A comparison of children's relative clause production in Georgian, Italian and Yoruba

Yangyu Sun[1], **Chiara Dal Farra**[1], **Aurore Gonzalez**[1], **Johannes Hein**[2], **Johnson F. Ilori**[3], **Tamar Makharoblidze**[4], **Chiara Saponaro**[1], **Kazuko Yatsushiro**[5], **Uli Sauerland**[5] **and Maria Teresa Guasti**[1]

[1]University of Milano-Bicocca
[2]Humboldt University of Berlin
[3]University of Lagos
[4]Ilia State University
[5]Leibniz-Zentrum Allgemeine Sprachwissenschaft

For info: yangyu.sun@unimib.it; mariateresa.guasti@unimib.it

Introduction

Prior research on the acquisition of **S(ubject)- and O(bject)-Relative Clauses (RCs)** suggests a (quasi-)universal **Subject advantage**. Studies that extended the investigation to Indirect Object (IO) and Genitive-RCs showed that these types of non-S RCs are also challenging for children. This raises the following questions:

(i) Does the RC acquisition follow the **typological universal Accessibility Hierarchy** (see below in "Background")?
(ii) What **strategies do children use** to overcome the complexity of RC formation in different languages?

We address these questions by comparing the production of **S, O, IO, and L(ocative)-RCs by 5- to 7-year-old children speaking Georgian, Italian and Yoruba with adult speakers as control**.

Background

Subject advantage:

- Across languages and populations, subject RCs are **easier to comprehend and produce** than object or other RCs (see Lau and Tanaka 2021, Tanaka et al. 2024 for a review of the state-of-the-art).
- Children often **transform** non-subject RCs into subject RCs (Labelle 1990; McDaniel et al. 1998; Guasti & Cardinaletti 2003).

Typological studies propose that the possibility of relativizing an argument/adjunct follow a universal **Accessibility Hierarchy** (Keenan & Comrie 1977):
SUBJECT > DIRECT OBJECT > INDIRECT OBJECT > OBLIQUE > GENITIVE > OBJECT OF COMPARISON

Languages under investigation

Georgian, Italian, and Yoruba are typologically unrelated languages—belonging respectively to the Kartvelian, Indo-European (Romance), and Niger-Congo families—and differ sharply in morpho-syntactic features, making them ideal for testing whether developmental patterns in RC production reflect universal rather than language-specific mechanisms.

	Georgian	Italian	Yoruba
Relativizer	*that*- and *wh*-relativizers	*that*-relativizer for S-/O-RCs; *wh*-relativizer for IO-/L-RCs	*that*-relativizer
Head position	external, **internal**, correlative	external	external
Case marking	external case of the matrix verb, internal case to *wh*-relative	no case marking	no case marking

(1) *Txam, **romelsac** ori k'at'a miatrevs.*
goat-ERG which-DAT-PROSOD-PART TWO-NOM cat PREV-VER-drag-TH-3SBJSG
'The goat which the two cats are dragging.'
(Georgian O-RC)

(2) *La capra **che** i gatti stanno tirando.*
the goat REL the cats are pulling
'The goat that the cats are pulling.'
(Italian O-RC)

(3) *Ewúrẹ́ kan **tí** olóńgbò méjì ń fà.*
goat one REL cat two PROG pull
'The goat that two cats are pulling.'
(Yoruba O-RC)

Experiment

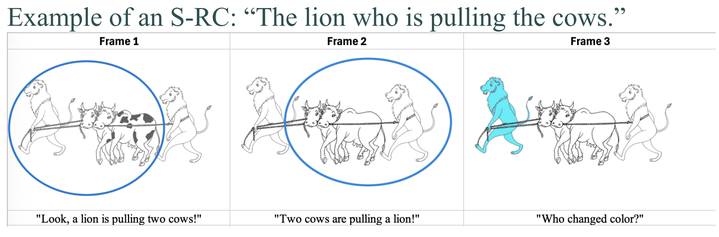
Participants

We collected data from 22 Georgian-speaking children (5;9-7;9) and 25 adults, 25 Italian-speaking children (5;2-6;3) and 28 adults, and 27 Yoruba-speaking children (5;9-7;0) and 28 adults. Participants who produced less than 8 RCs in the experimental phase are excluded.

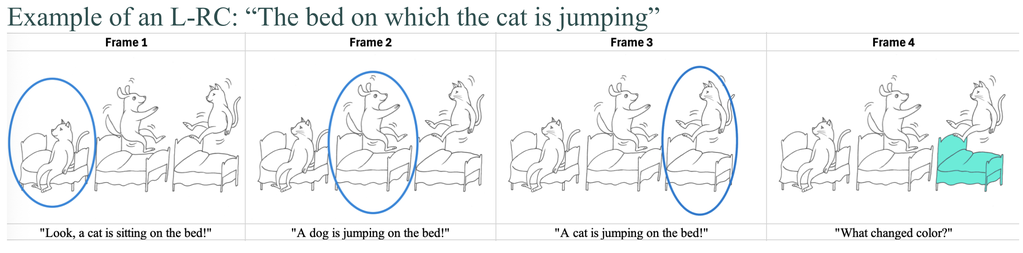
Materials

We used an elicitation task prompting the production of relative clauses. The task consists of a training phase and an experimental phase. During the training phase, participants are instructed on how to perform the task: 4 subject RCs, 3 including intransitive verbs and 1 including a transitive verb, are presented in a fixed order and feedback is provided to children (but not to adults). In the experimental phase, 4 types of RC are elicited: S-, O-, IO-, and L-RCs, with 4 items per type, resulting in a total of 16 items, presented fully randomly. Participants saw images depicting animals/objects involved in actions, with one entity changing colour. A character then asked, "Who/what changed colour?", prompting them to produce an RC (as in Zukowski 2009).

(4) Example of an S-RC: "The lion who is pulling the cows."



(5) Example of an L-RC: "The bed on which the cat is jumping"



Coding

Unconventionally, we did not code "**Correctness**"/ "**Accuracy**", or "**Referential success**", since the "target responses" have different morpho-syntactic features in the three languages, and some informative errors would be covered. Instead, WE CODED the following categories:

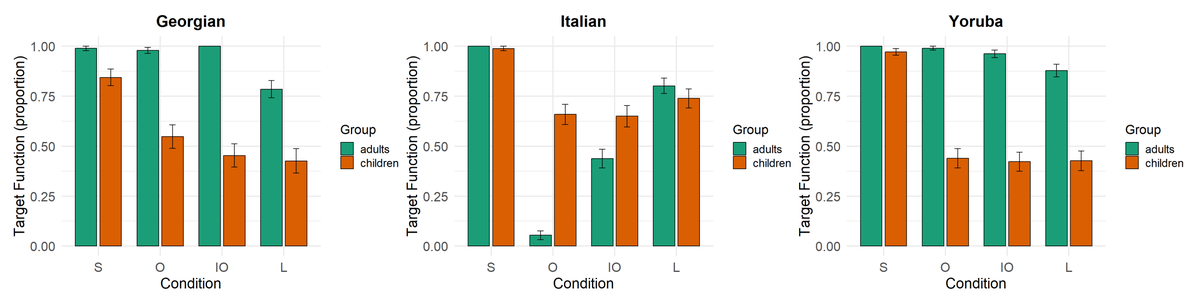
✓ **Target function**: if the RC preserves the targeted S/O/IO/L function of the item or not;
✓ **Function change**: if not the targeted function, what the response appears to be;
✓ **Resumptive**: if the RC contains a noun/pronoun co-referring to the head noun;
✓ **External/internal head**: if the head noun of the RC is outside or inside.

Combining these categories, we introduce the notion of "**circumvention**" as a comparative diagnostic for RC production, capturing all strategies that allow speakers to bypass derivational difficulty, which include function changes, resumptive elements and internal-headed RCs.

Results & Analysis

All analyses used mixed-effects logistic regression models with random intercepts for participants and items, and fixed effect for Group (adults/children) or Language (Georgian/Italian/Yoruba) and Condition (types of RCs). Across all three languages, **a clear group effect** emerged: children and adults differed systematically in how they distributed grammatical functions within RCs.

Target function: Adults vs. Children

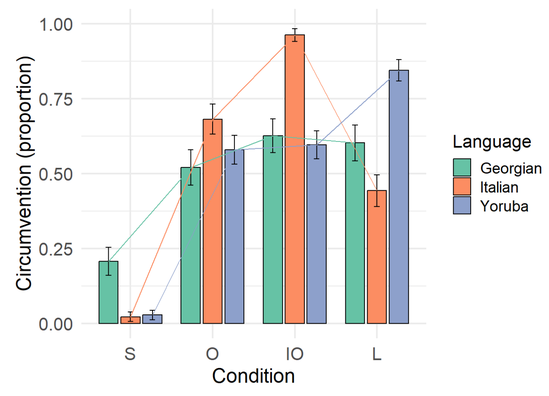


In Georgian and Yoruba, adults outperform children ($\chi^2 = -3.70, p < 0.001$, $\chi^2 = -56.9, p < 0.001$, respectively), whereas in Italian, the effect reverses–adults frequently use passives, reduced RCs and verb change to avoid producing non-S RCs ($\chi^2 = 19.5, p < 0.001$). Focusing on children, in all three languages we found that there is a significant difference between S-RCs and non-S-RCs (pairwise comparisons, all $p < 0.01$), but no difference among the three non-S-RC conditions (pairwise comparisons, all $p > 0.3$). Almost all function changes involve non-S to S in Italian and Yoruba, while in Georgian there are also few changes from non-O to O.

Resumptives: Adults vs. Children

Resumptives were (nearly) absent from Georgian child and adult data. Italian children produced resumptives in non-S RCs, but adults did not. In Yoruba, resumptive pronouns in L-RCs are grammaticalized, but children failed to produce them at the same level as the adults did ($\chi^2 = -25.28, p < 0.001$).

Circumvention: Children across the three languages



Across languages, instances of Circumvention in child data were virtually absent in subject RCs but rose sharply and comparably across all non-S types. Our model with Condition and Language as fixed effects revealed that Condition was statistically significant, suggesting that Circumvention in S was higher than in all the other RC types ($\chi^2 = 182.46, p < 0.001$), while Language did not raise any significant effect ($\chi^2 = 0.049, p = 0.98$), suggesting that children across the three languages exhibited parallel circumvention profiles, supporting a shared developmental tendency to simplify non-S relativization.

Children in all languages relied on function change to avoid producing difficult RCs. In addition, Georgian-speaking children produced internal-headed RCs, and Italian-speaking children produced resumptive elements. Thus, children's circumvention strategies also reflect language-specific features.

Conclusions

- **Adults diverge, children converge.** While adult speakers of Georgian, Italian, and Yoruba exhibit language-specific circumvention patterns — reflecting their grammar's available repair mechanisms (e.g., voice alternations and verb changes) — children show a strikingly uniform profile.
- **Subject advantage.** Children in all languages displayed a strong Subject advantage, producing S-RCs far more faithfully and accurately than O-, IO-, or L-RCs.

Regarding **RQ(i)**, we conclude that the RC acquisition does not clearly follow the typological universal Accessibility Hierarchy, only the subject advantage is confirmed.
Regarding **RQ(ii)**, children mainly used function change, resumptives and internal-headed RCs to overcome the difficulties, but the specific strategies differ across languages, depending on the morpho-syntactic features of the languages.

References

Adani et al. (2010). Grammatical feature dissimilarities make relative clauses easier: A comprehension study with Italian children. Lingua. **Guasti**, M. T., & **Cardinaletti**, A. (2003). Relative clause formation in Romance childs production. **Keenan**, E. L. & **Comrie**, B. (1977). Noun phrase accessibility and universal grammar. In: Linguistic Inquiry 8.1, pp. 63–99. **Lau**, E. and **Tanaka**, N (2021). The subject advantage in relative clauses: A review. In: Glossa 6.1. **Tanaka** et al. (2024). On the universality of the subject preference in the acquisition of relative clauses across languages. First Language. **Zukowski**, A. (2008). Elicited production of relative clauses in children with Williams syndrome. Language and Cognitive Processes, 24(1), 1–43.

Acknowledgments

This presentation is part of a project that has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 856421).