

DATABASE DESCRIPTION	Analysed dataset of gender identity labels and their lexical methods of word-formation together with their referential scope as identified in dictionaries and by the community.
Full reference	Filardo-Llamas, Laura & Alba Roldán-García. 2025. "Between "Queer" and "Cishet men". Lexical Creativity, Gender Identities and Sexualities in Linguistic Labels in Tumblr", <i>Journal of Language and Discrimination</i> 9:1, 1-27. DOI: https://doi.org/10.3138/jld-2025-050
Author and date of the database	Laura Filardo-Llamas and Alba Roldán-García; June 2023 for the original dataset; June 2024 for this analysis (see article for description of procedure)
Funding	Part of this research has been carried out as part of the SEMVARCOP grant (Semantic variation and communities of practice on social media) (PID2023-146434NB-100) funded by MICIU/AEI/10.13039/501100011033 and by European Regional Development Fund, EU. Alba Roldán-García has been funded by the Spanish Ministry of Science, Innovation, and Universities, as part of the predoctoral contract FPU23/00117.
Abstract (as taken from chapter)	Nowadays, social media has blurred the boundaries between offline and online reality, affecting every aspect of human experience, including individual and collective identity, as well as group formation and interaction (Zappavigna, 2011; Locher, 2015). Sometimes social media becomes intertwined to such a degree with the social fabric of specific communities that a deeper understanding of the latter requires a close examination of the former. This would be the case of the LGBTQ+ community and Tumblr (Byron et al., 2019). Tumblr, with its jargon and particular use of tags, has left an indelible mark in the current labels of the LGBTQ+ community (Oakley, 2016). In order to explore the semantic dimension of lexical creativity in relation with identity categorization, this article focuses on LGBTQ+ labels from a corpus of 100 Tumblr posts, from a previous study about figurative language and construals of gender on the platform (Filardo-Llamas & Roldán-García, 2024). This work aims at answering these research questions: First, which are the most common word formation processes in LGBTQ+ labels on Tumblr? Second, how do these lexical choices contribute to the characterization of different groups of people? And lastly, how does the activation of various mental images lead to creating inclusive or exclusive views of LGBTQ+ identities? The results seem to point to a tendency toward simple word formation processes to refer to new identities, while those labels, which have already settled in the jargon of the LGBTQ+ community on Tumblr often involve several word formation processes. Moreover, most LGBTQ+ labels entail multiple semantic dimensions triggering rather holistic views of gender and sexuality, as well as occasionally conflicting conceptualizations.
Keywords	Cognitive sociolinguistics, communities of practice, LGBTQ+ labels, semantics, categorization
NOTE	Following guidelines of research on Internet language, the original content upon which these labels are based have not been made available for anonymity reasons. These contents are available for research purposes upon demand to the corresponding authors.