



Separatum aus:

---

## THEMENHEFT 12

*Elisabeth Lienert / Joachim Hamm  
Albrecht Hausmann / Gabriel Viehhauser (Hrsg.)*

# Digitale Mediävistik

## Perspektiven der Digital Humanities für die Altgermanistik

Publiziert im November 2022.

Die BmE Themenhefte erscheinen online im BIS-Verlag der Carl von Ossietzky Universität Oldenburg unter der Creative Commons Lizenz [CC BY-NC-ND 4.0](https://creativecommons.org/licenses/by-nc-nd/4.0/). Die ›Beiträge zur mediävistischen Erzählforschung‹ (BmE) werden herausgegeben von PD Dr. Anja Becker (München) und Prof. Dr. Albrecht Hausmann (Oldenburg). Die inhaltliche und editorische Verantwortung für das einzelne Themenheft liegt bei den jeweiligen Heftherausgebern.

<http://www.erzaehlforschung.de> – Kontakt: [herausgeber@erzaehlforschung.de](mailto:herausgeber@erzaehlforschung.de)  
ISSN 2568-9967

*Zitiervorschlag für diesen Beitrag:*

Dimpel, Friedrich Michael/Blessing, Andre/Hinkelmanns, Peter/Ketschik, Nora/  
Zeppezauer-Wachauer, Katharina: Figuren und ihr Handeln. Eine computergestützte  
Untersuchung von Figurenaktivitäten im Kontext von Figurenreferenzen mit Hilfe des  
Begriffssystems der MHDBDB, in: Lienert, Elisabeth/Hamm, Joachim/Hausmann, Albrecht/  
Viehhauser, Gabriel (Hrsg.): Digitale Mediävistik. Perspektiven der Digital Humanities für die  
Altgermanistik, Oldenburg 2022 (BmE Themenheft 12), S. 283–328 (online).

*Friedrich Michael Dimpel / Andre Blessing  
Peter Hinkelmanns / Nora Ketschik  
Katharina Zeppezauer-Wachauer*

## Figuren und ihr Handeln

Eine computergestützte Untersuchung von  
Figurenaktivitäten im Kontext von Figurenreferenzen mit  
Hilfe des Begriffssystems der MHDBDB

*Abstract.* Im Kontext von Figurenreferenzen, die mit Hilfe der Annotationsumgebung CRETAnno teilautomatisch annotiert wurden, werden bestimmte Kategorien aus dem Begriffssystem der MHDBDB erfasst, die Aktivitäten von Figuren repräsentieren können. Diese Kategorien werden für vier Figurentypen (weibliche und männliche Hauptfigur, Zofen, Opponenten) ermittelt und mit einer Studie verglichen, die mit manueller Annotation Aktivitäten von Figurentypen erfasst hat – jeweils für ›Iwein‹, ›Tristan‹, ›Partonopier‹ und ›Mauritius von Craun‹. Es zeigt sich unter anderem, dass die automatische Erfassung von Aktivitäten eine Unterscheidung der Figurentypen in diesem Korpus ermöglicht. Mit Ausnahme der Aktivität ›Kopulieren‹, die meist nicht unmittelbar lexikalisch, sondern eher verhüllend oder metaphorisch adressiert wird, werden plausible Ergebnisse erzielt.

### 1. Rückblick: Die Paidia-Studie

In dem Paidia-Beitrag »Versuch einer quantitativen Analyse von Figurenaktivitäten in ›Iwein‹, ›Tristan‹, ›Partonopier‹ und ›Mauritius von Craun‹ in Analogie zu Computerspielen« (Dimpel 2018) wurde untersucht, inwieweit Aktivitäten wie ›Suchen‹, ›Töten‹ oder ›Argumentieren‹ bestimmten

Figurentypen zugeschrieben wurden. Die Studie konzentrierte sich dabei auf die weibliche und männliche Hauptfigur, ein bis drei Opponenten und je einen Adjuvanten in Gestalt einer Zofe. Auch wenn das Ziel eine quantitative Auswertung war, wurden die Aktivitäten doch zunächst manuell registriert: Gezählt wurde im Rahmen eines Lektürevorgangs, ob die jeweiligen Aktivitäten in einem bestimmten Textsegment den jeweiligen Figuren eingeschrieben sind, und zwar dann, wenn in den Texten explizit von den Aktivitäten berichtet wird. Darüber hinaus erfolgte eine Zählung mitunter auch dann, wenn der Kontext es erlaubt, mit einem hohen Grad an Offensichtlichkeit auf die fragliche Aktivität zu inferieren (vgl. zur Inferenzbildung Jannidis 2004, S. 206).

Auch wenn eine solche manuelle Erfassung in der Regel präziser sein wird als eine automatische, war und ist doch zu betonen, dass zwar viele, aber nicht alle Textbefunde einen eindeutigen Schluss darauf zulassen, ob etwa eine kommunikative Aktivität bereits unter ›Beklagen‹ fällt oder ob eine leichte Auseinandersetzung bereits das Etikett ›Kämpfen‹ verdient. Die mitunter vorhandenen Probleme, eine Textpassage einer Aktivität zuzuordnen, beruhen nicht nur auf der häufig konstatierten Polyvalenz literarischer Texte. Vielmehr hat man mit einem grundlegenden Problem der digitalen Literaturwissenschaft zu tun: Wenn man nicht nur alltagssprachlich umschreiben will, ob eine Figur kämpft, sondern digital über Sein oder Nichtsein einer Aktivität entscheiden soll, ist es nötig, die Aktivität möglichst eindeutig zu modellieren. Willard McCarty hat die Modellbildung und Modellweiterentwicklung als die zentrale Tätigkeit von digitalen Literaturwissenschaftler\*innen beschrieben, die nicht etwa nur einen Mangel wie unvollständige Explizitheit kompensiere, sondern die vielmehr häufig unmittelbar zu tiefergehenden Erkenntnissen über den Gegenstand des Modellierten führe (vgl. McCarty 2004, S. 254–270, insbes. S. 258, sowie McCarty 2005, S. 23–72.).

Die Ergebnisse der quantitativen Auswertung sind durchaus kompatibel zu einer konventionellen literaturwissenschaftlichen Einschätzung: Die

männlichen Hauptfiguren haben erhöhte Werte bei ›Sich-Bewegen‹; die weiblichen Hauptfiguren haben mit Ausnahme von Isolde im ›Tristan‹ verhältnismäßig niedrige Bewegungswerte und wären daher eher als statische Figuren einzustufen; Zofen helfen und beraten oft. Männliche Hauptfiguren essen häufig, während Opponenten und weibliche Hauptfiguren viel seltener bei der Nahrungsaufnahme gezeigt wurden. Bei den Opponenten wird lediglich Marke aus dem ›Tristan‹ bei einem Beilager gezeigt. Bei der Gegenüberstellung von sprachlichen und nichtsprachlichen Aktivitäten ergab sich ein überraschender Befund: Dass weibliche Hauptfiguren eher mittels Sprache handeln, war erwartbar, nicht hingegen, dass es sich bei den Opponenten ebenso verhält.

## 2. Automatisierungsversuch

Während in dem Paidia-Beitrag die Aktivitäten manuell gezählt wurden, soll hier erprobt werden, inwieweit man mit einer automatischen Zählung der Aktivitäten zu ähnlichen oder anderen Ergebnissen gelangen kann.

Bei diesem Projekt handelt es sich um eine Vorstudie, in der erstens für ein kleines Testkorpus ein *Procedere* etabliert wird, bei dem Informationen auf Basis des Begriffssystems der Mittelhochdeutschen Begriffsdatenbank (MHDBDB) und Figurenreferenzen mithilfe der Annotationsumgebung [CRETAnno](#) in ausgewählten Textausschnitten annotiert werden. Zweitens wird ein Auswertungssystem entwickelt, das eine Plausibilitätsprüfung der Daten erlaubt. Drittens erfolgen eine Prüfung der Ergebnisse und ein Abgleich der Ergebnisdaten mit den Daten der Paidia-Studie. Diese zeichnet sich zwar gegenüber automatischen Verfahren durch die größere Verlässlichkeit einer manuellen Annotation aus, sie hat aber den Nachteil, dass sie nicht ohne enormen Aufwand auf größere Textmengen und auf weitere Kategorien hochskaliert werden kann. Bei dem Vergleich mit den Paidia-Daten muss allerdings in Kauf genommen werden, Äpfel mit Birnen zu vergleichen – schon deshalb, weil im Begriffssystem der MHDBDB keine

exakten Entsprechungen zu den Aktivitäten vorliegen, die in der Paidia-Studie untersucht wurden. Mit dieser Vorstudie wird auch das Anliegen verfolgt, Probleme und Chancen auszuloten, wenn künftig in größerem Umfang Begriffe aus der MHDBDB im Umkreis von Figurenreferenzen in einem größeren Korpus und bei weiteren Figurentypen untersucht werden.

Bei einem solchen automatischen Unternehmen stellt sich erstens die Problematik, Referenzen auf Figuren zu identifizieren, zweitens, im Text erwähnte Aktivitäten einer Figur zuzuordnen, und drittens, eine Aktivität automatisch zu erkennen und zu klassifizieren.

Bei unserem Versuchsaufbau und bei der zugrundeliegenden Modellbildung gehen wir davon aus, dass die lexikalische Zuordnung von MHDBDB-Kategorien zu Aktivitäten, die wiederum auf einer Verknüpfung der Kategorien mit Lemmata beruht, zumindest eine Annäherung an die gesuchten Aktivitäten realisieren kann. Selbstverständlich kann man nicht davon ausgehen, dass die jeweilige Aktivität – wie auch immer man sie in traditionellen literaturwissenschaftlichen Studien konzeptualisieren wollte – tatsächlich vollständig und exakt durch diesen lexikalischen Ansatz abgebildet werden kann.

### 3. Figurenreferenzen erfassen mit CRETAnno

Das Erkennen von Figuren fällt dem menschlichen Leser meist relativ leicht. Figuren können mit ihrem Eigennamen (z. B. ›Parzival‹) bezeichnet werden, mit einem Gattungsnamen (z. B. *der knappe*) oder mit Pronomina (›er‹ oder ›dieser‹). All diese Figurenreferenzen stehen stellvertretend für eine bestimmte Figureninstanz und werden dieser bei der Lektüre eines Textes durch einen menschlichen Leser ›einfach‹ zugeordnet – Unsicherheiten bei der Zuordnung gibt es in seltenen Fällen bei unscharfen Referenzen auf Figurengruppen oder bei komplexen Gemengelagen. Eine automatische Erkennung von Figurenreferenzen sieht sich im Gegensatz zur menschlichen Lektüre mit vielfältigen Herausforderungen konfrontiert.

Systematisch zerfällt die Erkennung von Figurenreferenzen in zwei Teilaufgaben: Erstens die sog. Named Entity Recognition (NER), die zur Aufgabe hat, Eigennamen in einem Text zu finden und diese einer bestimmten Kategorie (die gängigen Kategorien sind Person, Ort, Organisation) zuzuordnen. Zweitens die Koreferenzresolution, die sowohl die Erkennung von Figurenerwähnungen (etwa Gattungsnamen wie ›Frau‹) als auch ihre Auflösung auf die entsprechende Entität (z. B.: ›Frau‹ referiert auf die Instanz ›Angela Merkel‹) umfasst. Für die vollautomatische NER gibt es zwar eine Reihe etablierter Werkzeuge (bekannt sind etwa die für verschiedene Sprachen verfügbaren Stanford [NER-Modelle](#)), für das Mittelhochdeutsche fehlen solche Werkzeuge jedoch. Zusätzlich funktionieren sie primär für Daten, die den Trainingsdaten ähneln (i. d. R. Zeitungskorpora), wohingegen ihr Einsatz auf literarische Texte mit einer erheblichen Verschlechterung der Erkennungsquote einhergeht. Für das hier angestrebte Untersuchungsvorhaben lag es daher nahe, eine teilautomatische Erkennung der Figurenreferenzen mit Unterstützung der Annotationsumgebung CRETAnno zu verfolgen. Der zugrundeliegende Workflow wurde im Rahmen des Stuttgarter Zentrums für reflektierte Textanalyse (CRETA) entwickelt und erprobt.<sup>1</sup> Indem die Figurenreferenzen zunächst in einem kleinen Textabschnitt manuell annotiert werden und anschließend auf Basis dieser Annotationen ein Vorhersagemodell trainiert wird, wird das Verfahren auf sprachliche und textliche Spezifika zugeschnitten. Der teilautomatische Ansatz hat den Vorteil, dass er einerseits die Annotationsarbeit gegenüber einer rein manuellen Vorgehensweise beschleunigt, andererseits aber eine Qualitätskontrolle – nämlich eine Korrektur der automatischen Erzeugnisse über die Benutzeroberfläche – erlaubt. Der Workflow zur Annotation der Figurenreferenzen für die hier vorgelegte Studie wird im Abschnitt ›Workflow 2‹ genauer erläutert.

Nachdem die Figurenreferenzen erfasst sind, gilt es, die Aktivitäten den jeweiligen Figuren zuzuordnen. Das naheliegende Mittel ist eine Zuordnung über Kookkurrenzen – wenn also eine Aktivität im Umkreis einer

bestimmten Anzahl an Wörtern von einer Figurenreferenz erwähnt wird. Allerdings bringt dieses Vorgehen gewisse Unschärfen mit sich, beispielsweise bei Formulierungen, die den Bezug von Begriffen auf Figuren einschränken, etwa durch Negationen oder durch Konjunktiv II. Während zudem ein menschlicher Leser bei einer Formulierung wie ›Isolde belügt Marke‹ keine Probleme haben dürfte, würde durch das Kookkurrenz-Verfahren auch Marke mit der Aktivität ›Lügen‹ in Verbindung gebracht. Um eine Zuordnung einer Textpassage zu einer Figur mit größerer Sicherheit automatisch prozessieren zu können, wäre es nötig, zuvor die Kategorie ›Figurenbezug‹ manuell zu annotieren, was wiederum mit einem erheblichen Aufwand verbunden wäre (vgl. Dimpel 2017, S. 94).

#### 4. Wortschatzbasierte Begriffe der MHDBDB als Aktivitäten

Bei wortschatzbasierten Zugriffen auf tatsächliche Phänomene trifft man ebenfalls auf Probleme – etwa bei ambigem Vokabular. Doch bis bessere Optionen realisierbar sind, scheint mit einem wortschatzbasierten Vorgehen immerhin eine gewisse Annäherung an die gesuchten Phänomene möglich zu sein.

Das Begriffssystem der MHDBDB geht zurück auf den Entwurf eines »Begriffssystems als Grundlage für die Lexikographie« Rudolf Halligs und Walther von Wartburgs und ergänzt diesen um Begriffe, die eine genauere Erfassung in Bezug auf die mittelalterliche Lebenswelt ermöglichen (vgl. Hallig/Wartburg 1963). Damit reiht es sich ein in die onomasiologischen Thesaurus-Ansätze von Roget 1864 bis hin zum ›[Historical Thesaurus of the OED](#)‹. Allen gemein ist, dass eine hierarchische Gliederung von Begriffen oder Kategorien dazu dient, den Wortschatz inhaltsseitig zu erschließen. In der MHDBDB werden die Bedeutungen eines Lemmas mittels Zusammensetzung aus Begriffen ausgedrückt. Diese Bedeutungen können von den Tokens der Korpus-texte referenziert werden, um die Bedeutung eines Wortes im Kontext zu kennzeichnen. Somit werden etwa auch Wort-

schatzanalysen ermöglicht, indem alle über Bedeutungspositionen mit einem Begriff verbundenen Wortartikel angezeigt werden können.

## 5. Aktivitäten und Begriffe – Tentatives Mapping

Den Aktivitäten aus dem Paidia-Beitrag wurden Begriffe aus der MHDBDB zugeordnet. Teilweise war es notwendig, diesen Aktivitäten mehrere Begriffe zuzuweisen, so dass insgesamt 62 Begriffe berücksichtigt werden.

Zu den Aktivitäten ›Suchen‹, ›Listhandeln‹, ›Verstecken‹, ›Warten‹, ›Gefangennahme‹, ›Erpressen‹, ›Stehlen‹, ›Zuhören‹, ›Unterweisen‹, ›Beraten‹, ›Drohen‹, ›Streiten‹, ›Verurteilen‹ und ›Abweisen‹ wurden keine plausiblen Entsprechungen im Begriffssystem gefunden. Bei den sprachlichen Aktivitäten konnte damit leider nur bei einem kleinen Teil ein Mapping zum Begriffssystem der MHDBDB erstellt werden. Da angestrebt wurde, jeden Begriff nur einer Aktivität zuzuordnen, wurde etwa darauf verzichtet, 22826000 (›Willensausübung auf andere‹) und 22642000 (›Wahrheit/Irrtum/Lüge/Täuschung ‹) auch dem nichtsprachlichen ›Listhandeln‹ zuzuordnen. Ein Blick in die Daten ergab, dass diese Begriffe eher zu den sprachlichen Aktivitäten ›Veranlassen‹ bzw. ›Lügen‹ passen.

a_Essen	Mahlzeiten	21111200
a_Essen	Ernährung	21111000
a_Essen	Gerichte/Speisen	21111306
a_Essen	Nahrungsmittel/Lebensmittel	21111300
a_Feiern	Festlichkeiten/Feiern	23135100
a_Feiern	Feste/Spiel/Unterhaltung	23135000
a_Feiern	Spiel/Zerstreuung	23135200
a_Feiern	Spielzeug	23135220
a_Feiern	Schachspiel	23135210
a_Feiern	Andere Spiele	23135230
a_Feiern	Tanz	25132000
a_Feiern	Tanz/Namen	25132100
a_Feiern	Instrumentalmusik	25153000
a_Feiern	Instrumentalmusik/Namen	25153100
a_Feiern	Gesang	25152000
a_Feiern	Gesang/Namen	25152100
a_Feiern	Musik/Allgemeines	25150000
a_Schlafen	Schlaf/Ausruhen	21080000
a_Helfen	Hilfeleistung/Dienst/Widerstand	23133000



Dimpel [u.a.]: Figuren und ihr Handeln

a_Hilfsmittel	Voraussetzungen zur Aktion - Mittel	22831500
a_Hilfsmittel	Objektbezogene Aktivität/Tätigkeit	21072000
a_Hilfsmittel	Werken/Werkzeuge/Utensilien/Ausrüstung	23304100
a_Jagen	Jagd/Fischerei	23302000
a_Jagen	Jagd/Fischerei/Waffen/Geräte	23302200
a_Jagen	Jagd/Fischerei/Waffen/Geräte/Namen	23302210
a_Jagen	Jagd/Fischerei/Gebräuche/Rituale	23302300
a_Jagen	Jagdtiere/Namen	23302400
a_Jagen	Fischzucht	23302500
a_Jagen	Raubvögel	14024000
a_Kämpfen	Kriegswesen/Kampf/Gewalt	23240000
a_Kämpfen	Kriegswesen/Kampf/Namen	23241000
a_Kämpfen	Heerfahrt	23242000
a_Kämpfen	Schlacht	23243000
a_Kämpfen	Aventiure	23244000
a_Kämpfen	Belagerung	23245000
a_Kämpfen	Zweikampf	23246000
a_Kämpfen	Rangordnung	23247000
a_Kämpfen	Ritterwesen/Allgemeines	23200000
a_Kämpfen	Waffen	23231000
a_Kämpfen	Waffen/Namen	23231100
a_Kämpfen	Kampf-/Waffentechnik	23231200
a_Kopulieren	Sexualleben/Erotik	21112000
a_Kopulieren	Geschlechtsmerkmale/Genitalien/Zeugung/ Geburt	21045000
a_Bewegen	Bewegung	31800000
a_Turnieren	Turnier/Organisation/Aufzug	23261000
a_Turnieren	Turnier/Schauplatz	23262000
a_Turnieren	Turnier/Kampfformen	23263000
a_Turnieren	Turnier/Rituale	23264000
a_Turnieren	Turnierwesen	23260000
a_Töten	Tod	21104000
a_Peinigen	Ärger/Zorn/Wut/Aggression	22705200
b_Reden	Mündliche Kommunikation	23121000
b_Reden	Eigenschaften/Mängel der stimm. Äußerung	23121100
b_Reden	Stimmlicher Ausdruck	23121200
b_Reden	Kommunikation und Sprache	23120000
b_Argumentieren	Beweis/Zustimmung/Widerspruch	22641000
b_Veranlassen	Autorität/Anordnung/Befehl	22826100
b_Veranlassen	Willensausübung auf andere	22826000
b_Beklagen	Gefallen/Missfallen	22702010
b_Beklagen	Glückseligkeit/Unglück	22702020
b_Beklagen	Freude/Leid	22702030
b_Lügen	Wahrheit/Irrtum/Lüge/Täuschung	22642000

Tabelle 1: Mapping Aktivitäten-Begriffe. Ein vorgestelltes »a\_« zeigt nichtsprachliche, ein vorgestelltes »b\_« zeigt sprachliche Aktivitäten an.

## 6. Figurenstudien in der digitalen Literaturanalyse

Es gibt eine Reihe von Studien zu Figuren in den DH, die jedoch nur geringes Anknüpfungspotential zum hier vorgelegten Vorhaben bieten. Die von Peer Trilcke und Frank Fischer vorgelegten Analysen zu Dramen bedienen sich der Methode der Sozialen Netzwerkanalyse, um u. a. offene und geschlossene Dramen miteinander zu vergleichen (vgl. grundlegend: Trilcke 2013, S. 201–247, sowie das Forschungsprojekt [dlina](#), Digital Literary Network Analysis). Damit basieren die Untersuchungen zwar auch auf einer Erfassung von Figurenvorkommen, diese bleibt aber auf Eigennamen beschränkt, die im Drama aus dem Nebentext extrahiert werden können. Auch die Forschungen des QuaDrama-Projekts ([Quantitative Drama Analytics](#)) beschäftigen sich mit Dramentexten, widmen sich hierbei aber auch Problemen der Koreferenz und ihrer Resolution (vgl. z. B. Pagel/Reiter 2020; Wiedmer [u. a.] 2020). Figurenbezogene Studien zu Erzähltexten wurden etwa hinsichtlich ihrer sozialen Interaktion (vgl. etwa: Elson [u. a.] 2010; Agarwal/Rambow 2010; Agarwal [u. a.] 2012; Dimpel 2020) oder ihres Raum-Bezugs (Viehhauser/Barth 2017) durchgeführt; spannend ist die Studie zu Figurentypen im modernen Drama von Krautter ([u. a.] 2020). Für mittelhochdeutsche Erzähltexte liegen erst wenige Studien vor; diese sind im Rahmen des DH-Zentrums CRETA entstanden und beschäftigen sich insbesondere mit sozialen Relationen (vgl. Braun/Ketschik 2019; Ketschik [u. a.] 2020, hier insb. Kap. 3.2). Der Zusammenhang von Figuren und Aktivitäten wurde bisher nicht in den Blick genommen. Gegenüber Topic Modeling-Ansätzen bietet das hier vorgestellte Verfahren den Vorteil, dass es gezielt ausgewählte Begriffe (hier: Aktivitäten) und ihre Ko-Okkurrenz mit ausgewählten Figureninstanzen untersucht. Über ein Topic Modeling-Verfahren werden hingegen dominierende Themenfelder für einen bestimmten Textabschnitt extrahiert. Zusätzlich bleibt hier – im Gegensatz zu vollautomatischen Verfahren – die Transparenz gewahrt, wie die

ausgewählten Begriffe zustande kommen, was eine kritische Reflexion der resultierenden Ergebnisse ermöglicht.

## 7. Workflow 1: sql-Abfrageergebnis nach xml konvertieren

In einem ersten Schritt wurden Daten zu den Texten und den gewählten Begriffen aus der MHDBDB erfasst. Exportiert worden sind die Begriffsrelationen und das POS-Tagging der Tokens, deren Bedeutungspositionen mindestens einen der zuvor ausgewählten Begriffe umfasst. Für Gottfrieds ›Tristan‹ ergab die Abfrage beispielsweise:

```
1,0,65537,6748,"gedaehte",32768,1992,"gedenken","TR",21072000.
```

Diese Daten wurden wieder in Texte konvertiert. Dazu wurden Perl-Programme entwickelt, die die Zeilennummer in ein xml-Tag umsetzen. Danach folgen die Wortformen des Textes und, falls vorhanden, das Kategorien-Tag <kat>, das kommasepariert die Kategorien bzw. Begriffe zur vorstehenden Wortform auflistet. ›Tristan‹, V. 9, erhielt damit folgende Form:

```
<l>9</l> ich hoere es velschen <kat>21072000,22642000,23120000,  
23304100</kat> harte vil.
```

Bereits dieser Schritt, der zunächst trivial anmutet, war mit einigem Aufwand verbunden, da die ebenfalls in Perl realisierte Konsistenzprüfung, die auch einen Abgleich mit vorhandenen Texten umfasste, eine Reihe an Zuordnungsproblemen und Abweichungen ergab, die teilweise aus verschiedenen Textgrundlagen resultierten und teilweise aus einer abweichenden Nummerierung.<sup>2</sup> Die mit den Kategorien-Tags versehenen Texte wurden sodann in CRETAnno eingespeist.

## 8. Workflow 2: Erkennung von Figurenreferenzen

Das Annotationstool CRETAnno war ursprünglich für eine vollständige Annotation bestimmter Entitätenklassen (etwa Personen, aber auch Orte oder Organisationen) gemäß den zugrundeliegenden Richtlinien ([online](#)) entwickelt. Für die neue Studie, die lediglich einige zuvor festgelegte Figuren in den Texten untersuchen möchte, wurde die Benutzeroberfläche so angepasst, dass das Korrekturmenü nur die Liste mit den zu untersuchenden Figuren enthielt.<sup>3</sup> Um den halbautomatischen Annotationsprozess nicht bei null beginnen zu müssen, wurde ein bereits entwickeltes Vorhersagemodell auf die neuen Daten angewandt. In einer ersten Sichtprüfung zeigte sich, dass das System, das an anderen Texten trainiert worden war, die auf diesen Daten erzielte Erkennungsquote (Recall von bis zu 76%, Precision von bis zu 91%; Blessing [u. a.] 2017) auf den neuen Texten nicht erreichen konnte. Zusätzlich galt es, das Modell für den neuen Anwendungsfall, nämlich die Beschränkung auf bestimmte Figuren, anzupassen. Deshalb wurden für den ›Tristan‹ und den ›Partonopier‹ jeweils ca. 1000 Verse manuell vorannotiert,<sup>4</sup> um das Erkennungsmodell neu zu trainieren. Der ›Iwein‹ lag in Stuttgart bereits annotiert vor, für den kurzen ›Mauritius‹ hätte sich der Aufwand dafür nicht amortisiert.

Nach diesem Schritt wurden die automatisch erzeugten Annotationsvorschläge vollständig kontrolliert und nachannotiert – für diese Arbeit danken wir Anne Faust (WHK, Darmstadt). Zu unklaren Textbefunden wurde eine Protokolldatei angelegt, die durch Friedrich Michael Dimpel überprüft wurde.<sup>5</sup>

Zusätzlich zu Figurenreferenzen über Eigennamen (etwa ›Isolde‹) oder Appellative (wie *mîn muoter* oder *der getriuwe*) wurden in bestimmten Fällen pronominale Referenzen auf eine Figur erfasst – und zwar dann, wenn eine Referenz auf diese Figur in den zehn Versen zuvor nicht bereits annotiert worden ist. Dies war notwendig, um auch Kategorien, die nur im

Kontext pronominaler Figurenreferenzen vorkommen, der entsprechenden Figur zuordnen zu können.

Nach Abschluss der Annotation wurden die Texte aus CRETAnno exportiert – einschließlich der xml-Informationen zu den Kategorien und zu den Figurenreferenzen. Mit einem Perl-Skript wurde die Konsistenz der Daten überprüft. Für Hartmanns ›Wein‹, der im Rahmen der CRETA-Arbeiten bereits mit Figurenreferenzen annotiert wurde, wurden die <kat>-Tags integriert sowie abschließend noch einige pronominale Referenzen nachannotiert.

### 9. Workflow 3: Bereinigung der Begriffskategorien aus der MHDBDB

Bei ersten Auswertungsversuchen fiel auf, dass insgesamt überraschend viele Wortformen mit Begriffen annotiert wurden. Ein Blick in die Daten zeigt, dass die Lemma-Begriff-Zuordnung in der MHDBDB eher großzügig vorgenommen wurde, da die verwendeten Korpus Texte noch nicht bedeutungsdisambiguiert sind. Das Lemma *ritter* ist beispielsweise in einer der drei differenzierten Bedeutungen dem Begriff 21111306 ›Gerichte/Speisen‹ zugeordnet. Diese Zuordnung hat durchaus ihre Berechtigung, listet doch der große Lexer unter *ritter* auch ›arme ritter, ein backwerk‹ als Bedeutungsangabe (Lexer 1992, Bd. 2, Sp. 466). Die Zuordnung ist für den primären Zweck der MHDBDB sinnvoll: Viele Anwender verwenden die MHDBDB, um herauszufinden, welche Begriffe oder Konzepte in welchen Texten anzutreffen sind, um nach diesem Rechreschritt manuell die Fundstellen auszuwerten, die dann wiederum etwa in eine Textanalyse einfließen können. Da die Texte der Begriffsdatenbank zwar lemmatisiert, aber bislang nicht vollständig nach verschiedenen Bedeutungspositionen disambiguiert worden sind, müssen bei nicht bedeutungsannotierten Texten zunächst ›falsche‹ Begriffe herausgefiltert werden.

Für die vorliegende Analyse ist es jedoch durchaus relevant, dass die Begriffe, die im Umkreis von Figurenreferenzen gefunden werden, nicht nur im Einzelfall tatsächlich für den gesuchten Begriff einschlägig sind, sondern idealerweise ausschließlich, zumindest aber möglichst überwiegend: Ritter werden im Artusroman eher selten verspeist, auch wenn es in ›Dietrichs Flucht‹ um Ortnit anders bestellt sein mag. Das Konzept 21111000 ›Ernährung‹ ist auch Bedeutungspositionen der Lemmata *starcken*, *twingen*, *vaste* zugeordnet, die in diesen Texten ebenfalls eher selten im Ernährungskontext verwendet werden.

Um die Begriff-Zuordnung um eher periphere Bedeutungen bereinigen zu können, wurden 2020 aus der MHDBDB alle hier untersuchten Begriffe, die einer Wortform in den vier hier untersuchten Texten zugeordnet sind, exportiert.<sup>6</sup> Ein Perl-Skript wurde entwickelt, das nach Wortform-Begriff-Zuordnungen sucht, die in den vier hier untersuchten Texten fünf Mal oder häufiger vorkommen. Wenn diese Zuordnung auf einer Wortbedeutung beruht, die im Korpus als eher selten eingestuft wurde, wurde die Zuordnung des zugehörigen Lemmas<sup>7</sup> in eine Löschliste aufgenommen – also dann, wenn ein deutlicher Verdacht bestand, dass das Lemma im Korpus überwiegend nicht in dieser Bedeutung vorkommt. Den Begriffen sind auch Lemmata zugeordnet, die nur ex negativo zur Kategorie passen. Auch solche Zuordnungen wurden bereinigt, beispielsweise *swîgen* bei ›Stimmlicher Ausdruck‹. Zwar hat *swîgen* durchaus mit einem stimmlichen Ausdruck zu tun, allerdings ist der Begriff 23121200 der Aktivität ›Reden‹ zugeordnet, so dass ein Vergleich mit den Paidia-Daten ins Leere laufen würde, würde man *swîgen* beibehalten.

Die Löschliste wurde anschließend erstens nach Lemmata und zweitens nach Begriffen sortiert und in beiden Fassungen jeweils manuell auf eine konsistente Entscheidungsstrategie hin überprüft. Es ist wohl unnötig zu betonen, dass trotzdem kaum objektivierbare Kriterien für diesen Vorgang, bei dem der Zeitfaktor doch auch eine kritische Größe war, festgelegt wer-

den konnten – ähnlich wie bei der älteren Editionsphilologie basiert die Entscheidung auf dem Iudicium des Anwenders.

Ein weiteres Skript hat die Zuordnungen, die in der Löschliste notiert sind, nach den jeweiligen Wortformen in den Texten eliminiert. Zur Wortform-Lemma-Zuordnung wurde dabei eine Liste verwendet, die aus dem Export aus der MHDBDB von 2020 erzeugt werden konnte. Im Rahmen dieser Bereinigung wurde bei 2.514 Lemmata eine manuelle Entscheidung getroffen. Davon wurden bei 1.569 Lemma-Begriff-Zuordnungen Inkonsistenzen bereinigt.<sup>8</sup> Beim Abarbeiten der Löschliste wurden zudem auch solche Begriff-Lemma-Zuordnungen gelöscht, die 2018 noch in der MHD-BDB eingetragen waren, die aber 2020 in der MHDBDB nicht mehr vorhanden waren. Bislang nicht zugeordnete Lemmata wurden auch im Rahmen dieser Bereinigung nicht zugeordnet.

Schwierig bleibt vor allem die Erfassung von Begriffen, die in den Texten meist nicht direkt lexikalisch abgebildet werden, sondern die oft metaphorisch oder verhüllend umschrieben werden, etwa beim Begriff ›Sexualleben/Erotik‹ (21112000). Mit einigen Bedenken wurden u. a. die Verben *umbevâhen*, *lieben*, *triuuten* und *minnen* beibehalten, auf die Löschliste wurden u. a. übernommen: *amîs*, *amor*, *brennen*, *gebern*, *gelieben*, *geluste*, *herzeliep*, *kiusche*, *küssen*, *lieplich*, *maget*, *minneclîch*, *munt*, *reinenen*, *sanfte*, *senender*, *süeze*, *tougen*, *tragen*, *triuwe*, *trôst*, *trûtgeselle*, *vriunt*, *vrouwe*.

## 10. Das Scoring-System zur Auswertung

In dem Paidia-Aufsatz ging es darum, festzustellen, ob eine Aktivität in einem bestimmten Textsegment überhaupt vorlag oder nicht. Anschließend wurde betrachtet, welche Aktivitäten in wie vielen Segmenten vorhanden waren; mehrfache Okkurrenzen im gleichen Segment wurden nicht berücksichtigt. Ein Vorteil dieses Verfahrens war, dass die Werte zu verschiedenen Aktivitäten eher vergleichbar sind. Erwartbar ist etwa, dass

›Reden‹ deutlich häufiger vorkommt als ›Töten‹, die Unterschiede konnten durch den Segmentbezug jedoch zueinander in einem aussagekräftigen Verhältnis verglichen werden.

Anders als in der Paidia-Studie wurde hier versucht, auch quantitativ zu berücksichtigen, wie häufig welche Begriffe im Figurenkontext vorkommen. Bei der Auswertung sind mehrere Problemfelder vorhanden: Einerseits ist die Häufigkeit der Aktivitäten/Kategorien sehr unterschiedlich verteilt, andererseits sind die Texte unterschiedlich lang. Außerdem muss überlegt werden, bei welchem Abstand eine Kategorie noch dem Figurenkontext zugerechnet wird und ab welchem Abstand nicht mehr.

Für die Zuordnung zum Figurenkontext wurde ein ternäres Abstands-Scoring-System etabliert. Bei einem geringen Abstand wurde ein Score von 1,3 festgelegt, bei einem mittleren Abstand ein Score von 1,0 und bei einem größeren Abstand ein Score von 0,66. Die Scores für mittlere und große Abstände wurden v. a. für die Reihenfolge ›Figurenreferenz‹ > ›Kategorie‹ angedacht; für die umgekehrte Reihenfolge ›Kategorie‹ > ›Figurenreferenz‹ wurde ein zusätzlicher ›Rückwärts‹-Grenzwert eingeführt, der dafür sorgt, dass Kategorien nur dann einer Figur zugeordnet werden, wenn die Figur kurz nach der Kategorie genannt wird, weil davon ausgegangen wurde, dass die Zuordnung einer Kategorie zu einer Figur lange vor deren Nennung größere Anforderungen an das Hörverstehen stellt und daher wohl eher die Ausnahme ist, weil der Rezipient hier rätseln muss, auf wen sich eine Aussage bezieht.

Geeignete Parameter für einen geringen, mittleren und großen Abstand wurden in einem Abgleich mit den Ergebnisdaten aus dem Paidia-Aufsatz ermittelt, der unten beschrieben wird. Überraschenderweise haben sich ausgesprochen niedrige Parameter als geeignet erwiesen: fünf, zehn und 15 Worte Abstand; als ›Rückwärts‹-Grenzwert für Kategorien vor der Figurenreferenz wurde ein Abstand von fünf Worten verwendet. Wenn eine Kategorie innerhalb dieser Abstände im Umkreis einer Figurenreferenz auftritt, wird dies im Weiteren kurz als Kookkurrenz bezeichnet.



Jede gefundene Kookkurrenz wird zunächst mit dem jeweiligen Abstandsparameter multipliziert, also maßvoll schwächer oder stärker gewichtet: Wenn auf eine Figurenreferenz eine Kategorie im Abstand von bis zu fünf Wörtern folgt, wird diese Kookkurrenz mit Faktor 1,3 ungefähr doppelt so stark gewichtet wie bei einer Kategorie, die erst nach bis zu 15 Wörtern folgt.

Zusätzlich zu diesem Abstands-Scoring-System wird noch ein System benötigt, das die Anzahl der Kookkurrenzen berücksichtigt. Wenn man jedoch lediglich diese gewichteten Kookkurrenzen anschließend je Figur und Kategorie aufaddiert, erhält man ein stark textlängenabhängiges Maß. Ein Vergleich der Anzahl der Kookkurrenzen zwischen ›Mauritius‹ (1.784 Verse) und ›Partonopier‹ (21.784 Verse) wäre auf diesem Weg sinnlos. Wenn man andererseits diesen Wert durch die Anzahl der Wörter im jeweiligen Text oder auch durch die Anzahl der Figurenreferenzen im jeweiligen Text bzw. die Referenzen der jeweiligen Figur im jeweiligen Text teilt, erhält man Werte, die die kurzen Texte und seltene Figuren ganz erheblich begünstigen. Auch diese Quotienten würden also längenabhängig bleiben.

Ebenfalls experimentell erprobt und verworfen wurde eine Z-Wert-Normalisierung,<sup>9</sup> bei der für jede Figur der Mittelwert zu allen gewichteten Kookkurrenzen und für alle Kategorien und auf dieser Grundlage Standardabweichung und Z-Werte berechnet wurden. Grundlage für diesen Weg war die Überlegung, dass hier eine Aussage möglich sein könnte, ob innerhalb der Werte zu den Kategorien bei den jeweils gleichen Figuren einer Kategorie erhöhte oder niedrigere Werte erreicht werden. Auch hier haben sich keine überzeugenden Daten ergeben: Im ›Mauritius‹ etwa erreicht der Ehemann-Graf bei Begriff 22642000 ›Wahrheit/Irrtum‹ auf der Grundlage von nur einer Kookkurrenz bei 15 Figurenreferenzen einen höheren Wert als die Zofe, die auf zehn Kookkurrenzen bei 29 Figurenreferenzen kommt. Vermutlich stellen dabei Nullwerte bei den Kookkurrenzen einen Störfaktor dar: Wenn eine Nebenfigur viele Nullwerte aufweist, sind dort, wo zumindest nur wenige Kookkurrenzen überhaupt vor-

handen sind, die Z-Werte recht hoch und wohl auch höher als bei häufigen Figuren, bei denen bei mehr Kategorien Kookkurrenzen überhaupt gezählt werden können. Das führt dazu, dass niedrige Distanzwerte bei Nebenfiguren zu höheren Z-Werten führen können als bei den Hauptfiguren in der gleichen Kategorie.

Bei der Anzahl der Kookkurrenzen ist also ein Scoring-System notwendig, das sowohl die Problematik der unterschiedlichen Textlänge als auch der unterschiedlichen Anzahl der Figurentags (also Haupt- versus Nebenfigur) berücksichtigt. Gleichzeitig sollen aber auch einzelne Kookkurrenzen, die trotz ihrer geringen Häufigkeit recht signifikant sein können (etwa ›Tod‹ oder ›Sexualleben/Erotik‹), statistisch nicht allzu stark nivelliert werden. Das Kookkurrenzen-Anzahl-Scoring-System erhält in sieben Stufen jeweils Scores von 5–15; weiterhin bleibt der Score bei null Punkten, wenn keine einschlägige Kookkurrenz gefunden wurde. Eine einzelne Kookkurrenz erhält die niedrigste Score-Stufe von fünf Punkten; wenn zwei oder drei Kookkurrenzen vorliegen, werden zunächst sieben Punkte vergeben. Für alle häufigeren Kookkurrenzen werden Scores vergeben, die davon abhängen, ob die relative Häufigkeit der Kookkurrenzen zu einer Figur und zu einer Kategorie eher als höher oder eher als niedriger betrachtet werden kann. Um zu ermitteln, ob die relative Häufigkeit eher als höher oder als niedriger betrachtet werden kann, werden zwei Vergleichswerte gebildet: einerseits ein normalisierter Kategorien-Mittelwert  $MW-K(Kat)$ , der nicht figurespezifisch ist, und andererseits ein figurespezifischer Wert zur jeweiligen Kategorie  $F-K(Kat)$ .

$$MW-K(Kat) = \frac{AK(Kat)}{AF} \quad > \text{normalisierter Kategorien-Mittelwert, nicht figurespezifisch}$$

AK(Kat): Anzahl aller Kookkurrenzen zur jeweiligen Kategorie in allen Texten zu allen Figuren  
 AF: Anzahl aller Figurenreferenzen in allen Texten

$$F-K(Kat, Fig) = \frac{AK(Kat, Fig)}{AF(Fig)} \quad > \text{figurespezifischer Wert zur jeweiligen Kategorie}$$

AK(Kat, Fig): Anzahl der Kookkurrenzen je einer Figur zu je einer Kategorie  
 AF(Fig): Anzahl der Figurenreferenzen zu dieser Figur

Formel 1:  $MW-K(Kat)$  und  $F-K(Kat)$

Der Wert MW-K(Kat) wird für alle Kategorien separat gebildet: Die Anzahl aller Kookkurrenzen zur jeweiligen Kategorie in allen Texten zu allen Figuren wird geteilt durch die Anzahl aller Figurenreferenzen in allen Texten. Dieser Wert besagt, wie häufig bei einer Kategorie Kookkurrenzen im Durchschnitt bei allen Texten in Relation zu allen Figurenreferenzen vorliegen. Es handelt sich also um einen Kategorien-Mittelwert, der jedoch nicht in Bezug auf die Textlänge, sondern in Bezug auf die Gesamtzahl der Figurenreferenzen normalisiert wird.

Gebildet wird weiterhin der Wert F-K(Kat). Dieser Wert wird für alle Kategorien und alle Figuren separat gebildet. Er wird berechnet aus der Anzahl der Kookkurrenzen je einer Figur zu je einer Kategorie, die geteilt wird durch die Anzahl der Figurenreferenzen zu dieser Figur. Dieser Wert besagt, wie häufig bei einer Kategorie Kookkurrenzen zu einer Figur in Relation zu der Häufigkeit der Figurenreferenzen dieser Figur im jeweiligen Text vorliegen.

Berücksichtigt werden auf diesem Weg auch relativ höhere oder relativ niedrige Werte, unabhängig davon, ob es sich um eine häufige Kategorie wie »Objektbezogene Aktivität« (hoher Wert MW-K(Kat): 290,7) oder um eine Kategorie handelt, die in allen Texten eher selten ist wie etwa ›Tod‹ (niedriger Wert MW-K(Kat): 68,7). Damit es zu einem Score von zehn Punkten kommt, muss der figurespezifische Wert F-K(Kat) mindestens den allgemeinen kategorienpezifischen Mittelwert erreichen. Über die Anzahl der Figurenreferenzen wird weiterhin berücksichtigt, ob eine Figur häufig oder selten vorkommt. Verwendet werden folgende Scores:

Score=5, wenn eine Kookkurrenz vorliegt (zur jeweiligen Figur und Kategorie)

Score=7, wenn 2-3 Kookkurrenzen vorliegen

Score=9, wenn  $F-K(Kat) > MW-K(Kat) \times 0,9$

Score=10, wenn  $F-K(Kat) \geq MW-K(Kat)$

Score=11, wenn  $F-K(Kat) > MW-K(Kat) \times 1,1$

Score=13, wenn  $F-K(Kat) > MW-K(Kat) \times 1,3$

Score=15, wenn  $F-K(Kat) > MW-K(Kat) \times 1,66$

Gewählt wird jeweils der höchste Score: Wenn eine Figur bspw. nur drei Kookkurrenzen aufweist, aber nur sehr selten vorkommt, kann bei einer seltenen Kategorie der relative Figurenwert  $F\text{-}K(\text{Kat})$  dennoch über dem allgemeinen Mittelwert  $MW\text{-}K(\text{Kat})$  liegen, so dass ein höherer Score als sieben Punkte möglich ist.

Abschließend wird der Anzahl-Score mit dem Abstands-Score kombiniert: Zu den einzelnen Abstand-Scores, die zu niedrigen Abständen (fünf Worte), mittleren (zehn) oder größeren Abständen (15) vergeben wurden, wird pro Figur und Kategorie ein Abstandsmittelwert gebildet, der zwischen 0,66 und 1,33 liegen kann. Mit diesem Abstands-Score wird der Anzahl-Score (Werte von 0–15) multipliziert, so dass der endgültige Score in einem Bereich von 0–19,95 liegen kann.

Ein Beispiel: Der allgemeine kategorienspezifische Mittelwert für die Kategorie ›Tod‹  $MW\text{-}K(\text{›Tod‹})$  liegt für alle Figuren und alle Texte bei 68,7. Bei Lunete wurden elf Kookkurrenzen zu ›Tod‹ erfasst. Der figurespezifische Wert  $F\text{-}K(\text{›Tod‹})$  für Lunete liegt bei 71,9. Dieser Wert ist größer als  $MW\text{-}K(\text{›Tod‹})$ , jedoch nicht größer als  $MW\text{-}K(\text{›Tod‹}) \times 1,1$ . Der Score beträgt also zehn Punkte – noch vor der Berücksichtigung der Abstand-Scores. Die Abstand-Scores betragen für jede der elf Kookkurrenzen entweder 0,66, 1,0 oder 1,3; aus diesen elf Abstand-Scores wird ein Mittelwert gebildet. Mit diesem Abstand-Score-Mittelwert (hier 0,87) wird der Kookkurrenzen-Anzahl-Score (hier zehn Punkte) multipliziert, so dass sich als kombinierter Score für Lunete und ›Tod‹ 8,7 Punkte ergeben.

Dieses Scoring-Verfahren, das auf der Anzahl der Kookkurrenzen beruht, stellt sicher, dass die Scores eher aussagekräftig und weniger textlängenabhängig sind; sie liegen zudem in einer vergleichbaren Größenordnung.

Während bei häufigen Figuren eher zu erwarten ist, dass sie im Kontext einer breiten Palette von Kategorien gezeigt werden, ist zu erwarten, dass selten vorkommende Figuren im Kontext einer weniger breiten Kategorien-Palette gezeigt werden; mitunter können an seltene Figuren spezifische

narrative Funktionen gebunden sein. Da in die Gewichtung auch die Häufigkeit der Referenzen auf die jeweilige Figur eingeht, ist das Scoring-Verfahren nicht vollständig längenunabhängig: Es weist eine leichte Tendenz dazu auf, seltenen Figuren bei den vorhandenen Kategorien höhere Werte zuzuerkennen als häufig vorkommenden Figuren. Diese Plausibilitätsüberlegung stimmt damit überein, dass etwa der Graf und die Gräfin im ›Mauritius‹ mitunter etwas höhere Scores aufweisen als vergleichbare Figuren in den Romanen.

## 11. Auswertung: Einzelne Kategorien und Figuren

Folgende Figuren wurden untersucht:

Text	Zofe	Hauptfigur ♂	Hauptfigur ♀	Opponenten
›Partonopier‹	Irekel	Partonopier	Meliur	Mutter, Sornagiur, Mareis
›Iwein‹	Lunete	Iwein	Laudine	Ascalon, Harpin, Truchsess
›Tristan‹	Brangaene	Tristan	Isolde	Marke, Irischer Truchsess, Marjodo
›Mauritius v.C.‹	Zofe-MvC	Mauritius	Graefin-MvC	Graf-MvC

Tabelle 2: Texte und Figuren

Die ausführlichen Ergebnis-Daten sind im Dariah-Repository unter [doi:10.20375/0000-000F-322E-6](https://doi.org/10.20375/0000-000F-322E-6) zu finden. Dort stehen zwei Tabellen: Erstens eine Tabelle für alle Figuren, zweitens eine Tabelle, die eine nach Aktanten gruppierte Ausgabe bietet.

Die Scores pro Kategorie und Figur sind dann mit einem (+) bzw. (-) markiert, wenn der Figurescore bei der jeweiligen Kategorie oberhalb oder unterhalb des Konfidenzintervalls (bei einem Konfidenzniveau von 80%) liegt – hervorgehoben werden also bereits mäßig über- oder unter-

durchschnittliche Werte. Auf Spitzenwerte kommt es hier zunächst nicht an; im Paidia-Aufsatz ging es nur darum, ob eine Aktivität in einem Segment überhaupt vorhanden war oder nicht. Bei Aktivitäten, die aus mehreren Kategorien bestehen, wird hier angegeben, bei wie vielen dieser einzelnen Kategorien sich solche erhöhten Werte für eine Figur ergeben – beispielsweise also, ob bei fünf von zehn Kategorien, die zu einer Aktivität gehören, bei einer Figur Werte oberhalb des Konfidenzintervalls vorhanden sind. Bei diesem Analyseschritt geht jedoch nicht ein, ob erhöhte Werte für eine Figur nur marginal oder ganz erheblich über dem Konfidenzintervall liegen.

Ein zweiter Analyseschritt bezieht die Höhe der Scores mit ein. Wenn eine Aktivität aus mehreren Kategorien besteht, wird der Mittelwert der Scores zu all diesen Kategorien gebildet. Dieses Verfahren ist nicht ganz unproblematisch: Wenn etwa aus den Mittelwerten der zwölf Kategorien, die der Aktivität ›Kämpfen‹ zugeordnet sind, wiederum ein Mittelwert gebildet wird, könnte damit theoretisch nivelliert werden, wenn eine Figur zwar ein enormes Kampfpensum absolviert, das sich etwa bei ›Zweikampf‹ dokumentiert, diesen hohen Wert aber durch niedrige Werte etwa bei ›Rangordnung‹ einbüßt. In der Praxis haben sich jedoch meist plausible Werte ergeben, die mehr oder weniger ähnliche Tendenzen dokumentieren wie die Betrachtung der Anzahl der Überschreitungen des Konfidenzintervalls.

Aktivität	alle <sup>10</sup>	Zofen	Hauptfigur ♂	Hauptfigur ♀	Opponenten
Essen	5,7	<b>6,4</b>	<b>10,2</b>	4,7	4,1
Feiern	2,3	1,5	<b>4,3</b>	<b>3,2</b>	1,5
Jagen	3,1	1,5	<b>5,6</b>	<b>3,8</b>	2,5
Schlafen	9,1	<b>10,2</b>	<b>10,1</b>	<b>10,2</b>	7,8
Helfen	11,4	<b>13,6</b>	9,9	<b>11,4</b>	11,0

Tabelle 3: Höhe der Scores je Aktantengruppe, Teil 1 (Summenwerte aller Kategorien, die einer Aktivität zugeordnet sind)

Essen: Der Aktivität ›Essen‹ sind vier Kategorien zugeordnet. Bei jeder Figur können sich also bis zu vier Werte ergeben, die oberhalb des Konfidenzintervalls liegen. Bei den folgenden Figuren ist mehr als eine Kategorie mit solcherart erhöhtem Wert vorhanden: Irekel (2), Partonopier (3), Meliur (2), Partonopiers Mutter (2), Mareis (2), Lunete (2), Iwein (3), Harpin (3), Brangäne (2), Tristan (3), Marke (2), Mauritius (3) – also vorwiegend männliche Hauptfiguren, Zofen und einzelne Opponenten.

In den Mittelwerten (Tabelle 3) würden demnach die Opponenten noch seltener als die weiblichen Hauptfiguren im Essens-Kontext genannt, obwohl einzelne Opponenten wie Marke und Partonopiers Mutter wohl aufgrund ihrer zentralen Stellung am Hof überdurchschnittliche Werte aufweisen.

Feiern: Der Aktivität ›Feiern‹ sind 13 Kategorien zugeordnet. Bei jeder Figur können sich also bis zu 13 Werte ergeben, die oberhalb des Konfidenzintervalls liegen. Bei den folgenden Figuren sind bei mehr als einer Kategorie erhöhte Werte vorhanden: Partonopier (6), Meliur (3), Sornagiur (5), Iwein (2), Brangäne (3), Tristan (7), Isolde (6), Marke (5), Irischer Truchsess (2), Mauritius (2). Feiern bzw. kulturelle Aktivitäten werden also insbesondere im Umkreis von männlichen Hauptfiguren vorgefunden; zudem sind zwei Könige (Sornagiur, Marke) beteiligt, deren Hofhaltung ebenfalls höfische Züge trägt. Die weiblichen Hauptfiguren sind nur teilweise in der Nähe solcher Kategorien angesiedelt – bei Isoldes hohem Wert darf man sofort an das gemeinsame Musizieren mit Tristan denken. Der leicht erhöhte Wert beim irischen Truchsess könnte mit seiner Präsenz am Königshof zusammenhängen.

In den Mittelwerten der 13 Kategorien (Tabelle 3) haben männliche und weibliche Hauptfiguren erhöhte Werte. Dass die weiblichen Hauptfiguren bei höfischen Festivitäten präsent sind, ist plausibel; Zofen werden bei solchen Gelegenheiten im Durchschnitt ebenso selten erwähnt wie die Opponenten – dies könnte gut mit der Fokusführung korrespondieren, die eher

Hauptfiguren als ihre Gegner bei diesen Aktivitäten zeigt. Insgesamt liegen damit ähnliche Tendenzen vor wie bei ›Essen‹ – mit der Ausnahme allerdings, dass Zofen weniger bei Festen als im Ernährungskontext genannt werden. Dies würde gut zu ihrer dienenden Funktion passen – man denke an Lunete, die den im Torraum gefangenen Iwein erst einmal mit Nahrung versorgt. – Eine ähnliche Tendenz wie bei ›Feiern‹ findet sich bei ›Jagen‹.

Schlafen: Hier gibt es bei den Aktanten-Mittelwerten nur geringe Unterschiede bei den Zofen, den männlichen und weiblichen Hauptfiguren; die Werte der Opponenten sind deutlich niedriger.

MW_alle	MW_Zofen	MW_mHF	MW_wHF	MW_Opp
9,1	10,2	10,1	10,2	7,8

Bei 23133000 (›Hilfeleistung‹) finden sich Werte oberhalb des Konfidenzintervalls bei Partonopiers Mutter, die immerhin meint, ihrem Sohn Gutes zu tun, sowie bei Sornagiur, der für die Freilassung von Partonopier sorgt, bei Lunete, Laudine, Brangäne sowie bei allen vier Figuren im ›Mauritius‹. Wie im Paidia-Aufsatz fallen in Tabelle 4 erwartungsgemäß die höheren Werte für die Zofen ins Auge. Die männlichen Hauptfiguren schneiden, denkt man an Iweins *helfe*-Aventiuren, überraschend schlecht ab, die weiblichen Hauptfiguren liegen nah am Mittelwert, die Opponenten knapp darunter.

Aktivität	alle	Zofen	Hauptfigur ♂	Hauptfigur ♀	Opponenten
Helfen	11,4	13,6	9,9	11,4	11,0
Hilfsmittel gebrauchen	7,8	7,1	9,7	8,4	7,1
Kämpfen	5,5	4,1	7,2	5,0	5,5
Turnieren	4,4	2,8	6,9	3,4	4,4

Tabelle 4: Höhe der Scores je Aktantengruppe, Teil 2 (Summenwerte aller Kategorien, die einer Aktivität zugeordnet sind)



Die drei Kategorien zur Aktivität ›Hilfsmittel gebrauchen‹ beziehen sich recht allgemein auf den Gebrauch von Gegenständen. Im Mittelwert haben die männlichen und weiblichen Hauptfiguren höhere Werte als die Opponenten und Zofen (Tabelle 4).

Der Aktivität ›Kämpfen‹ sind zwölf Kategorien zugeordnet. Mehr als ein Wert oberhalb des Konfidenzintervalls findet sich bei Irekel (3), Partonopier (8), Meliur (2), Sornagiur (8), Mareis (6), Iwein (5), Ascalon (4), Harpin (4), Laudines Truchsess (4), Brangäne (2), Isolde (2), Irischer Truchsess (7), Mauritius (3), Gräfin und Graf MvC (2). Erwartungsgemäß sind die Zofen und die weiblichen Hauptfiguren hier schwächer vertreten als die männlichen Hauptfiguren und die Opponenten; überraschend hoch ist der Wert des irischen Truchsesses, dessen Drachenkampf parodistische Züge trägt, und der Umstand, dass Tristan nur einen erhöhten Score aufweist. Der der mittlere Opponenten-Score (Tabelle 4) würde höher liegen, würde man Partonopiers Mutter (2,0) als weibliche Figur herausrechnen; auch Marjodo (2,4) zeichnet sich bekanntlich mehr durch Intrigen aus denn durch ritterliche Bewährung.

Ähnlich verhält es sich bei ›Turnieren‹. Dieser Aktivität sind fünf Kategorien zugeordnet. Mehr als ein erhöhter Wert findet sich bei Irekel (2), Partonopier (4), Sornagiur (3), Mareis (2), Iwein (2), Harpin (2), Laudines Truchsess (2), Irischer Truchsess (3), Mauritius (2), Gräfin MvC (1). Der niedrige Wert bei Tristan (1) könnte damit zusammenhängen, dass Tristan häufiger in der Nähe von anderen Aktivitäten erwähnt wird.

Die Rangfolge der Mittelwerte in Tabelle 4 ist analog zum Befund bei ›Kämpfen‹: Zofen haben deutlich unterdurchschnittliche Scores, weibliche Hauptfiguren leicht unterdurchschnittliche. Bei Opponenten finden sich durchschnittliche und bei männlichen Hauptfiguren deutlich überdurchschnittliche Werte.

Bei 31800000 (›Bewegung‹) haben erwartungsgemäß die weiblichen Hauptfiguren als eher statische<sup>11</sup> Figuren das Nachsehen gegenüber den übrigen Figuren; auffällig (Tabelle 5) ist, dass die Zofen in diesen Texten

noch häufiger im Bewegungskontext gezeigt werden als die männlichen Hauptfiguren. Während im Paidia-Aufsatz höhere Werte bei Isolde beobachtet wurden, weist hier die Gräfin im ›Mauritius‹ höhere Scores auf – womöglich spielt hier das Berechnungsmodell eine Rolle, das hier eher selten auftretende Figuren leicht bevorzugt.

Aktivität	alle	Zofen	Hauptfigur ♂	Hauptfigur ♀	Opponenten
Bewegung	10,6	12,4	11,0	8,9	10,5
Töten	10,1	8,3	9,4	9,1	11,5
Peinigen	7,4	8,1	10,9	7,1	5,8
Lügen	9,5	10,8	8,4	11,1	8,8
Reden etc.	9,8	11,5	9,7	10,4	8,9
Kopulieren	5,9	3,5	4,9	9,5	5,8

Tabelle 5: Höhe der Scores je Aktantengruppe, Teil 3 (Summenwerte aller Kategorien, die einer Aktivität zugeordnet sind)

›Töten‹ (21104000, ›Tod‹) ist erwartungsgemäß weniger Frauen- als Männersache, wobei auch die männlichen Hauptfiguren in Tabelle 5 unterhalb des Mittelwerts bleiben: Im Wesentlichen ist der Begriff an Opponenten gekoppelt.

›Peinigen‹ hat sich nur notdürftig zu der Kategorie 22705200 (›Ärger/Zorn/Wut‹) zuordnen lassen. Hier ist der niedrige Wert der Opponenten auf den ersten Blick kontraintuitiv. Denkbar ist jedoch, dass ›Ärger/Zorn/Wut‹ ausführlich bei den männlichen Hauptfiguren zur Darstellung kommt, auch wenn Ursache dafür andere Figuren waren, deren Referenz weiter zurückliegt. Dass die Darstellung von Leiderfahrungen bei den Opponenten nicht allzu ausgiebig erfolgt, ist wiederum erwartbar, wenn man eine Sympathiesteuerung voraussetzen darf, die eher die Hauptfiguren als die Opponenten begünstigt.

›Lügen‹ bzw. 22642000 (›Wahrheit/Irrtum‹) findet trotz des höheren Scores von Tristan im Kontext von männlichen Hauptfiguren unterdurch-

schnittlich statt. Die Daten in Tabelle 5 deuten eine genderspezifische Verteilung an – da Zofen und weibliche Hauptfiguren häufiger sprachlich handeln (s. u.), greifen sie womöglich häufiger zur Lüge. Allerdings kann bei den Opponenten ein breites Spektrum beobachtet werden: Während starke oder brutale Opponenten wie Ascalon oder Harpin sehr offen und direkt ihre Vorhaben ankündigen, findet man beim irischen Truchsess und bei Marjodo im ›Tristan‹ besonders hohe Werte.

Gruppiert man die zehn Kategorien, die ›Reden‹, ›Argumentieren‹, ›Veranlassen‹ oder ›Beklagen‹ zugeordnet sind, findet sich mit Ausnahme von Ascalon und Laudines Truchsess bei allen Figuren mehr als ein Wert oberhalb des Konfidenzintervalls. Der Mittelwert für alle Figurenwerte beträgt 9,8. Überdurchschnittliche Werte weisen die Zofen auf – gefolgt von den weiblichen Hauptfiguren (Tabelle 5). Minimal unterdurchschnittlich sind die männlichen Hauptfiguren, deutlich unterdurchschnittlich die Opponenten.

Im Paidia-Aufsatz wurde der überraschende Befund beobachtet, dass die Opponenten beim sprachlichen Handeln sehr hohe Werte aufweisen. Dieser Befund könnte auf dem Berechnungsverfahren beruhen, das dort nur berücksichtigt, ob eine Figur in einem Segment die Aktivität überhaupt ausübt, während hier die Häufigkeit der Aktivität stärker berücksichtigt wird. Hier bleiben Opponenten nunmehr unter dem Mittelwert.

Während die bislang vorstellten Daten zumindest im Großen und Ganzen nicht in Widerspruch zum Lektüreeindruck menschlicher Rezipienten stehen, zeigen die Mittelwerte der beiden Kategorien 21112000 und 21045000 (›Sexualleben/Erotik‹ und ›Geschlechtsmerkmale/Genitalien/Zeugung/Geburt‹) teilweise unplausible Befunde in Tabelle 5. Unter den Zofen erfährt man nur bei Brangäne explizit von einem Beilager in Markes Hochzeitsnacht. Dass die Opponenten höhere Werte aufweisen als die männlichen Hauptfiguren, wird nicht einleuchten; der Befund wird auch nicht allein etwa durch Harpins Drohung verständlich, der die

Tochter von Gaweins Schwager zur Frau nehmen oder sie durch seine Knechte vergewaltigen lassen will.

Abschließend sei deshalb die Problematik der Wortform-basierten Analyse mit einem detaillierteren Blick auf die Kategorie 21112000 (>Sexualleben/Erotik<) vorgeführt. Der Mittelwert für alle Figuren und alle Texte beträgt 7,27. Werte oberhalb des Konfidenzintervalls weisen auf: Meliur (13,1), Lunete (9,1), Laudine (9,1), Harpin (13,0), Brangäne (12,2), Tristan (11,6), Isolde (14,3), Marke (16,0), der irische Truchsess (13,1) und Marjodo (11,0). Bei Meliur, Laudine, Brangäne, Tristan, Isolde und Marke handelt es sich um erwartbare Befunde. In Widerspruch zu einer manuellen Erfassung der Aktivität >Kopulieren< stehen die Werte von Lunete, Harpin, dem irischen Truchsess und Marjodo.

Zu einem erhöhten Score führen bei Lunete vier Nennungen im Kontext der Wortformen *entwunge* (8089), *kuster* (7976), *beruorte* (5384) und *nacket* (3238); bei Harpin die Wortform *ruote* (5058), beim irischen Truchsess die Wortformen *damoysele* (9165), *gewaltesaere* (11027), *schranken* (11254) und *understan* (11050), bei Marjodo *geminne* (13467). Zwar unterdurchschnittliche, dennoch aber fragwürdige Werte größer null haben Irekel, Partonopiers Mutter und Sornagiur, die auf den Wortformen *umbevungen* (8827), *minnedieben* (11430), *triute* (6647), *minnetranc* (6957), *geminnet* (6905), *muotgelust* (5893) und *grans* (5489) beruhen. Während zu *nacket* und *umbevâhen* bei der Kategorienbereinigung die zwar auch diskussionswürdige Zuordnung zu Kategorie 21111200 beibehalten wurde, sind die anderen genannten Wortformen so selten im Korpus, dass sie bei der Kategorienbereinigung nicht überprüft worden sind.

Diese Zuordnungen sind ein gutes Beispiel für das Verfahren der Kategorienzuordnung in der MHDBDB: ein Verfahren, das suchende Benutzer im Zweifelsfall lieber zu einem Suchtreffer führt, als potentielle Suchtreffer bei unsicherer Relevanz zu übergehen. Die Berührung im >Iwein< (*beruorte* V. 5384) gilt den Panzerringen des Truchsessens durch den Löwen; Lunete wird im nächsten Vers erwähnt. Zur unsicheren

Zuordnung im Einzelfall kommt auch hier das Problem hinzu, dass Figuren nur zufällig im Kontext der Wortform stehen können, dass die Aktivität aber von anderen vollführt wird.

Unerwartet niedrige Werte finden sich bei Partonopier, bei Mauritius und der Gräfin im ›Mauritius‹. Hier zeigen sich die Grenzen von lexikalischen Zugriffen auf Phänomene, die oft umschreibend, verhüllend oder gar mit Unsagarkeitstopos auf der Textoberfläche realisiert werden. Im ›Mauritius‹ heißt es: *nû begunde ouch er erwarmen. / unde tete der vrouwen ichn weiz waz. / waz hulfe ez, sagete ich iu daz? ez ist ungesaget alsô guot / ir wizzet wol waz man dâ tuot.* (V. 1614–1618. ›Nun wurde auch ihm immer wärmer. Ich weiß nicht, was er mit der Dame anfang. Was würde es helfen, wenn ich Euch das sagen würde? Wenn das ungesagt bleibt, ist es genauso gut. Ihr wisst ja gut, was man da so macht.‹) Das erste Beilager im ›Tristan‹ ist metaphorisch umschrieben: *ietwederez schancte unde tranc / die süeze, diu von herzen gie. / sô sî die state gewonnen ie, / sô gie der wehsel under in / slîchende her unde hin* (V. 12042–12046. ›Jeder von ihnen schenkte und trank die Süße, die vom Herzen kam. Wann immer sie die Möglichkeit dazu finden konnten, so ging dieser Warentausch zwischen ihnen leise hin und her.‹)

Je weniger unmittelbar ein Phänomen auf der Textoberfläche adressiert wird, umso schwieriger wird es auch künftig bleiben, dieses Phänomen mit automatischen Verfahren zu erfassen. Bei der Beilager-Thematik ist dieses Problem besonders virulent, weil es sich erstens um ein Phänomen handelt, das im höfischen Roman nicht offen adressiert wird, und zweitens um eines, das etwa im Artusroman im Vergleich zu Aktivitäten wie ›Kämpfen‹ oder ›Reden‹ seltener vorkommt. Bei häufigeren Phänomenen ist eher zu erwarten, dass direkte Benennungen zumindest auch neben verhüllenden Umschreibungen stehen, so dass computerbasierte Verfahren weniger stark von spezifischen Einzelformulierungen abhängig sind, die manche Autoren gerade bei dieser Aktivität besonders elaboriert ausgestalten – man denke etwa an Wolframs *hirzwurz*-Metapher (›Parzival‹, 643,28).

## 12. Figurentypen mit Delta unterscheiden

In einem weiteren Analyseschritt werden alle Scores zu den 62 Kategorien in ein Delta-Verfahren eingespeist. Das Verfahren wird typischerweise für Studien zur Autorschaftsattribuion verwendet und gruppiert hierbei basierend auf Wortfrequenzen Texte nach Autoren. Für die vorliegende Untersuchung werden analog dazu die Figuren in Bezug auf ihre Kategorien in Figurencluster gruppiert bzw. voneinander unterschieden.<sup>12</sup>

Erwartungsgemäß weist die Zofe zur weiblichen Hauptfigur in ›Wein‹, ›Partonopier‹ und ›Mauritius‹ die größte Nähe auf (blau-rote Paare). Im ›Tristan‹ sind die Ähnlichkeiten von Tristan und Isolde derart groß, dass sie sogar stärker ins Gewicht fallen als die Gender-Differenz etwa zu Brangäne; eventuell könnte dieser Befund damit korrespondieren, dass Tristan und Isolde in vielen Szenen gleichzeitig im Erzählfokus stehen.

Die Opponenten Laudines Truchsess, Ascalon, Harpin, irischer Truchsess, Marjodo, Mareis und Sornagiur (schwarz in Abb. 1) weisen erhebliche Distanzen zu weiblicher Hauptfigur und Zofe auf – diese eher typischen Opponenten lassen sich also gruppieren.

Opponenten wie Marke und der Graf im ›Mauritius‹, die zugleich auch Ehemänner der weiblichen Hauptfigur sind, clustern relativ nahe zur weiblichen Hauptfigur; im grafischen Befund spiegelt sich damit, dass der Opponententyp ›Ehemänner‹ von anderen Opponenten wie Harpin oder Mareis unterschieden werden kann.

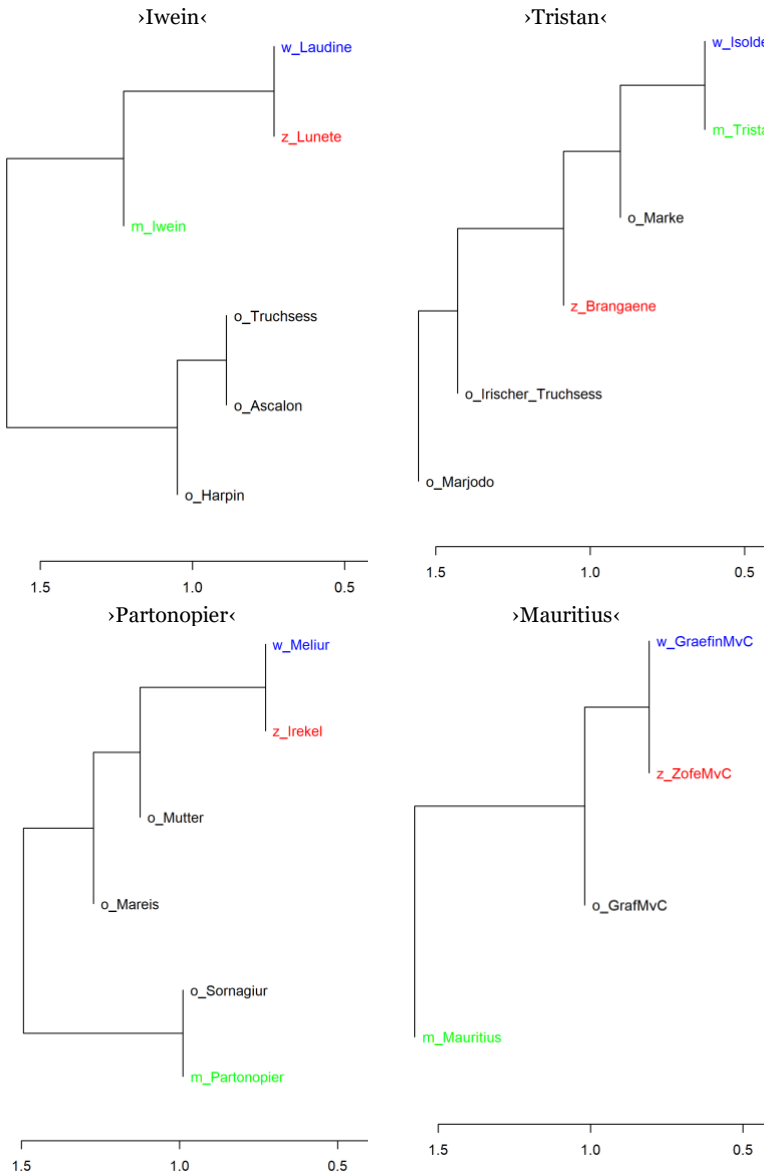


Abb. 1: Delta-Cluster zu den Aktanten

Bei den männlichen Hauptfiguren (grün) clustern Iwein und Tristan erwartungsgemäß nahe bzw. sehr nahe zur weiblichen Hauptfigur. Bei Partonopier überrascht auf den ersten Blick seine Nähe zum Opponenten Sornagiur. Allerdings ist Sornagiur kein klassischer Antagonist – nach dem Verrat von Mareis bewirkt er die Freilassung von Partonopier; die Gegnerschaft am Beginn wird schließlich in ein freundschaftliches Verhältnis überführt. In Bezug auf die Motive Tabubruch und Verrat kann Sornagiur auch als Komplementärfigur zu Partonopier aufgefasst werden (vgl. Dimpel 2015a, S. 63–67). Ein Delta-Plot ohne Sornagiur lokalisiert Partonopier wie etwa im ›Iwein‹ an den gleichen Ast, an dem auch die Zofe und weibliche Hauptfigur hängen; die beiden verbleibenden Opponenten sind deutlich von den übrigen Figuren getrennt:

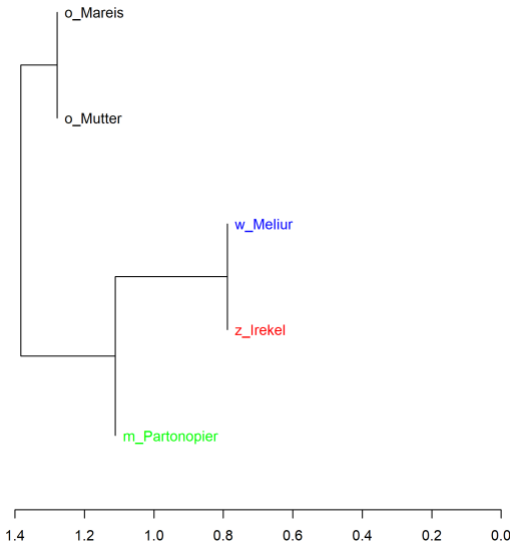


Abb. 2: Delta-Cluster zu den Aktanten im ›Partonopier‹ ohne Sornagiur

Leider sind im ›Mauritius‹ die Turniergegner derart marginal dargestellt, dass nur der Graf als singulärer Opponent, der noch dazu Ehemann der



Gräfin ist, in die Analyse eingehen kann; der Delta-Graph ist deshalb nur eingeschränkt mit den anderen Texten vergleichbar.

Jenseits der Figurenarmut des ›Mauritius‹ und der Einordnung von Tristan und Partonopier bei Isolde bzw. Sornagiur, die durch textspezifische Besonderheiten erklärbar ist, gelingt eine Erfassung der verschiedenen Figurentypen angesichts der niedrigen Zahl von nur 62 Textmerkmalen, die in das Delta-Verfahren eingehen, überraschend gut.

### **13. Abgleich mit der manuellen Annotation – Äpfel mit Birnen vergleichen?**

Nachdem die automatischen Verfahren erfreulich plausible Ergebnisse hervorbringen, sollte abschließend überprüft werden, wie hoch der Grad der Übereinstimmung mit den manuell annotierten Daten aus dem Paidia-Aufsatz ist. Immer dann, wenn automatische Verfahren einer manuell gefundenen Lösung nahekommen sollen, träumt man von einem F1-Wert von 1,0 – also von einem perfekten Recall bei perfekter Precision. Ein solches Ziel wäre bei dem vorliegenden Versuch von vornherein utopisch – die zahlreichen Gründe dafür seien hier kurz skizziert:

- Bereits die manuelle Annotation ist nicht vollständig objektiv, weil nicht alle Aktivitäten tatsächlich explizit in den Texten vorliegen.
- Die Zuordnung der Aktivitäten zu Kategorien der MHDBDB müsste perfekt passen; tatsächlich wurden die Beschreibung der Aktivitäten und die Konzeption der Kategorien der MHDBDB für ganz unterschiedliche Zwecke erstellt, so dass eine Zuordnung von Kategorien zu Aktivitäten nur tentativ möglich ist. So zielte etwa die Paidia-Aktivität ›Hilfsmittel‹ auf zentrale Gegenstände wie Ringe, Zaubergürtel oder Salben. In den zugeordneten MHDBDB-Kategorien sind Lemmata zu Gegenständen im weiteren Sinn (Werkzeuge bzw. gegenstandsgebundene Tätigkeiten jeder Art) enthalten. Auch die Unterschiede etwa

- zwischen ›Reden‹, ›Argumentieren‹ und ›Veranlassen‹ sind eine erhebliche Herausforderung für automatische Verfahren.
- Die Lemmata, die einer Kategorie zugeordnet sind, müssen diese Kategorie auch tatsächlich beschreiben; andere Bedeutungen bei Polysemie dürfen nicht überwiegen. Im Idealfall dürften Wortformen nur dann einer Kategorie zugerechnet werden, wenn sichergestellt ist, dass die im Text verwendete Bedeutung tatsächlich der Kategorie entspricht. Polysemie kann vereinzelt den Recall etwas verbessern, sie verursacht jedoch sehr häufig eine deutlich schlechtere Precision.
  - Die Wortform sollte tatsächlich auf die Figur bezogen sein. Text im Umkreis von Figurenreferenzen kann jedoch auch auf andere Figuren, auf Gegenstände oder abstrakte Entitäten bezogen sein. Systematische Fehler sind zu erwarten im Zusammenhang mit Aussagen in Negationen oder Konjunktiv II.
  - Im Paidia-Aufsatz wurden Aktivitäten-Zuordnungen zu Figuren in Analepsen nicht im Slot der eigentlichen erzählten Zeit erfasst, sondern sie wurden dem Segment zugeordnet, dem die nachgetragene Information eigentlich angehört. Für wichtige Analepsen wurden teilweise eigene Segmente gebildet. Hier werden Kookkurrenzen an der Textposition gezählt, an der sie vorgefunden werden.
  - Einige Kategorien involvieren mehrere Figuren. Während in einem Beispielsatz wie ›Iwein kämpft mit Harpin‹ die Zuordnung von ›Kämpfen‹ zu Iwein und Harpin gleichermaßen korrekt ist, wäre in einem Beispielsatz wie ›Lunete hilft Iwein‹ zwar bei manueller Annotation evident, dass die Aktivität ›Helfen‹ nicht Iwein, sondern Lunete eingeschrieben ist. Bei einer automatischen Erfassung von Kategorien, die die fünf Worte vor und die 15 Worte nach einer Figurenreferenz einbezieht, führen solche Fälle jedoch zwangsläufig dazu, dass es sich bei der Hälfte der Treffer um False-Positives handelt; entsprechend ist mit einer niedrigen Precision zu rechnen.

- In der Paidia-Studie wurden Häufigkeitswerte nicht in gleicher Weise wie hier gebildet. Annotiert wurde, ob in einem Textsegment eine Aktivität überhaupt vorhanden ist oder nicht – ob eine Aktivität in einem Segment einmal oder häufig vorkommt, wurde nicht berücksichtigt. Die Anzahl der Segmente, in der die jeweilige Aktivität vorlag, wurde dort durch die Zahl der Segmente geteilt. Dieses Verfahren führt dazu, dass Figuren, die nur in wenigen Segmenten überhaupt vorkommen, tendenziell schlechtere Werte erzielen. Hier wurde ein komplexes Scoring-System entwickelt, das von solchen Effekten weniger stark betroffen ist, aber leicht dazu neigt, selten vorkommende Figuren zu bevorzugen.

### 13.1 Segment-Vergleich

Im ersten Analyseversuch hat das Phänomen, das im letzten Spiegelstrich beschrieben wurde, keine Rolle gespielt: Ein Skript liest zunächst die Daten aus der Paidia-Studie dazu ein, in welchem Textsegment zu welcher Figur die jeweilige Aktivität vorhanden ist oder nicht. In einem zweiten Abschnitt werden die mit Figurenreferenzen und Kategorien annotierten Dateien eingelesen – sowie die Mapping-Tabelle, die die Zuordnungen von Kategorien zu Aktivitäten enthält. Überprüft wird nun, ob in den Segmenten, in dem zu einer Figur mindestens eine Aktivität vorhanden ist, auch mindestens eine zugeordnete Kategorie im Umkreis von fünf, zehn, 15 oder 20 Worten vor und nach einer Figurenreferenz vorhanden ist (hier ohne Rückwärtsschwelle). Hier die Ergebnisse:

Spannweite	Recall	Precision	F1
5	0,688	0,337	0,452
10	0,813	0,300	0,438
15	0,858	0,283	0,425
20	0,875	0,270	0,412

Tabelle 6: Vergleich mit manueller Annotation

Diese Werte sind zwar erheblich von den idealen Werten entfernt; gemessen an den hier skizzierten Problemen handelt es sich jedoch um ein Ergebnis, das in Relation zu dem Aufwand, der in dieser Studie möglich war, noch akzeptabel ist. Mit steigender Spannweite nimmt die Precision ab; der Recall steigt zwar etwas, aber nicht in einem so großen Ausmaß, dass dadurch bei größeren Intervallen das Absinken der Precision kompensiert werden könnte.

Auch mit Blick auf das Zeitbudget wurden in der Annotationsphase pronominale Referenzen auf eine Figur nur dann erfasst, wenn eine Figurenreferenz in den zehn Versen zuvor nicht bereits annotiert wurde. Retrospektiv hat sich diese Entscheidung als recht unglücklich herausgestellt, da sich Spannweiten, die 15 Worte übersteigen, als ungeeignet erwiesen haben; am Beginn dieser Studie stand die Annahme, dass man mit deutlich größeren Spannweiten operieren könnte. Hier die Einzelwerte bei Spannweite = 10:

	Recall	Precision	F1
a_Hilfsmittel	0,970	0,204	0,337
a_bewegen	0,879	0,548	<b>0,675</b>
a_essen	0,600	0,200	0,300
a_feiern	0,143	0,022	<b>0,038</b>
a_helfen	0,711	0,274	0,395
a_jagen	0,500	0,075	<b>0,131</b>
a_kopulieren	0,577	0,313	0,405
a_kämpfen	0,900	0,135	0,235
a_schlafen	0,636	0,103	<b>0,177</b>
a_turnieren	0,875	0,137	0,237
a_töten	0,700	0,080	<b>0,144</b>
b_argumentieren	0,500	0,327	0,396
b_beklagen	0,925	0,343	<b>0,500</b>
b_lügen	0,714	0,215	0,331
b_reden	0,906	0,783	<b>0,840</b>
b_veranlassen	0,935	0,274	0,424

Tabelle 7: Vergleich mit manueller Annotation; Einzelwerte

Während ›Bewegen‹, ›Beklagen‹ und ›Reden‹ offenbar recht zuverlässig erkannt werden können, ist die automatische Erfassung von ›Feiern‹, ›Ja-

gen«, »Schlafen« und »Töten« mit dem hier verwendeten Verfahren offenbar besonders fehleranfällig.

### 13.2 Vergleich von erhöhten Werten

Während im vorstehenden Abschnitt das gemeinsame Vorhandensein oder Fehlen einer Aktivität bzw. Kategorie im gleichen Segment untersucht wurde, wird hier verglichen, ob die Werte bzw. Scores für die Aktivitäten und die ihnen zugeordneten Kategorien jeweils erhöhte Werte aufweisen.

Bei den Aktivitäten geht dabei ein, ob die Aktivität einer Figur oberhalb des Konfidenzintervalls bei einem Konfidenzniveau von 80% liegt.<sup>13</sup> Da im Paidia-Aufsatz nur das Vorhandensein einer Aktivität in einem Segment berücksichtigt wurde, bleiben Unterschreitungen des Konfidenzintervalls hier unberücksichtigt. Wenn mehrere Kategorien einer Aktivität zugeordnet sind, genügt es für einen positiven Vergleich, wenn eine der Kategorien der jeweiligen Figur einen erhöhten Score oberhalb des Konfidenzintervalls aufweist: So, wie bei Erfassung der Aktivität »Kämpfen« nicht spezifiziert wurde, ob es sich um »Zweikampf« oder »Heerfahrt« handelt, genügt für den positiven Abgleich mit der Aktivität »Kämpfen«, wenn eine dieser Kategorien einen erhöhten Score aufweist.

Auch auf diesem Weg wurden letztlich Äpfel mit Birnen verglichen, so dass sich ähnliche Vergleichswerte ergeben:

Recall	Precision	F1
0,681	0,325	0,440

### 14. Bilanz und Ausblick

Wichtiger als die Analyseergebnisse war es, mit dieser Studie zunächst Konzepte zu erarbeiten, um semantische Kategorien im Kontext von Figurenreferenzen untersuchen zu können. Neben der Annotation von Figurenreferenzen und der Übertragung der Daten aus dem Begriffssystem der

MHDBDB in die digitalen Texte wurde ein Scoring-System zur Auswertung solcher Kookkurrenzen entwickelt.

Unter den Analyse-Ergebnissen waren viele Befunde, die sich gut mit intuitiven interpretativen Annahmen parallelisieren lassen – etwa Befunde zu mittelalterlichen Gender-Konzepten: Kämpfen oder Turnieren erfolgt eher im Kontext von männlichen Figuren, mündliche Kommunikation eher bei weiblichen Figuren. Kulturelle Aspekte wie Feiern oder Jagen sind eher an Hauptfiguren gekoppelt – eventuell eine Konsequenz der Fokusführung. Weibliche Hauptfiguren, die mitunter als statische Figuren betrachtet werden, kommen selten im Kontext von Bewegungsvokabular vor. Zunächst überraschende Befunde werden meist plausibel, wenn man betrachtet, welche einzelnen Figuren durch ihre Spezifika zu den überraschenden Befunden beitragen. Zugleich werden aber auch die Grenzen des Verfahrens offenkundig – etwa bei verhüllender Redeweise bei Sexualität.

Die Delta-Plots zu allen 62 Kategorien weisen meist eine Nähe von Zofe und weiblicher Hauptfigur aus, viele typische Opponenten lassen sich gut gruppieren, während spezielle Opponenten (wie Ehemänner der weiblichen Hauptfigur) sich von den übrigen Opponenten abgrenzen lassen.

Eine interessante Herausforderung bestand darin, Figuren und Texte zu analysieren, die im Rahmen der Paidia-Studie bereits mittels manueller Annotation unter zumindest ähnlichen Gesichtspunkten untersucht worden sind. Angesichts der erheblichen konzeptionellen Differenzen zwischen beiden Studien ist es nicht möglich, die Paidia-Daten als Messlatte für die Qualität der hier automatisiert erhobenen Daten zu verwenden – es ist eher erstaunlich, dass der Vergleich überhaupt einen F1-Wert von 0,45 erreichen kann. Vergleicht man allerdings die zentralen Tendenzen wie höhere oder niedrigere Werte von bestimmten Aktantengruppen (etwa: weniger ›Bewegung‹ bei weiblichen Hauptfiguren als bei den anderen Aktanten; vgl. oben, Abschnitte 1 und 11), so zeigt sich, dass die Ergebnisse der Paidia-Studie mit Ausnahme der Aktivität ›Sexualleben/Erotik‹ in hohem Maße mit den aktuellen Ergebnissen übereinstimmen.

Folgestudien können sich von der ursprünglichen Konzeption der Paidia-Studie freimachen und weitere Kategorien, weitere Figurentypen und weitere Texte einbeziehen. Zu fragen bleibt, ob eventuell nicht nur Figurentypen, sondern auch Gattungs- oder Epochenspezifika auszumachen wären. Dafür wäre es künftig wichtig, alle Figurenreferenzen auch bei pronominalen Referenzen zu annotieren, und die Kategorien-Bereinigung auch für Wortform-Häufigkeiten bereits ab einer niedrigeren Okkurrenzschwelle vorzunehmen. Zudem wäre eine Auflösung von Polysemien erfreulich. Zumindest bei höherfrequenten Wortformen sollte es machbar sein, typische von seltenen Bedeutungen im Korpus unterscheiden zu können.

## 15. Was braucht das Fach?

Die digitale Literaturanalyse hat in den letzten 20 Jahren enorme Fortschritte gemacht. In der Mediävistik sind die Fortschritte allerdings viel kleiner als in den neueren Philologien. Ein möglicher Grund dafür ist, dass sich die computeraffinen Mediävisten überwiegend um Editionen und Handschriften verdient machen. Ein anderer Grund ist unser Ressourcenproblem: Immerhin sind allmählich etwas mehr Texte digital frei verfügbar, doch sind wir noch weit davon entfernt, dass zumindest alle zentralen Texte in aktuellen Ausgaben digital zur Verfügung stehen. Dass bei der DFG die Pflicht zur digitalen Publikation von Editionen nur eine Soll- und keine Muss-Bestimmung ist, ist angesichts der öffentlichen Finanzierung dieser Editionen ein unhaltbarer Zustand.

Mit den vorhandenen digitalen Texten sind Studien zur Autorschafts-attribution, zu Gattungen und wortschatzbezogenen Fragen möglich. Wir können heute eine hohe Wahrscheinlichkeit dafür vorrechnen, dass die ›Halbe Birne‹ von Konrad von Würzburg ist und dass die Nürnberger Weingrüße von Rosenplüt sind (Dimpel [u. a.] 2019; Dimpel/Wagner 2022).

Die hier vorgestellte Studie ist ein Beispiel dafür, welche Möglichkeiten bereits bestehen, wenn man auch die Figurenebene in den Blick nehmen will. Es werden aber vor allem auch unsere Ressourcenprobleme sichtbar. Gerade die digitale Literaturanalyse kommt nicht mit reinen Textdateien aus: Man braucht Anreicherungen, etwa eine vollständige Annotation von Figurenreferenzen, die auch Pronomina erfasst. Solange keine Texte mit entsprechenden Anreicherungen verfügbar sind, bleibt die Automatisierbarkeit von figurenbezogenen Studien schwierig, was bedauerlich ist, da die Figur eine doch recht zentrale Kategorie darstellt.

Hätte man Syntax-Parser, könnte man Aussagen in Negation und Konjunktiv II identifizieren.<sup>14</sup> In der vorliegenden Studie könnte man damit etwa erfassen, ob Figurenprofile plausibler werden, wenn man negierte oder unsichere Kategorien im Figurenkontext übergeht. Für andere Studien bräuchte man die Annotation von Orten oder weitere narratologische Annotationen.

Das ist umso wichtiger, als wir gegenüber der Neugermanistik und der Anglistik massiv in Nachteil sind: Dort gibt es gute Parser und andere Tools, von denen Mediävist\*innen nur träumen können. Andererseits sind unsere zentralen Texte längst nicht so zahlreich, sodass es höchste Zeit ist, großangelegte Annotationsprojekte anzugehen: Wir brauchen eine breite Grundlage, damit man auch in der Mediävistik vernünftig digital arbeiten kann.

Weiterhin wäre es erfreulich, wenn gesprochene und gedachte Figurenrede einbezogen werden könnte. Dann wären auch Auswertungen denkbar, die einen höheren Grad an Sicherheit haben könnten: Welche Figur erwähnt welche Kategorie? Solche Annotationen wären auch für viele andere Zwecke hilfreich – SNA-Studien arbeiten beispielsweise häufig mit Figurenreferenzen und Figurenrede.

Für manche Zwecke benötigt man auch handschriftennahe Texte. Aber für quantifizierende und vergleichende literaturwissenschaftliche Studien sind normalisierte Texte notwendig. Und ideal wäre es, wenn alle digitalen



Texte konsequent nach den gleichen Regeln normalisiert wären; auch für die Parser-Entwicklung wäre das ein Vorteil.

Diese Ressourcen-Frage ist auch eine Frage der Nachwuchsförderung. W3-Professores können es sich eher leisten, Hilfskraftstunden in ein Projekt zu investieren, wenn man eine angereicherte Ausgangsbasis braucht. Im DH-Bereich ist aber überwiegend der wissenschaftliche Nachwuchs am Werk, der nicht über solche Ressourcen verfügt. Daher ist es essenziell, die Verfügbarmachung von Textkorpora sowie die Entwicklung geeigneter Tools in der digitalen Mediävistik voranzutreiben. Wenn hier keine wesentlichen Fortschritte stattfinden, wird sich der Rückstand der Mediävistik auf die neueren Philologien bei der digitalen Literaturanalyse noch weiter vergrößern.

Im Rahmen der Tagungsdiskussion hat ein Kollege bedauert, dass digitale Studien vielfach erwartbare Befunde vorstellen würden. Es wurde die Ansicht vertreten, dass Literaturanalyse wohl weiterhin Handarbeit bleiben würde. Eine solche Skepsis hat bereits Jannidis (1999) beobachtet:

Quantitative Verfahren werden in der Literaturwissenschaft nicht immer gern gesehen. Bestätigen sie gängige Einsichten, stehen sie im Verdacht, überflüssig zu sein. Widersprechen sie aber den üblichen Ansichten, schafft man sie sich mit dem leisen Hinweis vom Hals, daß man Statistiken ohnehin nicht trauen könne.

Mit Blick auf manche traditionellen Literaturwissenschaftler mag es auch 2022 noch hilfreich sein, wenn man Bekanntes bestätigen kann, um demonstrieren zu können, dass man mit digitalen Methoden verlässliche Ergebnisse hervorbringen kann, so dass man wenigstens nicht dem Verdikt verfällt, man würde fragwürdige Statistiken erzeugen. Doch eigentlich hat die digitale Literaturanalyse einige Fortschritte erlebt: Das Bestätigen von Bekanntem kann man mit Eibl (2013, S. 37), auch positiv als »Kontrollpeilung« einstufen. Vielfach wird (wie in der vorliegenden Studie) eine Methode zuerst anhand von bekanntem Material evaluiert, bevor man sie

in einem größeren Rahmen (hier: Korpus und weitere Kategorien) einsetzt. Andrea Rapp hat in der Diskussion darauf hingewiesen, dass man in den Naturwissenschaften geradezu froh ist, wenn man mit verschiedenen Methoden zu den gleichen Befunden kommt und so eine unabhängige Bestätigung findet.

Wenn man konventionelle Fragestellungen so modelliert, dass sie digital implementierbar werden, wird offensichtlich, dass man viel kleinschrittiger und präziser vorgehen muss als bei konventionellen oder narratologischen Beschreibungen: »Die kleinste Lücke, Ambiguität oder gar Widersprüchlichkeit im narratologischen Modell führt, sobald sie sich in der Software niederschlägt, zum Disaster.« (Meister 2013, S. 294; vgl. auch McCarty 2005, S. 46.) Insofern sind die Modellerzeugung und die Modellweiterentwicklung eine wesentliche Leistung der digitalen Literaturanalyse (zum »Modelling« vgl. McCarty 2005, S. 20–72); oft werden dabei Voraussetzungen und Vorannahmen präziser benannt als in konventionellen Studien. So wird in der vorliegenden Studie nicht die Aktivität ›Kämpfen‹ erfasst, sondern es werden ausgewählte lexikalische Referenzen erfasst. Die Lemmata, die den einschlägigen MHDBDB-Begriffen zugeordnet sind, realisieren dabei eine graduelle Annäherung zur Aktivität ›Kämpfen‹ (zum Begriff »proxy« als Stellvertretung für die gesuchten Phänomene bei digitalen Operationalisierungen vgl. etwa Moretti 2013, S. 2–5). Dabei ist anzumerken, dass viele traditionelle Studien keine Rechenschaft darüber ablegen, welches Konzept eigentlich gemeint ist, wenn sie über die Aktivität ›Kämpfen‹ schreiben. Dennoch bleibt es ein Vorzug einer digitalen Studie, dass nicht nur an ihrem Beginn interessegeleitete Fragestellungen und Modellierungen klar benannt werden und dass an ihrem Ende transparente Interpretationen von Ergebnisdaten stehen, sondern dass es dazwischen auch empirische Teilstrecken gibt (zum Computerphilologen als »Teilzeitempiriker« vgl. Dimpel 2015b, S. 349–354).

## Anmerkungen

- 1 Vgl. hierzu z. B. Ketschik [u. a.] 2020; Blessing [u. a.] 2017. Zu CRETA vgl. <https://www.creta.uni-stuttgart.de/>.
- 2 Beim ›Mauritius von Craun‹ folgt der Datenbankabzug der Pretzel-Ausgabe, abgeglichen wurde er mit einem digitalen Text, der auf der älteren Ausgabe von Edward Schröder beruht. Gegenüber dem Schröder-Text fehlten in dem Datenbankabzug knapp 20 Verse.
- 3 Vier Figuren im Falle des ›Mauritius‹.
- 4 Für das Vorannotieren wurden Passagen gewählt, in denen alle Figuren einigermaßen häufig auftreten: Partonopier: V. 5623-6134, 7435-7535, 7755-7908, 8595-9114; Tristan: V. 9365-9896, 12507-12674, 13451-13672.
- 5 Für die Annotationsphase wurden bei 14h pro Monat neun Monate benötigt.
- 6 Bei einem Plausibilitätscheck der exportierten Daten hat sich ergeben, dass seit dem Export der Texte im Jahr 2018 durch die fortlaufende Arbeit an der Datenbank einige Begriff-Lemma-Zuordnungen in der MHDDB entfernt worden sind.
- 7 Das Verfahren ist nicht vollständig zuverlässig, da Wortformen zu verschiedenen Lemmata gehören können.
- 8 Die MHDDB strebt eine Disambiguierung der Bedeutungen im Kontext an; wenn künftig disambiguierte Texte vorliegen, werden solche Zuordnungen weniger Probleme bereiten. Ebenfalls bislang nicht gelöst ist das Problem, dass eine Wortform zu mehreren Lemmata zugeordnet sein kann. Da die Bereinigung auf Lemma-Ebene und nicht auf Wortform-Ebene erfolgt, wurden Zuordnungen bei allen Wortformen entfernt, wenn ein zugehöriges Lemma in der Löschliste enthalten ist.
- 9 Vgl. etwa zur Z-Wert-Begrenzung bei Autorschaftsattributionsstudien Evert [u. a.] 2016, S. 62–64.
- 10 Im Gesamtmittelwert »MW-alle« für alle Figuren sind die Opponenten stärker repräsentiert als die anderen Aktantengruppen, da mit Ausnahme des ›Mauritius‹ drei Opponenten pro Text untersucht wurden, jedoch nur eine Zofe sowie eine männliche und weibliche Hauptfigur. »MW-alle« entspricht also nicht einem Mittelwert der Mittelwerte der Aktantengruppen (MW\_Zofen, MW\_mHF, ...), da es sich um den Mittelwert aller Einzelwerte handelt.
- 11 Zu dynamischen vs. statischen Figuren vgl. Lotman 1981.
- 12 Für die Delta-Analyse und die Cluster-Grafik wurde das stylo-R-Paket von Eder [u. a.] 2017 verwendet. Übergeben wurden die hier ermittelten Scores als ›table\_with\_frequencies.txt‹. Zu Delta bei Autorschaftsfragen vgl. etwa

Burrows 2003, Büttner [u. a.] 2017. Damit stylo-R den Figuren, die zum gleichen Aktantentyp gehören, die gleiche Farbe zuordnet, wurde den Figurenbezeichnungen hier ein z\_, m\_, w\_ oder o\_ vorangestellt.

- 13 Diese Schwelle wurde etwas niedriger gewählt als die Schwelle im Paidia-Aufsatz, damit eine breitere Vergleichsbasis vorliegt, in die auch leicht erhöhte Werte eingehen. Die im Paidia-Aufsatz mit + bzw. – markierten Werte wurden nur innerhalb der jeweiligen Aktantengruppe berechnet; hier werden die Abweichungen in Bezug auf alle Figuren berechnet.
- 14 Der Stuttgarter POS-Tagger für das Mittelhochdeutsche erreicht bei der Wortartenbestimmung eine Erkennungsgenauigkeit von  $F1=0,82$  sowie eine Accuracy von bis zu 0,91. Vgl. Echelmeyer [u. a.] 2017; Schulz/Ketschik 2019. Für eine Anwendung im Kontext quantitativer literaturwissenschaftlicher Studien wäre eine Steigerung der Erkennungsgenauigkeit wünschenswert. Ein erstes Experiment zu einem Tagger, der auch Flexionsformen bestimmt, hat Helmut Schmid (LMU) im Workshop »Automatische Annotation digitaler Editionen« (Stuttgart, 16./17.3.2022) vorgestellt; eine Evaluation auf einer ausreichenden Datenmenge steht noch aus.

## Literaturverzeichnis

### Primärliteratur

- Gottfried von Strassburg: Tristan und Isold, hrsg. von Friedrich Ranke, Berlin 1968.  
Hartmann von Aue: Iwein. Eine Erzählung von Hartmann von Aue, hrsg. von G.F. Benecke und Karl Lachmann, neu bearb. von Ludwig Wolff, siebente Ausgabe, Berlin 1968.  
Konrad von Würzburg: Partonopier und Meliur, hrsg. von Karl Bartsch, Berlin 1970.  
Moritz von Craûn, hrsg. von Ulrich Pretzel, 4. Aufl., Tübingen 1973 (ATB 45).  
Wolfram von Eschenbach: Parzival. Studienausgabe, 2. Aufl. Mittelhochdeutscher Text nach der sechsten Ausgabe von Karl Lachmann, Übersetzung von Peter Knecht. Mit Einführung zum Text der Lachmannschen Ausgabe und in Probleme der ›Parzival‹-Interpretation von Bernd Schirok, Berlin/New York 2003.

### Sekundärliteratur

- Agarwal, Apoorv/Corvalan, Augusto/Jensen, Jacob/Rambow, Owen: Social Network Analysis of ›Alice in Wonderland‹, in: Workshop on Computational Linguistics for Literature. Montréal, Canada, June 8, 2012, Montréal 2012, S. 88–96 ([online](#)).

- Agarwal, Apoorv/Rambow, Owen: Automatic Detection and Classification of Social Events, in: Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing, Massachusetts 2010, S. 1024–1034 ([online](#)).
- Blessing, André/Echelmeyer, Nora/John, Markus/Reiter, Nils: An End-to-end Environment for Research Question-Driven Entity Extraction and Network Analysis, in: Proceedings of the Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature, Vancouver, Canada 2017 ([online](#)).
- Braun, Manuel/Ketschik, Nora: Soziale Netzwerkanalysen zum mittelhochdeutschen Artusroman – oder: Vorgreiflicher Versuch, Märchenhaftigkeit des Erzählens zu messen, in: Das Mittelalter 24 (2019), S. 54–70 ([online](#)).
- Büttner, Andreas/Dimpel, Friedrich Michael/Evert, Stefan/Jannidis, Fotis/Pielström, Steffen/Proisl, Thomas/Reger, Isabella: »Delta« in der stilometrischen Autorschaftsattribuion, in: Zeitschrift für digitale Geisteswissenschaften (2017) ([online](#)).
- Burrows, John F.: Questions of Authorship: Attribution and Beyond. A Lecture Delivered on the Occasion of the Roberto Busa Award, ACH-ALLC 2001, New York, in: Computers and the Humanities 37 (2003), S. 5–32 ([online](#)).
- Dimpel, Friedrich Michael: Wertungsübertragungen und korrelative Sinnstiftung im ›Herzog Ernst B‹ und im ›Partonopier‹, in: DVjs 89 (2015a), S. 41–69.
- Dimpel, Friedrich Michael: Der Computerphilologe als Interpret – ein Teilzeit-Empiriker? in: Borkowski, Jan/Descher, Stefan/Ferder, Felicitas/Heine, Philipp David (Hrsg.): Literatur interpretieren: Interdisziplinäre Beiträge zur Theorie und Praxis, Münster 2015b, S. 339–359.
- Dimpel, Friedrich Michael: Novellenschätze narratologisch auszeichnen und analysieren am Beispiel Victor von Scheffels ›Hugideo‹ und der sozialen Netzwerkanalyse, in: Weitin, Thomas/Werber, Niels (Hrsg.): Scalable Reading, Siegen 2017 (LiLi 47), S. 87–108 ([online](#)).
- Dimpel, Friedrich Michael: Versuch einer quantitativen Analyse von Figurenaktivitäten in ›Iwein‹, ›Tristan‹, ›Partonopier‹ und ›Mauritius von Craun‹ in Analogie zu Computerspielen, in: Ascher, Franziska/Müller, Thomas (Hrsg.): Paidia Sonderausgabe: Vom ›Wigalois‹ zum ›Witcher‹ – Mediävistische Zugänge zum Computerspiel, 2018 ([online](#)).
- Dimpel, Friedrich Michael/Schlager, Daniel/Zeppezauer-Wachauer, Katharina: Der Streit um die Birne. Autorschafts-Attributionstest mit Burrows' Delta und dessen Optimierung für Kurztexte am Beispiel der ›Halben Birne‹ des Konrad von Würzburg, in: Bleier, Roman/Fischer, Franz/Hiltmann, Torsten/Viehauser, Gabriel/Vogeler, Georg (Hrsg.): Digitale Mediävistik, 2019 (Das Mittelalter. Perspektiven mediävistischer Forschung. Zeitschrift des Mediävistenverbandes, Band 24), S. 71–90.

- Dimpel, Friedrich Michael: Soziale Netzwerkanalyse und Erzählschemata. Eine explorative Vorstudie, in: Ernst, Marlene/Hinkelmanns, Peter/Zangerl, Lina Maria/Zeppezauer-Wachauer, Katharina (Hrsg.): *digital humanities austria 2018. empowering researchers*, Wien 2020, S. 95–111 ([online](#)).
- Dimpel, Friedrich Michael/Wagner, Silvan: Rosenplüt als Autor der Nürnberger Weingröße. Philologische und computerphilologische Analysen, in: Jurchen, Sylvia/Wagner, Silvan (Hrsg.): *Schlechtes Wetter und Grenzüberschreitungen*, Oldenburg 2022 (Zeitschrift *Brevitas* 2, BmE Sonderheft) ([online](#)).
- Echelmeyer, Nora/Reiter, Nils/Schulz, Sarah: Ein PoS-Tagger für »das« Mittelhochdeutsche, in: *Konferenzabstracts DHd 2017: Digitale Nachhaltigkeit*, Bern 2017, S. 141–147 ([online](#)).
- Eder, Maciej/Rybicki, Jan/Kestemont, Mike: *stylo* R package, in: 2017 ([online](#)).
- Eibl, Karl: Ist Literaturwissenschaft als Erfahrungswissenschaft möglich? Mit einigen Anmerkungen zur Wissenschaftsphilosophie des Wiener Kreises, in: Ajouri, Philip/Mellmann, Katja/Rauen, Christoph (Hrsg.): *Empirie in der Literaturwissenschaft*, Münster 2013 (Poetogenesis. Studien zur empirischen Anthropologie der Literatur 8), S. 19–45.
- Elson, David K./Dames, Nicholas/McKeown, Kathleen R.: Extraction Social Networks from Literary Fiction, in: *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, Uppsala 2010, S. 138–147.
- Evert, Stefan [u. a.]: Burrows' Delta verstehen, in: Burr, Elisabeth (Hrsg.): *Konferenzabstracts DHd 2016. Modellierung – Vernetzung – Visualisierung. Die Digital Humanities als fächerübergreifendes Forschungsparadigma*, Leipzig 2016, S. 62–65 ([online](#)).
- Hallig, Rudolf/Wartburg, Walther von: *Begriffssystem als Grundlage für die Lexikographie. Versuch eines Ordnungsschemas*, Berlin 1963.
- Jannidis, Fotis: Was ist Computerphilologie? in: *Jahrbuch für Computerphilologie* 1 (1999), S. 39–60 ([online](#)).
- Jannidis, Fotis: *Figur und Person. Beitrag zu einer historischen Narratologie*, Berlin/ New York 2004 (Narratologia 3).
- Ketschik, Nora/Blessing, André/Murr, Sandra/Overbeck, Maximilian/Pichler, Axel: Interdisziplinäre Annotation von Entitätenreferenzen, in: Reiter, Nils/Pichler, Axel/Kuhn, Jonas (Hrsg.): *Reflektierte algorithmische Textanalyse. Interdisziplinäre(s) Arbeiten in der CRETA-Werkstatt*, Berlin/Boston 2020, S. 203–263 ([online](#)).
- Krautter, Benjamin/Pagel, Janis/Reiter, Nils/Willand, Marcus: »[E]in Vater, dächte ich, ist doch immer ein Vater«. Figurentypen im Drama und ihre Operationalisierung, in: *ZfdG* (2020) ([online](#)).
- Lexer, Matthias: *Mittelhochdeutsches Handwörterbuch*. 3 Bde., Leipzig 1872–1878. Nachdr. Stuttgart 1992 ([online](#)).

- Lotman, Jurij M.: Die Entstehung des Sujets – typologisch gesehen, in: Ders. (Hrsg.): Kunst als Sprache. Untersuchungen zum Zeichencharakter von Literatur und Kunst, Stuttgart 1981, S. 175–204.
- McCarty, Willard: Modeling: A Study in Words and Meanings, in: Schreibman, Susan/Siemens, Ray/Unsworth, John (Hrsg.): A Companion to Digital Humanities, Oxford 2004, S. 254–270 ([online](#)).
- McCarty, Willard: Humanities Computing, London; New York 2005.
- Meister, Jan Christoph: Computerphilologie vs. Digital Text Studies. Von der pragmatischen zur methodologischen Perspektive auf die Digitalisierung der Literaturwissenschaften, in: Grond-Rigler, Christine/Straub, Wolfgang (Hrsg.): Literatur und Digitalisierung, Berlin/Boston 2013, S. 267–296.
- Moretti, Franco: »Operationalizing«: or, the function of measurement in modern literary theory, Stanford 2013 (Pamphlets of the Stanford Literary Lab 6).
- Pagel, Janis/Reiter, Nils: GerDraCor-Coref: A Coreference Corpus for Dramatic Texts in German, in: Proceedings of the 12th Language Resources and Evaluation Conference (LREC), Marseille 2020.
- Roget, Peter Mark: Thesaurus Of English Words And Phrases. Classified and arranged so as to facilitate the expression of ideas and assist in literary composition. Unter Mitarbeit von Barnas Sears. Revised and Edited, Boston 1864.
- Schulz, Sarah/Ketschik, Nora: From 0 to 10 million annotated words: part-of-speech tagging for Middle High German, in: Language Resources and Evaluation 53 (2019), S. 837–863 ([online](#)).
- Trilcke, Peer: Social Network Analysis (SNA) als Methode einer textempirischen Literaturwissenschaft, in: Ajouri, Philip/Mellmann, Katja/Rauen, Christoph (Hrsg.): Empirie in der Literaturwissenschaft, Münster 2013 (Poetogenesis. Studien zur empirischen Anthropologie der Literatur 8), S. 201–247.
- Viehhauser, Gabriel/Barth, Florian: Towards a Digital Narratology of Space, in: Digital Humanities 2017. Conference Abstracts. McGill University & Université de Montréal. Montréal, Canada 2017.
- Wiedmer, Nathalie/Pagel, Janis/Reiter, Nils: Semi-Automatische Extraktion von Beziehungen zwischen dramatischen Figuren, in: DHd 2020. Spielräume: Digital Humanities zwischen Modellierung und Interpretation. Konferenzabstracts, Paderborn 2020, S. 194–200 ([online](#)).

### Online-Ressourcen

Ausführliche Ergebnis-Daten zu diesem Beitrag im Dariah-Repositorium:  
<https://dx.doi.org/10.20375/0000-000F-322E-6>.

CRETA (Center For Reflected Text Analytics): <https://www.creta.uni-stuttgart.de/>,  
<https://www.creta.uni-stuttgart.de/tools/index.html>.

CRETAnno [web-basierte Website zur Annotation, in Arbeit]:

<http://hdl.handle.net/11022/1007-0000-0007-E1BE-5>.

CRETAnno Annotationsrichtlinien: [https://www.creta.uni-stuttgart.de/wp-content/uploads/2016/09/Annotationsrichtlinienv1\\_1.pdf](https://www.creta.uni-stuttgart.de/wp-content/uploads/2016/09/Annotationsrichtlinienv1_1.pdf).

dlina (Digital Literary Network Analysis): <https://dlina.github.io/>.

MHDBDB (Mittelhochdeutsche Begriffsdatenbank): <http://mhdbdb.sbg.ac.at/>.

NER (Named Entity Recognition): [https://fortext.net/routinen/methoden/named-entity-recognition-ner](https://fortext.net/routinen/methoden/named-entity-recognition-ner;);

Stanford NER-Modelle: <https://nlp.stanford.edu/software/CRF-NER.html>.

OED (Historical Thesaurus of the Oxford English Dictionary):

<https://www.oed.com/thesaurus/>.

QuaDramA-Projekts (QuaDramA: Quantitative Drama Analytics):

<https://www.ims.uni-stuttgart.de/forschung/projekte/quadrama/>.

## **Anschrift der Autorinnen und Autoren:**

Prof. Dr. Friedrich Michael Dimpel

Friedrich-Alexander-Universität Erlangen-Nürnberg

Bismarckstraße 1

91054 Erlangen

E-Mail: [mail@dimpel.de](mailto:mail@dimpel.de)

ORCID: 0000-0003-4833-4897

GND: 1057584525

Andre Blessing

Universität Stuttgart

Institut für Maschinelle Sprachverarbeitung

Pfaffenwaldring 5b

70197 Stuttgart

E-Mail: [andre.blessing@ims.uni-stuttgart.de](mailto:andre.blessing@ims.uni-stuttgart.de)

ORCID: <https://orcid.org/0000-0001-7573-578X>

GND: <https://d-nb.info/gnd/1058601865>

Peter Hinkelmanns

Paris-Lodron-Universität Salzburg

Mittelhochdeutsche Begriffsdatenbank MHDBDB



Erzabt-Klotz-Straße 1  
A-5020 Salzburg  
E-Mail: [peter.hinkelmanns@plus.ac.at](mailto:peter.hinkelmanns@plus.ac.at)  
ORCID: <https://orcid.org/0000-0001-8618-0185>

Nora Ketschik  
Universität Stuttgart  
Institut für Maschinelle Sprachverarbeitung  
Pfaffenwaldring 5b  
70569 Stuttgart  
E-Mail: [nora.ketschik@ilw.uni-stuttgart.de](mailto:nora.ketschik@ilw.uni-stuttgart.de)  
ORCID-ID: 0000-0001-8758-5432

Dr. Katharina Zeppezauer-Wachauer  
Paris-Lodron-Universität Salzburg  
Mittelhochdeutsche Begriffsdatenbank MHDBDB  
Erzabt-Klotz-Straße 1  
A-5020 Salzburg  
E-Mail: [Katharina.Wachauer@plus.ac.at](mailto:Katharina.Wachauer@plus.ac.at)  
ORCID: 0000-0001-9310-9029  
GND: 1140611232