The Corpus of Patient Documents (CoPaDocs): More Digging into Data 'from Below'

Markus Schiegg (Fribourg)

Medical records from psychiatric institutions of the 19th and early 20th century often contain texts, mostly letters, written by the former patients themselves (Schiegg 2022). As part of the Erlangen research project 'Flexible Writers in Language History' (2017–2023), over 4,000 these texts were published online (see http://copadocs.de; Brolich, Schiegg & Wiegand 2024). Currently, CoPaDocs is the largest digital corpus of historical, handwritten texts from German writers of diverse social backgrounds.

In this presentation, I will look back on more than ten years of work on data collection, transcription and corpus creation. Thereby, I illustrate the potentials of this corpus for historical sociolinguistics. I particularly dig into areas that go beyond existing research in historical linguistics in general and language history 'from below' in particular.

References

Brolich, Nina, Markus Schiegg & Frank Wiegand (2024): Edition historischer Patiententexte mit Präsenz im Deutschen Textarchiv und DWDS. In: *Digital Humanities im deutschsprachigen Raum (DhD) 2024*, Conference Poster University of Passau. https://doi.org/10.5281/zenodo.10706096

Schiegg, Markus (Ed.): *CoPaDocs – Korpus historischer Patiententexte*. http://copadocs.de

Schiegg, Markus (2022): *Flexible Schreiber in der Sprachgeschichte. Intraindividuelle Variation in Patientenbriefen (1850–1936)* (Germanistische Bibliothek 72). Heidelberg: Winter.