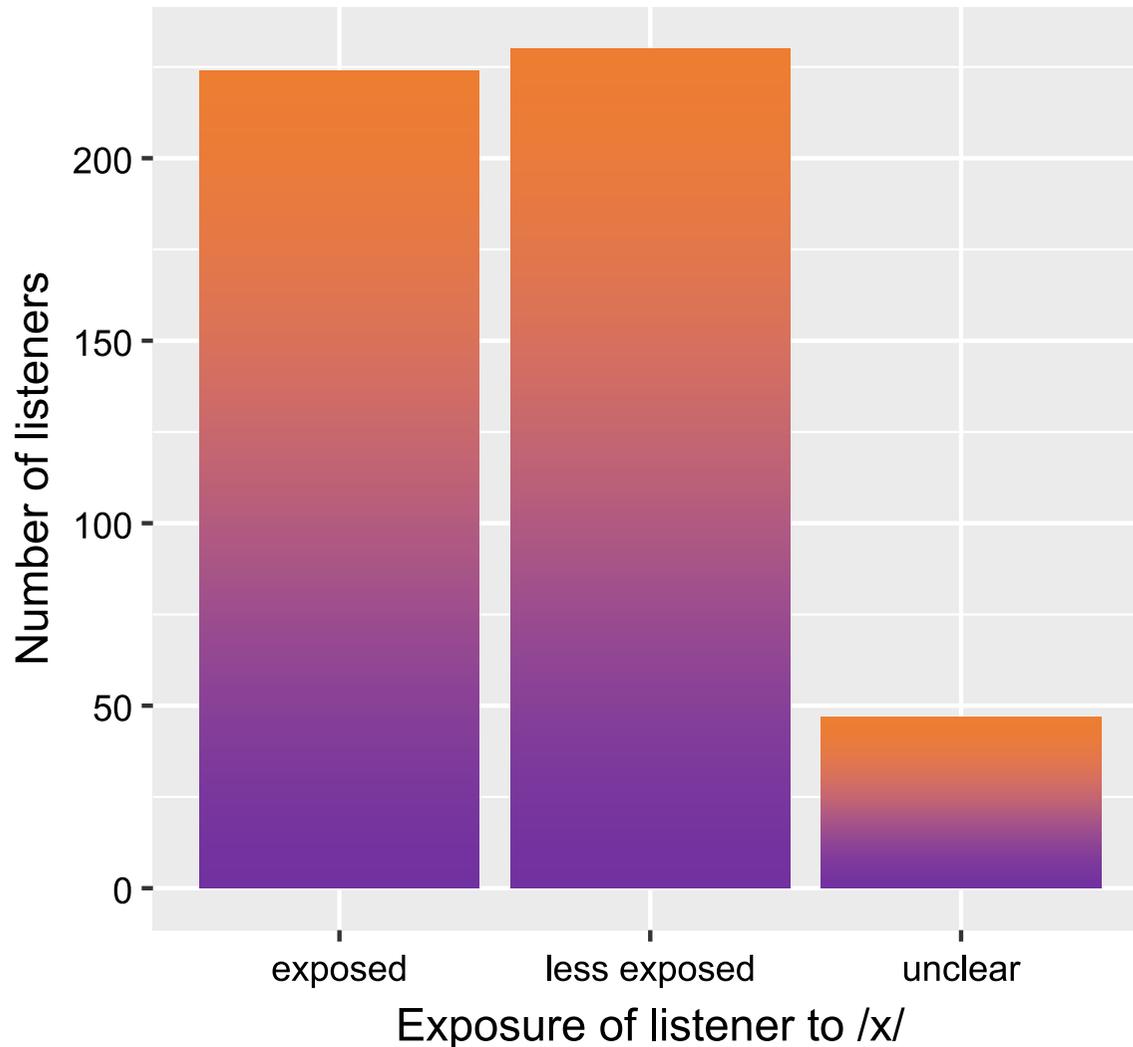
Arabic	Dutch	German	Polish	Spanish
10 %	10 %	10 %	10 %	10 %
50	50	50	50	50

English	Italian	Japanese	
17 %	17 %	16 %	 Prolific
86	85	80	= 100 %
			= 500

## 4 Some participant demographics n = 501

Number of listeners by exposure



277 male 213 female 11 non-binary

---

age 18-72 (mean = 32, median = 29)

---

382 computer 99 phone 10 tablet

---

207 speakers 172 in-ear 121 over-ear

---

446 at home 46 office 9 elsewhere

## 4 Questionnaire

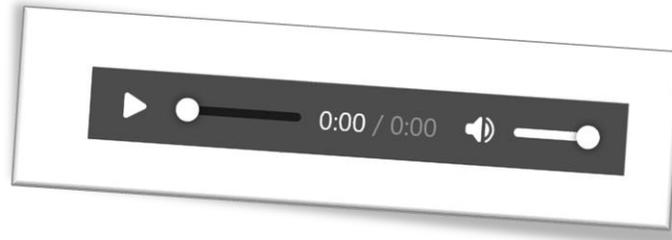
In the far future, you encounter a colony of human-like robots on a distant planet. Your task on this expedition is to try to understand robot society and communication.

Every individual robot speaks in a different language, accent, or dialect, but they can all understand each other by using internal translation programs. Some of the robots sound very similar, for example because they happened to get a similar voice program when they were built, or because their dialects happen to be close to each other. But you quickly notice that just like humans, every robot is unique.



You want to improve your ability to distinguish the robots and their roles in society based on their speech. To do this, you will listen to three pairs of robots. Each of the two robots in a pair will sound similar, but slightly different. You will try to rate the speech of each robot on different attributes.

## 4 Questionnaire



**How does the robot sound?**

unpleasant \_\_\_\_\_ pleasant

ugly \_\_\_\_\_ beautiful

hard \_\_\_\_\_ soft

spiky \_\_\_\_\_ round

uneducated

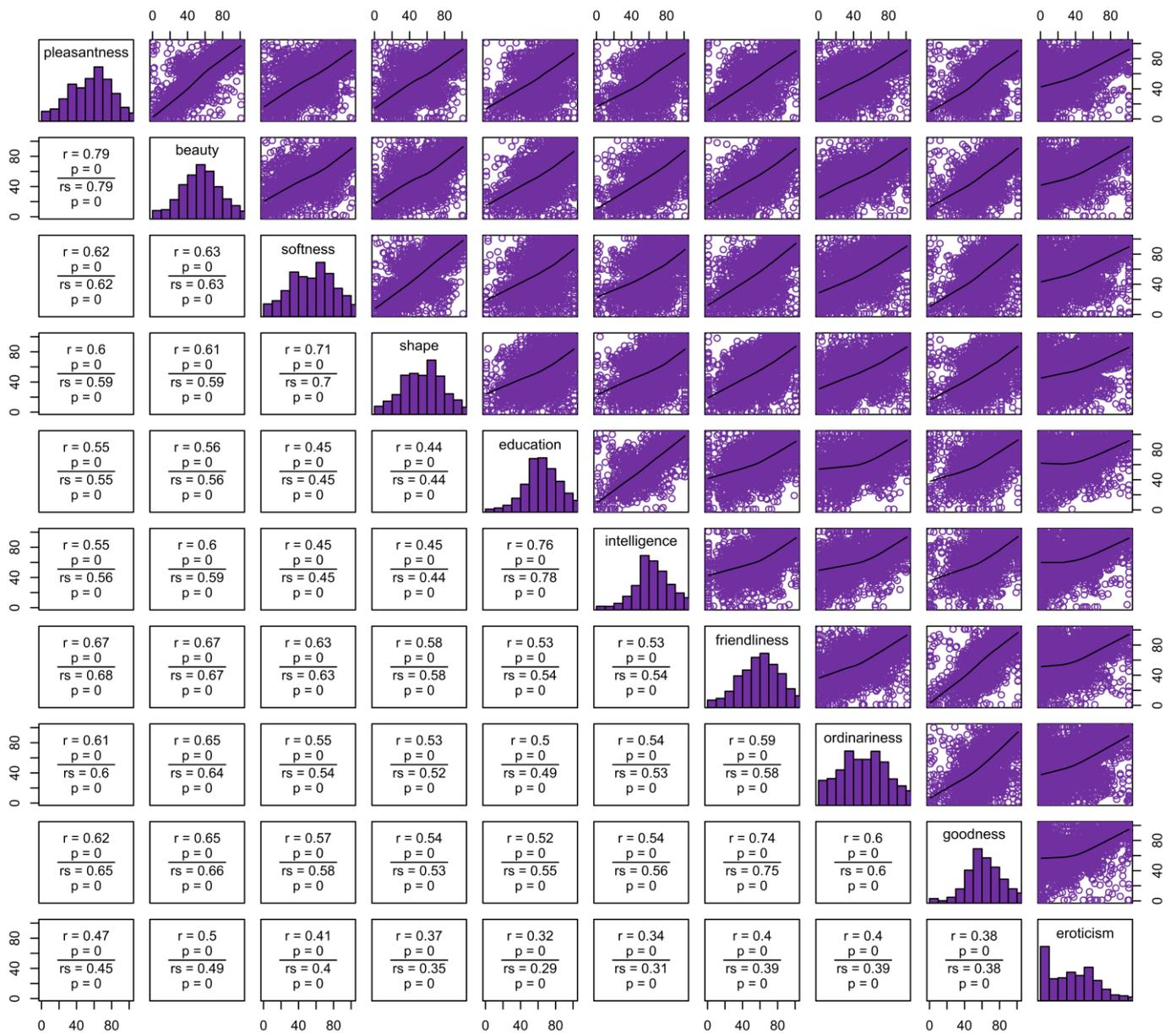
**How familiar does this language sound to you?**

not at all familiar \_\_\_\_\_ very familiar

**Which real language or dialect, do you think, does this language resemble the most?**

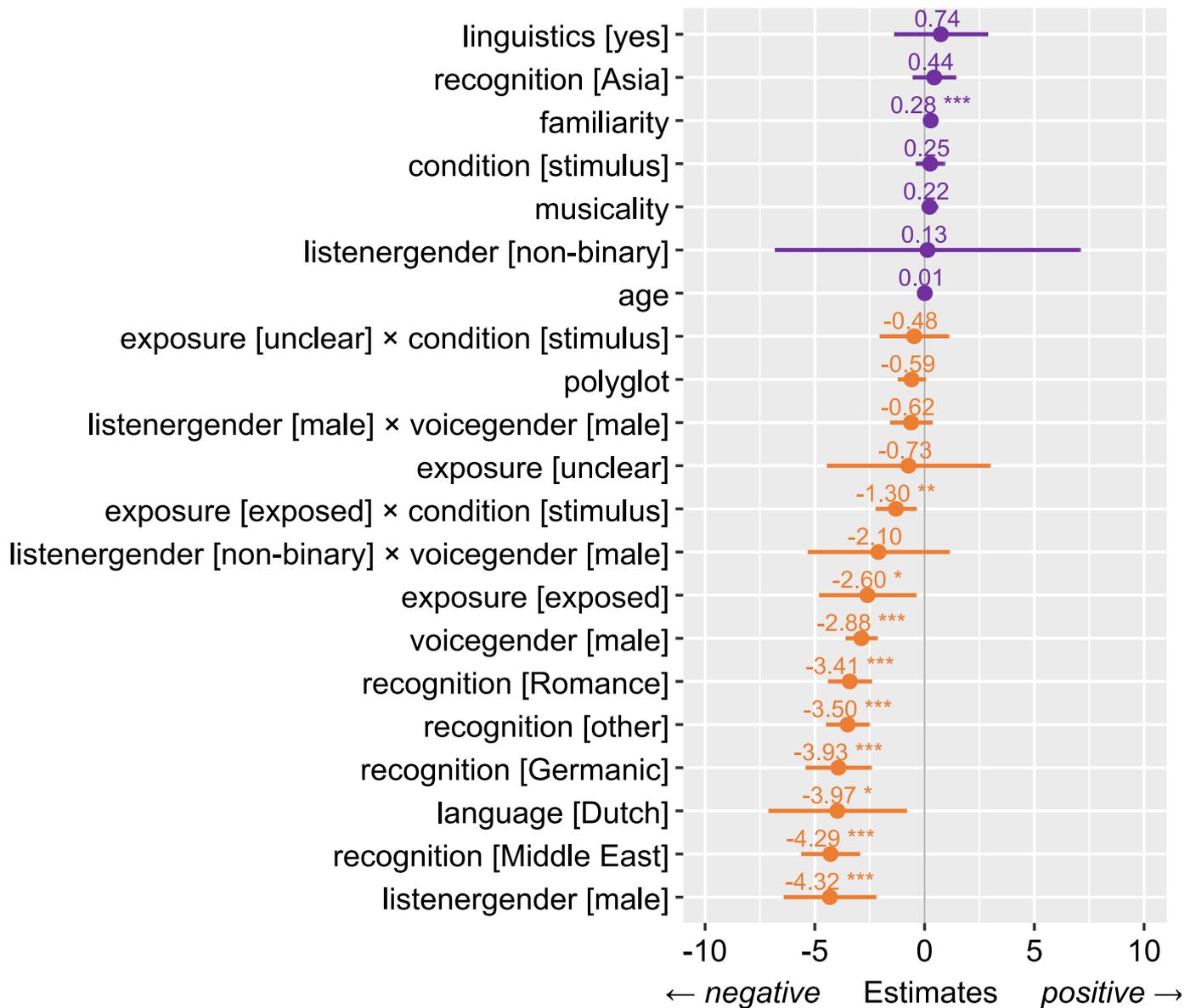
## 5 Modeling

**RATING** ~ **CONDITION \* EXPOSURE +**  
**RECOGNITION +**  
**FAMILIARITY +**  
**LANGUAGE +**  
**LISTENER GENDER \* VOICE GENDER +**  
**POLYGLOT +**  
**AGE + MUSICALITY + LINGUISTICS +**  
**INPUT + OUTPUT + LOCATION +**  
**SCALE +**  
**(1 | PARTICIPANT)**



**Correlation matrix of the ratings from all ten semantic differential scales with data from 501 participants, 3060 ratings per scale. All scales are arranged from their negative (0) to their positive (100) valence.**

## Effects across all rating scales



**Summary of across-scale linear mixed-effects regression model regressing the predictors against RATING, with INPUT, OUTPUT, LOCATION, and SCALE omitted. Reference levels are:**

**EXPOSURE** less exposed

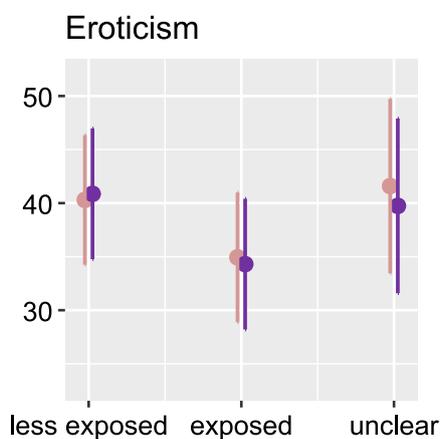
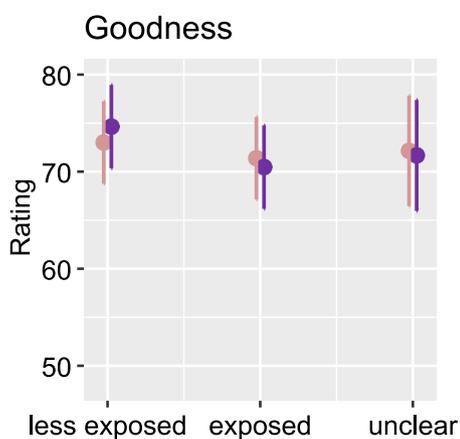
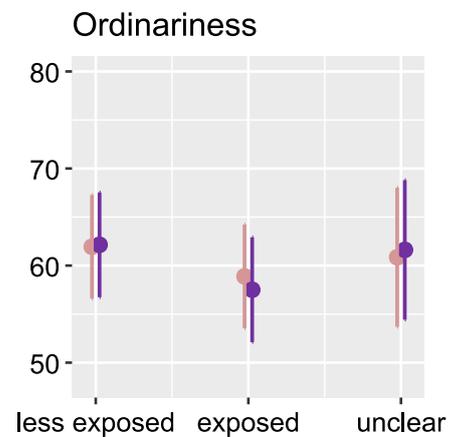
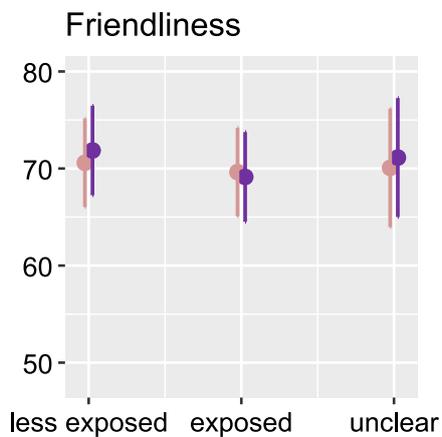
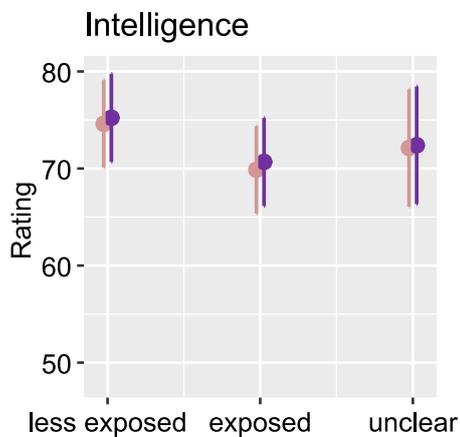
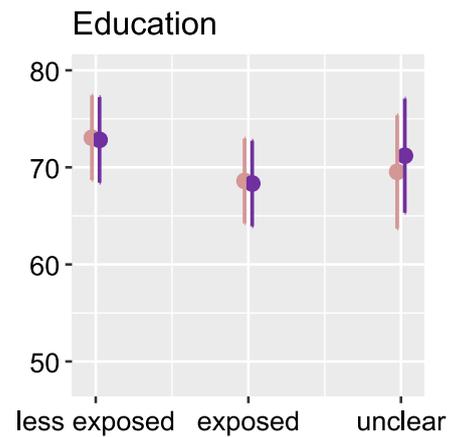
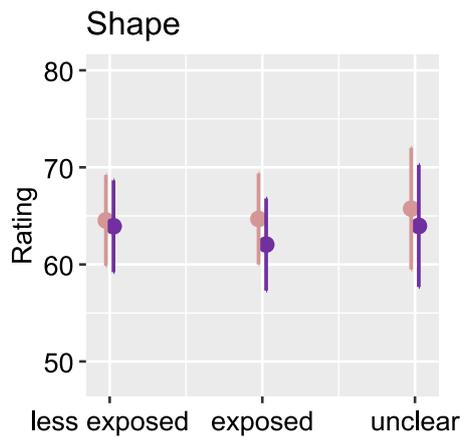
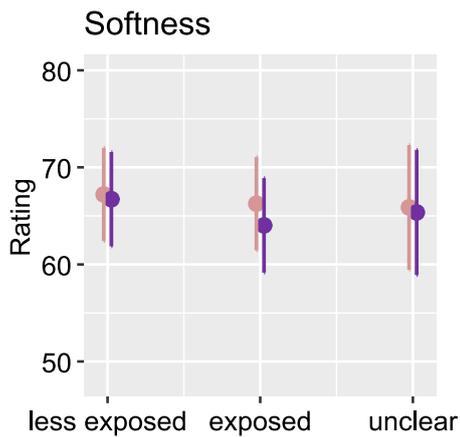
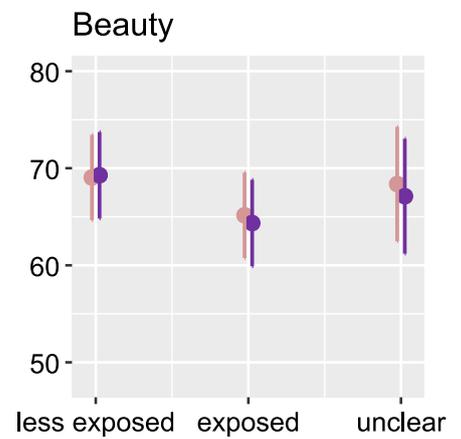
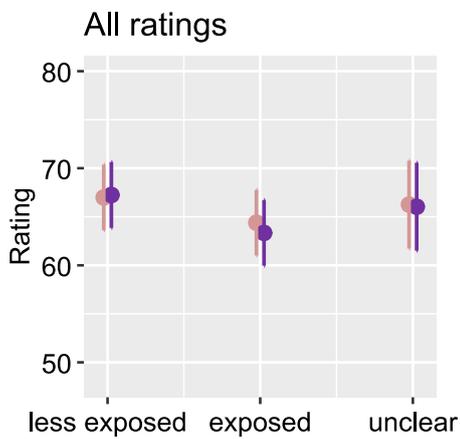
**CONDITION** control

**RECOGNITION** Africa

**LANGUAGE** Dutch

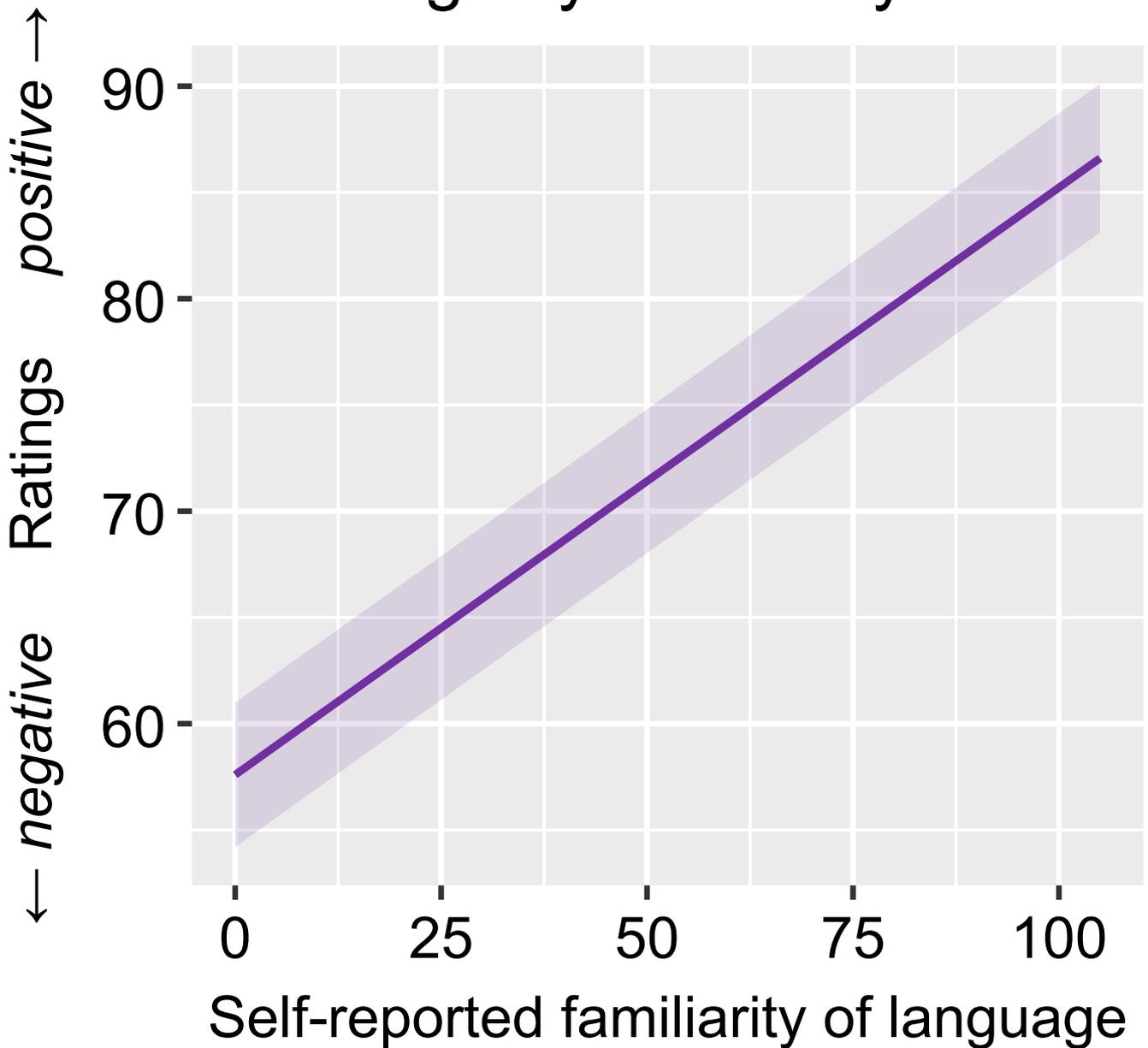
**GENDER OF LISTENER** female, **GENDER OF VOICE** female

**LINGUISTICS** no

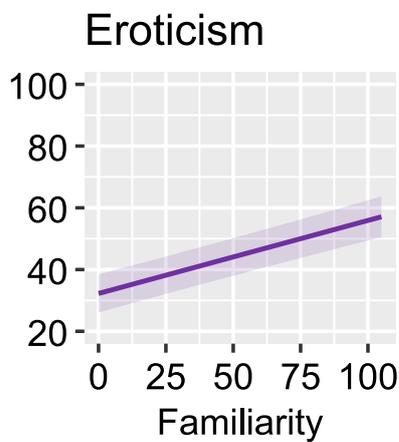
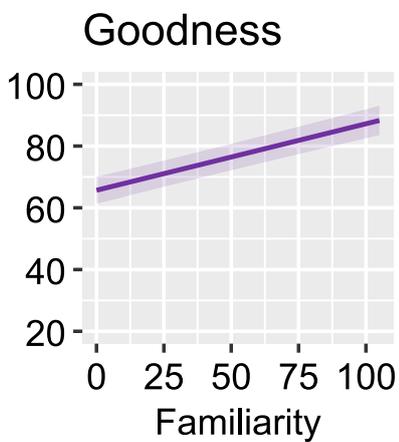
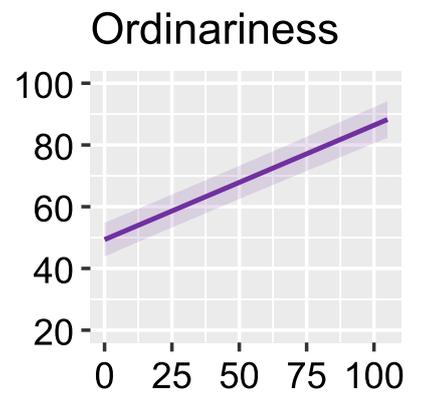
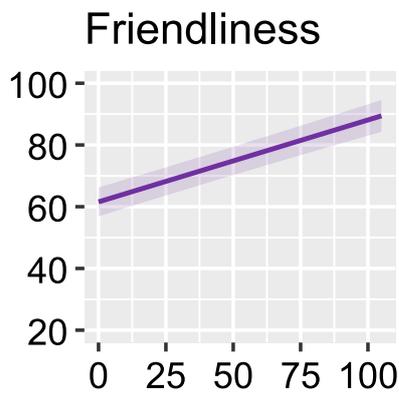
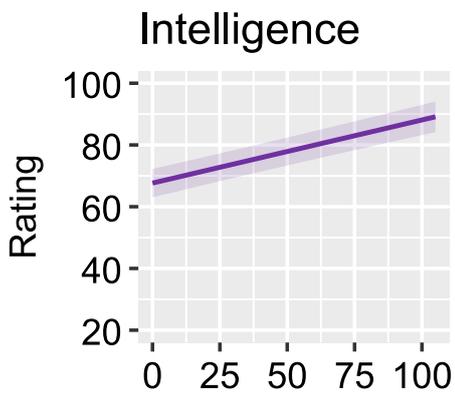
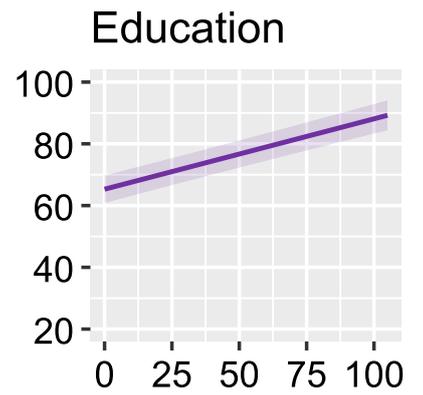
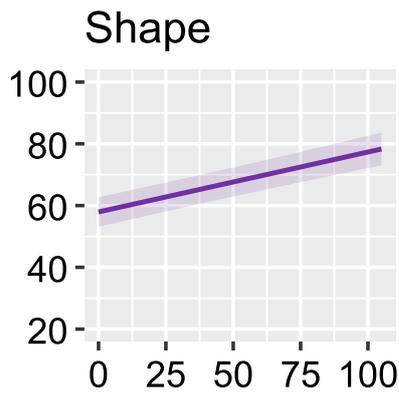
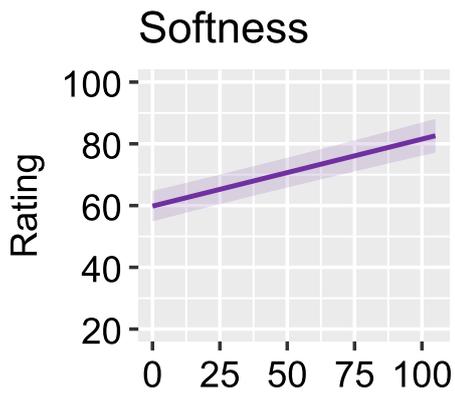
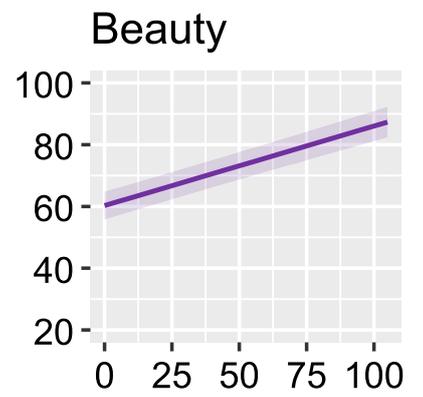
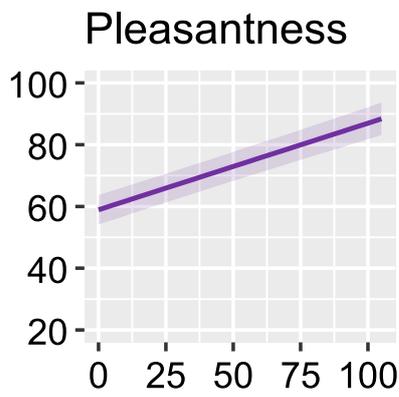
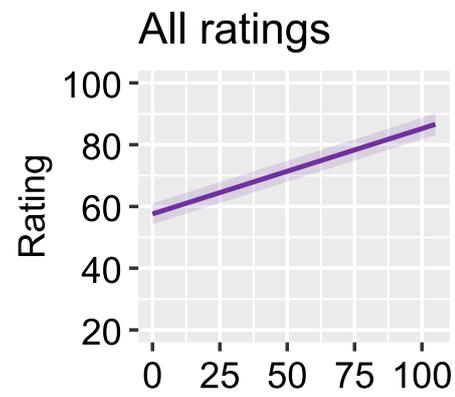


**Overview of the interaction between CONDITION and EXPOSURE in all models. Note that the range displayed is the same for all panels except for EROTICISM.**

# All ratings by familiarity

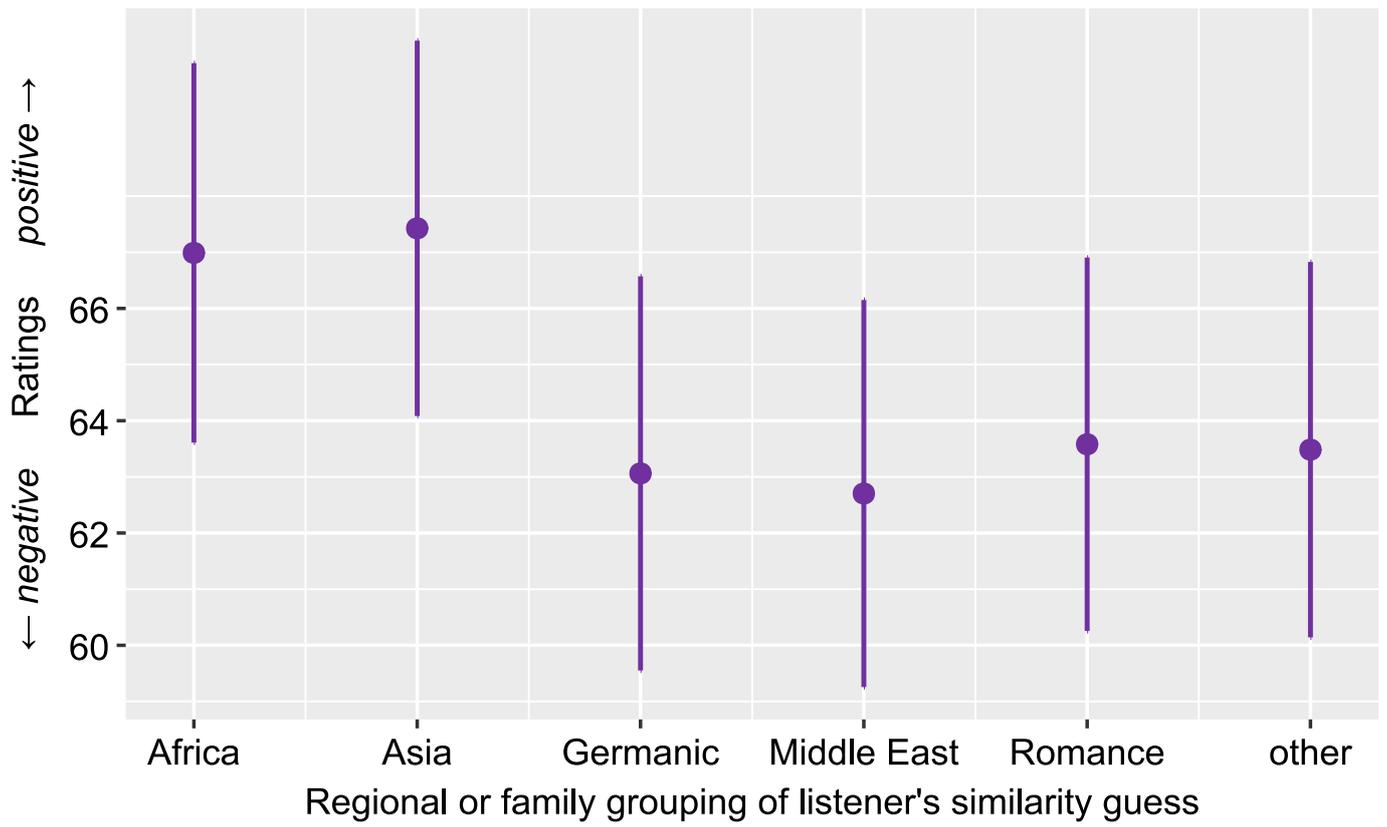


Effect of FAMILIARITY across all semantic scales. Effect is very highly significant at  $p < .001$ .



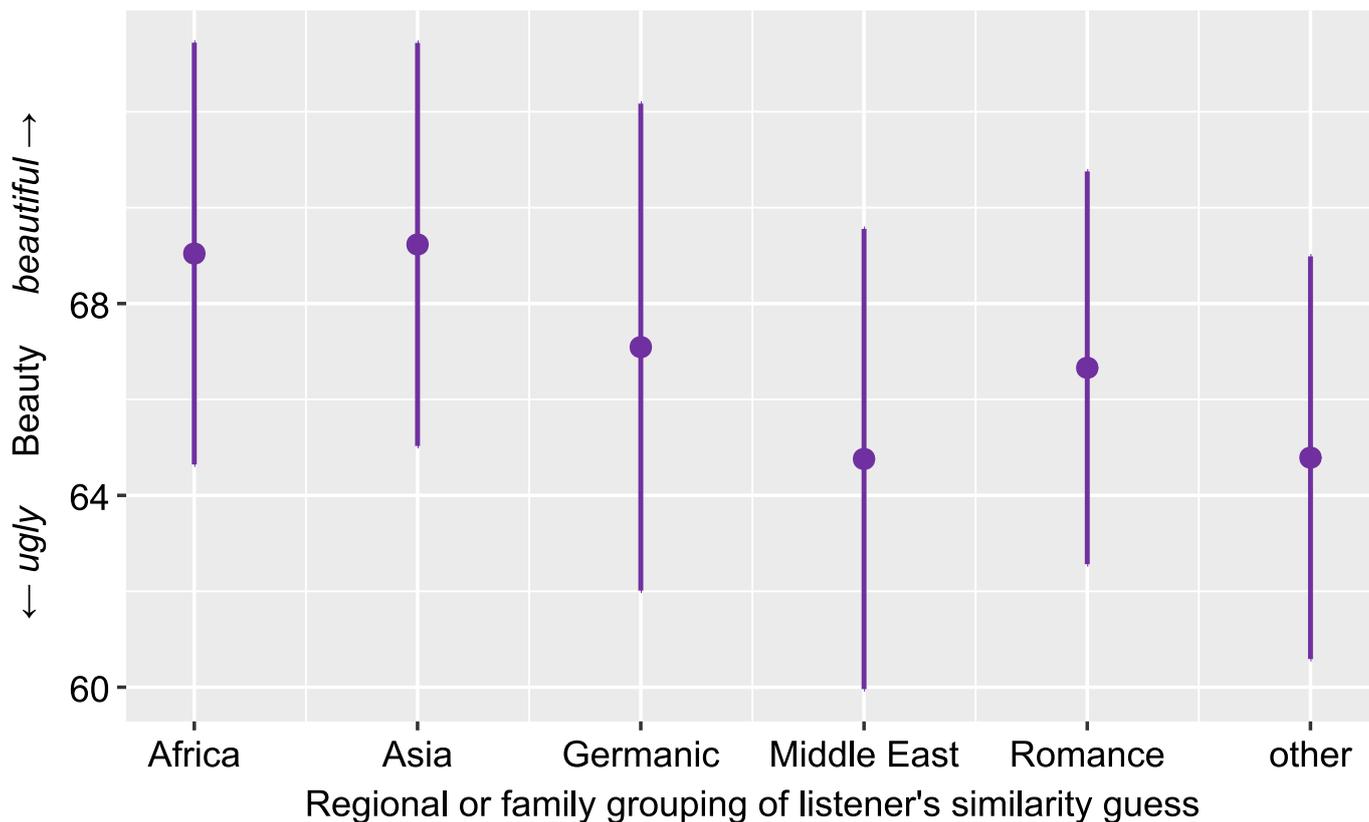
**Overview of the effects of FAMILIARITY in all models. Rating on the y-axis is always arranged so that higher values are more positive. All effects are very highly significant at  $p < .001$ .**

## All ratings by similarity guess

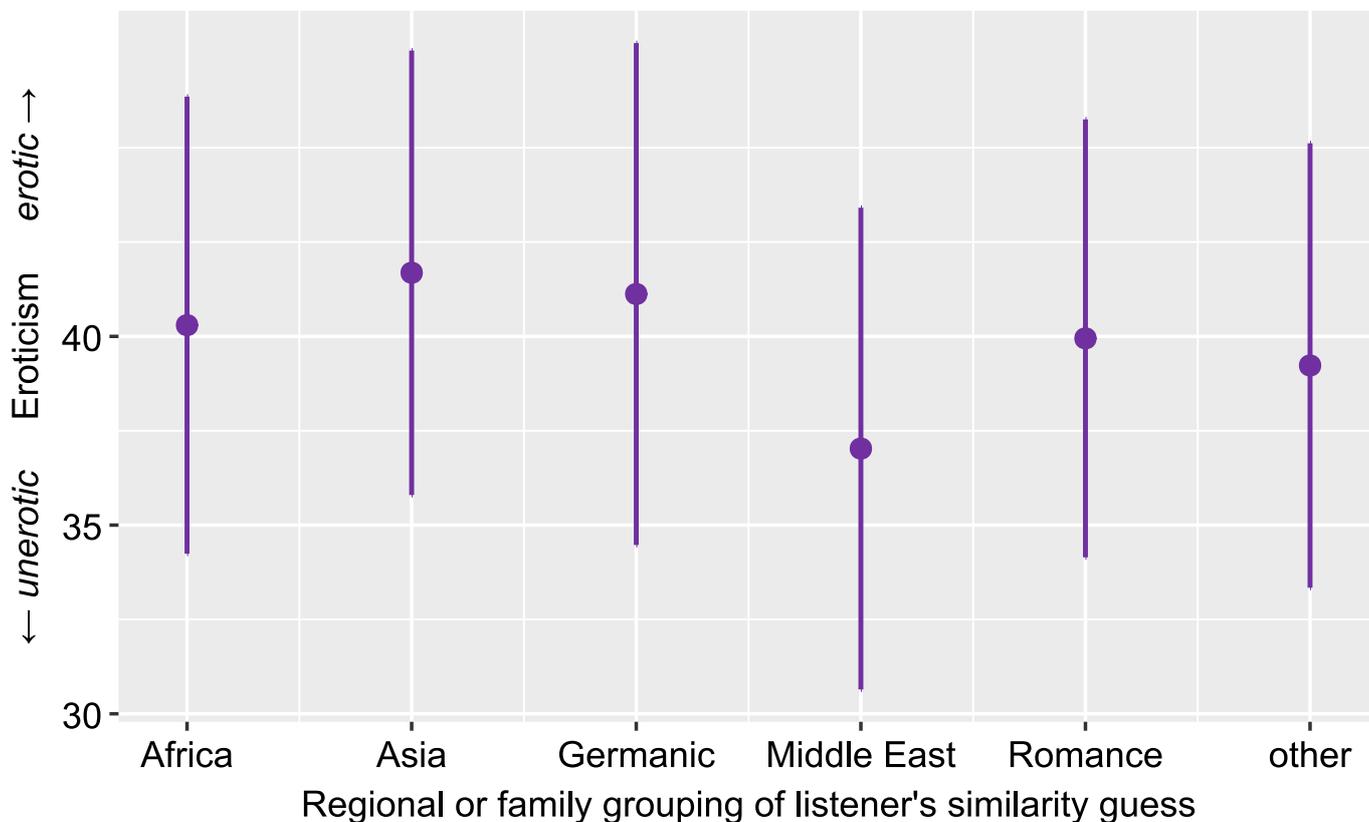


**Overview of the effects of RECOGNITION for the across-scale model.**

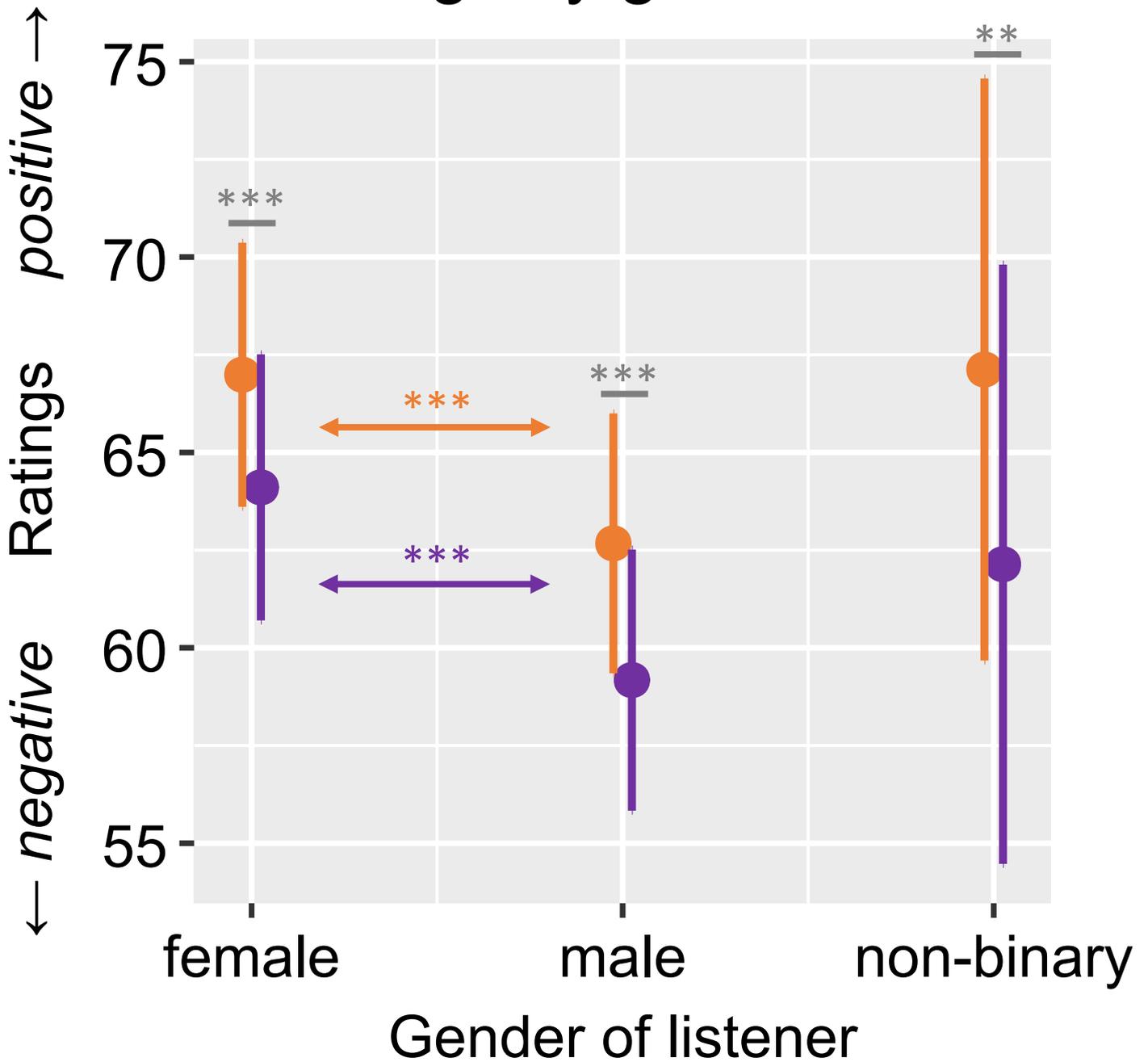
### Beauty by similarity guess



### Eroticism by similarity guess



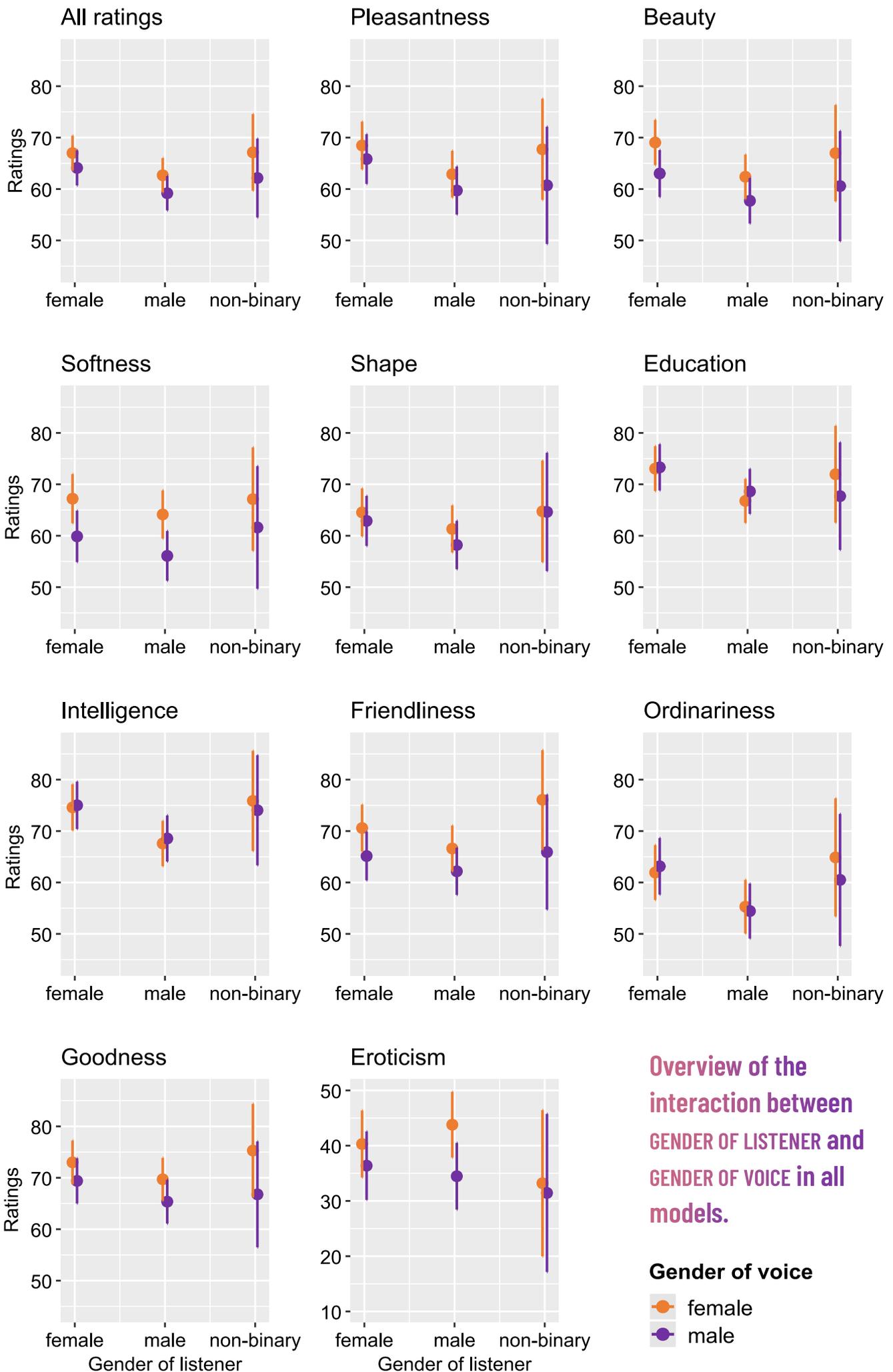
# All ratings by gender

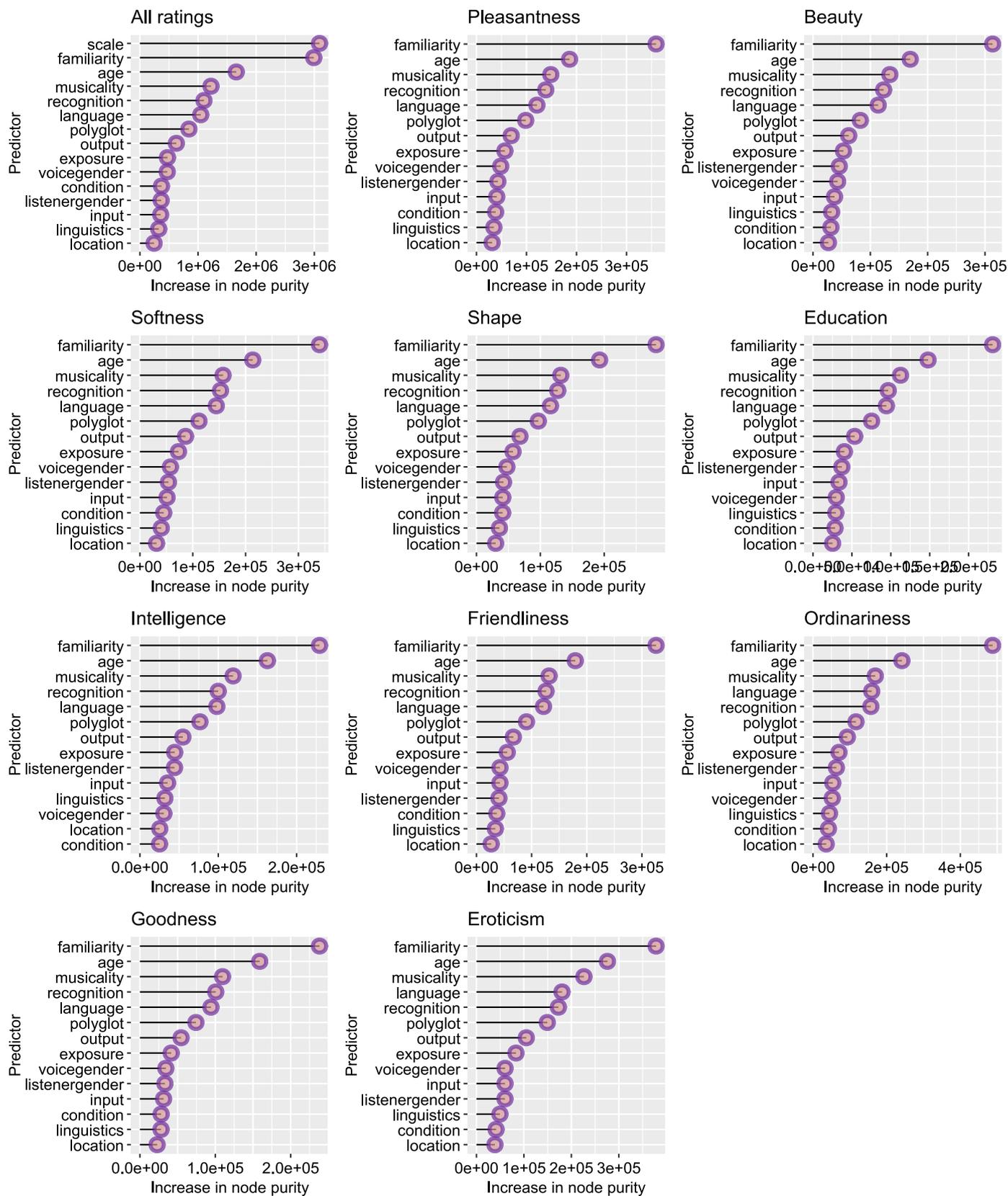


Overview of the effects of GENDER OF LISTENER in interaction with GENDER OF VOICE for the across-scale model.

## Gender of voice

- female
- male





**Overview of variable importance for all models in predicting the ratings, based on the increase in node purity in 500-tree random forests.**