



# Do LLMs Disambiguate Italian Relative Clause Attachment?

Michael Kamerath & Aniello De Santo

Department of Linguistics, University of Utah ✉: {u0773130, aniello.desanto}@utah.edu

## Broad Question

Can we leverage psycholinguistic studies on ambiguous sentences to probe LLMs’ linguistic representations?

## Background

### Relative clause attachment ambiguity

DP1 *of* DP2 [RC] construction

- They saw the daughter of the actress **who was on the balcony**.
  - The daughter was on the balcony. [**HA**: high attachment interpretation]
  - The actress was on the balcony. [**LA**: low attachment interpretation]

### Cross-linguistic variation

(Clifton Jr. & Frazier, 1996; Cuetos & Mitchell, 1988; Gibson et al., 1999; Shen, 2006; a.o.)

- LA preference: Basque, Chinese, **English**, Romanian
- HA preference: Dutch, Korean, **Italian**, Spanish

### Italian Attachment Preferences

- In Italian, main **verb type** affects attachment preferences (Grillo & Costa 2015; Lee & De Santo 2024)
  - Perceptual verbs (observe, smell, hear) lead to HA
  - Nonperceptual verbs (marry, know, cook) lead to LA

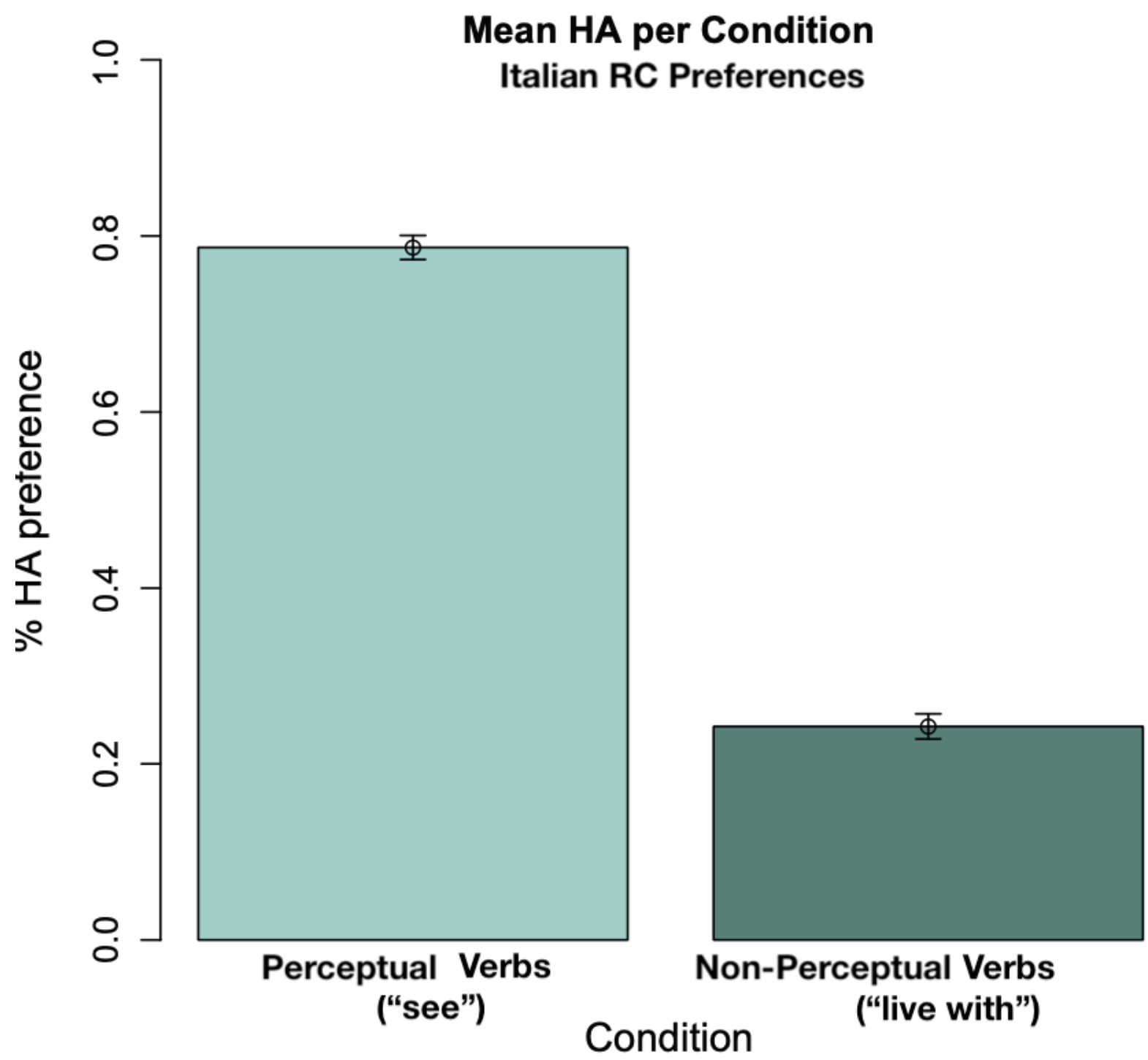


Figure: Adapted from (Grillo and Costa 2015)

### LLMs and Psycholinguistics

- Psycholinguistics tasks/datasets successfully used to probe LLM behavior (Linzen et al. 2016; Futrell et al. 2019)
- Little work on how LLMs handle RC ambiguity (Davis & Van Schijndel 2020; Hénot-Mortier 2023; Issa & Atouf 2024)

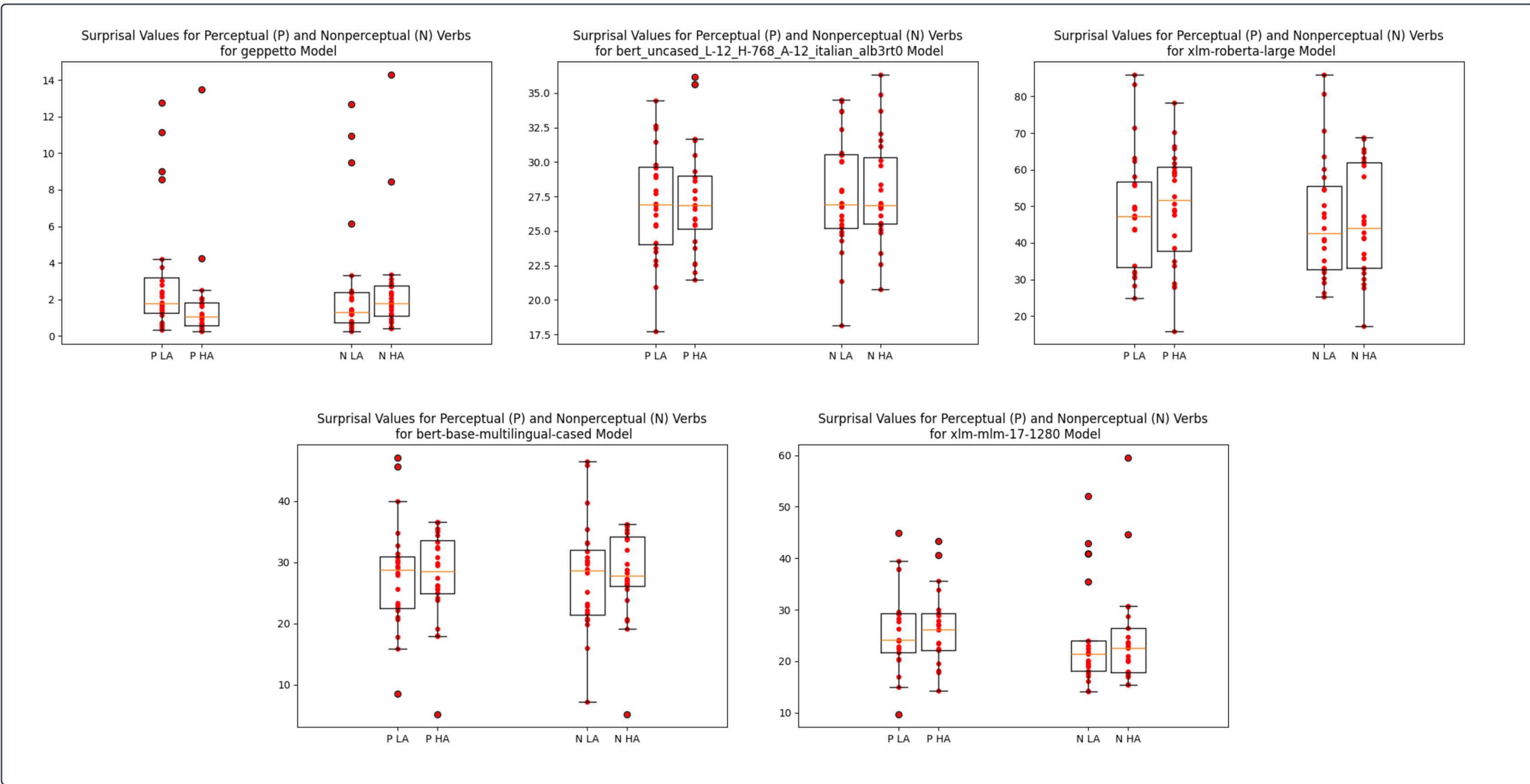
## Here: RC Attachment and LLMs

- Do LLMs tested on Italian show any type of attachment preference?
- Do they line up with Italian speakers?
- Is there a very type effect?

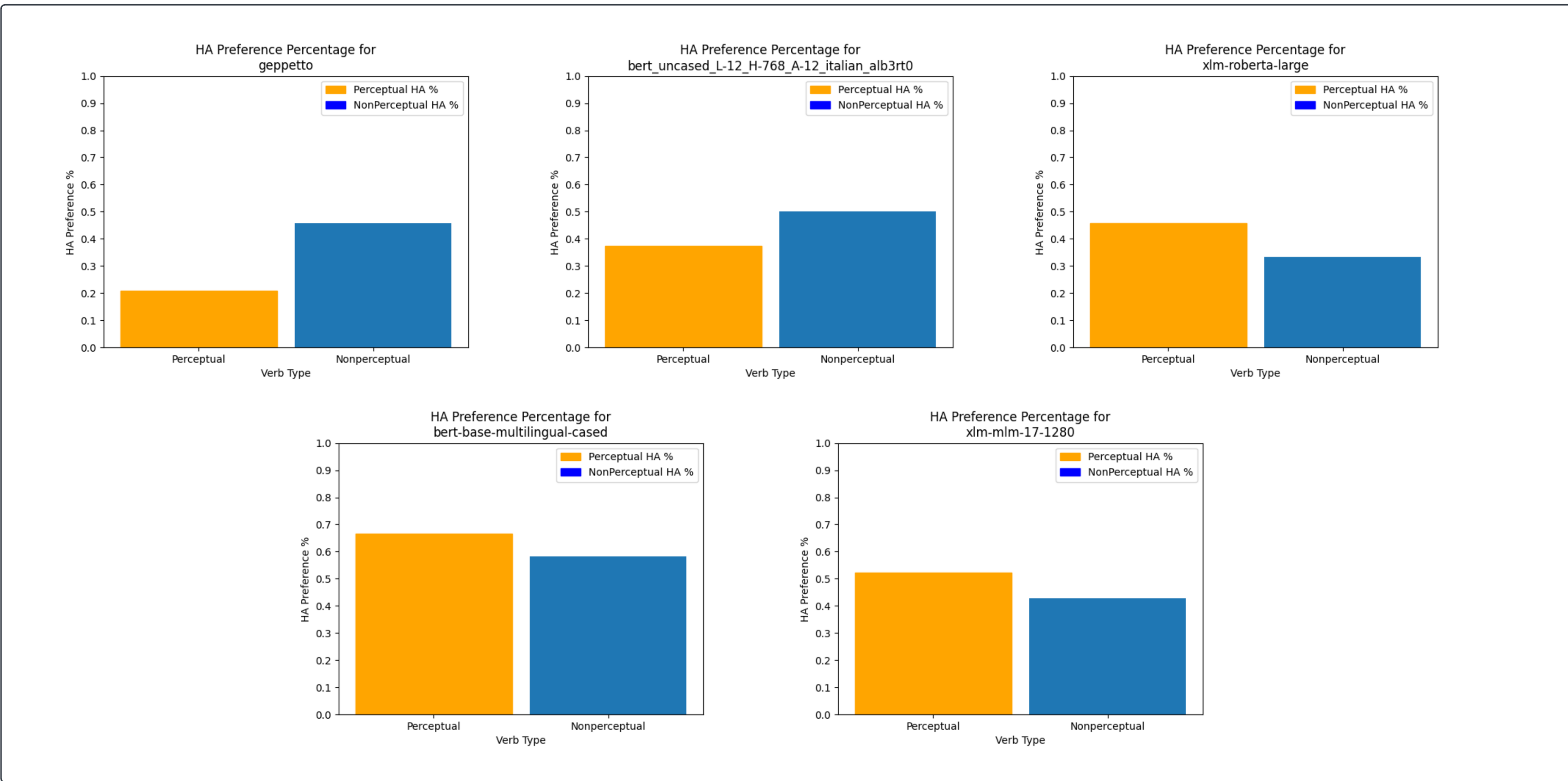
## Sample Stimuli

Verb (PR availability)		Attachment	Target			
a.	Perceptual (PR/RC)	LA	Gianni vide il figlio dei medici	che correvano	la maratona	
			Gianni saw the son-SG of the doctors-PL	who were running-PL	the marathon	
b.	Perceptual (PR/RC)	HA	Gianni vide il figlio dei medici	che correva	la maratona	
			Gianni saw the son-SG of the doctors-PL	who was running-SG	the marathon	
c.	Non-Perceptual (RC only)	LA	Gianni amò il figlio dei medici	che correvano	la maratona	
			Gianni loved the son-SG of the doctors-PL	who were running-PL	the marathon	
d.	Non-Perceptual (RC only)	HA	Gianni amò il figlio dei medici	che correva	la maratona	
			Gianni loved the son-SG of the doctors-PL	who was running-SG	the marathon	

## Results: Surprisal Values by Condition



## Results: Categorical Pairwise Comparisons



## Selected References

Forrest Davis and Marten Van Schijndel. 2020. Recurrent neural network language models always learn english-like relative clause attachment. *arXiv preprint* arXiv:2005.00165.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. Bert: Pre-training of deep bidirectional transformers for language understanding.

Nino Grillo and João Costa. 2015. A novel argument for the universality of parsing principles. *Cognition*, 133(1):156–187.

Adèle Hénot-Mortier. 2023. Do language models discriminate between relatives and pseudorelatives? *In Proceedings of the 2023 CLASP Conference on Learning with Small Data (LSD)*.

Elsayed Issa and Noureddine Atouf. 2024. Context-biased vs. structure-biased disambiguation of relative clauses in large language models. *Procedia Computer Science*, 244:425–431. 6th International Conference on AI in Computational Linguistics.

So Young Lee and Aniello De Santo. 2024. Online evidence for pseudo-relative effects on italian rc attachment resolution. *Language, Cognition and Neuroscience*, 39(9):1212–1229.

Marco Polignano, Pierpaolo Basile, Marco De Gemmis, Giovanni Semeraro, Valerio Basile, et al. 2019. Alberto: Italian bert language understanding model for nlp challenging tasks based on tweets. In CEUR workshop proceedings, volume 2481, pages 1–6. *CEUR*.

## Experiments

### Stimuli

- Temporarily ambiguous sentences (Lee & De Santo 2024)
- LA vs HA disambiguated by number agreement between nouns and the RC verb (*correva* vs *correvano*, counterbalanced across items)

### Procedure

- Tested two Italian-only models, and three multilingual models
- Dependent Measure: **surprisal** at the disambiguating verb
- Qualitatively: Compare surprisals pairwise (by verb type) to assign LA or HA categorically item-wise.

Model
GePpeTto (De Mattei et al. 2020)
Alberto (Polignano et al. 2019)
bert-base-multilingual-cased (Devlin et al 2019)
xlm-m1m-17-1280 (Conneau et al. 2019)
xlm-roberta-large (Conneau et al. 2020)

Table: Tested Models

## Attachment as Pairwise Comparisons

*Attachment Preference*  $\leftarrow$  *LOW* if *Verb Surprisal(a)* > *Verb Surprisal(b)*  
*Attachment Preference*  $\leftarrow$  *HIGH* if *Verb Surprisal(a)* < *Verb Surprisal(b)*  
*Attachment Preference*  $\leftarrow$  *LOW* if *Verb Surprisal(c)* > *Verb Surprisal(d)*  
*Attachment Preference*  $\leftarrow$  *HIGH* if *Verb Surprisal(c)* < *Verb Surprisal(d)*

## Results

- Fit a linear mixed effects model on raw surprisal:  
 $\text{Surprisal} \sim \text{Verb Type} + \text{Attachment Type} + \text{Verb Type} * \text{Attachment Type} + (1|\text{set})$
- No significant attachment or verb type effects found, nor their interaction
- LLMs lacked verb type sensitivity and attachment preferences
- Results consistent across both Italian and multilingual LLMs
- BUT: some trends on the qualitative contrasts  $\Rightarrow$  further investigations of differences between measures

## Conclusion

**TL;DR Ambiguity as a privileged ground for cross-linguistic evaluation!**

- Results contrast previous work showing LA for LLMs in English/Spanish (but different architectures and/or methods!)
- Next: extensive analysis of techniques (surprisal vs prompting vs categorical assignment), mechanisms (causal analysis; Want et al. 2023), and architectures
- Spoiler alert: Ask us about English!**

