

Andrea Abel/Jennifer-Carmen Frey/Aivars Glaznieks (Bozen)/
Maja Linthe/Carolyn Müller-Spitzer/Angelika Storrer/
Sascha Wolfer (Mannheim)

Wie misst man Textqualität im digitalen Zeitalter? (MIT.Qualität)

Schreibprodukte in sozialen Netzwerken entsprechen nicht immer den normativen Erwartungen, die man an redigierte Texte in der Presse, der Wissenschaft oder der Belletristik heranträgt. Die Frage, wie sich dies auf unsere Schriftsprache und auf die Entwicklung von Schreibkompetenzen auswirkt, wird in der Öffentlichkeit kontrovers diskutiert (vgl. Dürscheid/Brommer 2009; Storrer 2014; Dürscheid/Frick 2016). Allerdings gibt es bislang wenig Möglichkeiten, empirisch fundierte Antworten auf diese Frage zu geben. Das liegt einerseits daran, dass es aufwändig ist, Korpora mit Lernertexten aus dem schulischen Bereich und Schreibprodukten aus sozialen Medien zu sammeln und für die empirische Analyse aufzubereiten. Es liegt aber auch daran, dass eine „einheitliche Antwort auf die Frage, was eine gute Schreibkompetenz ist, noch immer aussteht“ (Dürscheid/Frick 2016, S. 125). Das Projekt „Wie misst man Textqualität im digitalen Zeitalter?“ (MIT.Qualität)¹ setzt an dieser Lücke an und möchte in zwei Arbeitsbereichen dazu beitragen, die Voraussetzungen für die empirische Untersuchung des Schreibgebrauchs und der Schreibkompetenzen in digitalen Medien zu verbessern.

Im Arbeitsbereich „Modellbildung und Kategorienfindung“ wird ein Modell entwickelt, das neue Qualitätsaspekte, die für das vernetzte, multimodale und dialogische Kommunizieren in sozialen Medien relevant sind, in Traditionen des Nachdenkens über Sprachqualität in Rhetorik und Textlinguistik (u.a. Fritz 2008; Sieber 2008) einordnet. Für uns steht außer Frage, dass der u.a. von Lobin (2014) beschriebene Wandel der Kulturtechniken des Lesens und Schreibens ein solches erweitertes Modell erforderlich macht. Einerseits kann beim dialogischen Schreiben in sozialen Medien die Schnelligkeit der Reaktion für den kommunikativen Erfolg wichtiger sein als das Einhalten von grammatischen und orthografischen Normen. Dies muss bei der Bewertung von Normabweichungen ebenso berücksichtigt werden wie der Einfluss von Schreib- und Korrekturhilfen. Andererseits werden in sozialen Medien andere Qualitäten relevant, z.B. Schlagfertigkeit, das

¹ Das Projekt wird von Dezember 2017 bis Mai 2019 im Rahmen der Ausschreibung „Originalitätsverdacht. Neue Optionen für die Geistes- und Kulturwissenschaften“ der VolkswagenStiftung gefördert. Beteiligt sind die Universität Mannheim, Eurac Research in Bozen und das Institut für Deutsche Sprache (IDS).

kreative Kombinieren von Inhalten, der adäquate Umgang mit Trollen und Shitstorms oder die Suchmaschinenoptimierung. Gesellschaftlich sind zunehmend Kompetenzen gefragt, die bislang größtenteils nicht über Bildungsinstitutionen, sondern über sprachliche Alltagspraktiken selbst angeeignet werden (vgl. Barton/Lee 2013, S. 137–152). Das erweiterte Modell dient einerseits als Ausgangspunkt, um die Veränderungen im Gebrauch von Schriftsprache im digitalen Zeitalter empirisch zu untersuchen und damit die bislang auf schmaler Datengrundlage geführte Debatte über mögliche Auswirkungen digitaler Medien auf Schreibkompetenzen zu versachlichen. Es soll aber auch dazu beitragen, die neuen Aspekte digitaler Textqualität in die Bildungsinstitutionen hineinzutragen und die bislang auf monologische Schrifttexte fokussierte Perspektive zu erweitern. Als Quellen zur „Entdeckung“ neuer Kategorien werten wir Ratgeber zum Schreiben in sozialen Medien aus; außerdem haben wir ein Blog² zum Projektthema eingerichtet mit dem Ziel, auch Menschen, die regelmäßig im Netz schreiben, in die Kategorienfindung mit einzubinden. Auf einem Workshop im Juni 2018 wurde ein erster Entwurf unseres Modells mit ExpertInnen aus der Wissenschaft und aus dem Social-Media-Bereich diskutiert, um frühzeitig Lücken erkennen und Anregungen berücksichtigen zu können. Weiterhin planen wir eine Fragebogenerhebung mit linguistischen Laien, in der wir Einschätzungen zur Angemessenheit konkreter Bausteine in Texten einholen. Auf diese Weise möchten wir die Urteile von ExpertInnen, wie wir sie u.a. auf unserem Workshop einholen, mit den intuitiven Einschätzungen einer breiter gestreuten Gruppe von SprecherInnen des Deutschen abgleichen. Im Sinne der Methodenreflexion ist dieser Projektteil insofern interessant, als die Befragungen von linguistischen Laien Ergebnisse hervorbringen können, die mit den Einschätzungen von linguistischen und Social-Media-ExpertInnen konvergieren, in bestimmten Fällen aber nicht immer deckungsgleich sein müssen.

Im Arbeitsbereich „Methodenexploration“ werden für die im Projekttitel genannte Leitfrage Messmethoden entwickelt und in Fallstudien evaluiert. Dabei konzentrieren wir uns auf den Bereich „Kohärenz und Kohäsion“. Zwei geplante Fallstudien widmen sich der kontrastiven Analyse des Gebrauchs sprachlicher Mittel zur Herstellung von Kohärenz in traditionellen und Social-Media-Texten anhand von Einwort-Konnektoren (wie *weil, eh, sprich*) und Mehrwort-Konnektoren (wie *im Übrigen, das heißt*) (vgl. Pasch et al. 2003). Dazu werden korpuslinguistische Methoden der automatischen Extraktion von Daten aus geeigneten Korpora angewandt, d.h. aus CMC-Korpora (DiDi: Frey/Glaznieks/Stemle 2016), Wikipedia (Margaretha/Lüngen 2014), L1-Lernerkorpora (KoKo: Abel et al. 2016) und Zeitungskorpora (DEREKO/IDS). Zudem werden relevante linguistische Phänomene

² Vgl. <https://mitqualitaet.com/blog-zur-textqualitaet-von-online-texten/> (Stand: 12.7.2018).

manuell in Subkorpora annotiert. Eine dritte Fallstudie schließlich befasst sich mit automatischen Methoden zur Bewertung von Textqualität. Dazu werden holistische Textqualitätsbewertungen in einem automatischen Textklassifizierungssystem modelliert und mit den Bewertungen in Beziehung stehende sprachliche Merkmale identifiziert und kritisch bewertet (vgl. McNamara/Crossley/McCarthy 2010).

In den Fallstudien wählen wir bewusst den Bereich „Kohärenz und Kohäsion“, weil er zentral ist für die bislang noch nie empirisch untersuchte Frage, ob und wie sich das häufige Lesen und Schreiben kurzer, in Dialoge eingebundener Botschaften auf die Kompetenz zum Planen und Verfassen monologischer, längerer Schrifttexte auswirkt. Nicht zufällig spielen Aspekte von Kohäsion und Kohärenz in Systemen zur automatischen Textbewertung für englischsprachige Texte eine wichtige Rolle und werden in Tools wie TAACO,³ Coh-Metrix⁴ oder CTAP⁵ aufgegriffen. Wir sind uns allerdings der Tatsache bewusst, dass die empirische Untersuchung von Kohärenz ohne bereits verfügbare und validierte Methoden für die deutsche Sprache mit großen methodischen Herausforderungen verbunden ist. Wir können in unserem Projekt diese Lücke nicht schließen, möchten aber in den Fallstudien Möglichkeiten und Grenzen zur automatischen Bewertung von Kohäsions- und Kohärenzaspekten explorieren.

Literatur

- Abel, Andrea et al. (2016): An extended version of the KoKo German L1 Learner corpus. In: Corazza/Montemagni/Semeraro (Hg.), S. 13–18.
- Barton, David/Lee, Carmen (2013): Language online. Investigating digital texts and practices. London/New York.
- Corazza, Anna/Montemagni, Simone/Semeraro, Giovanni (Hg.) (2016): Proceedings of the Third Italian Conference on Computational Linguistics CLiC-it 2016. 5.–6.12.2016, Napoli. Turin.
- Dürscheid, Christa/Brommer, Sarah (2009): Getippte Dialoge in neuen Medien. Sprachkritische Aspekte und linguistische Analysen. In: Linguistik online 37, 1. Internet: www.linguistik-online.de/37_09/duerscheidBrommer.pdf (Stand: 20.8.2013).
- Dürscheid, Christa/Frick, Karina (2016): Schreiben digital. Wie das Internet unsere Alltagskommunikation verändert. (= Einsichten 3). Stuttgart.
- Frey, Jennifer-Carmen/Glaznieks, Aivars/Stemle, Egon W. (2016): The DiDi Corpus of South Tyrolean CMC Data: A multilingual corpus of Facebook texts. In: Corazza/Montemagni/Semeraro (Hg.), S. 157–161.

³ Vgl. www.kristopherkyle.com/taaco.html (Stand: 12.7.2018).

⁴ Vgl. www.cohmetrix.com (Stand: 12.7.2018).

⁵ Vgl. www.ctapweb.com/ (Stand: 12.7.2018).

- Fritz, Gerd (2008): Bessere Texte schreiben. Überlegungen zur Textqualität aus der Sicht einer dynamischen Texttheorie. In: *Sprache und Literatur* 39/102, S. 75–105.
- Lobin, Henning (2014): *Engelbarts Traum. Wie der Computer uns Lesen und Schreiben abnimmt*. Frankfurt a.M./New York.
- Margaretha, Eliza/Lüngen, Harald (2014): Building linguistic corpora from Wikipedia articles and discussions. In: *Journal of Language Technology and Computational Linguistics* 29/2, S. 59–82. Internet: www.jlcl.org/2014_Heft2/3MargarethaLuengen.pdf (Stand: 27.3.2018).
- McNamara, Danielle/Crossley, Scott/McCarthy, Philip (2010): Linguistic features of writing quality. In: *Written Communication* 27, 1, S. 57–86.
- Pasch, Renate et al. (2003): *Handbuch der deutschen Konnektoren: linguistische Grundlagen der Beschreibung und syntaktische Merkmale der deutschen Satzverknüpfers (Konjunktionen, Satzadverbien und Partikeln)*. Bd. 1. (= *Schriften des Instituts für Deutsche Sprache* 9). Berlin/New York.
- Sieber, Peter (2008): Kriterien der Textbewertung am Beispiel *Parlando*. In: Janich, Nina (Hg.): *Textlinguistik*. 15 Einführungen. Tübingen, S. 271–290.
- Storrer, Angelika (2014): Sprachverfall durch internetbasierte Kommunikation? Linguistische Erklärungsansätze – empirische Befunde. In: Plewnia, Albrecht/Witt, Andreas (Hg.): *Sprachverfall? Dynamik – Wandel – Variation*. *Jahrbuch des Instituts für Deutsche Sprache* 2013. Berlin/New York, S. 171–196.