

Friedrich Michael Dimpel

Narratologische Textauszeichnung in Märe und Novelle

Abstract: In this paper, I suggest a system of annotation designed for the use in Computational Narratology. The levels of annotating I use in my study are as follows: space, time, focalisation, reported speech, narrator's speech, relation of the speech of a character/of the narrator to a certain character, evaluative remarks, negation, figurative speech, ambiguity. Alongside with the tagset, the guidelines for annotation are being refined in order to enable different persons to achieve consistent results (inter-annotator-agreement). In doing so, narratological models will be systematically reassessed and refined. The aim of this project is therefore to annotate about 100 short narratives. The corpus will render possible a great variety of follow-up aims and objectives. It permits systematic access to corpus segments annotated in the same way. Using ›Studentenabenteuer A‹ and ›Decamerone‹ IX,6, I discuss a number of details and problems in narratological annotation. Finally, I present some examples of quantitative analysis using among others MTLD and Social Network Analysis.

1 Quantitative Limitierungen und Computational Narratology

Digitale Korpora können aufgrund der Fortschritte in den Digital Humanities mit zahlreichen quantitativen Analyseverfahren untersucht werden:¹ Im Rahmen der Stilometrie beschäftigen sich Computerphilologen etwa mit Fragen der Autor-

¹ Vgl. exemplarisch Fotis Jannidis: »Methoden der computergestützten Textanalyse«, in: *Methoden der literatur- und kulturwissenschaftlichen Textanalyse. Ansätze – Grundlagen – Modellanalysen*, hg. v. Vera Nünning und Ansgar Nünning. Stuttgart 2010, S. 109–132. Zudem erfolgt auch die Textedition zunehmend mehr mit innovativen digitalen Methoden, vgl. etwa Malte Rehbein: »Vom Nutzen digitaler Editionen – das Göttinger ›kundige bok‹«, in: *Bibliothek und Wissenschaft* 42 (2009), S. 7–28.

schaftsattribution oder der Werkchronologie sowie mit Fragen nach den Spezifika von Gattungen oder Epochen.² Neben der Stilometrie erfährt derzeit die Computational Narratology intensive Aufmerksamkeit.³ Die vorliegende Projektstudie ist im Kernbereich der Computational Narratology situiert, es handelt sich dabei um Grundlagenforschung, die auch weitere vielfältige stilometrische Untersuchungen ermöglichen soll.

Meist ist ein vollständiger Text oder ein Abschnitt – beispielsweise ein Kapitel eines Romans – Gegenstand von digitalen Studien. Die Möglichkeiten für vollautomatische Analysen enden jedoch dort, wo spezifische Textebenen in den Blick zu nehmen wären; etwa dann, wenn Differenzen zwischen Erzählerrede und Figurenrede examiniert werden sollen. Derzeit ist es nicht möglich, verschiedene Ausprägungen von gesprochener oder gedachter Figurenrede (direkte oder indirekte Redewiedergabe) automatisch zu ermitteln⁴ und sie der jeweiligen Figur zuzuweisen – auch deshalb, weil die gleiche Figur mit verschiedenen Bezeichnungen aufgerufen werden kann (bspw.: »Orgeluse« oder »diu herzoginne«

2 Vgl. etwa John F. Burrows: »Questions of Authorship. Attribution and Beyond. A Lecture Delivered on the Occasion of the Roberto Busa Award«, in: *Computers and the Humanities* 37 (2003), S. 5–32; Hugh Craig: »Stylistic Analysis and Authorship Studies«, in: *A Companion to Digital Humanities*, hg. v. Susan Schreibman, Ray Siemens und John Unsworth. Oxford 2004, S. 273–288, <http://www.digitalhumanities.org/companion> (24. Januar 2015); Andreas Büttner, Friedrich Michael Dimpel, Stefan Evert, Fotis Jannidis, Steffen Pielström, Thomas Proisl, Isabella Reger, Christof Schöch und Thorsten Vitt: »Delta« in der stilometrischen Autorschaftsattribution«, in: *Konferenzabstracts DHd 2016. Modellierung – Vernetzung – Visualisierung. Die Digital Humanities als fächerübergreifendes Forschungsparadigma*, hg. v. Elisabeth Burr. Leipzig 2016, S. 61–74, <http://dhd2016.de> (08. September 2016).

3 Vgl. etwa das Exposé von Jan Christoph Meister, Fotis Jannidis und Christof Schöch: »Elemente einer Roadmap für das Forschungsfeld »Computational Narratology«« zum Workshop auf der DHd-Tagung 2015; online unter https://www.conftool.pro/dhd2015/index.php?page=browseSessions&form_session=5&presentations=show (27. Januar 2015); zum Begriff vgl. etwa Inderjeet Mani: »Computational Narratology«, in: *The living handbook of narratology*, hg. v. Peter Hühn, Jan Christoph Meister, John Pier und Wolf Schmid. Hamburg 2013, S. 1–12, <http://www.lhn.uni-hamburg.de/article/computational-narratology> (24. Januar 2015).

4 Annelen Brunner: *Automatische Erkennung von Redewiedergabe. Ein Beitrag zur quantitativen Narratologie*. Berlin, Boston 2015 (Narratologia 47), hat ein Verfahren vorgestellt, das direkte, indirekte und erlebte Rede in nhd. Texten automatisch analysiert. Dabei setzt Brunner sowohl auf regelbasierte computerlinguistische Verfahren als auch auf Maschinelles Lernen. Die besten Ergebnisse werden bei direkter Rede mit einer Fehlerquote von 13% erzielt (F1-Wert; »precision« und »recall« kombiniert). So beeindruckend diese Ergebnisse aus computerlinguistischer Sicht sind, setzen narratologische Auswertungen doch eine geringere Fehlerquote voraus. Für frühere Sprachstufen ist noch keine automatische Analyse von Figuren- und Gedankenrede in Sicht.

im ›Parzival‹), die gleiche Bezeichnung kann auf verschiedene Figuren referieren (»diu herzoginne« kann Orgeluse oder Jeschute bezeichnen).⁵

Wenn jedoch verschiedene Textschichten manuell annotiert werden, können solche Untersuchungen äußerst ertragreich sein. Ein Beleg dafür sind die Studien von Burrows zu Jane Austen: Mithilfe statistischer Analysen zu hochfrequenten Wörtern kann Burrows nachweisen, dass die gesprochene Figurenrede der jeweiligen Figuren erheblich variiert; Figurenrede unterscheidet sich zudem auch signifikant von der Erzählerrede. Die Sonderstellung einzelner Figuren bei Austen lässt sich auch an Besonderheiten des Figuren-Idiolekts nachweisen.⁶

Die Arbeiten von Meister zu ›Handlung‹, ›Ereignis‹ und ›Zeit‹ beruhen im ersten Schritt auf einer Textannotation, die in einem zweiten Schritt ausgewertet wird.⁷ Mit CATMA wurde in Hamburg eine webbasierte Arbeitsumgebung für kollaboratives Annotieren etabliert, in der Zeitstrukturen abgebildet werden. Diese Daten gehen in das Folgeprojekt HeureCléa als Trainingskorpus für Maschinelles Lernen ein; dort werden dem Anwender automatisiert Annotationsvorschläge zu Texten, die noch nicht annotiert sind, unterbreitet, die er akzeptieren oder abändern kann.⁸

5 Fotis Jannidis hat auf der DHD-Jahrestagung 2015 in Graz berichtet, dass neue Algorithmen bei der Named-Entity-Recognition nun zu einer Erkennungsrate >85% bei nhd. Texten führen (F1-Wert).

6 John F. Burrows: *Computation into Criticism. A Study of Jane Austen's Novels and an Experiment in Method*. Oxford 1987; vgl. zur Textauszeichnung S. 6–10. Um eine statistische Auswertung zu ermöglichen, hat Burrows zunächst das Korpus aufbereitet und Figurenrede sowie erlebte Rede ausgezeichnet. Wie wichtig eine Auszeichnung etwa von verschiedenen Formen der Redewiedergabe wäre, betonen Fotis Jannidis, Gerhard Lauer und Andrea Rapp: »Hohe Romane und blaue Bibliotheken. Zum Forschungsprogramm einer computergestützten Buch- und Narratologiegeschichte des Romans in Deutschland (1500–1900)«, in: *Literatur und Literaturwissenschaft auf dem Weg zu den neuen Medien*, hg. v. Lucas Marco Gisi, Jan Loop und Michael Stolz. germanistik.ch 2006, S. 2 und S. 12, http://www.germanistik.ch/publikation.php?id=Hohe_Romane_und_blaue_Bibliotheken (24. Januar 2015).

7 Vgl. insbes. Jan Christoph Meister: *Computing Action. A Narratological Approach*. Berlin, New York 2003 (Narratologia 2), S. 199–255; Jan Christoph Meister: »Computational Narratology: oder: Kann man das Erzählen berechenbar machen?«, in: *Erzählen, Archivieren, Beschreiben*, hg. v. Corinna Müller und Irina Scheidgen. Marburg 2007 (Schriftenreihe der Gesellschaft für Medienwissenschaft 5), S. 19–39, hier, S. 29; Jan Christoph Meister: »The Temporality Effect. Towards a Process Model of Narrative Time Construction«, in: *Time. From Concept to Narrative Construct: A Reader*, hg. v. dems. und Wilhelm Schernus. Berlin, Boston 2011 (Narratologia 29), S. 171–216.

8 Zu CATMA und heureCLÉA vgl. Evelyn Gius und Janina Jacke: »Informatik und Hermeneutik. Zum Mehrwert interdisziplinärer Textanalyse«, in: *Grenzen und Möglichkeiten der Digital Humanities*, hg. v. Constanze Baum und Thomas Stäcker. 2014 (ZfdG-Sonderband 1), www.ZfdG.de

Auch wenn die »hohe Präzision der strukturalistischen Narratologie« computerphilologische Zugriffe zu ermöglichen scheint, sind digitale Studien im Feld der Narratologie bislang relativ selten.⁹ Überwiegend stehen bei quantitativen Studien lexikalische Textmerkmale (etwa bei Burrows' Delta) im Mittelpunkt wie auch bei Analysen zu Funktionswörtern, Hapax Legomena, Kollokationen und n-Grammen.¹⁰ Der Aufwand, eigens für eine spezifische Fragestellung und für eine spezifische Methode ein Korpus manuell aufzubereiten, ist enorm; er kann oft aus ökonomischen Gründen nicht geleistet werden. Studien, die auf narratologischen Unterscheidungen oder Modellen wie Figurenrede/Erzählerrede beruhen, haben, wie die Studien von Burrows und Meister illustrieren, jedoch bislang eine solche manuelle Korpusaufbereitung zur Voraussetzung.

Vorgeschlagen wird daher, ein diachron angelegtes Korpus von Kurzerzählungen systematisch narratologisch auszuzeichnen – jedoch nicht mit Blick auf eine einzelne Fragestellung und auf eine konkrete Analysemethode, vielmehr wird hier eine generische Konzeption verfolgt, bei der Nachnutzbarkeit im Zentrum steht: Die Annotation erfolgt im Sinne einer Grundlagenarbeit für multiple Fragestellungen und multiple Methoden, die vielfältige Anschlussforschungen ermöglichen. Das Korpus erlaubt die Bearbeitung von Fragestellungen, die derzeit noch nicht absehbar sind, zudem können die Korpusdaten auch neue Fragestellungen generieren. Ein solches Korpus ist bislang ebenso wenig vorhanden wie eine narratologische Annotationskonzeption jenseits von Redewiedergabe

(24. Januar 2015) sowie Thomas Bögel, Michael Gertz, Evelyn Gius, Janina Jacke, Jan Christoph Meister, Marco Petris und Jannik Strötgen: »Gleiche Textdaten, unterschiedliche Erkenntnisziele? Zum Potential vermeintlich widersprüchlicher Zugänge zu Textanalyse«, in: *DHd 2015: Digital Humanities im deutschsprachigen Raum*. Graz 2015, <http://gams.uni-graz.at/archive/objects/o:dhd2015.nachlese.vortragsfolien/methods/sdef:HTML/get> (06. Juni 2017).

⁹ Bei der praktischen Anwendung narratologischer Explikationen ergeben sich gerade im Rahmen einer Textannotation rasch Probleme. Zitat: Fotis Jannidis: »Computerphilologie«, in: *Handbuch Literaturwissenschaft. Band 2: Methoden und Theorien*, hg. v. Thomas Anz. Stuttgart, Weimar 2007, S. 27–40, hier S. 35; vgl. zuletzt jedoch Evelyn Gius: *Erzählen über Konflikte. Ein Beitrag zur digitalen Narratologie*. Berlin, Boston 2015 (Narratologia 46).

¹⁰ Vgl. Fotis Jannidis: »Computerphilologie«, S. 35.

und Zeit.¹¹ Ein geeignetes XML-Tagset muss eigens entwickelt werden.¹² TEI-Kompatibilität wird angestrebt.¹³

2 Auszeichnen

1) *Figur am Fokussort*: <FigurFokusort>

<FigurFokusort>¹⁴ verzeichnet die Figuren, die sich an dem Ort aufhalten, von dem erzählt wird. Attribute sind @Bezeichnung (Figurennamen), @Figurengruppe (fakultativ), @level (fakultativ; Grad der Verschachtelung, falls ein Bericht über ein Geschehen an einem anderen Ort in eine Figurenrede/Gedankenrede eingebettet ist),¹⁵ @non-fact (fakultativ, falls eine Figur oder ein Erzähler auf Figuren an anderem Ort Bezug nimmt, die sich nicht faktisch dort befinden). Beispiel: <FigurFokusort Bezeichnung="Ritter; Tochter" Figurengruppe="">Nû lac ein juncvrouwelîn, edel, schoene unde fîn, [...] </FigurFokusort> (>Häslein< 61–186). Solche Informationen können eine Grundlage sein für Studien mit raumtheoretischen Fragestellungen – zumal in Verbund mit dem Element <BewegungLokal>, mit dem etwa Untersuchungen zu Transgressionen und statischen bzw. dynamischen Figuren möglich werden:

2) *Lokale Bewegungen*: <BewegungLokal>

Bei <BewegungLokal> geht es um lokale Bewegungen: Wechsel des Erzählortes;

¹¹ Evelyn Gius und Janina Jacke: *Zur Annotation narratologischer Kategorien der Zeit. Guidelines zur Nutzung des CATMA-Tagsets*. Hamburg 2014, <http://heureclea.de/publications/guidelines.pdf> (09. März 2015); Brunner: *Redewiedergabe*.

¹² Zu narratologischen Desideraten von TEI-P5 vgl. Fotis Jannidis: »TEI in a Crystal Ball«, in: *Literary and Linguistic Computing* 24 (2009), S. 253–265, hier S. 261f.

¹³ Eine TEI-Integration ist auch via ODD möglich; vgl. <http://www.tei-c.org/Guidelines/Customization/odds.xml> (13. März 2015).

¹⁴ Deutsche Bezeichnungen stehen bei neuen Tags, englische bei Tags, die aus TEI-P5 oder aus Annelen Brunner: *Redewiedergabe*, übernommen werden. Ich stelle dieses Projekt ebenfalls vor in »Novellenschätze narratologisch auszeichnen und analysieren am Beispiel Victor von Schefels ›Hugideo‹ und der sozialen Netzwerkanalyse«, in: *LiLi* 47 (2017), S. 87–108, DOI:10.1007/s41244-017-0044-8, sowie in »Narratologische Textauszeichnung in Märe und Novelle. Mit Annotationsbeispielen und exemplarischer Auswertung von ›Sperber‹ und ›Häslein‹ durch MTLD und Sozialer Netzwerkanalyse«, in: *ZfdG*, vsl. 2017. Da es sich bei Tagset und Workflow um den gleichen Gegenstand handelt, bestehen in diesen Passagen große und teils wörtliche Übereinstimmungen mit diesen Beiträgen.

¹⁵ Zu @level und @non-fact vgl. Brunner: *Redewiedergabe*, S. 72–76; hier zur Redewiedergabe.

Fortbewegung, die einer Figur folgt. Attribute sind @Typ (Fokuswechsel | Fortbewegung), @Bezeichnung (fakultativ), @level, @non-fact (fakultativ, s. o.). Beispiel: <BewegungLokal Typ="Fortbewegung" Bezeichnung="Ritter">*alsus reit er alzehant* [...].</BewegungLokal> (›Häslein‹ 56–60).

3) Abweichungen vom linearen Zeitverlauf <Zeit>

Bei <Zeit> werden Abweichungen vom linearen Zeitverlauf erfasst. Attribute sind @Typ (Prolepse_Beginn | Prolepse_Ende | Analepse_Beginn | Analepse_Ende | Ellipse | Pause), @level (fakultativ, s. o.). Beispiel: <Zeit Typ="Pause">*Tribe ich die zît vergebene hin*, [...].</Zeit> (›Häslein‹ 1–20). Eine Annotation von temporalen Phänomenen findet sich etwa im CATMA-Kontext; hier wird eine schlankere Modellierung verwendet.¹⁶

4) Fokalisierung <Fokalisiert>

Welche Figur ist intern oder extern fokalisiert: <Fokalisiert>. Attribute: @Typ: (intern | extern | Paralipse | Paralepse), @Bezeichnung. Beispiel: <Fokalisiert Bezeichnung="Tochter" Typ="intern">*und dô diu juncvrouwe zart der sumerzîte ginret wart*, [...] *daz lie diu juncvrouwe âne haz*.</Fokalisiert> (›Häslein‹ 159–166). Auch in narratologischen Darstellungen finden sich metaphorische Begriffe; etwa »Mitsicht« bei Genette.¹⁷ Grundlage für eine genauere narratologische Modellierung und für die Entwicklung der Annotationsrichtlinien sind die Fokalisierungskonzepte bei Hübner und Dimpel.¹⁸

5) Figurenrede und Gedankenrede: <Redewiedergabe>

Attribute sind @Typ (›Direkte Rede« | ›Indirekte/erzählte Rede« | ›Erlebte Rede« | ›Direkte Gedankenrede« | ›Bewusstseinsdarstellung« | ›Erlebte Gedankenrede«), @Bezeichnung, @level, @non-fact (fakultativ, s. o.),¹⁹ @narr (fakultativ):

¹⁶ Vgl. Gius, Jacke: *Guidelines*.

¹⁷ Gérard Genette: »Diskurs der Erzählung – ein methodologischer Versuch«, in: ders: *Die Erzählung. Dt. Übersetzung von ›Discours de récit‹ und ›Nouvelle Discours de récit‹*. Aus dem Französischen v. Andreas Knop, hg. v. Jochen Vogt. München [1972] 1998, S. 134.

¹⁸ Gert Hübner: *Erzählform im höfischen Roman. Studien zur Fokalisierung im ›Eneas‹, im ›Iwein‹ und im ›Tristan‹*. Tübingen 2003 (Bibliotheca Germanica 44), S. 25–63; Friedrich Michael Dimpel: *Die Zofe im Fokus. Perspektivierung und Sympathiesteuerung durch Nebenfiguren vom Typus der Confidante in der höfischen Epik des hohen Mittelalters*. Berlin 2011 (Philologische Studien und Quellen 232), S. 17–38.

¹⁹ Brunner: *Redewiedergabe*, S. 72f. Die Funktion von @ambig und @prag wird hier mittels @cert und @cert_comment realisiert.

Grenzbereich zwischen Sprache und Handlung,²⁰ @border [fakultativ): Grenzfälle von Rede (beten, zählen, singen etc.). Bei einem Wechsel der Erzählebene (Binnenerzählungen) wird weiterhin <embedded> verwendet.²¹ Beispiel<Redewiedergabe Typ="Bewusstseinsdarstellung" Bezeichnung="Ritter">des vröute sich der ritter dô</Redewiedergabe> (Häslein 39f.). Das Auszeichnungsmodell von Brunner kann mit einigen Vereinfachungen übernommen werden, da hier nicht die gleiche Auszeichnungstiefe notwendig ist.

6) Erzählerrede: <Erzählerrede>

Bei <Erzählerrede> werden als die Attribute erfasst: @Typ: (»Bericht Figurenaktivität« | »Sonstiger Bericht« | »Explanative Erzähleräußerung« | »Erzählerreflexion über erzählte Welt« | »Erzählerreflexion allgemein« | »Erzählerreflexion metanarrativ«); @Bezeichnung,²² @non-fact (fakultativ, s. o.). Beispiel: <Erzählerrede Typ="Bericht Figurenaktivität" Bezeichnung ="Ritter">alsus reit er alzehant gegen eime dorfe hin, </Erzählerrede> (Häslein 56f.). Die narrative Erzählerrede jenseits von Redewiedergabe wird differenziert in »Figurenaktivität« und »Sonstiger Bericht«. Die weitere Taxonomie folgt der Unterscheidung von Nünning;²³ zudem lässt sich hier auf die Annotationsrichtlinien von Gius / Jacke aufbauen.²⁴

7) Bezug von <Redewiedergabe> oder <Erzählerrede> auf Figur: <Figurenbezug>

Bei <Figurenbezug> geht es um die Zuordnung des Inhalts einer <Redewiedergabe> oder <Erzählerrede> zu einer Figur. Attribute sind @Unmittelbar (Bezeichnung der Figur bei unmittelbarem Bezug), @Mittelbar (Bezeichnung bei mittelbarem Bezug), @Vorläufig_implizit (bei Bezug, der auf Kontiguität, Äquivalenz etc. beruht, der jedoch später explizit oder stark offensichtlich wird) und @Impli-

20 @narr und @border nach Brunner: *Redewiedergabe*, S. 80f. sowie 82f.; @border hier jedoch ohne weitere Differenzierung mittels Werteliste (S. 87–89). Mit Hilfe von @non-fact, @border und @narr kann man Grenzbereiche als Grenzbereiche einer bestimmten Sorte auszeichnen.

21 Brunner: *Redewiedergabe*, S. 91f.

22 Bei Figurenaktivität erhält @Bezeichnung den Namen der Figur; sonst die Bezeichnung der Erzählinstanz.

23 Ansgar Nünning: Grundzüge eines kommunikationstheoretischen Modells der erzählerischen Vermittlung. Die Funktion der Erzählinstanz in den Romanen George Eliots. Trier 1989 (Horizonte 2), S. 51f.

24 Gius, Jacke: *Guidelines*, S. 10–14. Die feingliedrige Unterscheidung von <information_retrieval/writing_process>, <self_description>, <organization_of_discourse> und <address_to_the_reader> wird zunächst zu <Erzählerreflexion metanarrativ> zusammengefasst.

zit (bei Bezug, der auf Kontiguität, Äquivalenz etc. beruht); @non-fact (fakultativ; s. o.). Beispiel: <Figurenbezug Unmittelbar="Mutter; Tochter" Mittelbar="Ritter">des wart ir gelwez hâr zerrouft; darnâch ir liechten wangen begunde diu muoter zwanzen [...]. </Figurenbezug> (Häslein 198–207).

Hier geht es um den Gegenstand einer Äußerung: Figuren oder Erzähler können über eine andere Figur sprechen oder zu einer Figur; Figurenhandlungen können auf Figuren bezogen sein. Notiert wird, auf welche Figur Bezug genommen wird. Wenn etwa alles Wollen und Streben des Protagonisten darauf ausgerichtet ist, die Liebe einer Dame zu erwerben, so ist ein solcher Text häufig aus der Perspektive des Protagonisten erzählt. Häufig hat die Dame selbst nur wenig Anteil an Erzähler- und Figurenrede, obwohl es stets um sie geht. Solche Daten sind bei kulturwissenschaftlichen Analysen interessant: Die Frage »Über wen wird gesprochen?« kann hier wichtiger werden als Genettes Frage »Wer spricht?«. Neben unmittelbaren Bezügen gibt es auch mittelbare: Figuren wenden sich mitunter direkt an andere Figuren (Brautvater), um indirekt das eigentliche Ziel (Gewinnung der Braut) zu erreichen.

8) Evaluative Äußerungen: <Wertung>

Bei <Wertung> wird Erzähler- oder Figurenrede mit evaluativen Äußerungen zu einer Figur erfasst. Attribute sind: @Wertende (Bezeichnung des Erzählers oder der evaluierenden Figur), @Gewertete (Bezeichnung), @Typ (positiv explizit | positiv implizit | negativ explizit | negativ implizit), @level, @non-fact (fakultativ; s. o.). Beispiel: *dô wart dem ritter offenbâr einer juncvrouwen lîp gelobet vûr ein êlich wîp*, <Wertung Wertende="Erzähler" Gewertete="Verlobte" Typ="positiv explizit">diu was schoene unde kluoc,</Wertung> (Häslein 325). Erfasst werden explizite und implizite Bewertungshandlungen nach Winko.²⁵

9) Negation des Aussagegehalts: <Negation>

Zu <Negation> lauten die Attribute @Typ (»einfach« | »doppelt positiv« | »doppelt negativ« | »dreifach«). Beispiel <Negation Typ="doppelt negativ">ich engelebete nie sô lieben tac.</Negation> (Häslein 82). In quantitativen Studien kann es von Interesse sein, ob ein Wort wie »Treue« in Negation steht. Da in mittelhochdeutschen Texten eine Negation mit zwei Negationswörtern realisiert werden kann, wurde der Wert »doppelt negativ« eingeführt.

²⁵ Simone Winko: Wertungen und Werte in Texten. Axiologische Grundlagen und literaturwissenschaftliches Rekonstruktionsverfahren. Braunschweig 1991 (Konzeption Empirische Literaturwissenschaft 11), S. 134f.

10) *Uneigentliche Rede*: <UneigentlicheRede>

Mit <UneigentlicheRede> werden Abweichungen von literalem Sinn (metaphorische, ironische Rede etc.) annotiert. Attribut: @Bezeichnung. Beispiel: <UneigentlicheRede Bezeichnung="Erzähler"> *dô riet ime sîn herze stæte*, </UneigentlicheRede> (›Häslein‹ 44).

11) *Allgemeine Attribute*

@Datum und @Bearbeiter werden automatisch ergänzt. In @Regel wird die einschlägige Annotationsregel festgehalten. In @comment können Annotationsentscheidungen ggf. begründet werden. Zudem werden alle Figurenreferenzen wie Namen oder Pronomina annotiert. Das TEI-P5-Attribut @cert ist als fakultatives Attribut bei allen Elementen vorgesehen – z. B. als Attribut zu <FigurFokusort>, wenn nicht erzählt wurde, wann welche Figur den Raum verlassen hat.

3 Ambiguität und Annotationsrichtlinien

Textuelle Ambiguität und ambivalente Rezeptionsdispositionen in Texten haben Schriftgelehrte bereits in der Antike veranlasst, von einem mehrfachen Schriftsinn auszugehen. Vertreter der Empirischen Literaturwissenschaft haben die Polyvalenzkonvention zu einem wesentlichen Merkmal literarischer Texte erklärt. Allerdings ist es ein erheblicher Unterschied, ob es um mehrdeutige Sinnangebote oder um mehrere Zuordnungsmöglichkeiten einer mikrostrukturellen Passage zu einer narratologischen Kategorie geht: Auch wenn kaum je ein ›Inter-Annotator-Agreement‹ hinsichtlich des Sinngehalts eines Textes zu erwarten ist, sind die Probleme bei den hier skizzierten Kategorien deutlich geringer. Bei dem Versuch, narratologische Kategorien zu formalisieren, handelt es sich freilich nicht um eine voraussetzungsfreie Textbeschreibung;²⁶ Bereits der Umstand, dass auf die Frage »Was ist ein Wort?« verschiedene Antworten möglich sind, zeigt, dass auch konventionelle Literaturanalysen stets auf Vorannahmen beruhen, auch wenn diese Vorannahmen häufig nicht explizit mitgeführt werden. Insofern unterscheidet sich ein Annotationsprojekt von konventionellen Analysen u. a. dadurch, dass es hier notwendig wird, viele Vorannahmen und Festlegungen explizit festzuhalten.

²⁶ Vgl. Tom Kindt und Hans-Harald Müller: »Wieviel Interpretation enthalten Beschreibungen? Überlegungen zu einer umstrittenen Unterscheidung am Beispiel der Narratologie«, in: *Regeln der Bedeutung. Zur Theorie der Bedeutung literarischer Texte*, hg. v. Fotis Jannidis, Gerhard Lauder, Matías Martínez und Simone Winko. Berlin, New York 2003 (Revisionen 1), S. 286–304.

Dass Texte polyvalent und ambig sein können, ist eine zentrale Herausforderung für Annotationsprojekte:²⁷ Eine Annotation geht mit einer Disambiguierung einher: es werden eindeutige Festlegungen getroffen;²⁸ auch die Dokumentation von Ambiguität ist eine eindeutige Festlegung. Weil in der Regel interpretative Annahmen in Annotationsentscheidungen eingehen, plädiert Meister im Rahmen des Crowd-Sourcing-Projektes dafür, dass die Textauszeichnung hermeneutisch produktive Ambiguität und Widersprüchlichkeit abbilden soll.²⁹ Während etwa Meister, Gius, Jacke und Brunner mit einer mehrfachen Auszeichnung auf diese Problemlage reagieren und mehrere Lesarten abbilden,³⁰ wird hier angestrebt, nur eine Auszeichnung vorzunehmen, die gegebenenfalls als uneindeutig (@cert) markiert ist.³¹ Ein einfaches Markup hat den Vorteil, dass der Code gut lesbar und übersichtlich bleiben kann. Darüber hinaus wird die Auswertung im Rahmen der Anschlussforschung durch schlankere Tools erleichtert – etwa bei der Sample-Erstellung für stilometrische Untersuchungen, bei denen unsichere Auszeichnungen ausgeblendet werden können. Dagegen geht das Abbilden von alternativen Auszeichnungsvarianten mit dem »disadvantage of cumbersome processing«³² einher und soll daher möglichst vermieden werden. Angestrebt wird eine Auszeichnung zunächst in separaten Dateien je Element, die sich auto-

27 Vgl. Jan Christoph Meister: »Crowd sourcing ›true meaning‹. A collaborative markup approach to textual interpretation«, in: *Collaborative Research in the Digital Humanities. Festschrift for Harold Short*, hg. v. Willard McCarty und Marilyn Deegan. Surrey 2012, S. 105–122, hier S. 105–117. Fotis Jannidis: »Polyvalenz – Konvention – Autonomie«, in: *Regeln der Bedeutung. Zur Theorie der Bedeutung literarischer Texte*, hg. v. ders., Gerhard Lauer, Matías Martínez und Simone Winko. Berlin, New York 2003 (Revisionen 1), S. 305–328, hier S. 323–327, zeigt, wie man der Polyvalenzfrage mit Hilfe des Konzeptes der Offensichtlichkeit begegnen kann.

28 Vgl. Jerome McGann: »Marking Texts of Many Dimensions«, in: *A Companion to Digital Humanities*, hg. v. Susan Schreibman, Ray Siemens und John Unsworth. Oxford 2004, S. 198–217, <http://www.digitalhumanities.org/companion> (24. Januar 2015).

29 Meister: »Crowd sourcing«, S. 121.

30 Vgl. Jan Christoph Meister: »Tagging Time in Prolog: The Temporality Effect Project«, in: *Literary and Linguistic Computing* 20 (2005), S. 107–124, hier S. 117f.; Gius, Jacke: »Informatik und Hermeneutik«, Brunner: *Redewiedergabe*, S. 63–66.

31 In @cert_comment wird die Art der Ambiguität bzw. der Grund für die unsichere Zuschreibung dokumentiert. Falls es sich entgegen der Zielsetzung nicht als praktikabel erweist, auf eine parallele Auszeichnung von mehreren Auszeichnungsoptionen zu verzichten, können die mit @cert ausgezeichneten Passagen rasch aufgesucht und alternativ ausgezeichnet werden; damit würde die Auszeichnung von mehreren separaten Dateien je Element oder eine Umstellung auf ein stand-off-markup nötig, wie es in CATMA etabliert ist.

32 <http://www.tei-c.org/release/doc/tei-p5-doc/en/html/NH.html>, hier Abschn. 20.2 (13. März 2015).

matisiert in eine fragmentierte Repräsentation der Daten überführen lassen – dieser Weg dient einerseits der Lesbarkeit des Codes; andererseits wird damit das Problem der Überlappung umschifft.³³

Ziel ist ein einheitlich annotiertes Korpus, das auf einem ›Inter-Annotater-Agreement‹ beruht. Dazu sind Annotationsregeln nötig: So, wie man in Editionsprojekten Kollationsregeln festlegen muss, werden hier Annotationsregeln erarbeitet, die regelmäßig weiterentwickelt werden. Solche Regeln werden auch deshalb benötigt, weil narratologische Definitionen häufig nicht so exakt formuliert sind, dass sich Textphänomene damit eindeutig einordnen lassen. Zu erwarten ist, dass diese Annotationsrichtlinien über einen langen Zeitraum weiterentwickelt werden müssen. Darin liegt jedoch ein Vorteil, denn im Prozess der Entwicklung der Annotationsrichtlinien erwarte ich Erkenntnisse über mögliche narratologische Modellierungen und über die Probleme, die mit der Anwendung von narratologischen Modellierungen auf historische Texte einhergehen. Modellkreation und Modellkorrektur sind, so Willard McCarty, bei computerphilologischen Studien nicht nur ein Standardprocedere; vielmehr trägt die Modellmodifizierung zentral zum Erkenntnisgewinn bei.³⁴

Insoweit bei Entscheidungen nicht auf interpretative Momente verzichtet werden kann, sollen solche Interpretationsentscheidungen möglichst bei der Formulierung der Annotationsrichtlinien stattfinden und nicht bei der konkreten Textauszeichnung. Die einzelne Annotationsentscheidung kann so in Übereinstimmung mit den Richtlinien erfolgen. Falls dies gelingt, kann eine Teilstrecke der Annotation als empirisch beschrieben werden: die Anwendung der Richtlinie auf ein Textsegment.³⁵ Dieses Verfahren führt dazu, dass in die Annotationsricht-

³³ Neben einer Auszeichnung mit leeren Elementen (milestones) und stand-off-markup kann das Problem einer überlappenden Annotation durch separates oder fragmentiertes Markup gelöst werden; vgl. den Abschnitt »20 Non-hierarchical Structures« der TEI-Guidelines; zur Separierung hier 20.1 und zur Fragmentierung 20.3, <http://www.tei-c.org/release/doc/tei-p5-doc/en/html/NH.html> (13. März 2015).

³⁴ Willard McCarty: »Modeling: A Study in Words and Meanings«, in: *A Companion to Digital Humanities*, hg. v. Susan Schreibman, Ray Siemens und John Unsworth. Oxford 2004, S. 254–270, hier S. 258, <http://www.digitalhumanities.org/companion> (24. Januar 2015). Ausführlich zum Erkenntnisgewinn durch »modeling« Willard McCarty: *Humanities Computing*. London, New York 2005, insbes. S. 23–72.

³⁵ Vgl. zu empirischen Teilstrecken Friedrich Michael Dimpel: »Der Computerphilologe als Interpret – ein Teilzeit-Empiriker?«, in: *Literatur interpretieren: Interdisziplinäre Beiträge zur Theorie und Praxis*, hg. v. Jan Borkowski, Stefan Descher, Felicitas Ferder und Philipp Heine. Münster 2015, S. 339–359.

linien interpretative Entscheidungen eingehen, die einerseits auf narratologischen Forschungsergebnissen beruhen, die jedoch andererseits auch Spezifikationen und Setzungen enthalten, die sich in der Annotationspraxis ergeben. Diese Entscheidungen sind in den Annotationsrichtlinien dokumentiert und somit kontrollierbar; sie dienen der Einheitlichkeit im Annotationsvorgang. Ein zentrales Kriterium ist Konsistenz: Eine einzige »richtige« Annotation kann es ebenso wenig geben, wie es etwa eine einzige richtige literaturwissenschaftliche Analyse etwa zur Fokalisierung einer Textpassage geben kann: Wenn man sich für Genettes Modell entscheidet, wird man zu anderen Ergebnissen kommen als mit Bals Fokalisierungsmodell.

Angaben zum Bearbeiter, zum Bearbeitungsdatum und zur angewendeten Annotationsrichtlinie werden im XML-Code dokumentiert. Wenn im Annotationsvorgang auffällt, dass die narratologische Modellierung und die Annotationsrichtlinien modifiziert werden müssen, kann mit Hilfe dieser Dokumentation rasch auf Fälle zugegriffen werden, die nach der Regelkorrektur neu zu annotieren sind. Zudem kann das bislang annotierte Material auch als Fallsammlung dienen, an der sich ein Annotator orientieren und eine konkordante Auszeichnung vornehmen kann.

4 Eine Beispielstudie: ›Studentenabenteuer A‹ und ›Decamerone‹ IX,6

Bevor es im Weiteren anhand des ›Studentenabenteuer A‹ (Mitte 13. Jahrhundert) und von ›Decamerone‹ IX,6 um Annotationsprobleme gehen soll, sei kurz der Inhalt der beiden Beispieltexte skizziert:³⁶ In beiden Texten geht es um das Motiv

36 Textausgaben: Wilhelm Stehmann: *Die mittelhochdeutsche Novelle vom Studentenabenteuer*. Berlin 1909 (Palaestra 67) und Giovanni di Boccaccio: *Das Dekameron*. Aus dem Italienischen übertragen von Albert Wesselski. Band III. Leipzig 41912. Da hier die Darstellung von methodischen Problemlagen im Fokus steht, verwende ich in der Konzeptionsphase vorläufig eine neu-hochdeutsche Übersetzung; eine quantitative Auswertung lässt freilich Aussagen über die Übersetzung und nur eingeschränkt über Boccaccios Text zu. Zu beiden Texten vgl. Hans-Joachim Ziegeler: »Boccaccio, Chaucer, Mären, Novellen: ›The Tale of the Cradle‹«, in: *Kleinere Erzählformen im Mittelalter. Paderborner Colloquium 1987*, hg. v. Klaus Grubmüller, Peter L. Johnson und Hans-Hugo Steinhoff. Paderborn 1988, S. 9–31; Hans-Joachim Ziegeler: »›Studentenabenteuer A und B‹ (FM 129 u. 107)«, in: *Kleinepik, Tierepik, Allegorie und Wissensliteratur*, hg. v. Fritz Peter Knapp. Berlin, Boston 2013 (GLMF VI), S. 100–110; weiterhin Klaus Grubmüller: *Die Ordnung, der Witz und das Chaos. Eine Geschichte der europäischen Novellistik im Mittelalter: Fabliau – Märe*

der verstellten Wiege: Ein junger Mann möchte mit einem Mädchen schlafen. Deshalb nimmt er gemeinsam mit einem Freund Quartier bei dem Vater des Mädchens, der nur über eine Schlafkammer verfügt. In dieser Kammer schläft der Vater in einem Bett mit seiner Frau, davor steht die Wiege mit dem Säugling. Die Tochter schläft in einem eigenen Bett, die beiden Freunde teilen sich ein drittes Bett. Der erste Freund schleicht sich nachts zu dem Mädchen und schläft mit ihm. Die Ehefrau geht etwas später kurz aus dem Raum, weil ein Geräusch zu hören war. In dieser Zeit verstellt der zweite Freund die Wiege so, dass sie vor seinem Bett steht – bei Boccaccio aus Versehen, im mittelhochdeutschen Text absichtlich. Die Ehefrau kehrt zurück, findet aber die Wiege nicht bei ihrem Bett; sie legt sich dann in das Bett, vor dem die Wiege steht, und sie ist durchaus angetan davon, dass es dort zum Beilager kommt – allerdings ist sie der Meinung, sie würde mit ihrem Mann schlafen.

Der Jüngling, der mit der Tochter geschlafen hat, kehrt danach zurück, er lässt sich aber ebenfalls durch die verstellte Wiege irritieren. Er legt sich zum Ehemann und berichtet über seinen Erfolg bei der Tochter – in der Meinung, er sei bei seinem Freund. Es kommt zum Streit und zu Handgreiflichkeiten, bevor es der Ehefrau gelingt, dem Mann zu suggerieren, es sei gar nichts passiert. Bei Boccaccio wird der erste Jüngling als Schlafwandler ausgegeben; im mittelhochdeutschen Text ist der Mann der Meinung, dass es die Mutter gewesen sei, die behauptet habe, mit ihrer Tochter intim gewesen zu sein, und dass sie völlig betrunken sei. Die Jünglinge stellen sich schlafend, so dass auch hier jede Konsequenz ausbleibt.

5 Annotationsbeispiele und Annotationsprobleme

A) Konkrete raumzeitliche Verortung bei <FigurFokusort>

Beim Element <FigurFokusort> sollen die Figuren benannt werden, die im Fokus stehen – prima facie ein einfaches Unterfangen. Doch *Decamerone* IX,6 beginnt denkbar allgemein:

– *Novelle*. Tübingen 2006, S. 127–131; Susanne Reichlin: »Zeitperspektiven. Das Beobachten von Providenz und Kontingenz in der ›Buhlschaft auf dem Baume««, in: *Kein Zufall. Konzeptionen von Kontingenz in der mittelalterlichen Literatur*, hg. v. Cornelia Herberichs und ders. Göttingen 2010 (Historische Semantik 13), hier S. 245–251.

In der Ebene des Mugnone lebte, es ist noch nicht lange her, ein biederer Mann, der den Reisenden für ihr Geld zu essen und zu trinken gab; und obwohl er arm war und nur ein kleines Häuschen hatte, beherbergte er doch dann und wann im Notfalle, zwar nicht jedermann, aber doch manchen Bekannten.

Es bleibt offen, wer genau in der Mugnone-Ebene anwesend war, zudem bleibt die Zeitangabe vage. Als Kriterium für das Element verwende ich eine konkrete zeitliche Verortung, hier wird das Tag <FigurFokusort> also nicht vergeben. Als man nach dem Abendessen zur Schlafkammer geht, heißt es:

<FigurFokusort Bezeichnung="UNDEF"> Nun hatte der Wirt nur eine sehr kleine Kammer, worein er, so gut es sich hatte tun lassen, drei Betten gestellt hatte, zwei an die eine Seite, das dritte gegenüber an die andere, so daß nur so wenig Raum geblieben war, daß man gerade noch dazwischen durchgehn konnte. Von diesen drei Betten ließ der Wirt das am wenigsten schlechte für die beiden Freunde herrichten,</FigurFokusort>

Hier ist unklar, wer tatsächlich damit betraut ist, die Schlafkammer herzurichten, und ob sich der Wirt und die Gäste ebenfalls bereits in der Schlafkammer befinden. Auch wenn die raumzeitliche Verortung eindeutig ist, ist es nicht möglich, bestimmte Figuren auszuzeichnen: Als Wert wird »UNDEF« vergeben.

B) Kleinteiliges Auszeichnen bei Redewiedergabe

Das Ziel, eine möglichst eindeutige Auszeichnung zu erreichen, macht oft ein sehr kleinteiliges Taggen nötig (hier Synopse aus <Redewiedergabe> und <Erzählerrede>):

<Redewiedergabe Typ="Bewusstseinsdarstellung" Bezeichnung="Jü2"> Als das Adriano,</Redewiedergabe><Erzaehlerrede Typ="Sonstiger Bericht" Bezeichnung="Binnenerzähler"> der noch nicht eingeschlafen war, </Erzaehlerrede> <Redewiedergabe Typ="Bewusstseinsdarstellung" Bezeichnung="Jü2"> merkte, </Redewiedergabe> <Erzaehlerrede Typ="Bericht Figurenaktivität" Bezeichnung="Jü2">empfang er sie freundlich </Erzaehlerrede><Redewiedergabe Typ="Bewusstseinsdarstellung" Bezeichnung="Jü2">und froh</Redewiedergabe><Erzaehlerrede Typ="Bericht Figurenaktivität" Bezeichnung="Jü2"> und lud ihr Schiffchen, </Erzaehlerrede><Erzaehlerrede Typ="Sonstiger Bericht" Bezeichnung="Binnenerzähler"> ohne ein Wort zu sagen, </Erzaehlerrede> <Erzaehlerrede Typ="Bericht Figurenaktivität" Bezeichnung="Jü2"> mehr als einmal voll </Erzaehlerrede><Redewiedergabe Typ="Bewusstseinsdarstellung" Bezeichnung="Mut">zu ihrem größten Vergnügen.</Redewiedergabe>

Hier muss mitunter für ein einziges Wort wie »merkte« ein eigenes Tag gesetzt werden – dass Adriano »noch nicht eingeschlafen war«, ist Erzählerbericht. Dass

der Jüngling-2 die Mutter freundlich empfängt, ist eine Aktivität; bei dem folgenden »und froh« handelt es sich dagegen um einen Blick ins Figurennere.³⁷ Ein kleinteiliges Auszeichnen erlaubt hier eine sichere Auszeichnung. Kleinteiliges Auszeichnen ist jedoch nicht bei allen Elementen die Politik der Wahl: Beim Element <Figurenbezug> bietet sich eine gröbere Segmentierung an, weil es oft um größere Bezüge auf Gründe und Ziele geht – weite Teile der Handlungen und der listigen Figurenrede der Jünglinge sind etwa auf das Ziel ausgerichtet, mit der Tochter zu schlafen.

C) Abgrenzungsprobleme bei Bewusstseinsdarstellung

Als der zweite Jüngling mit der Mutter schläft, umschreibt der Erzähler den Vorgang mit den Worten: *der schuolære vil gemeit / tete rehte als ein man / der vrouwen wol dienen kan, / und lie si niht müezic ligen.* (358–361) Das Wort *gemeit* kann aufgrund der semantischen Ambiguität als »wacker« oder »stattlich« übersetzt werden, jedoch auch als »freudig« oder »vergnügt«. Es kann also eine Schilderung »von außen« oder Bewusstseinsdarstellung vorliegen – oder beides zugleich. In den Annotationsrichtlinien lässt sich etwa festlegen, dass bei einer derartigen Ambiguität der Text sowohl in <Erzählerrede> als Erzählerbericht als auch in <Redewiedergabe> als Bewusstseinsdarstellung ausgezeichnet wird, obwohl sonst angestrebt wird, Textsegmente, die als Redewiedergabe annotiert sind, nicht auch zugleich als Erzählerrede zu annotieren. Alternativ wird geprüft, ob ein eigenes Tag für Formulierungen vergeben wird, die hinsichtlich innerer bzw. äußerer Figurenrede ambig sind.

6 Möglichkeiten der quantitativen Auswertung

Da die Annotationsrichtlinien noch in den Kinderschuhen stecken, sind Annotationen und Auswertungsdaten vorläufig. Sie sollen nur exemplarisch andeuten, welche Auswertungsmöglichkeiten und Fragestellungen denkbar sind.

³⁷ Diese vorläufige Auszeichnung bezieht sich auf den übersetzten Text. Bei Übersetzungen kommt es mitunter gerade bei einer knappen Erwähnung von inneren bzw. äußeren Vorgängen dazu, dass der Innen/Außen-Bezug nicht in gleicher Weise wie in der Ausgangssprache realisiert wird.

A) <Figurenbezug>

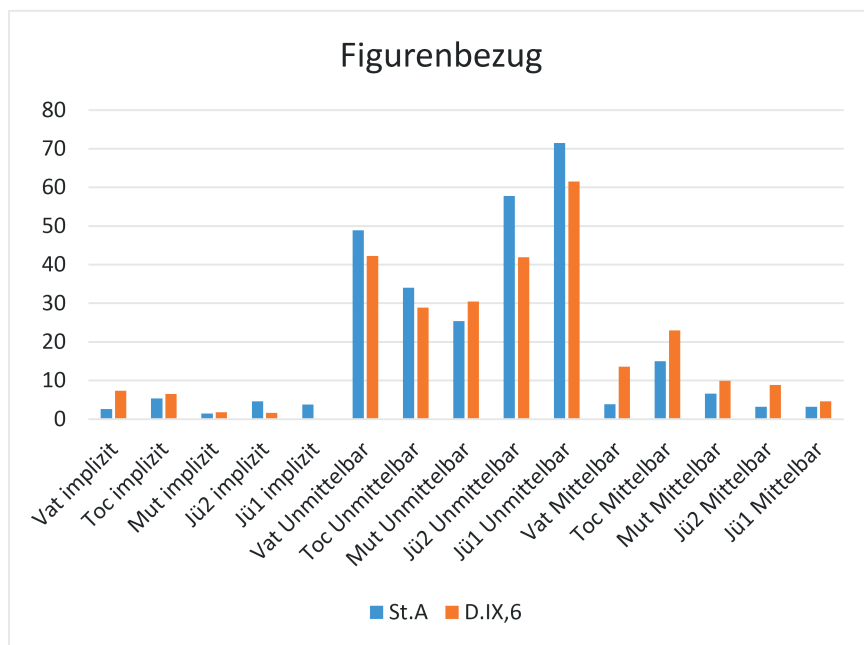


Abb. 1: Jü1: Jüngling-1; Jü2: Jüngling-2; Toc: Tochter; Vat: Vater; Mut: Mutter. Werte: Wörterzahl in %; nur Hauptfiguren.

Bei <Figurenbezug> hat die Mutter niedrige Werte; das Beilager ergibt sich recht kurzfristig. Die unmittelbaren Bezüge auf die Tochter sind ebenfalls relativ niedrig, obwohl alle vorbereitenden Aktionen und Reden auf sie bezogen sind; diese Bezüge sind in den relativ hohen mittelbaren Bezügen repräsentiert. Während die Jünglinge als Subjektaktanten selbstverständlich im Vordergrund stehen, ist weniger selbstverständlich, dass zahlreiche Reden und Handlungen unmittelbar auf die Hindernisfigur (Vater) bezogen sind.

B) Fokalisierung

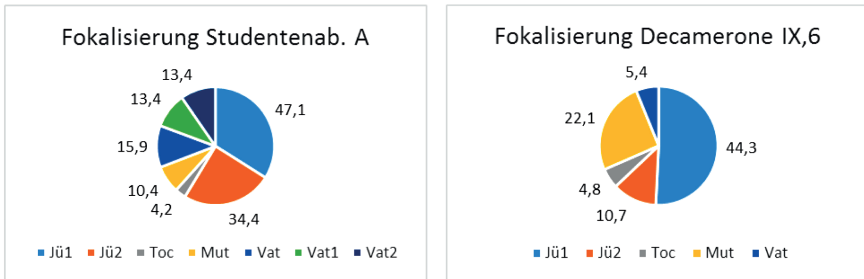


Abb. 2: Fokalisierung

Bei Boccaccio herrscht eine relativ eindeutige Fokalisierungspolitik: Jüngling-1 ist beinahe zur Hälfte allein die fokale Figur. Im mittelhochdeutschen Text verteilt sich die Fokalisierung gleichmäßiger auf mehrere Instanzen – auch auf die Väter der Jünglinge (Vat1 / Vat2); der zweite Jüngling ist hier viel häufiger ebenfalls fokale Figur als bei Boccaccio. Während der Vater bei Boccaccio von der Fokalisierungspolitik deutlich benachteiligt wird (5,4%), ist die Hindernisfigur im mhd. Text in einer ähnlich Größenordnung fokalisiert wie die Väter der Jünglinge und häufiger als die Mutter. Das Interesse der Fokalisierungspolitik für die Tochter, um die es den Jünglingen eigentlich geht, ist nahezu auf die Beilagerszene begrenzt – mit Blick auf Gender-Fragen ein bemerkenswerter Befund, wenn auch kaum überraschend. Während sich insgesamt bei Boccaccio eine stärker hierarchische Fokalisierungspolitik zeigt, ist der mittelhochdeutsche Text eher variabel fokalisiert und eher polyperspektivisch erzählt.³⁸

³⁸ Die Summen der fokalen Passagen können mehr als 100% ergeben, da kollektive Fokalisierungen zugleich auf zwei Figuren den einzelnen Figuren jeweils separat zugerechnet werden. Bei Boccaccio sind einige Passagen nicht fokalisiert.

C) Auswertung: Figurenrede / Gedankenrede

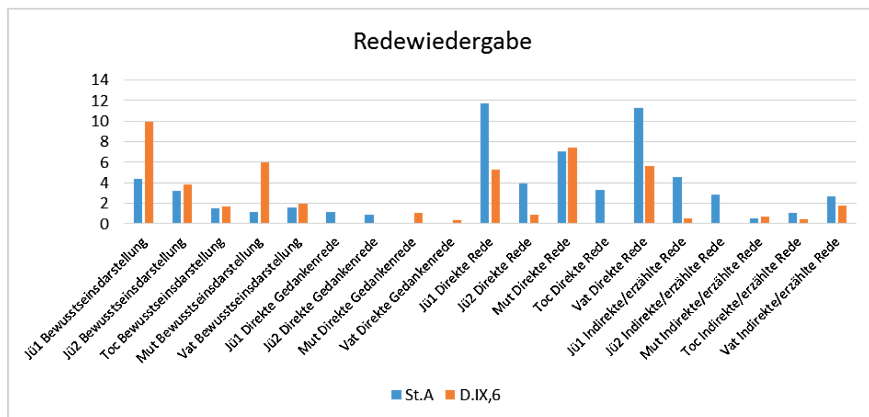


Abb. 3: BWD: Bewusstseinsdarstellung; G-Rede: Gedankenrede

Bei direkter Rede sind die hohen Werte von Jüngling-1 und Vater im mhd. Text auffällig.³⁹ Auf den ersten Blick hat es den Anschein, dass der Vater durch Begrenzung der Gedankenrede bei Boccaccio weniger stark diskriminiert wird als im mhd. Text; allerdings zeigt das folgende Diagramm, dass der Anteil der inneren Figurenrede bei Boccaccio insgesamt fast doppelt so hoch ist wie im mhd. Text, der mehr von außen erzählt, so dass der Anteil der Innenweltdarstellung des Antagonisten in Relation zu den anderen Figuren auf niedrigem Niveau bleibt.

³⁹ Hier nur Level-1-Redewiedergabe bei Hauptfiguren; also ohne eingebettete Redewiedergabe.

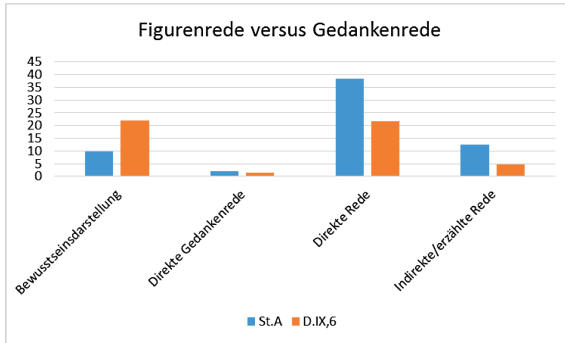


Abb. 4: Figurenrede versus Gedankenrede

Der mhd. Text besteht zur Hälfte aus Wiedergabe von gesprochener Figurenrede; Boccaccio gibt der Innenweltdarstellung breiteren Raum.

D) Wertungen

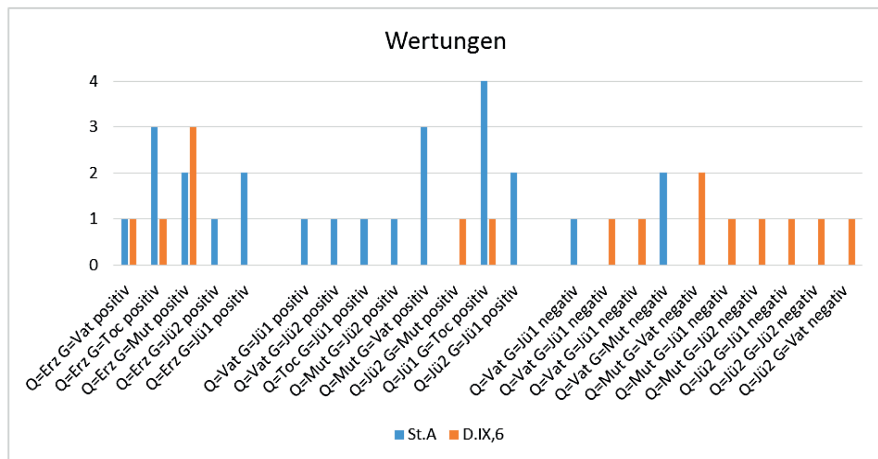


Abb. 5: Q: Quelle / Bewertende Figur; G: Gewertete Figur. Werte: Anzahl Taginstanzen. Nur Wertungen des Erzählers von Hauptfiguren bzw. Hauptfiguren, die andere Hauptfiguren bewerten. Implizit und explizit aufsummiert; ohne non-fact="true". Kollektive Bewertungen sind bei den gewerteten Figuren jeweils separat ausgewiesen.

Beide Erzähler bewerten durchgehend positiv – auch den Vater, wenn auch etwas seltener als andere Figuren. Bei den mittelhochdeutschen Figuren sind weit mehr positive Bewertungshandlungen zu verzeichnen; auch der Vater wird hier positiv bewertet. Negative Wertungen fehlen hier fast ganz, anders als bei Boccaccio, bei dem weit mehr negative als positive Wertungen stehen.

E) MTLD

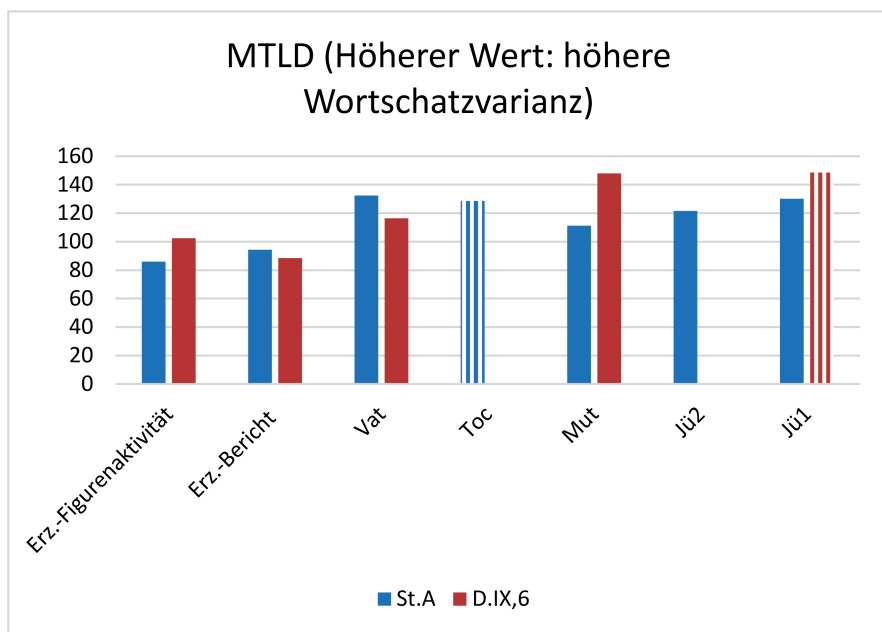


Abb. 6: MTLD

MTLD ist ein Indikator für die lexikalische Vielfalt, der anders als die Type-Token-Ratio längenunabhängig ist.⁴⁰ In beiden Texten variiert die Erzählerrede,

⁴⁰ Vgl. zu Philip M. McCarthy: *An Assessment of the Range and Usefulness of Lexical Diversity Measures and the Potential of the Measure of Textual, Lexical Diversity (MTLD)*. Dissertation Memphis. Available from Proquest Dissertations and Theses. (UMI No. 3199485), Philip M. McCarthy und Scott Jarvis: »MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment«, in: *Behavior Research Methods* 42 (2010), S. 381–392.

obschon sie hier um dixit-Formeln bereinigt wurde, weniger als die Figurenrede. Man könnte zunächst vermuten, dass der Erzähler die gebildeteren Studenten bzw. Edelmänner variationsreicher sprechen lassen würde als die Familie des Wirtes; diese Vermutung lässt sich jedoch nicht bestätigen. Während der Antagonist im mhd. Text sogar den höchsten MTLD-Wert verzeichnen darf, wird der Vater bei Boccaccio gegenüber der Mutter und, falls eingeschränkt vergleichbar,⁴¹ gegenüber Jüngling 1 benachteiligt.

F) Literarische Netzwerkanalyse

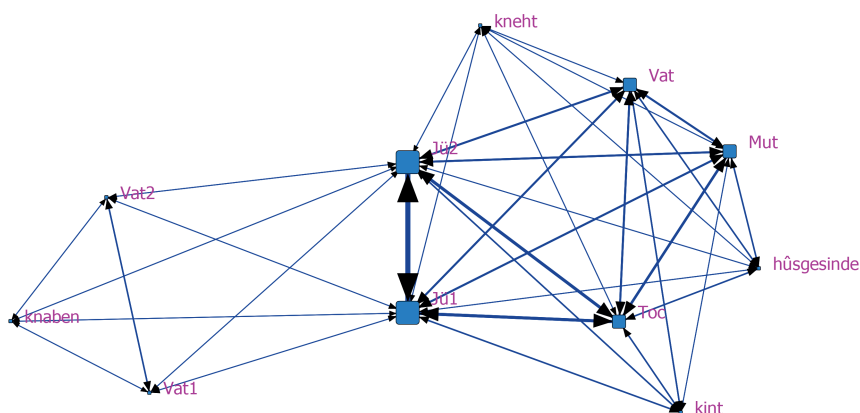


Abb. 7: St.A – <FigurFokusort>

⁴¹ Die schraffierten Werte zu Jü1 (D.IX,6) und zu Toc (St.A) sind nicht belastbar, da weniger als 100 Wörter direkte Rede vorliegen (95 und 83 Wörter). Vgl. hierzu Rie Koizumi: »Relationships Between Text Length and Lexical Diversity Measures: Can We Use Short Texts of Less than 100 Tokens?«, in: *Vocabulary Learning and Instruction* 1 (2012), S. 60–69, DOI:<http://dx.doi.org/10.7820/vli.v01.1.koizumi>. Für Toc und Jü2 liegen in D.IX,6 keine bzw. deutlich zu wenig direkte Rede vor.

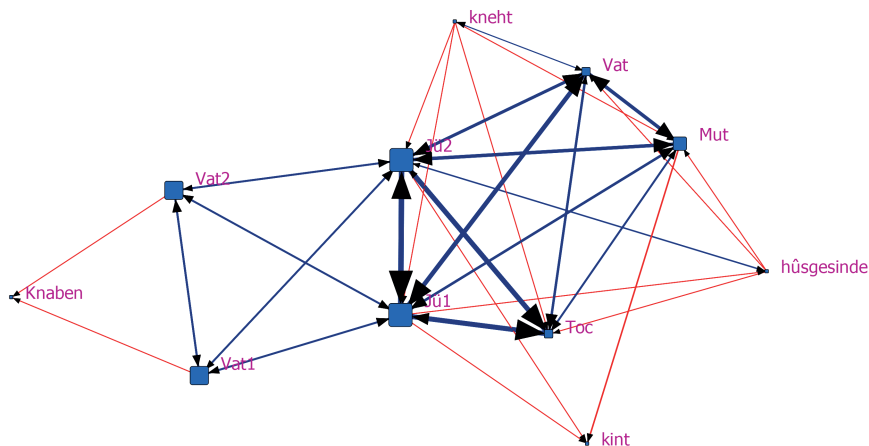


Abb. 8: St.A – <Redewiedergabe> & <Erzählerrede> => <Figurenbezug>

In der literarischen Netzwerkanalyse werden die Figurenbeziehungen untersucht. Ähnlich wie in der Studie von Trilcke, die auf der gemeinsamen Bühnenpräsenz im Drama beruht,⁴² lässt sich auswerten, welche Figuren gemeinsam am Fokusort präsent sind (links). Die vorliegenden Annotationen erlauben auch einen Abgleich von annotierter Redewiedergabe und Erzählerrede mit annotiertem Figurenbezug, so dass auch abgebildet werden kann, welche Rede oder Figurenaktivität auf welche Figur bezogen wird; möglich wird eine gewichtete und gerichtete Analyse (rechts). Während bei <FigurFokusort> die Beziehungen zwischen Mutter und beiden Jünglingen gleich gewichtet sind – sie sind im gleichen Raum –, ist durch den annotierten Figurenbezug unterscheidbar, dass Jüngling 1 nur in 22 Taginstanzen auf die Mutter Bezug nimmt. Jüngling 2, der mit der Mutter schläft, nimmt dagegen in 33 Taginstanzen auf sie Bezug. Die Beziehungen von Nebenfiguren sind meist nicht erwidert.

⁴² Peer Trilcke: »Social Network Analysis (SNA) als Methode einer textempirischen Literaturwissenschaft«, in: *Empirie in der Literaturwissenschaft*, hg. v. Philip Ajouri, Katja Mellmann und Christoph Rauen. Münster 2013 (Poetogenesis. Studien zur empirischen Anthropologie der Literatur 8), S. 201–247, ist grundlegend für die SNA im Bereich der deutschen Literatur. Vgl. weiterhin Friedrich Michael Dimpel: »Novellenschätze narratologisch auszeichnen«. Im vorliegenden Plot wurden einige Nebenfiguren und Rahmenfiguren entfernt.

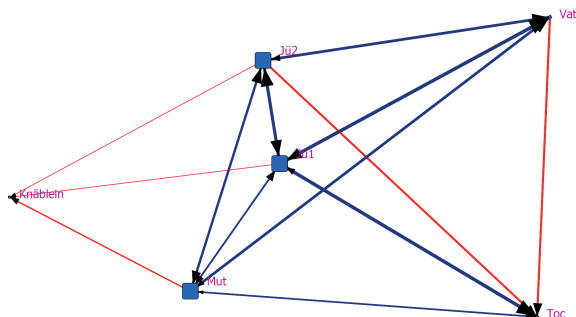


Abb. 9: D. IX,6 – <Redewiedergabe> & <Erzählerrede> => <Figurenbezug>

Bei Boccaccio sind die Beziehungen von Jüngling 2 zur Tochter und vom Vater zur Tochter nicht erwidert; die Tochter ist weniger eingebunden. Auch hier ist die Beziehung von Jüngling 2 zur Mutter stärker als die von Jüngling 1 (17 versus 11 Taginstanzen).

Bilanziert werden kann bisher, dass der Antagonist ausweislich der erfassten Daten nicht auf der Ebene der Wertungen diskreditiert wird; vielmehr wird er auf anderen Ebenen benachteiligt, etwa bei der Gedankenrede. Bei Boccaccio wird der Vater auch beim Variationsreichtum der Figurenrede und bei der Fokalisierung schlechter gestellt – anders im mhd. Text. Insgesamt können mit diesen Zahlen literaturhistorische Urteile wie etwa Neuschäfers Auffassung, dass Boccaccio gegenüber älteren Formen eher zur Offenheit oder Multiperspektivität neige, nicht bestätigt werden.⁴³ Allerdings haben diese Ergebnisdaten noch vorläufigen Charakter.

7 Ausblick

Wenn erst ein narratologisch annotiertes historisches Korpus vorliegt, werden vielfältige Studien zu multiplen Fragestellungen möglich, beispielsweise: Wie steht es um die diachrone Entwicklung von Fokalisierung, um die Eigenschaften

⁴³ Vgl. zu dieser Frage auch Friedrich Michael Dimpel: »Sprech- und Beißwerkzeuge, Kunsthandwerk und Kunst in Kaufringers ›Rache des Ehemanns‹«, in: *Daphnis* 42 (2013), S. 1–27, insbes. S. 22–27.

von Erzähler- und Figurenrede, um chronologische und temporale Alternationen, wie verteilt sich der Redebezug auf verschiedene Figurentypen wie Protagonist oder Antagonist, wie steht es um quantitative Parameter etwa bei metaphorischer Rede? Welche Spezifika können bei bestimmten Aktanten (Subjekt, Objekt, Adjuvant, Opponent) ausgemacht werden? Wie steht es um eine Korrelation kulturwissenschaftlich relevanter Terme und aktantieller Rolle? Möglich werden beispielsweise zahlreiche Gender-bezogene Auswertungen durch eine Sample-Analyse mit Figuren- oder Erzählerrede, die jeweils auf weibliche oder männliche Figuren bezogen ist oder durch eine Sample-Analyse mit Figurenrede, die jeweils von weiblichen oder männlichen Figuren stammt. Wie sind evaluative Äußerungen auf Erzählerrede und Figurenrede verteilt? Welche Aktanten bewerten bevorzugt, welche werden bevorzugt bewertet? Welche Unterschiede finden sich bei Wertungen in gesprochener und in innerer Figurenrede? Stimmen Bewertungen von Erzählern und Figuren überein im Sinne eines monoperspektivischen Erzählens? Wie und wann verbreiten sich multiperspektivische Erzähltechniken – wie steht es historisch wann um Schwarzweißmalerei? Lassen sich epochenspezifische Verteilungen ausmachen? Lassen sich Theoriebildungen überprüfen, die das Verhältnis vom Märe zur Novelle in Anschluss an die Unterscheidungskriterien von Hans-Jörg Neuschäfer unter dem Gesichtspunkt eines »Noch-Nicht« beschreiben?⁴⁴

Bibliographie

- Boccaccio, Giovanni di: *Das Dekameron*. Aus dem Italienischen übertragen von Albert Wesselski. Band III. Leipzig 1912.
- Bögel, Thomas, Michael Gertz, Evelyn Gius, Janina Jacke, Jan Christoph Meister, Marco Petris und Jannik Strötgen: »Gleiche Textdaten, unterschiedliche Erkenntnisziele? Zum Potential vermeintlich widersprüchlicher Zugänge zu Textanalyse«, in: *DHD 2015: Digital Humanities im deutschsprachigen Raum*. Graz 2015, <http://gams.uni-graz.at/archive/objects/o:dhd2015.nachlese.vortragsfolien/methods/sdef:HTML/get> (06. Juni 2017).

⁴⁴ Hans-Jörg Neuschäfer: *Boccaccio und der Beginn der Novelle. Strukturen der Kurzerzählung auf der Schwelle zwischen Mittelalter und Neuzeit*. München 1969 (Theorie und Geschichte der Literatur und der schönen Künste 8), insbes. S. 90–99. Zu Neuschäfer vgl. etwa Jan-Dirk Müller: »Noch einmal: Maere und Novelle. Zu den Versionen des Maere von den ›Dreilistigen Frauen‹«, in: *Philologische Untersuchungen gewidmet Elfriede Stutz zum 65. Geburtstag*, hg. v. Alfred Ebenbauer. Wien 1984 (Philologica Germanica 7), S. 289–311, hier S. 305, Hans-Joachim Ziegeler: »Boccaccio«, S. 14, Klaus Grubmüller: *Die Ordnung, der Witz und das Chaos*, S. 257–267, Friedrich Michael Dimpel: »Sprech- und Beiwerkzeuge«, insbes. S. 22–27.

- Brunner, Annelen: *Automatische Erkennung von Redewiedergabe. Ein Beitrag zur quantitativen Narratologie*. Berlin, Boston 2015 (Narratologia 47).
- Burrows, John F.: »Questions of Authorship: Attribution and Beyond. A Lecture Delivered on the Occasion of the Roberto Busa Award«, in: *Computers and the Humanities* 37 (2003), S. 5–32.
- Burrows, John F.: *Computation into Criticism: A Study of Jane Austen's Novels and an Experiment in Method*. Oxford 1987.
- Büttner, Andreas, Friedrich Michael Dimpel, Stefan Evert, Fotis Jannidis, Steffen Pielström, Thomas Proisl, Isabella Reger, Christof Schöch und Thorsten Vitt: »Delta« in der stilometrischen Autorschaftsattributions, in: *Konferenzabstracts DHd 2016. Modellierung – Vernetzung – Visualisierung. Die Digital Humanities als fächerübergreifendes Forschungsparadigma*, hg. v. Elisabeth Burr. Leipzig 2016, S. 61–74, <http://dhd2016.de> (08. September 2016).
- Craig, Hugh: »Stylistic Analysis and Authorship Studies«, in: *A Companion to Digital Humanities*, hg. v. Susan Schreibman, Ray Siemens und John Unsworth. Oxford 2004, S. 273–288, <http://www.digitalhumanities.org/companion> (24. Januar 2015).
- Dimpel, Friedrich Michael: »Novellenschätze narratologisch auszeichnen und analysieren am Beispiel Victor von Scheffels »Hugideo« und der sozialen Netzwerkanalyse«, in: *LiLi* 47 (2017), S. 87–108, DOI:10.1007/s41244-017-0044-8.
- Dimpel, Friedrich Michael: »Der Computerphilologe als Interpret – ein Teilzeit-Empiriker?«, in: *Literatur interpretieren: Interdisziplinäre Beiträge zur Theorie und Praxis*, hg. v. Jan Borkowski, Stefan Descher, Felicitas Ferder und Philipp Heine. Münster 2015, S. 339–359.
- Dimpel, Friedrich Michael: »Narratologische Textauszeichnung in Märe und Novelle. Mit Annotationsbeispielen und exemplarischer Auswertung von »Sperber« und »Häslein« durch MTLD und Sozialer Netzwerkanalyse«, in: *ZfdG*, vsl. 2017.
- Dimpel, Friedrich Michael: »Sprech- und Beißwerkzeuge, Kunsthandwerk und Kunst in Kaufringers »Rache des Ehemanns««, in: *Daphnis* 42 (2013), S. 1–27.
- Dimpel, Friedrich Michael: *Die Zofe im Fokus. Perspektivierung und Sympathiesteuerung durch Nebenfiguren vom Typus der Confidante in der höfischen Epik des hohen Mittelalters*. Berlin 2011 (Philologische Studien und Quellen 232), S. 17–38.
- Genette, Gérard: »Diskurs der Erzählung – ein methodologischer Versuch«, in: ders.: *Die Erzählung. Dt. Übersetzung von »Discours de récit« und »Nouvelle Discours de récit«*. Aus dem Französischen v. Andreas Knop, hg. v. Jochen Vogt. München [1972] 1998.
- Gius, Evelyn und Janina Jacke: »Informatik und Hermeneutik. Zum Mehrwert interdisziplinärer Textanalyse«, in: *Grenzen und Möglichkeiten der Digital Humanities*, hg. v. Constanze Baum und Thomas Stäcker. 2014 (ZfdG-Sonderband 1), www.ZfdG.de (24. Januar 2015).
- Gius, Evelyn und Janina Jacke: *Zur Annotation narratologischer Kategorien der Zeit. Guidelines zur Nutzung des CATMA-Tagsets*. Hamburg 2014, <http://heureclea.de/publications/guidelines.pdf> (09. März 2015).
- Gius, Evelyn: *Erzählen über Konflikte. Ein Beitrag zur digitalen Narratologie*. Berlin, Boston 2015 (Narratologia 46).
- Grubmüller, Klaus: *Die Ordnung, der Witz und das Chaos. Eine Geschichte der europäischen Novellistik im Mittelalter: Fabliau – Märe – Novelle*. Tübingen 2006, S. 127–131.
- Hübner, Gert: *Erzählform im höfischen Roman. Studien zur Fokalisierung im »Eneas«, im »Iwein« und im »Tristan«*. Tübingen 2003 (Bibliotheca Germanica 44), S. 25–63.
- Jannidis, Fotis, Gerhard Lauer und Andrea Rapp: »Hohe Romane und blaue Bibliotheken. Zum Forschungsprogramm einer computergestützten Buch- und Narratologiegeschichte des

- Romans in Deutschland (1500–1900)«, in: *Literatur und Literaturwissenschaft auf dem Weg zu den neuen Medien*, hg. v. Lucas Marco Gisi, Jan Loop und Michael Stolz. germanistik.ch 2006, http://www.germanistik.ch/publikation.php?id=Hohe_Romane_und_blaue_Bibliotheken (24. Januar 2015).
- Jannidis, Fotis: »Computerphilologie«, in: *Handbuch Literaturwissenschaft. Band 2: Methoden und Theorien*, hg. v. Thomas Anz. Stuttgart, Weimar 2007, S. 27–40.
- Jannidis, Fotis: »Methoden der computergestützten Textanalyse«, in: *Methoden der literatur- und kulturwissenschaftlichen Textanalyse. Ansätze – Grundlagen – Modellanalysen*, hg. v. Vera Nünning und Ansgar Nünning. Stuttgart 2010, S. 109–132.
- Jannidis, Fotis: »Polyvalenz – Konvention – Autonomie«, in: *Regeln der Bedeutung. Zur Theorie der Bedeutung literarischer Texte*, hg. v. ders., Gerhard Lauer, Matías Martínez und Simone Winko. Berlin, New York 2003 (Revisionen 1), S. 305–328.
- Jannidis, Fotis: »TEI in a Crystal Ball«, in: *Literary and Linguistic Computing* 24 (2009), S. 253–265.
- Kindt, Tom und Hans-Harald Müller: »Wieviel Interpretation enthalten Beschreibungen? Überlegungen zu einer umstrittenen Unterscheidung am Beispiel der Narratologie«, in: *Regeln der Bedeutung. Zur Theorie der Bedeutung literarischer Texte*, hg. v. Fotis Jannidis, Gerhard Lauer, Matías Martínez und Simone Winko. Berlin, New York 2003 (Revisionen 1), S. 286–304.
- Koizumi, Rie: »Relationships Between Text Length and Lexical Diversity Measures: Can We Use Short Texts of Less than 100 Tokens?«, in: *Vocabulary Learning and Instruction* 1 (2012), S. 60–69, DOI:<http://dx.doi.org/10.7820/vli.v01.1.koizumi>.
- Mani, Inderjeet: »Computational Narratology«, in: *The living handbook of narratology*, hg. v. Peter Hühn, Jan Christoph Meister, John Pier und Wolf Schmid. Hamburg 2013, <http://www.lhn.uni-hamburg.de/article/computational-narratology> (24. Januar 2015), S. 1–12.
- McCarthy, Philip M., Scott Jarvis: »MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment«, in: *Behavior Research Methods* 42 (2010), S. 381–392.
- McCarthy, Philip M.: *An Assessment of the Range and Usefulness of Lexical Diversity Measures and the Potential of the Measure of Textual, Lexical Diversity (MTLD)*. Dissertation Memphis. Available from Proquest Dissertations and Theses.
- McCarty, Willard: »Modeling: A Study in Words and Meanings«, in: *A Companion to Digital Humanities*, hg. v. Susan Schreibman, Ray Siemens und John Unsworth. Oxford 2004, <http://www.digitalhumanities.org/companion> (24. Januar 2015), S. 254–270.
- McCarty, Willard: *Humanities Computing*. London, New York 2005.
- McGann, Jerome: »Marking Texts of Many Dimensions«, in: *A Companion to Digital Humanities*, hg. v. Susan Schreibman, Ray Siemens und John Unsworth. Oxford 2004, <http://www.digitalhumanities.org/companion> (24. Januar 2015), S. 198–217.
- Meister, Jan Christoph, Fotis Jannidis, Christof Schöch: »Elemente einer Roadmap für das Forschungsfeld, Computational Narratology«. *DHd-Tagung 2015*, https://www.conftool.pro/dhd2015/index.php?page=browseSessions&form_session=5&presentations=show (27. Januar 2015).
- Meister, Jan Christoph: »Computational Narratology: oder: Kann man das Erzählen berechenbar machen?«, in: *Erzählen, Archivieren, Beschreiben*, hg. v. Corinna Müller und Irina Scheidgen. Marburg 2007 (Schriftenreihe der Gesellschaft für Medienwissenschaft 5), S. 19–39.

- Meister, Jan Christoph: »Crowd sourcing ›true meaning‹. A collaborative markup approach to textual interpretation«, in: *Collaborative Research in the Digital Humanities. Festschrift for Harold Short*, hg. v. Willard McCarty und Marilyn Deegan. Surrey 2012, S. 105–122.
- Meister, Jan Christoph: »Tagging Time in Prolog: The Temporality Effect Project«, in: *Literary and Linguistic Computing* 20 (2005), S. 107–124.
- Meister, Jan Christoph: »The Temporality Effect. Towards a Process Model of Narrative Time Construction«, in: *Time. From Concept to Narrative Construct: A Reader*, hg. v. ders. und Wilhelm Schernus. Berlin, Boston 2011 (Narratologia 29), S. 171–216.
- Meister, Jan Christoph: *Computing Action. A narratological approach*. Berlin, New York 2003 (Narratologia 2).
- Müller, Jan-Dirk: »Noch einmal: Maere und Novelle. Zu den Versionen des Maere von den ›Dreilistigen Frauen‹«, in: *Philologische Untersuchungen gewidmet Elfriede Stutz zum 65. Geburtstag*, hg. v. Alfred Ebenbauer. Wien 1984 (Philologica Germanica 7), S. 289–311.
- Neuschäfer, Hans-Jörg: *Boccaccio und der Beginn der Novelle. Strukturen der Kurzerzählung auf der Schwelle zwischen Mittelalter und Neuzeit*. München 1969 (Theorie und Geschichte der Literatur und der schönen Künste 8).
- Nünning, Ansgar: *Grundzüge eines kommunikationstheoretischen Modells der erzählerischen Vermittlung. Die Funktion der Erzählinstanz in den Romanen George Eliots*. Trier 1989 (Horizonte 2).
- Rehbein, Malte: »Vom Nutzen digitaler Editionen – das Göttinger ›kundige bok‹«, in: *Bibliothek und Wissenschaft* 42 (2009), S. 7–28.
- Reichlin, Susanne: »Zeitperspektiven. Das Beobachten von Providenz und Kontingenz in der ›Buhlschaft auf dem Baume‹«, in: *Kein Zufall. Konzeptionen von Kontingenz in der mittelalterlichen Literatur*, hg. v. Cornelia Herberichs und ders. Göttingen 2010 (Historische Semantik 13).
- Stehmann, Wilhelm: *Die mittelhochdeutsche Novelle vom Studentenabenteuer*. Berlin 1909 (Palaestra 67).
- Trilcke, Peer: »Social Network Analysis (SNA) als Methode einer textempirischen Literaturwissenschaft«, in: *Empirie in der Literaturwissenschaft*, hg. v. Philip Ajouri, Katja Mellmann und Christoph Rauen. Münster 2013 (Poetogenesis. Studien zur empirischen Anthropologie der Literatur 8), S. 201–247.
- Winko, Simone: *Wertungen und Werte in Texten. Axiologische Grundlagen und literaturwissenschaftliches Rekonstruktionsverfahren*. Braunschweig 1991 (Konzeption Empirische Literaturwissenschaft 11).
- Ziegeler, Hans-Joachim: »›Studentenabenteuer A und B‹ (FM 129 u. 107)«, in: *Kleinepik, Tierepik, Allegorie und Wissenskultur*, hg. v. Fritz Peter Knapp. Berlin, Boston 2013 (GLMF VI), S. 100–110.
- Ziegeler, Hans-Joachim: »Boccaccio, Chaucer, Mären, Novellen: ›The Tale of the Cradle‹«, in: *Kleinere Erzählformen im Mittelalter. Paderborner Colloquium 1987*, hg. v. Klaus Grubmüller, Peter L. Johnson und Hans-Hugo Steinhoff. Paderborn 1988, S. 9–31.

TEI-Guidelines

<http://www.tei-c.org/Guidelines/Customization/odds.xml> (13. März 2015).

<http://www.tei-c.org/release/doc/tei-p5-doc/en/html/NH.html>

