

Zeitschrift für digitale Geisteswissenschaften

Artikel aus:

Sonderband 1 der ZfdG: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum und Thomas Stäcker. 2015. DOI: [10.17175/sb01](https://doi.org/10.17175/sb01)

Titel:

Historische Stilometrie? Methodische Vorschläge für eine Annäherung textanalytischer Zugänge an die mediävistische Textualitätsdebatte.

Autor/in:

Gabriel Viehhauser

Kontakt:

viehhaus@germ.unibe.ch

Institution:

Universität Bern

GND:

[139919783](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-39919783)

ORCID:

[0000-0001-6372-0337](https://orcid.org/0000-0001-6372-0337)

DOI des Artikels:

[10.17175/sb001_009](https://doi.org/10.17175/sb001_009)

Nachweis im OPAC der Herzog August Bibliothek:

[83016927X](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-83016927X)

Erstveröffentlichung:

19.02.2015

Lizenz:

Sofern nicht anders angegeben 

Medienlizenzen:

Medienrechte liegen bei den Autoren

Letzte Überprüfung aller Verweise:

24.05.2016

GND-Verschlagwortung:

[Mediävistik](#) | [literarischer Stil](#) | [Textanalyse](#) |

Zitierweise:

Gabriel Viehhauser: Historische Stilometrie? Methodische Vorschläge für eine Annäherung textanalytischer Zugänge an die mediävistische Textualitätsdebatte.. In: Grenzen und Möglichkeiten der Digital Humanities. Hg. von Constanze Baum / Thomas Stäcker. 2015 (= Sonderband der Zeitschrift für digitale Geisteswissenschaften, 1). PDF Format ohne Paginierung. Als text/html abrufbar unter DOI: [10.17175/sb001_009](https://doi.org/10.17175/sb001_009).

Gabriel Viehhauser

Historische Stilometrie? Methodische Vorschläge für eine Annäherung textanalytischer Zugänge an die mediävistische Textualitätsdebatte.

Abstracts

Da die mediävistische Textualitätsdebatte von einer deutlichen Zurückhaltung gegenüber der Instanz des Autors geprägt ist, stehen die Voraussetzungen für eine Akzeptanz des Einsatzes stilometrischer Methoden auf dem Gebiet der mittelhochdeutschen Literatur nicht günstig. Doch eröffnet gerade das Spannungsfeld von Tradition und Individualität, das für diese Texte von Bedeutung ist, der Stilometrie Anwendungsmöglichkeiten, die zu einer gegenseitigen Erhellung von Methode und Fachdisziplin führen können. Der Artikel zeigt anhand von Beispielanalysen, wie mittelhochdeutsche Texte sowohl durch individuelle als auch durch gattungsspezifische Stilfaktoren bestimmt sind und was bei der Quantifizierung dieser Kategorien zu beachten ist.

The debate on medieval textuality is shaped by an obvious reservation about the category of authorship. Therefore, the conditions for the acceptance of stylometric methods in the field of Middle High German literature might not seem too favorable. However, it is exactly this tension between tradition and individuality, which bears great significance for those texts, that offers opportunities for the application of stylometric methods, which can lead to a mutual elucidation of method and discipline. The paper provides examples of the ways in which Middle High German texts are influenced by individual as well as genre-specific style factors. Furthermore, it elaborates on the issues that must be addressed for a proper quantification of these categories.

1. Zur Problematik des Stilbegriffs in der Mediävistik

Im Rahmen der Digital Humanities stellt die Stilometrie, also die Anwendung quantitativer Methoden zur Erfassung und Klassifizierung stilistischer Merkmale von Texten, eine jener Unterdisziplinen dar, die zur Zeit eine gesteigerte Aufmerksamkeit der Forschung für sich beanspruchen kann. Das Potential und die Beliebtheit der Methode belegt nicht zuletzt eine große Zahl von Publikationen, die sich auch für den deutschsprachigen Raum nachweisen lässt.¹ Aufgrund ihrer Aktualität bietet sich die Stilometrie demnach als Untersuchungsfeld geradezu an, wollte man, wie dies im Folgenden geschehen soll, exemplarisch der Frage nachgehen, ob die Digital Humanities einen methodischen Brückenschlag zu den traditionellen Geisteswissenschaften ermöglichen können oder aber die Gefahr einer »feindlichen Übernahme« der Geisteswissenschaften durch die Computerphilologie besteht.² Was die Methode diesbezüglich vor allem interessant macht, ist der Umstand, dass sie in letzter Zeit offenkundig auch in der traditionellen Literaturwissenschaft verstärkt wahrgenommen wird. Dies scheint zumindest für die germanistische Mediävistik zuzutreffen, dem Fachgebiet,

¹ Ungeachtet ihrer derzeitigen Konjunktur handelt es sich bei der Stilometrie freilich nicht um eine gänzlich neue Methode. Einen Überblick über die ältere Forschung bietet Schöch 2014, S. 133f.

² Ich greife damit die vom Verband [Digital Humanities im deutschsprachigen Raum](#) (DHD) auf der Passauer Jahrestagung ausgegebene Grundfrage auf.

dem ich mich näher zuwenden will. Als symptomatisch für das Interesse kann hier etwa die Aufnahme eines einschlägigen Aufsatzes von Mike Kestemont in die traditionsreiche *Zeitschrift für deutsches Altertum* gewertet werden³ oder der programmatische Artikel von Manuel Braun in der letztjährigen Ausgabe der *Zeitschrift für Literaturwissenschaft und Linguistik*, die unter dem Titel »Turn, Turn, Turn? Oder: Braucht die Germanistik eine germanistische Wende?« den gegenwärtigen Stand der Fachdisziplin reflektiert.⁴

Fast sieht es also so aus, als könnte sich nach der mittlerweile recht gut etablierten digitalen Editorik die nächste Teildisziplin der Digital Humanities anschicken, ihren Platz auf dem Feld der althergebrachten Germanistik zu erobern. Jedoch stehen die Voraussetzungen für eine Akzeptanz der Stilometrie wesentlich ungünstiger, als dies bei der Editionswissenschaft der Fall war. Denn rückblickend lässt sich feststellen, dass die Etablierung der digitalen Editorik nicht zuletzt deswegen gelingen konnte, weil sich die Methode in vorzüglicher Weise zur Beantwortung von Fragestellungen eignete, die bereits im Fokus der literaturwissenschaftlichen Diskussion standen: Die Möglichkeiten einer dynamischen, von Autorpositionen freigehaltenen Betrachtungsweise von Texten trafen sich mit den Anforderungen einer stark poststrukturalistisch beeinflussten Philologie, die Konzepte von Subjekt, Autorschaft und Textfestigkeit zunehmend in Frage stellte.⁵ Selbst nach dem Abklingen der heftig geführten Debatte um die sogenannte *New Philology* und der partiellen Rücknahme allzu radikaler Extrempositionen ist nicht von der Hand zu weisen, dass die digitale Darstellung gerade für mittelalterliche Texte besondere Vorteile bietet, da diese nur handschriftlich und in den allermeisten Fällen nur in späteren Abschriften, nicht aber in einem vom Autor selbst angefertigten Original überliefert sind.⁶ Auch wenn man aufgrund dieser Umstände die Bedeutung von Autorschaft nicht gleich völlig in Abrede stellen möchte, wie dies poststrukturalistische Extrempositionen tun, so bleibt doch der Befund, dass das Autororiginal in der Überlieferung nicht mehr direkt greifbar ist. Statt nun dessen ungeachtet gezwungenermaßen nur einen, möglicherweise sogar hypothetisch rekonstruierten ›Original-Text bieten zu müssen, wie dies früher bei den nur gedruckt vorliegenden Editionen der Fall war, lässt sich mittels einer digitalen Editionsdarstellung die gesamte handschriftliche Überlieferung mit einbeziehen. Dies ermöglicht den Blick auf eine dynamische Entwicklungsgeschichte des Textes im Verlauf seiner Tradierung, die nicht allein durch den Autor bestimmt wird, wodurch aber grundlegende theoretische Fragen wie jene, was die Integrität eines Textes oder die Rolle des Autors überhaupt ausmachen, nicht nur neu angegangen werden können, sondern sich regelrecht aufdrängen.

Demgegenüber müssen die primären Anwendungsgebiete der Stilometrie aus der Sicht der traditionellen Literaturwissenschaft zumindest auf den ersten Blick deutlich rückwärtsgewandt, wenn nicht gar altmodisch anmuten. Nur allzu leicht etwa könnten stilistische Untersuchungen

³ Kestemont 2013, S. 1–33.

⁴ Braun 2013, S. 83–90. Braun beschäftigt sich in dem Artikel mit den Möglichkeiten computerunterstützter Zugangsweisen in der Literaturwissenschaft insgesamt. Sein Fazit, dass der Einfluss der Digital Humanities zu einer Wiederannäherung von Sprach- und Literaturwissenschaft führen kann, betrifft auch die Stilometrie im engeren Sinn.

⁵ Von grundlegendem Einfluss waren hier etwa Foucault 1988 sowie Cerquiglini 1989. Vgl. zu den methodisch-technologischen Voraussetzungen digitaler Editionen den Überblick bei Sahle 2013, insbesondere Teil 1, S. 191ff.

⁶ Zur Debatte um die *New Philology* vgl. Stackmann 1994 sowie Glessgen / Lebsanft 1997.

zur Identifizierung von Autorschaft, die bislang einen der ergebnisträchtigen Zweige der Stilometrie darstellen, Assoziationen an mittlerweile als überholt geltende Ansätze zur Klärung von ›Echtheitsfragen‹ und anderen wertästhetischen Problemstellungen hervorrufen.⁷ Denn die Frage nach der Beschaffenheit von Autorstilen wurde in der traditionellen Philologie durchaus bereits gestellt, gerade nämlich bei der heute verpönten Rekonstruktion von Originaltexten, zu der sich die ältere Forschung aufgrund ihrer Ausrichtung auf feste Textgestalten genötigt sah. Da es den Vertretern der klassischen Textkritik vor allem darum ging, den einen, ›richtigen‹ Autortext aus der ihrer Ansicht nach entstellten handschriftlichen Überlieferung zu rekonstruieren, erfolgte die Einrichtung der Ausgabentexte nicht selten unter der Maßgabe, ob eine bestimmte Formulierung dem jeweiligen Autor stilistisch zuzutrauen sei oder nicht. Die Frage nach dem ›richtigen‹ Text traf sich dabei mit einem überhöhten Stilideal, das davon ausging, dass die besten Autoren sich auch durch einen besonders qualitätsvollen Stil auszeichnen sollten, und daher stilistisch Minderwertiges aus dem Text auszuschneiden sei. Auf dieser Grundlage wurden Entscheidungen getroffen, die mitunter höchst subjektiv waren, was diese Methode nicht zu Unrecht in Verruf gebracht hat.

Es zeigt sich also, dass die Wissenschaftsgeschichte der Mediävistik nicht die besten Anschlussvoraussetzungen für eine autorzentrierte Stilometrie bietet. Gerade eine Fixierung auf die Instanz ›Autor‹ könnte den erkenntnistheoretischen Mehrwert in Frage stellen, der sich durch die Anwendung der Computertechnologie mit ihrer Möglichkeit zur Öffnung und Perspektivierung von Texten ergeben hat. In dieser Hinsicht erscheint die Stilometrie auf den ersten Blick der erfolgreichen digitalen Editorik fast schon diametral entgegengesetzt.

Dieser Befund überrascht umso mehr, wenn man bedenkt, dass gerade der vielleicht wirkmächtigste Ansatz zur Rehabilitierung quantitativer textanalytischer Methoden in den Geisteswissenschaften, das von Franco Moretti in die Diskussion eingebrachte Konzept des *Distant Reading*, eigentlich mit dem Anspruch angetreten ist, die Fokussierung auf Höhenkammliteratur und Autorgenies zu unterlaufen: Durch die Anwendung quantitativer Methoden, so paraphrasiere ich Moretti, werde es nämlich möglich, ein realistischeres Bild der gesamten Schriftproduktion zu erhalten, das sich nicht auf wenige elitäre Spitzenprodukte beschränken muss, sondern, aufgrund der erhöhten Verarbeitungskapazitäten des Computers, die breite Masse der tatsächlich vorhandenen Texte erschließen kann.⁸ Dieser reizvolle, provokante Ansporn ginge verloren, würde sich die Stilometrie allzu sehr auf die Aufdeckung von Autorschaftszugehörigkeiten beschränken.

Doch wird gerade an den Arbeiten von Moretti deutlich, dass sich die Stilometrie durchaus auch zur Klärung nicht-autorbezogener Fragestellungen verwenden lässt. So hat sich die stilometrische Forschung etwa schon bald – und nicht erst seit Moretti – an der Klassifizierung von Texten in Hinblick auf ihre Gattung versucht, und zwar allein schon aus methodischen Gründen.⁹ Der Einbezug dieser Kategorie liegt deshalb nahe, weil sich gerade die Zugehörigkeit

⁷ Dieses forschungsgeschichtliche Erbe erschwert den Umgang der Mediävistik mit dem Stilbegriff insgesamt, vgl. hierzu Haustein 2011, S. 46. Zum Verhältnis von Stilometrie und Stilfeorschung jetzt grundlegend Jannidis 2014, insbesondere S. 178–183.

⁸ Vgl. Moretti 2000; Moretti 2005, S. 3f.

⁹ Vgl. etwa Biber 1992, Moretti 2005 und jüngst insbesondere Jockers 2013, Jannidis / Lauer 2014 sowie Schöch 2014.

von Texten zu unterschiedlichen Gattungen als größter Störfaktor bei der Differenzierung von Individualstilen erwiesen hat. Ganz offenbar ist der Stil eines Textes nicht nur von der Persönlichkeit des Autors abhängig, sondern auch von überindividuellen Faktoren, wie sie etwa die Gattungen darstellen.

Gerade dieses Spannungsfeld zwischen Tradition und Individualität, welches sich in der Determiniertheit der Texte zwischen Autor- und Gattungsstil zeigt, eröffnet nun aber sehr wohl Anschlussmöglichkeiten an Fragestellungen, die wieder mehr ins Zentrum aktueller Diskussionen in den traditionellen Literaturwissenschaften führen. Denn Morettis Versuch einer Entdifferenzierung zwischen Spitzenprodukten und breiter Masse an Texten erscheint bei näherer Hinsicht der spezifisch mittelalterlichen Literaturauffassung vielleicht gar nicht völlig unangemessen. So wurde bei der Beschreibung mittelalterlicher Literatur immer wieder die Notwendigkeit betont, die einzelnen Texte nicht nach dem Maßstab einer genieästhetischen, auf Originalität abzielenden Literaturproduktion zu beurteilen, sondern den Eigenwert einer speziellen »Ästhetik der Identität«¹⁰ anzuerkennen: Mittelalterliche Texte zielen, anders als wir dies vom modernen Literaturbetrieb seit der Goethezeit gewohnt sind, gar nicht darauf ab, völlig Neues oder Individuelles zu schaffen, sondern knüpfen bewusst an Traditionen an. Statt um das Neuerfinden einer Geschichte geht es vielmehr um das *Wiedererzählen* althergebrachter Stoffe.¹¹ So ist beispielsweise der *Iwein* von Hartmann von Aue, einer der wichtigsten Artusromane der mittelhochdeutschen Literatur, über große Passagen hinweg eine ziemlich genaue Übertragung seiner französischen Quelle, des *Yvain* von Chrétien de Troyes. Die Eingriffe Hartmanns in den Text sind zwar durchaus aussagekräftig, führen aber nicht zu einer völligen Lösung von seiner Vorlage. Dieser Mangel an Originalität wird dabei nicht als Manko empfunden, sondern im Gegenteil, gerade die Wiederaufnahme des Altbewährten, immer schon Gültigen, rechtfertigt erst das literarische Tun. Aufgrund dieser grundsätzlichen Ausrichtung mittelalterlicher Literatur kommt der Einordnung der Texte in Traditions- und damit Gattungszusammenhänge tendenziell eine größere Bedeutung zu als bei modernen Texten -- und dementsprechend, so ließe sich zumindest vermuten, verringert sich die Relevanz des Individualstils eines einzelnen Autors. Dass die Instanz ›Autor‹ deswegen aber nicht gleich völlig verloren geht, hat die sich am Poststrukturalismus abarbeitende Autorschaftsdebatte in der Mediävistik ebenfalls deutlich gezeigt.¹² Auch das Bewusstsein für literarische ›Spitzenprodukte‹ fehlt im Mittelalter nicht, so wird etwa – um wieder bei Hartmann als Beispiel zu bleiben – dieser im bekannten ›Literaturkatalog‹ aus dem *Tristan* Gottfrieds von Straßburg neben anderen als herausragender Autor gepriesen und gerade für seinen besonders klaren Stil gelobt.¹³

2. Fallstudien

¹⁰ Der Ausdruck »Ästhetik der Identität« wurde von Lotman 1993, S. 410ff. geprägt.

¹¹ Vgl. Worstbrock 1999, S. 128–130.

¹² Vgl. exemplarisch Schnell 1998.

¹³ Vgl. die *Tristan*-Verse 4589–4852. Im ›Literaturkatalog‹ unterbricht Gottfried seine Erzählung von Tristan, um eine Reihe seiner Meinung nach vorbildlicher Autoren zu nennen. Hartmann werden dabei insbesondere seine »kristallinen wortelin«, also seine kristall(klaren) Worte, zugutegehalten (Vers 4629).

2.1 Voraussetzungen

Vor diesem Hintergrund könnte es nun gerade besonders interessant erscheinen, auch mittelalterliche Texte in Hinblick auf ihre stilometrische Auswertbarkeit zu überprüfen.

Dem stehen jedoch gewichtige praktische Gründe entgegen: Stilometrische Verfahren beruhen im Wesentlichen auf der computerunterstützten Auszählung von Worthäufigkeiten, auf deren Grundlage Texte mit ähnlichem Wortgebrauch nach statistischen Verfahren sortiert werden. Zwar sind mittlerweile immer mehr mittel- und frühneuhochdeutsche Texte in elektronischer Form verfügbar, was die Grundlage für die Wortfrequenzfassung bietet, die Vergleichbarkeit des Wortgebrauchs in diesen Texten ist jedoch mit erheblichen Schwierigkeiten behaftet. Denn im mittelalterlichen Deutsch gibt es keine festgelegte Orthographie, die Schreibung ein- und desselben Wortes kann variieren, einerseits regional, weil die Schreiber aus unterschiedlichen Dialektgebieten stammen, andererseits auch zeitbedingt, weil sich die Schreibsprache im Verlauf des Mittelalters erheblich verändert hat. Darüber hinausgehend können die Schreiber der Handschriften auch unabhängig von Dialekten und Zeitstufen unterschiedliche Schreibgewohnheiten aufweisen, was sogar soweit geht, dass ein und dasselbe Wort in derselben Handschrift unterschiedlich geschrieben auftritt. Zwar werden diese Differenzen in den meisten der heute verfügbaren Textausgaben durch den Herausgeber ausgeglichen und die Schreibung bis zu einem gewissen Grad vereinheitlicht, doch stellt gerade diese Normalisierung einen weiteren Störfaktor dar, denn auch für diese Vereinheitlichung gibt es keine bis in alle Details festgelegten Regeln, weshalb sie je nach Herausgeber unterschiedlich ausfallen kann. Schließlich ist noch ein weiterer Punkt zu nennen, der bei der Untersuchung vieler mittelhochdeutscher Texte ins Gewicht fällt: Die wichtigsten literarischen Werke sind in Versen abgefasst, also auch epische bzw. erzählende Texte, die zumeist in Reimpaarversen, manchmal sogar in Strophen gebunden sind. Die Besonderheiten der metrischen Struktur und der Reimbindung wirken sich ebenfalls auf den Wortgebrauch aus, und auch dies kann die Vergleichbarkeit der Texte erschweren.¹⁴

2.2 Höfische Epik

Trotz dieser Schwierigkeiten, mit denen eine stilometrische Analyse mittelalterlicher Texte konfrontiert ist, habe ich in einer ersten Annäherung zu eruiieren versucht, welche Resultate sich mit den bereits vorhandenen Textkorpora erzielen lassen. Ziel der im Folgenden beschriebenen Fallstudien ist es also ausdrücklich nicht, endgültige Ergebnisse zu liefern, sondern zunächst heuristisch zu ermitteln, welche Probleme genauer in den Blick zu nehmen sind. Die Grundlage für meinen ersten Versuch bildete ein Korpus von einigen wichtigen epischen Texten der mittelhochdeutschen Literatur, das ich mithilfe des von Maciej Eder, Mike Kestemont und Jan Rybicki entwickelten *Stylo-Packages*¹⁵ für das Statistikprogramm R¹⁶ einer

¹⁴ Dazu in **Abschnitt 2.2** ein Beispiel. Allgemein zu möglichen Fehlertypen in Textkorpora und deren Auswirkungen auf die stilometrische Analyse Eder 2013, S. 604ff. Eders Versuche zeigen, dass sich stilometrische Auswertungen erstaunlich robust gegenüber *Noise* in den herangezogenen Textkorpora erweisen (vgl. insbesondere Eder 2013, S. 610ff.).

¹⁵ Eder / Kestemont / Rybicki 2013.

¹⁶ R Core Team 2013.

stilometrischen Analyse ausgesetzt habe. Konkret herangezogen habe ich die heute erhaltenen Werke der drei Klassiker der höfischen Epik, den *Tristan* Gottfrieds von Straßburg, den *Parzival* und den *Willehalm* Wolframs von Eschenbach sowie die epischen Werke Hartmanns von Aue, nämlich die beiden Artusromane *Erec* und *Iwein* sowie die Legendendichtungen *Der arme Heinrich* und *Gregorius*.¹⁷ Um eine weitere zeitliche Streuung zu erreichen, wurden zudem noch der *Eneas*-Roman Heinrichs von Veldeke, der eine Art Vorläufer der höfischen Klassik darstellt, in die Analyse mit einbezogen sowie die Werke Konrads von Würzburg und Ulrichs von Etzenbach, die zeitlich etwas später anzusiedeln sind.

Eine Clusteranalyse dieser Texte, die auf der Grundlage der 200 häufigsten Wörter erstellt wurde,¹⁸ bietet ein auf den ersten Blick erstaunlich eindeutiges Bild (Abbildung 1):

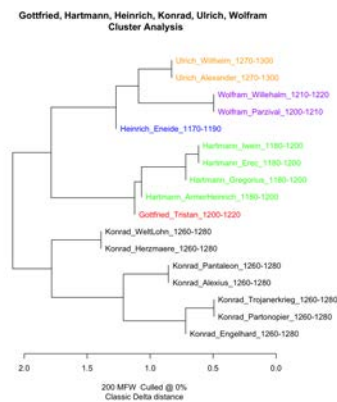


Abb. 1: Clusteranalyse der höfischen Epik (Graphik: Autor).

Die Werke der Autoren werden vom Computer ohne Ausnahme richtig sortiert, zum Teil scheinen sich sogar Texte derselben Gattung zusammenzuordnen, so bilden etwa die beiden Artusromane *Erec* und *Iwein* einen eigenen Unterzweig in der Hartmann-Gruppe.¹⁹ Grundsätzlich scheint jedoch das Autorensignal das Gattungssignal zu überwiegen, so befinden sich etwa der *Parzival* und der *Willehalm*, die unterschiedlichen Gattungen angehören, auf ein und demselben Wolfram-Zweig. Diese Ergebnisse sind freilich mit äußerster Vorsicht zu behandeln: Denn natürlich spielen die bereits erwähnten Einflussfaktoren wie Schreiber, Schreibdialekt und Herausgeber bei dieser Sortierung eine Rolle, und ich werde im Folgenden noch näher auf ein Beispiel für eine solche Verzerrung eingehen. Allerdings wäre selbst unter Berücksichtigung dieser Faktoren nicht unbedingt eine solch klare Verteilung zu

¹⁷ Zu den verwendeten Ausgaben vgl. das Literaturverzeichnis.

¹⁸ Die Analyse beruht auf einer Auszählung der prozentual am häufigsten auftretenden Wörter in den einzelnen Texten. Je ähnlicher zwei Texte sich in Bezug auf ihren Wortgebrauch verhalten, desto näher werden sie zusammensortiert. Zur Berechnung der Nähe kommen statistische Methoden zum Einsatz, im vorliegenden Fall das von Burrows 2002 entwickelte Delta-Distanzmaß, das mit standardisierten Wortfrequenzen arbeitet.

¹⁹ Versuche mit anderen Distanzmaßen (Manhattan Distance, Eder's Delta, vgl. dazu die Dokumentation bei Eder / Rybicki / Kestemont 2014, S. 14ff.) liefern jeweils ein ähnliches Bild. Eine signifikante Änderung ergibt sich aber bei der Berechnung der Darstellung auf der Grundlage von 1000 Wörtern: Hier sprengt der *Tristan* die Gruppe der Hartmann-Werke und bildet mit dem *Armen Heinrich* einen eigenen Zweig. Zu den möglichen Gründen hierfür siehe [Abschnitt 2.2](#).

erwarten, da ja die Werke mancher der hier aufgeführten Autoren in ganz unterschiedlichen Handschriften überliefert sind und auch von unterschiedlichen Editoren herausgegeben wurden. So ist beispielsweise die Ausgabe von Hartmanns *Iwein* nach Handschriften des 13. Jahrhunderts erstellt, der nebengeordnete *Erec* jedoch nach einer Handschrift, die erst aus dem 16. Jahrhundert stammt. Diese Differenz in der handschriftlichen Basis spielt also offensichtlich keine genügend große Rolle, als dass sie eine Zusammenordnung der beiden Texte verhindern könnte.²⁰

2.3 Wolfram und Hartmann

Die Ergebnisse erscheinen jedenfalls vielversprechend genug, um diesen Zusammenhängen genauer nachzugehen. In einem weiteren Versuch habe ich daher Wolframs *Parzival* und die beiden Artusromane Hartmanns, den *Erec* und den *Iwein*, mithilfe des von John Burrows und Hugh Craig entwickelten Zeta-Tests einer kontrastiven Analyse unterzogen. Dabei werden die einzelnen Texte in Abschnitte gleicher Länge zerteilt (hier in Abschnitte zu 2000 Wörtern), die jeweils von einer Textgruppe (Wolfram bzw. Hartmann) im Vergleich zur anderen konstant bevorzugten Wörter ermittelt und schließlich die einzelnen Textpartien in eine graphische Darstellung gebracht, bei der wieder Textabschnitte mit ähnlichem Wortgebrauch näher sortiert werden.²¹ Als Testgruppe sind der Analyse noch Wolframs *Willehalm*, Gottfrieds *Tristan* und Hartmanns Legendedichtungen hinzugefügt worden (Abbildung 2).

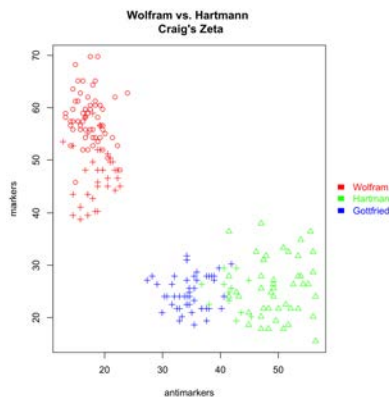


Abb. 2: Kontrastive Analyse der Werke Hartmanns, Wolframs und Gottfrieds (Graphik: Autor).

In der Darstellung repräsentiert jedes Symbol (Kreis, Kreuz und Dreieck) jeweils einen Textabschnitt. Bei den roten Kreisen handelt es sich um Textpartien aus Wolframs *Parzival* (also aus der ersten Textgruppe), bei den grünen Dreiecken um Abschnitte aus Hartmanns

²⁰ Freilich sind die Ausgaben des *Iwein* und des *Erec* nach ähnlichen Maßgaben normalisiert worden, doch auch in Bezug auf den Herausgeberfaktor gilt ähnliches wie für die Handschriften: Trotz zum Teil unterschiedlicher Ausgabenkonventionen werden die Werke gleicher Autoren sortiert.

²¹ Die Auswertung erfolgte wieder mit dem Stylo-Package von Eder / Kestemont / Rybicki 2013. Zum Zeta-Test vgl. Burrows 2007 sowie Craig / Kinney 2009. Anders als das in [Abschnitt 2.2](#) zur Anwendung gebrachte Delta-Distanzmaß zielt Craig's Zeta weniger auf die absolute Worthäufigkeit ab als auf die Konstanz des Wortgebrauchs über die einzelnen Abschnitte hinweg.

Artusromanen (der zweiten Textgruppe). Kreuze stellen Abschnitte aus der Testgruppe dar (*Willehalm*, *Tristan* und Hartmanns Legenden), wobei die Partien aus dem *Willehalm* rot, die aus dem *Tristan* blau und die aus Hartmanns Legenden grün eingefärbt sind.

Überwiegend zeigt sich wieder ein relativ klarer Unterschied zwischen den Autoren, die Textpartien sind nicht weit über das Raster verteilt, sondern ordnen sich entsprechend den Gruppierungen zusammen. Der *Willehalm* erscheint zusammen mit dem *Parzival* deutlich von Hartmann abgesetzt und die Legenden Hartmanns finden sich bei seinen Artusromanen. Einzig beim *Tristan* gibt es Unschärfen, da er sich teilweise mit den Werken Hartmanns vermischt.

Spätestens an diesem Punkt wäre es nun interessant zu wissen, welche Wörter für diese Sortierung verantwortlich sind. Eine Auflistung der für die Differenzierung der beiden Korpora aussagekräftigsten Wörter sieht folgendermaßen aus (Abbildung 3):

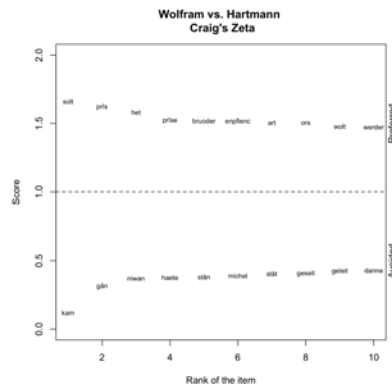


Abb. 3: Liste der von Wolfram gegenüber Hartmann bevorzugten bzw. vermiedenen Wörter (Graphik: Autor).

Ausgangspunkt der Darstellung ist das Wolfram-Korpus,²² über dem horizontalen Strich in der Mitte liegen die von Wolfram im Vergleich zu Hartmann konstant häufig gebrauchten Wörter, unter dem Strich die von Wolfram vermiedenen Wörter. Die erste Position unter den vermiedenen Wörtern nimmt das Wort ›kam‹ ein, das im Mittelhochdeutschen die 1. und 3. Pers. Sg. Präteritum des neuhochdeutschen Verbs ›kommen‹ ausdrückt. Das Wort ist deutlich von der Mittellinie abgesetzt, wird also von Wolfram gegenüber Hartmann in besonders auffälliger Weise gemieden.²³ Doch ist gerade dies kein Beleg dafür, dass Hartmann etwa öfter von ›kommen‹ sprechen würde als Wolfram, denn die 1. und 3. Pers. Sg. von ›komen‹ kann im Mittelhochdeutschen auch ›kom‹ lauten. Und genau diese Form ›kom‹ tritt nun bei Wolfram anstelle von ›kam‹ auf, wodurch sich der scheinbare Mangel an ›kam‹-Formen in den Wolfram-Texten erklärt. Die Verteilung der Formen ›kom/kam‹ ist regional bedingt, ›kam‹ ist nämlich die

²² Im Unterschied zur vorigen Darstellung (Abbildung 2) wurde bei dieser Analyse der *Willehalm* gemeinsam mit dem *Parzival* in die erste Textgruppe eingeordnet. Die zweite Textgruppe bilden wie zuvor *Erec* und *Iwein*.

²³ Der Zahlenwert auf der y-Achse gibt den Zeta-Score der aufgelisteten Wörter an. Grundsätzlich kann der Zeta-Score einen Wert von 0 bis 2 annehmen. Ein Zeta-Score von 0 würde bedeuten, dass das Wort in allen Abschnitten der Hartmann-Texte vorhanden ist und in jedem Abschnitt der Wolfram-Texte fehlt, ein Zeta-Score von 2 den genau umgekehrten Fall.

Variante, die tendenziell im südwestdeutschen (alemannischen) Raum bevorzugt wird, aus dem Hartmann stammt.²⁴ Dass die Verwendung von ›kam‹ tatsächlich auf Hartmann zurückgeht und nicht erst durch spätere Abschreiber seiner Werke eingeführt wurde, lässt sich in einigen Fällen sogar sichern, nämlich dann, wenn das Wort auch im Reim auftritt, und dort auf Wörter reimt, die nicht nur im Alemannischen, sondern auch überregional auf die Silbe #am enden.²⁵ Die Schreiber des Mittelalters hatten zwar grundsätzlich keine Bedenken, den Text, den sie kopierten, ihrem eigenen Dialekt anzupassen, im Fall eines Reimpaars hätten sie jedoch den Text regelrecht umschreiben müssen, um etwa die Form ›kom‹ in ›kam‹ zu ändern.

Nun eignet sich so gesehen das Wortpaar ›kom‹ versus ›kam‹ bei der direkten Gegenüberstellung tatsächlich hervorragend zur Differenzierung zwischen Hartmann und Wolfram,²⁶ allerdings ist damit über einen möglichen Individualstil Hartmanns noch wenig ausgesagt.²⁷ Zudem dürfte dieser letztlich regionale Einflussfaktor eine mögliche Klassifizierung nach Autorschaften bei einem Korpus mit mehreren Autoren unterlaufen. So erscheint es etwa durchaus denkbar, dass das vom Computer berechnete Näheverhältnis von Gottfried und Hartmann, das in den oben gezeigten Clusterdarstellungen ersichtlich ist, darauf zurückgeht, dass auch Gottfried – und weite Teile der *Tristan*-Überlieferung – wie Hartmann aus dem südwestdeutschen Raum stammen.

Und schließlich bleibt die Verteilung der Formen ›kom‹ versus ›kam‹ anfällig für Schreiber- und Herausgeber-Einflüsse, zumindest solange sie nicht im Reim auftreten. Gleich dieses erste Wort in der Rangliste bietet also einen Beleg für die oben beschriebenen Störfaktoren, die die Autorzuordnung behindern können. Für weitergehende stilometrische Untersuchungen könnte es sich daher als sinnvoll erweisen, mit lemmatisierten Texten zu arbeiten, um diese Verzerrungen auszuschließen. Leider gibt es dafür aber noch nicht ausreichend offen zugängliche lemmatisierte E-Texte, die zur Analyse herangezogen werden können. Zudem lässt sich am Beispiel zeigen, dass die Lemmatisierung noch nicht alle systematischen Probleme löst, die die unterschiedliche Schreibung mit sich bringt: Blickt man nämlich zurück auf die Reimbindung des Wortes und betrachtet man nach der Gruppe der Hartmann-Texte nun jene der Wolfram-Texte, dann zeigt sich, dass Wolfram ›kom‹ im Reim nie gebraucht. Dies hat seinen guten Grund, der darin liegt, dass im Mittelhochdeutschen ein Reimwort auf #om sehr viel schwerer zu finden ist als ein Reim auf #am.²⁸ Das heißt also, dass ein Autor, der die Form ›kom‹ verwendet, diese im Reim weniger zum Einsatz bringen kann. Eine hohe Frequenz von ›kom‹ wäre so gesehen noch auffälliger als eine gleich hohe Frequenz von ›kam‹ -- und dieser Unterschied ginge bei der Lemmatisierung verloren.²⁹

²⁴ Vgl. hierzu Paul 2007, S. 253 (§M 79, Anm. 4); zu Unschärfen in der regionalen Verteilung der Formen vgl. Klein 1985, S. 137f.

²⁵ Auffälligerweise hat Hartmann den Reim auf ›kam‹ jedoch in seinem zweiten Artusroman, dem *Iwein*, im Verlauf des Textes immer mehr vermieden. Vgl. hierzu Schirokauer 1923, S. 13.

²⁶ Dieser Umstand ist natürlich der älteren Stilforschung nicht entgangen, vgl. etwa Schirokauer 1923, S. 13.

²⁷ Zur grundsätzlichen Problematik, dass sich die in der stilometrischen Analyse ermittelten Stilmarker nicht zu einem Gesamtbild des Individualstils eines Autors zusammenfügen lassen vgl. Jannidis 2014, S. 191.

²⁸ Auf diesen Umstand hat mich Prof. Klaus Schmidt von der [mittelhochdeutschen Begriffsdatenbank](#) aufmerksam gemacht. Ich danke Herrn Prof. Schmidt für seine umfangreichen Anregungen zum vorliegenden Thema.

²⁹ Ein weiterer Störfaktor ergibt sich dadurch, dass ›kom‹ anders als ›kam‹ auch die 2. Pers. Sg. des Imperativs ausdrücken kann.

Neben diesen problematischen Fällen gibt es in der Wortliste nun aber auch Marker, die eindeutig nicht auf regionale oder Schreibereinflüsse zurückzuführen sind. Ein Beispiel hierfür wäre das Wort ›prîs‹, das in der Liste der von Wolfram bevorzugten Wörter etwas überraschend an zweiter Stelle aufscheint. Für ›prîs‹, was dem Neuhochdeutschen ›Preis‹ im Sinne von Lob, Ruhm und Anerkennung entspricht, gibt es keine doppelte Wortform, der ›prîs‹ wird also tatsächlich viel häufiger in den Wolfram-Texten thematisiert als bei Hartmann.

Die Frage ist nun, ob ›prîs‹ also ein Wort darstellt, das besonders typisch für den Individualstil von Wolfram ist oder ob andere Faktoren bei dieser Verteilung eine Rolle spielen. Zur Beantwortung dieser Frage ist es zunächst von Relevanz, dass ›prîs‹ ein inhaltlich bedeutendes Substantiv ist und kein lediglich füllendes Funktionswort. Die jüngere stilometrische Forschung präferiert für die Unterscheidung von Autorstilen bekanntlich eher inhaltsleere Funktionswörter wie Artikel, Pronomen oder Konjunktionen, da diese von den Autoren tendenziell unbewusst eingesetzt werden.³⁰ ›prîs‹ als semantisch aussagekräftiges Wort hingegen wäre so gesehen eher unter den topic-bezogenen Markern einzuordnen, also unter den Worten, deren Verwendung beispielsweise von der Thematik oder der Gattung beeinflusst wird.³¹ Nun ist ›prîs‹ im Sinne von Belohnung ritterlicher Bewährung zweifelsohne ein Wort, das man in höfischer Artusliteratur erwarten würde, aber es bleibt immer noch auffällig, dass das Wort in den Artusromanen Hartmanns viel seltener auftaucht. Und ebenso erklärungsbedürftig wäre, wieso Wolfram ›prîs‹ sowohl im *Parzival* als auch im *Willehalm* häufiger gebraucht, die streng genommen gar nicht zur selben Gattung gehören.³² Viel mehr als eine Gattungsvorgabe scheint sich hier also eine Art Dichterkonzept Wolframs anzudeuten, der den höfischen Leitbegriff ›prîs‹ gattungsübergreifend zum Einsatz bringt.³³

3. Gattungsstil

Diese Zwischenstellung des Wortes ›prîs‹ zwischen den einzelnen Stilebenen bringt mich nun wieder zurück zu der Frage, ob mittelalterliche Literatur vermehrt von Traditionsvorgaben beeinflusst ist und zurück zu dem vielschichtigen Begriff der Gattung. Die Verwendung von ›prîs‹ bei Wolfram hat gezeigt, dass es offensichtlich gewisse thematische oder konzeptuelle Vorgaben gibt, die quer zu diesen Gattungen liegen. *Parzival* und *Willehalm* gehören aus

³⁰ Vgl. Kestemont 2014, S. 60–61. Für einen Überblick über die bei der Autorschaftsattribuierung in der Forschung verwendeten Marker vgl. Holmes 1994.

³¹ Zur Problematik der Differenzierung vgl. jedoch Schöch 2014, S. 147ff.

³² Der *Willehalm* ist anders als der Artusroman *Parzival* ein Text aus der Chanson de Geste-Tradition. Freilich sind beide Texte von derselben höfischen Grundhaltung getragen.

³³ Dass sich Wolframs Wortschatz gegenüber anderen zeitgenössischen Dichtern durch verschiedene Besonderheiten auszeichnet, ist schon früh in der Forschung bemerkt worden (vgl. dazu zusammenfassend Hartmann 2011, S. 147), ebenso, »daß Wolframs Dichtung von terminologischen Leitmotiven durchzogen ist« (Mohr 1954, S. 175). Im Zentrum von Wolframs Wortschatzes stehen »bestimmte Leitwörter [...], die gewissermaßen das ethisch-theologische Rückgrat seines Erzählprogramms bilden und immer wieder aufgerufen werden« (Hartmann 2011, S. 150). Heiko Hartmann nennt als Beispiel für solche Leitwörter »kiusche« (Keuschheit), »riuwe« (Reue), »triuwe« (Treue) und »zwîvel« (Zweifel), deren Stellung im Werk schon genauer untersucht wurde. Insofern ist es überraschend, dass in der Liste der von Wolfram bevorzugten Worte nun gerade ›prîs‹ als erstes Substantiv auftritt, das bislang, soweit ich sehe, noch vergleichsweise wenig Beachtung gefunden hat. Hier lassen sich Anknüpfungspunkte für eine qualitative Auswertung der stilometrischen Methode finden: Natürlich können die Leitbegriffe nicht unabhängig von ihrem Kontext und ihrer Semantik betrachtet werden und die Häufigkeit eines Wortes sagt noch nicht alles über seinen Stellenwert aus, dennoch wäre es den Versuch wert, zunächst einmal rein quantitativ festzustellen, welche Wörter sich aufgrund ihrer Frequenz als Leitwörter aufdrängen.

literaturwissenschaftlicher Sicht zwar unterschiedlichen Gattungen an, sie sind aber aus sprachlicher Sicht durch ein ähnliches Vokabular geprägt, mit dem Wolfram auf seine spezielle Art ein höfisches Idealbild entwirft. Das heißt aber, dass der stilistische Befund nicht immer ausreichend zur Differenzierung von Gattungen ist, zumindest von Gattungen, wie sie in der Literaturwissenschaft eingeführt sind.

Aus literaturwissenschaftlicher Sicht sind Gattungen nämlich vielschichtige Gebilde, die sich nicht oder nicht nur durch einen ähnlichen Stil auszeichnen, sondern auch durch ähnliche Thematik, ähnliche Figurenkonstellationen oder andere Faktoren bestimmt sein können. Gattungen sind so gesehen Sammelbegriffe, die ganz unterschiedliche Aspekte vereinen, die nicht immer auf derselben kategorialen Ebene liegen müssen. So kann etwa das rein äußerlich-formal hervorgehobene Sonett (eine Gedichtform mit zweimal vier und zweimal drei Zeilen) ebenso als Untergattung gelten wie beispielsweise der sich durch seine Hauptfigur auszeichnende Alexanderroman (ein episches Werk, welches über das Leben Alexanders des Großen erzählt). Letztlich – so hat die literaturwissenschaftliche Gattungstheorie bereits seit den 1970er Jahren herausgearbeitet – sind Gattungen in erster Linie institutionell bestimmt, sie sind Ordnungsbegriffe, die in die literarische Diskussion eingeführt werden und den Erwartungshorizont von Autoren und Lesern determinieren.³⁴

Das heißt aber auch, dass der literaturwissenschaftliche Gattungsbegriff nicht unbedingt mit dem sprachwissenschaftlichen Begriff von Textsorten deckungsgleich sein muss. Dass es ratsam ist, bei der textstatistischen Analyse zwischen eher literaturwissenschaftlich bestimmten *genres* und eher linguistisch bestimmten *text types* zu differenzieren, hat bereits Douglas Biber erkannt, der sich in den 90er Jahren mit der stilometrischen Gattungsbestimmung beschäftigt hat und damit als Vorreiter auf dem Gebiet gelten kann. *Genres* sind für Biber, »text varieties that are readily recognized and ›named‹ within a culture (e.g., letters, press editorials, sermons, conversation)«, entsprechen also dem institutionell bestimmten, literaturwissenschaftlichen Modell, während der Begriff *text types* für »varieties that are defined linguistically (rather than perceptually)«³⁵ vorbehalten ist. Zwar können sich sowohl *genres* als auch *text-types* durch gewisse stilistische Eigenarten auszeichnen, aber nur *text-types* sind per definitionem dadurch determiniert:

»Both genres and text types can be characterized by reference to co-occurring linguistic features, but text types are further defined quantitatively such that the texts in a type all share frequent use of the same set of co-occurring linguistic features.«³⁶

Genres müssen dagegen nicht unbedingt sprachlich kohärent sein: »Genres have a perceptual basis in a given culture, but they are not necessarily linguistically coherent.«³⁷

³⁴ Vgl. grundlegend Voßkamp 1977 sowie Jauß 1977.

³⁵ Biber 1992, S. 332.

³⁶ Biber 1992, S. 332.

³⁷ Biber 1992, S. 339.

Das Problem scheint mir nun zu sein, dass die stilometrische Analyse eher dafür geschaffen ist, *text types* zu untersuchen, dass es aus literaturwissenschaftlicher Sicht aber eigentlich aufschlussreicher wäre, *genres* bestimmen zu können. Aus literaturwissenschaftlicher Sicht sind Gattungen hochgradig unfeste Größen, die vor allem aufgrund ihrer historischen Veränderlichkeit nur schwer zu fassen sind. Nur allzu oft scheinen sich Texte der gattungsmäßigen Einordnung zu widersetzen bzw. die einmal festgelegten Gattungsgrenzen zu überschreiten. Diese Schwierigkeiten bei der Kategorisierung hat die Gattungsforschung dazu gebracht, Gattungen als historisch offene Kategorien mit prototypensemantischen Kernen zu definieren. Und einmal mehr ist die Situation im Mittelalter besonders prekär, denn anders als wir das vom heutigen Literaturbetrieb gewöhnt sind, ist Literatur im Mittelalter kein ausdifferenziertes, institutionalisiertes Teilsystem der Gesellschaft. Es fehlen die festen Rahmenbedingungen, die wir von einem modernen Literaturbetrieb kennen, wo die Zugehörigkeit von Texten zu Gattungen viel stärker reflektiert werden, als dies im Mittelalter der Fall war. Mittelalterliche volkssprachige Textsorten hingegen sind eher durch eine sich immer wieder neu formierende, mündliche Aufführungssituation determiniert als durch eine feste Vorstellung davon, wie Gattungen aussehen müssen.³⁸

Ich möchte an meinem letzten Beispiel zeigen, was dieser Umstand für die Kategorienbildung bedeuten kann. Ich gehe dazu von der Gattung der Epik zum Bereich der Lyrik über. Die mittelhochdeutsche Lyrik wird traditionell in zwei Unterarten eingeteilt, in den Minnesang und den Sangspruch. Die Abgrenzung zwischen diesen beiden Unterarten ist äußerst unscharf und letztlich nur thematisch bedingt. Minnesang ist grundsätzlich Dichtung, in der von höfischer Liebe, also der so genannten Minne, die Rede ist. Demgegenüber kann der Sangspruch nur negativ definiert werden, er gilt als alles das im Rahmen mittelalterlicher Lieddichtung, was nicht Minnesang ist.³⁹

Wollte man sich nun mit stilometrischen Methoden an einer genaueren Differenzierung dieser beiden Formen versuchen, so wird ersichtlich, dass eine Auszählung der häufigsten Wörter für den Minnesang ein durchaus interessantes Ergebnis bringt, das schon in einer Darstellung der bekannten Minnesang-Anthologie *Minnesangs Frühling* als Wordcloud auf dem ersten Blick ins Auge fällt (*Abbildung 4*):⁴⁰

³⁸ Vgl. Grubmüller 1999, S. 195; Bleumer / Emmelius 2011, S. 1.

³⁹ Vgl. den Überblick bei Tervooren 1995, S. 2. Den aktuellsten in einer langen Reihe von Abgrenzungsversuchen zwischen Sangspruch und Minnesang bietet Schnell 2013.

⁴⁰ Bei der Wordcloud, die mithilfe des Programms *Voyant-Tools* erstellt wurde (Sinclair / Rockwell 2012) werden die häufigsten im Text auftretenden Wörter in einer Schriftgröße entsprechend ihrer Frequenz dargestellt: Je häufiger ein Wort auftritt, umso größer erscheint es in der Darstellung.

4. Schluss

Ich breche hier meine Beispielreihe ab und fasse meine Überlegungen zusammen: Ich hoffe mit meinen Explorationsen gezeigt zu haben, dass sowohl traditionelle Literaturwissenschaft als auch Stilometrie bei gegenseitiger Kenntnisnahme voneinander profitieren können. Die traditionelle Literaturwissenschaft wird wohl erst dann bereit sein, sich auf quantifizierende Verfahrensweisen einzulassen, wenn diese den Anschluss an ihre aktuellen Fragestellungen suchen. Wenn diese Verbindung aber hergestellt ist, dann – so legen die hier gezeigten vorläufigen Ergebnissen zumindest nahe – könnte die Stilometrie durchaus ihren Beitrag zu theoretischen Debatten der Mediävistik leisten, etwa wenn es um die Frage geht, ob der Individualstil und damit die Profilierung von Autorschaft vielleicht nicht doch eine größere Rolle spielt als mitunter in der mediävistischen Forschung angenommen. Umgekehrt sollte die Stilometrie darauf achten, was die traditionelle Literaturwissenschaft in Hinblick auf die verwendeten Kategorien zu sagen hat. Denn tut sie das nicht und überblendet leichtfertig quantitative Textanalyse und Hermeneutik, dann lassen sich zwar messbare Ergebnisse erzielen, die aber an den Objekten der Literaturwissenschaft vorbeigehen. Erst die Herausarbeitung der Differenzen schafft die Voraussetzung für eine gegenseitige Befruchtung der Disziplinen und ermöglicht einen ›methodischen Brückenschlag‹, der über die bloße ›feindliche Übernahme‹ hinausgeht.⁴⁴

⁴⁴ Vgl. den Titel der [DHD-Tagung](#) in Passau, auf die dies Bezug nimmt: »Digital Humanities – methodischer Brückenschlag oder ›feindliche Übernahme‹? Chancen und Risiken der Begegnung zwischen Geisteswissenschaften und Informatik«.

Bibliographische Angaben

Primärtexte

- Gottfried von Straßburg: Tristan. Nach dem Text von Friedrich Ranke neu hg., ins Neuhochdeutsche übersetzt, mit einem Stellenkommentar und einem Nachwort von Rüdiger Krohn. Stuttgart 1980. [\[Nachweis im GBV\]](#)
- Hartmann von Aue: Der arme Heinrich. Mittelhochdeutsch/Neuhochdeutsch. Hg. von Ursula Rautenberg, übersetzt von Siegfried Grosse. Stuttgart 1993. [\[Nachweis im GBV\]](#)
- Hartmann von Aue: Erec. Mittelhochdeutscher Text und Übertragung von Thomas Cramer. Frankfurt/Main 1972. [\[Nachweis im OPAC\]](#)
- Hartmann von Aue: Gregorius. Hg. von Hermann Paul. 13., neu bearbeitete Auflage, besorgt von Burghart Wachinger. Tübingen 1984. [\[Nachweis im GBV\]](#)
- Hartmann von Aue: Iwein. Eine Erzählung. Hg. von Georg F. Benecke / Karl Lachmann. Neu bearbeitet von Ludwig Wolff. 7. Ausgabe. Berlin 1968. [\[Nachweis im OPAC\]](#)
- Heinrich von Veldeke: Eneasroman. Mittelhochdeutsch/Neuhochdeutsch. Nach dem Text von Ludwig Ettmüller ins Neuhochdeutsche übersetzt, mit einem Stellenkommentar und einem Nachwort von Dieter Kartschoke. Stuttgart 1986. [\[Nachweis im GBV\]](#)
- Konrad von Würzburg: Engelhard. Hg. von Ingo Reiffenstein. 3., neu bearbeitete Auflage der Ausgabe von Paul Gereke. Tübingen 1982. [\[Nachweis im GBV\]](#)
- Konrad von Würzburg: Kleinere Dichtungen. Bd. 1: Der Welt Lohn – Das Herzmaere – Heinrich von Kempten. Hg. von Edward Schröder. 9. Auflage. Berlin 1968. [\[Nachweis im GBV\]](#)
- Konrad von Würzburg: Die Legenden. Bd. 2. Hg. von Paul Gereke. Tübingen 1926. [\[Nachweis im OPAC\]](#)
- Konrad von Würzburg: Pantaleon. Hg. von Winfried Woesler. 2. Auflage. Tübingen 1974. [\[Nachweis im OPAC\]](#)
- Konrads von Würzburg: Partonopier und Meliur. Aus dem Nachlasse von Franz Pfeiffer, hg. von Karl Bartsch. Nachdruck der Ausgabe von 1871. Berlin 1970. [\[Nachweis im OPAC\]](#)
- Konrad von Würzburg: Der Trojanische Krieg. Nach den Vorarbeiten K. Frommanns und F. Roths zum ersten Mal hg. von Adelbert von Keller. Stuttgart 1858. [\[Nachweis im OPAC\]](#)
- Jean L. C. Putmans: EDV-Text von ›Des Minnesangs Frühling‹. Lorch 1993 (= Göppinger Arbeiten zur Germanistik 600). [\[Nachweis im GBV\]](#)
- Ulrich von Etzenbach: Alexander. Hg. von Wendelin Toischer. Tübingen 1888. [\[Nachweis im OPAC\]](#)
- Ulrich von Etzenbach: Wilhelm von Wenden. Hg. von Hans-Friedrich Rosenfeld. Berlin 1957. [\[Nachweis im OPAC\]](#)
- Wolfram von Eschenbach: Willehalm. Nach der gesamten Überlieferung kritisch hg. von Werner Schröder. Berlin, New York 1978. [\[Nachweis im OPAC\]](#)
- Wolfram von Eschenbach: Werke. Hg. von Karl Lachmann. 5. Auflage. Berlin 1891. [\[Nachweis im OPAC\]](#)

Sekundärliteratur

- Alte und neue Philologie. Hg. von Martin-Dietrich Glessgen / Franz Lebsanft. Tübingen 1997. [\[Nachweis im OPAC\]](#)
- Douglas Biber: The Multi-Dimensional Approach to Linguistic Analyses of Genre Variation: An Overview of Methodology and Findings. In: Computers and the Humanities 26 (1992), H. 5/6, S. 331–345. [\[Nachweis im GBV\]](#)
- Hartmut Bleumer / Caroline Emmelius: Generische Transgressionen und Interferenzen. Theoretische Konzepte und historische Phänomene zwischen Lyrik und Narrativik. In: Lyrische Narrationen – narrative Lyrik. Gattungsinterferenzen in der mittelalterlichen Literatur. Hg. von Hartmut Bleumer / Caroline Emmelius. Berlin, New York 2011, S. 1–39. [\[Nachweis im OPAC\]](#)
- Manuel Braun: Kodieren, Annotieren, Theoretisieren. Zur Wiederannäherung von Literatur- und Sprachwissenschaft über Korpora. In: Lili. Zeitschrift für Literaturwissenschaft und Linguistik 43 (2013), H. 172, S. 83–90. [\[Nachweis im OPAC\]](#)
- John Burrows: ›Delta‹: a Measure of Stylistic Difference and a Guide to Likely Authorship. In: Literary and Linguistic Computing 17 (2002), H. 3, S. 267–287. [\[Nachweis im GBV\]](#)
- John Burrows: All the Way Through: Testing for Authorship in Different Frequency Strata. In: Literary and Linguistic Computing 22 (2007), H. 1, S. 27–47. [\[Nachweis im GBV\]](#)

- Bernard Cerquiglini: *Eloge de la variante. Histoire critique de la philologie*. Paris 1989. [[Nachweis im OPAC](#)]
- Hugh Craig / Arthur Kinney: *Shakespeare, Computers, and the Mystery of Authorship*. Cambridge 2009. [[Nachweis im GBV](#)]
- Maciej Eder / Mike Kestemont / Jan Rybicki: *Stylometry with R: a suite of tools*. In: *Digital Humanities 2013: Conference Abstracts*. University of Nebraska-Lincoln, NE, S. 487–489. [[Nachweis im GBV](#)]
- Maciej Eder: *Mind Your Corpus: systematic errors in authorship attribution*. In: *Literary and Linguistic Computing* 28 (2013), H. 4, S. 603–614. [[Nachweis im GBV](#)]
- Maciej Eder / Jan Rybicki / Mike Kestemont: *Stylo: a package for stylometric analyses*. August 17, 2014. [[online](#)]
- Michel Foucault: *Was ist ein Autor?* In: *Michel Foucault: Schriften zur Literatur*. Aus dem Französischen von Karin von Hofer / Anneliese Botond. Frankfurt/Main 1988, S. 7–31. [[Nachweis im OPAC](#)]
- Klaus Grubmüller: *Ich als Rolle. »Subjektivität« als höfische Kategorie im Minnesang? In: Höfische Literatur -- Hofgesellschaft – Höfische Lebensformen um 1200*. Hg. von Gert Kaiser / Jan-Dirk Müller. Düsseldorf 1986, S. 387–408. [[Nachweis im OPAC](#)]
- Klaus Grubmüller: *Gattungskonstitution im Mittelalter*. In: *Mittelalterliche Literatur und Kunst im Spannungsfeld von Hof und Kloster. Ergebnisse der Berliner Tagung, 9.–11. Oktober 1997*. Hg. von Nigel F. Palmer / Hans-Jochen Schiewer. Tübingen 1999, S. 193–210. [[Nachweis im OPAC](#)]
- Heiko Hartmann: *Darstellungsmittel und Darstellungsformen in den erzählenden Werken*. In: *Wolfram von Eschenbach. Ein Handbuch*. Hg. von Joachim Heinze. Berlin, Boston 2011, S. 145–220. [[Nachweis im OPAC](#)]
- Jens Hausteine: *Mediävistische Stilforschung und die Präsenzkultur des Mittelalters. Mit einem Ausblick auf Gottfried von Straßburg und Konrad von Würzburg*. In: *Textprofile stilistisch. Beiträge zur literarischen Evolution*. Hg. von Ulrich Breuer / Bernhard Spies. Bielefeld 2011, S. 43–60. [[Nachweis im GBV](#)]
- David I. Holmes: *Authorship Attribution*. In: *Computers and the Humanities* 28 (1994), H. 2, S. 87–106. [[Nachweis im GBV](#)]
- Fotis Jannidis / Gerhard Lauer: *Burrow's Delta and Its Use in German Literary History*. In: *Distant Readings. Topologies of German Culture in the Long Nineteenth Century*. Hg. von Matt Erlin / Lynne Tatlock. Rochester 2014, S. 29–54. [[Nachweis im GBV](#)]
- Fotis Jannidis: *Der Autor ganz nah. Autorstil in Stilistik und Stilometrie*. In: *Theorien und Praktiken der Autorschaft*. Hg. von Matthias Schaffrick / Marcus Willand. Berlin, Boston 2014, S. 169–195. [[Nachweis im GBV](#)]
- Hans Robert Jauf: *Theorie der Gattungen und Literatur des Mittelalters*. In: *Hans Robert Jauf: Alterität und Modernität der mittelalterlichen Literatur. Gesammelte Aufsätze 1956–1976*. München 1977, S. 327–358. [[Nachweis im OPAC](#)]
- Matthew Jockers: *Macroanalysis. digital methods and literary history*. Champaign 2013. Champaign 2013. [[Nachweis im OPAC](#)]
- Mike Kestemont: *Arthur's Authors. A Quantitative Study of the Rhyme Words in the Middle Dutch Arthurian Epic*. In: *Zeitschrift für deutsches Altertum* 142 (2013), H. 1, S. 1–33. [[Nachweis im OPAC](#)]
- Mike Kestemont: *Function Words in Authorship Attribution. From Black Magic to Theory?* In: *forProceedings of the 3rd Workshop on Computational Linguistics Literature (ClfL)*. Göteborg 2014, S. 59–66. [[online](#)]
- Thomas Klein: *Ermittlung, Darstellung und Deutung von Verbreitungstypen in der Handschriftenüberlieferung mittelhochdeutscher Epik*. In: *Deutsche Handschriften 1100–1400. Oxforder Kolloquium 1985*. Hg. von Volker Honemann / Nigel F. Palmer. Tübingen 1988, S. 110–167. [[Nachweis im OPAC](#)]
- Erich Kleinschmidt: *Minnesang als höfisches Zeremonialhandeln*. In: *Archiv für Kulturgeschichte* 58 (1976), S. 35–76. [[Nachweis im OPAC](#)]
- Jurij M. Lotman: *Die Struktur literarischer Texte*. 4., unveränderte Auflage. München 1993. [[Nachweis im GBV](#)]
- Wolfgang Mohr: *Hilfe und Rat in Wolframs Parzival*. In: *Festschrift für Jost Trier zu seinem 60. Geburtstag am 15. Dezember 1954*. Hg. von Benno von Wiese / Karl Heinz Borch. Meisenheim 1954, S. 173–197. [[Nachweis im GBV](#)]
- Franco Moretti: *Conjectures on World Literature*. In: *New Left Review* 1 (2000), S. 54–68. [[Nachweis im GBV](#)]
- Franco Moretti: *Graphs – Maps – Trees*. London, New York 2005. [[Nachweis im GBV](#)]
- Hermann Paul: *Mittelhochdeutsche Grammatik*. 25. Auflage. Tübingen 2007. [[Nachweis im OPAC](#)]
- R Core Team: *A language and environment for statistical computing*. R Foundation for Statistical Computing. Wien 2013. [[online](#)]
- Patrick Sahle: *Digitale Editionsformen. Zum Umgang mit der Überlieferung unter den Bedingungen des Medienwandels*, 3 Bände. Norderstedt 2013 (= *Schriften des Instituts für Dokumentologie und Editorik* 7–9). [[Nachweis im OPAC](#)]
- Arno Schirokauer: *Studien zur mittelhochdeutschen Reimgrammatik*. In: *Beiträge zur Geschichte der deutschen Sprache und Literatur (PBB)* 47 (1923), S. 1–126. [[Nachweis im OPAC](#)]
- Rüdiger Schnell: *»Autor« und »Werk« im deutschen Mittelalter. Forschungskritik und Forschungsperspektiven*. In: *Neue Wege der Mittelalter-Philologie. Landshuter Kolloquium 1996*. Hg. von Joachim Heinze / L. Peter Johnson / Gisela Vollmann-Profe. Berlin 1998 (= *Wolfram-Studien* 15), S. 12–73. [[Nachweis im OPAC](#)]

Rüdiger Schnell: Minnesang und Sangspruch im 13. Jahrhundert. Gattungsdifferenzen und Gattungsinterferenzen. In: Transformationen der Lyrik im 13. Jahrhundert. Wildbader Kolloquium 2008. In Verbindung mit Eckart Conrad Lutz und Klaus Ridder hg. von Susanne Köbele. Berlin 2013 (= Wolfram-Studien 21), S. 287–348. [\[Nachweis im OPAC\]](#)

Christoph Schöch: Corneille, Molière et les autres. Stilometrische Analysen zu Autorschaft und Gattungszugehörigkeit im französischen Theater der Klassik. In: Literaturwissenschaft im digitalen Medienwandel. Hg. von Christof Schöch / Lars Schneider. Philologie im Netz, Beiheft 7 (2014), S. 130–157. [\[online\]](#)

Stéfan Sinclair / Geoffrey Rockwell / the Voyant Tools Team: Voyant Tools (web application). 2012. [\[online\]](#)

Karl Stackmann: Neue Philologie? In: Modernes Mittelalter. Neue Bilder einer populären Epoche. Hg. von Joachim Heinze. Frankfurt/Main 1994, S. 398–427. [\[Nachweis im GBV\]](#)

Peter Strohschneider: Tanzen und Singen. Leichs von Ulrich von Winterstetten, Heinrich von Sax sowie dem Tannhäuser und die Frage nach dem rituellen Status des Minnesangs. In: Mittelalterliche Lyrik. Probleme der Poetik. Hg. von Thomas Cramer / Ingrid Kasten. Berlin 1999, S. 197–231. [\[Nachweis im OPAC\]](#)

Helmut Tervooren: Sangspruchdichtung. Stuttgart, Weimar 1995. [\[Nachweis im OPAC\]](#)

Wilhelm Voßkamp: Gattungen als literarisch-soziale Institutionen. Zu Problemen sozial- und funktionsgeschichtlich orientierter Gattungstheorie und -historie. In: Textsortenlehre – Gattungsgeschichte. Hg. von Walter Hinck. Heidelberg 1977, S. 27–44. [\[Nachweis im OPAC\]](#)

Franz Josef Worstbrock: Wiedererzählen und Übersetzen. In: Mittelalter und frühe Neuzeit. Übergänge, Umbrüche und Neuansätze. Hg. von Walter Haug. Tübingen 1999, S. 128–142. [\[Nachweis im OPAC\]](#)

Abbildungslegenden und -nachweise

Abb. 1: SClusteranalyse der höfischen Epik (Graphik: Autor).

Abb. 2: Kontrastive Analyse der Werke Hartmanns, Wolframs und Gottfrieds (Graphik: Autor).

Abb. 3: Liste der von Wolfram gegenüber Hartmann bevorzugten bzw. vermiedenen Wörter (Graphik: Autor).

Abb. 4: Wordcloud Minnesangs Frühling (Graphik: Autor).