

Christopher Georgi, Susanne Haaf, Linda Kirsten & Frauke Thielert  
(Paderborn und Berlin)

## **Mixed Methods in Aktion: Das Projekt t.evo und die mehrdimensionale Annäherung an das Phänomen der Textsorte**

Textsorten sind ein Begriff der Alltagssprache: Was darunter zu verstehen ist, ist Sprecher\*innen grob geläufig. Geht es allerdings darum, das Wesen von Textsorten und die Kriterien ihrer Differenzierung zu beschreiben, bleiben auch nach langjähriger fachwissenschaftlicher Diskussion innerhalb der Linguistik noch immer viele Fragen offen.

Auseinandersetzungen mit dem Gegenstand erfolgten zunächst anhand sehr überschaubarer Mengen von Textbeispielen und nur selten im Kontrast verschiedener Textsorten. Mit dem Vorhandensein großer Korpora jedoch ist es nun möglich, viele Vertreter einer Textsorte mit automatisierten Verfahren und im Vergleich mit großen Korpora anderer Textsorten auf ihre konstitutiven Merkmale hin zu untersuchen. Zudem stehen mittlerweile zahlreiche Tools und Werkzeuge zur Verfügung, die die automatische Sprachverarbeitung ermöglichen bzw. die manuelle Erschließung des Materials mittels Annotation unterstützen. Trotzdem dürfen die daraus sich ergebenden methodischen Möglichkeiten in ihrer Leistungsfähigkeit auch nicht überschätzt werden. Quantitatives muss an qualitatives, automatisches an manuelles Arbeiten rückgekoppelt werden und vice versa, um belastbare Ergebnisse zu erzielen – und es bleibt das Spannungsfeld zwischen der kritischen Masse und den Grenzen der Fehlertoleranz.

Im Projekt t.evo (DFG, 2019–2022) wird genau diese Synthese angestrebt. Ziel ist, anhand von Korpora erbaulicher Textsorten des 17. und 18. Jhs. sowie der Pressekommunikation des 19. und 20. Jhs. (die jeweils als Massenmedium ihrer Zeit gelten können) möglichst allgemeingültige Kriterien zur Textsortendifferenzierung herauszuarbeiten. Hierzu wurden in Rückgriff auf den fachwissenschaftlichen Diskurs verschiedene Dimensionen festgelegt, in denen die musterhaften Erscheinungen verschiedener Textsorten sichtbar und somit analysierbar gemacht werden sollen. Dabei bedienen wir uns sowohl automatisierter als auch rein manueller Zugänge zum Korpusmaterial und kombinieren quantitative mit qualitativen Auswertungsweisen.

Unser Vortrag wird in diese Arbeiten einführen und dabei nicht nur die einzelnen Arbeitsschritte näher skizzieren, sondern auch Fallstricke der einzelnen Methoden und schließlich deren Verzahnung zugunsten einer zusammenhängenden Textsortenbeschreibung eingehen.