

Processing Effects of Code-Switching in Humans and LLMs

Marina Sokolova¹, Natalia Moskvina¹, Nayara Mirio e Silva²

¹Autonomous University of Barcelona; ²University of Barcelona

This study investigates the processing effects of code-switching (CS) between typologically distant languages in humans and large language models (LLMs). A central question is whether CS induces a prosodic boundary and how its placement affects syntactic parsing. While this remains an empirical question in human sentence processing, it provides a strong test case for human-machine comparison: we hypothesize that the position of CS should affect parsing in both humans and LLMs, although potentially for different reasons. In humans, CS may induce a prosodic break that guides attachment preferences, whereas in LLMs the language switch itself may function as a structural cue marking constituent boundaries. We therefore examine whether the position of CS affects the resolution of relative clause (RC) attachment ambiguities, exemplified in (1), as predicted by the Implicit Prosody Hypothesis (IPH), outlined in (2).

Under the IPH (Fodor 2002), a prosodic break before the relative clause favors high attachment (HA) (2a), whereas a break before the preposition favors low attachment (LA) (2b). Previous auditory studies have shown that manipulating prosodic boundaries affects RC resolution (Goad et al., 2020; Fromont et al., 2017); here, we extend this work to silent reading and bilingual contexts involving CS. In the human experiment, a 1×2 design manipulated the position of CS favoring HA vs. LA attachment (3 a/b), while maintaining CS in the L1→L2 direction across language pairs and establishing baseline attachment preferences through unilingual controls (3 c/d). Adult L2 speakers of Norwegian-English, Russian-Norwegian, Portuguese-English, Spanish-English, and Catalan-English participated in a self-paced reading experiment in which sentences were presented one word at a time; participants advanced through the sentence by pressing a button and answered a comprehension question after each trial.

Parallel experiments were conducted with GPT-4.1 and DeepSeek-V4-Flash, which represent distinct transformer-based LLM families with different architectural and training paradigms. The dataset was evaluated over 40 runs, allowing direct comparison between human parsing and machine processing under identical ambiguity conditions. Unlike human participants, the LLMs were tested in a 2×2 design manipulating both the position and direction of code-switching, using English→Spanish and Spanish→English CS alongside unilingual baselines.

Human results remain stable across groups: participants generally favor low attachment (LA) in CS sentences, but the position of CS produces a modest shift toward the attachment type cued by its placement, with stronger effects in speakers whose languages differ in baseline attachment preferences (e.g., English-Spanish, Norwegian-Russian).

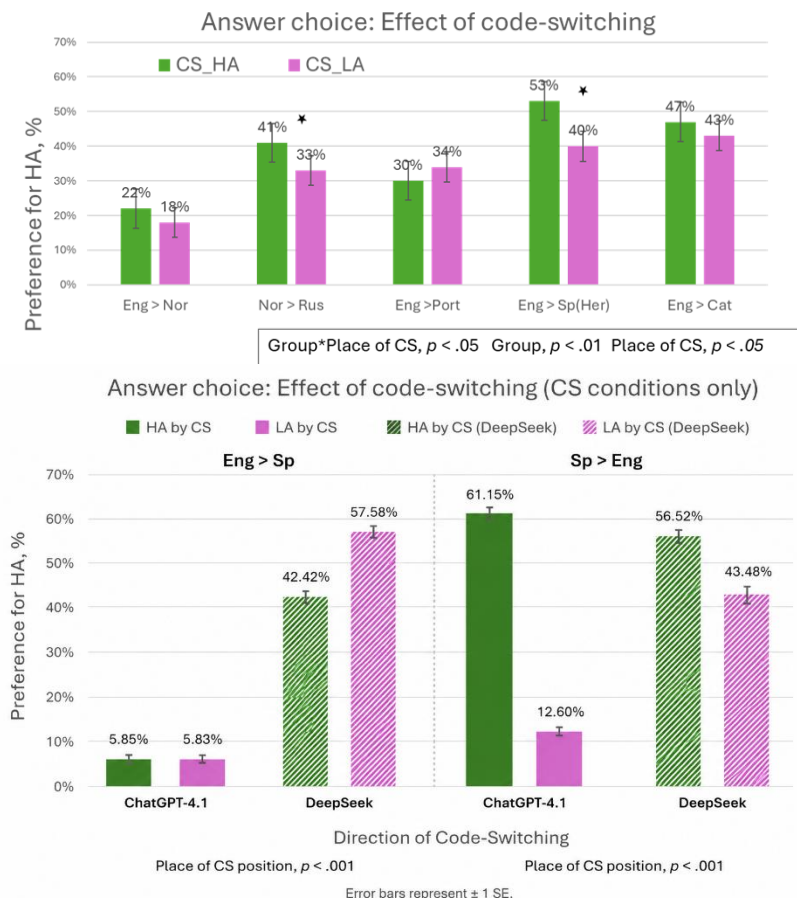
The LLM results reveal both convergences and divergences across models. GPT-4.1 shows an effect of CS position only in Spanish→English switching, while maintaining a strong overall LA preference in English→Spanish CS and in unilingual English and Spanish. DeepSeek, by contrast, shows sensitivity to the position of CS in both switching directions. In Spanish→English CS, its behavior parallels GPT-4.1 and follows the predictions of the IPH, whereas in English→Spanish CS it shows a robust effect in the opposite direction. The two models also differ in their unilingual baselines: GPT-4.1 shows near-categorical LA in both English and Spanish, whereas DeepSeek reproduces the human-like LA preference in English but shows no stable attachment preference in Spanish. DeepSeek additionally demonstrates increased instability in code-switched sentences, with glitches rising from 1.48% in unilingual conditions to 6.03% under CS, suggesting that language switches constitute a salient structural disruption for the model.

Overall, the findings suggest that humans use CS as a probabilistic prosodic cue, whereas LLMs treat language switches as stronger and model-dependent structural signals, producing asymmetric parsing behavior across switching directions and architectures.

Appendix:

- (1) Bill saw [DP the friend of the neighbor [RC that was drinking coffee]]
- [DP [DP the friend of the neighbor] [RC that was drinking coffee]]
(HA, the friend was drinking coffee)
 - [DP the friend of [DP the neighbor [RC that was drinking coffee]]]
(LA, the neighbor was drinking coffee)
- (2) Bill saw [DP the friend of the neighbor [RC that was drinking coffee]]
- [DP [DP the friend of the neighbor] pause [RC that was drinking coffee]] (HA)
 - [DP the friend pause of [DP the neighbor [RC that was drinking coffee]]] (LA)
- (3) *italics marks a language change*
- Bill arrested the friend of the neighbor *that was drinking coffee* (HA by CS)
 - Bill arrested the friend *of the neighbor that was drinking coffee* (LA by CS)
- c. Eng: Bill arrested the friend of the neighbor that was drinking coffee (LA by Language)
- d. Rus: Ivan arestoval druga soseda kotoryj pil kofe (HA by Language)
- Nor: Bill arresterte venner til naboen som drakk kaffe (LA by Language)
- Port: O Guilherme prendeu o amigo do vizinho que estava a beber café (LA by Language)

(4) Results



- RC attachment, humans: Language and place of CS, reference category HA; * $p < .05$; $^{\circ}$ $p < .1$
- There is no effect of the place of CS on response time

References:

- Goad, Heather, Guzzo, Natalia, and Lydia White. 2020. Parsing ambiguous relative clause in L2 English. *Learner sensitivity to prosodic cues. Studies in Second Language Acquisition* 43(1): 83-108.
- Fodor, Janet D. 2002. Prosodic disambiguation in silent reading. In *Proceedings of the North East Linguistics Society*. Edited by M. Hirotsani: 32, vol. 1., pp. 113-132.
- Fromont, Laurent A, Soto-Faraco, Salvador, and Emmanuel Biau. 2017. Searching high and low: Prosodic break disambiguate relative clauses. *Frontiers in Psychology* 8: article 96 [doi:10.3389/fpsyg.2017.00096](https://doi.org/10.3389/fpsyg.2017.00096)