Do LLMs Disambiguate Italian Relative Clause Attachment?

Michael Kamerath and **Aniello De Santo**
University of Utah
{michael.kamerath,aniello.desanto}@utah.edu

1 Introduction

An important component of human sentence comprehension is how to judge multiple simultaneously correct interpretations for a single input. For instance, consider the case (as in 1) of a relative clause (RC) *that was running* following a complex noun phrase *son of the doctor*:

(1) I saw the son of the doctor that was running.

There are two possible interpretations of this sentence based on whether the RC *that was running* modifies *the son* or *the doctor*. The interpretation in which the RC modifies *the doctor* is referred to as low attachment (LA) while the case of the RC modifying *the son* is referred to as high attachment (HA).

RC attachment preferences thus present an interesting way of probing LLMs' syntactic knowledge, given that speakers' preferences for HA or LA vary cross-linguistically, and has been reported to be affected by a variety of syntactic and semantic factors (Grillo and Costa, 2014). However, RC attachment preferences seem to be understudied in the LLM syntactic evaluation literature (Davis and Van Schijndel, 2020; Issa and Atouf, 2024). Here, we aim to add to this scarce literature, and evaluate a variety of LLMs trained to determine their disambiguation strategies over relative clauses in Italian.

2 Italian RC Attachment

Modulo other variables, English speakers generally exhibit a low attachment RC preference with Italian speakers preferring a high attachment interpretation. Recently, it has been argued that one important predictor of attachment preference in Italian RCs is whether the verb in the main clause is non-perceptual (*marry, know, cook, etc*) or perceptual (*observe, hear, smell, etc*). When other semantic and syntactic aspects are controlled for, RCs of sentences containing non-perceptual verbs lead to a LA preference while perceptual verbs lead to a HA preference (Grillo and Costa, 2014; Lee and De Santo, 2024). Focusing on French, Hénot-Mortier (2023) has shown that monolingual and multilingual transformer architectures exhibit some sensitivity to non-perceptual/perceptual verb type modulations in non-RC contexts. Building on the psycholinguistics literature and these past LLM results on RC attachment and verb type effects, here we ask:

Sentence	Verb Type	Attachment
a	perceptual (P)	HA
b	perceptual (P)	LA
c	non-perceptual (N)	HA
d	non-perceptual (N)	LA

Table 1: Summary of Italian Stimuli by Group

1. whether LLMs tested on Italian show any type of attachment preference;
2. whether these preferences conform to those of Italian speakers;
3. whether these preferences show sensitivity to verb type.

3 Experiment and Results

We build on Grillo and Costa (2014), testing Italian sentences with a structure as in (1) but with a matrix verb manipulation, which is either perceptual or non-perceptual in a 2×2 design summarized in Table 1. As a proxy for a model's attachment preferences, we measure the surprisal value at the embedded verb (Davis and Van Schijndel, 2020). Therefore, we depart from Grillo and Costa (2014) in being unable to use fully ambiguous relative clauses. Instead, we leverage the stimuli used in Lee and De Santo (2024), testing our models on sentences that are disambiguated for HA or LA by gender agreement between one of the two nouns in the complex DP (*son* or *doctors*) and the embedded verb. We can then think of contrasting sentence types pairwise, which leads to the following LA/HA predictions:

Proceedings of the Society for Computation in Linguistics (SCiL) 2025, pages 359-360.
Eugene, Oregon, July 18-20, 2025

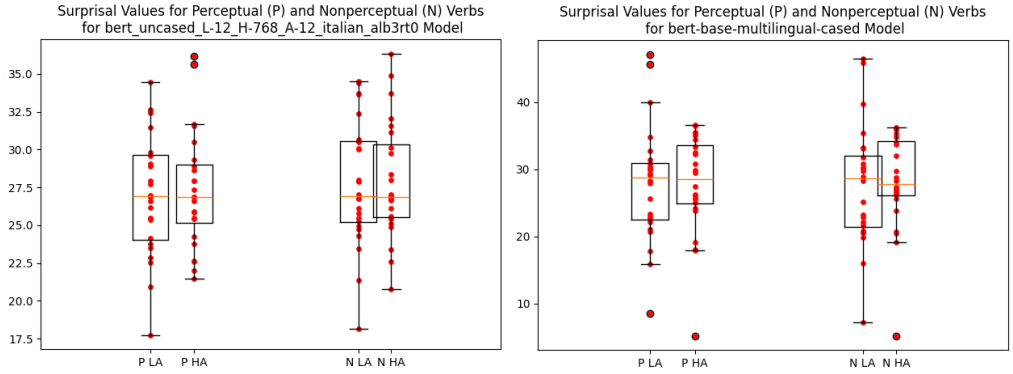

Figure 1: Surprisal Values by Attachment Type and Verb Type for one of the Italian-only (GePpeTto; Polignano et al., 2019) and one of the multilingual (bert-base-079 multilingual-cased; Devlin et al., 2019) models tested.

`Attachment Preference` ← `LOW` if Verb Surprisal(a) > Verb Surprisal(b)
`Attachment Preference` ← `HIGH` if Verb Surprisal(a) < Verb Surprisal(b)
`Attachment Preference` ← `LOW` if Verb Surprisal(c) > Verb Surprisal(d)
`Attachment Preference` ← `HIGH` if Verb Surprisal(c) < Verb Surprisal(d)

We evaluated two Italian-only models, and three multilingual models (GePpeTto; AlBerto; bert-base-multilingual-cased; bert-base-multilingual-cased; xlm-roberta-large) in line with those tested for French by Hénot-Mortier (2023). For each LLM, pairwise contrasts do not reveal a strong tendency towards either LA or HA. We then fit linear mixed-effect models using Surprisal at the embedded verb as the dependent variable, and Verb Type and Attachment Type as fixed effects. Our analyses show no significant attachment or verb type effects, again consistent with the absence of attachment preferences (in line with Italian speakers or not) in each of the models, (see Figure 1 for results for two of the models).

4 Discussion and Further Work

In this work, we measured the difference in surprisal of locally ambiguous sentences at the point of disambiguation to determine whether a variety of LLMs learn human-like attachment preferences in Italian. Our results indicate that none of the models we tested exhibits any attachment preference at all, somewhat in contrast to previous results for Spanish and Arabic (Davis and Van Schijndel, 2020; Issa and Atouf, 2024). However, Davis and Van Schijndel (2020) tested models with an LSTM architecture, while Issa and Atouf (2024) used prompting methods as opposed to the surprisal measurements used here. Future work should explore these differences more in depth, and suggest a primary role for RC disambiguation in the study of LLMs' capabilities cross-linguistically.

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. 2014. 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)*, pages 55–61.

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.