

# Testing the interaction of gendered sound symbolism and morphology in Urdu names

Hassan Khan & Yoonjung Kang

Department of Linguistics

University of Toronto

SULC 2024



# Acknowledgements

University of Toronto Work Study Program (\$)

SSHRC Insight Development Grant (\$)

UTSC Department Research Fund (\$)

Professor Yoonjung Kang (😊)

# Outline

- Background and motivation
  - Gendered sound symbolism in Urdu names
  - Gendered morphology as a confound
- Our contribution
  - Morphology as an interaction effect: corpus findings
  - Future experimental directions

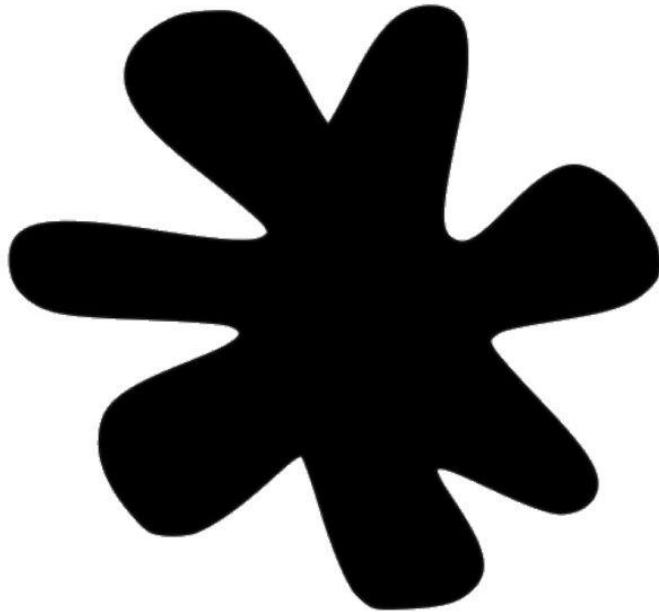
# Outline

- Background and motivation
  - **Gendered sound symbolism in Urdu names**
  - Gendered morphology as a confound
- Our contribution
  - Morphology as an interaction effect: corpus findings
  - Future experimental directions

(Köhler 1929; Ramachandran & Hubbard 2001; Ówiek et al. 2021)

# Sound symbolism

bouba



kiki



# Gendered sound symbolism

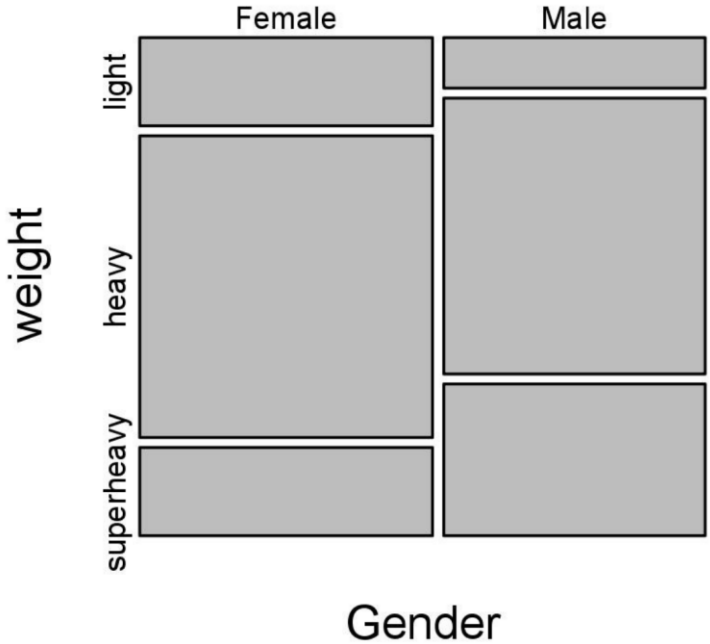
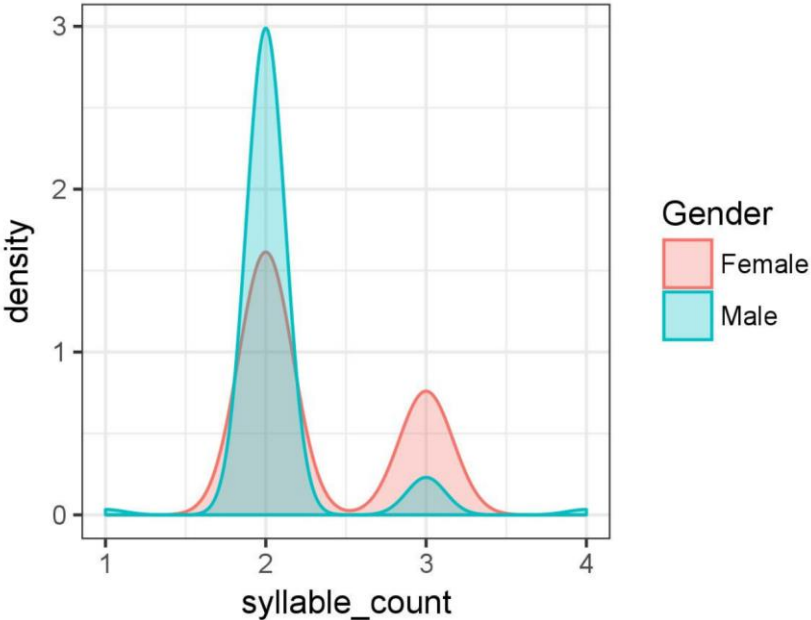


- sonorants
- longer
- open syllables
- front vowels
- nonfinal stress



- obstruents
- shorter
- closed syllables
- back vowels
- final stress

# Sound-gender patterns in Urdu names



$p < 0.001$   
(highly significant)

Longer → feminine

Heavier syllables → masculine

# Outline

- Background and motivation
  - Gendered sound symbolism in Urdu names
  - **Gendered morphology as a confound**
- Our contribution
  - Morphology as an interaction effect: corpus findings
  - Future experimental directions



# Gender morphology

Italian: Giorgi<sub>o</sub><sup>♂</sup> / Giorgi<sub>a</sub><sup>♀</sup>

Russian: Alexander (Александр) / Alexandr<sub>a</sub><sup>♀</sup> (Александр<sub>a</sub><sup>♀</sup>)

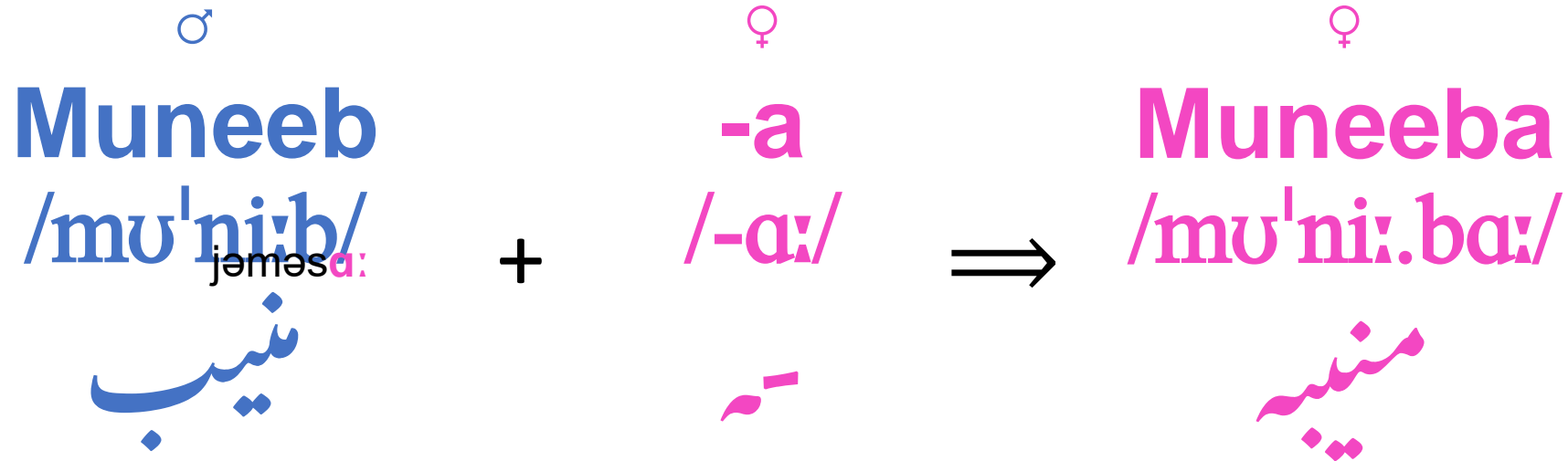
Korean: Tae-shik (태식) / Yoon-jung (윤정)

Urdu:

Muneeb (منیب) / Muneeb<sub>a</sub><sup>♀</sup> (منیب<sub>a</sub><sup>♀</sup>)

Arslan (ارسلان) / Iman (ایمان)

# Confounds from gender morphology



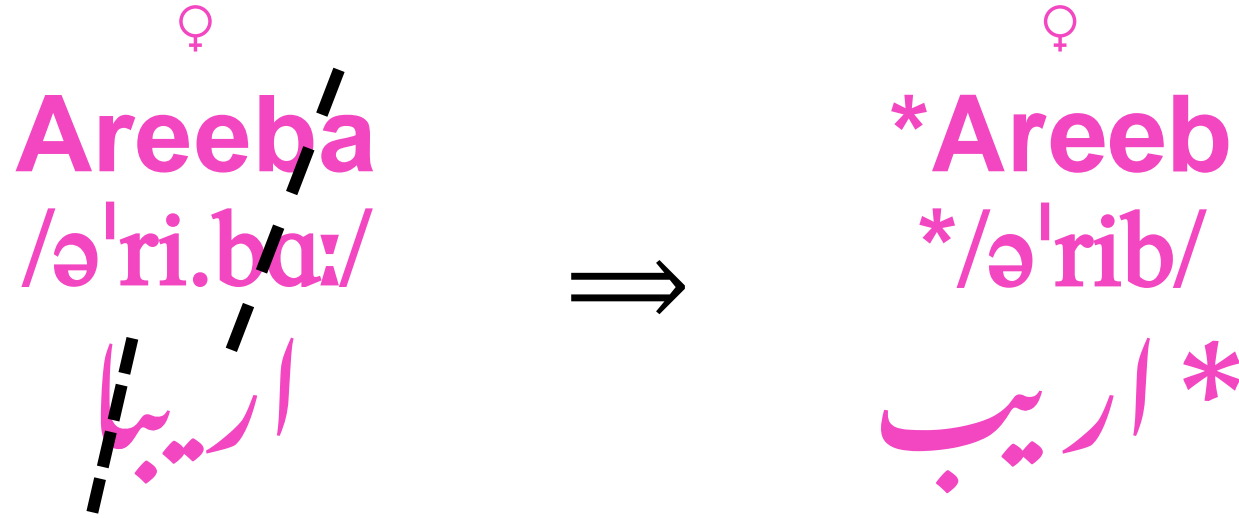
# Avoiding confounds from morphology

“Average sonority (till final sound)” and “Sonority of final sound”  
(Ackermann & Zimmer 2021)

“Final phoneme”  
(Cassidy, Kelly, & Sharoni 1999)

“The second scale, which applies to the last phoneme...”  
(Barry & Harper 1995)

# Controlling for morphology



**Longer → masculine**

Heavier syllables → masculine

**Low vowels → masculine**

# Outline

- Background and motivation
  - Gendered sound symbolism in Urdu names
  - Gendered morphology as a confound
- Our contribution
  - **Morphology as an interaction effect: corpus findings**
  - Future experimental directions

# Communicative functions of gendered sound symbolism and morphology

(Oelkers 2003)

**Areeba**

/ə'ri.baː/

أريبا



**Noor**

/nur/

نور

Amalia Kadji, [Vecteezy.com](https://www.vecteezy.com)

# Communicative functions of gendered sound symbolism and morphology

(Oelkers 2003)

**Areeba**

/ə'ri.baː/

أريبا



**Noor**

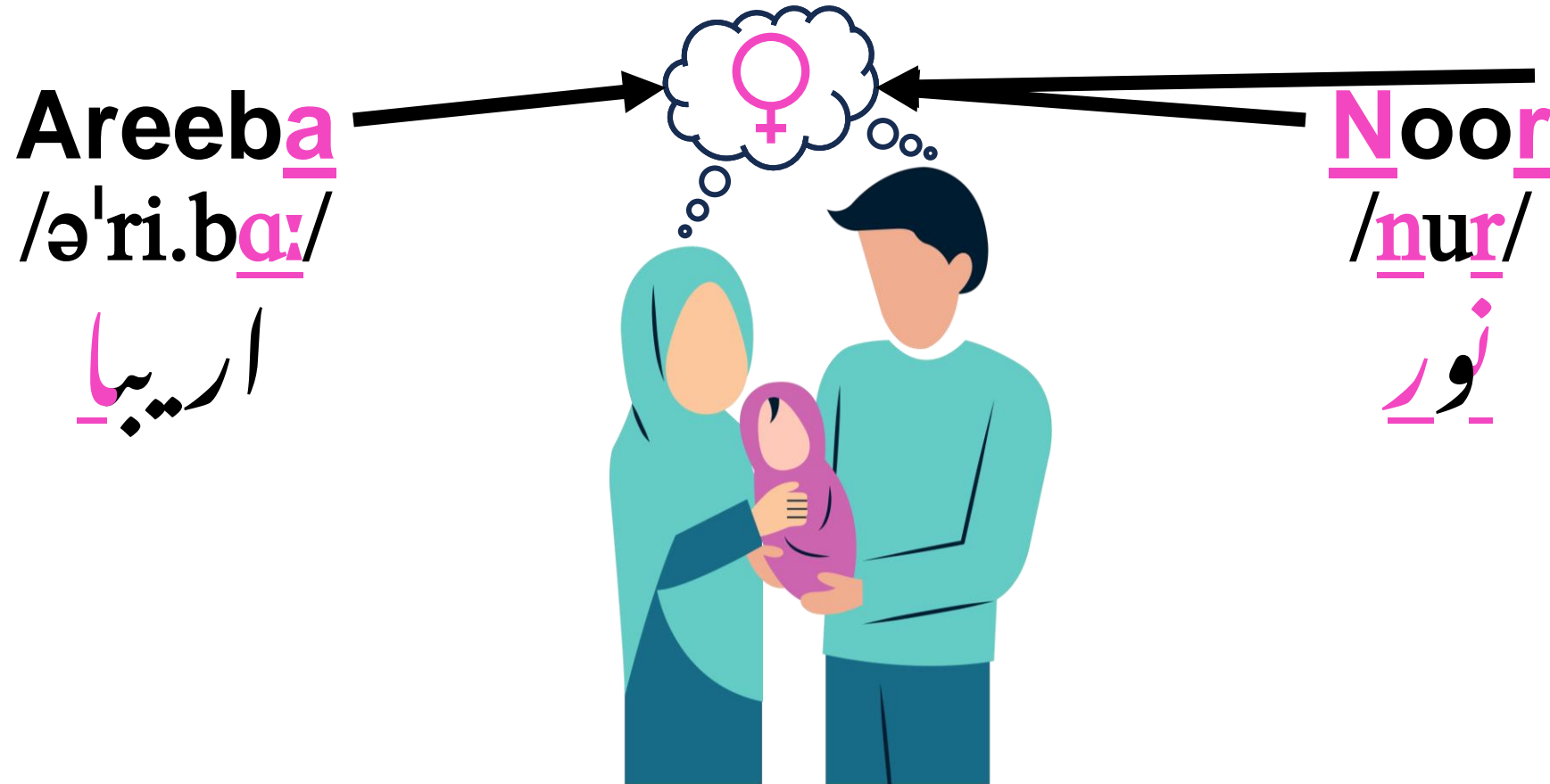
/nur/

نور

Amalia Kadji, [Vecteezy.com](https://www.vecteezy.com)

# Communicative functions of gendered sound symbolism and morphology

(Oelkers 2003)



Amalia Kadji, [Vecteezy.com](https://www.vecteezy.com)



# Name data

~100 boys' names and ~100 girls' names

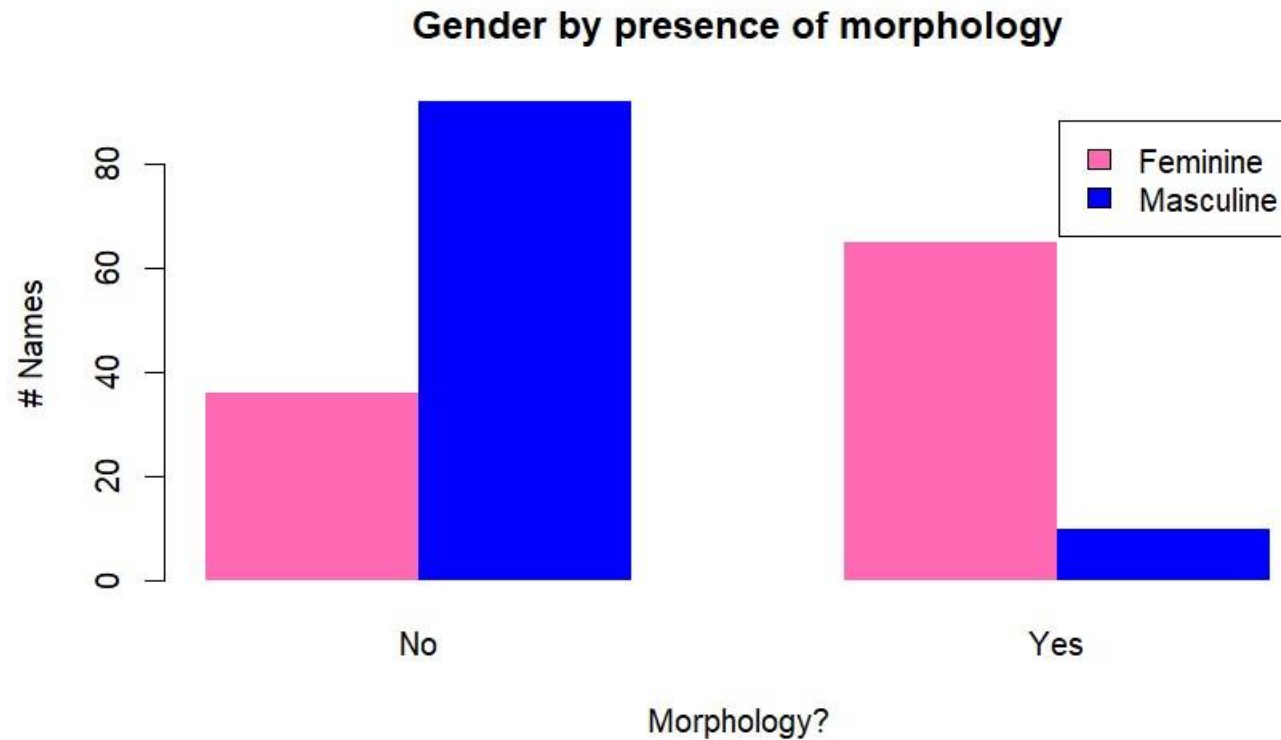
2022 matriculation results

Board of Intermediate and Secondary Education, Gujranwala

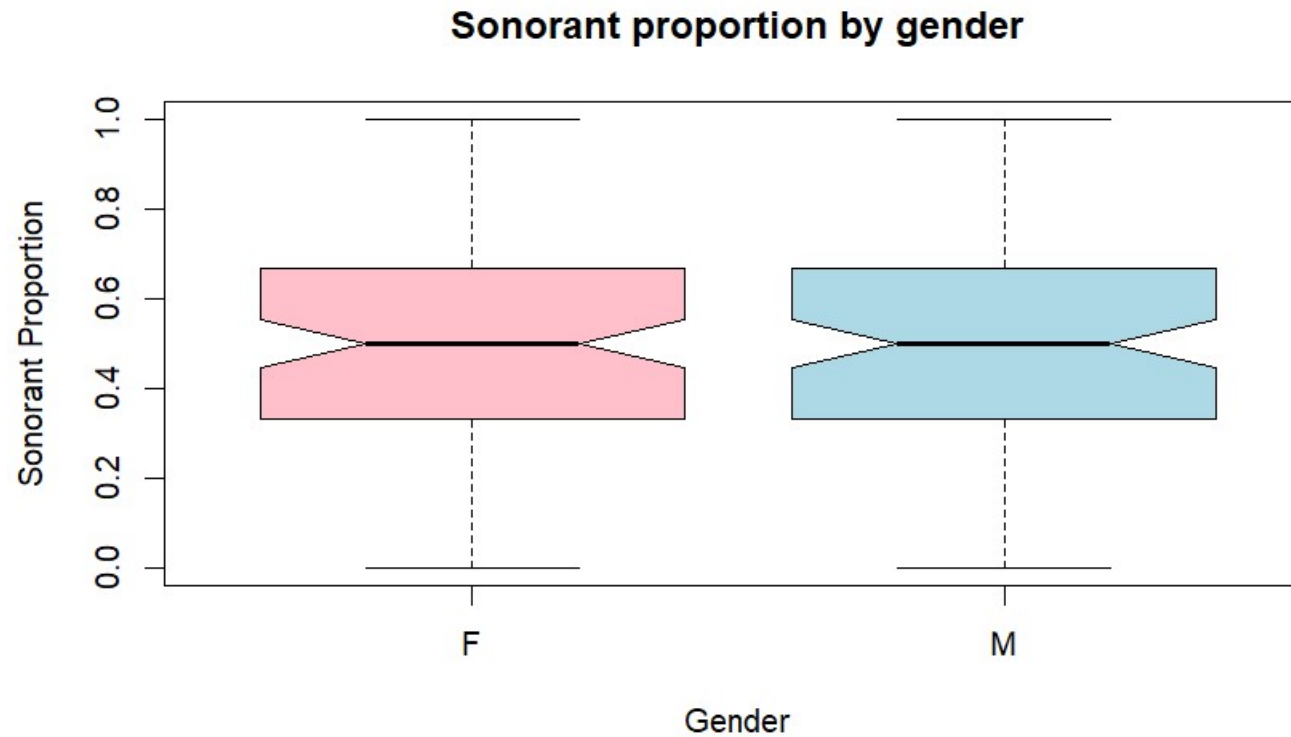
14- and 15-year-olds

16 900 boys, 11 485 girls

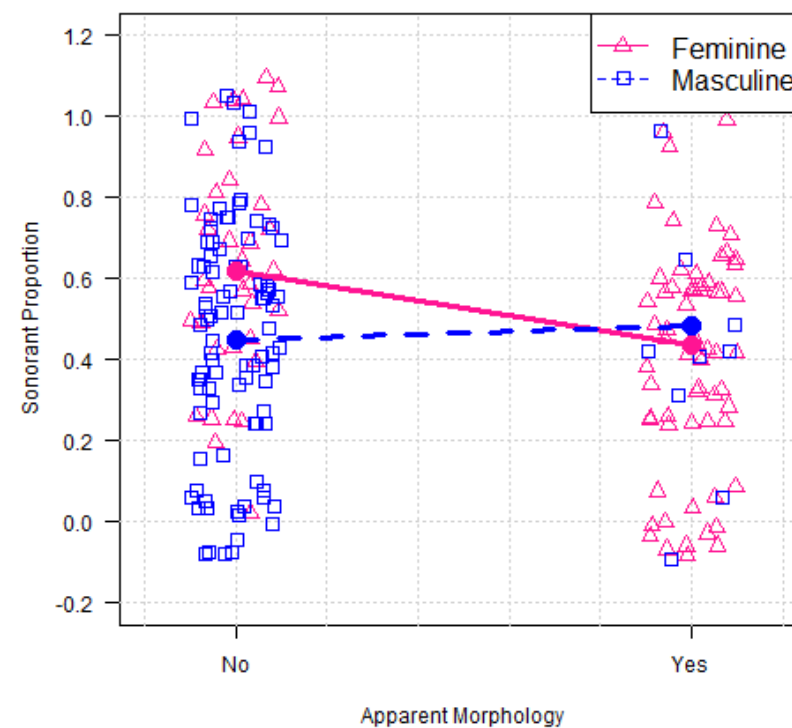
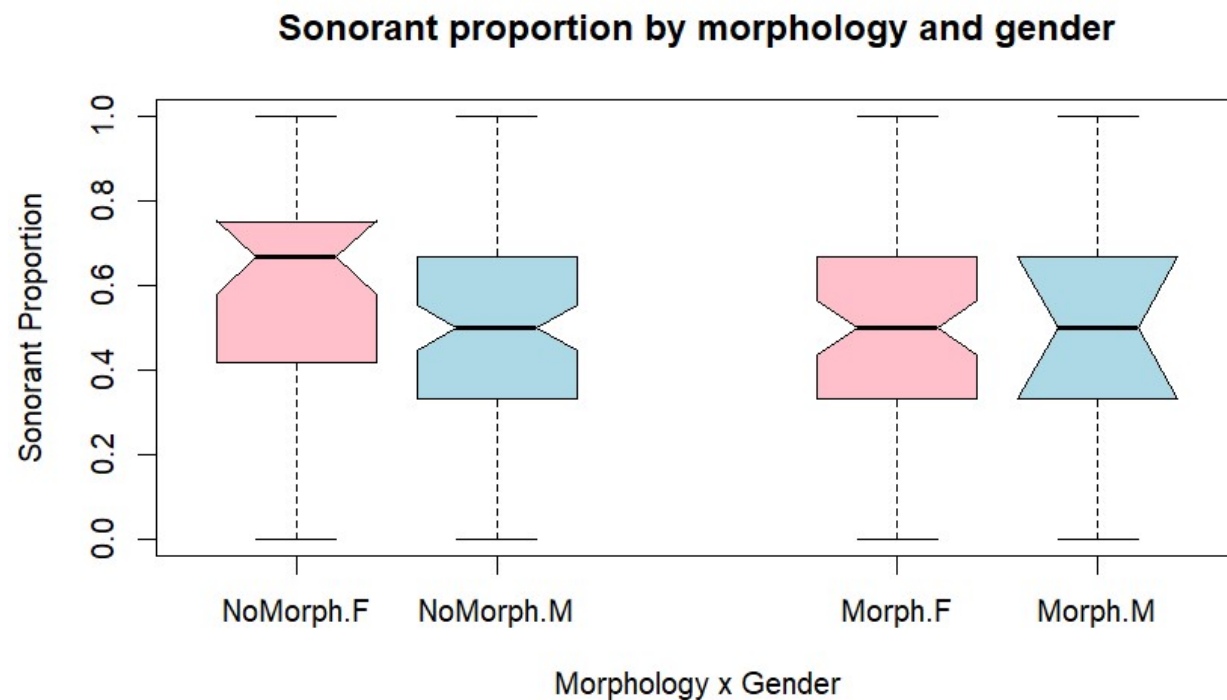
# Feminine morphology as a feminine cue



# Sonorants as a feminine cue...?

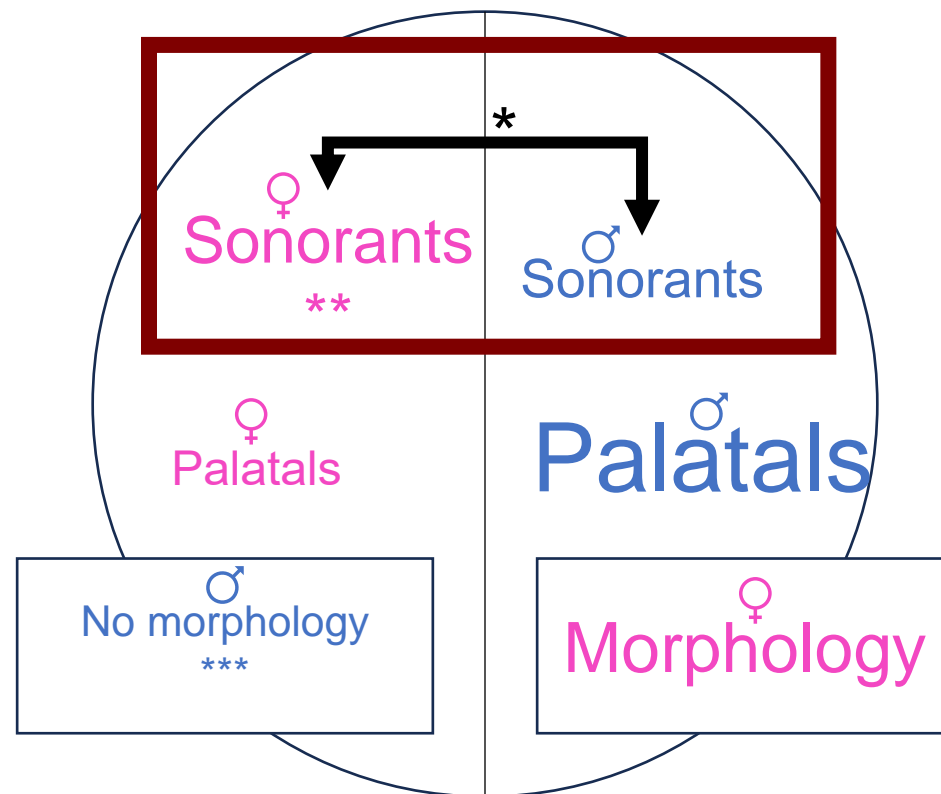


# Sonorants as a feminine cue!



# These results are statistically significant

## Minimal model



\*  $p < 0.05$   
\*\*  $p < 0.01$   
\*\*\*  $p < 0.001$

# Outline



- Background and motivation
  - Gendered sound symbolism in Urdu names
  - Gendered morphology as a confound
- Our contribution
  - Morphology as an interaction effect: corpus findings
  - **Future experimental directions**

# Experiment design 1: Simulating robust/no morphology | Procedure

**Training trials**  
Real Urdu names, matching the morphology condition

**Audio-only stimulus**

Does this sound like a girl's name, or a boy's name?  
کیا یہ نام لڑکی کے نام کی طرح لگتا ہے یا لڑکے کے نام کی طرح؟



**Forced-choice scale**  
Definitely a girl's (یقیناً لڑکی کا نام)  
Maybe a girl's (شاید لڑکی کا نام)  
Maybe a boy's (شاید لڑکے کا نام)  
Definitely a boy's (یقیناً لڑکے کا نام)

**Between-subject design: morphology conditions**

**No morphology:**  
No names end in a feminine suffix

**All morphology:**  
All names end in -a (ا-)

**Mixed morphology:**  
Some names end in -a (ا-), others do not

# Experiment design 1: Simulating robust/no morphology | Stimuli

Stimuli vary by:  
 Sonorant proportion  
 Length in syllables  
 Palatal consonant

Start with a  
 "root"

## No morphology

<b>təbəs</b>	ʃabas
təməs	ʃəməs
rəməs	jəməs
rəmən	jəmən
təbəsʊ:	ʃəbəsʊ:
təməsʊ:	ʃəməsʊ:
rəməsʊ:	jəməsʊ:
rəmənʊ:	jəmənʊ:

No palatals

Palatals

Sonorant  
 proportion  
 increases

## All morphology

təb <b>a:</b>	ʃəb <b>a:</b>
təm <b>a:</b>	ʃəm <b>a:</b>
rəm <b>a:</b>	jəm <b>a:</b>
təbəs <b>a:</b>	ʃəbəs <b>a:</b>
təməs <b>a:</b>	ʃəməs <b>a:</b>
rəməs <b>a:</b>	jəməs <b>a:</b>
rəmən <b>a:</b>	jəmən <b>a:</b>

Two  
 syllables

Three  
 syllables

## Mixed morphology

Stimuli from both conditions (exclude minimal pairs)

təbəs	
	rəmən
ʃəb <b>a:</b>	
	jəməs <b>a:</b>



# Experiment design 1: Simulating robust/no morphology | Predictions

## No morphology

Sonorants remain a feminine cue

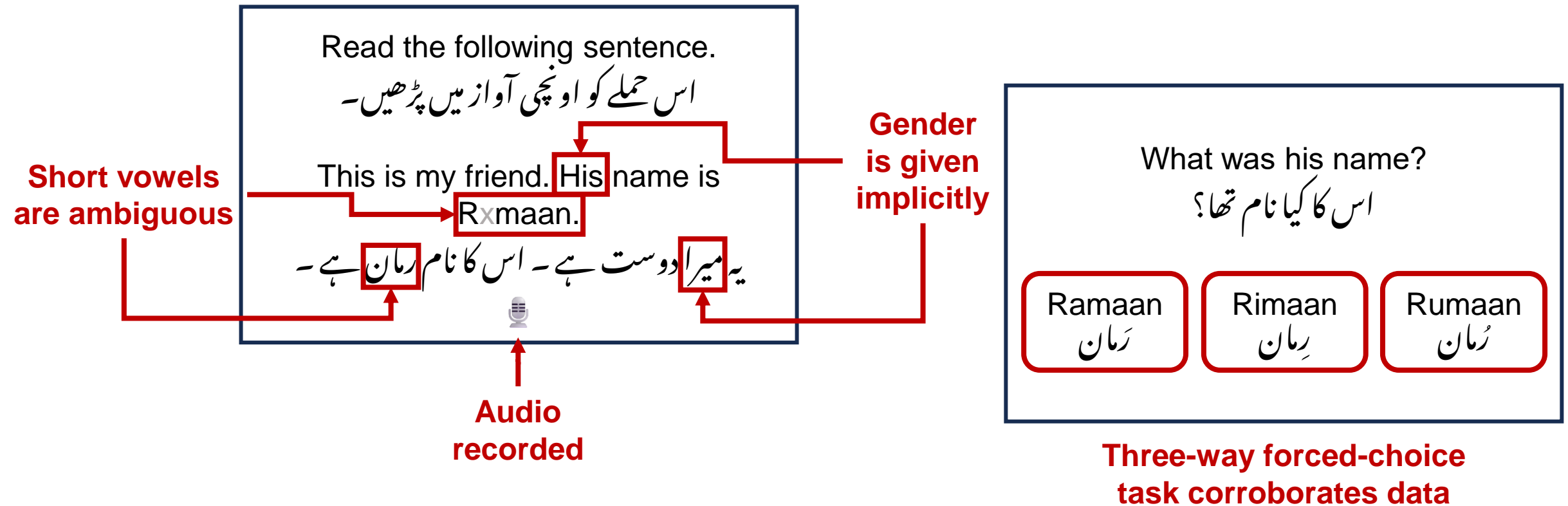
## All morphology

**Sonorants become a feminine cue when morphology is no longer effective**

## Mixed morphology

Mimicking corpus results (sonorants and suffixes both lean feminine)

# Experiment design 2: Using writing system to test vowel height | Procedure



# Experiment design 2: Using writing system to test vowel height | Predictions

Vowel height effects may still be present in speaker intuitions, even if not reflected in corpus data

Higher, fronter vowels should be produced more frequently in feminine name contexts

# Implications in the search for sound-meaning associations

A methodological problem

A cross-linguistic account

An increasingly intricate human communication system

# Conclusions

Feminine Urdu names have more sonorants **when morphology is not present**

Experimental evidence is needed to understand the **synchronicity and robustness** of these results

Future sound symbolism research needs to **account for these confounding effects**

[hassankhan.net/  
research](http://hassankhan.net/research)



# References

Ackermann, Tanja & Christian Zimmer. 2021. The sound of gender – correlations of name phonology and gender across languages. *Linguistics* 59(4). 1143–1177.

Board of Intermediate and Secondary Education. 2022. *Secondary School Certificate Second Annual Examination Result Gazette*. <https://bisegrw.edu.pk/result-gazatte.html>. Gujranwala, Pakistan.

Ćwiek Aleksandra et al 2022. The bouba/kiki effect is robust across cultures and writing systems. *Philosophical Transactions of the Royal Society B* 377(20200390).

Kang, Yoonjung. 2021. Phonology of gender in personal names. talk. Department of Linguistics, University of Southern California

Köhler, Wolfgang. 1947. *Gestalt psychology*. New York, NY: Liveright.

Mohsin, Nazia & Yoonjung Kang. 2018. Gender phonology of Urdu first names. Summer Phonology Forum 2018. University of Toronto.

Nubling, Damaris. 2009. Von Monika zu Mia, von Norbert zu Noah: Zur Androgynisierung der Rufnamen seit 1945 auf prosodisch-phonologischer Ebene. *Beiträge zur Namenforschung* 44. 67–110.

Oelkers, Susanne. 2003. *Naming gender: Empirische Untersuchungen zur phonologischen Struktur von Vornamen im Deutschen*. Frankfurt am Main: Lang.

Ohala, John J. 1994. Towards a universal, phonetically-based, theory of vowel harmony. *3rd International Conference on Spoken Language Processing [ICSLP 1994]*. 491-494.

Saussure, Ferdinand de. 1916. *Course in General Linguistics*. London: Duckworth.

Ramachandran, Vilayanur Subramanian & Edward M. Hubbard. 2001. Synaesthesia — a window into perception, thought and language. *Journal of Consciousness Studies* 8. 3–34.

# Appendix A: Data collection and annotation

# Name sources

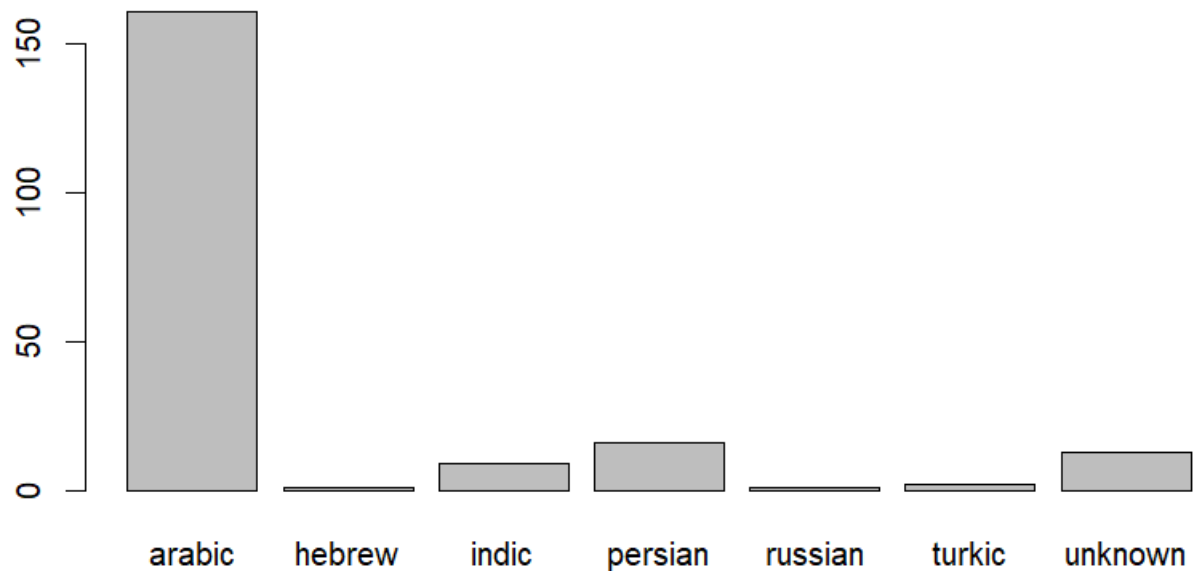
- 102 boys' names, 101 girls' names; unisex names were not removed
- Board of Intermediate and Secondary Education 2022 matriculation results
- Pronunciations from Urdu-language baby name YouTube channels
  - *ZAHID INFO HUB, Ali Bhai, Urdusy, Smart Life Tube...*
  - Coded using Saleem's (2002) phonemic inventory of Urdu



# Variables

- Independent variable gender (M or F)
- Dependent variables coded automatically in *R* (R Core Team 2023)
- Continuous phonological dependent variables (centred)
  - Length in syllables
  - Proportions of high vowels, front vowels, sonorant consonants, palatal consonants, light syllables
- Novel categorical dependent variable “apparent morphology”
  - True for names ending in feminine suffixes /-ɑ/, /-i/, /-ət̪/

# Urdu name origins

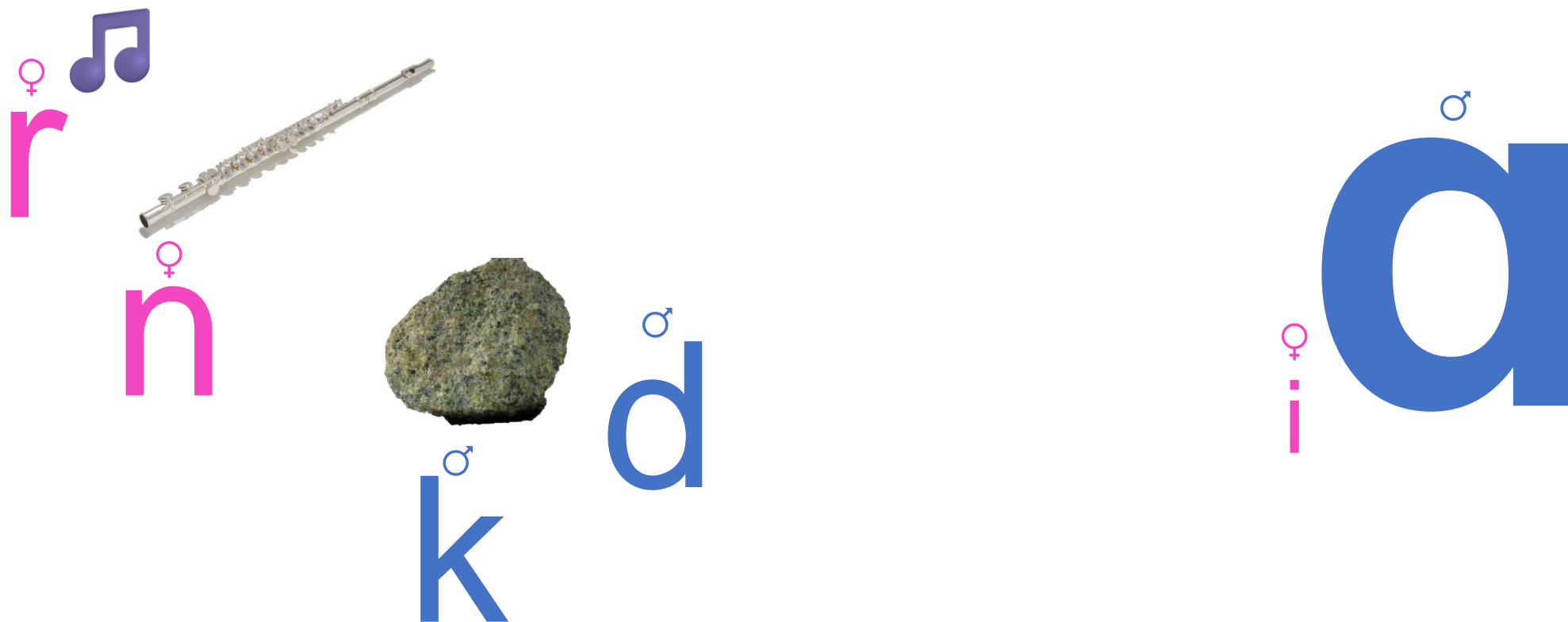


-a  
/-ɑ:/  
ا

# Appendix B: Conventional vs synesthetic sound symbolism

Competing accounts of (gendered) sound symbolism

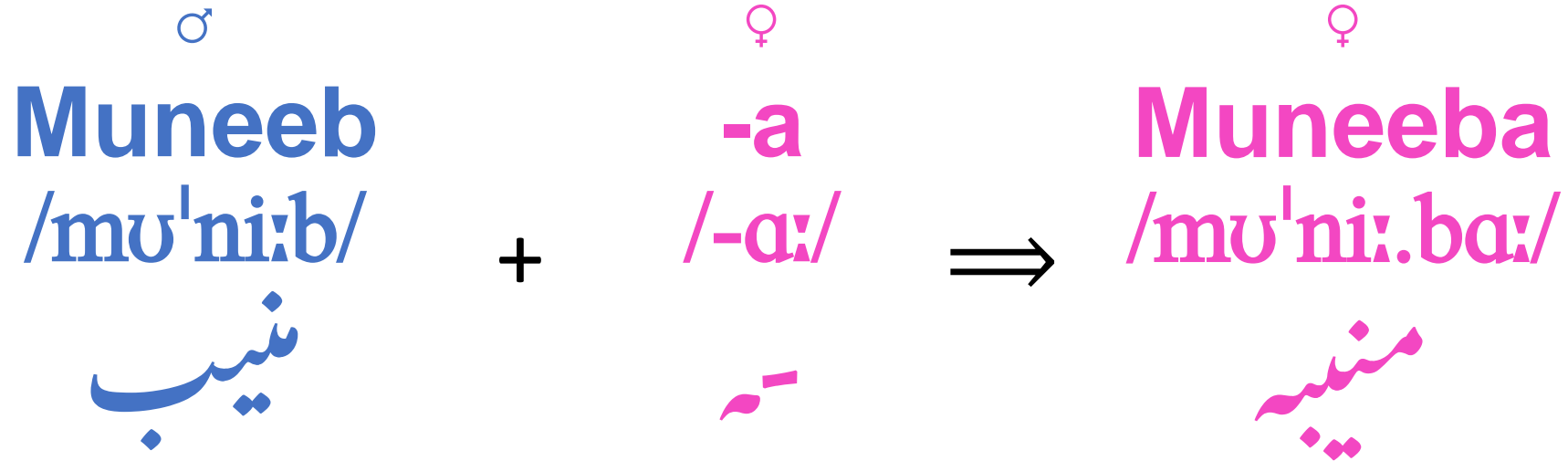
# Synesthetic sound symbolism



Flute: Thea Paraskevaides via Wikimedia, 2010. CC-BY-SA

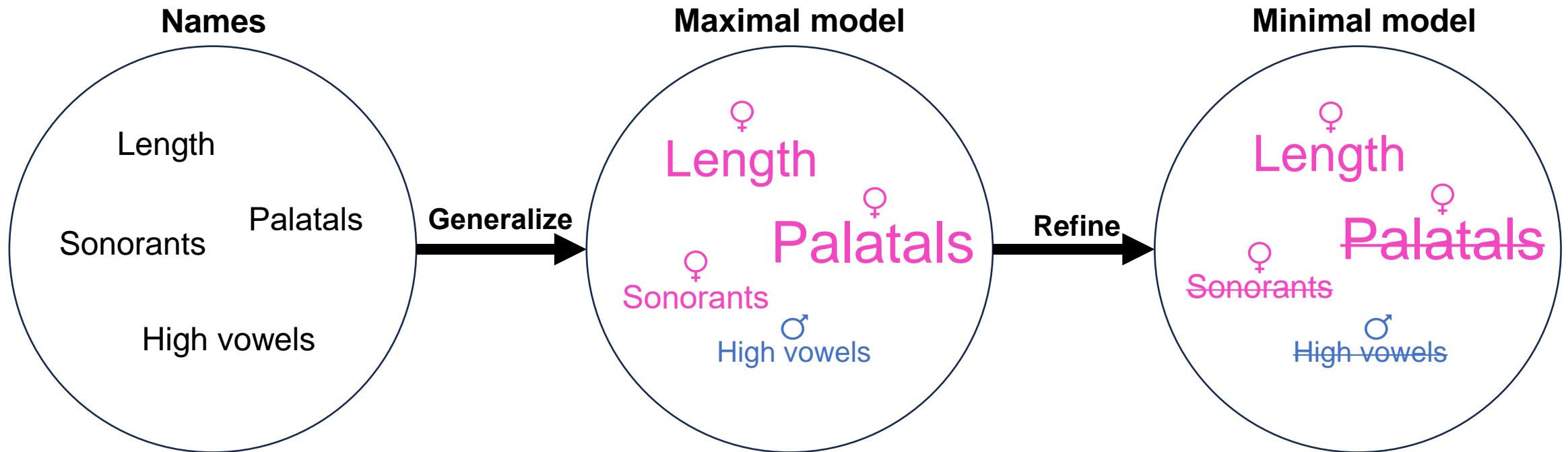
Rock: Harouy Jean-Michel via Wikimedia, 2019. Background removed. CC-BY-SA

# Conventional sound symbolism

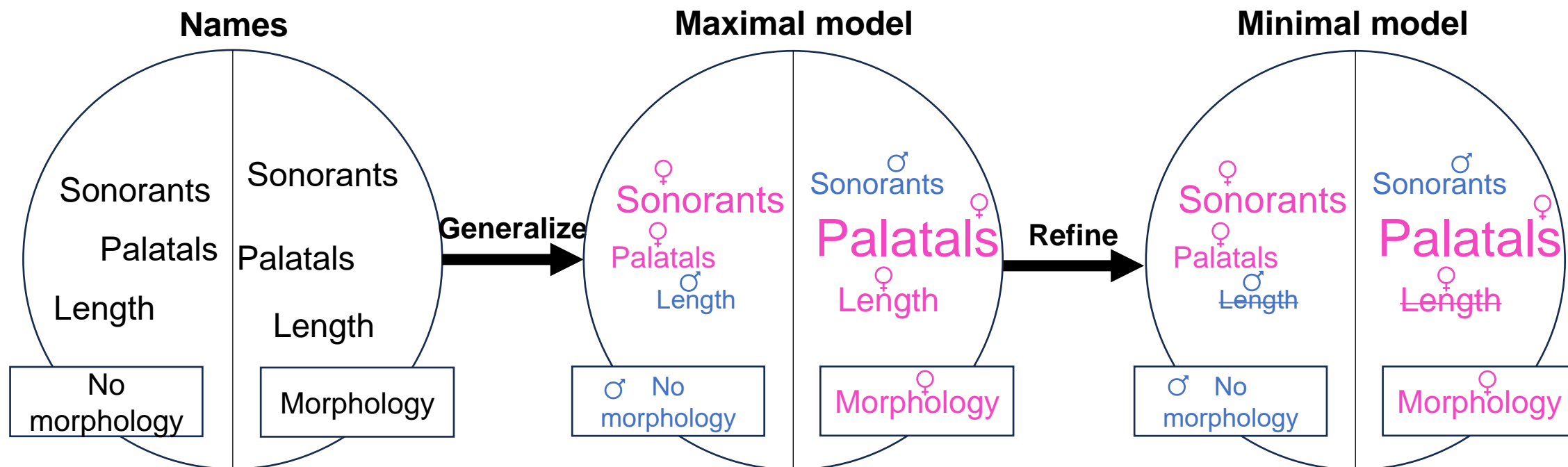


# Appendix C: Explaining logistic regression models

# Logistic regression models

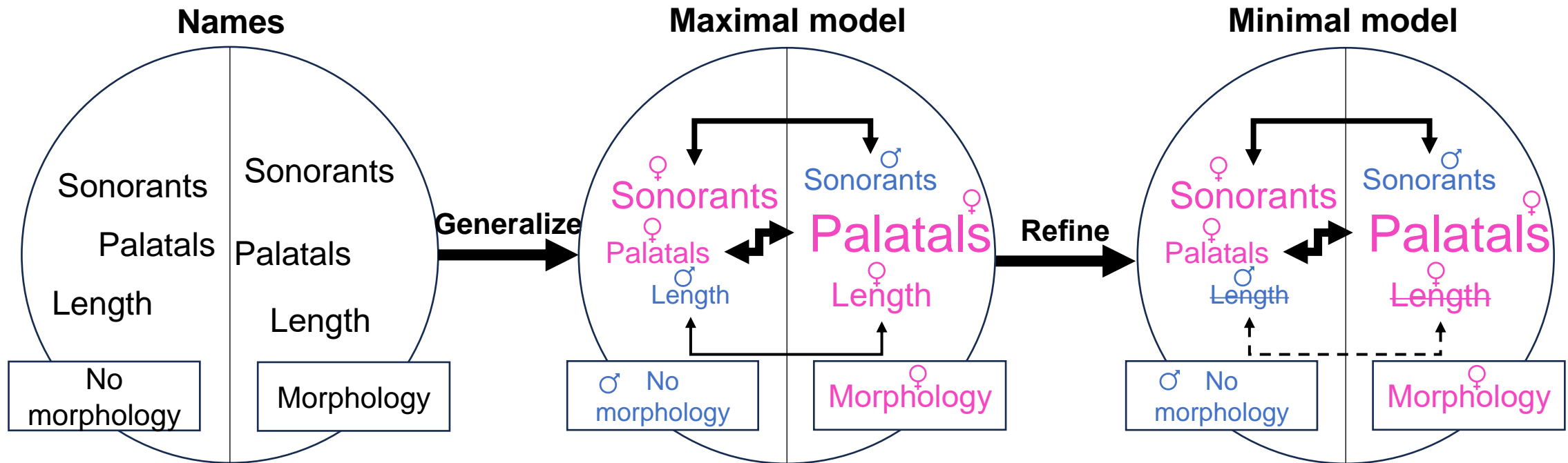


# Interaction effects

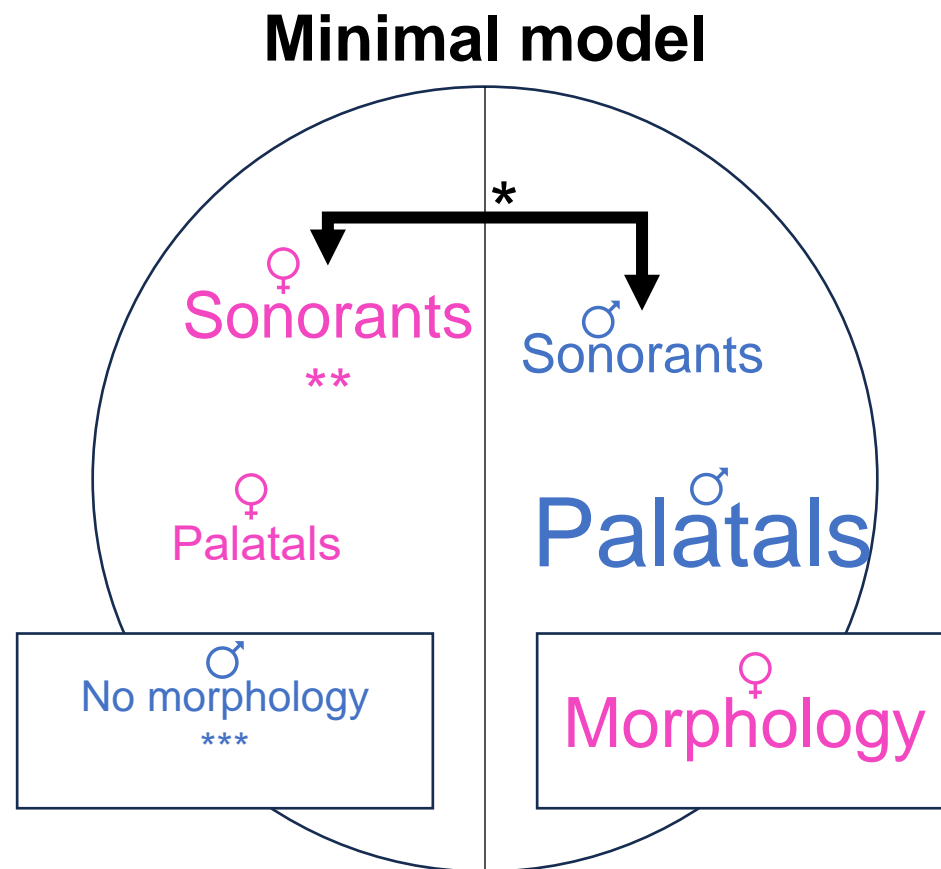




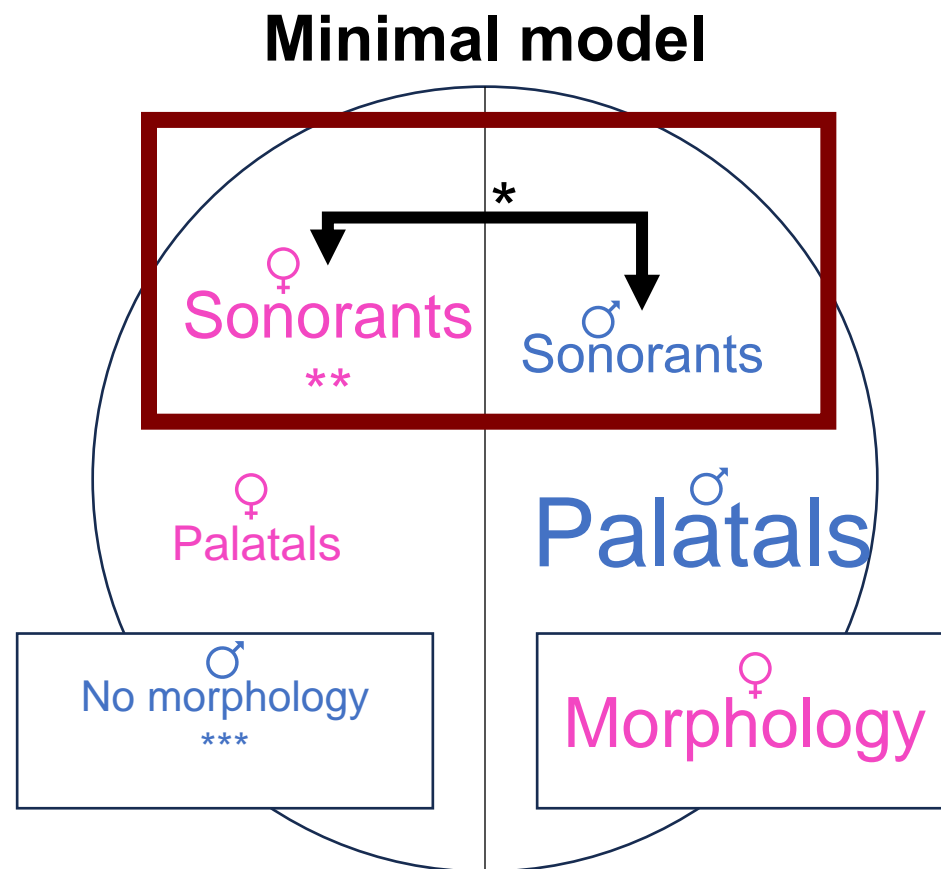
# Interaction effects



# Our findings



# Our findings



\*  $p < 0.05$   
\*\*  $p < 0.01$   
\*\*\*  $p < 0.001$

# Appendix D: Models

Negative coefficients lean feminine.

# Minimal model, only phonological variables

	<b>Coefficient</b>	<b>Std. Error</b>	<b>Z value</b>	<b>Pr(&gt;  Z )</b>
(Intercept)	-0.0125	0.1485	-0.084	0.933
Syllables	-1.3892	0.3368	-4.124	3.72e-05
<b>Log Likelihood</b>			-130.30	
<b>Akaike Information Criterion</b>			264.61	

# Maximal model, only phonological variables

	Coefficient	Std. Error	Z value	Pr(>  Z )
(Intercept)	-0.0212	0.1510	-0.141	0.88812
High Vowels	0.4336	0.6880	0.630	0.52855
Front Vowels	-0.2242	0.6402	-0.350	0.72615
Sonorants	-0.7072	0.5375	-1.316	0.18829
Palatals	-1.9820	0.9268	-2.139	0.03247*
Light Syllables	-0.2344	0.6687	-0.351	0.72594
Syllables	-1.2695	0.3569	-3.557	0.00038***
<b>Log Likelihood</b>			-127.40	
<b>Akaike Information Criterion</b>			268.79	

# Minimal model, with morphology as an interaction effect

	<b>Coefficient</b>	<b>Std. Error</b>	<b>Z value</b>	<b>Pr(&gt;  Z )</b>
(Intercept)	1.0443	0.2163	4.827	1.38e-6***
Apparent Morphology	-6.7789	286.5978	-0.024	0.98113
Sonorants	-2.2891	0.7779	-2.943	0.00325**
Palatals	-1.2661	1.4150	-0.895	0.37091
App Morph:Sonorants	2.8054	1.4189	1.977	0.04802*
App Morph:Palatals	-47.4993	3262.3920	-0.015	0.98838
	<b>Log Likelihood</b>		-96.73	
	<b>Akaike Information Criterion</b>		205.46	

# Maximal model, with morphology as an interaction effect

	Coefficient	Std. Error	Z value	Pr(>  Z )
(Intercept)	1.0719	0.2410	4.446	8.74e-06***
Apparent Morphology	-7.4064	232.3169	-0.032	0.97457
High Vowels	0.2298	0.8147	0.282	0.77786
Front Vowels	-0.0934	0.7612	-0.123	0.90230
Sonorants	-2.3470	0.8092	-2.901	0.00373**
Palatals	-1.3221	1.4382	-0.919	0.35792
Light Syllables	-0.7021	0.8742	-0.803	0.42191
Syllables	0.1963	0.5704	0.344	0.73075
App Morph:High Vowels	9.7718	4.9479	1.975	0.04828*
App Morph:Front Vowels	-11.506	5.1371	-2.240	0.02510*
App Morph:Sonorants	3.4167	1.6176	2.112	0.03468*
App Morph:Palatals	-54.8929	2644.4998	-0.021	0.98344
App Morph:Light Syllables	0.4160	2.1430	0.194	0.84609
App Morph:Syllables	-1.9817	1.3500	-1.468	0.14214

---

**Log Likelihood** -92.28

**Akaike Information Criterion** 212.55

---



# Appendix E: Post-hoc interaction effects testing

Done using the *phia* package (Posit team 2023)

# Minimal model, sonorants

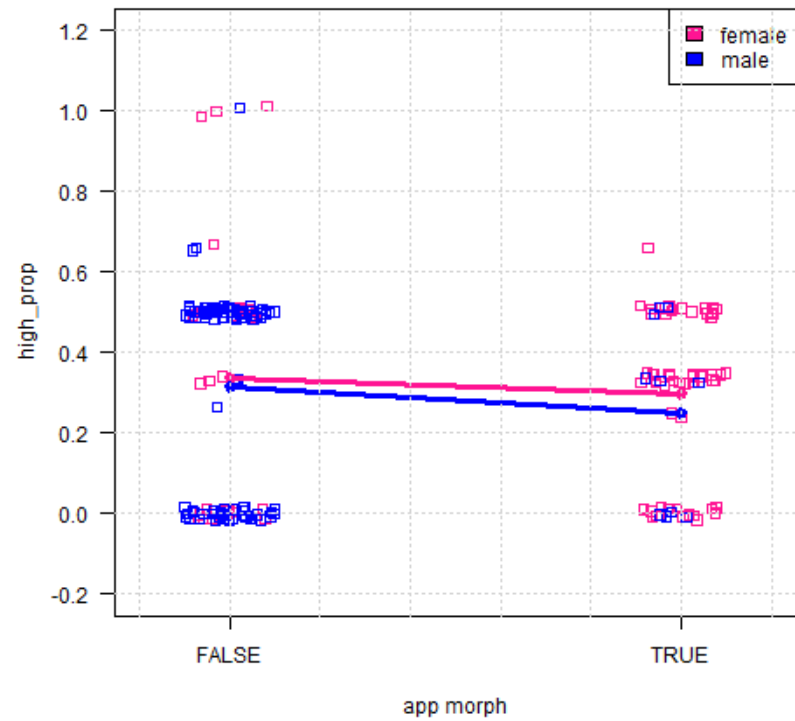
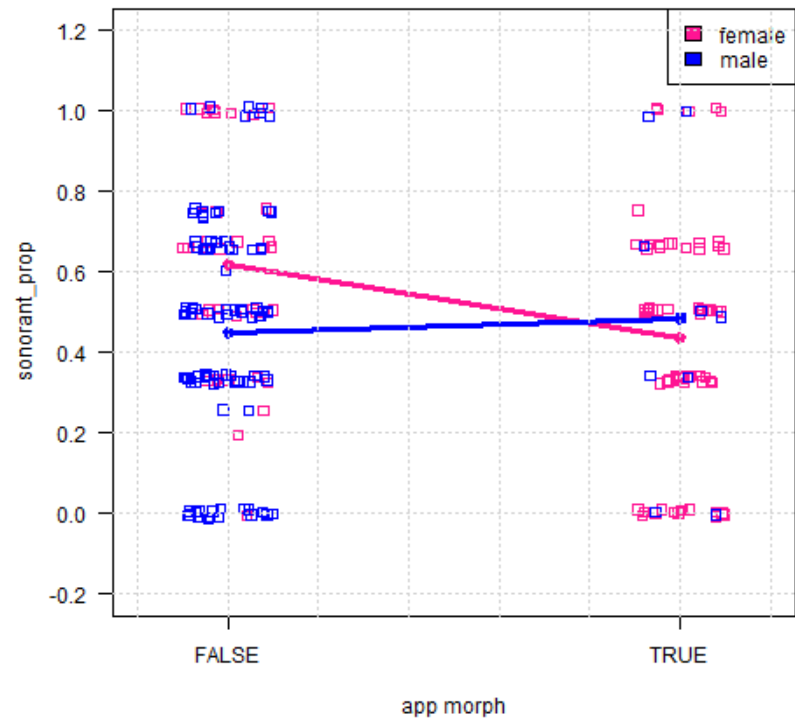
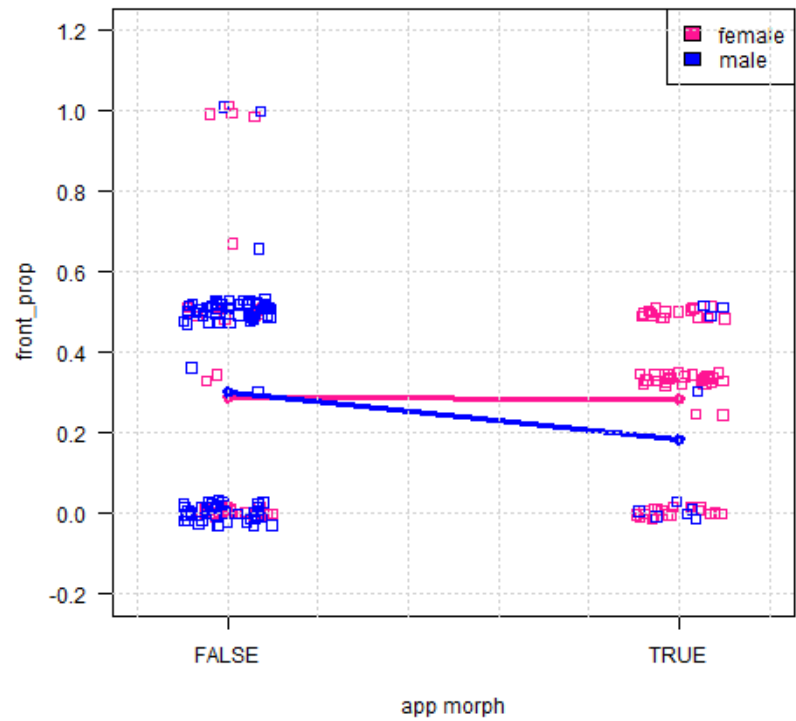
	Value	$\chi^2$	Pr(> $\chi^2$ )
No apparent morphology	-2.28914	8.6591	0.006509**
Apparent Morphology	0.51624	0.1893	0.663518

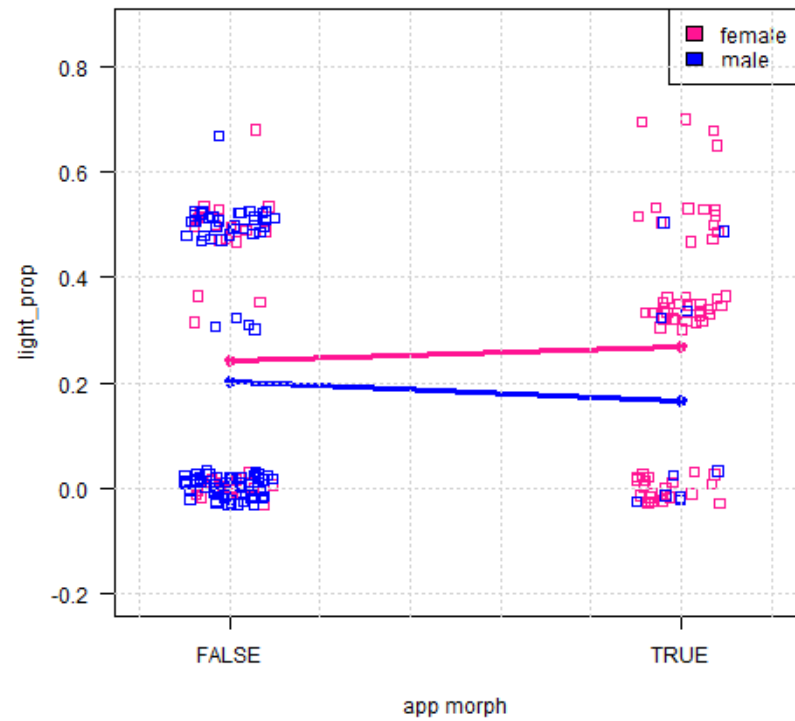
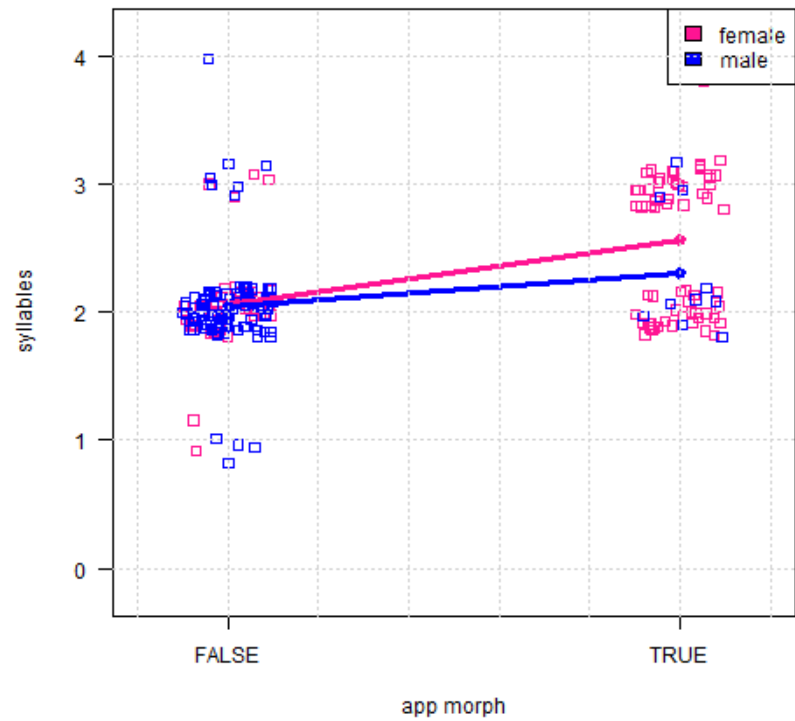
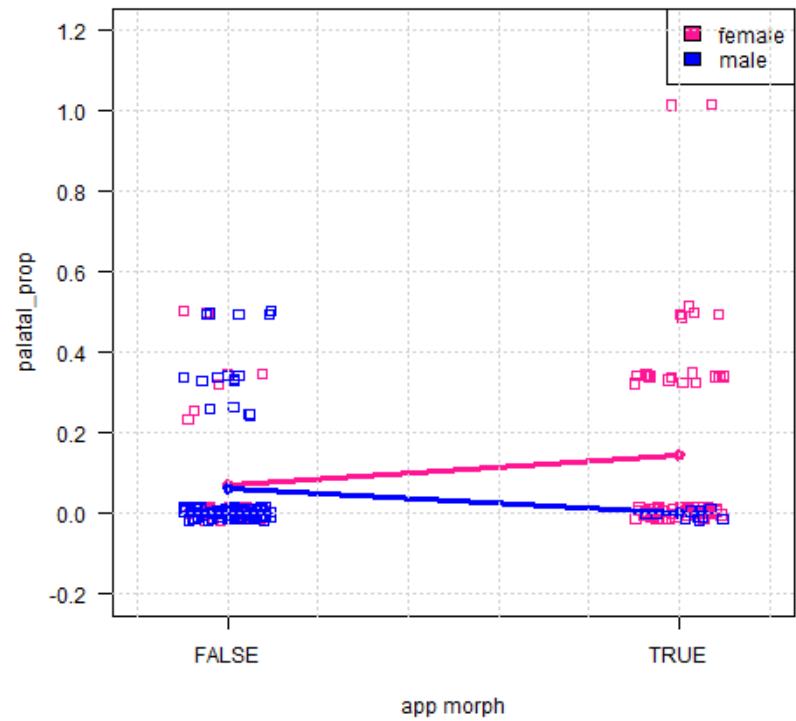
# Minimal model, palatals

	Value	$\chi^2$	Pr(> $\chi^2$ )
No apparent morphology	-1.266	0.8006	0.7418
Apparent Morphology	-48.765	0.0002	0.9881

# Appendix F: Stripcharts of phonological variables across apparent morphology

x-axis is morphology, colour/shape indicates gender, y-axis is phonological variable





# Appendix G: Drawing generalizations in logistic regression models

# Names are marked for variables

Muneeb

-Length  
+Sonorants  
-Palatals  
+High vowels

Maryam

+Length  
++Sonorants  
+Palatals  
High vowels

Ayesha

+Length  
Sonorants  
++Palatals  
High vowels

Sohaib

-Length  
+Sonorants  
Palatals  
High vowels



# Patterns are found

Muneeb

-Length  
+Sonorants  
-Palatals  
+High vowels

Maryam

+Length  
++Sonorants  
+Palatals  
High vowels

Ayesha

+Length  
Sonorants  
++Palatals  
High vowels

Sohaib

-Length  
+Sonorants  
Palatals  
High vowels

# Generalizations are drawn

++Sonorants

+High vowels

+Length

++Palatals