



GABRIEL VIEHHAUSER

# Mittelalterliche Texte als Modellierungsaufgabe

## Digitale Zugänge zur mittelhochdeutschen Literatur

Die Frage, ob digitale Methoden in den Geisteswissenschaften etwas qualitativ Neues und Eigenes mit sich bringen oder ob sie bloß performantere Verfahren sind, um bereits bestehende Forschung zu unterstützen, hat die Theoriediskussion in den Digital Humanities lange beschäftigt.<sup>1</sup> Ich möchte im Folgenden darauf aus mediävistischer Sicht mögliche Antworten geben sowie zugleich einige digitale Verfahren vorstellen und diskutieren, die sich im Bereich der Altgermanistik einsetzen ließen.<sup>2</sup> Als Einstieg beginne ich hierfür mit einem Beispiel aus der Parzival-Philologie.

### I.

Abbildung 1 zeigt die erste Seite des Kassler Exemplars einer frühen Druckausgabe von Wolframs *Parzival*, nämlich den Abdruck der Sankt Galler Epenhandschrift (Cod. 857), die der Bodmer-Schüler Christoph Heinrich Myller im Jahr 1784 publiziert hat.<sup>3</sup> Wie die zahlreichen im Faksimile ersichtlichen handschriftlichen Eintragungen zwischen und neben den Zeilen verraten, handelt es sich beim Kassler Exemplar nicht

---

<sup>1</sup> Vgl. hierzu etwa zusammenfassend Manfred THALLER: Digital Humanities als Wissenschaft, in: Digital Humanities. Eine Einführung, hrsg. v. Fotis Jannidis/Hubertus Kohle/Malte Rehbein, Stuttgart 2017, S. 13–18, oder die Beiträge in: Debates in the Digital Humanities, hrsg. v. Matthew K. Gold, Minneapolis/London 2012.

<sup>2</sup> Der Aufsatz stellt eine überarbeitete Fassung meines Vortrags auf der Bamberger Tagung dar, die Vortragsform wurde im Wesentlichen beibehalten.

<sup>3</sup> WOLFRAM VON ESCHENBACH: Parzival. Ein Ritter-Gedicht aus dem dreizehnten Jahrhundert von Wolfram von Eschilbach. Zum zweiten Male aus der Handschrift abgedruckt, weil der erste Anno 1477 gemachte Abdruck so selten wie Manuscript ist, in: Sammlung deutscher Gedichte aus dem XII., XIII. und XIV. Jahrhundert [hrsg. v. Christoph Heinrich Müller (Myller)], Berlin 1784.

um eine beliebige Ausfertigung des Drucks, sondern um eines von zwei Handexemplaren, die niemand geringerer als Karl Lachmann auf seiner berühmten Wolfram-Reise in Verwendung hatte.



Um seine Ausgabe der Werke Wolframs vorzubereiten, war Lachmann von Berlin nach Heidelberg, München und Sankt Gallen gereist,

um die dort befindlichen Parzival-Handschriften einzusehen.<sup>4</sup> Die Lesarten der Handschriften hat Lachmann dabei per Hand in den Myller'schen Druck eingetragen. Das Handexemplar wurde damit letztlich zur Grundlage für Lachmanns editorisches Meisterstück, für seine Wolfram-Ausgabe aus dem Jahr 1833.

In seiner Ausgabe hat Lachmann bekanntlich die Fülle der Überlieferung von Wolframs *Parzival* auf einen rekonstruierten Editions-Text reduziert. Insgesamt sind heute noch 16 vollständige Handschriften des *Parzival* bekannt, dazu kommen noch ca. 70 Fragmente. Ein nicht unbedeutlicher Teil davon war Lachmann bereits bekannt;<sup>5</sup> dennoch ist in seiner Ausgabe, die in nur gering modifizierter Form bis heute in Verwendung ist, bloß ein einzelner Editionstext zu lesen, den der Herausgeber aus der handschriftlichen Überlieferung destilliert hat. Die Lesarten der Handschriften sind lediglich im Apparat verzeichnet, und dies auch nur, insoweit sie die editorischen Entscheidungen des Herausgebers nachvollziehbar machen.

Dieses Verfahren der Rekonstruktion von letztlich als original gedachten Texten hat bekanntlich im Laufe der Forschungsgeschichte viel Widerspruch hervorgerufen, da seine Angemessenheit für die grundsätzlich variable Überlieferung volkssprachiger mittelalterlicher Texte in

---

<sup>4</sup> Vgl. hierzu Friedrich NEUMANN: Karl Lachmanns Wolframreise, in: Jahrbuch der Albertus-Universität zu Königsberg / Pr. 2, Freiburg/Frankfurt a.M. 1952, S. 138–158; Bernd SCHIROK: Der Aufbau von Wolframs Parzival, Freiburg 1972, hier S. 591–628; Marc R. MCCULLOH: Myller's Parzival and Lachmann's Critical Method: the 'Wolfram-Reise' Revisited, in: MLN 98 (1983), S. 484–491; Robert SCHÖLLER: Von der Handschrift zum Druck: zur Editions-geschichte des Parzival, in: Schachzabel, Edelstein und der Gral. Spätmittelalterliche Handschriftenschatze der Burgerbibliothek Bern, hrsg. v. der Burgerbibliothek Bern (Passepartout 1), Bern 2009, S. 73–77; Volker MERTENS: Die Wiederentdeckung Wolframs und die Anfänge der Forschung, in: Wolfram von Eschenbach. Ein Handbuch, hrsg. v. Joachim Heinze, Berlin / Boston 2011, S. 706–741, insbesondere S. 724–730.

<sup>5</sup> Vgl. Bernd SCHIROK: Einführung zum Text der Lachmannschen Ausgabe, in: Parzival. Studienausgabe. Mhd. Text nach der sechsten Ausgabe von Karl Lachmann. Übersetzung v. Peter Knecht. Mit Einführungen zum Text der Lachmannschen Ausgabe und in Probleme der Parzival-Interpretation v. Bernd Schirok, 2. Aufl., Berlin / New York 2003, S. LXIX.

Frage gestellt wurde.<sup>6</sup> Dies ist hinlänglich bekannt, worauf es mir in diesem Zusammenhang aber insbesondere ankommt, ist der Konnex von Lachmanns Methode zum Kassler Reise-Exemplar: Denn dieses beweist meiner Ansicht nach recht schlagend, dass Lachmanns Idealvorstellung, wie man zu edieren hat, offensichtlich auch Hand in Hand ging mit den technischen Voraussetzungen, denen Lachmann unterworfen war: Denn wenn man sozusagen nur einmal im Leben die Möglichkeit hat, die entsprechenden Handschriften einzusehen und deren Lesarten auch noch von Hand in einen Abdruck von zweifelhafter Qualität eintragen muss, dann liegt es auch nahe, dass man sich schon zu Beginn einer solchen Reise wie der Wolfram-Reise gut überlegen muss, wie man die Überlieferung vorsortiert. Angesichts solcher Produktionsbedingungen ist es eigentlich nur zu verständlich, dass Lachmann das Quellenmaterial nicht in seiner gesamten Fülle berücksichtigt hat. Seine Entscheidung hierfür mag aus konzeptionellen Überlegungen getroffen worden sein, sie wird durch die medialen Bedingungen aber auch notwendig gemacht bzw. befördert.

Dass sich heute, unter digitalen Bedingungen, andere Wege ergeben, zeigen Editionen wie etwa die digitale Parzival-Edition, die zurzeit unter der Leitung von Michael Stolz an der Universität Bern entsteht.<sup>7</sup> Diese Edition bietet nicht nur einen Parzival-Text, sondern gleich vier Parzival-Fassungen, die gleichberechtigt nebeneinanderstehen (vgl. hierzu die Editionsprobe in Abbildung 2).

---

<sup>6</sup> Vgl. hierzu zusammenfassend Thomas BEIN: Textkritik. Eine Einführung in Grundlagen germanistisch-mediävistischer Editionswissenschaft. Lehrbuch mit Übungsteil. 2., überarb. u. erw. Auflage, Frankfurt a.M. u.a. 2011, hier: S. 73–97.

<sup>7</sup> Vgl. hierzu die Homepage <http://www.parzival.unibe.ch> (Parzival-Projekt).

D	M	G	T
*D 453.01 Swer mich dār von è vrāgte	*M 453.03 ob ich es ime niht esaget,	*G 453.01 Swer mich drumbe vrāgte	*T 453.01 Swer mich dār von è vrāgte
*D 453.02 unt dār umbe mit mir blāgte,	*M 453.04 unprīs er dār an bejāgete.	*G 453.02 unt drumbe mit mir blāgte,	*T 453.02 und drumbe mit mir blāgte,
*D 453.03 ob ichs im niht sagte,	*M 453.05 mich bat ez heln Kyot,	*G 453.03 ob ichs im niht sagete,	*T 453.03 ob ichs im niht esagetē,
*D 453.04 unprīs der dran bejāgte.	*M 453.06 wan im diu āventiure gebōt,	*G 453.04 unbrīs er dran bejāgete.	*T 453.04 unprīs er dran bejāgete.
*D 453.05 mich batz heln Kyot,	*M 453.07 daz es immer man gedaht,	*G 453.05 mich bat ez heln Kyot,	*T 453.05 mich bat ez heln Kyot
*D 453.06 wand im diu āventiure gebōt,	*M 453.08 è ez diu āventiure braht	*G 453.06 wande im diu āventiure gebōt,	*T 453.06 wand in diu āventiure gebōt,
*D 453.07 daz es immer man gedahte,	*M 453.09 mit worten an der mār gruoz,	*G 453.07 daz es immer man gedahte,	*T 453.07 daz ers iemer man gedahte,
*D 453.08 è ez diu āventiure brahte	*M 453.10 daz man dā von doch sprechen muoz.	*G 453.08 è ez diu āventiure brahte	*T 453.08 è ez diu āventiure brahte
*D 453.09 mit worten an der mære gruoz,	*M 453.11 Kyot, der meister wol bekant,	*G 453.09 mit worten an der mære gruoz,	*T 453.09 mit worten an der mære gruoz,
*D 453.10 daz man dār von doch sprechen muoz.	*M 453.12 zuo Dolet verworfen ligen vant	*G 453.10 daz man dār von nū sprechen muoz.	*T 453.10 daz man dā von doch sprechen muoz.
*D 453.11 Kyot, der meister wol bekant,	*M 453.13 in heidenischer schrift	*G 453.11 Kyot, der meister wol bekant,	*T 453.11 Kyot, der meister wol bekant,
*D 453.12 ze Dolet verworfen ligen vant	*M 453.14 diser āventiure gestift.	*G 453.12 ze Dolet verworfen ligen vant	*T 453.12 ze Dolet verworfen ligen vant
*D 453.13 in heidenischer schrifte	*M 453.15 den karakter a b c	*G 453.13 in heidenischer schrifte	*T 453.13 in heidenischer schrifte
*D 453.14 dirre āventiure gestifte.	*M 453.16 muos er hān gelernet è,	*G 453.14 dirre āventiure stifte.	*T 453.14 dirre āventiure gestifte.
*D 453.15 der karakter a b c	*M 453.17 āne den list von nigromanzi.	*G 453.15 der karakter a b c	*T 453.15 den karakter a b c
*D 453.16 muoser hān gelernet è,	*M 453.18 ez half, daz im der touf was bl.	*G 453.16 müese er haben gelernet è,	*T 453.16 mües er hān gelernet è,

D 5	m n o	G I O L M Z	T U V W Q R
Apparat 1	Apparat 1	Apparat 1	Apparat 1
453.01 <i>Initiale FfS</i>	453.01 <i>Überschrift:</i> Also parzial gen treuende dem einseidel kam m Also parzial gon treuende zū dem einseydel kam in	453.01 <i>Überschrift:</i> Hie ist parzial zv dem kloesner: zv fontane komen der sagt im alle gelegenheit vmb den gral wie er dar zv	453.01 <i>Überschrift:</i> Awentewr wie parzial bericht wart vmb den gral Q • <i>Initiale V W • Großinitiale Q R</i>

Damit ist das von Joachim Bumke noch analog entwickelte Konzept der Fassungsedition<sup>8</sup> digital umgesetzt. Auch eine Fassungsedition stellt im gewissen Sinne eine Reduktion der Überlieferung dar, werden doch die heute bekannten ca. 86 Textzeugen auf vier Fassungen heruntergebrochen. Das hat im Fall des *Parzival* gute konzeptionelle Gründe, überlieferungsgeschichtliche Studien haben gezeigt, dass sich alle Textzeugen diesen vier Gruppen zuordnen lassen. Die Unterschiede innerhalb der

<sup>8</sup> Das Konzept wurde theoretisch entwickelt in Joachim BUMKE: Epenhandschriften. Vorüberlegungen und Informationen zur Überlieferungsgeschichte der höfischen Epik im 12. und 13. Jahrhundert, in: Philologie als Kulturwissenschaft. Studien zur Literatur und Geschichte des Mittelalters. Festschrift für Karl Stackmann zum 65. Geburtstag, Göttingen 1987, S. 45–59; Joachim BUMKE: Die vier Fassungen der ‚Nibelungenklage‘. Untersuchungen zur Überlieferungsgeschichte und Textkritik der höfischen Epik im 13. Jahrhundert, Berlin / New York 1996; sowie umgesetzt in der Edition: Die ‚Nibelungenklage‘. Synoptische Ausgabe aller vier Fassungen, hrsg. v. Joachim Bumke, Berlin / New York 1999.

Handschriften der Gruppen sind in der Regel nicht allzu groß, zumindest nicht so groß wie die Unterschiede zwischen den vier Fassungen.<sup>9</sup>

In der digitalen Edition des *Parzival* sind die Differenzen zwischen den Fassungen relativ schnell erkennbar, sie sind in den vier Editionstexten fett markiert. Die einzelnen Handschriften, die zu den vier Fassungen gehören, sind unterhalb der Haupttexte aufgelistet, und es gibt für jede Fassung einen Variantenapparat, der jetzt aber nur mehr die vergleichsweise geringen Abweichungen der Handschriften zu ihrer Fassung verzeichnet.

Da die Edition digital angelegt ist, bleibt die Fassungsansicht jedoch nicht die einzige Option der Darstellung: Durch einen Klick auf den Link in den Variantenapparat ist es möglich, auch die einzelnen Handschriften mit Transkription und Faksimiles aufzurufen und so den Blick auf die ganze Überlieferung auszuweiten. Zwar bildet die Beschränkung auf Fassungstexte den konzeptionellen Ausgangspunkt der Edition, doch muss die Benutzerin oder der Benutzer nicht dabei stehen bleiben, sondern kann auf die ganze Überlieferung zugreifen.

Digitale Editionen bringen also eine Ausweitung der Perspektiven mit sich, und das selbst in mehrfacher Hinsicht:<sup>10</sup> Zum einen ermöglicht das digitale Medium den Blick auf die Gesamtüberlieferung, auf alle Handschriften. Zum anderen erlaubt es darüber hinausgehend aber auch, unterschiedliche editorische Herangehensweisen nebeneinander zu stellen, im Fall der *Parzival*-Edition etwa auf der einen Seite

---

<sup>9</sup> Vgl. hierzu Michael STOLZ: Intermediales Edieren am Beispiel des *Parzival*-Projekts, in: *Wege zum Text. Beiträge des Grazer Kolloquiums über die Verfügbarkeit mediävistischer Editionen im 21. Jahrhundert (17.–19. September 2008)*, hrsg. v. Wernfried Hofmeister / Andrea Hofmeister-Winter, Tübingen 2009 (Beihefte zu editio 30), S. 213–228; Robert SCHÖLLER: Die Fassung \*T des *Parzival* Wolframs von Eschenbach. Untersuchungen zur Überlieferung und zum Textprofil (Quellen und Forschungen zur Literatur- und Kulturgeschichte 56 [290]), Berlin/New York 2009; Gabriel VIEHHAUSER-MERY: Die *Parzival*-Überlieferung am Ausgang des Manuskriptzeitalters. Handschriften der Lauberwerkstatt und der Straßburger Druck (Quellen und Forschungen zur Literatur- und Kulturgeschichte 55 [289]), Berlin/New York 2009.

<sup>10</sup> Dieser Umstand wird ausführlich beschrieben bei Patrick SAHLE: *Digitale Editionsformen. Zum Umgang mit der Überlieferung unter den Bedingungen des Medienwandels* (3 Bde.), Norderstedt 2013, Bd. 2, S. 157–280.

rekonstruierte Texte wie die vier Fassungstexte, auf der anderen Seite ein radikal handschriftennahes Tableau aller einzelnen Textzeugen. Man könnte demnach geradezu sagen, dass das Digitale damit die Grundlage dafür schafft, die widersprüchlichen Positionen der letzten 200 Jahre Editions wissenschaft zu versöhnen, nämlich eine mehr auf das Original und eine mehr auf die Überlieferung ausgerichtete Art zu edieren.

Ausgangspunkt dafür – und das ist wieder das Entscheidende für meine Argumentation – sind die technischen Voraussetzungen, das digitale Medium. Zwar zwingen digitale Editionen nicht unbedingt dazu, unterschiedliche Perspektiven auf Texte einzunehmen, das digitale Medium mit seiner Hyperlinkstruktur und seiner potentiellen Grenzenlosigkeit befördert diese Herangehensweise aber; es „prämiert“ diese Herangehensweise, wie Patrick Sahle das mit einem an Michael Giesecke angelehnten Begriff bezeichnet hat.<sup>11</sup> Mediale Bedingungen und konzeptionelle Überlegungen gehen Hand in Hand; bei Lachmann bestand schon aus pragmatischen Gründen die Notwendigkeit zur Konzentration: Wenn aus arbeitstechnischen Gründen die ganze Überlieferung des *Parzival* in ein Handexemplar gezwungen werden muss, dann ergibt sich auch die Notwendigkeit, vorher zu überlegen, auszuwählen und zu selektieren. Im Ergebnis der Edition finden dementsprechend nur die besten Lesarten Berücksichtigung. Bei der digitalen Edition hingegen gibt es fast unbegrenzt Platz, daher können alle Texte und sogar bis zu einem gewissen Grad alle möglichen editorischen Zugänge nebeneinander präsentiert werden. Die technischen Möglichkeiten, wie man edieren kann, und die Vorstellungen, wie man das sollte, bedingen sich offenbar bis zu einem gewissen Grad gegenseitig.

Freilich sind damit Tendenzen beschrieben, die nicht immer im Einzelnen voll ausgeprägt erscheinen müssen. Es ließe sich zum Beispiel einwenden, dass im digitalen Medium die Darstellungsmöglichkeiten auch nicht zur Gänze unbeschränkt sind: Mehr als vier Fassungen auf einem Bildschirm darzustellen, wäre etwa recht unübersichtlich, und

---

<sup>11</sup> SAHLE: Digitale Editionsformen, Bd. 2, S. 172; sowie Patrick SAHLE: Zwischen Medien gebundenheit und Transmedialisierung. Anmerkungen zum Verhältnis von Edition und Medien, in: editio 24 (2010), S. 23–36, hier: S. 23.

dieser Umstand hat sicher mit dazu beigetragen, dass für die Parzival-Edition ausgerechnet vier Fassungen zum Tragen kommen.<sup>12</sup> Aber gerade dies bestätigt den Befund: Auch hier bestimmt das Medium mit, wie eine Edition aussieht.<sup>13</sup> Sieht man aber von der Bildschirmbeschränkung einmal ab (die ohnedies durch die Möglichkeit der Verlinkung von Seiten zum Teil ausgehebelt wird), dann hat man im Digitalen zumindest potentiell die Möglichkeit, immer weitere Texte an einen Ausgangstext anzulagern. Das Internet ist durch keine Marginalräume der Buchseite beschränkt<sup>14</sup> und damit so grenzenlos, dass man in der Fülle der Möglichkeiten fast die Orientierung verlieren könnte: Das Problem im Digitalen ist mithin sogar eher jenes, nicht zu ausufernd zu werden, und es ist bezeichnend, dass eine nicht untypische Reaktion auf digitale Editionen die Befürchtung zu sein scheint, einen Text nun tatsächlich mehrmals in unterschiedlichen Fassungen lesen zu müssen. Und welcher dieser Texte ist dann überhaupt der ‚richtige‘ Text, an den man sich halten und den man referenzieren kann? Etwas polemisch überspitzt könnte man sagen, dass wir uns sicherer fühlen, wenn wir Bücher in die Tasche stecken und unser Wissen zwischen den begrenzten Raum von Buchdeckeln bannen können. Es mag so manche Leserin oder so manchen Leser geben, die vor einer neuen Lektüre darauf schielen, wie viel Seiten das zu bewältigende Buch denn überhaupt oder, bei fortgeschrittener Lektüre, ‚noch‘ hat.

---

<sup>12</sup> Auch in der noch analog erstellten Musteredition Joachim Bumkes zu den vier Fassungen der ‚Nibelungenklage‘ (BUMKE: Die ‚Nibelungenklage‘) ist eine Zahl von Fassungen gewählt, die gerade noch auf einer aufgeschlagenen Buchseite synoptisch präsentiert werden kann.

<sup>13</sup> Mediale Formen haben zudem bekanntlich auch ein gewisses Beharrungselement (vgl. hierzu Michael GIESECKE: Von den Mythen der Buchkultur zu den Visionen der Informationsgesellschaft. Trendforschungen zur kulturellen Medienökologie, Frankfurt a. M. 2002, S. 271ff). Manches, was wir am Computer tun, ist noch vom alten Medium abhängig: In der Editionsprobe aus Abbildung 2, die einen frühen Entwurf der Editionsdarstellung im Parzival-Projekt wiedergibt, befindet sich zum Beispiel der Apparat noch immer am Seitenende, was zwar dem Usus von Printausgaben entspricht, am Bildschirm aber gar nicht nötig wäre. Dieses unnötige Relikt ist in neueren Versionen der Edition beseitigt, hier erscheinen die Apparateinträge beim Überfahren des Haupttextes direkt neben der entsprechenden Zeile.

<sup>14</sup> Vgl. zum regulierenden Effekt der Marginalräume Hans Ulrich GUMBRECHT: The Powers of Philology: Dynamics of Textual Scholarship, Urbana/Chicago 2003, S. 41–53.



Seiten zählen kann man aber nun paradoxerweise gerade im Digitalen nicht mehr, in diesem undefinierten Bereich mit der Tendenz zur dauernden Abschweifung.

Dass es aber nun vielleicht sogar gut ist, dass Editionen keine Sicherheit erzeugen, sondern im Gegenteil sogar ein gewisses Maß an Unsicherheit herausstellen sollen, ist ein Gedanke, den es schon lange vor dem digitalen Zeitalter gab. Sehr prominent findet er sich etwa bei Karl Stackmann in dem bereits 1964 publizierten Aufsatz „Mittelalterliche Texte als Aufgabe“, der als bahnbrechend für eine neue editionstheoretische Ausrichtung der Altgermanistik gilt, die sich vom Lachmann'schen Ideal der Textkritik hin zur Anerkennung der Eigenart mittelalterlicher Überlieferung bewegt hat.<sup>15</sup> Vieles in der Überlieferung der mittelhochdeutschen Literatur lässt sich nicht genau sagen und festlegen, weshalb Stackmann es als vordringlichste Aufgabe von (zu seiner Zeit) modernen Textausgaben ansah, eben nicht Sicherheit über den Text, sondern im Gegenteil „ein höchstes Maß an Unsicherheit“ zu erzeugen.<sup>16</sup> Gerade dies können, wie das Beispiel gezeigt hat, aber nun digitale Editionen besonders gut.

Nimmt man diesen Gedanken ernst, dann zeigt sich, dass etwas überraschend und entgegen den klischeehaften Erwartungen, die man vielleicht an technokratische Zukunftsvisionen von der Allmacht formalistischer Rechenmaschinen knüpft, der Computer ganz im Gegenteil zunächst einmal etwas mit sich bringt, das man als genuin geisteswissenschaftlich bezeichnen könnte, nämlich ein gesteigertes Maß an Multiperspektivität, die immer wieder als das genuine Feld der Expertise der Geisteswissenschaften bezeichnet wurde.<sup>17</sup> Komplexität statt Reduktion, vier

---

<sup>15</sup> Karl STACKMANN: Mittelalterliche Texte als Aufgabe, in: Festschrift für Jost Trier zum 70. Geburtstag, hrsg. v. W. Foerste/K. H. Borck, Köln/Graz 1964, S. 240–267, wieder in: Karl STACKMANN: Mittelalterliche Texte als Aufgabe. Kleine Schriften 1, hrsg. v. Jens Haustein, Göttingen 1997, S. 1–23.

<sup>16</sup> STACKMANN: Mittelalterliche Texte, S. 23.

<sup>17</sup> Vgl. etwa Odo MARQUARD: Über die Unvermeidlichkeit der Geisteswissenschaften, in: Apologie des Zufälligen: Philosophische Studien, hrsg. v. Odo Marquard, Stuttgart, 1986, S. 98–116.

Parzivale statt ein Parzival, die noch dazu unterschiedlich tief rekonstruierbar sind: Fast scheint es, als hätten ausgerechnet die Digital Humanities ein besonderes Potential, als Bewahrer der alten, fachübergreifenden geisteswissenschaftlichen Idee der Komplexität zu fungieren.

## II.

Um die Generalisierbarkeit dieses Gedankens zu überprüfen, möchte ich mich im Folgenden vom Feld der Editorik wegbewegen und der digitalen Textanalyse zuwenden, die in den letzten Jahren in den Digital Humanities vermehrt Aufmerksamkeit auf sich gezogen hat. Die digitale Textanalyse wurde wesentlich durch eine geschickt ausformulierte theoretische Perspektive befördert, die der amerikanische Komparatist Franco Moretti in die Diskussion mit eingebracht hat. Die Rede ist vom sogenannten ‚Distant Reading‘, das als Schlagwort die Vorstellung von digitalen Analysezugängen wesentlich geprägt hat.<sup>18</sup>

Die Grundkonstellation, von der Moretti ausgeht, ist eine Art Entsetzen über die Komplexität der Welt, und zwar genauer gesagt, ein Entsetzen über die Menge an Büchern, die es eigentlich gibt. Laut einer Schätzung der Firma Google sind dies zurzeit ziemlich genau 129.864.880 Bücher<sup>19</sup> und es versteht sich von selbst, dass selbst ein eifriger Leser oder auch eine eifrige Leserin nicht einmal einen Bruchteil dieser Bücher überblicken kann. Die traditionelle Literaturwissenschaft, so Moretti, hat sich über diesen Umstand hinweggeholfen, indem sie immer nur um einen kleinen Ausschnitt von Texten gekreist ist, den Kanon, der auf subjektiven und ideologischen Kriterien beruht. Wenn man demgegenüber die Wahrheit über die ganze Textproduktion erfahren will, dann, so Moretti in seiner berühmten provokativen Formulierung, nützt es nichts, zu versuchen, immer mehr Bücher zu lesen, sondern man muss im

---

<sup>18</sup> Vgl. die Ausformulierung des Konzepts in Franco MORETTI: *Conjectures on World Literature*, in: *New Left Review* 1 (2000), S. 54–68.

<sup>19</sup> Vgl. Leonid TAYCHER: *Books of the World, Stand up and Be Counted! All 129,864,880 of you*, 2010 [Web Log-Eintrag], abgerufen von: <http://booksearch.blogspot.de/2010/08/books-of-world-stand-up-and-be-counted.html>.

Gegenteil gerade damit aufhören zu lesen.<sup>20</sup> Nur wer Bücher nicht liest, behält den Überblick.<sup>21</sup> Der Begriff ‚Distant Reading‘, den Moretti als Alternative vorschlägt, versteht sich daher als Gegenbegriff zum klassischen ‚Close Reading‘, das Moretti aus seiner Sicht als gängiges Verfahren der traditionellen Literaturwissenschaft identifiziert.

Was nun Distant Reading genau ist und in der Praxis bedeutet, ist letztlich nicht klar abgrenzbar. Klar ist, dass sich das Verfahren tendenziell auf die Makroperspektive richtet, auf größere strukturelle Muster, die für den Neo-Marxisten Moretti immer auch Muster gesellschaftlicher Machtmechanismen sind.<sup>22</sup> Man könnte sogar sagen, dass Moretti Texte in erster Linie als Material für die Aufdeckung von solchen Machtmechanismen ansieht, weshalb ihm der Verzicht auf einen ästhetisch definierten Kanon natürlich leichtfällt.<sup>23</sup>

Zum Distant Reading eignet sich nun der Einsatz von Computermethoden besonders gut. Und auch beim Distant Reading ließe sich so gesehen also beobachten, dass die ursprüngliche Stoßrichtung der Methode ebenso wie bei der digitalen Editorik darauf abzielt, eine Ausweitung der Perspektiven zu befördern: Statt den Blick lediglich auf den bildungsbürgerlichen Kanon zu richten, soll der Untersuchungsgegenstand ausgeweitet werden auf außerliterarische, außereuropäische, außerpatriarchale Texte. Beste Voraussetzungen also eigentlich für eine Anwendung von Distant Reading-Verfahren auf mittelalterliche Literatur, die ja schon seit

---

<sup>20</sup> Vgl. MORETTI: *Conjectures*, S. 57.

<sup>21</sup> Damit gleicht Morettis Credo dem des Bibliothekars im 100. Kapitel („General Stumm dringt in die Staatsbibliothek ein und sammelt Erfahrungen über Bibliothekare, Bibliotheksdiener und geistige Ordnung“) in Robert Musils ‚Mann ohne Eigenschaften‘. Zu dieser Koinzidenz vgl. Gabriel VIEHHAUSER: Digital Humanities als Geisteswissenschaften. Zur Auflösung einer Tautologie, in: *Digital Humanities. Perspektiven der Praxis*, hrsg. v. Peggy Bockwinkel / Beatrice Nickel / Gabriel Viehhauser, Berlin 2018 (*Digitalisierung und Globalisierung der Wissenschaften* 1), S. 17–41, hier: S. 32–33.

<sup>22</sup> Vgl. MORETTI: *Conjectures*, S. 66.

<sup>23</sup> Zur Einbettung des Distant Reading-Ansatzes in eine literatursoziologische Tradition vgl. auch Ted UNDERWOOD: A Genealogy of Distant Reading, in: *Digital Humanities Quarterly* 11 (2017).

Langem mit einem erweiterten, nicht der Genieästhetik verhafteten Literaturbegriff rechnet.

Sieht man sich aber nun die Methoden an, die zurzeit in der digitalen Textanalyse zur Anwendung kommen, dann muss man diese Erwartungen wohl ein wenig zurückschrauben. Ein Verfahren, das vor wenigen Jahren in den Digital Humanities intensiv erprobt wurde, ist etwa die sogenannte Stilometrie. Stilometrie meint die automatische Klassifikation von Texten aufgrund von quantitativ auszählbaren stilistischen Eigenschaften. Ihre besonderen Erfolge hat die Stilometrie bei der Autorschaftserkennung gefeiert, also bei der Frage, ob ein anonym überlieferter Text einem namentlich bekannten Autor aus einem Vergleichskorpus zugeordnet werden kann. Dazu hat sich ein sehr einfaches Feature als sehr zielführend erwiesen, nämlich die simple Auszählung der *most frequent words*, also der häufigsten Wörter, die in einem Text vorkommen.<sup>24</sup> Diese sind nun nicht thematisch bedeutende Wörter, sondern sogenannte Funktionswörter, also Konjunktionen, Artikel, Präpositionen usw. Die Häufigkeit, wie oft wir diese Wörter verwenden, ist also offensichtlich sehr verräterisch für unsere spezifische Art zu schreiben und folgt einer unbewusst eingelernten Signatur.

Nun ist aber gerade Autorschaftsbestimmung ein Thema, das in der Literaturwissenschaft des 21. Jahrhunderts alles andere als aktuell erscheint.<sup>25</sup> Das hat gerade in der Altgermanistik gute Gründe, zum einem forschungsgeschichtlich, weil die Klärung von Echtheitsfragen und Zuschreibungen von Texten zu bestimmten Autoren an dunkle Zeiten der

---

<sup>24</sup> Vgl. Jan HORSTMANN: Stilometrie, in: forTEXT. Literatur digital erforschen (2018), <https://fortext.net/routinen/methoden/stilometrie>; Christof SCHÖCH: Quantitative Textanalyse, in: Digital Humanities. Eine Einführung, hrsg. v. Fotis Jannidis/Hubertus Kohle/Malte Rehbein, Stuttgart 2017, S. 279–298, hier: S. 292–296; Fotis JANNIDIS: Der Autor ganz nah – Autorstil in Stilistik und Stilometrie, in: Theorien und Praktiken der Autorschaft, hrsg. v. Matthias Schaffrick/Marcus Willand, Berlin 2014, S. 169–195.

<sup>25</sup> Vgl. allgemein zum schwierigen Umgang mit der Stilistik in der Altgermanistik Jens HAUSTEIN: 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, hrsg. v. Ulrich Breuer/Bernhard Spies, Bielefeld 2011, S. 43–60.

Philologie erinnert, in der divinatorische Herausgeber genau zu wissen vermeinten, wie Wolfram oder Walther geschrieben haben und auf dieser Basis großzügig Texte und Autorschaften aberkannt hatten. Zum anderen kommt es uns vielleicht seltsam vor, Autorsignale aus mittelalterlicher Literatur herauszufiltern zu wollen, wo wir doch wissen, dass die Texte, die wir heute vor uns haben, eine lange handschriftliche Überlieferungsgeschichte durchlaufen haben, bei der sich vermutlich nicht wenige Beteiligte, etwa Redaktoren oder Schreiber, in den Text eines Autors eingeschrieben haben.

Lässt man diese Bedenken zunächst beiseite, dann kann man jedoch immerhin konstatieren: Stilometrie funktioniert etwas überraschenderweise auch für mittelhochdeutsche Texte.

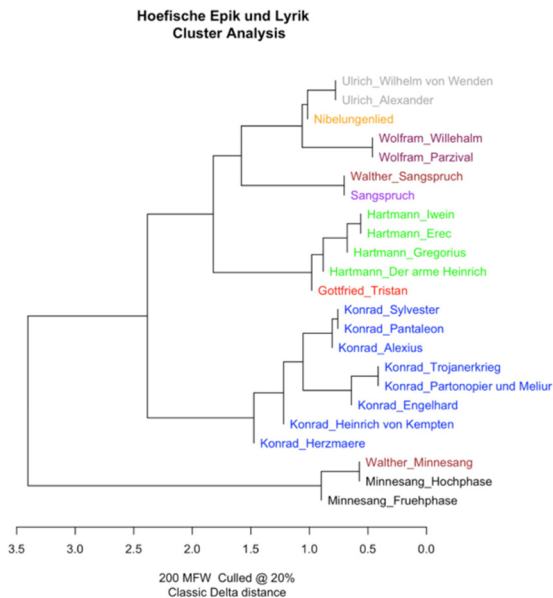


Abbildung 3 zeigt eine Clusterdarstellung klassischer höfischer Texte, die mit stilometrischen Mitteln erstellt wurde.<sup>26</sup> Die Darstellung ist so zu

<sup>26</sup> Zur Erstellung dieser und der nächsten Visualisierung habe ich auf das stylo-Paket (Maciej EDER/Mike KESTEMONT/Jan RYBICKI: Stylometry with R: a Suite of Tools, in: Digital

lesen, dass Texte, die auf den äußersten Zweigen zusammensortiert werden, vom Computer als stilistisch am nächststehenden angesehen werden. Wolframs *Parzival* ist also laut Computer dem *Willehalm* am ähnlichsten, auch die Werke Hartmanns sortieren sich zusammen, ebenso wie die Texte Konrads von Würzburg.<sup>27</sup> Wie es scheint, ist der Computer also durchaus in der Lage, allein durch quantifizierende Auszählungen die Autorkorpora richtig zusammenzuordnen.

Die Frage, ob dieser Befund nun bedeutet, dass es trotz Cerquigli<sup>28</sup> und anderen, die den Autor als nicht dem Mittelalter entsprechende Größe verabschiedet haben, also doch ein sehr starkes Autorsignal in mittelhochdeutscher Literatur gibt, ist aus dieser Darstellung aber doch nicht so einfach zu beantworten. Das liegt vor allem an der großen Uneinheitlichkeit des Mittelhochdeutschen, das ja vor allem in dialektalen und sprachgeschichtlich unterschiedlichen Ausprägungen vorliegt, was die Anwendung digitaler Analyseverfahren, die auf der Auszählung von konstanten Wortformen beruhen, grundsätzlich schwierig macht. So könnte es etwa sein, dass die Zusammensortierung von *Parzival* und *Willehalm* darauf beruht, dass diese Texte beide von Karl Lachmann nach der St. Galler Epenhandschrift herausgegeben wurden; in das Autorsignal könnten hier also sowohl ein Schreiber- als auch ein Herausgebersignal hereingespielt haben. Dem widerspricht aber immerhin der Befund bei Hartmann, bei dem die Texte ja aus unterschiedlichen Handschriften stammen.

---

Humanities 2013: Conference Abstracts, University of Nebraska-Lincoln, NE, S. 487–489) für die Programmiersprache R zurückgegriffen. Die Texte sind aus der Mittelhochdeutschen Begriffsdatenbank (<http://mhdadb.sbg.ac.at/>) bezogen und wurden zur Abstimmung geringfügig angeglichen. Genauere Angaben zur Zusammenstellung der Korpora finden sich in meinem Aufsatz Gabriel VIEHHAUSER: Digitale Gattungsgeschichten. Minnesang zwischen generischer Konstanz und Wende, in: Zeitschrift für digitale Geisteswissenschaften 2017, text/html Format. DOI: 10.17175/2017\_003.

<sup>27</sup> Vgl. hierzu ausführlicher Gabriel VIEHHAUSER: Historische Stilometrie? Methodische Vorschläge für eine Annäherung textanalytischer Zugänge an die mediävistische Textualitätsdebatte, in: Grenzen und Möglichkeiten der Digital Humanities, hrsg. v. Constanze Baum/Thomas Stäcker (= Sonderband der Zeitschrift für digitale Geisteswissenschaften 1), 2015.

<sup>28</sup> Bernard CERQUIGLINI: *Éloge de la variante. Histoire critique de la philologie*. Paris 1989.

Schwerer wiegt vielleicht noch die Tatsache, dass es sich bei der Stilometrie um ein stildifferenzierendes und nicht ein stilbestimmendes Verfahren handelt.<sup>29</sup> Schaut man sich etwa die Wörter genauer an, die für eine Differenzierung zwischen den Wolframtexten auf der einen und den Hartmantexten auf der anderen Seite verantwortlich sind, so fällt zum Beispiel auf, dass eines der distinktivsten Wörter zwischen den beiden Korpora die Präteritumform *kom* bzw. *kam* ist.<sup>30</sup> Beide Formen sind Formen vom Wort *komen*, *kam* ist jedoch die südwestdeutsche Variante, *kom* ist in den anderen Gebieten des deutschsprachigen Raums verbreitet.<sup>31</sup> Dass sich nun der Alemanne Hartmann in dieser Beziehung von Wolfram unterscheidet, wussten schon die Germanisten des 19. Jahrhunderts.<sup>32</sup>

Es wäre also wohl etwas phantasielos, sich allzu sehr auf solche stilometrischen Autorschaftszuweisungen einzulassen. Ein solches Unterfangen würde zudem erst recht darauf führen, sich wieder über kanonische Autorgenies zu unterhalten und mithin einer Stoßrichtung zu folgen, die ja gerade Moretti mit seinem Distant Reading-Konzept vermeiden wollte.

Sucht man nach anderen auswertbaren Befunden aus der Darstellung, so zeigt sich bei genauerer Darstellung ein auffälliges Korpus, bei dem die Autorschaftssortierung nicht funktioniert hat: Die Texte Walthers von der Vogelweide, die ich grob in Sangspruch und Minnesang

---

<sup>29</sup> Vgl. hierzu auch JANNIDIS: Autor ganz nah.

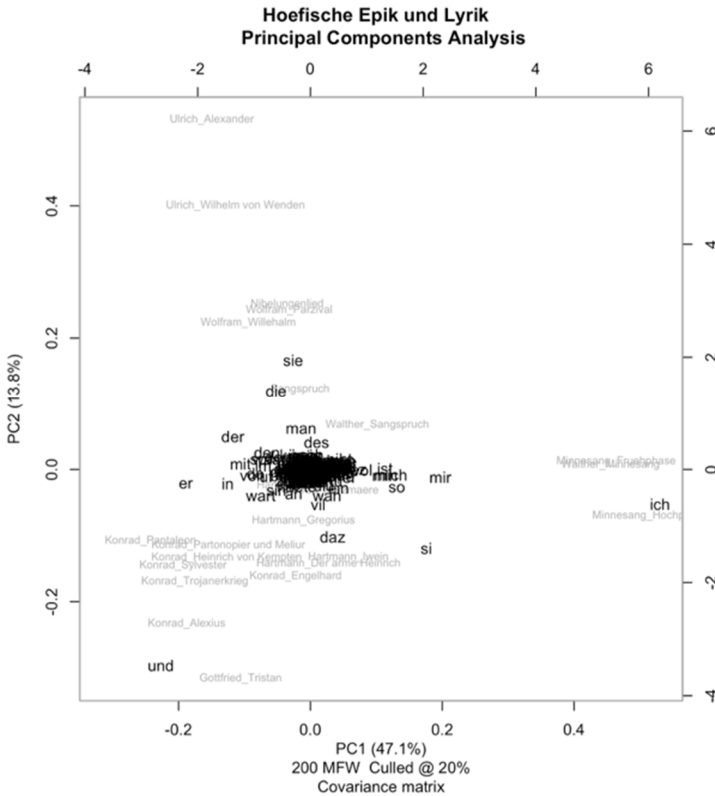
<sup>30</sup> Dies bringt zumindest ein Vergleich zu Tage, der mit einem anderen stilometrischen Verfahren, nämlich Craigs Zeta durchgeführt wurde: Zeta zielt nicht auf die Wortfrequenz, sondern auf die Konstanz in der Wortverwendung ab, vgl. John BURROWS: All the Way Through: Testing for Authorship in Different Frequency Strata, in: Literary and Linguistic Computing 22 (2007), H. 1, S. 27–47; sowie Hugh CRAIG/Arthur KINNEY: Shakespeare, Computers, and the Mystery of Authorship, Cambridge 2009.

<sup>31</sup> Vgl. hierzu Hermann PAUL: Mittelhochdeutsche Grammatik, 25. Aufl., Tübingen 2007, S. 253 (§M 79, Anm. 4); zu Unschärfen in der Verteilung Thomas KLEIN: Ermittlung, Darstellung und Deutung von Verbreitungstypen in der Handschriftenüberlieferung mittelhochdeutscher Epik, in: Deutsche Handschriften 1100–1400. Oxforder Kolloquium 1985, hrsg. v. Volker Honemann/ Nigel F. Palmer, Tübingen 1988, S. 110–167.

<sup>32</sup> Etwa Arno SCHIROKAUER: Studien zur mittelhochdeutschen Reimgrammatik, in: Beiträge zur Geschichte der deutschen Sprache und Literatur (PBB) 47 (1923), S. 1–126, hier: S. 13.

eingeteilt habe,<sup>33</sup> clustern nicht zusammen, sondern werden offensichtlich nach Gattungskriterien sortiert: Walthers Sang gruppiert sich zum Minnesang, sein Sangspruch zu einem Korpus mit anderer Sangspruchdichtung, die ich für diese Analyse zusammengestellt habe.

Offenbar gibt es im Fall des Minnesangs ein starkes Gattungssignal in mittelhochdeutscher Literatur, das sich mit stilometrischen Methoden eruieren lässt. Welches das ist, wird deutlicher, wenn man sich die Unterschiede zwischen den Texten in einer sogenannten ‚Principle Component Analysis‘, kurz PCA, oder Hauptkomponentenanalyse ansieht.



<sup>33</sup> Sangspruch und Minnesang sind die beiden großen Spielarten mittelhochdeutscher Liedlyrik.



Die PCA ist ein recht komplexes statistisches Verfahren, bei der die Varianz des Datenmaterials auf wenige Dimensionen reduziert wird; hier auf die beiden Hauptkomponenten, die auf der x- und der y-Achse dargestellt sind.<sup>34</sup> Abbildung 4 ist also so zu lesen, dass die x-Achse die erste Hauptkomponente abbildet, die für 47,1% der Varianz im Datenmaterial einsteht. Texte, die ganz links stehen, unterscheiden sich von jenen, die ganz rechts stehen (in diesem Fall etwa Walthers und anderer Minnesang), in Bezug auf diese Hauptkomponente. Die y-Achse gibt die zweite Hauptkomponente wieder und unterscheidet die Texte ganz unten von denen ganz oben.

In der PCA der mittelhochdeutschen Texte zeigt sich nun, dass der Gegensatz zwischen Minnesang, und zwar Walthers Sang und dem hohen und frühen Sang, und den restlichen Texten vor allem durch die erste Hauptkomponente erklärt wird, und, auch das kann die PCA zeigen, sich vor allem aus dem Gebrauch des Personalpronomens in der ersten Person, also dem Ich ergibt. Minnesang ist demnach sehr stark stilistisch durch die Ich-Form geprägt, und zwar so stark, dass diese generische Trennung auch das Autorsignal im Walther-Korpus überspielt. Dieser Befund ist aus traditioneller literaturwissenschaftlicher Sicht keineswegs abwegig, denn im Minnesang, das ist bekannt, spielen die Ich-Aussage und das Sänger-Ich eine entscheidende Rolle.<sup>35</sup>

Stilometrische Verfahren können also gewisse Aussagen über den Stil von Texten treffen, und um daraus Erkenntnisse zu ziehen, muss man nicht immer emphatische Auffassungen von Autorstilen in Anschlag bringen, sondern man kann diese Methoden durchaus explorativ nutzen. Natürlich bleiben diese Techniken sehr auf der Oberfläche, Distant Reading erfolgt nun einmal aus der Vogelperspektive, und je weiter von oben man Dinge betrachtet, desto eher gehen Details verloren. Das ist auch gar nichts Ungewöhnliches, sondern stellt eigentlich eine

---

<sup>34</sup> Vgl. zum Verfahren Matthew Lee JOCKERS: *Macroanalysis: Digital Methods and Literary History*, Urbana / Chicago / Springfield 2013.

<sup>35</sup> Vgl. hierzu etwa Klaus GRUBMÜLLER: Ich als Rolle, ‚Subjektivität‘ als höfische Kategorie im Minnesang?, in: *Höfische Literatur – Hofgesellschaft – Höfische Lebensformen um 1200*, hrsg. v. Gert Kaiser / Jan-Dirk Müller, Düsseldorf 1986, S. 387–408.

Grundfigur des wissenschaftlichen, aber auch des Alltags-Erkennens dar: Um etwas erkennen zu können, ist es notwendig, Komplexität zu reduzieren und Details beiseite zu lassen. Schon allein deshalb ist es aber auch gar nicht nötig und wohl auch nicht ratsam, bei der Makroperspektive des Distant Readings stehen zu bleiben. Ein Distant-Reading-Befund wie jener zum Minnesang könnte beispielsweise zum Ausgangspunkt genommen werden, wieder in die Texte sozusagen hineinzuzoomen und genauer hinzusehen, was es z. B. mit der Ich-Haltigkeit des Minnesangs auf sich hat. Der Anglist Martin Mueller hat einen solchen Ansatz, der als Kompromissposition Distant- und Close-Reading-Verfahren verbindet, und dabei wie mit einem Zoomobjektiv sich zwischen Mikro- und Makroperspektive hin und her bewegt, ‚Scalable Reading‘ genannt.<sup>36</sup> Nimmt man einen solchen Ansatz ernst, dann schließt sich der Kreis zurück zum Beginn meiner Argumentation: Wie bei der digitalen Editorik haben wir auch bei der Textanalyse die Möglichkeit, unterschiedliche Perspektiven auf Texte einzunehmen.

Mir scheint es nun entscheidend für die Arbeit in den Digital Humanities zu sein, dass man zum einem diese unterschiedlichen Perspektiven auch wahrnimmt und nebeneinander stellt, und zum anderen, dass man sich immer auch bewusst ist, dass es sich eben um unterschiedliche Perspektiven handelt, unterschiedliche Brillen sozusagen, aus bzw. mit denen die Dinge betrachtet werden. Die Position des Betrachters spielt eine Rolle, das ist die Lehre, die man unter diesen Voraussetzungen gerade aus dem Einsatz des Computers ziehen könnte - und diese Berücksichtigung der Betrachterperspektive ist eigentlich wieder ein ganz klassisch geisteswissenschaftliches Thema.

### III.

Um nun z. B. etwas weiter in die Geschichte des Minnesangs hineinzuzoomen, sollen zum Abschluss noch einige Methoden aus der

---

<sup>36</sup> Martin MUELLER: Shakespeare His Contemporaries: Collaborative Curation and Exploration of Early Modern Drama in a Digital Environment, in: Digital Humanities Quarterly 8 (2014), H. 3.

distributionellen Semantik vorgestellt werden, die sich bislang als am geeignetsten erwiesen haben, über die Textoberfläche hinauszukommen und dem Computer so etwas wie Aussagen über Bedeutung zu entlocken. Die bislang vorgestellten Analysemethoden beruhen im Wesentlichen auf der Auszählung von Einzelwörtern und der sogenannten Bag-of-Words-Methode: Die Wörter werden isoliert und unabhängig von ihrem Kontext betrachtet, der Text also gleichsam in einzelne Wortbestandteile ausgelöst, die in einem Wörter-Sack gesammelt und durcheinandergemischt werden können. An welcher Stelle etwa oder in welcher Umgebung das Personalpronomen ‚ich‘ im Minnesang verwendet wird, ist beispielsweise für das stilometrische Verfahren nicht von Belang, wesentlich ist nur die aggregierte Frequenz der Einzelbelege.

Die Idee der distributionellen Semantik wäre demgegenüber, dass man dem Kontext besondere bzw. sogar entscheidende Relevanz beimisst, da sich die Bedeutung eines Wortes nach dieser Auffassung aus dem Kontext überhaupt erst ergibt. Der englische Linguist John Rupert Firth hat diesen Zusammenhang in der berühmt gewordenen Formel „You shall know a word by the company it keeps“ auf den Punkt gebracht.<sup>37</sup>

Eine gängige Technik der distributionellen Semantik, die in der digitalen Textanalyse häufig zum Einsatz gebracht wird, ist das ‚Topic Modeling‘.<sup>38</sup> Topic Modeling dient dazu, die thematische Struktur in großen Textkorpora darzustellen. Die Idee hinter dem Verfahren ist jene, dass das häufige gemeinsame Auftreten von Wörtern in einem Abschnitt auf eine zugrundeliegende Thematik dieses Abschnitts hinweisen könnte. Wenn also z. B. in einem Text besonders häufig von Schiffen, Netzen und Fischen die Rede ist, liegt der Schluss nahe, dass es hier um ein Thema wie Fischerei geht.

---

<sup>37</sup> John R. FIRTH: A Synopsis of Linguistic Theory 1930-1955, in: Studies in Linguistic Analysis, Special volume of the Philological Society, hrsg. von John R. Firth, Oxford 1957, S. 1–32, hier: S. 11.

<sup>38</sup> Vgl. die Einführung bei Jan HORSTMANN: Topic Modeling, in: forTEXT. Literatur digital erforschen (2018), <https://fortext.net/routinen/methoden/topic-modeling> .

Der Computer kann allerdings nur das auffällige gemeinsame Auftreten von Wörtern erkennen (also z. B. feststellen, dass Schiffe, Angeln und Fische zu einem gemeinsamen Topic gehören) – die Inferenz von diesen Wörtern auf das verbindende Thema bleibt dem menschlichen Benutzer überlassen.

In einem Text können bzw. werden mehrere solcher Topics auftreten und der Computer trifft Voraussagen, zu welchem Prozentsatz diese Topics vorkommen. Zudem wird vorhergesagt, welche Wörter einen gesteigerten Anteil an einem Topic aufweisen.



Ein Topic-Modell des Minnesangs ist in Abbildung 5 dargestellt.<sup>39</sup> Ich habe hier den Computer 15 Topics berechnen lassen, zu denen er die Wortverteilungen zusammensortiert hat.<sup>40</sup> Funktionswörter, die bei den

<sup>39</sup> Eine interaktive Darstellung des Modells ist unter <https://gabvie.github.io/Distant-Reading-Minnesang/TopicModel1/index.html> abrufbar, weiterführend zum Verfahren VIEHHAUSER: Digitale Gattungsgeschichten.

<sup>40</sup> Der Analyse liegt der verbreitete Latent Dirichlet Allocation-Algorithmus (LDA) zugrunde (vgl. David M. BLEI/Andrew Y. NG/Michael I. JORDAN: Latent Dirichlet Allocation, in: Journal of Machine Learning Research 3 (2003), S. 993–1022, der mit dem Java-Tool MALLET prozessiert wurde (Andrew Kachites MCCALLUM: MALLET: A Machine Learning for Language Toolkit, hrsg. v. d. University of Massachusetts, Amherst 2002). Zur Erstellung der Topics wurden 2.000 Iterationen vorgenommen. Zur Datenaufbereitung und zur Erstellung des Modells wurde auf das bei Matthew Jockers dokumentierte Verfahren und den von

stilometrischen Verfahren noch von größter Wichtigkeit waren, habe ich dabei aus der Analyse ausgeschlossen, da sie thematisch zu wenig aussagekräftig sind.<sup>41</sup> Die übrig gebliebenen häufigsten Wörter der Topics sind in der Visualisierung in Word-Clouds dargestellt, bei denen die Häufigkeit der Wörter proportional ihrer Schriftgröße entspricht.<sup>42</sup>

Die einzelnen Topics sind, das ist bei dem Verfahren üblich, unterschiedlich spezifisch ausgeprägt.<sup>43</sup> Ein paar dieser Topics sollen im Folgenden herausgegriffen werden.



Abbildung 6 zeigt ein Topic, das Ausdrücke für Jahreszeiten mit Naturausdrücken zusammenbringt. Es handelt sich hier augenscheinlich um die Motivik des Natureingangs, also die Manier, Minnelieder damit zu beginnen, dass sich der Sänger mit seiner Stimmungslage in Relation zur Jahreszeit setzt.<sup>44</sup>

---

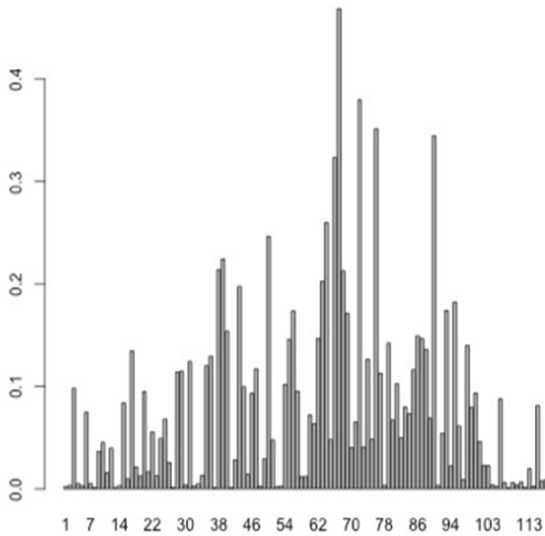
ihm erstellten R-Code zurückgegriffen (Matthew Lee JOCKERS: *Text Analysis with R for Students of Literature*. Cham 2014), welcher an die Fragestellung angepasst wurde. Dementsprechend habe ich das Textkorpus in einzelne Abschnitte gleicher Größe gegliedert, wobei ich aufgrund der Kürze des Korpus Abschnittslängen von 500 Wörtern definiert habe.

<sup>41</sup> Es wurde keine Lemmatisierung und kein POS-Tagging vorgenommen.

<sup>42</sup> Die Word-Clouds wurden mit dem R-Package `wordcloud` erstellt, vgl. Ian FELLOWS: `wordcloud`: Word Clouds. R package version vom 2.5.2014.

<sup>43</sup> Vgl. hierzu Christof SCHÖCH: *Topic Modeling Genre: An Exploration of French Classical and Enlightenment Drama*, in: *Digital Humanities Quarterly* 11.2 (2017).

<sup>44</sup> Vgl. Ernst Robert CURTIUS: *Rhetorische Naturschilderung im Mittelalter*, in: *Romanische Forschungen* 56 (1942), S. 219–256; Daniel EDER: *Der Natureingang im Minnesang. Studien zur Register- und Kulturpoetik der höfischen Liebeskanzone*, Tübingen 2017.



Bei diesem Topic lohnt es sich, auf die zeitliche Verteilung im Minnesang zu sehen, die in Abbildung 7 auf der X-Achse als Verlaufskurve angeordnet ist: Zu Beginn der x-Achse sind die Autorkorpora aus der Frühzeit des Minnesangs sortiert, dann folgt die Hoch- und die Spätphase, am Schluss steht ohne zeitliche Ordnung der Sangspruch. Jede Säule steht also für das Oeuvre eines Sängers.

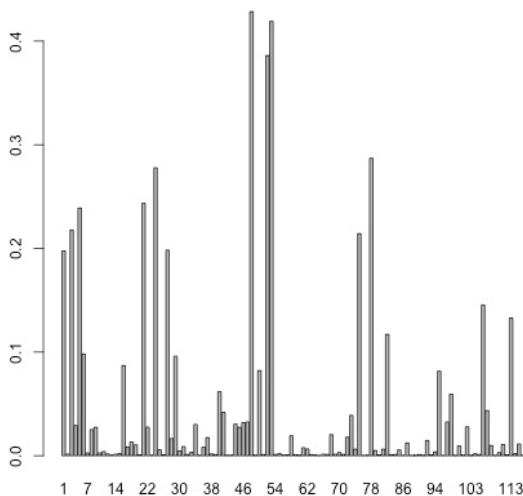
Die Abbildung zeigt nun, dass der Natureingang eine deutliche Konjunktur in der Spätphase aufweist. Auch dieser Befund steht durchaus im Einklang zur traditionellen Forschung, die den Natureingang fallweise sogar geradezu als Gattungskennzeichen des späten Sangs bezeichnet hat.<sup>45</sup>

---

<sup>45</sup> Rüdiger SCHNELL: Minnesang und Sangspruch im 13. Jahrhundert. Gattungsdifferenzen und Gattungsinterferenzen, in: Transformationen der Lyrik im 13. Jahrhundert, hrsg. v. Susanne Köbele in Verb. mit Eckart Conrad Lutz/Klaus Ridder (Wildbader Kolloquium 21, Rothenburg o. d. T., 30.09.–03.10.2008)(= Wolfram-Studien 21), Berlin 2013, S. 287–347, hier: S. 326.



Ein weiteres leicht zu deutendes Topic zeigt Abbildung 8, die wichtigsten Wörter sind *scheiden*, *ritter*, *tac* und *klage*. Diese Wörter konstituieren die bekannte Tageliedthematik, eine sehr auffällige Variation des Minnethemas, da hier für einmal nicht das Sänger-Ich im Mittelpunkt steht, sondern von einem Ritter und einer Dame erzählt wird, die sich nach einer gemeinsamen Liebesnacht am Morgen trennen müssen.



Im zeitlichen Verlauf in Abbildung 9 sind hier vereinzelte Spitzen zu sehen, die die *Œuvres* typischer Tagelieddichter repräsentieren. Laut dem Topic Model ist der tageliedhafteste Sänger des Minnesangs etwa der

Burggraf von Lienz, mit einem Anteil von 42,84 Prozent des Topics am Gesamtoeuvre. Danach folgen Günther von dem Forste und der von Wissenlo. Einer der bekanntesten Tagelieddichter, Wolfram von Eschenbach, wird immerhin noch mit 24,35 Prozent ausgewiesen. Doch was bedeuten diese Prozentangaben überhaupt? Natürlich nicht, dass genau 24,35 von 100 Wörtern bei Wolfram Tagelied-Wörter sind. Vielmehr zeigen die Zahlen eine Tendenz an, letztlich Wahrscheinlichkeiten. Das ist in mehrerer Hinsicht praktisch und vielleicht auch durchaus geisteswissenschaftlich: Denn trotz des Klischees fällt der Computer hier gerade nicht eine binäre Entscheidung, ob ein Text Tagelied ist oder nicht, sondern er gibt eben nur einen gewissen Grad an Wahrscheinlichkeit hierfür an. Gerade der Computer bleibt hier also unscharf, was aber wie gesagt nichts Schlechtes ist, denn diese Unschärfe lässt sich nutzen, Dinge vielleicht angemessener mit der Formel ‚mehr oder weniger‘ zu beschreiben als mit ‚genau so und nicht anders‘. Ein gutes Beispiel dafür wären etwa die hohen Werte an Tagelied-Motivik im frühen, donauländischen Minnesang (also innerhalb der ersten sechs Korpora auf der x-Achse): Bekanntlich gibt es im frühen Minnesang (bis Dietmar von Eist zumindest) keine Tagelieder, doch hat der frühe Minnesang etwas Erzählerisches an sich, das vielleicht durchaus Verwandtschaft zum Tageliedmotiv zeigt. Und auf diese Verwandtschaft führt die Auswertung des Computers, der gerade keine binäre Entscheidung trifft, sondern nur Tendenzen anzeigt. Auf der Grundlage dieser Tendenzen entscheiden und Kategorien einziehen hingegen bleibt immer noch dem menschlichen Nutzer bzw. der menschlichen Nutzerin überlassen.<sup>46</sup>

---

<sup>46</sup> Dabei bleibt offen, wie weit so eine Entscheidung getroffen werden muss oder soll: Statt ein übergeordnetes Bild zusammenzufügen und damit etwa ein Verlaufsnarrativ zu erstellen, ist es durchaus auch eine Option, bei punktuellen Befunden stehenzubleiben (oder, ähnlich dem Scalable-Reading-Ansatz zwischen Narrativ und punktuelltem Befund hin und herzuzoomen). Die Konstruktion von ‚Meistererzählungen‘ ist grundsätzlich mit erkenntnistheoretischen Problemen behaftet, da letztlich immer interpretativ. Dies gilt gerade für die mittelhochdeutsche Literatur und den Minnesang, bei dem eine zeitliche literaturgeschichtliche Ordnung immer nur ungefähr vorgenommen werden kann, da die genauen Lebensdaten der meisten Dichter fehlen bzw. keine Informationen über die Abfassungszeit der Texte vorhanden sind. So gesehen ist schon die hier vorgenommene Gliederung des



## IV.

Wie die hier angeführten Beispiele zeigen, dürfte sich mit Hilfe von Topic Modeling doch einiges an semantischem Potential des Minnesangs fassen lassen. Dass die Methode trotz vergleichsweise geringem Datenmaterials hier so gut anwendbar ist, mag auch mit der starken Konventionalität zusammenhängen, die dem Minnesang inhärent ist: In der Forschung wurde immer wieder betont, dass die Themen des Minnesangs sehr limitiert sind, es im Sinne einer Ästhetik der Identität mehr um die Variation von Konstellationen geht, als diese neu zu erfinden. Dem entspricht, dass der Minnesang auf der Textoberfläche sehr stark durch Leitbegriffe wie etwa *minne*, *leit*, *klage* usw. geprägt ist.<sup>47</sup>

Dieser Befund kann einmal mehr auch quantitativ bestätigt werden: Schließt man die Funktionswörter aus und nimmt man eine simple Frequenzauszählung vor, dann werden die häufigsten Wörter im Minnesang *minne* und *wîp*, aber auch *vroude*, *herze* sowie *güete* und schließlich auch negative Aspekte wie *not*, *leit* und *sorge* sichtbar.

Vor diesem Hintergrund erscheint die Frage interessant, wie sich die Geschichte des Minnesangs anhand dieser Leitbegriffe sortieren lässt. Eine Möglichkeit, dies darzustellen, wär ein sogenannter Wörterbuch-basierter Ansatz. Dieses Verfahren haben z.B. Nils Reiter und Marcus Willand in ihrem QuaDrama-Projekt genutzt, das sich mit der quantitativen Auswertung von Dramentexten befasst.<sup>48</sup> Reiter und Willand haben dabei den semantischen Gehalt von Figurenreden nach Wortfeldern aufgeschlüsselt, um so typische Redeinhalte bestimmter Charaktere aufzuzeigen.

---

Minnesangs in Früh-, Hoch- und Spätphase und die Zuordnung der einzelnen Autorœuvres eine Konstruktion, die auch anders vorgenommen werden kann. Auch hier gilt, dass Grundvoraussetzung für die Erkenntnis eine Komplexitätsreduktion ist, derer man sich als Forscher oder als Forschende bewusst sein sollte.

<sup>47</sup> Vgl. hierzu etwa Manuel BRAUN: Typus und Variation im Minnesang des 13. Jahrhunderts, in: Ulrich von Liechtenstein. Leben – Zeit – Werk – Forschung, hrsg. v. Christopher Young/Sandra Linden, Berlin/New York 2010, S. 398–441, hier: S. 417.

<sup>48</sup> Nils REITER/Marcus WILLAND: Ottokar Capulet und Julia Schroffenstein – Liebeskonflikte und Gattungskontexte bei Kleist und Shakespeare, <https://quadrama.github.io/blog/2016/10/07/ottokar-capulet>.

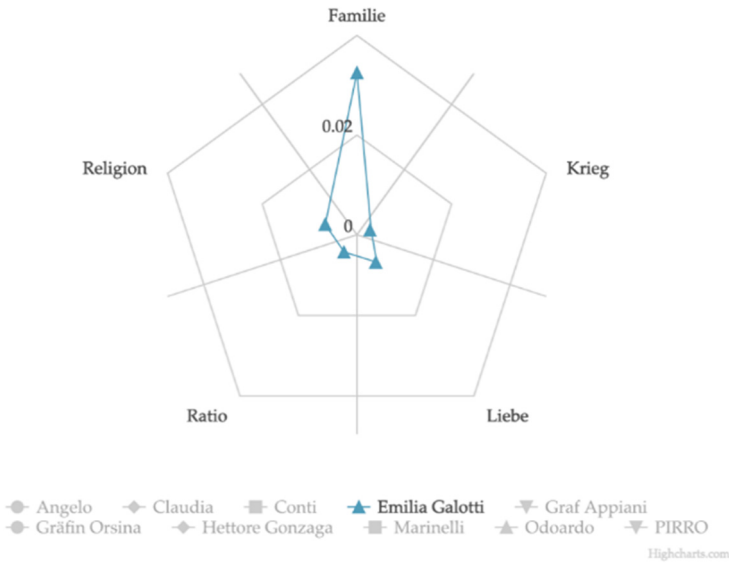


Abbildung 10 zeigt einen sogenannten Spider-Plot der Figurenrede von Emilia Galotti, bei der fünf Dimensionen herausgegriffen wurden, Familie, Krieg, Liebe, Ratio und Religion, die jeweils konzentrisch in der Visualisierung in unterschiedliche Richtungen aufgetragen sind. Mit dieser Methode lässt sich nun etwa zeigen, dass Emilias Figurenrede einen großen Anteil an Wörtern hat, die aus dem Wortfeld ‚Familie‘ stammen, während Wörter aus anderen Themenfeldern seltener sind.

Für meinen Transfer dieser Methode habe ich mir die häufigsten Wörter des Minnesangs angesehen und fünf Themenfelder festgelegt, die sich in diesen Wörtern abzeichnen:

- Feld 1 ist aus den positiven Aspekten der Liebe gebildet, mit Ausdrücken wie *genâde*, *güete*, *gruoz*, *hult*, *lôn*, *saelde*.
- Feld 2 beinhaltet negative Aspekte wie *kumber*, *leit*, *leide*, *nôt*, *sorge*.
- Feld 3 sind Naturausdrücke, die die weite Verbreitung des Natureingangs reflektieren, also z.B Wörter wie *bluomen*, *heide*, *winter*, *vogel*, *meien*, *walt*.
- Feld 4 betrifft Ausdrücke der Wahrnehmung, *ougen*, *sinne*, *mun*t, *schîn*, *sach*.

Feld 5 enthält Ausdrücke, die man zur Beschreibung der besonderen Qualität der Dame einsetzen könnte, *guot, schoen, reine, süeze*.<sup>49</sup>

Die Erstellung der Wortfelder habe ich in einer Kombination von automatischen und manuellen Verfahren vorgenommen. Zunächst wurden die häufigsten Wörter im Minnesangkorpus (abzüglich der Stoppwörter) manuell den fünf Kategorien zugewiesen. Um das Wortmaterial zu vermehren, habe ich die Wortfelder mit einer weiteren Methode der distributionellen Semantik erweitert, nämlich mit sogenannten ‚Word Embeddings‘.<sup>50</sup> Word Embeddings berechnen – verkürzt gesagt - auf einem großen Korpus (mit Hilfe eines einfachen neuronalen Netzes) die Auftrittswahrscheinlichkeiten, dass ein Wort im Kontext eines anderen vorkommt.<sup>51</sup> Dadurch erhält man für jedes Wort einen Vektor, also eine Zahlenreihe, die diese Wahrscheinlichkeiten enthält. Und diese Vektoren lassen sich dann wiederum mit Distanzmaßen auf ihre Ähnlichkeit hin überprüfen.

Da man Vektoren beliebig aufaddieren und auch wieder abziehen kann, kann man mit diesen Vektoren sehr gut rechnen und z. B. nicht nur die ähnlichsten Wörter eines Wortes im Korpus berechnen, sondern auch die ähnlichsten Wörter zu einem ganzen Wortfeld. Nimmt man etwa die Vektoren von *kumber, leit, leide, nôt, sorge* und *sorgen* zusammen, dann lassen sich mit Distanzmaßen eine ganze Reihe von Wörtern eruieren, die die Maschine für ähnlich hält.

---

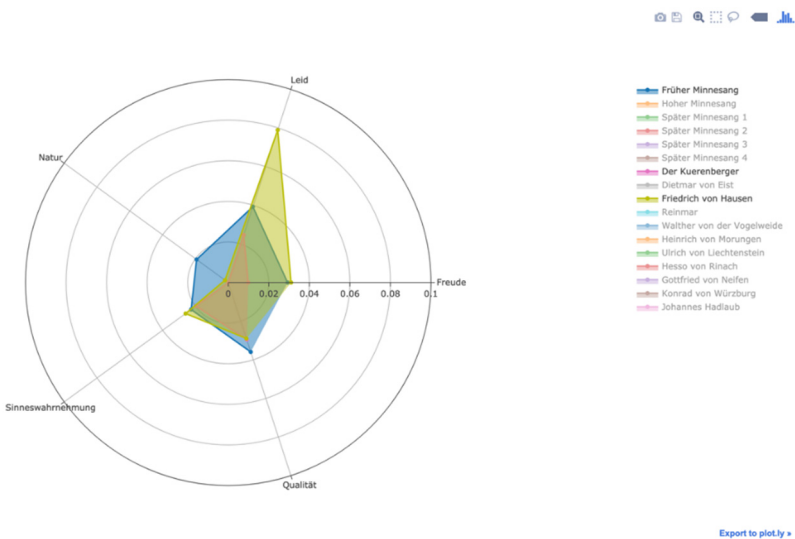
<sup>49</sup> Natürlich sind einige dieser Ausdrücke ambig und ließen sich mehreren oder anderen Wortfeldern zuordnen. Auch der Wörterbuch-basierte Ansatz bietet keine endgültigen Antworten, sondern einen explorativen Ausgangspunkt, der mit qualitativer Interpretation ergänzt werden muss.

<sup>50</sup> Die Erstellung des Modells erfolgte mit Hilfe der Daten der mittelhochdeutschen Begriffsdatenbank (<http://mhdadb.sbg.ac.at/>) und der gensim-Library für Python (Radim ŘEHŮREK/Petr SOJKA, Software Framework for Topic Modelling with Large Corpora, in: Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks, 2010).

<sup>51</sup> Vgl. zur Einführung <https://machinelearningmastery.com/what-are-word-embeddings/>.

```
In [13]: model.wv.most_similar(positive=["kummer", "leit", "leide", "nöt", "sorge", "sorgen"], topn=20)
Out[13]: [('swære', 0.7795711360397339),
('klage', 0.7424070835113525),
('pin', 0.7372208833694458),
('jâmer', 0.7343716621398926),
('swære', 0.7320749759674072),
('smerzen', 0.7109144926071167),
('kummer', 0.701533317565918),
('pine', 0.6822977662086487),
('arebeit', 0.680925726890564),
('quäle', 0.6732896566390991),
('trüren', 0.6611348986625671),
('smerze', 0.6554744839668274),
('ungemach', 0.654428892881775),
('ungemiete', 0.6485596895217896),
('vröide', 0.6451141834259033),
('fröude', 0.6424039602279663),
('riuwe', 0.6320071220397949),
('vröude', 0.6300934553146362),
('ungenâde', 0.6296942234039307),
('noete', 0.6225767135620117)]
```

Abbildung 11 zeigt, dass diese Ergebnisse auch für die menschliche Einschätzung wohl durchaus nachvollziehbar erscheinen.



Zur Visualisierung des Auftretens dieser Wortfelder in den Minnesangtexten habe ich einen interaktiven Spider-Plot erstellt<sup>52</sup> und die

<sup>52</sup> Die Visualisierung wurde mit Hilfe des Plot.ly-Python-Packages, <https://plot.ly/>, erstellt. Eine interaktive Version ist unter [https://gabvie.github.io/Distant-Reading-Minnesang/Wortfeld/temp-plot\\_word2vec\\_largescale.html](https://gabvie.github.io/Distant-Reading-Minnesang/Wortfeld/temp-plot_word2vec_largescale.html) abrufbar.

Frequenzen der Begriffe aus den Wortfeldern zunächst für die sechs zeitlichen Phasen des Minnesangs ausgezählt sowie für einzelne, exemplarisch ausgewählte Autorkorpora.

Betrachtet man mit Hilfe dieses Verfahrens etwa den Minnesang aus der Makroperspektive, dann zeigt sich, dass im Spider Plot für den Frühen Minnesang die Dimensionen der Wortfelder recht ausbalanciert sind. Im hohen Minnesang nehmen die Ausdrücke für Freude und Leid zu, was zum Bild passt, dass im hohen Minnesang die romanische Vorstellung von unerfüllbarer, leidvoller höfischer Liebe immer wichtiger wird. Der späte Minnesang, den ich in vier zeitliche Phasen unterteilt habe, bringt zunächst ein ähnliches Bild wie der Hohe Minnesang, lediglich in der dritten Phase des späten Sanges gibt es einen auffälligen Anstieg bei den Naturexpressionen, der offensichtlich auf die Blütezeit des Natureingangs hinweist.

Zoomt man im Sinne des Scalable Readings etwas näher an die einzelnen Autorkorpora heran, so zeigt sich, dass etwa in der Grafik des Kürenbergers, der im gängigen Verlaufsnarrativ der Minnesangforschung als der früheste Minnesänger gilt, die Dimensionen (noch) nicht stark ausgeprägt sind. Dieser Befund wird besonders im Vergleich zum Gesamtkorpus des frühen Minnesangs deutlich, er könnte darauf hindeuten, dass das Œuvre des Kürenbergers eben noch nicht so stark von Vorstellungen höfischer Liebe bestimmt ist.<sup>53</sup>

---

<sup>53</sup> Für weitere mögliche Auswertungen verweise ich auf die interaktive Grafik: Interessant ist etwa Friedrich von Hausen, der eine hohe Frequenz von Ausdrücken für Leid aufweist. Das Bild würde man auch für Reinmar erwarten, der geradezu als Meister der Minneklage gilt. Sein Spider-Plot zeigt zwar etwas weniger Klage-Wörter als bei Hausen, die Dimension ist aber immer noch stark ausgeprägt, zudem gibt es auch mehr Freude-Ausdrücke. Walther, der oft als Gegenspieler von Reinmar angesehen wird, weil bei ihm die Vorstellung der gegenseitigen Liebe eine größere Rolle spielt, hat tatsächlich weniger Leid-Ausdrücke als Reinmar und die Qualität der Dame scheint bei ihm eine größere Rolle zu spielen. Als letztes bekanntes Beispiel sei noch Ulrich von Liechtenstein erwähnt, ein Dichter aus der Spätphase. Freudeausdrücke sind bei ihm ganz dominant und häufiger als Leidausdrücke. Auch dies könnte man konform zu traditionellen Forschungsmeinungen sehen, die davon ausgehen, dass die Dichter des späten Sanges wie Ulrich den Freude-Aspekt der Liebe stärker betonen, da für sie die Minne-Konzeption unproblematisch wird.

Bei der Auswertung des Spider-Plots habe ich bislang vor allem damit argumentiert, dass sich die Ergebnisse der quantitativen Analyse im Einklang mit der traditionellen Forschung befinden. Dabei offenbart sich ein altes Dilemma der Digital Humanities, das auch oft gegen sie angewendet wird: Eine Validierung der Distant-Reading-Methode, die ja qua Definition für den einzelnen unüberschaubare Textmengen durchforsten soll, ist schwierig, da man geneigt ist, quantitativen Befunden nur insofern zu trauen, als sie das Althergebrachte bestätigen; Abweichungen sind hingegen als mögliche Fehler des quantitativen Verfahrens verdächtig. Eine Möglichkeit, diesen unbefriedigenden Alternativen zu entgehen, wäre allerdings der genauere Blick ins Detail. So lohnt es sich durchaus, auch weniger bekannte Dichter anzusehen, etwa Hesso von Rinach, den ich hier zufällig herausgreife: Hesso hat eine stark ausgeprägte Leid-Dimension, was eigentlich der konventionellen ‚Meistererzählung‘ der Minnesanggeschichte widerspricht, dass späte Dichter weniger auf die leidvollen als auf die freudvollen Aspekte der Minne abzielen. Um aber zu beurteilen, ob und warum das Verlaufsnarrativ des Minnsangs an dieser Stelle nicht greift, müsste man nun wieder auf die Textebene zurückkehren und sich die Texte von Hesso genauer ansehen, mithin also die quantitativen Makro- und Mesoperspektiven durch die qualitative Mikroperspektive ergänzen.

## V.

Damit möchte ich meine Beispielreihe abschließen und ein Fazit ziehen: Ich bin davon ausgegangen, dass die Art, wie man Wissenschaft betreibt und zu welchen Ergebnissen man dabei kommt, nicht unabhängig von den materiellen und medialen Bedingungen ist und habe diesen Zusammenhang zunächst mit dem Lachmann'schen Reise-Parzival zu belegen versucht. Das digitale Medium bringt also schon vor diesem Hintergrund auf alle Fälle etwas Neues mit sich, auch etwas qualitativ Neues.

Die Frage wäre jetzt noch, ob dieses Neue tatsächlich auch wissenschaftlich ertragreich ist oder nicht. Ich habe zu zeigen versucht, dass einer der Mehrwerte digitaler Methoden überraschenderweise darin liegen könnte, dass es mit ihnen möglich wird, einen multiperspektivischen

Blick auf die Dinge zu eröffnen, und dass gerade dadurch das Bewusstsein dafür gefördert werden kann, dass man Dinge notwendigerweise immer aus einer gewissen Perspektive sieht. Wenn wir etwas sehen wollen, dann müssen wir das aus einem bestimmten Blickwinkel tun, wir müssen Komplexität reduzieren, aber das ist nichts Schlechtes, solange wir uns dessen bewusst sind. Wir konstruieren in der Wissenschaft (und auch der Alltagserkenntnis) Modelle der Dinge, und Modelle sind nicht identisch mit der Wirklichkeit, sondern sie sind abstrahierende Abbildungen der Wirklichkeit, die bestimmten Zwecken dienen. Wie die rezente Theoriediskussion in den Digital Humanities beweist, erscheint die Reflexion dieses Umstandes angesichts von computerunterstützten Methoden besonders naheliegend.<sup>54</sup> Und gerade mit diesem Thema der Perspektivität und der Beobachterabhängigkeit des Standpunkts berühren die Digital Humanities eines der ureigensten Themen der Geisteswissenschaften.

Das Ermöglichen von Multiperspektivität lässt sich vor allem bei digitalen Editionen gut beobachten. Aber auch auf dem Feld der digitalen Analyse ließe sich das Prinzip der Multiperspektivität weiter verfolgen: Scalable Reading besteht im Hin- und Herzoomen zwischen Makro- und Mikroperspektive, bis zu einem gewissen Grad auch zwischen quantitativer Analyse und qualitativer Interpretation. Wie bei der digitalen Editorik lassen sich auch bei der Textanalyse also nicht nur unterschiedliche Betrachtungsweisen auf den Text nebeneinanderstellen, sondern auch unterschiedliche methodisch-konzeptionelle Zugänge vereinen. Von besonderer Wichtigkeit wäre dabei aber jedenfalls, dass man beide Perspektiven auch berücksichtigt und sich nicht von der Evidenz von Grafiken und Zahlen täuschen lässt. „Numbers are a lens“, so hat Andrew Piper diesen

---

<sup>54</sup> Der Modellcharakter von Erkenntnis ist in der Theoriediskussion der Digital Humanities in letzter Zeit intensiv diskutiert worden, vgl. grundlegend Willard McCARTY: Modeling: A Study in Words and Meanings, in: A Companion to Digital Humanities, hrsg. v. Susan Schreibman/Ray Siemens/John Unsworth, Oxford, 2004, <http://www.digitalhumanities.org/companion/>; sowie The Shape of Data in Digital Humanities. Modeling Texts and Text-based Resources, hrsg. v. Julia Flanders/Fotis Jannidis, London 2018.

Sachverhalt formuliert.<sup>55</sup> Ich erweitere: Zahlen sind auch nichts anderes als eine Perspektive, die uns nicht dazu verleiten sollte, mehr Sicherheit herzustellen, sondern im Gegenteil, die Unsicherheiten herauszustellen, mit denen zum Beispiel Aussagen über größere literaturgeschichtliche Verläufe oder über Überlieferungsgeschichtliche Textrekonstruktionen getroffen werden können.

Mit dem Ausstellen von Unsicherheiten bin ich aber nun wieder bei Karl Stackmann angelangt, dessen Studie über ‚mittelalterliche Texte als Aufgabe‘ für meinen Aufsatztitel Pate stand: Meine Beispiele zielten darauf ab zu zeigen, dass digitale Geisteswissenschaften, die den Modellierungsgedanken ernst nehmen, von der Vorstellung nur profitieren können, Texte als Aufgabe anzusehen. Auch bei der Anfertigung von digitalen Modellen geht es meines Erachtens nicht darum, absolute Sicherheit herzustellen und Textzustände oder Interpretationen auf statische Modelle zu fixieren, sondern letztlich darum, ein höchstes Maß an Unsicherheit zu erzeugen. Wie schon Stackmann möchte ich damit keinem völligen Relativismus das Wort reden: Unsicherheit ausstellen heißt nicht, dass man gar nichts Genaues sagen kann, aber möglicherweise ist es für so manchen geisteswissenschaftlichen Gegenstand erhellender zu zeigen, wo digitale Modelle ihre Grenzen haben, als diese Modelle absolut zu setzen.

---

<sup>55</sup> Andrew PIPER: *Enumerations*, Chicago/London 2018, S. 39.



## Bibliographische Hinweise

Alle Internetlinks wurden am 05.06.2020 überprüft.

### Primärtexte

BUMKE, Joachim (Hrsg.): Die ‚Nibelungenklage‘. Synoptische Ausgabe aller vier Fassungen, Berlin / New York 1999.

WOLFRAM VON ESCHENBACH: *Parcival*. Ein Ritter-Gedicht aus dem dreizehnten Jahrhundert von Wolfram von Eschilbach. Zum zweiten Male aus der Handschrift abgedruckt, weil der erste Anno 1477 gemachte Abdruck so selten wie Manuscript ist, in: *Samlung deutscher Gedichte aus dem XII., XIII. und XIV. Iarhundert* [hrsg. v. Christoph Heinrich Müller (Myller)], Berlin 1784.

### Forschungsliteratur

BEIN, Thomas: *Textkritik. Eine Einführung in Grundlagen germanistisch-mediävistischer Editionswissenschaft. Lehrbuch mit Übungsteil*. 2., überarb. u. erw. Auflage, Frankfurt a.M. u.a. 2011.

BLEI, David M./NG, Andrew Y./JORDAN, Michael I.: *Latent Dirichlet Allocation*, in: *Journal of Machine Learning Research* 3 (2003), S. 993–1022.

BRAUN, Manuel: *Typus und Variation im Minnesang des 13. Jahrhunderts*, in: *Ulrich von Liechtenstein. Leben – Zeit – Werk – Forschung*, hrsg. v. Christopher Young/Sandra Linden, Berlin / New York 2010, S. 398–441.

BUMKE, Joachim: Die vier Fassungen der ‚Nibelungenklage‘. Untersuchungen zur Überlieferungsgeschichte und Textkritik der höfischen Epik im 13. Jahrhundert, Berlin / New York 1996.

BUMKE, Joachim: *Epenhandschriften. Vorüberlegungen und Informationen zur Überlieferungsgeschichte der höfischen Epik im 12. und 13. Jahrhundert*, in: *Philologie als Kulturwissenschaft. Studien zur Literatur und Geschichte des Mittelalters. Festschrift für Karl Stackmann zum 65. Geburtstag*, hrsg. v. Ludger Grenzmann, Göttingen 1987, S. 45–59.

BURROWS, John: *All the Way Through: Testing for Authorship in Different Frequency Strata*, in: *Literary and Linguistic Computing* 22 (2007), H. 1, S. 27–47.

CERQUIGLINI, Bernard: *Éloge de la variante. Histoire critique de la philologie*. Paris 1989.

CRAIG, Hugh/KINNEY, Arthur: *Shakespeare, Computers, and the Mystery of Authorship*, Cambridge 2009.

CURTIVS, Ernst Robert: *Rhetorische Naturschilderung im Mittelalter*, in: *Romanische Forschungen* 56 (1942), S. 219–256.

EDER, Daniel: *Der Natureingang im Minnesang. Studien zur Register- und Kulturpoetik der höfischen Liebeskanzone*, Tübingen 2017.

EDER, Maciej/KESTEMONT, Mike /RYBICKI, Jan: *Stylometry with R: a Suite of Tools*, in: *Digital Humanities 2013: Conference Abstracts*, University of Nebraska-Lincoln, NE, S. 487–489.

FELLOWS, Ian: *Wordcloud: Word Clouds*. R package version 2.5, 2014.

- FIRTH, John R.: A Synopsis of Linguistic Theory 1930-1955, in: *Studies in Linguistic Analysis*. Special volume of the Philological Society, hrsg. von John R. Firth, Oxford 1957, S. 1–32.
- GOLD, Matthew K.: *Debates in the Digital Humanities*, Minneapolis/London 2012.
- GIESECKE, Michael: *Von den Mythen der Buchkultur zu den Visionen der Informationsgesellschaft. Trendforschungen zur kulturellen Medienökologie*, Frankfurt a. M. 2002.
- GRUBMÜLLER, Klaus: Ich als Rolle, ‚Subjektivität‘ als höfische Kategorie im Minnesang?, in: *Höfische Literatur – Hofgesellschaft – Höfische Lebensformen um 1200*, hrsg. v. Gert Kaiser/Jan-Dirk Müller, Düsseldorf 1986, S. 387–408.
- GUMBRECHT, Hans Ulrich: *The Powers of Philology. Dynamics of Textual Scholarship*. Urbana/Chicago 2003.
- HAUSTEIN, Jens: 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*, hrsg. v. Ulrich Breuer/Bernhard Spies, Bielefeld 2011, S. 43–60.
- HORSTMANN, Jan: Stilometrie, in: *forTEXT. Literatur digital erforschen*, 2018, <https://fortext.net/routinen/methoden/stilometrie>.
- HORSTMANN, Jan: Topic Modeling, in: *forTEXT. Literatur digital erforschen*, 2018, <https://fortext.net/routinen/methoden/topic-modeling>.
- JANNIDIS, Fotis: Der Autor ganz nah – Autorstil in Stilistik und Stilometrie, in: *Theorien und Praktiken der Autorschaft*, hrsg. v. Matthias Schaffrck/Marcus Willand, Berlin 2014, S. 169–195.
- JOCKERS, Matthew Lee: *Macroanalysis: Digital Methods and Literary History*, Urbana/Chicago/Springfield 2013
- JOCKERS, Matthew Lee: *Text Analysis with R for Students of Literature*. Cham 2014.
- KLEIN, Thomas: Ermittlung, Darstellung und Deutung von Verbreitungstypen in der Handschriftenüberlieferung mittelhochdeutscher Epik, in: *Deutsche Handschriften 1100–1400. Oxforder Kolloquium 1985*, hrsg. v. Volker Honemann/Nigel F. Palmer, Tübingen 1988, S. 110–167.
- MARQUARD, Odo: Über die Unvermeidlichkeit der Geisteswissenschaften, in: *Apologie des Zufälligen: Philosophische Studien*, hrsg. v. Odo Marquard, Stuttgart, 1986, S. 98–116.
- MCCARTY, Willard: Modeling: A Study in Words and Meanings, in: *A Companion to Digital Humanities*, hrsg. v. Susan Schreibman/Ray Siemens/John Unsworth, Oxford, 2004, <http://www.digitalhumanities.org/companion/>.
- MCCULLOH, Marc R.: Myller’s Percival and Lachmann’s Critical Method: the ‚Wolfram-Reise‘ Revisited, in: *MLN* 98 (1983), S. 484–491.
- MCCALLUM, Andrew Kachites: *MALET: A Machine Learning for Language Toolkit*, hrsg. v. d. University of Massachusetts, Amherst 2002.
- MERTENS, Volker: Die Wiederentdeckung Wolframs und die Anfänge der Forschung, in: *Wolfram von Eschenbach. Ein Handbuch*, hrsg. v. Joachim Heinzle, Berlin/Boston 2011, S. 706–741.
- MORETTI, Franco: Conjectures on World Literature, in: *New Left Review* 1 (2000), S. 54–68.

- MUELLER, Martin: Shakespeare His Contemporaries: Collaborative Curation and Exploration of Early Modern Drama in a Digital Environment, in: Digital Humanities Quarterly 8 (2014), H. 3.
- NEUMANN, Friedrich: Karl Lachmanns Wolframreise, in: Jahrbuch der Albertus-Universität zu Königsberg / Pr. 2, Freiburg / Frankfurt a.M. 1952, S. 138–158.
- PAUL, Hermann: Mittelhochdeutsche Grammatik, 25. Aufl., Tübingen 2007.
- PIPER, Andrew: Enumerations, Chicago / London 2018.
- ŘEHŮŘEK, Radim / SOJKA, Petr: Software Framework for Topic Modelling with Large Corpora, in: Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks, 2010.
- REITER, Nils / WILLAND, Marcus: Ottokar Capulet und Julia Schroffenstein – Liebeskonflikte und Gattungskontexte bei Kleist und Shakespeare, Vortrag am 07.10.2016 im Kleist-Museum, Frankfurt (Oder),  
<https://quadrama.github.io/blog/2016/10/07/ottokar-capulet>.
- SAHLE, Patrick: Digitale Editionsformen. Zum Umgang mit der Überlieferung unter den Bedingungen des Medienwandels (3 Bde.), Norderstedt 2013.
- SAHLE, Patrick: Zwischen Mediengebundenheit und Transmedialisierung. Anmerkungen zum Verhältnis von Edition und Medien, in: editio 24 (2010), S. 23–36.
- SCHIROK, Bernd: Der Aufbau von Wolframs *Parzival*, Freiburg 1972.
- SCHIROK, Bernd: Einführung zum Text der Lachmannschen Ausgabe, in: Parzival. Studienausgabe. Mhd. Text nach der sechsten Ausgabe von Karl Lachmann. Übersetzung v. Peter Knecht. Mit Einführungen zum Text der Lachmannschen Ausgabe und in Probleme der Parzival-Interpretation v. Bernd Schirok, 2. Aufl., Berlin / New York 2003.
- SCHIROKAUER, Arno: Studien zur mittelhochdeutschen Reimgrammatik, in: Beiträge zur Geschichte der deutschen Sprache und Literatur (PBB) 47 (1923), S. 1–126.
- SCHNELL, Rüdiger: Minnesang und Sangspruch im 13. Jahrhundert. Gattungsdifferenzen und Gattungsinterferenzen, in: Transformationen der Lyrik im 13. Jahrhundert, hrsg. v. Susanne Köbele in Verb. mit Eckart Conrad Lutz / Klaus Ridder (Wildbader Kolloquium: 21, Rothenburg o.d.T., 30.09.–03.10.2008) Berlin 2013 (= Wolfram-Studien 21), S. 287–347.
- SCHÖCH, Christof: Quantitative Textanalyse, in: Digital Humanities. Eine Einführung, hrsg. v. Fotis Jannidis / Hubertus Kohle / Malte Rehbein, Stuttgart 2017, S. 279–298.
- SCHÖCH, Christof: Topic Modeling Genre: An Exploration of French Classical and Enlightenment Drama, in: Digital Humanities Quarterly 11.2 (2017).
- SCHÖLLER, Robert: Die Fassung \*T des Parzival Wolframs von Eschenbach. Untersuchungen zur Überlieferung und zum Textprofil (Quellen und Forschungen zur Literatur- und Kulturgeschichte 56 [290]), Berlin / New York 2009.
- SCHÖLLER, Robert: Von der Handschrift zum Druck: zur Editionsgeschichte des Parzival, in: Schachzabel, Edelstein und der Gral. Spätmittelalterliche Handschriftenschatze der Burgerbibliothek Bern, hrsg. v. der Burgerbibliothek Bern (Passepartout 1), Bern 2009, S. 73–77.
- STACKMANN, Karl: Mittelalterliche Texte als Aufgabe, in: Festschrift für Jost Trier zum 70. Geburtstag, hrsg. v. W. Foerste / K. H. Borck, Köln / Graz 1964, S. 240–267, wieder in: Karl

- STACKMANN: Mittelalterliche Texte als Aufgabe. Kleine Schriften 1, hrsg. v. Jens Haustein, Göttingen 1997, S. 1–23.
- STOLZ, Michael: Intermediales Edieren am Beispiel des Parzival-Projekts, in: Wege zum Text. Beiträge des Grazer Kolloquiums über die Verfügbarkeit mediävistischer Editionen im 21. Jahrhundert (17.–19. September 2008), hrsg. v. Wernfried Hofmeister/Andrea Hofmeister-Winter, Tübingen 2009 (Beihefte zu editio 30), S. 213–228.
- TAYCHER, Leonid: Books of the World, Stand up and Be Counted! All 129,864,880 of you, 2010 [Web Log–Eintrag], abgerufen von:  
<http://booksearch.blogspot.de/2010/08/books-of-world-stand-up-and-be-counted.html>.
- THALLER, Manfred: Digital Humanities als Wissenschaft, in: Digital Humanities. Eine Einführung, hrsg. v. Fotis Jannidis/Hubertus Kohle/Malte Rehbein, Stuttgart 2017, S. 13–18.
- The Shape of Data in Digital Humanities. Modeling Texts and Text-based Resources, hrsg. v. Julia Flanders/Fotis Jannidis, London 2018.
- UNDERWOOD, Ted: A Genealogy of Distant Reading, in: Digital Humanities Quarterly 11 2017.
- VIEHHAUSER-MERY, Gabriel: Die Parzival-Überlieferung am Ausgang des Manuskriptzeitalters. Handschriften der Lauberwerkstatt und der Straßburger Druck (Quellen und Forschungen zur Literatur- und Kulturgeschichte 55 [289]), Berlin / New York 2009.
- VIEHHAUSER, Gabriel: Historische Stilometrie? Methodische Vorschläge für eine Annäherung textanalytischer Zugänge an die mediävistische Textualitätsdebatte, in: Grenzen und Möglichkeiten der Digital Humanities, hrsg. v. Constanze Baum/Thomas Stäcker (= Sonderband der Zeitschrift für digitale Geisteswissenschaften 1), 2015.
- VIEHHAUSER, Gabriel: Digitale Gattungsgeschichten. Minnesang zwischen generischer Konstanz und Wende, in: Zeitschrift für digitale Geisteswissenschaften 2017, text/html Format. DOI: 10.17175/2017\_003.
- VIEHHAUSER, Gabriel: Digital Humanities als Geisteswissenschaften. Zur Auflösung einer Tautologie, in: Digital Humanities. Perspektiven der Praxis, hrsg. v. Peggy Bockwinkel/Beatrice Nickel/Gabriel Viehhauser, Berlin 2018 (Digitalisierung und Globalisierung der Wissenschaften 1), S. 17–41.

## Internetlinks

<https://machinelearningmastery.com/what-are-word-embeddings/>.

Mittelhochdeutsche Begriffsdatenbank (MHDBDB), Universität Salzburg, Koordination: Katharina Zeppezauer-Wachauer, 1992-2019;  
<http://mhdbdb.sbg.ac.at/>.

<http://www.parzival.unibe.ch> (Parzival-Projekt).