

# m\*w Figurengender zwischen Stereotypisierung und literarischen und theoretischen Spielräumen. Genderstereotype und - bewertungen in der Literatur des 19. Jahrhunderts

**Schumacher, Mareike**

mareike.schumacher@uni-hamburg.de  
University of Hamburg, Deutschland

**Flüh, Marie**

marie.flueh@uni-hamburg.de  
University of Hamburg, Deutschland

Während in den Digital Humanities bereits erste korpusbasierte Analysen von Figurengender in der Literatur vorgelegt wurden (Underwood 2019: 111 ff), wird in den Kulturwissenschaften zu diesem Thema selten korpusbasiert gearbeitet. Stattdessen sind Theorien zur Genderthematik häufig philosophisch-soziologisch (z. B. bei Beauvoir oder Bourdieu), diskurstheoretisch (z. B. Foucault) oder dekonstruktivistisch (z. B. Butler)<sup>1</sup> motiviert. Die Lücke zwischen einer an technischen Methoden ausgerichteten Modellierung und der theoretischen Betrachtung der Genderthematik schließen wir mit dem Projekt m\*w, von dem wir erste Pilotstudien hier vorstellen. In einem theoriegeleiteten Mixed-Methods-Ansatz operationalisieren wir zunächst klassische Ansätze aus den Genderstudies und wenden das entstandene Modell sowohl mit Hilfe von Named Entity Recognition als auch mittels digitaler Annotation auf ein literarisches Korpus an. Das Zusammenspiel von theoretischer und datenbasierter Modellierung und die Bereitstellung von Zwischenergebnissen und "losen Enden" auf der *Projektwebseite* (<https://msternchenw.de>) eröffnen Spielräume, um weiter- und umzudenken und sich mit anderen Projekten zusammenzuschließen.

## Konzeptioneller Rahmen und methodische Desiderata

Die leitende Fragestellung des m\*w-Projektes ist: Wie werden Genderrollen in der Literatur des 19. Jahrhunderts dargestellt und bewertet? Um uns dieser Fragestellung zu nähern, erstellen wir mithilfe einer Auswahl theoretischer Ansätze ein Modell. Dieses nutzen

wir, um im überwachten Machine-Learning-Verfahren der Named Entity Recognition (NER) ein Tool darauf zu trainieren, Figuren und ihre Genderzuschreibungen automatisch zu erkennen. Die Ergebnisse des NER-Verfahrens nutzen wir um unser Modell weiter zu schärfen. Im abschließenden Close Reading der Novellen werden schließlich Genderbeschreibungen und -bewertungen analysiert und mit dem Modell abgeglichen. Grundsätzlich nehmen wir sowohl Modell als auch Korpus als variable bzw. dynamische Größen wahr, die sich für den Praxistest der Operationalisierung theoriebasierter Modelle zur Erforschung literarischer Genderrollen eignen.

## Korpus

Forschungsgegenstand ist der *Deutsche Novellenschatz*, der 1871–1876 von Paul Heyse und Hermann Kurz herausgegeben wurde und 87 Novellen umfasst.<sup>2</sup> Da in dieser Textsammlung die Thematik der Ehe von großer Bedeutung ist (vgl. Weitin/Herget 2017), gehen wir davon aus, dass stereotype Genderrollen sich potentiell häufen. Zur tatsächlichen, möglicherweise diversen Genderausrichtung der Autor\*innen ist nichts bekannt. Um einerseits einen möglichst ausgewogenen Forschungsgegenstand zu bekommen, andererseits aber möglichst wenig auf das Korpus einzuwirken, wurden zunächst alle 12 aus der Feder von Schriftstellerinnen\* stammenden Novellen in ein Teilkorpus übernommen. Ergänzend dazu wurden 12 per Zufallsgenerator ausgewählte Novellen von Autoren\* ergänzt. Die verbleibenden Novellen wurden als Trainingsmaterial genutzt. Bei der Zusammenstellung des Korpus war uns bewusst, dass wahrscheinlich überwiegend Autorenpersönlichkeiten einbezogen wurden, die sich einem binären Geschlechtermodell zugehörig fühlten. Um Trainingsmaterial für Anschlussforschung über die Pilotstudien hinaus zur Verfügung zu stellen, planen wir ein weiteres Korpus zu erstellen, das Texte enthält, die von Personen geschrieben wurden, die sich nachweislich nicht in ein binäres System integrieren lassen (wollten).<sup>3</sup>

## Genderstereotype

Davon ausgehend, dass sich sowohl in der Theorie als auch in Erzähltexten Spielräume als Zwischenräume auftun, die dadurch sichtbar werden, dass sie sich von den sie umgebenden normierten Räumen unterscheiden<sup>4</sup>, untersuchen wir stereotype Darstellungen und solche, die sich davon abheben in einem relationalen Ansatz. Anschließend an Butler (vgl. Butler 1990: 190-219) differenzieren wir den Begriff Gender nach Geschlecht, Gender Identität und Gender Performanz und übersetzen diese Trias für die Anwendung auf das literarische Korpus in Geschlechtszuordnung durch Personalpronomen, Ausdruck der Gender-Identität durch Figureneigenschaften und Beschreibung der Gender-Performanz durch

Aufzeigen (nicht) rollenkonformer Figurenhandlungen. Die Betrachtung der Geschlechtszuordnung durch Personalpronomen schließen wir aus, da der dieser binären Zuordnung zu Grunde liegende Biologismus fraglich ist, wie z.B. die Ausführungen Foucaults (vgl. Foucault 1998: 7-18) zu Herculine Barbin<sup>5</sup> sowie dessen eigene Lebenserinnerungen (vgl. ebd. 19-126) zeigen.

Um entscheiden zu können, ob Eigenschaften und Handlungen von Figuren stereotypen Rollenbildern zugeschrieben werden können oder nicht, haben wir zunächst möglichst viele Rollenbilder in unser Modell integriert. Jede dieser Rollen kann sowohl im Sein (Gender-Identität) als auch im Handeln (Gender-Performanz) von Figuren verankert sein. Die sechs in Abb. 1 abgebildeten Oberkategorien von Eigenschaften sind ebenfalls der Theorie entnommen. Allerdings wurden hier die Beschreibungen stärker kondensiert, um die zahlreichen genannten Einzeleigenschaften für die digitale Annotation besser handhabbar zu machen.

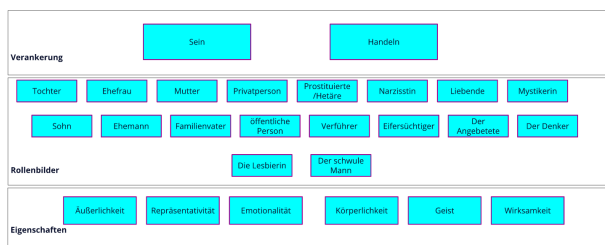


Abbildung 1: Theoriebasiertes Gender-Modell

## Erste Ergebnisse der Analysen

Alle drei eingesetzten Methoden - NER, Annotation von Stereotypen und Emotionsanalyse haben sich im ersten Proof of Concept als fruchtbar für die digitale Erforschung von Figurengender erwiesen. Besonders eklatante Zwischenergebnisse fassen wir im Folgenden zusammen.

### NER

Im Laufe des NER-Trainingsprozesses erwies sich die Kategorie "divers" nicht als dem Korpus angemessen, weshalb wir die NER-Kategorien auf "männlich", "weiblich", "genderneutral"<sup>6</sup> festlegten. Mit einer schrittweise Ausweitung des Trainingsmaterials von 40.000 auf 68.000 Tokens<sup>7</sup> erreichten wir eine Steigerung der anfänglich bei rund 0,47 liegenden F1-Score-Werte um 0,03 in einem ersten und 0,07 in einem zweiten Testtext (*Irrwisch-Fritze* von Reinhold und *Die drei Schwestern* von Kähler). In beiden Testtexten zeigte sich, dass genderneutrale Figuren am zuverlässigsten und weibliche Figuren am schlechtesten erkannt werden (für detailliertere

Beschreibungen der ersten NER-Tests vgl. Schumacher 2020 [2]).

Um Hinweise auf eine mögliche Verzerrung durch die Zusammensetzung des Trainingskorpus zu bekommen, haben wir einen dritten Testtext hinzu genommen (*Eine fromme Lüge* von Gall). Die Ergebnisse des Tests weichen von den vorherigen ab. Der F1-Score liegt bei 0,6202 und ist damit um 0,08 höher als der höhere der beiden vorherigen Testtexte. Außerdem werden weibliche Figuren besser erkannt als männliche (F-Scores von 0,7093 und 0,4293). Ein unerwarteter Nebenfund aus dem wir nach intensiver Prüfung des Ergebnisses (mehr darüber in Schumacher 2020 [1]) schließen, dass das Tool hier einen Text mit sehr stereotypen Rollenbeschreibungen erkannt hat. Dass das im Hinblick auf Autorengender zu homogene Trainingskorpus verzerrend wirkt, konnte durch diesen Test hingegen nicht endgültig bestätigt werden. Ein inhaltlicher Vergleich der NER-Ergebnisse zeigt aber, dass die Anzahl männlicher Figurenbenennungen insgesamt höher ist (Abb. 3).



Abbildung 3: Die NER-Ergebnisse der Testtexte visualisiert; links steht Hellblau, mittig Dunkelblau und rechts rosa für männliche Figuren.

## Digitale Annotation von Genderstereotypen

Im Annotationstool CATMA (Meister et al. 2019) wurden die NER-Ergebnisse zuerst ergänzt, sodass alle tatsächlich vorkommenden weiblichen, männlichen und neutralen Figurenbezeichnungen im Testtext *Die drei Schwestern* annotiert wurden. Der Eindruck, dass es hier am meisten männliche Figuren gibt, wird bestätigt (Abb. 4).

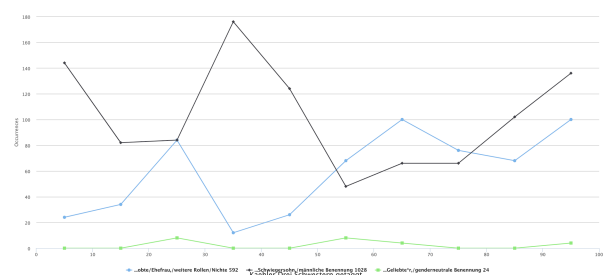


Abbildung 4: Gender-Kategorien im Verlauf des Beispiel-Textes (männlich (schwarz), weiblich (blau) und genderneutral (grün)).

Anschließend wurden die NER-Kategorien mit Unterkategorien versehen, die den Benennungen der Genderrollen des Modells entsprechen (Abb. 5).

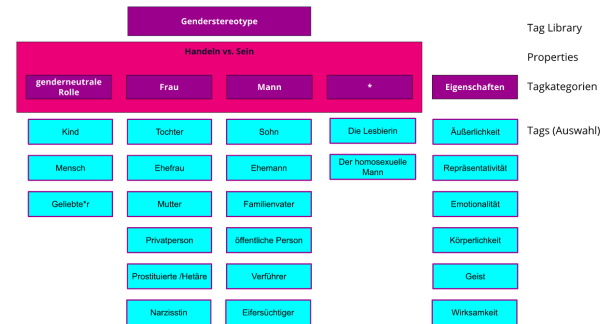


Abbildung 5: Vorläufiges Tagset (Auswahl der angewandten weiblichen und männlichen Rollen/ nur Oberkategorien der Eigenschaften).

Sofern es keine adäquate, in der Theorie erwähnte Rolle oder Eigenschaft gab, wurden zusätzliche Annotationskategorien erstellt und den Oberkategorien "unsortierte Rollen" und "unsortierte Eigenschaften" zugewiesen.

Die digitale Annotation des ersten Beispieltextes zeigt, dass stereotype Beschreibungen vor allem zu Beginn der Erzählung häufig und somit besonders für die Etablierung der Figuren von Bedeutung sind. Stereotype Eigenschaften sind hier zwar divers, für männliche und weibliche Figuren gibt es aber jeweils einige wenige, die quantitativ herausstechen. Für weibliche Figuren ist das vor allem Äußerlichkeit/Schönheit für männliche sind es Körperlichkeit/Trinkfestigkeit und Wirksamkeit/Herrschaft. Unsortierte Rollen und Eigenschaften sind zumeist in einem binären Rollensystem verankert und dennoch werden in dieser Novelle zum Teil Stereotype aufgebrochen. Dies wird hauptsächlich durch die Zuschreibung einzelner Eigenschaften zu einer Figur eines Genders erreicht, die in der theoretischen Literatur eher dem Stereotyp des anderen zugeschrieben werden (ausführlichere Auswertung des ersten Beispieltextes in Schumacher 2020 [3]).

## Emotionsanalyse mit CATMA

Die Auswertung der digitalen Annotation des Beispieltextes *Gemüth und Selbstsucht* von Margarethe von Wolff macht folgende Einsichten besonders deutlich: Die Emotionen Zorn, Ekel, Trauer und Liebe – die sich wiederum aus unterschiedlichen Vertretern dieser Emotionsfamilien zusammensetzten – treten am häufigsten auf.

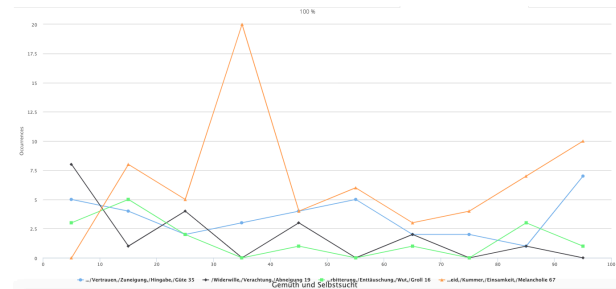


Abbildung 6: Der Verteilungsgraph zeigt den Emotionsverlauf in Gemüth und Selbstsucht: TRAUER (orange), LIEBE (blau), EKEL (schwarz), ZORN (grün)

Die Emotionen werden in den allermeisten Fällen verbal (407 Annotationen) von den Figuren ausgedrückt. Über die Veränderung des körperlichen Zustands (19 Annotationen), und nonverbal (107 Annotationen) werden Emotionen vergleichsweise selten repräsentiert. Die Auswertung der Properties ergibt genderspezifische Emotionsinformationen (s. Abb. 7 und 8). Weibliche Figuren treten ängstlicher auf als männliche. Männliche Figuren reagieren häufiger zornig als weibliche. Sie empfinden außerdem häufiger Ekel – hier meistens im Sinne von Abneigung – als weibliche Figuren. Diese leiden häufiger unter gedrückter Stimmung und empfinden deutlich häufiger negative Basisemotionen als männliche Figuren. Diese treten im Schnitt fröhlicher auf als die weiblichen Figuren. Die männlichen Figuren zeigen häufiger positive Basisemotionen als die weiblichen Figuren und auch die positive Basisemotion LIEBE überwiegt seitens der männlichen Figuren. Scham – als Unterkategorie der Problemfälle – ist seitens der weiblichen Figuren deutlich stärker ausgeprägt (für eine ausführlichere Auswertung des ersten Beispieltextes vgl. Flüh 2020).

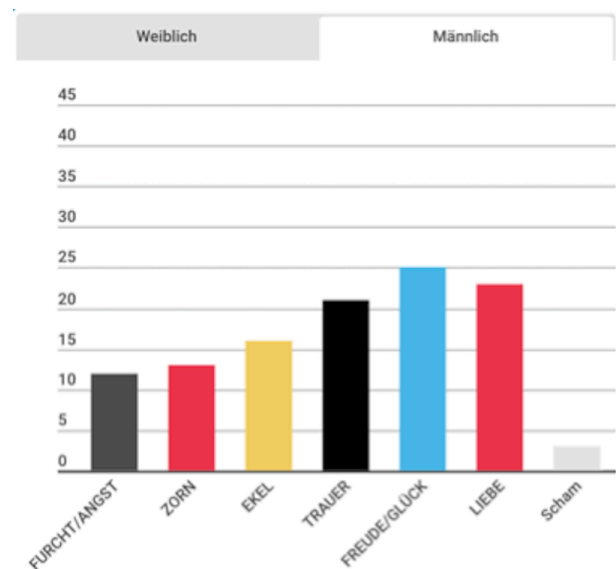


Abbildung 7: Gesamtvorkommen der mit der Value männlich ausgezeichneten Emotionen

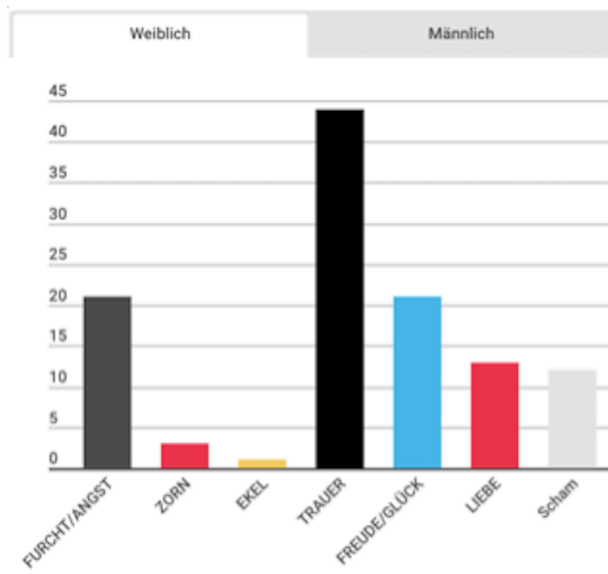


Abbildung 8: Gesamtheit der mit der Value weiblich ausgezeichneten Emotionen

## Theorien hinterfragen, aufbrechen und weiterdenken

Im Hinblick auf die leitende Fragestellung können zu diesem Zeitpunkt drei Zwischenergebnisse festgehalten werden. Der Trainingsprozess des NER-Modells zeigt, dass im deutschen Novellenschatz hauptsächlich weibliche, männliche und genderneutrale Figurenbeschreibungen eine Rolle spielen. Um diverse Genderrollen in der Literatur des 19. Jahrhunderts erforschen zu können, muss weiteres Trainingsmaterial erstellt werden. Die Tests des NER-Modells machen deutlich, dass die Methode der CRF-Vorhersage geeignet ist, um innerhalb des Novellenschatzes Texte mit besonders stereotypen Genderzuschreibungen ausfindig zu machen.

Erste Auswertungen der digitalen Annotation von Genderstereotypen zeigen, dass in Beschreibungen von Figuren zwar eine Vielzahl von Eigenschaften genutzt wird, dass genderspezifisch aber jeweils einige wenige Kategorien vorherrschen. Für unseren Beispieltext konnten wir belegen, dass stereotype Genderrollen nicht durch die Einführung diverser Figuren aufgebrochen werden, sondern durch die Kennzeichnung einzelner Figuren mit einzelnen stereotypen Eigenschaften eines anderen Genders. Die digitale Annotation zeigt auch, dass genderneutrale Figurenbezeichnungen eher selten sind. Erste Auswertungen der Emotionsanalysen zeigen, dass genderspezifische Unterschiede v. a. hinsichtlich der Wertung der referierten Emotionen bestehen; männliche Figuren greifen auf ein positiveres Gefühlskonzept bzw. Bewertungskonzept zurück als weibliche. Die strukturorientiert fundierten Tagsets eignen sich, um

die vielfältigen Ausdrucksweisen genderspezifischer Emotionsmanifestationen und -familien aus literarischen Texten herauszufiltern. Im weiteren Verlauf gilt es, die übrigen Texte auszuwerten und die Analyseergebnisse miteinander in Bezug zu setzen.

## Fußnoten

1. Die hier ausgewerteten Ansätze der Genderstudies beziehen sich nicht primär auf literarische Korpora, die in diesen Theorien ausgewerteten Beispiele sind aber häufig literarischer Art. Die daran erkennbare Nähe zu den Literaturwissenschaften erleichtert die im m\*w vorgenommene Domänenadaption der Ansätze.
2. Wir nutzen die digitale Version des deutschen Novellenschatzes, die vom Discourse Lab herausgebracht wurde und unter <https://www.discourselab.de/cqpweb/> heruntergeladen werden kann. Die von uns zum Zweck des Machine Learning Trainings erstellen Daten (sowohl Trainingsdaten als auch Modell) werden auf der Projektwebseite unter [http://msternchenw.de/m\\*w-insiders/](http://msternchenw.de/m*w-insiders/) zur Nachnutzung bereitgestellt.
3. Dieses "Diversitäts-Korpus" wird ebenfalls auf der Projektseite unter [http://msternchenw.de/m\\*w-insiders/](http://msternchenw.de/m*w-insiders/) zur Verfügung gestellt werden.
4. Deutlich wird dies in der Genderforschung z.B. bei Beauvoir, die zunächst eine Reihe traditioneller Rollenmuster der Frau beschreibt, um sich dann der lesbischen Frau zuzuwenden, die hier ausbricht (vgl. Beauvoir 1949) oder bei Bourdieu, der zunächst ein binär organisiertes Herrschaftssystem beschreibt, um sich dann abschließend der Liebe als Möglichkeit, dieses zu überwinden und der Homosexuellen-Bewegung, die in seinen Augen die männliche Herrschaft eher stützt, zuzuwenden (vgl. Bourdieu 2005). Interessant ist in diesem Zusammenhang auch der gar nicht auf Geschlechterfragen ausgerichtete Ansatz Foucaults zur Heterotopologie, in dem er sich der Identifizierung von Heterotopien als Räume des anderen widmet, die Ausdifferenzierung des "normalen" Raumes aber unterlässt, wodurch eine Leerstelle entsteht (vgl. Foucault 2005). Letzteres soll hier vermieden werden.
5. Die Lebenserinnerungen Herculine Barbins zeigen einen Fall von Hermaphroditismus im 19. Jahrhundert in Frankreich.
6. Auch diese Kategorien sind für uns nur als vorläufig gedacht und können jederzeit an unterschiedliche Korpora angepasst bzw. mit Hilfe von anderen, zusätzlichen Korpora weiterentwickelt werden.
7. Das Trainigskorpus wird auch weiterhin laufend erweitert.

## Bibliographie

Argyle, Michael (1996): *Körpersprache und Kommunikation*. Paderborn: Junfermann.

- Barbin, Herculine / Foucault, Michel** (1998): *Über Hermaphroditismus*. Frankfurt am Main: Suhrkamp.
- Beauvoir, Simone** (1949): *Das andere Geschlecht*. Reinbek: Rowohlt.
- Bourdieu, Pierre** (2005): *Die männliche Herrschaft*. Frankfurt am Main: Suhrkamp.
- Borod, Joan C.** (2000): *The Neuropsychology of Emotion*. New York: Oxford University Press.
- Butler, Judith** (1990): *Das Unbehagen der Geschlechter*. Frankfurt am Main: Suhrkamp.
- Connell, Raewyn** (1999): *Der gemachte Mann. Konstruktion und Krise von Männlichkeiten*. Wiesbaden: Springer.
- Darwin, Charles** (1884): *Der Ausdruck der Gemütsbewegungen bei dem Menschen und den Thieren*. Stuttgart: Schweizerbart.
- Ekman, Paul** (1999): *Basic Emotions*. In: Dalglish, Tim und Michael J. Power (Hrsg.): *Handbook of Cognition and Emotion*. Chichester: John Wiley & Sons, 45–60.
- Erhart, Walter** (2001): *Familienmänner. Über den literarischen Ursprung moderner Männlichkeit*. München: Wilhelm Fink.
- Finkel, Jenny Rose / Grenager, Trond / Manning, Christopher** (2005): Incorporating Non-local Information into Information Extraction Systems by Gibbs Sampling. In: *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL 2005)*, Seiten 363-370. <http://nlp.stanford.edu/~manning/papers/gibbscrf3.pdf> [zuletzt geprüft: Dezember 21, 2019]
- Flüh, Marie** (2020): "Emotionsinformationen analysieren: 3. Erste Ergebnisse – Gemüth und Selbstsucht (1787) von Margarethe von Wolff." In: *m\*w*, Januar 4, 2020, <https://msternchenw.de/?p=249>, [zuletzt geprüft: Januar 4, 2020].
- Foucault, Michel** (2005): *Die Heterotopien. Der utopische Körper*. Frankfurt am Main: Suhrkamp.
- Foucault, Michel** (2008): *Die Ordnung der Dinge*. In: Michel Foucault. Die Hauptwerke. Frankfurt am Main: Suhrkamp.
- Heydebrand, Renate von / Simone Winko** (1996): *Einführung in die Wertung von Literatur*. Paderborn/München/Wien u. a.: Ferdinand Schöningh.
- Irigaray, Luce** (1980): *Speculum. Spiegel des anderen Geschlechts*. Frankfurt am Main: Suhrkamp.
- Izard, Carroll, E.** (1992): "Basic emotion, relations among emotions, and emotion-cognition relations". In: *Psychological Review* 99 (3), 561–565.
- Luhmann, Niklas** (2008): *Schriften zur Kultur und Literatur*. Frankfurt am Main: Suhrkamp.
- Meister, Jan Christoph / Horstmann, Jan / Petris, Marco / Jacke, Janina / Bruck, Christian / Schumacher, Mareike / Flüh, Marie** (2019): CATMA 6.0.0 (Version 6.0.0). Zenodo. DOI: .
- Moretti, Franco** (2013): *Operationalizing: or the function of measurement in modern literary theory*. <https://litlab.stanford.edu/LiteraryLabPamphlet6.pdf> [zuletzt geprüft: August 8, 2019].
- Oatley, Keith / P.N. Johnson-Laird** (1987): "Towards a cognitive theory of emotions". In: *Cognition and Emotion* 1, 1987, 29–50.
- Plutchik, Robert** (1984): "Emotions. A general psycho-evolutionary theory". In: Klaus Scherer und Paul Ekman (Hrsg.): *Approaches to emotion*. Hillsdale: Erlbaum, 197–200.
- Schumacher, Mareike** (2020 [1]): "Named Entity Recognition und Reverse Engineering". In: *Lebe lieber literarisch*, Januar 2, 2020, <http://lebelieberliterarisch.de/named-entity-recognition-und-reverse-engineering/>, [zuletzt geprüft: Januar 3, 2020].
- Schumacher, Mareike** (2020 [2]): "Automatische Erkennung von Figuren-Gender – das erste Modell." In: *m\*w*, Januar 3, 2020, <https://msternchenw.de/automatische-erkennung-von-figuren-gender-das-erste-modell>, [zuletzt geprüft: Januar 3, 2020].
- Schumacher, Mareike** (2020 [3]): "Genderstereotype in der Literatur - erste Analysen". In: *Lebe lieber literarisch*, Januar 7, 2020, <http://lebelieberliterarisch.de/genderstereotype-in-der-literatur-erste-analysen/>, [zuletzt geprüft: Januar 7, 2020].
- Schwarz-Friesel, Monika** (2017): "Das Emotionspotenzial literarischer Texte." In: Betten, Anne, Ulla Fix und Berbeli Wanning (Hrsg.): *Handbuch Sprache in der Literatur*. Berlin, Boston: De Gruyter, 351-370.
- Simmel, Georg (1986): "Das Relative und das Absolute im Geschlechter-Problem." In: *Philosophische Kultur. Über das Abenteuer, die Geschlechter und die Krise der Moderne*. Berlin: Klaus Wagenbach, S. 64–93.
- Stephan, Inge** (2003): "Im toten Winkel. Die Neuentdeckung des ‚ersten Geschlechts‘ durch men's studies und Männlichkeitsforschung." In: dies., Claudia Benthien (Hrsg.): *Männlichkeit als Maskerade. Kulturelle Inszenierung vom Mittelalter bis zur Gegenwart*. In: Stephan, Inge, Sigrid Weigel: *Literatur – Kultur – Geschlecht. Studien zur Literatur- und Kulturgeschichte*. Kleine Reihe, Band 18. Köln/Weimar/Wien: Böhlau Verlag, 11–35.
- Underwood, Ted** (2019): *Distant Horizons*. Chicago: The University of Chicago Press.
- Weitin, Thomas / Katharina Herget** (2017): "Falkentopics. Über einige Probleme beim Topic Modeling literarischer Texte." In: Hartmut Bleumer, Rita Franceschini, Staphan Habscheid, Constanze Spieß und Niels Weber (Hrsg.): *Scalable Reading*. Zeitschrift für Literaturwissenschaft und Linguistik. 1/2017 (47), 1–20.