THE STORY OF AN AFRICAN LAB

CLICK PERCEPT!ON

EXPERIMENTS IN XHOSA AND ZULU

Aaron Braver†‡, Wm G Bennett†,
Camilla Christie†, Tyler Miller†, Khethani Yende†

† Rhodes University
‡ Texas Tech University

bit.ly/lsa2025clicks

A BIT OF A CONUNDRUM

- Why are clicks…?
  - Hand 1: super salient (the only Cs consistently louder than Vs, e.g.)
  - Hand 2: rare, restricted, limited in distribution qua consonants
- Contradictory intuitions afoot
  - Clicks are very easy to recognize as different from nonclicks
  - Click vs click distinctions seem much less easy to discern

P-MAPPING CLICKS

- Steriade (2001/2008) P-map: confusability of a given contrast…
  - is different in different contexts
  - projects relative ranking of faithfulness constraints
    (less distinct ~> less important to distinguish)

- Oddball expectations for clicks:
  - faithfulness for clickiness is supreme
    (predicts clicks are hard to get rid of)
  - faithfulness among clicks less crucial
    (because less salient)

a.

Obstruent voicing	V_V	C_V	V_R	V_]	V_T	C_T
p/ b	p/b	p/b	p/b	p/b	p/b	p/b
t/ d	t/d	t/d	t/d	t/d	t/d	t/d
k/ g	k/g	k/g	k/g	k/g	k/g	k/g
s/z	s/z	s/z	s/z	s/z	s/z	s/z

CLICKS SHOULD BE EASY TO PERCEIVE

- Ladefoged & Traill (1994:45), !Xóõ
  - “clicks are probably the most salient consonants”
  - clicks easier to ID than non-clicks
- masking level test: confounds loudness and spectral distribution
- only looked at plain [ʘ | ! ǁ ǂ] vs. pulmonic consonants; doesn’t establish click~click baselines

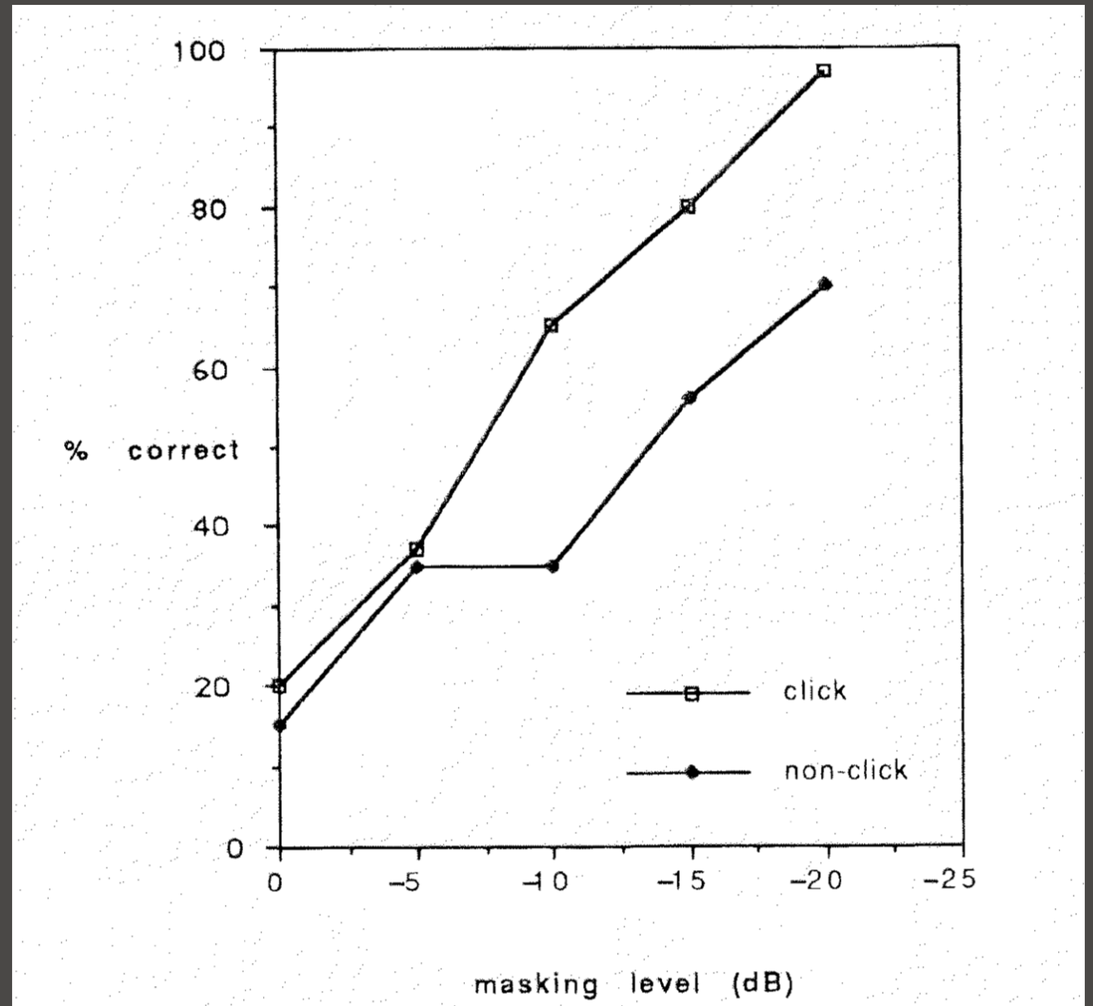


Figure 6. The perceptual saliency of clicks and non-clicks.

CLICKS SHOULD BE EASY TO PERCEIVE

- Most previous studies of click perception focus on non-native listeners (Best et al. 1988, 1999, 2003, 2008, 2020, among others)
  - American click-naïve listeners do not perceive clicks as speech sounds
  - Click-naïve listeners still extremely good at AXB discrimination (worst participant still 81% correct, cf. Zulu listener avg. 87%)
- Best et al. (2003) compare Sesotho & Zulu listeners, but both groups listened only to fricated clicks from !Xóõ, which are an obviously non-native category for them
- Point: none of these establish clear baseline expectations for the perceptibility of, e.g., different Zulu click contrasts by Zulu-speaking listeners

CLICKS ARE PERCEIVED AS DIFFICULT

- Anecdotal evidence abounds that people regard clicks as complex and difficult to master
  - probably not just an artifact of eurocentrism in descriptions (“*exotic!*”)
  - L1 Xhosa & Zulu speaking undergraduates report this impression too

- Try it and see
  - click contrast test 1
  - click contrast test 2

FAILED ATTEMPTS AT REPLICATING PREVIOUS FINDINGS

- Pilot 1: ABX task, attempting to probe type vs accompaniment contrasts, using v!v intervals excised from real words, both Xhosa and click-experienced English speakers (Miller 2020)
  - all participants at chance (top performer ~54%) 😵

- Pilot 2: AXB task, new stimuli recordings of nonce a!a sequences
  - all participants still at chance; they can't discriminate
  - very much at odds with Best's findings; even L1 listeners failed to recover the contrasts from this set of stimuli

TODAY'S EXPERIMENT: DESIGN

- AXB paradigm, audio presentation, all CVCV nonce sequences
- Bipartite structure, stimuli recorded by Khethani Yende
- Side A: click type contrasts (c ~ q ~ x) (dental ~ alveolar ~ lateral)
  - intended to probe for differences based on dialect/accent
  - is c ~ q variation rooted in perceptual difference? (Yende 2023, in prep)
- Side B: two tranches
  - "Count" (!ada vs !a!a) (cf. Gallagher 2010)
  - "Site" (!ada vs da!a)
- Each side served as distractors from the other

PARTICIPANTS

- 12 L1 speakers of Zulu recruited by Khethani Yende using word of mouth + snowball sampling
- Targeted recruitment to probe for regional/dialect variation
  - one group of speakers from Gauteng (=urban Zulu, "Sowetan" Zulu)
  - one group of speakers from KZN ("proper" Zulu, "deep" Zulu)
- All could also speak multiple languages (English, etc)
- No other participant factors considered here

PERCEPTIBILITY VARIATION BY CLICK TYPE

c ~ q	😩
c ~ x	😶
x ~ q	😶

- “Click type”: synonymous with front closure place distinctions

  [IPA key: c = | x = ǁ q = !]

- c ~ q: attested as free variation (Gunnink 2014)

  ~> so maybe is harder to discern?

- c ~ x: predictions unclear
- x ~ q: predictions unclear

PERCEPTIBILITY VARIATION BY CLICK CONTEXT

da!a	
!ada	
!a!a	

- Contradictory intuitions here
  - !ada > da!a would make sense bc initial prominence
  - But: da!a > !ada makes sense if the preceding vowel carries some of the cues
- Cross-interaction likely; different contrasts are made with different cues (burst vs VOT, e.g.)
- Also, count is a thing we looked at (!ada vs !a!a)

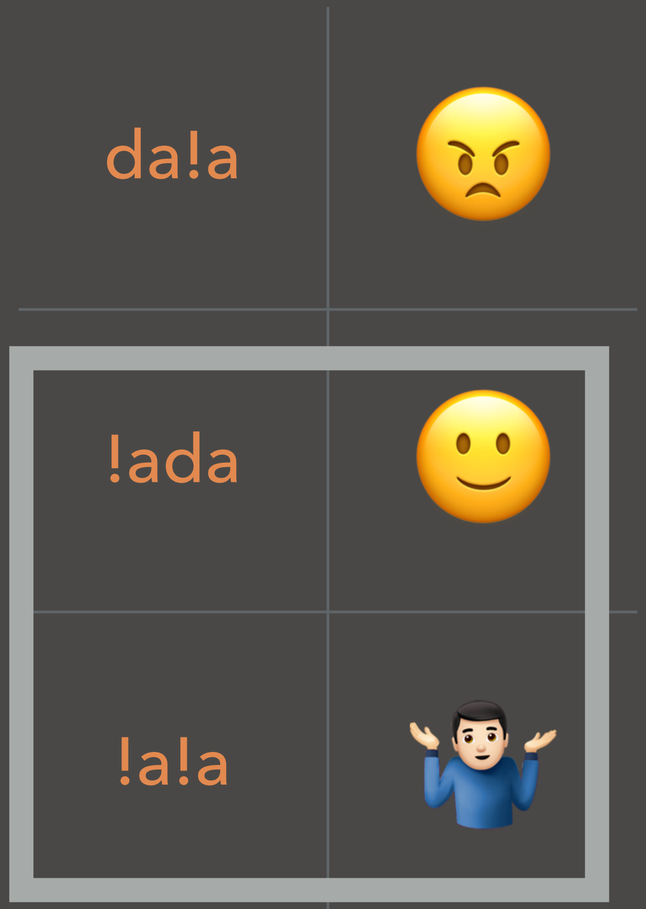
PERCEPTIBILITY VARIATION BY CLICK CONTEXT

"Site"

da!a	
!ada	
!a!a	

- Contradictory intuitions here
  - !ada > da!a would make sense bc initial prominence
  - But: da!a > !ada makes sense if the preceding vowel carries some of the cues
- Cross-interaction likely; different contrasts are made with different cues (burst vs VOT, e.g.)
- Also, count is a thing we looked at (!ada vs !a!a)

PERCEPTIBILITY VARIATION BY CLICK CONTEXT



"Count"

- Contradictory intuitions here
  - !ada > da!a would make sense bc initial prominence
  - But: da!a > !ada makes sense if the preceding vowel carries some of the cues
- Cross-interaction likely; different contrasts are made with different cues (burst vs VOT, e.g.)
- Also, count is a thing we looked at (!ada vs !a!a)

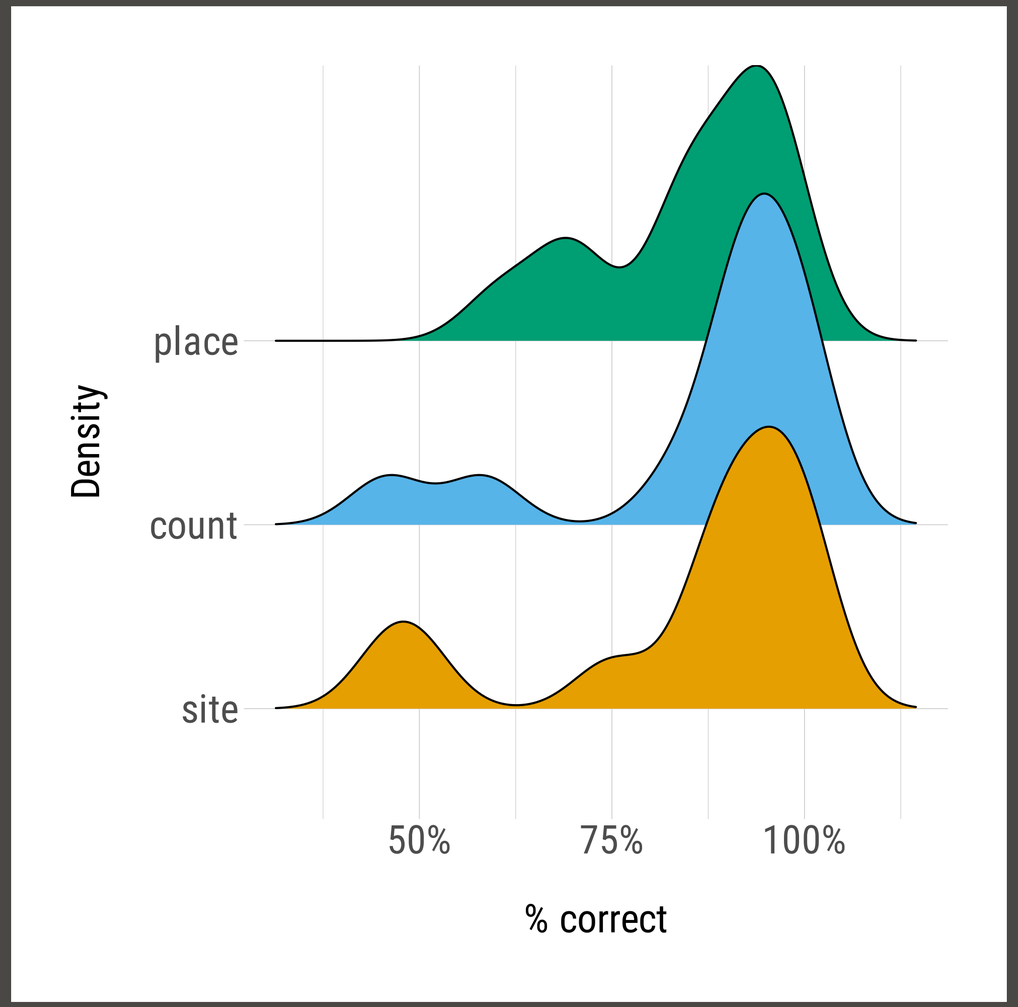
RESULTS

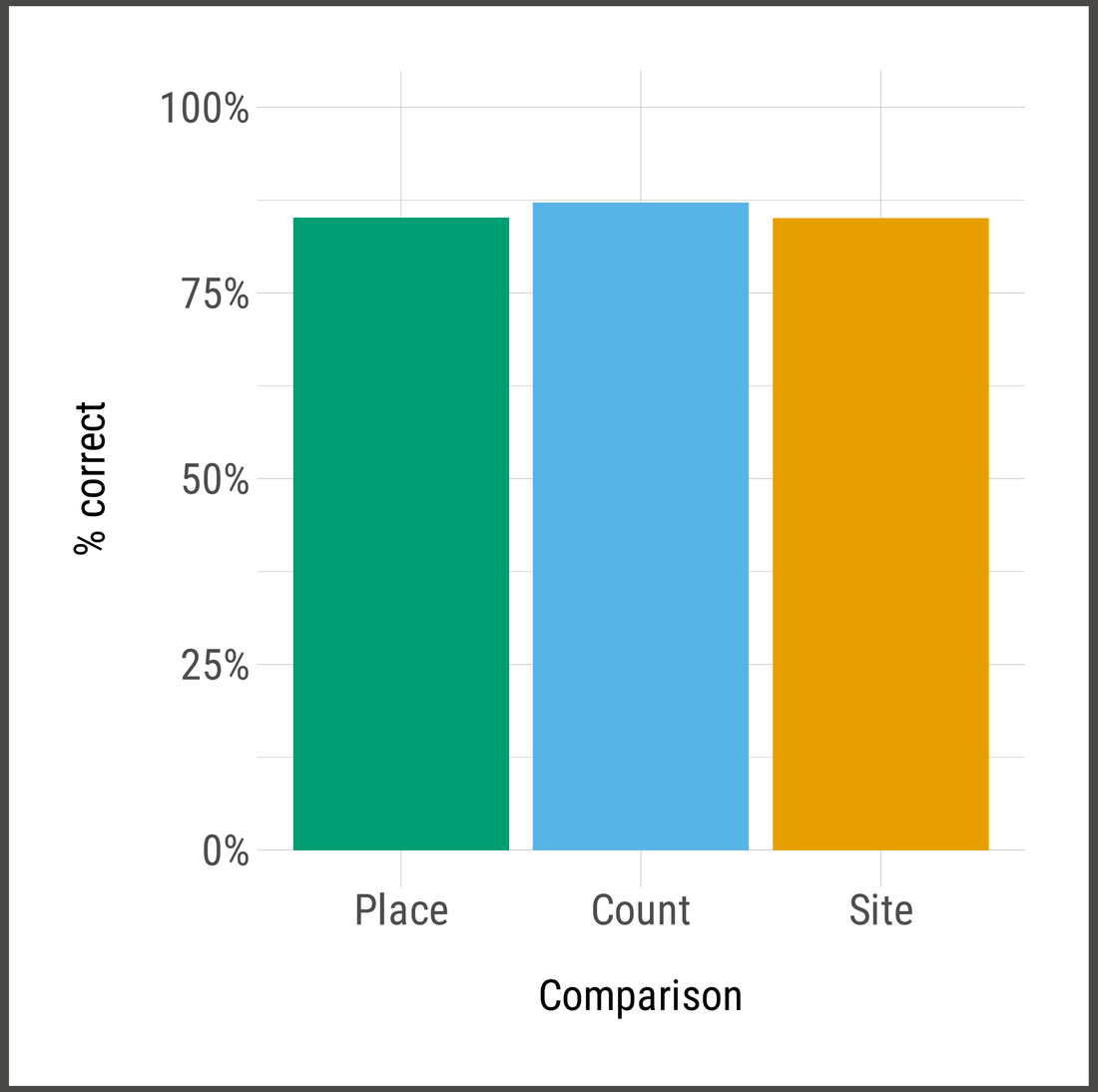
ANALYSIS

- Logistic mixed effects model (via `glmer`)
  - Fixed effects: tranche, place, site, correct answer, correct button
  - Random effects: intercepts for item and participant
- Marginal means computed and pairwise tests (via `marginaleffects`)

BIRD'S EYE VIEW

- Correct perception is similar in each tranche:
  - place 85.14%
  - count 87.15%
  - site 85.07%
- More in line with Best's prior work





OVERALL RESULTS BY TRANCHE

- place 85.14%
- count 87.15%
- site 85.07%

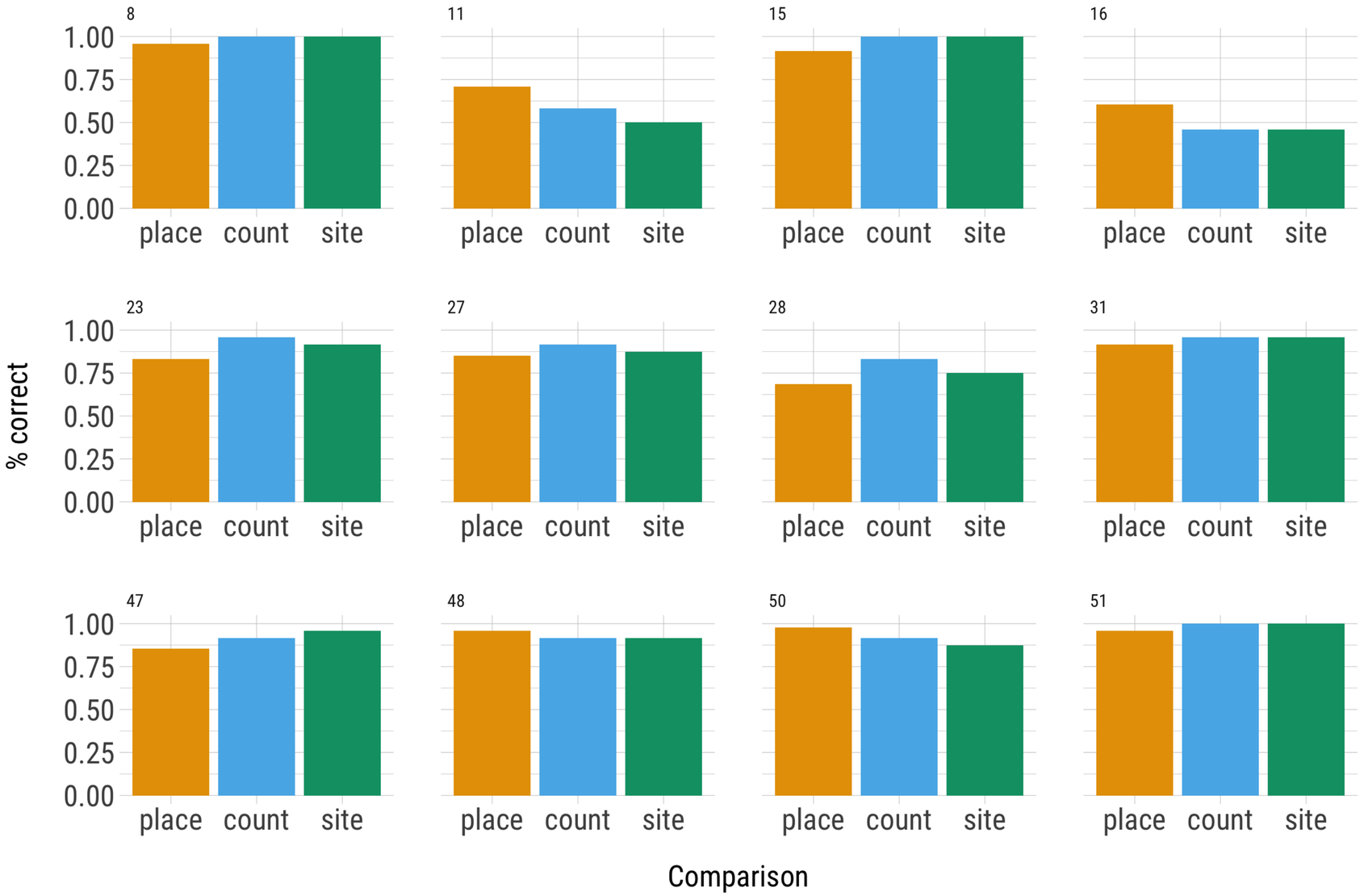
Model

Parameter	Coeff.	CI	z	p
(Intercept)	2.23	[1.36, 3.09]	5.05	0.00
correct ans [B]	0.29	[-0.14, 0.73]	1.32	0.19
tranche [count]	0.28	[-0.46, 1.01]	0.73	0.46
tranche [site]	-0.05	[-0.58, 0.48]	-0.19	0.85
place [q]	-0.04	[-0.58, 0.51]	-0.13	0.90
place [x]	-0.51	[-1.04, 0.02]	-1.88	0.06
site [2]	0.25	[-0.25, 0.75]	0.99	0.32
site [b]	0.07	[-0.82, 0.96]	0.15	0.88

Marginal means for tranche

	Est.	statistic	p	CI
count - place	0.03	0.75	0.45	[-0.04, 0.10]
site - place	-0.01	-0.19	0.85	[-0.06, 0.05]

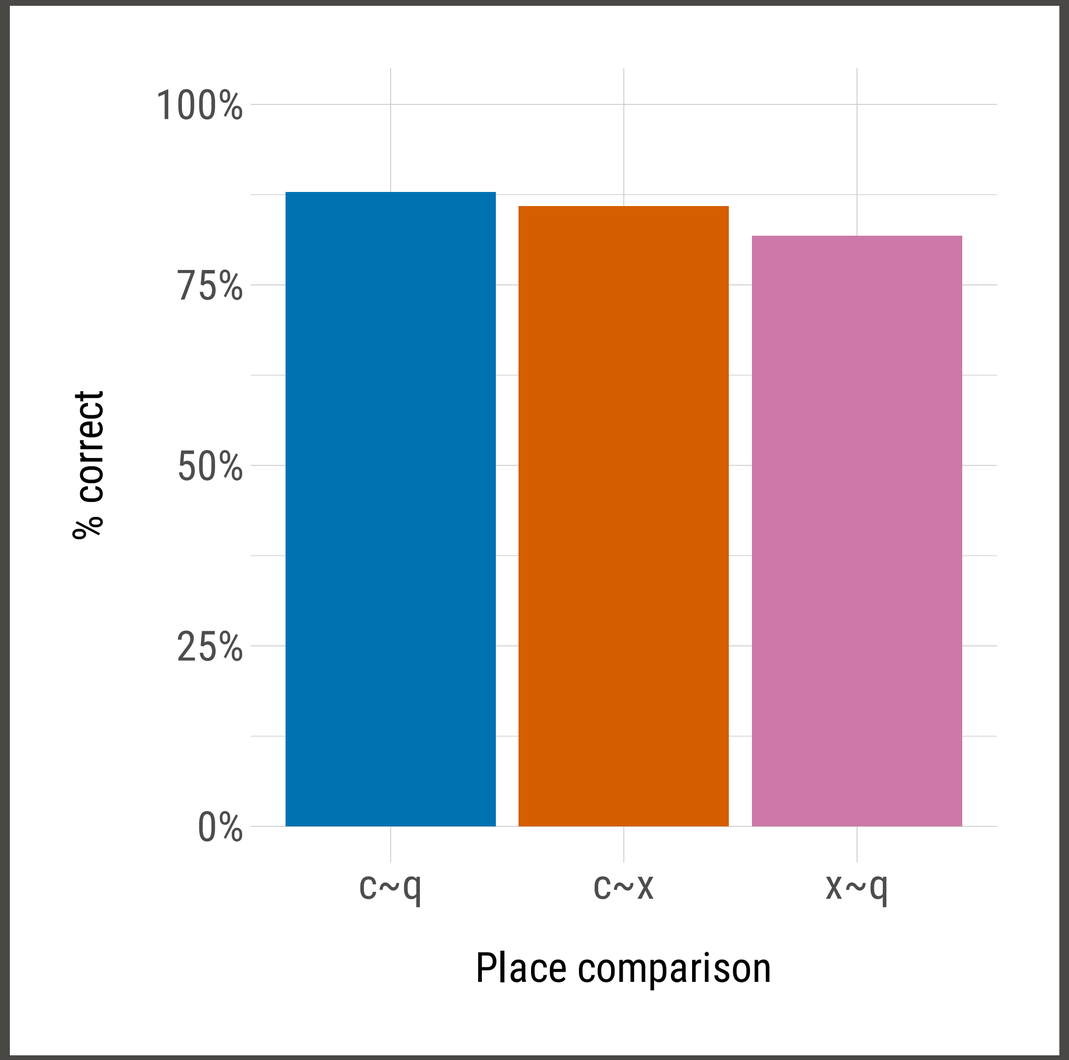
RESULTS BY TRANCHE AND SPEAKER



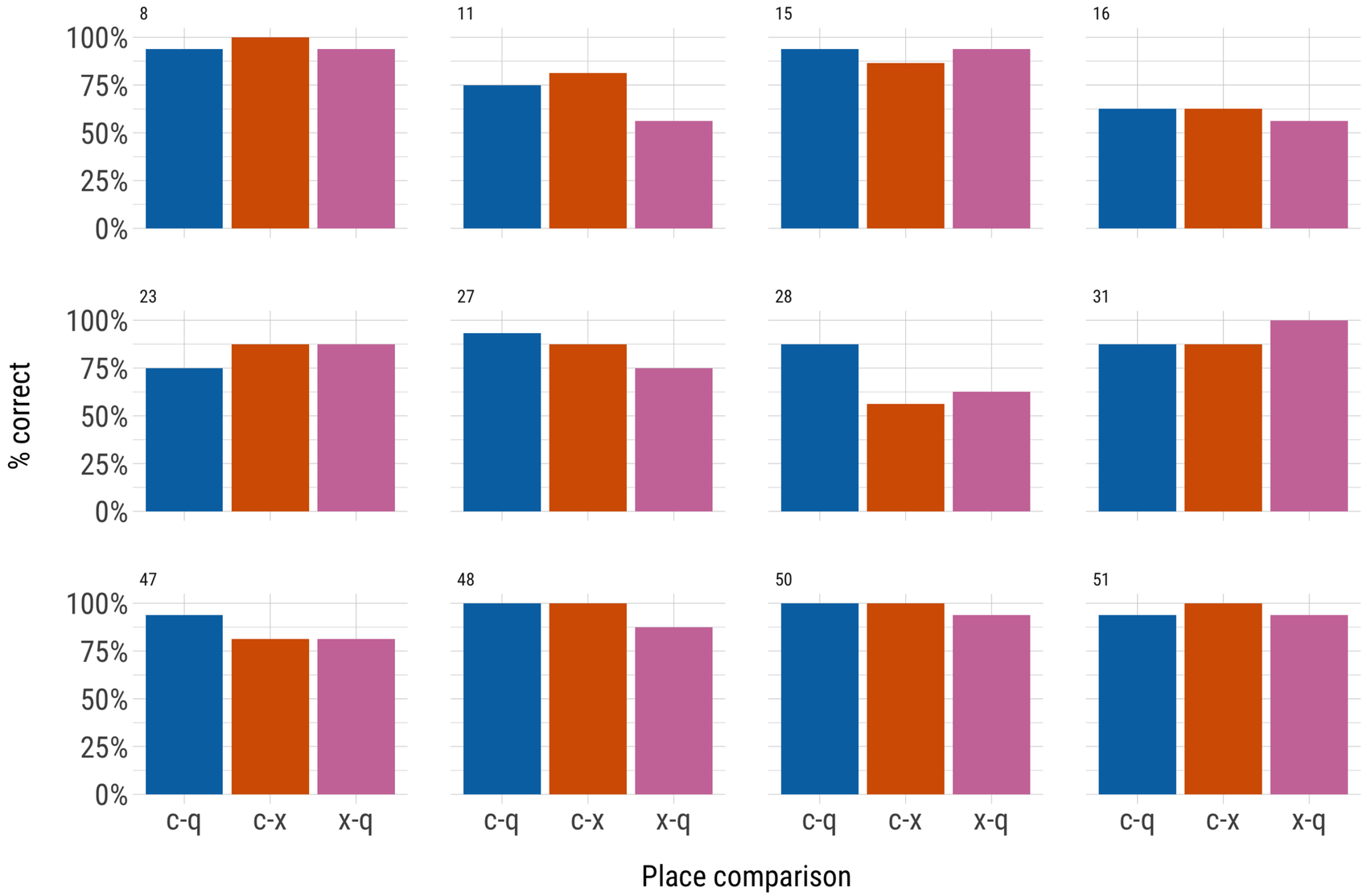
PLACE COMPARISONS

- Mean percent correct: 85.14%
- Place comparison type is not a significant predictor in the model
- Pairwise comparison of marginal means show no differences between any pair of place comparisons

Comparisons	Contrast	conf.low	conf.high	p.value
x-q - c-q	-0.06	-0.16	0.04	0.22
x-q - c-x	-0.04	-0.14	0.06	0.46
c-q - c-x	0.02	-0.07	0.11	0.62

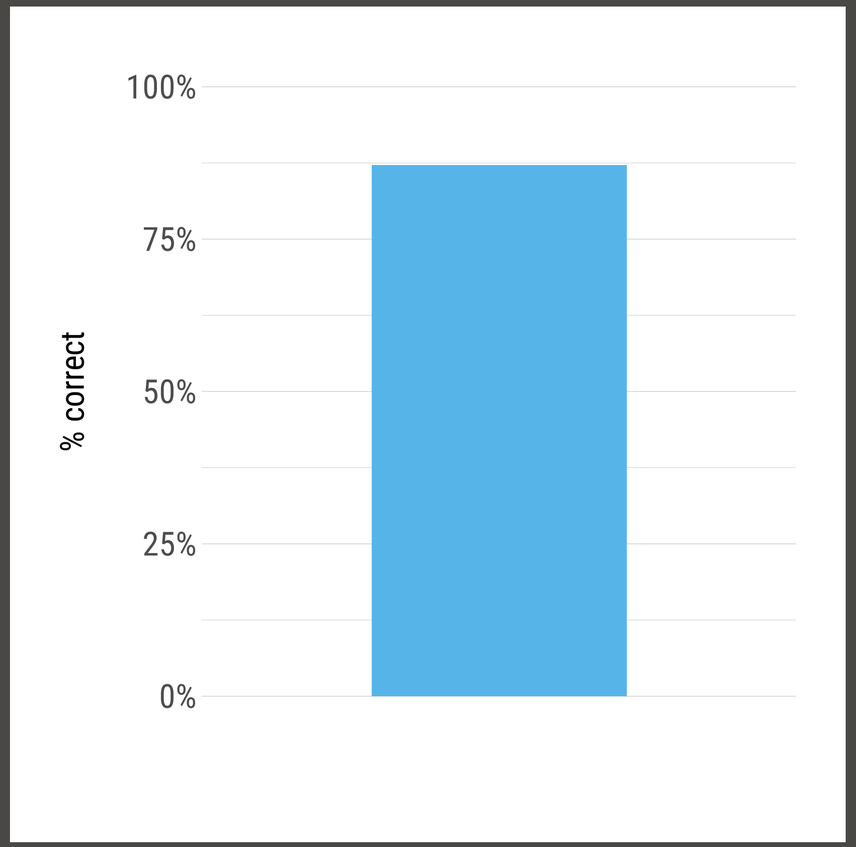


RESULTS BY PLACE COMPARISON AND SPEAKER - PLACE TRIALS



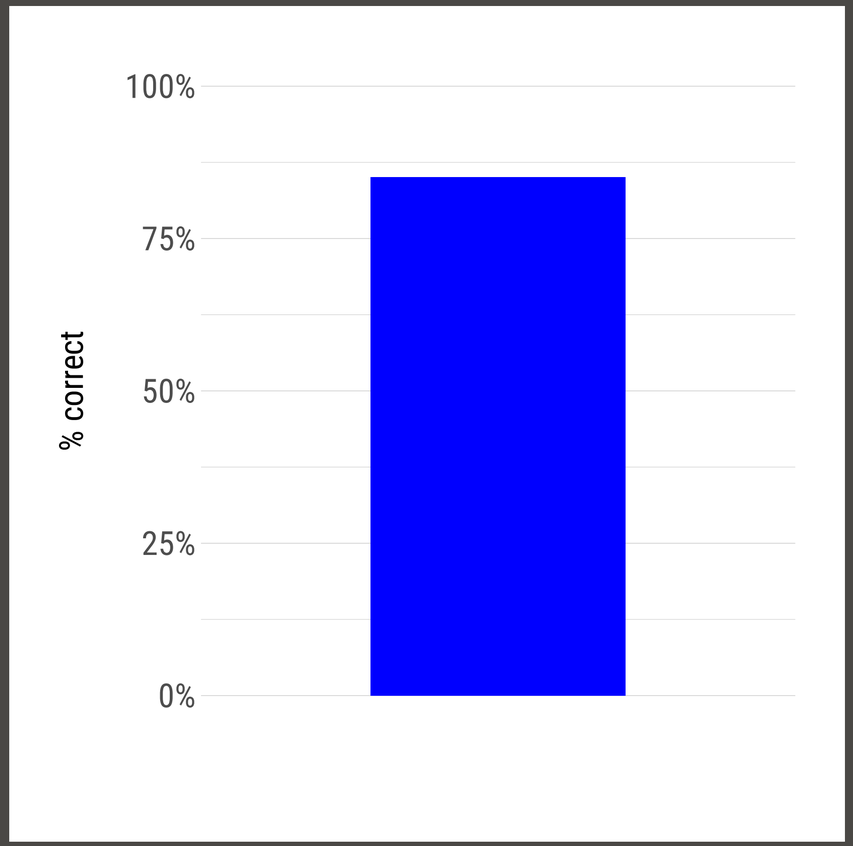
COUNT COMPARISONS

- Mean percent correct: 87.15%



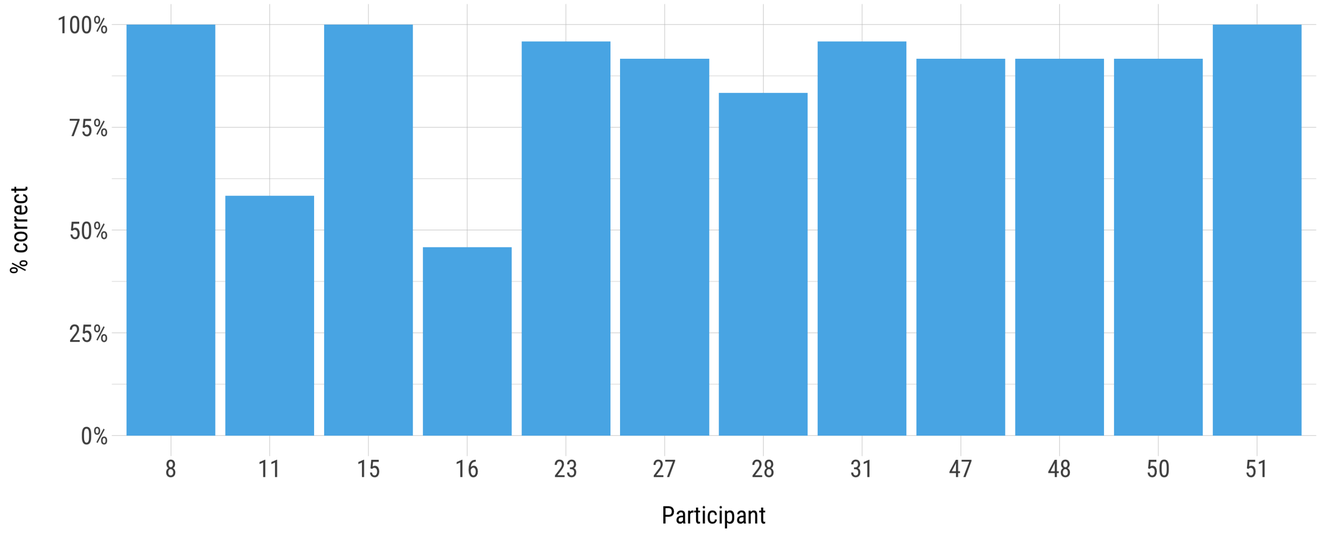
SITE COMPARISONS

- Mean percent correct: 85.07%

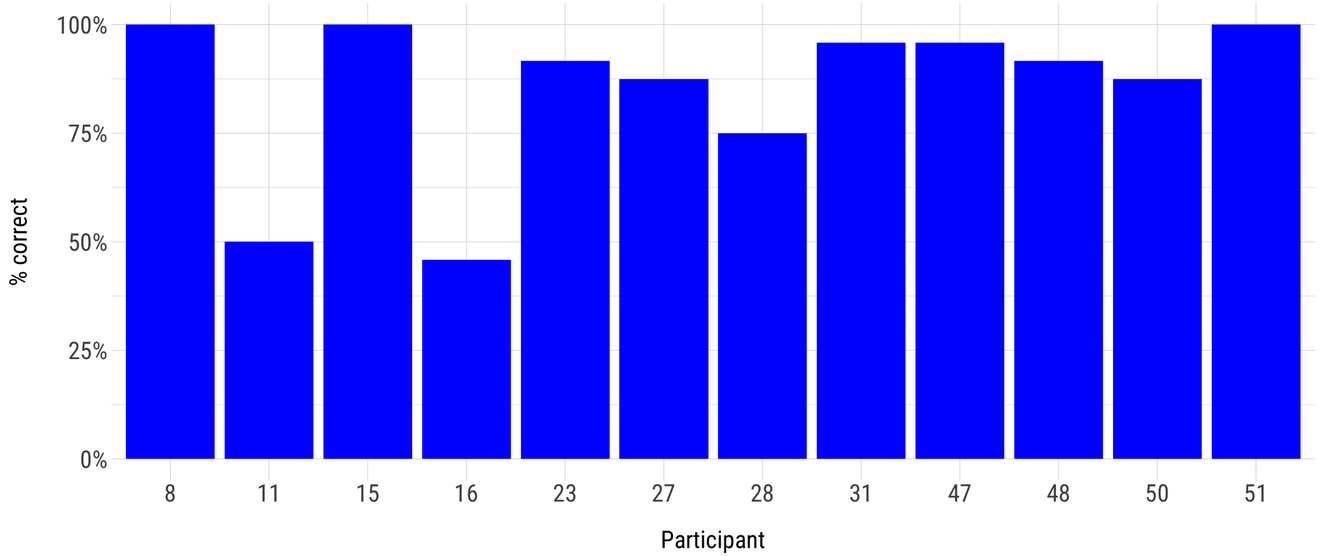


COUNT AND SITE BY SPEAKER

- Count



- Site



CONCLUSION

- We have established now that the procedure we used in expt 3 is a good way to get some functional baselines
- Pilot experiment kinks seem mostly worked out
  - AXB > ABX
- Further task effects remain to be studied
  - CV vs. VCV vs. CVCV stimuli

NEXT QUESTIONS

- Place (c ~ q ~ x) vs. accompaniment (q ~ nq ~ ngq ~ gq)
- Does accompaniment modify place perception?
- Do clicks pattern like ejectives with respect to perceptual biases of site and count?
- How common is metathesis as a perceptual illusion with clicks (ǫada perceived as daǫa)?
- Click-naive vs. click-experienced listeners (SA vs. US English L1s)
- Cue weighting within each accompaniment (e.g. pitch effects)

THANKS!

- Thanks are due to many people who assisted with or contributed to this work, including Bonny Sands, Catherine T. Best, Didier Demolin, Eva-Marie Bloom-Ström, Mark de Vos, Senamile Mchunu, Nancy Klaas, Zoleka Maqwili, Hlumi Kondile, Nezi Gangani, Sive Bavuma, Alaric Heim’s pet chameleon, and all of you fine folks.

REFERENCES

Best, Catherine T., Gerald W. McRoberts, and Nomathemba M. Sithole. 1988. “Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants.” Journal of Experimental Psychology: Human Perception and Performance 14: 45–60.

Best, Catherine T., and Robert A. Avery. 1999. “Left hemisphere advantage for click consonants is determined by linguistic significance.” Psychological Science 10: 65–69.

Best, Catherine T., Anthony Traill, Allyson Carter, K. David Harrison, and Alice Faber. 2003. “!Xóõ click perception by English, Isizulu, and Sesotho listeners.” Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona, Spain, August 3–9, 2003. Universitat Autònoma de Barcelona, pp. 853–856.

Best, Catherine T. 2020. Perception of Non-native Click Consonant Contrasts: Implications for Theories of Speech Perception. In: Sands, Bonny (ed.) Click consonants. Leiden: Brill.

Gallagher, Gillian. 2010. “Perceptual distinctness and long-distance laryngeal restrictions.” Phonology 27: 435-480.

Gunnink, Hilde. 2014. “The grammatical structure of Sowetan tsotsitaal.” Southern African Linguistics and Applied Language Studies 32: 161–171.

Ladefoged, Peter, and Anthony Traill. 1994. “Clicks and their accompaniments.” Journal of Phonetics 22: 33–64.

Miller, Tyler. 2020. isiXhosa Click Perception in Non-Native Speakers. Unpublished Honours thesis, Rhodes University.

Steriade, D. 2001. The Phonology of Perceptibility Effect: The P-map and Its Consequences for Constraint Organization. Manuscript, UCLA. Version updated 2008.

Yende, Khethani. 2023. Rural-Urban IsiZulu Click Phoneme Variation. Unpublished Honours thesis, Rhodes University.