Exploring AI-Resistance to Degendering Strategies in Italian and Spanish

Antonella BOVE, Ca' Foscari University of Venice
Federica MARENGHI, Ca' Foscari University of Venice

This contribution addresses the area *Digital Platforms and Technology – AI language models*, examining how generative AI systems respond to gender-respectful language use and where resistance to such forms emerges.

In Italian and Spanish, gender-fair strategies range from binary doublets (*los/as niños/as*; *i/le bambini/e*) to more recent nonbinary neomorphemes (-ə in Italian; -e in Spanish) and typographic signs (-@, -*, -x) [1][2]. These strategies aim either to explicitly include gender-expansive individuals or to degenderize the language altogether. Nevertheless, experimental studies [3] show that role nouns remain culturally biased, as measured by Misersky and Carreiras scores [4][5], revealing the persistence of stereotypical associations even under degendering strategies.

Generative AI models enable new textual mediation but risk amplifying stereotypes by resisting inclusive practices [6]. This study undertakes an applied analysis of two generative models – ChatGPT-4o and ChatGPT-5 – focusing on their responses to standard (STD) and nonstandard (Non-STD) degendering strategies in Italian and Spanish. Twelve prompts were designed around professional role nouns classified according to their Misersky (Italian) and Carreiras (Spanish) scores, reflecting stereotypically male, female, or neutral associations. Each prompt was encoded with common-gender (CG) nouns in either STD or Non-STD forms, and model outputs were compared across languages.

To minimize stochastic variation, the *temperature* parameter was controlled for ChatGPT-4o, enabling systematic observation of patterns. The study investigates two dimensions of resistance: (i) whether generative outputs replicate or challenge cultural expectations associated with professional role nouns, and (ii) whether models resist degendering strategies themselves, for example, by renormalizing neutral prompts into binary or stereotypical formulations.

Through a 2×2×2 factorial design (Italian vs. Spanish; standard vs. non-standard strategies; ChatGPT-4o vs. ChatGPT-5), the study aims to identify potential sites of resistance and backlash, providing a cross-linguistic and cross-model perspective on how generative systems may interact with gender-inclusive language.

References

[1] Giusti, G. (2022) "Inclusività nella lingua italiana: come e perché. Fondamenti teorici e proposte operative", *DEP. DEPORTATE, ESULI, PROFUGHE 48*:1-19.
[2] Escandell-Vidal, M.V. (2020) "En torno al género inclusivo", *IgualdadES 2*: 223-249. https://doi.org/10.18042/cepc/IgdES.2.08
[3] Ducoli, M., Giusti, G., Lebani, G. (2024) "Districare gli stereotipi dal genere semantico in italiano. Un'indagine psicolinguistica", In De Cesare and Giusti (eds) *Lingua inclusiva: forme, funzioni, atteggiamenti e percezioni*, LIVVAL. LINGUAGGIO E VARIAZIONE, ECF: 155-178.

[4] Misersky, J. et al. (2014) “Norms on the gender perception of role nouns in Czech, English, French, German, Italian, Norwegian, and Slovak”, *Behavior research methods 46*: 841-871.
[5] Carreiras, M., Garnham, A., Oakhill, J., & Cain, K. (1996) “The use of stereotypical gender information in constructing a mental model: Evidence from English and Spanish”, *Quarterly Journal of Experimental Psychology 49A*: 639-663.
[6] Wan, Y., Pu, G., Sun, J., Garimella, A., Chang, K.-W., & Peng, N. (2023) “«Kelly is a Warm Person, Joseph is a Role Model»: Gender Biases in LLM-Generated Reference Letters” (No. arXiv:2310.09219). arXiv. https://doi.org/10.48550/arXiv.2310.09219

SHORT BIO

Antonella Bove holds a BA in Linguistic and Cultural Mediation with a focus on tourism and business translation from the University for Foreigners of Siena, and an MA in Interpreting and Translation for Publishing and for Special Purposes from Ca’ Foscari University of Venice. She spent two semesters abroad studying specialized Translation at the University of Granada and Computational Linguistics and Machine Translation at the Complutense University of Madrid. She is a PhD student in Modern Languages, Cultures and Societies and Linguistics at Ca’ Foscari. Her research investigates the optimal integration of AI into specialized translation processes in Italian and Spanish.

Federica Marenghi is a PhD student in Modern Languages, Cultures and Societies and Linguistics at Ca’ Foscari University of Venice. Her doctoral project, *Assessing Fitness of Degendering Strategies in Molding Unbiased Gender Representations as a Result of an Adaptive Process*, is supervised by Professor Giuliana Giusti. She holds an MA in Language Sciences from Ca’ Foscari, specializing in linguistics for deafness and language impairments. Her MA thesis examined the interpretation of the *-ə* suffix as representative of the nonbinary population in Italian. She also earned a BA in European Languages and Cultures from the University of Modena and Reggio Emilia.