MICHAEL BENDER

Vom Algorithmic Criticism in der digitalen Forschungsumgebung zur linguistischen Korpushermeneutik

Abstract

Im Beitrag wird die Entwicklung einer theoretischen und methodischen Forschungsperspektive in der digitalen Linguistik beschrieben, die mit dem Begriff Korpushermeneutik bezeichnet wird. Dabei wird der interdisziplinäre Entstehungshintergrund in den Digital Humanities dargestellt. Fokussiert werden Verfahren der Annotation bzw. Kategorienbildung und Theorieentwicklung sowie erkenntnistheoretische Diskussionen um datengeleitete vs. theoriegeleitete Verfahren, die in verschiedenen Disziplinen geführt werden – insbesondere vor dem Hintergrund der Digitalisierung.

This article describes the development of a theoretical and methodological research perspective in digital linguistics, which is referred to as corpus hermeneutics. It describes the interdisciplinary background to its development in the digital humanities. The focus is on methods of annotation, category formation and theory development as well as epistemological discussions about data-driven vs. theory-driven methods that are conducted in various disciplines - especially against the background of digitization.

Keywords: Digital Humanities, Bedarfsanalysen, Annotation, Kategorienbildung, Theoriebildung, datengeleitet vs. theoriegeleitet, Korpushermeneutik, digitale Linguistik

Digital Humanities, Requirements Engineering, Usability Studies, Annotation, Category Formation, Theory Development, data-driven vs. theory-driven, Corpus Hermeneutics, Digital Linguistics

1 Einleitung

In der Ich-Form eine Publikation für eine sprachwissenschaftliche Zeitschrift zu verfassen, ist gewöhnungsbedürftig, zumindest für den Autor des vorliegenden Beitrags. Sogar für sein früheres, journalistisches Ich war das Schreiben in der ersten Person Singular – da bezogen auf Zeitungsartikel – eine Ausnahme, allenfalls in Selbsterfahrungsreportagen erlaubt. Ansonsten galt (gilt?) auch dort offenbar eine Art Ich-Tabu bzw. ein angenommenes „Ich-Verbot", wie es Wein-

Osnabrücker Beiträge zur Sprachtheorie 104 (2025), 157-171.
Published online 20-05-25; DOI https://doi.org/10.17192/obst.2025.104.8810

rich (1989, 132-133) im Hinblick auf Wissenschaftstexte konstatiert hat. Dagegen soll hier verstoßen werden. So wurde es von den Herausgeber/inne/n dieses Hefts explizit gefordert. Da muss ein wenig Reflexion vorab erlaubt sein.

Dieses wahrgenommene Phänomen der Vermeidung der Ich-Form, das bereits von Polenz (1981, 105) im Zuge der Untersuchung von grammatischen Deagentivierungsformen in wissenschaftlichen Texten beobachtet hat, lässt sich – sowohl im wissenschaftlichen als auch im journalistischen Kontext – wie folgt erklären: Einerseits wird mit der unpersönlichen Form versucht, dem Objektivitätspostulat gerecht zu werden, sich selbst zurückzunehmen, für die Aufgabe der intersubjektiven Wissensvermittlung von der Kommunikationssituation und der individuellen Perspektive zu abstrahieren und eine neutral referierende Rolle einzunehmen. Andererseits spielt in der Wissenschaftssprache auch eine Rolle, dass Ich-vermeidende Passivformen, Reflexivkonstruktionen oder Nominalisierungen zum „common sense" (Feilke/Steinhoff 2003, 118) gehören, Wissenschaftlichkeit kontextualisieren und deshalb routinemäßig häufig verwendet werden (vgl. Steinhoff 2007, 5). Selbst so umständliche Konstruktionen wie „Autor des vorliegenden Beitrags" (siehe erster Satz des vorliegenden Beitrags) lassen sich auf diese Weise erklären. Diese Spannung zwischen (auch inszenierter) Objektivierung als Kriterium für Wissenschaftlichkeit, intersubjektiv-diskursiver Vermittlung/Aushandlung und der zunächst individuellen Verstehens- und Erkenntnis-Perspektive des/der Forschenden, die der Ich-oder-nicht-Ich-Frage inhärent ist, stellt auch in Diskussionen über den geisteswissenschaftlichen Umgang mit der Digitalität einen zentralen Aspekt dar. Immer wieder Thema ist etwa das als Gegensatz konstruierte Verhältnis von hermeneutisch-interpretativen und quantifizierend-statistischen Methoden. Insofern lässt sich mit der Ich-Form der rote Faden dieses Beitrags entwickeln und an das Hauptthema anknüpfen.

Doch bleiben wir zunächst noch kurz bei der Verbots- bzw. Tabu-Annahme. Empirische Studien (vgl. Steinhoff 2007, 7-10) zeigen, dass sie nicht pauschal bestätigt werden kann. ‚Ich' wird zwar nicht durchgängig, aber durchaus verbreitet verwendet. Steinhoff hat drei funktionale Typen der Ich-Verwendung herausgearbeitet: erstens das Verfasser-Ich, das vor allem für die metatextuelle Deskription der Vorgehensweise des Verfassers verwendet wird (vgl. Steinhoff 2007, 13-16), zweitens das Forscher-Ich, das im Rahmen von argumentativen und referentialisierenden Prozeduren zum Tragen kommt, etwa bei Begriffsbestimmungen oder Hypothesenformulierungen (vgl. Steinhoff 2007, 17-21), und drittens das Erzähler-Ich, das in autobiographisch-narrativen Textpassagen Verwendung findet, Steinhoff zufolge in wissenschaftlichen Texten jedoch selten ist und vor allem in studentischen Arbeiten, die in der Studie ebenfalls untersucht wurden, zu finden ist (vgl. Steinhoff 2007, 21-23). Auto-Narration stellt also das eigentliche Tabu in wissenschaftlichen Texten dar. Doch auch dagegen darf/soll

hier verstoßen werden, „narrativ" und „forschungsbiographisch" soll es werden – allerdings an Leitfragen orientiert und nicht auf Selbstbespiegelung ausgerichtet.

Also gut, dann wollen wir uns mal vorsichtig über das Verfasser-Ich herantasten. Ich werde im Folgenden die Leitfragen nach disziplinären und interdisziplinären Herausforderungen im Bereich der computergestützten Textanalyse, aber auch allgemeiner auf dem Gebiet der digital gestützten geisteswissenschaftlichen Forschungsarbeit, vor dem Hintergrund meiner Forschungsbiographie reflektieren. Das werde ich vor allem in Form des Erzähler-Ichs tun. Ich behalte mir aber vor, auch mal das Forscher-Ich zu Wort kommen zu lassen, wenn es gerade passend erscheint. Die Leitfragen werde ich aus drei verschiedenen Perspektiven reflektieren.[1] Die erste Perspektive bilden Erfahrungen aus meiner Begleitforschung – insbesondere der Bedarfserhebung bzw. dem Requirements-Engineering – zur Entwicklung virtueller Forschungsumgebungen und Forschungsdateninfrastrukturen in den Digital Humanities. Die zweite umfasst Reflexionen aus dem interdisziplinären Forschungsprojekt ePoetics (gefördert im Rahmen der eHumanities-Förderlinie des BMBF), in dem ich als Linguist und Computerphilologe mit Forschenden aus der Literaturwissenschaft, Computerlinguistik und Informatik an der Erschließung und Analyse sprach- und literaturtheoretischer Schriften zusammengearbeitet habe. Drittens reflektiere ich aus Sicht meiner aktuellen Forschung im Rahmen der digitalen Linguistik computergestützte Textanalyse als korpushermeneutischen Prozess.

2 Bedarfsanalysen zum digital gestützten Forschen

Als ich die Leitfragen im Exposé zu diesem Heft las, musste ich gleich an die Leitfragen denken, die ganz am Beginn meiner Forschungsarbeit standen. Sie gingen in eine ähnliche Richtung, waren ebenfalls an Forschende aus verschiedenen Disziplinen gerichtet, allerdings nicht nur an solche mit hoher Affinität zum computergestützten Arbeiten wie hier, sondern mit unterschiedlich ausgeprägtem, aber zumindest ansatzweise vorhandenem Digital-Hintergrund. Insofern waren die Leitfaden-Interviews, die ich geführt habe, nicht darauf ausgerichtet, retrospektiv eine fachgeschichtliche Entwicklung zu reflektieren, sondern prospektiv die Potenziale des computergestützten, wissenschaftlichen

1 Diese Perspektiven sind geprägt durch meine verschiedenen wissenschaftlichen Arbeitsstellen (nach journalistischer Tätigkeit während des Erststudiums). Zunächst habe ich mit der Promotion im Bereich der Digital Humanities parallel zu einem Zweitstudium begonnen (gefördert mit einem Stipendium), anschließend auf DH-Projektstellen geforscht, mich dabei immer mehr auf linguistische Aspekte konzentriert und nach der Promotion eine PostDoc-Stelle in der (digitalen) Linguistik angetreten.

Arbeitens. Dies war einerseits dem Zweck der Interviews geschuldet, Bedarfserhebung bzw. Requirements-Engineering im Rahmen der Entwicklung von digitalen Forschungsumgebungen und Forschungsdateninfrastrukuren zu leisten, und andererseits dem schon ein wenig zurückliegenden Zeitpunkt der Interviews in einer früheren Entwicklungs- und Etablierungsphase der Digital Humanities. Der Methode geschuldet war die Offenheit der Fragen, die darauf ausgerichtet waren, durch die Erweiterung von Antwortspielräumen unverfälscht die von den Befragten selbst konstruierten und thematisierten Bezugsrahmen, Erfahrungshintergründe, Relevanzstrukturen und letztlich Bedarfsschwerpunkte zu erfassen. Ich wollte wissen, welche Konzepte die Befragten von sich aus ansprechen und nicht, wie sie auf Fragen etwa nach ‚Methodengetriebenheit' oder ‚Innovationsdruck' reagieren. Als Leitfaden für ein solches Heft sind das natürlich spannende Impulse, eine Befragung würden diese Thematisierungen aber verzerren. Immerhin bestand ja auch die Möglichkeit, dass durch die Digitalisierung getriebener Innovationsdruck in bestimmten disziplinären Kulturen gar keine Rolle spielt. Ich habe also als ganz allgemeinen Erzähl-Stimulus Fragen gestellt nach bisherigen Erfahrungen, Anwendungen und Forschungsaktivitäten der Interviewten im Bereich der digital gestützten Forschungsarbeit sowie nach Erwartungen im Hinblick auf Funktionalitäten und Erkenntnismöglichkeiten – nach Hands-On-Workshops zum Testen von Software auch konkreter auf Tools und Datenrepositorien bezogen, z. B. TextGrid. Ausgewertet habe ich die erhobenen Daten in Form einer qualitativen Inhaltsanalyse mit induktiver Kategorienbildung und Kodierung (vgl. Bender 2016), einer sozialwissenschaftlichen Methode, die ich mit Blick auf die sprachliche Oberfläche sowie diskurs- und pragmalinguistische Analysekategorien angepasst und später auch im Rahmen der Entwicklung digitaler linguistischer Annotationsverfahren einbezogen habe. Insofern hat die Bedarfsforschung in diesem Gebiet meine eigene digitale Methodenentwicklung evoziert.

Diesen Zugang zu Fragestellungen, die den hier gestellten Leitfragen ähnlich bzw. hier auch relevant sind, habe ich auch in weiteren Projekten und Gremien weiterverfolgt. Im Rahmen von Digital-Humanities-Forschungsverbünden wie DARIAH (Digital Research Infrastructure for the Arts and Humanities)[2] war ich in unterschiedliche Workshop- und Schulungsformate und die integrierten Befragungen, Bedarfserhebungen und Feedback-Schleifen involviert. Im Verbundprojekt HUMANIST (Humanist Computer Interaction auf dem Prüfstand)[3], 2017–2021 gefördert vom BMBF im Programm „Validierung des technologischen und gesellschaftlichen Innovationspotenzials wissenschaftlicher Forschung – VIP+", war ich in die Antragstellung und im Mentoring eingebunden.

2 https://de.dariah.eu/web/guest/startseite

3 https://www.hs-mainz.de/microseiten/humanist-computer-interaction/projekt/

Ziel dieses Projekts war es, herauszufinden, welche digitalen Methoden, Tools, Ressourcen bzw. Merkmale von digitalen Forschungsumgebungen das Potenzial haben, innovative geisteswissenschaftliche Anwendungen und Kollaborationsformen zu ‚triggern'. Und auch heute noch, vor dem Hintergrund und auch trotz der auch in der Wissenschaft als allgegenwärtig wahrgenommenen Digitalisierung, spielen diese Fragen nach dem Bedarf an und dem möglichen Mehrwert von digitalen Forschungsmöglichkeiten nach wie vor eine wichtige Rolle, werden etwa in den Arbeitsgruppen und Gremien von Text+ als Teil der Nationalen Forschungsdateninitiative (NFDI) diskutiert. Gerade habe ich in diesem Rahmen an einem Meeting teilgenommen, in dem wir uns wieder bewusst gemacht haben, dass es im Umgang mit potenziellen Nutzer/inne/n wichtig ist, nicht missionarisch bestimmte Anwendungen zu empfehlen, sondern den Balanceakt zu schaffen, einerseits zu zeigen, was möglich ist, und andererseits ganz offen zuzuhören und möglichst unvoreingenommen und unverfälscht zu erfassen, was wirklich gebraucht wird, welche Vorstellungen, Fragestellungen, Zielsetzungen und praktischen Anforderungen es in verschiedenen Disziplinen gibt.

Nun kann ich hier nicht im Detail über die Ergebnisse unterschiedlicher, über viele Jahre verteilter Bedarfserhebungen und -analysen berichten. Aber eine Zusammenfassung der entscheidenden Punkte, die ich über die Jahre beobachtet habe, im Hinblick auf die hier gestellten Fragen ist schon möglich. Ganz klar im Vordergrund stand und steht die Problemorientierung der Forschenden, um es mit der im Exposé zu diesem Heft verwendeten Terminologie auszudrücken. Oder anders gesagt: Der formulierte Bedarf der Befragten ist darauf ausgerichtet, genau das computergestützt tun zu können, was sie für ihr jeweiliges Erkenntnisinteresse bzw. ihre Forschungsziele brauchen. Dafür wird auch eine gewisse Komplexität von digitalen Werkzeugen, Methoden und Ressourcenzugriff in Kauf genommen. Zwar besteht grundsätzlich Interesse an digitaler Methodeninnovation, aber „Methodengetriebenheit" oder „Innovationsdruck" durch digitale Möglichkeiten war keine Kategorie der Befragten. Nicht gefragt ist die ‚eierlegende Wollmilchsau', wenn es auf Kosten der disziplinspezifischen Spezialanforderungen geht. Anpassungsmöglichkeiten an die spezifischen Anforderungen, aber auch Interoperabilität durch Datenstandards und damit verbundene Kollaborationsmöglichkeiten werden als wichtig erachtet. Wohlgemerkt beziehen sich diese Tendenzen allgemein auf Geisteswissenschaftler/innen, die digital gestützt arbeiten oder Interesse daran haben, nicht speziell auf diejenigen, die ihre Disziplin- bzw. Forschungs-Identität durch computergestützte Analysen definieren. Hier gibt es ein breites Spektrum an Selbstbildern, an Vorstellungen von der Relation der geisteswissenschaftlichen Disziplin zur computergestützten Methodik und an fachlichen Verortungen – von der Integration des Digitalen in die geisteswissenschaftlichen Fächer über die Konzeption der Digital Humanities als separate Disziplin und Community mit eigenen

Tagungen bis hin zu Verortungen der DH als Teilfach der Informatik. In diesem Spektrum ist die Frage nach dem durch digitale Entwicklung getriebenen Innovationsdruck natürlich naheliegend – und gerade in der informatiknahen oder computerlinguistischen Richtung – meiner Beobachtung (allerdings als Nicht-Informatiker) nach sicher ein relevantes Thema. Die Bedeutung dieses Themas hängt sicherlich mit der forschungskulturellen Prägung zusammen, bei der es zwischen Geisteswissenschaften und der Informatik (eventuell auch der informatiknahen Computerlinguistik) immer noch große Unterschiede gibt. Diesen Clash der Forschungskulturen habe ich im Rahmen des eHumanities-Projekts ePoetics miterlebt, worauf ich im nächsten Abschnitt eingehe.

3 Der Algorithmic Criticism im Clash der Forschungskulturen

Das Projekt ePoetics war eine Kooperation zwischen der TU Darmstadt, für die ich als Mitarbeiter im Fachgebiet Computerphilologie beteiligt war, und den Instituten für Literaturwissenschaft, Computerlinguistik und Informatik der Universität Stuttgart. Gegenstand des Projekts waren 20 Poetiken (eine zweibändig, eine dreibändig, insgesamt daher 23 Bücher), allerdings keine Regelpoetiken, sondern spätere sprach- und literaturtheoretische Schriften aus dem Zeitraum von 1804 bis 1959 (vgl. Richter 2010, I.2, 19-24 und I. 15, 288-298.). Diese befassen sich nicht normativ, sondern deskriptiv und analytisch mit Dichtung und mit Sprache im Allgemeinen. Dabei werden auch theoretische Ansätze, Konzepte und Begriffe vor dem Hintergrund des Forschungsstandes und Fachdiskurses der Zeit diskutiert und kritisch auf antike sowie zeitgenössische Poetiken Bezug genommen.

Der Clash der Forschungskulturen war beabsichtigt und schon durch die Projektziele – ja: programmiert. Ein Ziel war die Digitalisierung und computerphilologisch erschließende Annotation der Poetiken für die Publikation im deutschen Textarchiv, was auch umgesetzt wurde (vgl. Alscher/Bender/Rapp et al. 2016). Die sprach- und literaturwissenschaftliche Perspektive zielte vor allem auf Diskursreferenzierungen zwischen den Poetiken untereinander und Verweisungsstrukturen zur Primärliteratur, aber auch auf die Entwicklung sprach- und literaturtheoretischer Begriffe wie dem der *Erhabenheit* oder der *Metapher*. Kritische Diskursivität und die damit verbundenen Bezugnahmen und intertextuellen Verknüpfungen stellen eine Besonderheit dieser Schriftensammlung (im Vergleich zu Regelpoetiken) dar. Wir haben verschiedene explizite und implizite Formen der Referenz auf dichterische Primärliteratur und die poetologische Sekundärliteratur in einem Kategoriensystem modelliert. Inhaltlich haben wir exemplarisch die Entwicklung der definitorischen Merkmale des Metaphernbegriffs annotierend nachvollzogen. Wir waren auf der Suche nach dem „goldnen Baum“, der passenden taxonomischen Baumstruktur als Annotationsschema für

Referenzstrukturen und die Erfassung der Entwicklung theoretischer Begriffe. Dieses Schema hat letztlich ermöglicht, die Entwicklung der Verweisungslinien auf verschiedene Schulen der Metapherntheorie sichtbar zu machen (vgl. Alscher/Bender 2016).

Die Computerlinguistik arbeitete parallel am automatisierten Erkennen von Diskursreferenzierungen durch Namen oder Werktitel, Phrasen, die eine Attribuierung enthalten, und der Klassifizierung von Text in Anführungszeichen (als Zitat vs. Hervorhebung z. B.). Aufgrund der sehr heterogenen Verwendung von Diskursreferenzierungen, teilweise auch nur in Form von Anspielungen, sowie von Zitations- bzw. Literaturverweisungspraktiken in den sehr eigenständigen, formal nicht konventionalisierten Poetiken im über 150 Jahre umfassenden Korpus konnte zwar dessen Exploration computerlinguistisch unterstützt werden. Die Entwicklung entsprechender Algorithmen war jedoch nur durch Training mittels manueller Annotationen möglich. Von informatischer Seite wurde im Bereich interaktiver Visualisierung ein Tool namens *Varifocal Reader* für die Navigation in den digitalisierten Poetiken entwickelt. Es ermöglichte den Zugriff auf verschiedene strukturelle Hierarchie-Ebenen – das Faksimile, einzelne Abschnitte und Kapitel-Ebenen sowie die Inhaltsverzeichnis-Ebene, die Anzeige von bereits bestehenden Annotationen auf verschiedenen Abstraktionsebenen, außerdem die Visualisierung von Suchergebnissen in unterschiedlichen Perspektiven sowie Darstellungen von Worthäufigkeiten. Auch dieses Tool hat zunächst basal die Annotationsebene unterstützt, war aber ebenfalls auf maschinelles Lernen ausgelegt und auf das Training durch sukzessive differenziertere, manuelle Annotation angewiesen. Computerlinguistik und Informatik haben also in dem Projekt Sprach- und Literaturwissenschaftler mit Tools unterstützt und neue entwickelt. Ihr Hauptziel war dabei allerdings die Optimierung der Algorithmen (vgl. John et al. 2018).

Wir hingegen – der Kollege aus der Literaturwissenschaft und ich – waren neben dem Ziel der Publikation eines digitalisierten Korpus bzw. einer digitalannotierten Edition zunächst einmal an der geisteswissenschaftlichen Forschungsfrage nach Diskursreferenzierungs-Praktiken und der Terminologieentwicklung interessiert, aber auch an deren Operationalisierbarkeit und Modellierung in einem Annotationsschema mit dem Ziel der Automatisierung. Es gab also schon eine konstruktive Schnittstelle zwischen den Disziplinen, aber auch einen forschungskulturellen Kontrast und unterschiedliche Prioritäten hinsichtlich der Erkenntnisziele.

Ein wichtiger theoretischer bzw. methodenreflektierender Hintergrund war für uns – die philologisch geprägte Projekt-‚Fraktion' – der *Algorithmic Criticism* nach Steven Ramsay, ein Forschungsparadigma, das Wechselwirkungen zwischen hermeneutischen und algorithmischen Methoden beschreibt und ursprünglich auf die kritische Interpretation (nicht Literaturkritik) literarischer

Texte bezogen war. Ramsay grenzt seine Idee explizit von Ansätzen ab, die nur auf die computergestützte Überprüfung hermeneutisch erstellter Hypothesen abheben oder quantitative Methoden der traditionellen hermeneutischen Textanalyse im empirisch-objektivierenden Sinne entgegensetzen. Er formuliert sein Forschungsparadigma hingegen als vermittelnden Ansatz, der das Gegensatzmodell überwinden soll. Ramsay geht davon aus, dass Bedeutungskonstitution im Rezeptions- bzw. Verstehensprozess stattfindet und die komplexen, interaktiv und interpretativ konstruierten Bedeutungsstrukturen nicht direkt als Texteigenschaften mit empirisch-objektiven Methoden erfasst werden können. Quantitativ-empirische Aspekte fließen aber durchaus in hermeneutische Überlegungen ein (und umgekehrt), stellen Zwischenschritte im Forschungsprozess dar, können Argumente beitragen, neue Perspektiven eröffnen und Theoriebildungsprozesse generieren. Sie haben vor allem die Aufgabe der Transformation der Ausgangstexte, um so neue multiperspektivische Rezeptionsmöglichkeiten zu schaffen. Ein entscheidender Punkt in Ramsays Idee ist, dass wissenschaftliche Fragestellungen, Interpretationen und Erkenntnisse – zumindest in den Geisteswissenschaften – letztlich nicht allein durch empirisch objektivierte Mess-Ergebnisse verifiziert, sondern in der wissenschaftlichen Community diskursiv verhandelt und anerkannt werden:

> 'Verification' occurs in a social community of scholars whose agreement or disagreement is almost never put forth without qualification (Ramsay 2007, 54).

Voraussetzung dafür ist, dass Kategorisierungskriterien und Interpretationen explizit und somit intersubjektiv nachvollziehbar gemacht werden. Dies wurde im Projekt im Zuge der Aushandlung des „goldnen Baums", der Goldstandard-Annotation bis hin zum fertigen Kategorien-Schema in Form von Richtlinien mit Annotationskriterien (Guidelines) umgesetzt. Die Kategorienbildung und die damit verbundenen Entscheidungen über die Auswahl der zu annotierenden Segmentgröße (Wort, Phrase, Satz etc.), Granularität, Interpretationstiefe, Ausgewogenheit und Trennschärfe der Kategorien im Rahmen einer distinktiven Taxonomie wurde mit Blick auf die Operationalisierbarkeit für algorithmische Analysen sozusagen vor der Hintergrundfolie des Digitalen entwickelt. Dies zeigte zugleich auf, was automatisierbar ist und was nicht, wo die Möglichkeiten und Grenzen des Algorithmischen zu verorten sind, also auch einen „Criticism" im Hinblick auf die computerlinguistischen und informatischen Methoden. In diesem Projekt gab es also durchaus eine Herausforderung in Form des Umgangs mit algorithmischen Textanalyse- bzw. Machine-Learning-Methoden, die man als Innovationsdruck beschreiben könnte. Meines Erachtens war Innovationsdruck in dem Sinne, dass man die neuesten algorithmischen bzw. informati-

schen Methoden anwenden muss, eher als typisches Phänomen der computerlinguistischen und informatischen Forschungskultur wahrzunehmen. Neben dem Aspekt, dass man in der Wissenschaft grundsätzlich unter Innovationsdruck steht, traf auf die Erfahrung in diesem Projekt aus geisteswissenschaftlicher oder besser philologischer Perspektive eher der im Konzept zu diesem Heft aufgeführte Punkt des Achtens auf „Einhaltung rigoroser disziplinärer Qualitätskriterien [als] notwendiges Korrektiv" zu. Dies wirkte jedoch nicht als „Hindernis für inter- und transdisziplinäre Innovationen", sondern als methodengetriebene, aber kritische Reflexion und Operationalisierung nach fachlichen Kriterien. Bezüglich dieser Kriterien haben sich im Zuge der Kategorienbildung und des Annotierens wiederum Unterschiede zwischen der linguistischen und der literaturwissenschaftlichen Perspektive gezeigt. Diese waren nicht nur auf fachlich bedingte unterschiedliche Wissensbestände und Relevanzstrukturen als Hintergrund für die Kategorisierung zurückzuführen. Auch Differenzen in der Orientierung an der sprachlichen Oberfläche (wie z. B. unterschiedliche, fachlich geprägte Segmentierungseinheiten) und der Interpretationstiefe flossen in die Aushandlungsprozesse im Zuge der Entwicklung des Kategoriensystems ein.

4 Korpushermeneutik in der digitalen Linguistik

Die in den beiden vorigen Abschnitten beschriebenen, vor allem interdisziplinären Erfahrungen im Bereich der computergestützten Textanalyse in den Digital Humanities haben meine aktuelle korpuslinguistische Forschung methodisch mitgeprägt. Mein methodischer Schwerpunkt auf der Kombination manuell-interpretativer Annotationsverfahren mit frequenz- und distributionsorientierten Verfahren der Korpusstatistik hat sich vor diesem Hintergrund entwickelt. Dabei habe ich vor allem auf dem Gebiet der Kategorienbildung und Annotation Ansätze aus verschiedenen Disziplinen integriert.

Annotation ist in der Linguistik vor allem durch computerlinguistische Verfahren des routinemäßigen Preprocessing, der Aufbereitung von Daten für sprachstatistische Analysen, geprägt worden. In diesen Prozessen werden in einem distinktiven und exhaustiven Kategoriensystem Tags festgelegt, wobei es sich vor allem um kategoriale (z. B. Wortarten) oder funktionale (z. B. Koreferenzstrukturen) systemlinguistische Klassen handelt, aber auch weniger basale linguistische Theoriemodelle operationalisiert werden (Überblick vgl. Ide 2017, 1-18). Dazu werden entsprechende Guidelines formuliert. Nach Probeannotationen durch mehrere Annotierende und Überprüfung des Inter-Annotator-Agreement (Artstein 2017) wird ein Goldstandard hergestellt, mit dem Tools für die automatisierte Auszeichnung großer Textkorpora trainiert werden. Das ist die Hauptzielrichtung des computerlinguistischen Annotierens.

> The past four decades have seen a great deal of evolution in strategies and ‚best practices' (de facto or otherwise) for linguistic annotation, spurred in particular by the need for gold standard data to train machine learning algorithms (Ide 2017, 15).

Insofern lässt sich das linguistische Annotieren auch als durch die Entwicklung algorithmischer Methoden getrieben ansehen. Manuelles digitales Annotieren wird in der Linguistik im Hinblick auf implizite Phänomene bzw. Merkmale insbesondere in der linguistischen Pragmatik praktiziert (vgl. z. B. Weisser), auch da mittlerweile mit Blick auf maschinelles Lernen (Bender 2023).

Dabei muss besonders die Frage nach der angemessenen Segmentierung, also inwiefern implizite Phänomene mit linguistischen Einheiten der sprachlichen Oberfläche aligniert werden können, wieder neu diskutiert werden (Becker/Bender/Müller 2020; Bender/Becker/Kiemes/Müller 2023). In der digitalen Literaturwissenschaft sind Segmentierungseinheiten wie in der Linguistik bspw. Wortarten, Phrasenkategorien etc. gar nicht standardmäßig vorhanden, werden in einigen Annotationsansätzen gar nicht einbezogen. Das Fehlen standardisierter Segmentierungseinheiten stellt aber unter Umständen für die quantifizierenden Verfahren der computationellen Literaturwissenschaft, deren Messungen auf Unterteilungen von Texten angewiesen sind, ein Problem dar. Sie müssen deshalb je nach Untersuchungsgegenstand, Fragestellung und Interpretationstiefe ausgehandelt und festgelegt werden. Ein Beispiel wären Szenen(grenzen) in Novellen (vgl. Bartsch et al. 2023). Annotation wird in der digitalen Literaturwissenschaft aber auch mit offen konzipierten und sukzessive entwickelten Annotationsschemata als Akt der interpretativen Aneignung von Texten angewendet, ohne dass statistische Auswertbarkeit im Mittelpunkt steht. Abweichungen beim kollaborativen Annotieren können auch der Identifikation von deutungsoffenen Passagen dienen, die mehrere Lesarten erlauben (vgl. Gius/Jacke 2016, 8). Wie mit der Idee des *algorithmic criticism* oben beschrieben, spielt auch in diesen Ansätzen die intersubjektive Aushandlung von Kategorien, Kriterien und Interpretationen eine zentrale Rolle. Diese Perspektiven haben mich dazu angeregt, auch in linguistischen Annotationsprojekten stärker die Reflexion, Innovation und intersubjektive Aushandlung von Segmentierungen und Kategorisierungen zu bedenken.

Der Aspekt der Segmentierung bzw. der Relation zwischen sprachlicher Oberfläche und kategorisierender Interpretation der dahinterstehenden Fachperspektive stellt meines Erachtens auch den entscheidenden Unterschied zwischen dem linguistischen (bzw. philologischen) Annotieren und dem sozialwissenschaftlichen Kodieren dar, das ich wie oben beschrieben im Rahmen der Interviewanalyse angewendet habe. In sozialwissenschaftlichen Analysen schriftlich verfasster, transkribierter oder protokollierter Kommunikation in Inhaltsanalysen (vgl. z. B. Mayring 2010) oder Grounded-Theory-Studien (vgl.

Strauss/Corbin 1996) werden verschiedene Aspekte der Kommunikation untersucht, nicht nur thematisch-propositionale Elemente, die mit dem problematischen Begriff *Inhalt* beschrieben werden, sondern auch Intentionen, Wirkungen etc., aus denen Rückschlüsse auf soziale Strukturen gezogen werden. Die zahlreichen verschiedenen Ansätze in diesem Bereich unterscheiden sich im Hinblick auf die Tendenz zum theoriegeleiteten oder datengeleiteten Vorgehen. Gemeinsamkeiten der Verfahren bestehen aber in der Selektion von relevanten Textpassagen, wobei hier durchaus auch Segmentgrößen festgelegt werden können, die aber nicht unbedingt an sprachwissenschaftlichen Kategorien und Kriterien orientiert sind. So werden in der qualitativen Inhaltsanalyse nach Gläser und Laudel (2010, 209–212) beispielsweise durch Themenwechsel abgegrenzte „Sinnabschnitte" als „Analyseeinheiten" segmentiert, nach Mayring werden genauer Kodiereinheit (minimale Segmentierungsgröße), Kontexteinheit (maximal bei der Kategorisierung berücksichtigte Texteinheit) und Auswertungseinheit (nacheinander auszuwertende Textteile) bestimmt. In der Grounded Theory wird auch die Segmentierung zunächst variabel gehandhabt und offengehalten. Sie wird an den relevanten Phänomenen ausgerichtet, die identifiziert werden müssen, bevor passende Segmente festgelegt werden können. In den induktiv ausgerichteten Ansätzen wird von diesen selektierten Segmenten ausgehend abstrahiert. Die identifizierten Phänomene werden nach Gemeinsamkeiten gruppiert, nach Relationen geordnet und in ein Theoriemodell überführt, wobei nicht die quantitative Auswertung, sondern die qualitative Verdichtung zur Theorie das Ziel ist und eine Ablösung von der empirischen, sprachlichen Oberfläche stattfindet. Diese unterschiedlichen Verfahrensweisen haben mich ebenfalls dazu inspiriert, zur Segmentierung und Kategorisierung im Theoriebildungsprozess im Rahmen von Annotationsstudien auch aus linguistischer Perspektive weiter zu forschen und Möglichkeiten und Grenzen zu reflektieren.

Eine erkenntnistheoretische Dauerdebatte in diesem Zusammenhang und allen hier thematisierten Disziplinen betrifft die Gegenüberstellung von datengeleitetem und theoriegeleitetem Forschen. Die in den Geisteswissenschaften auch im Zuge der Digitalisierung wieder an Bedeutung gewinnende datengeleitete Perspektive ist vor allem darauf ausgerichtet, als Korrektiv zu theoriegeleiteten Vorannahmen zu wirken, die den Blick auf Erkenntnisse, die nicht mit Vorwissen kompatibel sind, verstellen können. Gerade in den Digital Humanities wird versucht, auf der Basis von größeren Datenmengen und der daraus gewonnenen Evidenz Phänomene sichtbar zu machen, die etwa durch den konventionell lesenden Zugriff nicht wahrnehmbar wären. In der digitalen Literaturwissenschaft ist das von Moretti geprägte Begriffspaar *close reading* vs. *distant reading* besonders prominent geworden (Moretti 2013, 48-49, 211). In der Korpuslinguistik wurden im Zuge der erkenntnistheoretischen Diskussion unter den Bezeichnungen *cor-*

pus-based und *corpus-driven* (Tognini-Bonelli 2001) theorie- und datengeleitete Ansätze einander gegenübergestellt. Mittlerweile ist deren integrative Kombination verbreitet.

In diesen Ansätzen wird in den verschiedenen Disziplinen zwar reflektiert, dass auch datengeleitete Forschung theoretisches Vorwissen einbezieht und Datenanalysen auch interpretativ sind. Die Explizierung und Reflexion theoretischer Vorüberlegungen und Interpretationskriterien wird in der datengeleiteten Forschung jedoch oft weniger fokussiert, steht zumindest nicht im Vordergrund. Eher werden die Chancen und Grenzen quantitativer Methoden reflektiert und Möglichkeiten ausgelotet, das als zu stark subjektiv geprägte und zu wenig repräsentative Hermeneutische zu überwinden und eine Annäherung an Objektivität herbeizuführen. Doch nicht nur in der digitalen Linguistik, sondern gerade auch in der interdisziplinären Forschung ist es meines Erachtens von besonderer Relevanz, Vorwissenskomponenten, Verstehens- und Interpretationsprozesse und den Prozess der Theoriebildung explizit und bewusst zu reflektieren und intersubjektiv auszuhandeln. Erst diese Explizierung ermöglicht die fundierte und reflektierte Überprüfung von Vorannahmen mit datengeleiteten Verfahren und das Aufdecken von potenziellen Bias-Problematiken sowie die Entdeckung neuer methodischer und theoretischer Ansätze. Insofern versuche ich, methodenentwicklungsgetriebene Innovation, korpusstatistische und Machine-Learning-Verfahren nicht nur als Methoden zu reflektieren, sondern mit der Reflexion hermeneutischer Interpretations- und Theoriebildungsprozesse vor der Folie des Digitalen und Algorithmischen zu verbinden. Diese Perspektive knüpft an die hermeneutischen Traditionen an, die in den verschiedenen geisteswissenschaftlichen Disziplinen vorhanden sind und insofern auch eine interdisziplinäre Verbindung darstellen. Eine ausführlichere Reflexion und Vertiefung dieser hermeneutischen Perspektive habe ich jüngst zusammen mit Katharina Jacob, einer der Herausgeber/innen dieses Bandes, mit Blick auf die germanistischen Teildisziplinen unter dem Begriff der *Korpushermeneutik* erarbeitet (vgl. Bender/Jacob/Spieß 2024). Dies stellt also eine Schnittstelle zu einer anderen forschungsbiographischen Perspektive in diesem Band dar. Ich entwickele diesen Ansatz aktuell spezifischer in meinen linguistischen Projekten weiter. Mit dieser ausblickartigen Perspektive komme ich zum Ende. Wir schreiben hier ja forschungsbiographische Reflexionen mitten aus dem laufenden Betrieb, keine Memoiren. Für ein richtiges Fazit wäre es also zu früh, finde ich. Deshalb beendet der Autor den vorliegenden Beitrag ohne obligatorischen Schlussteil.

Literatur

Alscher, Stefan/Bender, Michael (2016) Auf der Suche nach dem „goldnen Baum“. Digitale Annotation als Erkenntnisprozess und 'tertium comparationis' – am Beispiel der Auszeichnung des Metaphernbegriffs in Poetiken. In: *Zeitschrift für digitale Geisteswissenschaften*. DOI: https://doi.org/10.17175/2016_004

Alscher, Stefan/Bender, Michael/Rapp, Andrea et al. (2016) ePoetics – 20 Poetiken (1770-1960), TEI-konformes XML, HTML-Lesefassung, synoptische Faksimile-Ansicht, inkl. Annotationen des BMBF-Projekts epoetics. https://www.deutschestextarchiv.de/search/metadata?corpus=epoetics [zuletzt aufgerufen am 12.03.2024].

Artstein, Ron (2017) Inter-annotator agreement. In: Die, Nancy/ Pustejofsky, James (Hrsg.) *The Handbook of Linguistic Annotation – Introduction*. Dordrecht: Springer, 297–313.

Bartsch, Sabine/Gius, Evelyn/Müller, Marcus/Rapp, Andrea/Weitin, Thomas (2023) Sinn und Segment. Wie die digitale Analysepraxis unsere Begriffe schärft. In: *Zeitschrift für digitale Geisteswissenschaften* 8. DOI: https://doi.org/10.17175/2023_003

Becker, Maria/Bender, Michael/Müller, Marcus (2020) Classifying Heuristic Textual Practices in Academic Discourse: A Deep Learning Approach to Pragmatics. In: *International Journal of Corpus Linguistics* 4, 426–460.

Bender, Michael (2016) *Forschungsumgebungen in den Digital Humanities: Nutzerbedarf, Wissenstransfer, Textualität*. Berlin/Boston: De Gruyter. DOI: https://doi.org/10.1515/9783110463927

Bender, Michael (2023) Pragmalinguistische Annotation und maschinelles Lernen. In: Bülow, Lars/Marx, Konstanze/Meier-Vieracker, Simon/ Mroczynski, Robert (Hrsg.) *Digitale Pragmatik*. Stuttgart: Metzler, 267–286. https://link.springer.com/chapter/10.1007/978-3-662-65373-9_12 [zuletzt aufgerufen am 08.03.2024].

Bender, Michael/Becker, Maria/Kiemes, Carina/Müller, Marcus (2023) Category Development at the Interface of Interpretive Pragmalinguistic Annotation and Machine Learning – Annotation, Detection and Classification of linguistic routines of discourse referencing in political debates. In: *Digital Humanities Quarterly* 17(3) (Special Issue: Working on and with Categories for Text Analysis: Challenges and Findings from and for Digital Humanities Practices). http://www.digitalhumanities.org/dhq/vol/17/3/000720/000720.html [zuletzt aufgerufen am 12.03.2024].

Bender, Michael/Jacob, Katharina/Spieß, Constanze (2024) Korpushermeneutik. Verstehen, Interpretieren und Erklären in der digitalen germanistischen Forschung. In: *Zeitschrift für Literaturwissenschaft und Linguistik* (LiLi) 54(4). DOI: https://doi.org/10.1007/s41244-024-00329-2

Feilke, Helmuth/Steinhoff, Torsten (2003) Zur Modellierung der Entwicklung wissenschaftlicher Schreibfähigkeiten. In Ehlich, Konrad/Steets, Angelika (Hrsg.) *Wissenschaftlich schreiben - lehren und lernen*. Berlin/New York: De Gruyter, 112–128.

Gius, Evelyn/Jacke, Janina (2016) Zur Annotation narratologischer Kategorien der Zeit. Guidelines zur Nutzung des CATMA-Tagsets. http://heureclea.de/wp-content/uploads/2016/11/guidelinesV2.pdf. [zuletzt aufgerufen am 12.03.2024].

Gläser, Jochen/Laudel, Grit (2010) *Experteninterviews und qualitative Inhaltsanalyse – als Instrumente rekonstruierender Untersuchungen*. Wiesbaden: Springer.

Ide, Nancy (2017) Introduction: The Handbook of Linguistic Annotation. In: Die, Nancy/Pustejovsky, James (Hrsg.) *Handbook of Linguistic Annotation*. Band 1. Dordrecht: Springer, 1–18.

John, Markus/Bender, Michael/Alscher, Stefan/Müller, Andreas/Koch, Steffen/Kuhn, Jonas/Richter, Sandra/Rapp, Andrea/Ertl, Thomas (2018) Korpuserschließung und Visualisierung deutschsprachiger Poetiken. In: Bockwinkel, Peggy/Nickel, Beatrice/Viehhauser, Gabriel (Hrsg.) *Digital Humanities. Perspektiven der Praxis*. Berlin: Frank & Timme.

Mayring, Philipp (2010) *Qualitative Inhaltsanalyse. Grundlagen und Techniken*. Weinheim/Basel: Beltz.

Moretti, Franco (2013) *Distant reading*. London/New York: Verso.

von Polenz, Peter (1981) Über die Jargonisierung von Wissenschaftssprache und wider die Deagentivierung. In: Bungarten, Theo (Hrsg.) *Wissenschaftssprache. Beiträge zur Methodologie, theoretischen Fundierung und Deskription*. München, 85–110.

Ramsay, Stephen (2007) Algorithmic Criticism. In: Siemens, Ray/Schreibman, Susan (Hrsg.) *A Companion to Digital Literary Studies*. Oxford, 477–491.

Richter, Sandra (2010) *A History of Poetics. German Scholarly Aesthetics and Poetics in International Context, 1770–1960*. With Bibliographies by Zenk, Anja/Azazmah, Jasmin/Jost, Eva/Richter, Sandra. Berlin/New York: De Gruyter.

Steinhoff, Torsten (2007) Zum ich-Gebrauch in Wissenschaftstexten. In: *Zeitschrift für germanistische Linguistik* 35(1), 1–26. DOI: https://doi.org/10.1515/ZGL.2007.002

Strauss, Anselm/Corbin, Juliet (1996) *Grounded Theory: Grundlagen Qualitativer Sozialforschung*. Weinheim: Beltz/Psychologie Verlags Union.

Tognini-Bonelli, Elena (2001) *Corpus Linguistics at Work*. Amsterdam/Philadelphia: John Benjamins.

Weinrich, Harald (1989) Formen der Wissenschaftssprache. In: *Jahrbuch der Akademie der Wissenschaften zu Berlin*, 119–158.

Weisser, Martin (2018) *How to Do Corpus Pragmatics on Pragmatically Annotated Data*. Amsterdam/Philadelphia: John Benjamins.