
Computational Modeling of Register Variation in Positive Online Language

Maria Berger¹, Yulia Clausen¹ & Hannah Seemann²

¹Ruhr-Universität Bochum, ²Universität Tübingen

[maria.berger-a21, yulia.clausen]@rub.de, hannah.seemann@uni-tuebingen.de

While offensive language on social media has been extensively studied from both linguistic and computational perspectives, positive language – though equally prevalent – has received considerably less attention. Understanding and detecting positive language is essential for fostering inclusive online spaces. In this study, we examine three previously identified forms of positive online language: hope speech (Chakravarthi & Muralidaran 2021), candy speech (Clausen & Scheffler 2025) and empowering language (Njoo et al. 2023). Although these categories represent distinct phenomena, each characterized by unique linguistic properties and communicative functions, they exhibit a notable conceptual and functional overlap. Clarifying the distinctions between them is crucial for strengthening the theoretical foundation of positive language research and for developing effective computational models for their detection and classification.

Building on Biber and Conrad’s (2019) framework, we conceptualize these forms of positive language as distinct registers defined by their communicative purposes. To examine the similarities and differences among these registers, we employ both traditional machine learning methods, such as SVMs, and BERT-based language models. We extract features expected to be indicative of the studied registers (e.g., sentiment and social media-specific features) and apply clustering techniques to contextual embeddings. This data-driven approach allows us to refine the theoretical boundaries of these closely related concepts and supports a more accurate automatic identification of distinct forms of positive language in online contexts.

References: • Biber, D. & S. Conrad (2019). *Register, Genre, and Style* (2nd edn.). Cambridge University Press. • Chakravarthi, Bh. R. & V. Muralidaran (2021). Findings of the Shared Task on Hope Speech Detection for Equality, Diversity, and Inclusion. In *Proceedings of the First Workshop on Language Technology for Equality, Diversity and Inclusion*. Kyiv, 61–72. • Clausen, Y. & T. Scheffler (2025). Annotating candy speech in German YouTube comments. In *Proceedings of the 19th Linguistic Annotation Workshop (LAW XIX)*. • Njoo, L., Park, Ch., Stappart, O., Thielk, M., Chu, Y. & Y. Tsvetkov (2023). TalkUp: Paving the way for understanding empowering language. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, 9334–9354, Singapore.