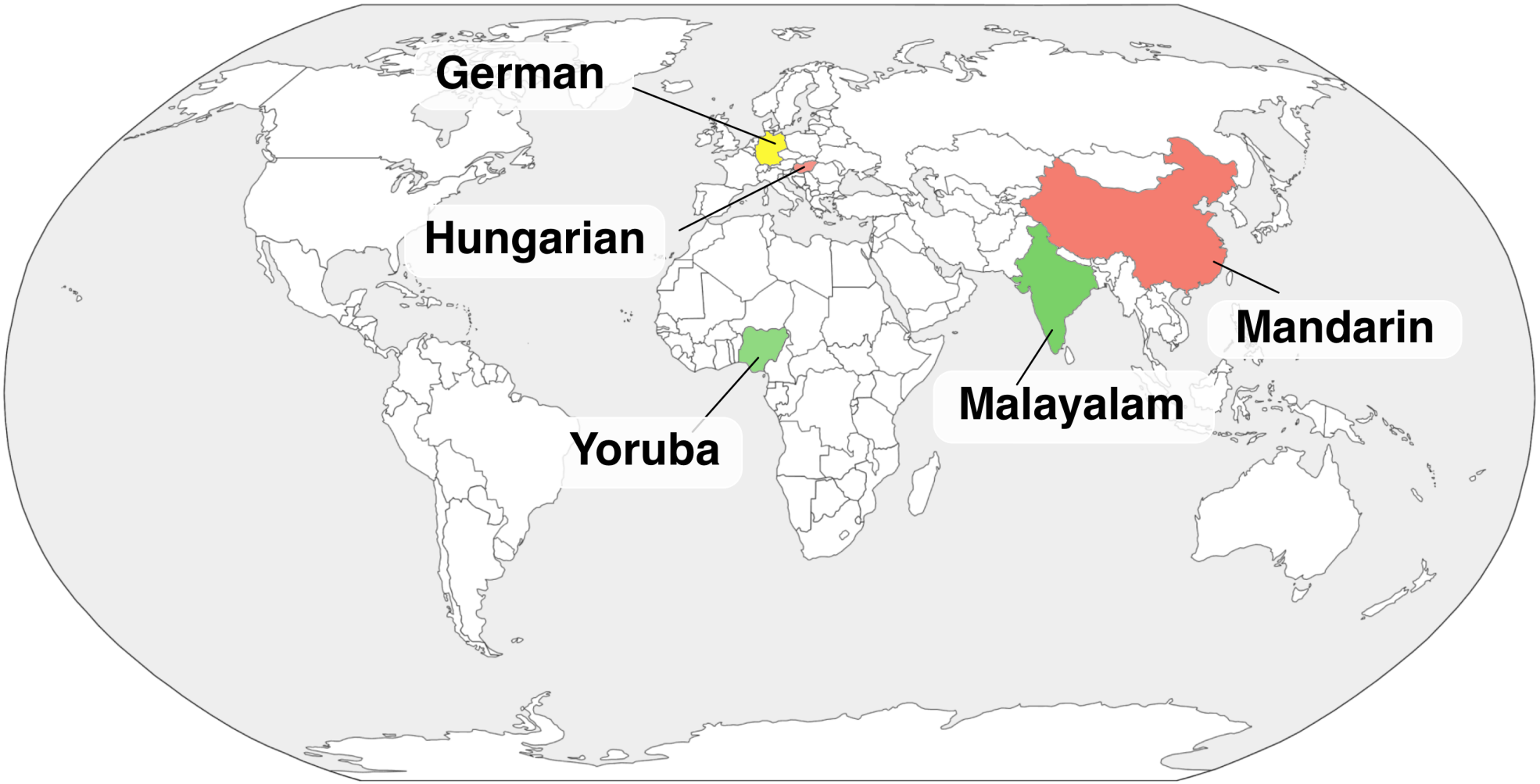
How children ask questions across languages

BUCLD 46, November 4-6, 2021

K. Yatsushiro, C. Dal Farra, A. Gonzalez, J. Hein, S. Silleresi, A. Avellana, G. Gayathri, A. Huang, J. Ilori, L. Pintér, M. Guasti, U. Sauerland

Geographical diversity



Linguistic Diversity

	Yes	No
Obligatory wh-movement	German, Hungarian, Yoruba	Malayalam, Mandarin
Case on wh-phrase	German, Hungarian, Malayalam	Mandarin, Yoruba
Verb final	German, Malayalam	Hungarian, Mandarin, Yoruba
Null argument	Hungarian, Mayalayam, Mandarin	German, Yoruba

Goals of the study

A. Linguistic Goals

1. Check to see whether we observe previously identified asymmetries between

   - Subject wh-questions vs. Object wh-questions (Ervin-Tripp 1970, Yoshinaga 1996, and others)

   - who-type questions vs. which N-type questions (Friedmann et al. 2009, Guasti et al. 2012)

   - non-animate (what) vs. animate (who) wh-phrases (Guasti 1996)

2. Check whether previously observed features of languages that affect comprehension of wh-questions (Sauerland et al. 2016 and others) also affect production.

B. Methodological goal:

Investigate less studied languages, in collaboration with local investigators.

Experiment

Participants

- Participants: between (roughly) 4;0 to 6;0

	#	Age Range	mean score of parental education level
German	22	3;10-6;0 (*M*=5;0)	
Hungarian	20	5;3-6;2 (*M*=4;7)	4.87
Mandarin	16	4;3-6;2 (*M*=5;10)	4.76
Malayalam	22	4;1-5;0 (*M*=4;7)	4.47
Yoruba	12	3;2-6;7 (*M*=5;1)	4.25

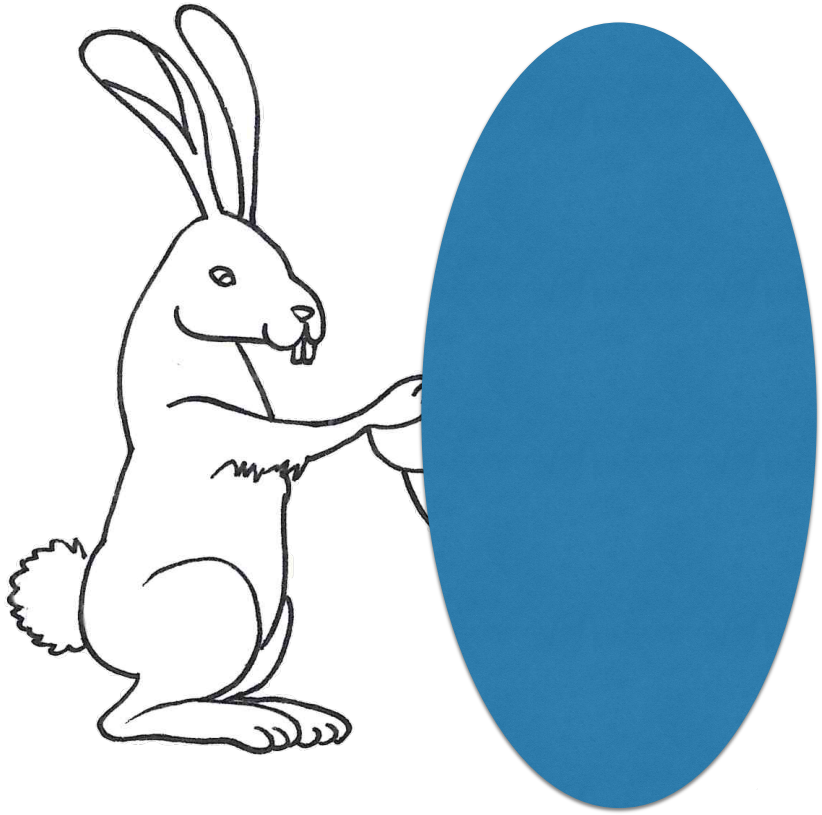
1: up to 2nd grade
2: up to 6
3: up to 10
4: more than 10
5: college

Material

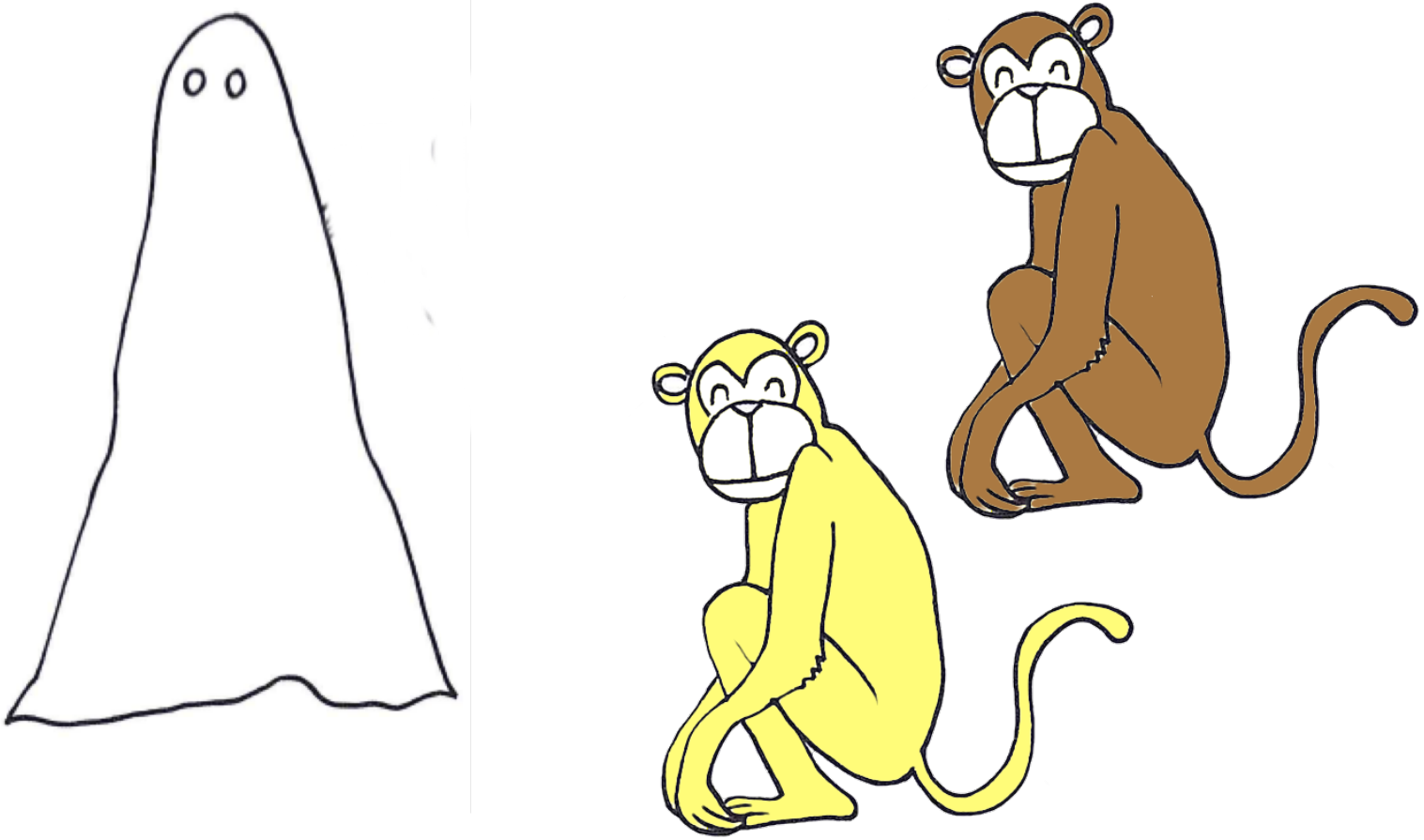
- 5 types of target questions (5 items each):
  - Who-subject question: “**Who** is scratching the monkey?”
  - Who-object question: “**Who** is the rabbit scratching?”
  - Which-subject question: “**Which monkey** is scratching the ghost?”
  - Which-object question: “**Which frog** is the mouse scratching?
  - What-object question: “**What** is the boy hiding?”

- The experimental design: modeled after Guasti et al. (2012)

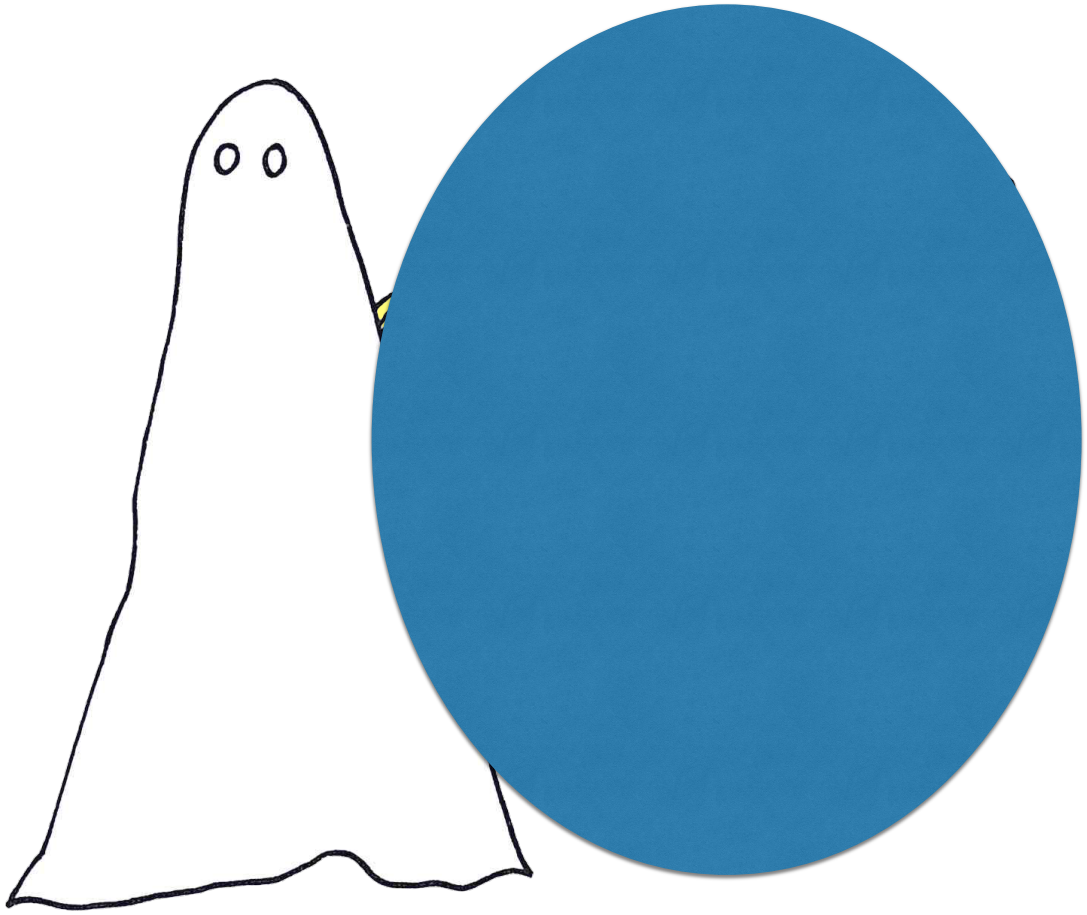
An example: Who is the rabbit scratching?



An example: Which monkey is scratching the ghost?

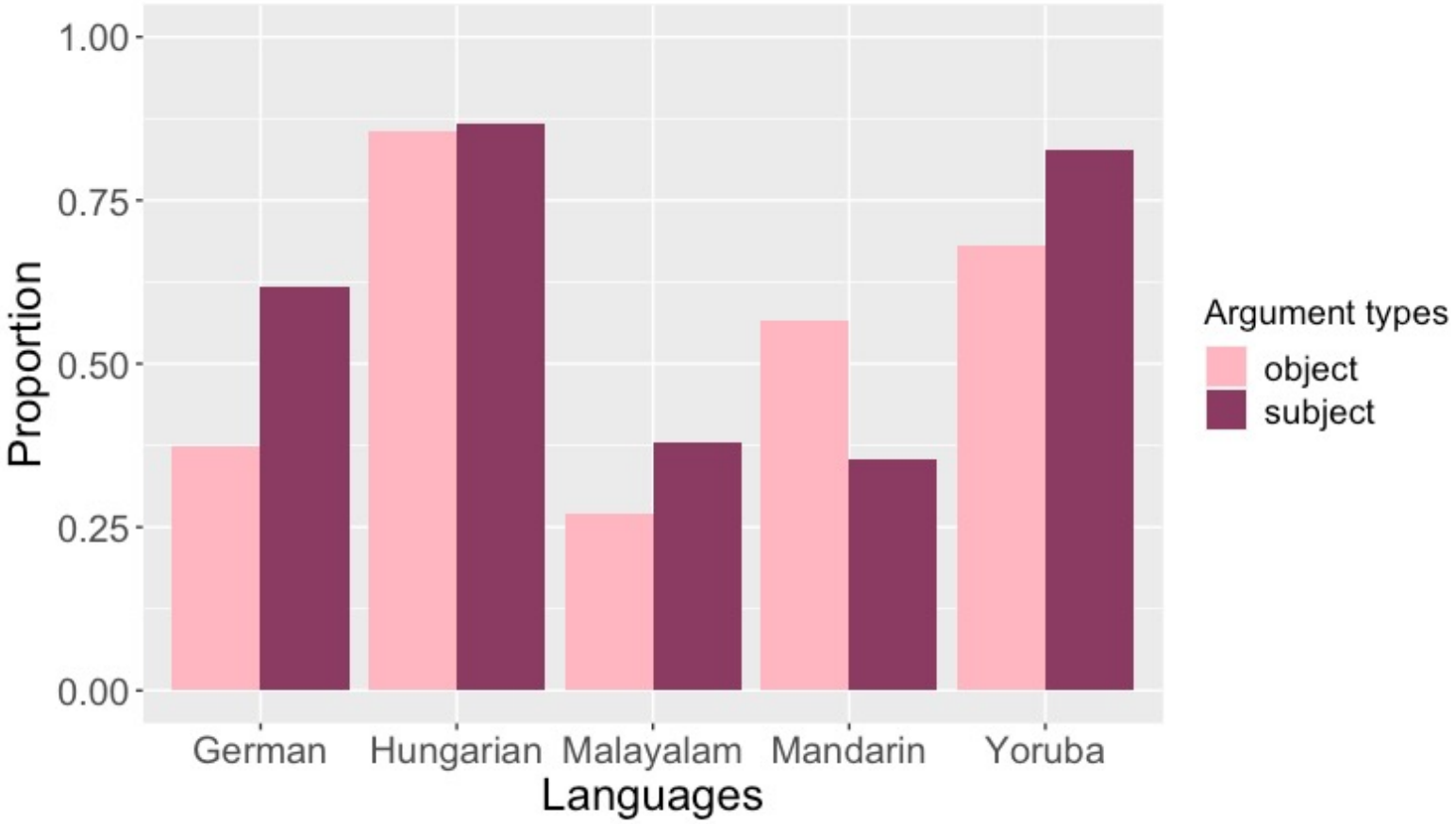


An example: Which monkey is scratching the ghost?



S8

Results: proportion of target responses



Generalized Mixed Effect Models (lmer)

- fixed effect: argument type, language
- random effect: participant
- argument type: t-value: 2.909, $p<.01$
- Effect of languages, except for between Mandarin and German

Some “errors” were not errors: *correct* responses

We defined **Target** as containing a wh-phrase, use of definite NP (e.g. the cat), and use of active voice.

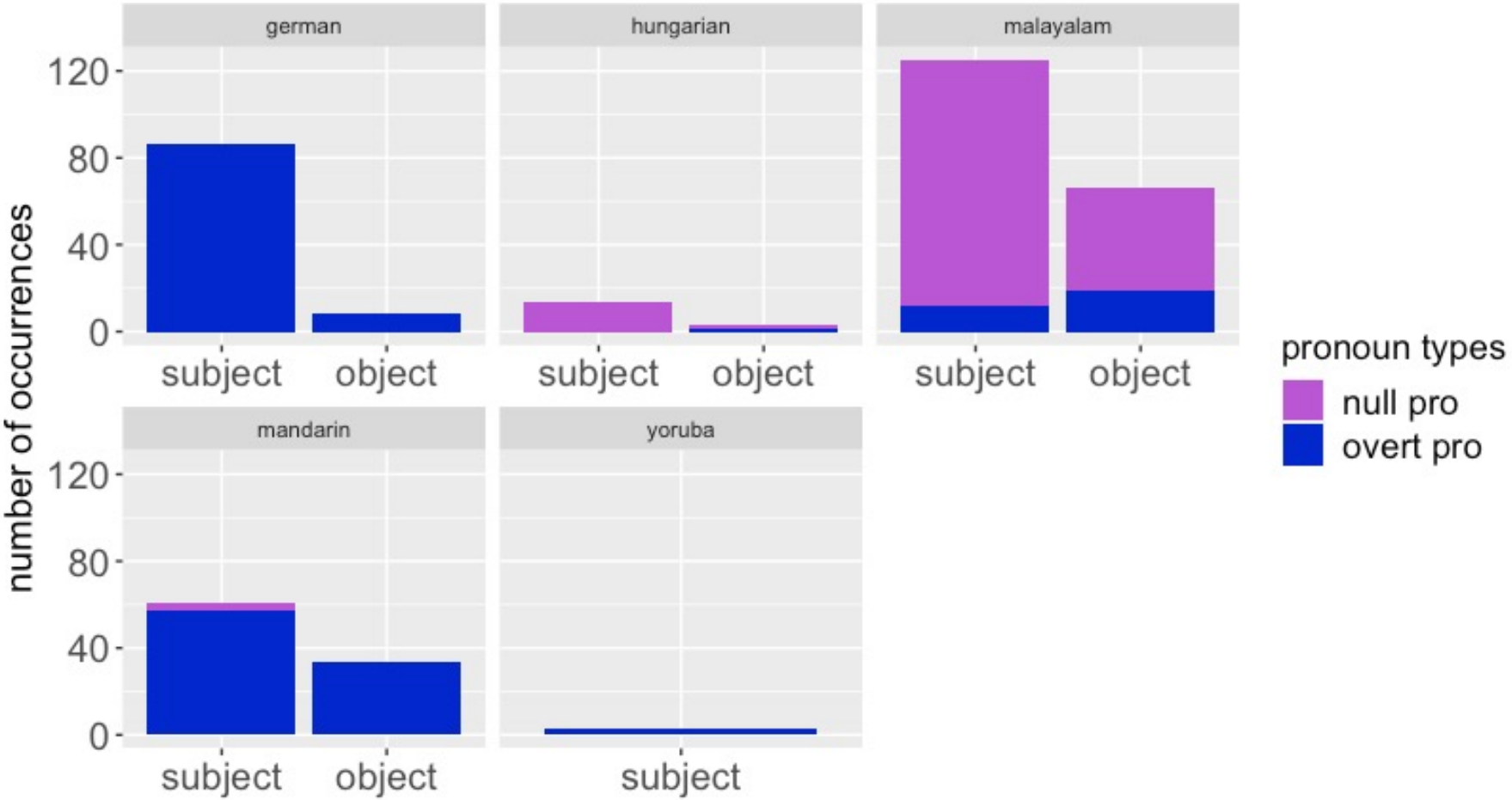
Some non-target structures:

- Use of **overt** pronouns (270 instances)
  - *Wen wecken **die**?* (who.Acc is the cat awaking?) —German
- Use of **covert** pronouns in Hungarian and Malayalam (197 instances)
  - *Ki-t húz-nak?* (who-ACC pull-3PL) — Hungarian
- Passive structures (12 instances):
  - Welcher Frosch wird gekratzt? (which frog is (being) scratched?) — German

Most frequent errors per language

German	overt pronoun as subject/object, passive
Hungarian	null pro as subject
Mandarin	cleft, pronouns as subject/object, “NP V is wh”-structure
Malayalam	null arguments, overt pronoun as subject/object, additional demonstrative
Yoruba	use of additional demonstrative, *what*-questions instead of *who*

Use of pronouns per language



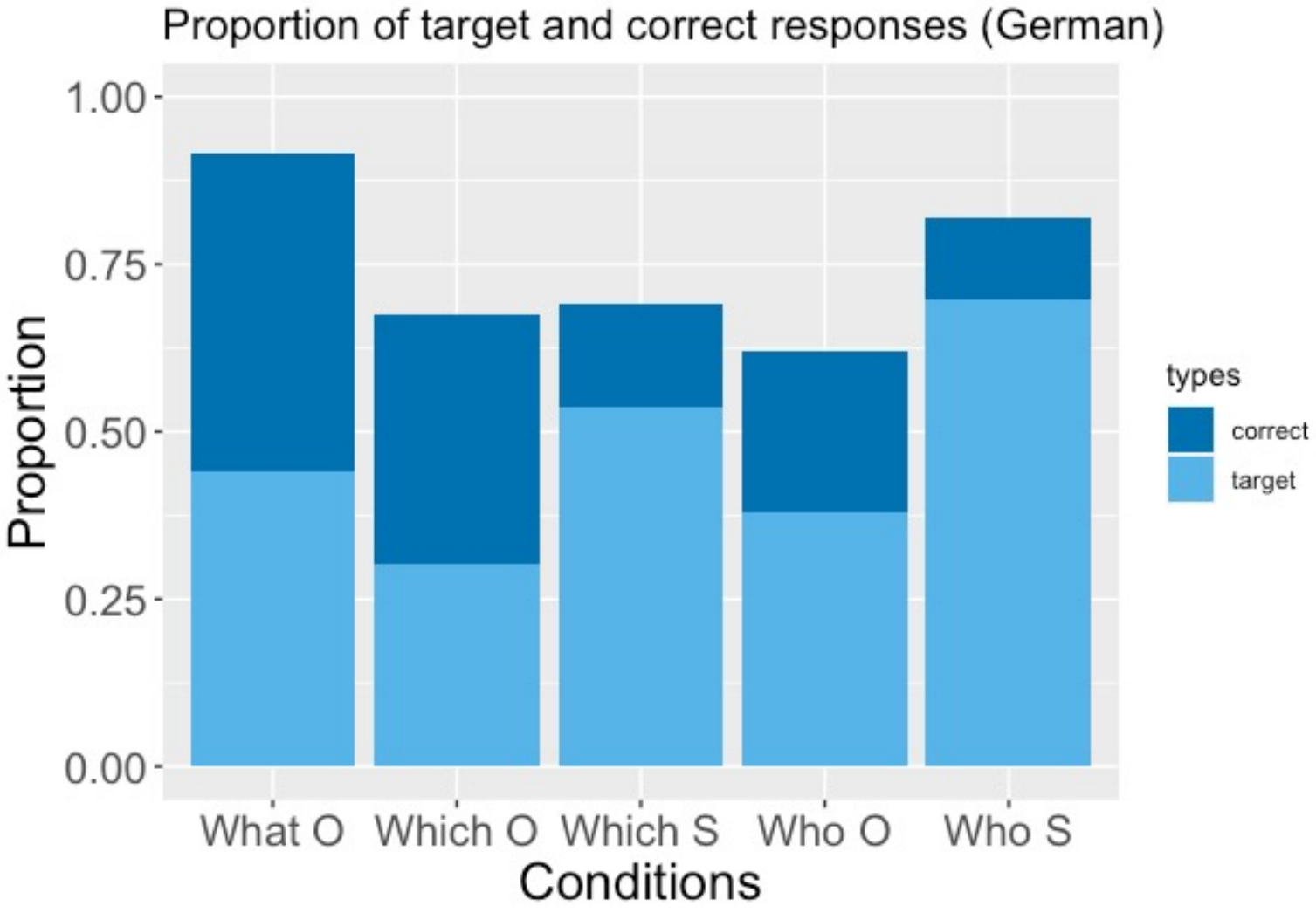
Results and analysis

- Generalized linear mixed models for each language separately, with
  - dependent variable: Correct
  - fixed effect:

    argument type (subject vs. object)

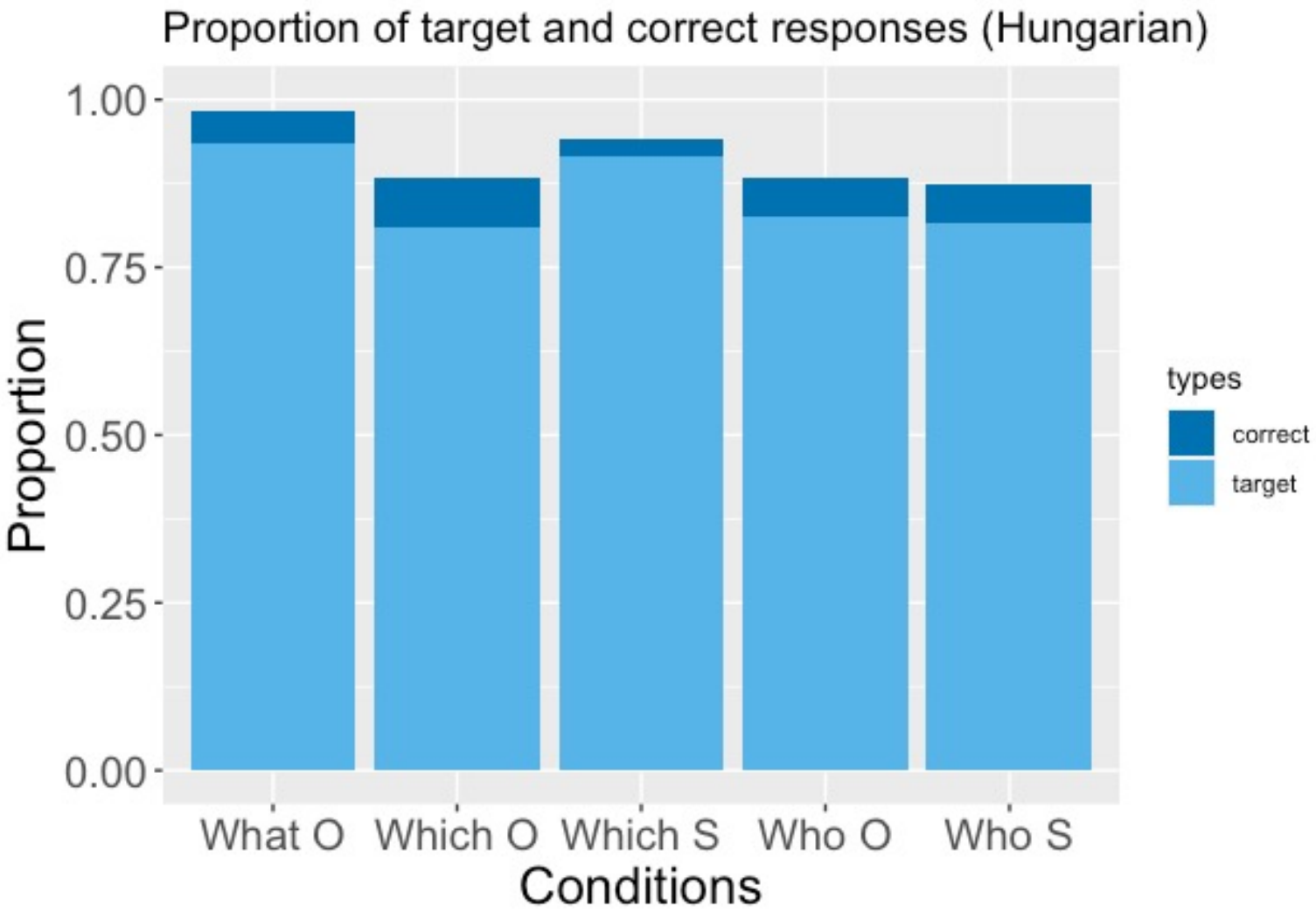
    type of wh-phrase
  - random effect: participant

German



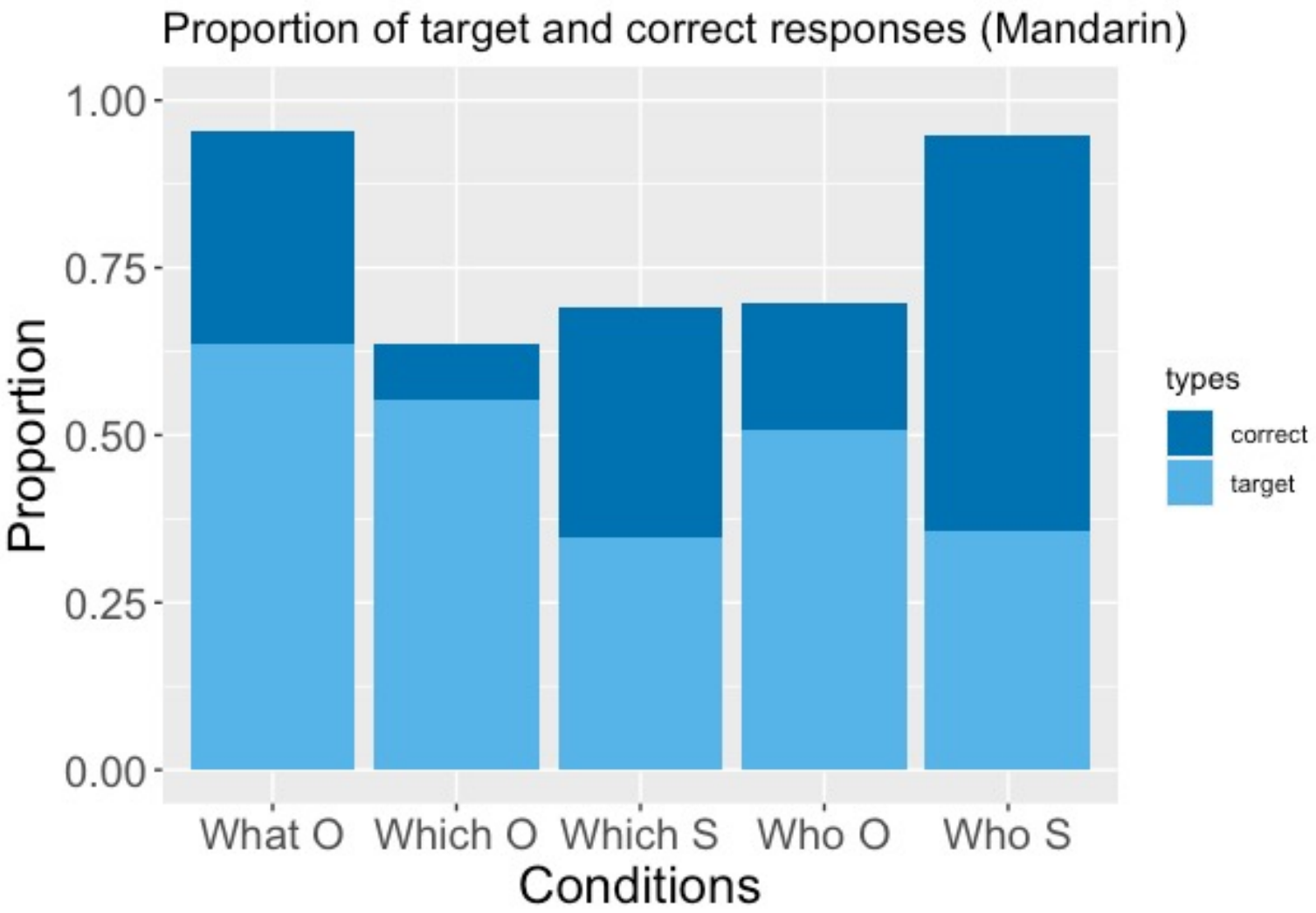
- Effect of argument type t-value=3.056, $p < .01$
- Main effect of type of wh-phrase:
  - significant difference between what-questions and *who*-questions (t-value = -5.445, $p < .01$)
  - significant difference between what-questions and which-questions (t-value = -6.270, $p < .01$)
  - Interaction between argument type and type of wh-phrase
- no main effect but interaction between who and which questions

Hungarian



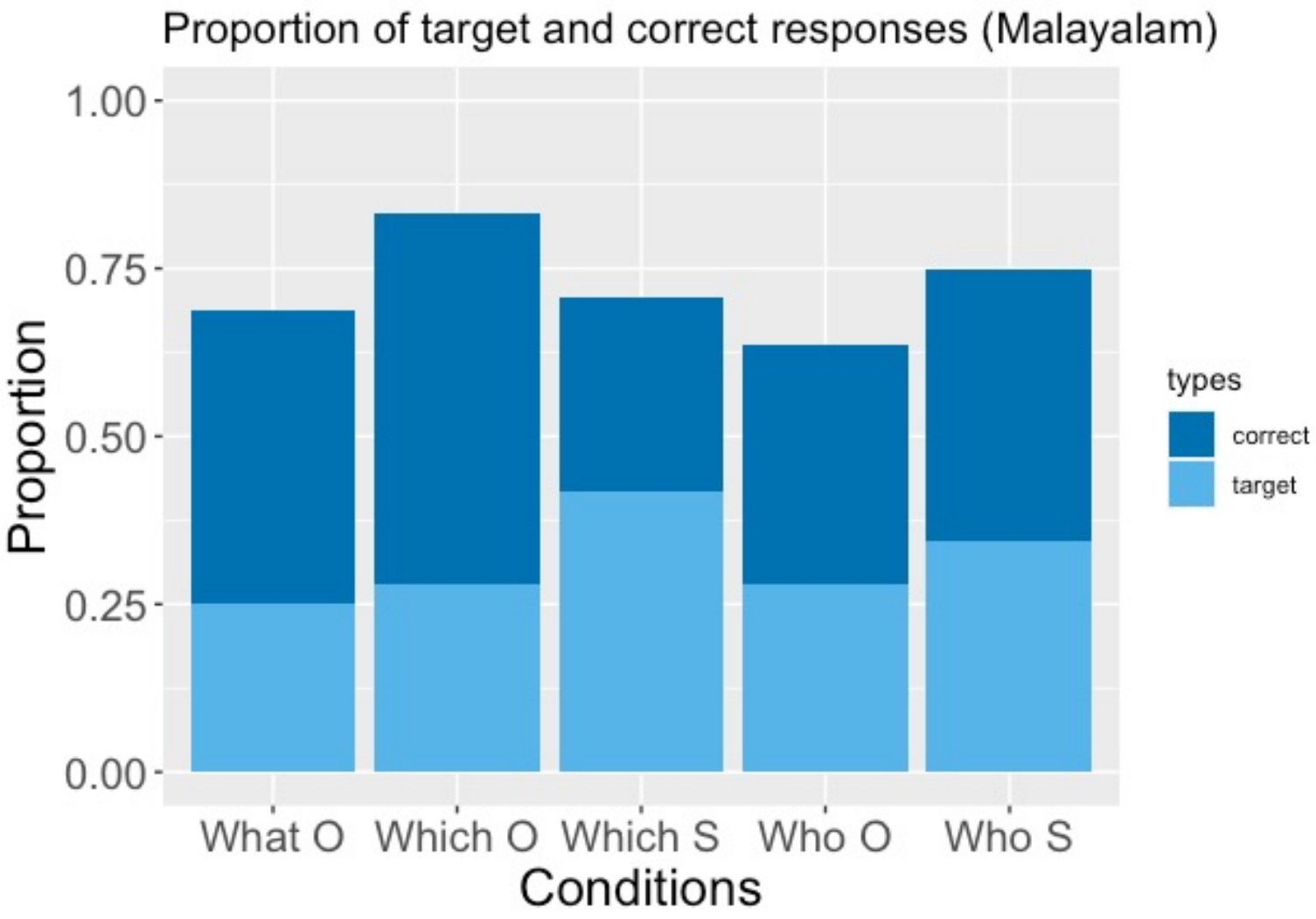
- No effect of argument type t-value=1.311, p = 0.81676
- Main effect of type of wh-phrase:
  - significant difference between what-questions and *who*-questions (t-value = -2.782, p < .01)
  - significant difference between what-questions and which-questions (t-value = -0.2782, p < .01)
- No difference between who questions and which questions

Mandarin



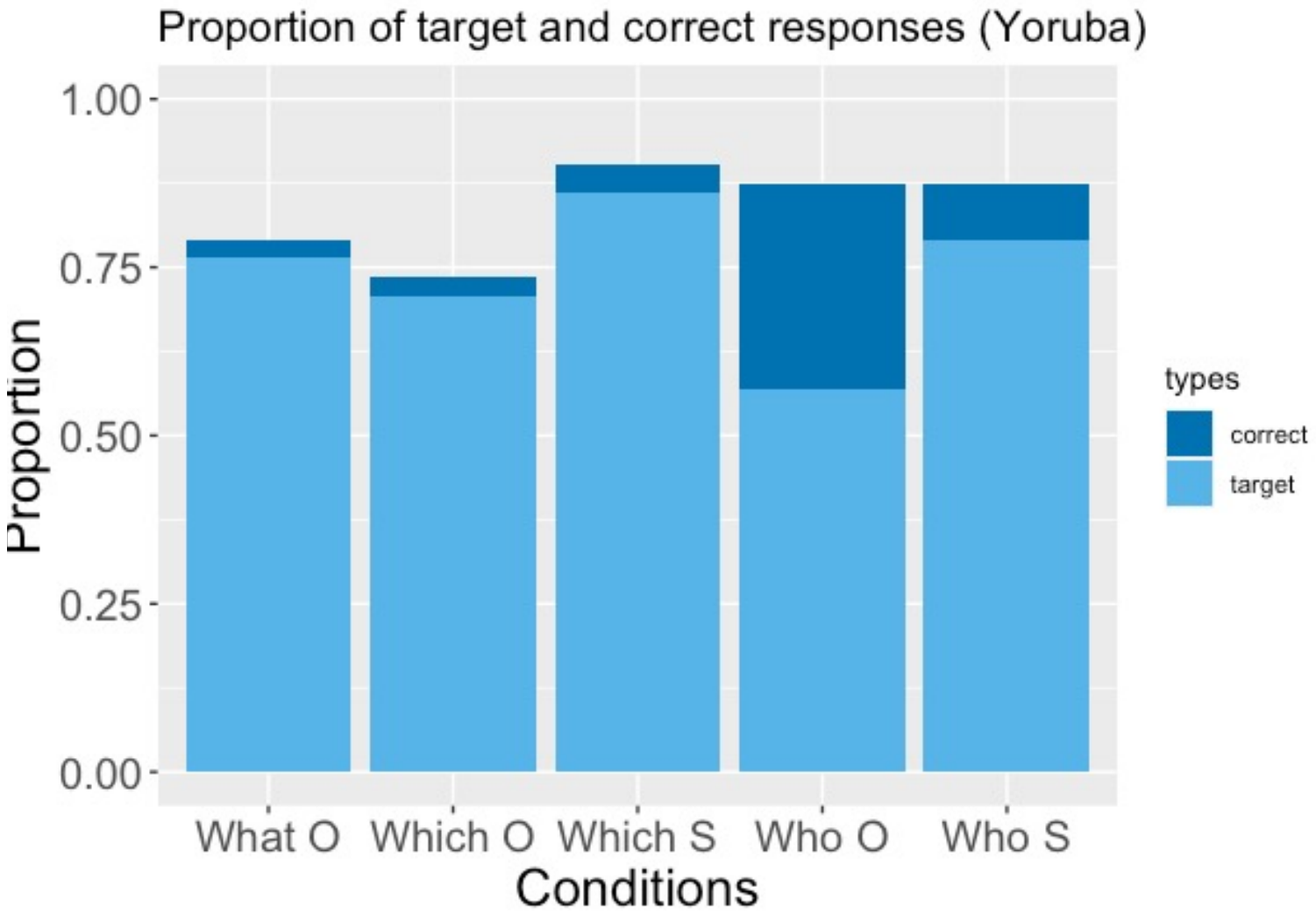
- Effect of argument type t-value=4.709, $p < .01$
- Main effect of type of wh-phrase:
  - significant difference between what-questions and *who*-questions (t-value = -4.906, $p < .01$)
  - significant difference between what-questions and which-questions (t-value = -8.465, $p < .01$)
  - Interaction between argument type and type of wh-phrase
  - main effect and interaction between who and which questions

Malayalam



- Effect of argument type t-value=2.336, $p < .05$
- No significant difference between what-questions and *who*-questions (t-value = 0.232, p= .816346)
- No significant difference between what-questions and which-questions (t-value = 0.883, p= .377672)
- No interaction between argument type and type of wh-phrase
- main effect ($p< .01$) and interaction ($p< .01$) between who and which questions

Yoruba



- Effect of argument type t-value=4.092, p < .01
- Significant difference between what-questions and *who*-questions (t-value = -2.921, p< .01)
- No significant difference between what-questions and which-questions (t-value = -1.203, p= .22988)
- No interaction between argument type and type of wh-phrase
- no difference between who and which questions (p=0.1521) but interaction (p< .05) between who and which questions

Back to other grammatical characteristics:

A new model with the whole dataset, with

- correct responses as the dependent variable
- participant as a random effect
- each characteristic as a fixed effect

1. Obligatory wh-movement language?
   - t-value=2.308, p< .05
2. Case info on wh-phrases?
   - t-value=-0.48, p= .632
3. SOV vs. SVO
   - t-value= -2.304 p< .05
4. Null-argument language?
   - t-value=1.61, p= .111

Conclusion

1. We found overall effect of argument type (subject Q vs. object Q).
2. Within each language, the effect of argument type was found in all languages except for Hungarian
3. Use of pronouns, overt of covert, more frequent in subject position.
4. Who-questions vs. which-questions: no clear advantage of who-questions compared to which-questions with correct responses.
5. What vs. who: what-questions elicited more correct wh-questions than who-and which-questions.
6. Obligatoriness of wh-movement and the verb-final clause structure had an effect on children's production of wh-questions.
7. Educational level of guardians did not have an effect within our data set.

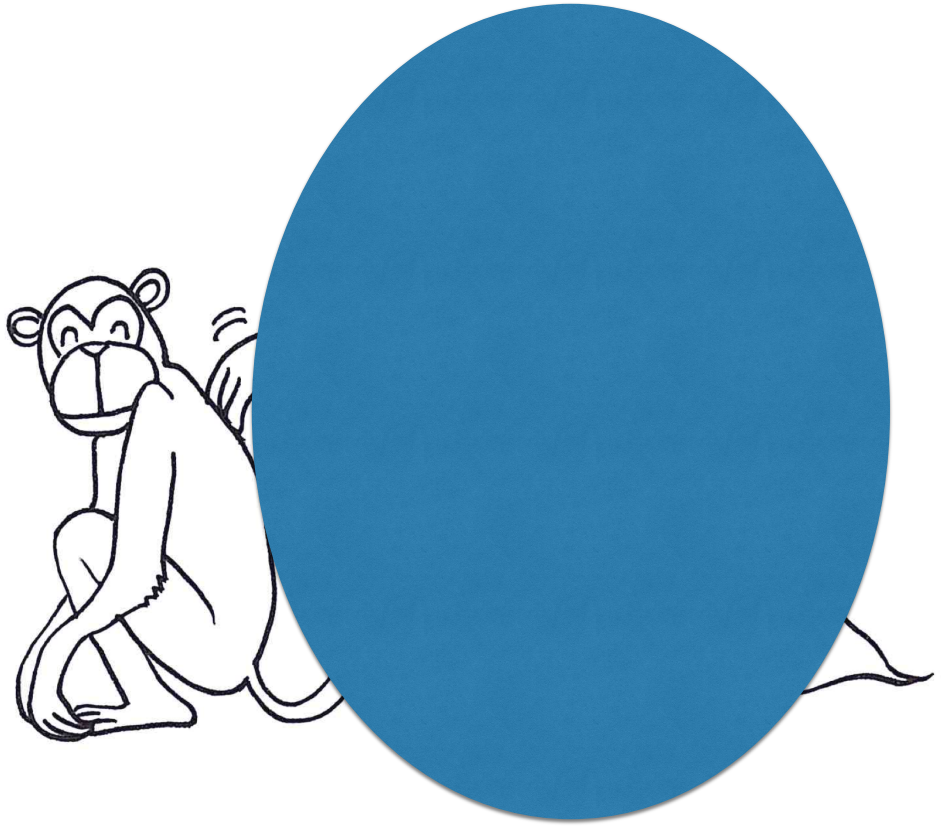
Thank you!

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 856421 and No 787929.

Procedure

1. Experiments were conducted by either one or two experimenters.
   A. 1 experimenter: conduct the experiment while manipulating the puppet
   B. 2 experimenters: one led the experiment, the other manipulated the puppet
2. Picture with a bubble, occluding either the agent or the patient was presented.
3. Lead-in sentence in English:
   “Look, someone is scratching the monkey. Ellie (the puppet) knows who. Ask Elli who.”
4. Target: “Who is pushing the ant?”
5. 6 familiarization items, 25 target items

An example: Who is scratching the monkey?



S2

Steps

1. Designing the experiment (non-local investigators)
2. Adjustment of the material when necessary (local)
3. Data collection (local investigators)
4. Coding of the data (collaboration between local and non-local investigators)
    - Did the utterance have the target question structure?
        - A question was classified as target if:
            - it contains the correct wh-phrase
            - it uses a definite NP for the other argument
            - the verb is in active voice
    - If no, how did the utterance diverge from the target form?
5. Analysis