

Michael Bender

## Annotation als Methode der digitalen Diskurslinguistik

*Abstract*

*Annotation hat sich von einer analog geprägten Kulturtechnik, deren Geschichte sich bis in die Antike zurückverfolgen lässt, hin zu einer digitalen Praktik entwickelt, die in verschiedenen Spannungsfeldern verortet werden kann. Sie wird einerseits als alltagsweltliche Rezeptions-, aber auch Kommunikationsform vollzogen (und ist somit potenzieller Gegenstand von linguistischen Untersuchungen, z.B. Hashtag-Analysen), andererseits auch als wissenschaftliche Methode angewendet. Methodisch bildet sie eine Schnittstelle zwischen hermeneutischer und automatisiert-quantifizierender Analyse. Aus diskurslinguistischer Perspektive verbindet das Annotieren wissenschaftliche Metadiskursivität mit diskursiver Involviertheit. Diese Verortungen werden im vorliegenden Aufsatz diskutiert und anhand von konkreten Annotationsprojekten veranschaulicht. Im Mittelpunkt steht dabei das kollaborativ-diskursive Annotieren als wissenschaftliche Methode in der digitalen (Diskurs-)Linguistik.*

### 1. Einleitung

Im Zuge der Digitalisierung hat sich ein vielfältiges Anwendungsspektrum des Annotierens entwickelt – von alltäglichen Mediennutzungspraktiken des Taggens bis hin zu digitalen Methoden in unterschiedlichen (nicht nur) geisteswissenschaftlichen Disziplinen. Längst geht Annotation über die schon als analoge Kulturtechnik praktizierten Verfahren des individuellen Textstellen-Markierens und Verschlagwortens und/oder Kommentierens hinaus. So kann etwa die Alltagspraktik des ‚Social Tagging‘ in sozialen Medien als Diskursebene angesehen werden, auf der Interessen gebündelt bzw. bestimmte Inhalte ausgehandelt und Strukturierungen sowie Kontextualisierungen in Form von Hashtags und Timelines umgesetzt werden. Wissenschaftliche Annotationen umfassen eben-

falls meist mehr als individuelle Texterschließungsprozesse. Sie dienen – manuell oder (teil-) automatisiert generiert – als Grundlage für intersubjektive computergestützte Analysen und/oder werden schon kollaborativ entwickelt und umgesetzt, sodass im Zuge einer sukzessiven Tagset-Erstellung und Annotation deduktiv und induktiv kategorisiert und dabei Wissen ausgehandelt wird. Annotation kann insofern sowohl Untersuchungsgegenstand (kommunikative und kollektiv-kollaborative Praktik) als auch Methode (systematische Kategorisierung, Korpuserschließungs- und -analyseverfahren) digitaler (Diskurs-)Analyse sein. Im Folgenden steht nicht die Alltagspraktik, sondern die (diskurs-)linguistische methodische Praktik im Vordergrund. Diesbezüglich ist auch die Frage relevant, inwiefern digitale Annotation als Methode lediglich eine Übertragung ihrer analogen Form auf das Digitale darstellt, nach der Terminologie der ‚Digital Methods Initiative‘ also als lediglich „digitized“ anzusehen ist – in Abgrenzung von Methoden, die als „natively digital“, „methods of the medium“ bzw. einfach als „digital methods“ beschrieben werden (vgl. Rogers 2013: 1, 5, 19 und 24). Dieser Unterscheidungs-Ansatz ist im Zusammenhang mit digitalen Diskursen bereits gewinnbringend angewendet worden (vgl. Gredel 2018) und wird im vorliegenden Beitrag zur Reflexion genutzt.

Die wissenschaftlichen Anwendungsmöglichkeiten, individuell oder kollaborativ umgesetzt, umfassen manuelle sowie (teil-)automatisierte Workflows, mit denen unterschiedliche Aufgaben bearbeitet werden. Das Spektrum reicht von der Erschließung und Anreicherung verschiedener Daten(-formate) bzw. digitaler Korpora und Editionen bis hin zu komplexen Analyseprozessen und Automatisierungs- bzw. Machine Learning-Modellen. Diese Aufgaben können auf sehr unterschiedliche Gegenstände (sowohl genuin digitale als auch digitalisierte), Fragestellungen und Zielsetzungen vor verschiedenen fachlichen Hintergründen ausgerichtet sein.

Speziell in der Linguistik ist allerdings nach wie vor ein methodisches Verständnis etabliert, nach dem Annotation vor allem als automatisiertes Verfahren angesehen wird, das hauptsächlich als basale Routine zur Erschließung systemlinguistischer Kategorien in der Korpus- und Computerlinguistik genutzt wird. Diese Art von Annotation stellt eine Voraussetzung für korpuslinguistisch-statistische Auswertungen auf unterschiedlichen Komplexitätsebenen dar und ist auch für digitale Diskursanalysen als Zugriffsmöglichkeit auf sprachliche Indikatoren von

großer Bedeutung. Mit der ursprünglichen Kulturtechnik der Textrezeption und -kommentierung hat diese digital-transformierte Annotationsanwendung jedoch nur noch teilweise etwas zu tun. Automatisierte Tagging-Methoden kommen an ihre Grenzen, wenn Phänomene annotiert und auf dieser Basis erforscht werden sollen, die nicht ohne Weiteres und kontextunabhängig maschinell an der sprachlichen Oberfläche erkennbar sind, etwa aus bestimmten Forschungsperspektiven der Semantik oder Pragmatik. In diesen Bereichen gewinnt die manuelle Annotation in der Linguistik an Bedeutung, die hermeneutische Zuweisungsschritte und intersubjektiv-diskursive Wissensaushandlung digital ermöglicht. Damit ist allerdings oft die Frage verbunden, wie solche Ansätze für maschinelle Verfahren operationalisiert bzw. an Automatisierungen angeschlossen werden können. Insbesondere in kollaborativen Szenarien kann Annotation einerseits einen Metadiskurs bilden, andererseits zugleich aber auch eine Fortsetzung bzw. Ausweitung der Diskurse bedeuten, die in den Annotationsgegenständen enthalten sind. Hier liegt der Bezug zur Debatte über Deskription und Kritik in der Diskurslinguistik nahe (vgl. Reisigl/Warnke 2013: 7-36).

Annotation stellt also nicht einfach eine im Analogen geprägte Kulturtechnik dar, die in einer mimetischen Form ins Digitale überführt wurde, sondern eine Praxis, die sich unter den spezifischen Bedingungen des Digitalen weiterentwickelt hat. Aus Sicht der digitalen Diskurslinguistik ist sie zu verorten in verschiedenen Spannungsfeldern: zwischen alltagsweltlicher Rezeptions- und Kommunikations-Praktik und wissenschaftlicher Methode, zwischen hermeneutischer und automatisiert-quantifizierender Analyse sowie zwischen wissenschaftlichem Metadiskurs und Fortführung der untersuchten Diskurse (ob digitalisiert oder genuin digital). Diese Spannungsfelder und möglichen Verortungen werden im Folgenden beschrieben und diskutiert sowie anhand von Beispielen veranschaulicht. Fokussiert wird dabei das kollaborativ-diskursive Annotieren als wissenschaftliche Methode in der digitalen (Diskurs-) Linguistik.

## 2. Annotation zwischen Untersuchungsgegenstand und Methode in der digitalen (Diskurs-) Linguistik

Annotation lässt sich als Kulturtechnik bis in die Antike zurückverfolgen. Glossen und Scholien werden als frühe Vorläufer angesehen (vgl. Agosti et al. 2007: 4, Stricker 2009: 23). Die Entwicklung im Analogen hat bereits unterschiedliche Formen des Annotierens hervorgebracht, von der individuellen Selektion und Markierung von Inhalten (Textstellen, Bild- bzw. Objektausschnitten u.a., von anderen verfasste sowie eigene) und ihrer Erschließung bzw. Anreicherung bis hin zur systematisierten Verschlagwortung und weiterführenden Analyse oder Publikation von Kommentierungen – sowohl in unterschiedlichen Alltagskontexten als auch in verschiedenen wissenschaftlichen Disziplinen. Insofern stellen schon analoge Annotationen einerseits einen potenziellen Untersuchungsgegenstand für die Linguistik dar, andererseits aber auch eine methodische Zugriffsmöglichkeit.

### 2.1 Annotation als sozial-kommunikative Alltagspraxis

Im Digitalen hat sich das Anwendungsspektrum von Annotation auf verschiedenen Ebenen ausdifferenziert und erweitert, bedingt durch die hypermediale Einbindung und algorithmische Erfassbarkeit von Annotationen. In sozialen Medien stellt das ‚Social Tagging‘ – am bekanntesten in Form von Hashtags – eine typische Nutzungsform dar und ist insofern als alltagsweltliche, digitale Kommunikationspraktik sehr weit verbreitet. Dabei handelt es sich um Annotation in Form von individuellen Verschlagwortungen bzw. Indexierungen, denen zunächst keine festgelegten Vokabulare oder spezifischen bzw. expliziten Kategorisierungssysteme (Taxonomien oder Ontologien) zugrunde liegen. Es wird lediglich ein Operator – in Form des Raute-Zeichens # – von der jeweiligen Plattform technisch implementiert, der auch Sortierungen nach Hashtags in Timelines ermöglicht. Erst im Zuge der gemeinschaftlichen Nutzung entwickeln sich soziale Dynamiken und damit auch eine Vielfalt an Verwendungsweisen und Funktionalitäten (Überblick vgl. Zappavigna 2015: 274–291). Durch themen- bzw. interessenbezogene Aggregations- und Aushandlungsprozesse können sich dabei Ad-hoc-Öffentlichkeiten bzw. Communities konstituieren und sukzessive konventionalisierte Tag-

Vokabulare ausbilden. Hashtags werden so – per Timelines – zur Diskursstrukturierung und -organisation, aber auch zur Kontextualisierung bzw. Rekonfigurierung von Kontexten verwendet (vgl. zur Kontextualisierungstheorie: Gumperz 1982, Auer 1986, Müller 2015: 78, in Bezug auf Hashtags: Dang-Anh/Einspänner/Thimm 2013: 146). Konventionalisiert – im kommunikativ-funktionalen Sinne – sind außerdem Formen der Bewertung, Perspektivierung oder Explizierung bestimmter sprachlicher Handlungen oder rhetorischer Mittel sowie Formen der Metakommunikation durch Hashtags (vgl. z.B. Marx 2019, Zappavigna 2011: 803, Überblick auch in Dang-Anh/Einspänner/Thimm 2012: 84). Bezüge zwischen Hashtags und etwa Microblog-Posts können wechselseitig bestehen, also nicht nur in der Etikettierung von Posts durch Hashtags als Metadaten, sondern auch umgekehrt in der Bezugnahme von Tweets auf Tags (vgl. Gruber 2016: 6). Diese Vernetzungskomponente kommt zu der Funktion hinzu, die Hashtags als sichtbare Text-Bestandteile von Posts haben. Somit stehen sie „auf der Schwelle zwischen Text und Metatext“ und bilden ein Bindeglied zwischen „Alltagskommunikation und Computercode“ (Bernard 2018: 2).

Diese drei Faktoren sind ausschlaggebend dafür, dass Annotationen sowohl Gegenstand als auch Methode digitaler Diskursanalysen sein können: 1. die diskursive Einbindung und kollektiv-kollaborative Effekte durch die hypermediale Vernetzung, 2. die Eigenschaft der Schnittstelle zwischen Text und Metatext und 3. die Verbindung mit Algorithmen bzw. von Kommunikation und Code.

Als Untersuchungsgegenstand bieten sich bspw. Twitter-Hashtags zur Analyse verschiedener funktionaler Ebenen an, Dang-Anh et al. (2013: 138) unterscheiden die Operatorebene (technische Indexierungsfunktion), die Textebene (sprachliche Verschlagwortungsfunktion) und die Handlungsebene (Kontextualisierungs- und Diskursorganisationsfunktion), weitere Funktionen führt – wie erwähnt – Zappavigna (2015: 274–291) auf. Twitter-Hashtags können aber auch als Zugriffspunkt für die Datenerhebung und zur Gruppenbildung nach Themen oder auch Hashtag-Nutzern in diskurslinguistischen Korpusstudien dienen (vgl. Müller/Stegmeier 2016: 507–508).

In Lern- oder Wissensmanagement-Umgebungen (z.B. im bibliographischen Bereich) werden Annotationen ebenfalls unter dem Begriff ‚Social Tagging‘ (vgl. Müller-Prove 2008: 18) oder in diesem Kontext auch als ‚Collaborative Tagging‘ (vgl. Golder/Huberman 2006: 199)

untersucht. Auch hier werden individuell gewählte Tags innerhalb der Nutzergemeinschaft eines Tools vergeben. Die daraus entstehenden individuellen Tag-Clouds stellen Klassifikationen dar, die – in Abgrenzung zur Taxonomie – als „Folksonomy“ bezeichnet werden (vgl. Vander Wal 2005). Sie basieren nicht auf einer z.B. von Expert/inn/en vorgegebenen Ordnung oder Bedeutungszuweisung, sondern gehen aus einem „dezentralen, unkoordinierten, sozial-kumulativen Prozess“ (Müller-Prove 2008: 20) hervor. Graduelle Konventionalisierungs-, Aggregations- und Rückkopplungsprozesse (vgl. Müller-Prove 2008: 18) können z.B. durch Clustering-Verfahren sichtbar gemacht und in die Nutzergemeinschaft zurückgespiegelt, aber auch als Untersuchungsgegenstand genutzt werden.

## 2.2 Annotation als systematische Methode

Eine Analyseform, in der Annotationen zwar auch als Untersuchungsgegenstand dienen können, aber keine Praktik der Alltagskommunikation mehr sind, sondern die Funktion eines methodischen Instruments haben, stellen Crowdsourcing-Analysen dar, die in der Marktforschung entwickelt wurden. Bekanntestes Beispiel sind Amazons Mechanical Turks<sup>1</sup>. Das Basismodell dieser Verfahren besteht darin, dass Annotations-Aufgaben an Nicht-Expert/inn/en bzw. mehr oder weniger anonyme Crowd-Arbeiter/innen vergeben werden, die in ihrer großen Gesamtheit vor allem durchschnittliche potenzielle Käufer/innen oder Nutzer/innen von Angeboten repräsentieren sollen. Auch bei komplexeren Aufgaben wird davon ausgegangen, dass – analog zum Wikipedia-Prinzip – ein Kollektiv statt einzelner Expert/inn/en diese lösen kann und Fehler durch das Korrektiv der Vielzahl an Beteiligten aufgedeckt werden (vgl. Giles 2005).

Meist handelt es sich jedoch um einfachere Aufgaben, die die Zuweisung weniger Annotationskategorien zu bestimmten Inhalten umfassen, die ohne ein aufwändiges Training oder gar nachgewiesene Expertise bewältigt werden können. Die Ergebnisse werden aggregiert. Gemessen wird auch der Grad der Übereinstimmung der Annotationen von mehreren Crowd-Arbeiter/inne/n an derselben Aufgabe. Typische Anwen-

---

<sup>1</sup> <https://www.mturk.com/>

dungsfelder sind die Relevanzbewertung der Ergebnisse von Suchmaschinen-Anfragen oder Sentiment-Analysen, also im einfachsten Fall die Annotation positiver oder negativer Wahrnehmung von Inhalten bzw. Inhaltsbestandteilen. Spezifisch linguistische Anwendungsszenarien gibt es viele verschiedene, von der Annotation von Grammatikalität bis hin zur Erstellung von Wortbedeutungsinventaren (vgl. Juska-Bacher/Biemann/Quasthoff 2013 und Biemann 2012). Annotationen werden in Crowdsourcing-Modellen also meist als vorgegebene Kategoriensysteme, aber auch in Form von sukzessive gebildeten Tag-Inventaren eingesetzt. Insofern werden sie sowohl als methodische Operationalisierungen als auch als crowd-generierte Untersuchungsgegenstände genutzt. Dadurch können Rückschlüsse gezogen werden auf Merkmale der annotierten Inhalte (und z.B., wie diese durch mehr oder weniger bestimmte Nutzergruppen wahrgenommen werden), aber auch auf Eigenschaften von Annotator/inn/en (im Sinne von potenziellen Nutzer/inne/n), z.B. Einstellungen gegenüber den annotierten Inhalten.

Von Wissenschaftler/inn/en selbst durchgeführte Annotationen basieren in der Regel auf einem Kategorienschema, taxonomisch oder ontologisch, und dienen der systematischen Analyse und Interpretation von Untersuchungsgegenständen. In solchen Modellen können Annotationen im Sinne der Reflexion von Forschungsprozessen, Theoriebildungen und Operationalisierungen von Forschungsfragen selbst als Untersuchungsgegenstände betrachtet werden. In diesem Sinne können sie auch als kommunikative Praktiken der internen Wissenschaftskommunikation angesehen werden, wenn etwa über kollaborative Tools Aushandlungsprozesse in Annotationsprojekten stattfinden. Eine weitere Möglichkeit besteht darin, dass Annotationen publiziert und anderen Forschungsperspektiven zugänglich gemacht werden, sei es als Tagging oder – in einem weiteren Verständnis von Annotation – als erläuternd-interpretierende, frei formulierte textuelle Anreicherungen in der Art von Stellenkommentaren. Solche Mikropublikationen können bspw. im Rahmen von digitalen Editionen oder Korpuspublikationen Forschungsergebnisse darstellen und anderen wiederum als Untersuchungsgegenstand dienen (vgl. Bradley 2012, Rapp 2017: 255–262). Darüber hinaus können sie neben dem Text z.B. Literaturangaben oder Verlinkungen zu anderen digitalen Quellen in Repositorien enthalten und auf diese Weise eine Form inter- bzw. hypertextueller Kontextualisierung darstellen (vgl. Bender 2016b: 1–8, konkretes Beispiel in Alscher/Bender 2016).

In diesem Aufsatz steht jedoch – als Untersuchungsgegenstand – Annotieren als Methode der digitalen (Diskurs-) Linguistik im Mittelpunkt. Aus dieser Perspektive ist vor allem die Unterscheidung, die schon im Zusammenhang mit Crowdsourcing-Modellen beschrieben wurde, relevant – nämlich die zwischen dem Annotieren mit vorab fixierten Kategoriensystemen, die vor allem in der Computerlinguistik unter dem Begriff „Goldstandard“ gängig sind und hauptsächlich mit dem Ziel der Automatisierung entwickelt werden, einerseits und der sukzessiven Entwicklung von Schemata im Zuge des Annotationsprozesses andererseits. Im erstgenannten Ansatz wird das festgelegte Goldstandard-Schema als Metaebene für die automatisierte Auswertung ange-setzt. Im letztgenannten hingegen bleiben Annotationen als auszudifferenzierende Kategorien im Rahmen eines Analyse- und Aushandlungsprozesses in die Auseinandersetzung mit den Untersuchungsgegenständen involviert – auch in der Art, dass auf dieser Ebene nicht nur Metadiskurse geführt, sondern auch Diskurse fortgesetzt werden, die in den Untersuchungsgegenständen selbst enthalten sind – z.B. die Reflexion von Bedeutungen von Fachbegriffen und ihrem dynamischen Wandel (siehe Abschnitt 3.2, Beschreibung des Projekts zur Terminologiedynamik). Insofern sind Annotationen der letztgenannten, diskursiven Art des Forschungsprozesses an der Schnittstelle bzw. im Spannungsverhältnis zwischen Gegenstand und Methode zu verorten. Im folgenden Abschnitt wird diese in der Linguistik neue Methodenvariante des Annotierens im Hinblick auf ihr Innovationspotenzial reflektiert und diskutiert, insbesondere hinsichtlich ihrer Wechselwirkung zwischen hermeneutischer bzw. diskursiv ausgehandelter Interpretation bzw. Kategorisierung und automatisiert-algorithmischen Analysemöglichkeiten. Dies wird anhand von Beispielen veranschaulicht.

### 3. Annotation als digitale Methode zwischen Automatisierung und diskursiver Hermeneutik

In wissenschaftlich-methodischen Anwendungsszenarien von Annotation spielen die drei oben herausgearbeiteten Faktoren ebenfalls eine wichtige Rolle: Wissenschaftliche Annotationen sind Schnittstellen zwischen Text (bzw. Inhalt, Gegenstand) und Metatext, und sie sind in einen diskursiven Analyseprozess eingebunden – auch schon als analoge

Inhaltserschließungs- und Anreicherungsverfahren, wenn man bspw. an die Überlieferung und Weiterführung von Anmerkungsapparaten denkt. Durch die Digitalisierung sind diese Eigenschaften dahingehend transformiert worden, dass die Text-Metatext-Schnittstelle auch technisch-operational vorhanden ist und über die hypermediale Vernetzung kollektiv-kollaborative Dynamiken ermöglicht. Als dritter wichtiger Faktor ist im Digitalen die Erfassbarkeit durch Algorithmen, d.h. die erwähnte Verbindung von Text und Code, hinzugekommen.

Die zentrale Praktik des analogen Annotierens, also das Markieren und Kommentieren bzw. Verschlagworten von Textstellen (oder anderen Rezeptionsgegenständen) wurde durch diese Faktoren auch im wissenschaftlich-methodischen Bereich in mehrerlei Hinsicht erweitert. Hinzu kamen die Anbindung an mehr oder weniger komplexe, mittels Annotationstools modellierte und dadurch explizierte Kategoriensysteme, also taxonomische oder ontologische Tagsets, und die entsprechenden Annotationsrichtlinien. Als zusätzliche Dimensionen sind – neben der hypermedialen Vernetztheit – die (Durch-)Suchbarkeit von Annotationen sowie der mögliche Zugriff durch algorithmische Analysefunktionen und maschinelles Lernen anzusehen. Letzteres wird eingesetzt, um Algorithmen dahingehend zu trainieren, komplexe, bisher nur manuell annotierbare Kategorien zuweisen zu können. Solche Verfahren erweitern die bisherigen algorithmischen Möglichkeiten zur annotationsbasierten Analyse, die bislang – neben der Auszählung von manuellen Annotationen – vor allem in der Verknüpfung mit automatisierten Annotationen (wie z.B. Part-of-Speech-Tagging) und statistisch-korpuslinguistischen Methoden gesehen wurden, welche nach wie vor ebenfalls eine wichtige und gewinnbringende Analysemöglichkeit von Annotationen darstellen (vgl. Bubenhofer 2013: 128–129). Wissenschaftliches bzw. methodisches Annotieren findet dadurch in komplexen digitalen Lern- bzw. Forschungsumgebungen statt (vgl. Bender 2016: 263–276). In diesen interagieren Annotierende rezeptiv und produktiv nicht nur mit einem Annotationsgegenstand, z.B. einem einzelnen Text, sondern auch mit anderen Texten bzw. Quellen im jeweiligen Korpus oder extern, z.B. Sekundärliteratur, angebundenen Vokabularien oder Datenbanken. Darüber hinaus findet Rezeption und Produktion auf der Ebene der Tagsets und Guidelines statt. Auf all diesen Ebenen interagieren Annotierende außerdem mit anderen Annotator/inn/en in kollaborati-

ven Szenarien (vgl. Bender 2016a: 263–277, Bender 2016b, Bender/Kollatz/Rapp 2018, Alscher/Bender 2016).

Prinzipiell sind solche Interaktionsformen und Relationen in analogen Umgebungen zwar ebenfalls möglich. Grundlegend verändert werden die Bedingungen jedoch durch die direkte technische Verknüpfung und die entsprechende Skalierung, was die Datenmenge, die Kollaborationsmöglichkeiten und die algorithmische Verarbeitung angeht. Das Ziel der maschinellen Verarbeitbarkeit und der kollaborativen, computervermittelten Nutzung von Annotationen stellt veränderte Anforderungen an das Annotieren. Vor allem die Explizierung, Systematisierung und Operationalisierung von Fragestellungen, Hypothesen und Interpretationen ist ganz entscheidend. Konkret umgesetzt werden diese Aspekte 1.) durch die Festlegung der zu annotierenden Einheiten bzw. Segmentierungen relativ zu entsprechenden Kategorien, 2.) die trennscharfe Kategorienbildung sowie 3.) die präzise und eindeutige Formulierung von Annotationsrichtlinien. Auf diesen drei Ebenen finden in kollaborativen Annotationsprojekten entsprechende Aushandlungsprozesse statt.

Diese Aspekte bedingen in der (Diskurs-) Linguistik im Speziellen und in den Geisteswissenschaften im Allgemeinen verschiedene methodische Ausrichtungsmöglichkeiten. Als Extrempunkte des Spektrums können Verfahren mit einem geschlossenen Tagset bzw. ‚Goldstandard‘ einerseits und Ansätze mit ganz offenen Kategoriensystemen andererseits angesehen werden.

‚Goldstandard‘-Ansätze sind vor allem in der Computerlinguistik gängig. Dabei werden in einem distinktiven und exhaustiven Tagset Kategorien festgelegt – vor allem systemlinguistische Klassen, kategoriale (Wortarten, Phrasenkategorien) oder funktionale (Satzglieder, Abhängigkeitsrelationen, Koreferenzstrukturen) – und Zuweisungsrichtlinien (Guidelines) formuliert. Nach Probeannotationen wird das hierarchische Schema nicht mehr verändert. Es wird anschließend von mehreren Annotierenden unabhängig voneinander und ohne wechselseitige Abstimmung auf Untersuchungskorpora angewendet. Übereinstimmungen und Abweichungen, das so genannte Inter-Annotator-Agreement bzw. Inter-Coder-Agreement (Artstein/Poesio 2008: 555–596), werden geprüft. Meistens ist in diesem Modell das Trainieren von Tools das Ziel der Annotation. Es wird also die algorithmische Erfassung von Annotationskategorien und die Automatisierung der Annotationsprozesse angestrebt. Dadurch wird die Auszeichnung und Anreicherung großer Text-

korpora ermöglicht, die dann als getagte Daten vorgehalten und/oder als Mikro-Publikationen veröffentlicht werden können. Annotationen stellen dann einen Zwischenschritt in Forschungsprozessen und eine Basis für automatisierte Analyseverfahren dar (vgl. McCarty 2004: 254–271).

In diesem Ansatz findet zwar auch ein vorgeschalteter diskursiver Aushandlungsprozess statt, aber eben nur bis zu dem Punkt der Goldstandard-Fixierung. In so einem Vordiskurs z.B. zu einem Part-of-Speech-Tagset (PoS) muss bspw. in Auseinandersetzung mit verschiedenen linguistischen Ansätzen eine Wortarten-Klassifikation festgelegt werden. Dabei spielen, vor dem Hintergrund des Ziels der Automatisierung, aber auch computerlinguistisch-anwendungsorientierte Kriterien eine Rolle, die sich an der maschinell erfassbaren, sprachlichen Oberfläche orientieren. Nicht alle Kategorien sind daher in einem PoS-Tagset unbedingt nur nach einem bestimmten Wortklassen-Modell definiert. Im Stuttgart-Tübingen-Tagset (STTS) gibt es beispielsweise Kategorien wie ‚abgetrennter Verbzusatz‘, ‚Partikel bei Adjektiv oder Adverb‘ oder ‚Kompositions-Erstglied‘ (vgl. Schiller et al. 1999).

Solche systemlinguistischen Klassifizierungen sind als Basis-Annotation Standard in der Korpus- und Computerlinguistik und stellen das dar, was allgemein unter Annotation in diesen Disziplinen verstanden wird. In diesen linguistischen Forschungsrichtungen wurden über solche basalen Ebenen hinaus viele verschiedene theoretische Ansätze in Form von Annotationsschemata zu operationalisieren versucht (Überblick vgl. Ide 2017: 1–18), allerdings immer mit dem Ziel, einen Goldstandard für das Training von Machine-Learning-Anwendungen zu entwickeln:

„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).

Immer öfter wird allerdings auch das digitale, aber manuelle Annotieren von Phänomenen bzw. Merkmalen, die (noch) nicht algorithmisch erfassbar sind, in der Linguistik praktiziert, z.B. die Annotation von Sprechakten (vgl. Weisser 2015 und Weisser 2018). Neben der Pragmalinguistik gibt es aber auch mögliche Anwendungsfelder in der Semantik sowie der Sozio- und Diskurslinguistik (Beispiele siehe Abschnitt 3). In anderen geisteswissenschaftlichen Disziplinen gewinnt manuelles Annotieren in verschiedenen Formen und mit unterschiedlichen Zielen im

Hinblick auf algorithmische Analysemöglichkeiten ebenfalls an Bedeutung (vgl. Rapp 2017: 258–262).

Am entgegengesetzten Extrempunkt des oben erwähnten Spektrums werden offen konzipierte und sukzessive entwickelte Annotationsschemata bspw. aus literaturwissenschaftlicher Perspektive als Akt der interpretativen Aneignung von Texten angewendet. In solchen Forschungsprozess-Modellen annotieren mehrere Personen unabhängig voneinander auf der Basis eines vorläufigen Tagsets, wie z.B. Gius und Jacke narratologische Kategorien der Zeit (vgl. Gius/Jacke 2016). Die Ergebnisse werden mit der Absicht verglichen, das Annotationsschema zu präzisieren und auszudifferenzieren, nicht mit dem Ziel einer statistischen Analyse – zumindest nicht unbedingt und in erster Linie. Dabei spielt die Prüfung von Abweichungen zwar eine wichtige Rolle, es wird aber nicht unbedingt auf übereinstimmende Annotationen gezielt, sondern divergierende Annotationen werden für die Anpassung des Tagsets und der Guidelines genutzt – bis auf den Fall, dass eine Abweichung durch einen offensichtlichen Annotationsfehler zustande gekommen ist. In diesem Modell kann es durchaus vorkommen, dass eine Abweichung insofern zur Erkenntnis beiträgt, als dadurch eine deutungsoffene, „polyvalente“ (Gius/Jacke 2016: 8) Passage identifiziert werden kann, die verschiedene Lesarten erlaubt.

Der Unterschied der beiden Ansätze liegt also in der Zielrichtung des Annotationsprozesses und der Funktion des Annotierens im Forschungsprozess. Die Automatisierung von Annotation und die Analyse großer Korpora steht im korpus- und computerlinguistischen Ansatz im Vordergrund, die kollaborative Textanalyse und die Weiterentwicklung der Kategorienschemata, also auch die theoretische Ausdifferenzierung mit Blick auf die Operationalisierung bestimmter Forschungsfragen, ist im literaturwissenschaftlichen Ansatz das Wesentliche. Beide beinhalten jedoch als zentrale Komponente, dass mit Guidelines geregelt wird, nach welchen Kriterien die Analysekategorien auf Segmente der Untersuchungsgegenstände (Wörter, Sätze, Passagen usw.) bezogen werden – auch wenn sich die Qualitätssicherungsmaßnahmen unterscheiden.

Im vorliegenden Beitrag wird ein dritter Ansatz vorgestellt, der im Spektrum zwischen den beiden beschriebenen Extrempositionen zu verorten ist und einige Aspekte der beiden gezeigten vereint. Entwickelt wurde er vor allem für diskurslinguistische Fragestellungen und solche der linguistischen Pragmatik mit diskurslinguistischer Perspektive. Er

umfasst das kollaborative, manuelle Annotieren mit Überprüfung des Inter-Annotator-Agreements, aber auch die sukzessive Ausdifferenzierung und Weiterentwicklung von Tagsets. Allerdings sind Annotations-Schemata in diesem Verfahren nicht als dauerhaft offen für das Erfassen verschiedener Lesarten konzipiert, sondern auf letztlich möglichst trennscharfe Differenzierungen ausgerichtet. Es wird also eine Art „Goldstandard“ erarbeitet, der aber an einem späteren Punkt im Forschungsprozess angesetzt wird und eher eine Zieldimension als eine Startvoraussetzung darstellt.

Ein entscheidender Unterschied zum computerlinguistischen ‚Goldstandard‘-Ansatz liegt auch in der Ausrichtung der Qualitätssicherung in Form der Überprüfung des Inter-Annotator-Agreements. Im computerlinguistisch geprägten Ansatz werden statistische Signifikanz-Messungen bei Mehrfachannotationen angewendet, wenn z.B. von Standardsprecher/innen einer bestimmten Sprache als Annotator/innen ausgegangen sowie deren Sprachkompetenz als Objektivierungskriterium herangezogen wird. In dem Fall geht es also um die Verallgemeinerbarkeit von Kategorien, z.B. auf der Sprachsystemebene. Dieses Prinzip wurde oben schon im Kontext von Crowd-Sourcing-Ansätzen kurz erwähnt. Der hier vorgestellte dritte Ansatz bezieht sich jedoch auf Phänomene, die nicht einfach auf der Basis von Standardkenntnissen von nicht-geschulten Annotierenden erfasst werden können. Spezifische Expertise auf Seiten der Annotierenden ist in solchen Projekten erforderlich, d.h. die Annotation muss entweder von Expert/innen durchgeführt werden oder es müssen Annotator/innen entsprechend geschult werden. Dieses expertise-geleitete und zunächst Phänomene explorierende Annotationsverfahren bedingt auch, dass durch die Überprüfung des Inter-Annotator-Agreements Zweifelsfälle identifiziert, entschieden und die Kategorien-Definitionen und Annotations-Guidelines erst entsprechend präzisiert werden können, um die fachspezifische Komplexität im Tagset distinktiv erfassen zu können. Fachspezifische Komplexität muss dabei nicht unbedingt bedeuten, dass ein besonders feingranulares bzw. ausdifferenziertes Kategoriensystem erstellt werden muss – wenige Kategorien, deren Definition, Abgrenzung und Zuordnung zu Segmenten aber schwierig ist und die sich nicht durch immer detailliertere Subkategorien erfassen lassen, stellen ebenfalls eine Herausforderung dar.

Gerade Phänomene, die aus pragmalinguistischer Sicht – mit diskurslinguistischer Perspektive – interessant sind, müssen über Zuweisungs-

richtlinien (Guidelines) detailliert definiert werden, um die Operationalisierung in Form von Schemata trennscharf umzusetzen. Eine generelle Herausforderung bei der Annotation pragmlinguistischer Phänomene besteht darin, dass inferenzielle kommunikative Verfahren erfasst werden müssen, also Äußerungen, die vor dem Hintergrund des Kontextwissens der Kommunizierenden gemacht werden und Ausgangspunkte für komplexe Schluss- bzw. Interpretationsprozesse sind. In diesem Bereich muss über die Indikatorik an der sprachlichen Oberfläche und den explizit bekannten Ko- und Kontext auch Implizites erschlossen werden. Eine weitgehend automatisierte Annotation, wie sie im Bereich der Systemlinguistik mittlerweile Standard ist, ist auf dieser Ebene daher nicht ohne Weiteres möglich. Zwar wird versucht, bestimmte pragmatische Diskursmarker automatisch zu erfassen – wie z.B. Höflichkeitsmarker, Hedging oder Interjektionen, meist in gesprochener Sprache (vgl. Aijmer 2015). Aber komplexere Inferenzen müssen zunächst hermeneutisch und manuell annotierend erschlossen werden. Daher werden manuelle Annotationsverfahren in diesem Zusammenhang weiterentwickelt (Archer/Culpeper/Davies 2008: 613). Bspw. gibt es in der Sprechakttheoretischen Annotation schon fortgeschrittenere und feingranularere Ansätze (vgl. Weisser 2015, 2018), auch mit Blick auf Automatisierungsmöglichkeiten.

Eine weitere Problematik lässt sich auf der Ebene pragmlinguistischer Annotationen zeigen: die Problematik der Segmentierung bzw. der Wahl der Einheiten, die jeweils annotiert werden. Man kann z.B. einzelne Sätze als relevante Äußerungsgröße setzen, wie es in der Sprechakttheorie bspw. üblich ist. Oder man taggt bestimmte Indikatoren, bspw. Lexeme, um deren Frequenz zu messen, wenn es der Beantwortung der Fragestellung dient. Oft legen Operationalisierungen von theoretischen Ansätzen aber nahe, ganze Passagen bzw. Zonen zu annotieren. Mit Blick auf die späteren Analyseschritte muss dieser Aspekt entschieden und vor dem Hintergrund der jeweiligen theoretischen Ansätze und Forschungsfragen begründet werden. Beim Annotieren müssen also auf mehreren Ebenen (Segmentierung der annotierten Abschnitte, Benennung von Tags bzw. Bildung von Kategorien, Strukturierung in Schemata und Formulierung von Guidelines) präzise und trennscharfe Differenzierungen ausgehandelt und festgelegt werden. Dies wirkt sich vor der Folie des Digitalen umso stärker aus, weil in diesen Szenarien auch mit der Perspektive der maschinellen Analysierbarkeit annotiert wird.

Diese Aspekte werden im Folgenden anhand von drei Beispiel-Annotationsprojekten veranschaulicht. Die Projekte können im vorliegenden Rahmen nicht ausführlich beschrieben werden. Dies ist aber auch nicht das Ziel der folgenden Ausführungen. Vielmehr soll das Spektrum der Formen des kollaborativen Annotierens als Methode der digitalen (Diskurs-) Linguistik veranschaulicht und ihr Potenzial aufgezeigt werden. Dies erfolgt anhand einer kurzen Darlegung der wesentlichen Punkte zu den Untersuchungsgegenständen und Forschungsfragen der Projekte sowie der aus annotationsmethodischer Perspektive relevanten Aspekte.

### 3.1 Beispiel-Projekt 1: Wissenschaftliche Textroutinen

Am ersten Projekt lässt sich der beschriebene Ansatz von der Segmentierungsentscheidung über die kollaborative Tagset-Entwicklung bis hin zu Automatisierungsansätzen zeigen. Gegenstand des Projekts sind wissenschaftliche Textroutinen in 65 online publizierten Dissertationen aus den 13 verschiedenen Fachrichtungen der TU Darmstadt. Im Rahmen einer Pilot-Annotationsstudie wurden die Einleitungskapitel untersucht. Wir verwenden auch den Begriff „heuristische Textpraktiken“ (Bender/Müller 2020). Darunter verstehen wir Formulierungsverfahren, mit denen in institutionell verankerten Routinen neue, in der Aushandlung befindliche Aspekte an ausgehandeltes Wissen angeschlossen werden. Dazu gehören z.B. ‚die Relevanz eines Forschungsthemas markieren‘, ‚einen Begriff definieren‘ oder ‚eine Aussage argumentativ stützen‘. Ziel ist es herauszufinden, in welchem Ausmaß, in welcher Verteilung und mit welcher Kombinatorik solche Textpraktiken in unterschiedlichen wissenschaftlichen Disziplinen vorkommen und welche Effekte fachkulturelle Kontexte auf heuristische Textpraktiken haben.

Dazu ist es notwendig, bei der Operationalisierung unseres theoretischen Ansatzes zu heuristischen wissenschaftlichen Textpraktiken in Form eines sukzessive ausdifferenzierten Kategorienschemas eine Quantifizierungsmöglichkeit zu entwickeln, die auch das Kriterium der Vergleichbarkeit erfüllt. Wir haben uns dafür entschieden, grundsätzlich einzelne Sätze als Annotationseinheiten zuzuweisen, legen also jeweils die primäre Funktion eines Satzes vor dem Hintergrund unseres taxonomischen Schemas fest. Teilweise oder überlappende Annotationen

schließen wir aus, um eine konsistente, quantitative Auswertbarkeit und Vergleichbarkeit zu gewährleisten. Dies bedeutet, dass eine trennscharfe Definition und Abgrenzung von Kategorien erforderlich ist, die sich auf einzelne Sätze als Segmente der Redekette anwenden lässt. Anders als bei der Kategorisierung systemlinguistischer Einheiten ist es bei textpragmatischen Kategorien nicht immer einfach zu bestimmen, wo die Grenzen eines zu annotierenden Segments liegen, bspw. bei komplexen Argumentationen oder Definitionen, die über Koreferenzketten aufgebaut werden. Insofern muss man bei einer Operationalisierung über das Tagging einzelner Sätze eindeutige Zuweiskriterien erstellen, die auch eine Kategorisierung des jeweiligen Satzes als Vertreter genau einer textpragmatischen Kategorie erlaubt, wenn er auch als Teil einer satzübergreifenden Praktik anzusehen ist.

Betrachten wir als Beispiel einen Satz, der die heuristische Funktion ‚Relevanzmarkierung durch ungerichtete Diskursreferenzierung‘ hat. In dieser Kategorie werden Sätze erfasst, mit denen auf eine Klasse bzw. Gruppe an vorgängigen Forschungsäußerungen, die nicht genau spezifiziert oder definiert ist, referenziert wird, bspw. Relevanzmarkierung über die Existenz, Dauer und Publikationsdichte des entsprechenden Forschungsdiskurses realisiert wird (vgl. Bender/Müller 2020: 24). Als exemplarischer Beleg lässt sich der folgende Satz aus der Dissertation von Braun (2016: 3) anführen: „In der über vierzigjährigen internationalen Forschung zum Prüferwechsel sind weit über 100 Studien zu Gründen und Auswirkungen von Prüferwechseln zu verzeichnen.“

Im Umfeld dieses Satzes, in dem die Relevanz des Forschungsthemas durch eine Diskursreferenzierung ohne Nennung konkreter Autoren markiert wird, können weitere Sätze mit dieser Funktion oder etwa mit der Funktion der ‚Relevanzmarkierung durch gerichtete Diskursreferenzierung‘ mit Autorennennung stehen und eine ganze Diskursreferenzierungs-Zone bilden, was bei diesem Beispiel der Fall ist. Es können aber auch Sätze im Umfeld stehen, die über Sachverhaltsbeschreibungen Relevanz markieren und mit dem Beispielsatz zusammen allgemeiner eine Relevanzmarkierungszone bilden (hier im erweiterten Umfeld ebenfalls zutreffend). Auch andere Umfelder sind natürlich möglich. Solche komplexeren Einheiten erschließen wir aus den annotierten Einzelsätzen in ihrer linearen Abfolge. Für die Klassifizierung ist allerdings oft Ko- bzw. Kontextwissen Voraussetzung, das über entsprechende Kriterien in den Annotationsrichtlinien regelgeleitet und gestützt durch die entsprechen-

de Expertise der Annotator/inn/en in die Kategorienzuordnung einbezogen werden muss. Auf der Basis der Setzung der Segmentierungseinheiten werden also ein distinktives und exhaustives Kategoriensystem sowie die entsprechenden Zuweisungsrichtlinien in unserem Ansatz kollaborativ ausgehandelt – und zwar zunächst deduktiv auf der Basis der einbezogenen theoretischen Ansätze und dann induktiv durch die datengeleitete Ausdifferenzierung der vorläufig erstellten Kategorien. Das Kategoriensystem und seine Beschreibung mit Definitionen der verschiedenen Kategorien und Klassifikationsrichtlinien stellt – als ausdifferenzierte Typologie heuristischer Textroutinen in wissenschaftlichen Diskursen – bereits eine wichtige Ergebnisdimension des Projekts dar.

Darüber hinaus werden durch die Analyse der Verteilung und Reihenfolge der annotierten Sätze Muster erkennbar, die bspw. eine Typisierung von Einleitungen ermöglichen. So lassen sich z.B. Einleitungen, die tendenziell stärker Diskursreferenzierung zur Relevanzmarkierung einsetzen, von solchen unterscheiden, die Relevanzmarkierung eher durch Sachverhaltsbeschreibungen umsetzen. Solche Muster können hinsichtlich ihrer Disziplinspezifik untersucht werden. Es kann auf diese Weise etwa gezeigt werden, dass (bzw. – differenzierter betrachtet – inwiefern) in geistes-, sozial- und kulturwissenschaftlichen Arbeitern mehr Diskursreferenzierung betrieben wird und in natur- und ingenieurwissenschaftlichen mehr Sachverhaltsbeschreibung. Die folgenden Visualisierungen einzelner, pro Satz farblich annotierter Einleitungen, zeigen solche Verhältnisse auf (Abb. 1). Die Linearität der Satzabfolge wird hier von links nach rechts abgebildet, die Länge der Einleitungen ist ebenfalls repräsentiert. Die dunkelgrünen Zonen stehen für Sachverhaltsbeschreibungen, die hellgrünen für Formen der Diskursreferenzierung, die Blautöne für verschiedene Formen der Zielformulierung und die magentafarbenen für Thesenformulierungen. Bei der ersten Abbildung mit einem hohen Anteil an Diskursreferenzierung handelt es sich um eine Einleitung aus dem Fachbereich ‚Gesellschafts- und Geschichtswissenschaften‘, bei der zweiten, die nur Sachverhaltsbeschreibungen und Zielformulierungen aufweist, um eine aus dem Fachbereich ‚Mathematik‘ (an der TU Darmstadt den Naturwissenschaften zugeordnet).



Abb. 1: Visualisierung der Abfolge heuristischer Textpraktiken in einer Einleitung aus dem Fachbereich ‚Gesellschafts- und Geschichtswissenschaften‘ (oben) und Mathematik (unten)

Eine weiterführende Perspektive des Projekts ist die der (halb-) automatischen Klassifizierung bzgl. des Textroutinen-Schemas und die dadurch mögliche Untersuchung größerer Datenmengen. Erste Tests haben gezeigt, dass unsere Kategorien trotz bisher relativ geringer Datenmenge im Pilotkorpus robust genug für die Anwendung bestimmter Deep-Learning-Verfahren sind. Zumindest ergab eine Pilotstudie sehr vielversprechende Ergebnisse (vgl. Becker/Bender/Müller: in Vorb.). Dies zeigt, dass die Anwendung algorithmischer Methoden auf Annotationen, die auf der Erkennung und Klassifizierung von teilweise impliziten bzw. inferierten, kontextabhängigen Aspekten beruhen, eine vielversprechende Forschungsrichtung in der Pragma- und Diskurslinguistik darstellt.

### 3.2 Beispiel-Projekt 2: Terminologiedynamik in Poetiken

Das zweite Projekt zielt sowohl bzgl. der Segmentierung als auch der kollaborativen Kategorienentwicklung sowie der quantifizierenden Analyse in eine andere Richtung. Es baut auf dem abgeschlossenen BMBF-Projekt ePoetics auf, in dem 20 Poetiken, sprach- und literaturtheoretische Schriften aus den Jahren 1770–1960, digital erschlossen und im Deutschen Text Archiv (DTA) publiziert wurden (als TEI-konformes XML mit zusätzlicher HTML-Lesefassung und synoptischer Faksimile-Ansicht, vgl. Alscher/Bender/Rapp/Richter et al. 2016<sup>2</sup>). Ein Schwerpunkt dabei war die Annotation des theoretischen Begriffs ‚Metapher‘ und die Analyse seiner Entwicklung. Diese Forschungsrichtung – wir nennen sie Terminologiedynamik (vgl. Müller/Mell 2020) – verfolgen wir aus (diskurs-)linguistischer Perspektive auf Wissenschaftssprache weiter. Bzgl. der Segmentierung haben wir uns dafür entschieden, nur die

---

<sup>2</sup> <http://www.deustextarchiv.de/search/metadata?corpus=epoetics>

jeweiligen Fachbegriffe zu annotieren, nicht etwa zusätzlich Kontext-Segmente zu selektieren, in denen Definitionen, ein spezifisches Begriffsverständnis, verschiedene Explikationsaspekte usw. explizit gemacht werden. Stattdessen sind diese Aspekte in die Kategorienbildung als Unterkategorien oder Guideline-Kriterien eingeflossen (vgl. Alscher/Bender 2016). So wurde z.B. annotiert, ob der ‚Metapher‘-Begriff von dem jeweiligen Poetiken-Autor nach Quintilian als verkürzte Vergleichung oder nach Aristoteles als Übertragung definiert wird und ob demnach ‚Metapher‘ als Ober-, Unter- oder Parallelkategorie von ‚Vergleichung‘, ‚Metonymie‘ usw. angesehen wird (Abb. 2).

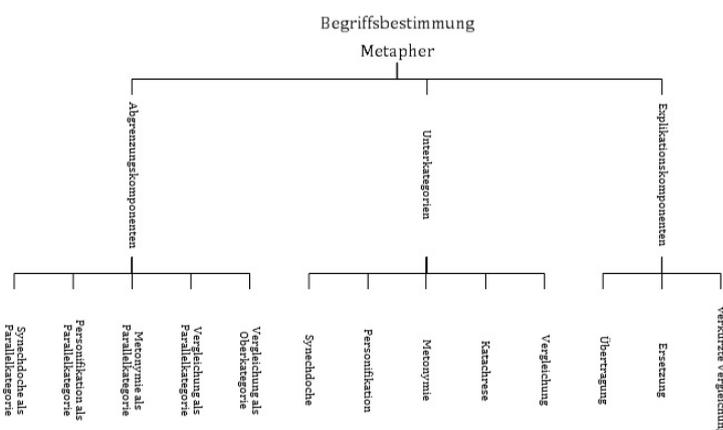


Abb. 2: Annotationsschema zum Terminus ‚Metapher‘ in Poetiken (Alscher/Bender 2016)

Zusätzlich wurden Verweisungsstrukturen ausgezeichnet, um zu unterscheiden, ob es sich bei annotierten Stellen um eigene Gedankengänge des Poetiken-Autors handelt oder um zitierte oder paraphrasierte Überlegungen anderer, denen sich der Autor anschließt, die evtl. verworfen werden oder einfach nur als Überblick dargestellt werden, wie hier (Abb. 3):

Zumpt (latein. Grammatik) nennt die Metapher ein zusammengezogenes Gleichnis; Wackernagel eine abgefürzte Vergleichung; Gottschall eine konzentrierte Vergleichung; Vischer die Herbeiziehung einer Erscheinung aus einer andern Sphäre; Max Müller die Übertragung des Namens eines Gegenstands auf andre Gegenstände.

Abb. 3: Verweisungen auf Definitionen des Begriffs ‚Metapher‘ bei verschiedenen Poetiken-Autoren in Beyers Poetik (Beyer 1882: 157)

Es handelt sich also um ein Korpus, das nicht aus präskriptiven Regelpoetiken besteht, sondern in dem die diskursive Aushandlung sprach- und literaturtheoretischer Konzepte, bspw. der Metapherntheorie, verfolgt werden kann, etwa die Orientierung an ‚Schulen‘ wie der nach Quintilian vs. der nach Aristoteles. Unser Annotationsansatz ermöglicht die Erschließung dieser Theorieentwicklung und das Nachvollziehen von Dynamiken bis hin zu den moderneren Metapherntheorien, bspw. kognitiven Ansätzen. Ähnliche Ansätze, in denen bestimmte Varianten des Begriffsverständnisses annotiert werden, um Terminologieentwicklung zu erforschen, wenden wir in unserer Forschungsgruppe in verschiedenen Projekten auch auf andere Untersuchungsgegenstände an, z.B. auf die Entwicklung des Terminus *Regime* in politikwissenschaftlichen Zeitschriften zum Thema internationale Beziehungen (vgl. Steffek/Müller/Behr: 2020).

### 3.3 Beispiel-Projekt 3: Empathie in der Wissenschaftskommunikation

Das dritte Beispiel-Projekt befindet sich noch in der Entwicklung, eignet sich aber, um anhand einer ersten Konzeption eine weitere Form der kollaborativen Annotation von Aspekten, die Ko- bzw. Kontext- und Expertenwissen voraussetzen, zu beschreiben. Es zielt auf die Erfassung von komplexen sprachlichen Phänomenen durch eine Mehrebenen-Annotation und die Korrelation der verschiedenen Schichten in der Analyse. Ausgangspunkt der Annotationsstudie ist ein Forschungsprojekt zum Thema ‚Empathie in der Wissenschaftskommunikation‘ (Bender/Janich 2020). Untersuchungsgegenstand sind Wissenschaftsblogs. In einer Pilotstudie wurde ein Auszug aus der Themensektion *KlimaLounge* des Portals *ScieLogs – Tagebücher der Wissenschaft*, betrieben vom Spektrum Verlag, untersucht. In dem Beispiel-Auszug folgt auf einen kritischen Blog-Beitrag über eine Expertenanhörung in einem Landtagsausschuss zum Thema Erderwärmung eine Debatte, die 112 Einträge umfasst und schwerpunktmäßig im Juni 2014 stattfand, sich aber bis ins Folgejahr fortsetzte.<sup>3</sup> Im Mittelpunkt der Diskussion stehen physikalische Fragen. Eng verknüpft mit der Argumentation sind jedoch auch Selbst- und

---

<sup>3</sup> <http://www.scilog.de/klimalounge/der-anti-treibhauseffekt-herrn-ermecke/>

Fremdzuschreibungen von Akteursrollen und dem jeweiligen Experten-/Laienstatus. Hinzu kommen metakommunikative Aushandlungsprozesse darüber, inwiefern die Diskussion ‚gut‘ geführt wird. Dabei spielt neben Faktoren wie der Sachlichkeit und Fachlichkeit auch der Aspekt der Empathie eine wichtige Rolle (vgl. auch Janich 2020). Diese Konstellation im Rahmen der Wissenschaftskommunikation in Blog-Kommentarbereichen ermöglicht die Analyse von Sprachhandlungen mit Empathie-Bezug relativ zu bestimmten sozialen Rollen in einem Experten-Laien-Spektrum. Unter Empathie wird dabei eine Komponente bzw. Möglichkeitsbedingung von Verständigungs- und Verstehensprozessen (vgl. Hermanns 2007: 145) im Rahmen der kommunikativen Koordination (vgl. Feilke 1994: 23) verstanden, wobei die Einnahme der Perspektive des/der Anderen und das Hineinversetzen in den entsprechenden Verweisungsraum als zentrale kognitive (ggf. auch emotive) Operation angesehen wird. Verbalisierungen, die Indikatoren für solche Operationen, ihre Anbahnung oder Metakommunikation darüber darstellen, sind Ziel des Annotationsprojektes. Weiter ausdifferenziert wird bspw. bzgl. des Immersionsgrades der Perspektiveneinnahme. Da die vorliegenden Ausführungen jedoch nur als Entwurf eines Beispielszenarios für die Anwendung der Annotationsmethode dienen sollen, wird auf die differenzierteren theoretischen Hintergründe zum Empathie-Begriff hier nicht vertiefend eingegangen (vgl. diesbzgl. Überblick aus linguistischer Perspektive: Bender/Janich 2020, aus philosophischer Sicht: Breyer 2013: 17–44). Im Folgenden wird ansatzweise das Annotationskonzept beschrieben, das auf die Erfassung der oben beschriebenen Empathie-Aspekte im Bezug zur sozialen Rollenkonstellation in der Wissenschaftskommunikation auf Blogs zielt. Der erste Entwurf dieses Annotationskonzepts lässt sich an einem Ausschnitt darlegen, der einen einzelnen Diskussions-Beitrag umfasst.

Kommentator 1 (am 25. Juni 2014 um 14:21 Uhr als Reaktion auf eine Blogautor-Antwort vom 24. Juni 2014, 17:49 Uhr):

Sehr geehrter Herr [Klimaforscher/Blogautor], Eine Reaktion wie diese bringt Ihnen genau den Gegenwind über den Sie sich (mittelbar – siehe Ihren Artikel) beklagen. Das haben Sie doch eigentlich gar nicht nötig. Um die Akzeptanz zu erreichen, die Sie gern hätten, wäre es meiner Meinung nach sinnvoll, wenn Sie ein wenig über Ihren Schatten springen und der Allgemeinheit mehr Zuwendung entgegen bringen würden. Zum Beispiel wäre eine In-

formation darüber hilfreich - und würde sicher einiges an Missverständnissen ausräumen - ob und, wenn ja, warum ich im Irrtum bin. [...] Schon allein das würde, aus meiner naiven Nichtfachmensch-Sicht [sic] heraus, den Treibhauseffekt hinreichend erklären. [...] Ich würde Sie bitten, mir (und eventuell auch dem einen oder anderen interessierten Leser) zu erklären, warum Sie das (offensichtlich) für falsch halten. Vielleicht könnten Sie ergänzend diverse Quellen angeben [...] Vielen Dank.

Zunächst wird der gesamte Beitrag mit Metadaten ausgezeichnet, die ihn einem Sprecher zuordnen, Zeit- und Datums-Angaben enthalten und die Interaktionseinbettung erfassen, also die Verortung in der Struktur des Kommentarbereiches und die Bezugnahme auf Voräußerungen, in dem Fall z.B. als Reaktion auf eine bestimmte Antwort des Blogautors.

Innerhalb des Beitrags als annotierte Zone werden die eigentlich kategorisierenden Tags vergeben – und zwar bezogen auf Rollenzuschreibungen, hier z.B. „aus meiner naiven Nichtfachmensch-Sicht [sic] heraus“, und auf Sprachhandlungen mit Bezug zu Empathie-Aspekten.

Letztere umfassen verschiedene Kategorien. Naheliegend ist eine Kategorie, die Verbalisierungen bzw. Markierungen von Perspektivwechseln erfasst, hier z.B. „Das haben Sie doch eigentlich gar nicht nötig.“, wobei diesbzgl. weiter in Unterkategorien differenziert werden müsste, inwieweit graduell tatsächlich ein Perspektivwechsel veranschlagt werden kann oder ob es sich eher um eine Projektion oder gar eine Floskel handelt, die eigentlich nichts mit wirklicher Empathie zu tun hat. Dies müsste anhand der theoretischen Grundlagen zum Phänomen Empathie aus linguistischer Sicht (vgl. Bender/Janich 2020) weiter ausgeführt werden.

Eine zweite Kategorie erfasst Äußerungen, die Verständnisbemühungen bzw. Verständnissicherungen darstellen, hier z.B. die retrospektive Äußerung „Ich würde Sie bitten, mir (und eventuell auch dem einen oder anderen interessierten Leser) zu erklären, warum Sie das (offensichtlich) für falsch halten.“

Einen dritten Klassifikationsbereich bildet die – mehr oder weniger explizite – Metakommunikation über Empathie. Ein Beispiel für das metakommunikative Einfordern von Empathie ist im Auszug oben ebenfalls enthalten: „Um die Akzeptanz zu erreichen, die Sie gern hätten, wäre es meiner Meinung nach sinnvoll, wenn Sie ein wenig über Ihren Schatten springen und der Allgemeinheit mehr Zuwendung entgegen

bringen würden.“ Auch dieser Bereich ist in Unterkategorien auszudifferenzieren, bspw. ist auch die Unterscheidung zwischen Empathieerwartung, Empathieerzeugung und Empathiebereitschaft zu berücksichtigen.

Im Vergleich zu den beiden vorher beschriebenen Projekten ist dieser Ansatz also sowohl im Hinblick auf die Segmentierung als auch mit Blick auf die Auswertungsmöglichkeiten anders konzipiert. Es wird eine strukturelle Segmentierungsebene angesetzt, auf der ganze Sprecher/innen-Beiträge und deren interaktionale Einbindung erfasst werden, und es müssen Segmentierungsentscheidungen auf den einzelnen kategorialen Ebenen getroffen werden, in Abhängigkeit von der Indikatorik der zu erfassenden Phänomene. Bei der Auswertung stellt die Korrelierung der verschiedenen Ebenen den besonderen Mehrwert über die quantifizierende Häufigkeitsanalyse hinaus dar. Bspw. besteht die Möglichkeit, ausgeben zu lassen, inwiefern Zusammenhänge zwischen Rollenindikatoren, Empathie indizierenden Sprachhandlungen und der metakommunikativen Einforderung von Empathie bestehen. Im Beispiel oben liegt z.B. die Selbstzuschreibung der Laienrolle in Kombination mit Formen der zumindest empathieanbahnenden Verständnissicherung und metakommunikativen Empathieeinforderung gegenüber einem Experten vor. Solche Muster sind durch den beschriebenen Mehrebenen-Annotationsansatz erfassbar. Aber auch hier stellen bereits das sukzessive entwickelte Kategoriensystem und die entsprechenden Annotationsrichtlinien wichtige Ergebnisdimensionen dar, die Erkenntnisgewinn über sprachliche Handlungen mit Empathie-Bezug und ihre Indikatoren ermöglichen.

#### 4. Annotation zwischen wissenschaftlicher Metadiskursivität und diskursiver Involviertheit

Kollaborative Annotation weist als Praktik auf mehreren Ebenen Diskursivität auf. Diese These bzw. Feststellung ist hier wohlgermerkt nicht hauptsächlich auf die kommunikative Alltagspraktik des Taggens bezogen, die im vorliegenden Aufsatz ja bereits als Diskurskomponente in sozialen Medien und Untersuchungsgegenstand der digitalen Diskurslinguistik beschrieben wurde. Im Folgenden soll vielmehr die Diskursivität der wissenschaftlichen Methode des digitalen, kollaborativen Annotierens diskutiert werden.

Zunächst einmal können durch Annotation Diskurse bzw. diskursive Muster in digitalen Korpora sichtbar gemacht und für die Analyse aufbereitet oder schon durch das Annotieren selbst in Form der systematischen Kategorisierung analysiert werden. Beim Annotieren als Methode, wie sie hier vor allem im Rahmen der korpuslinguistischen Forschung beschrieben wurde, werden bestimmte sprachliche Einheiten aus Diskursausschnitten (im Sinne von als Korpora aufbereiteten Texten) selektiert, markiert und regelgeleitet Kategorien zugewiesen. Annotationen bilden Schnittstellen zwischen Text-Segmenten und Sinn-Systemen. Durch die kategorisierende Auszeichnung und Zuweisung zu einem textübergreifenden Annotationsschema werden transtextuelle Bezüge und musterhafte Strukturen bzw. diskursive Ordnungen sichtbar gemacht und analysierbar. Insofern liegt das Diskurslinguistische der Methode des Annotierens auf der Hand. Sie ist anwendbar aus verschiedenen diskurslinguistischen Perspektiven und auf verschiedenen für die Diskurslinguistik relevanten Untersuchungsebenen. Bspw. lassen sich auf allen Ebenen des DIMEAN<sup>4</sup>-Modells (vgl. Spitzmüller/Warnke 2011: 201) jeweils Annotations-Layer und -Tagsets je nach Erkenntnisinteresse modellieren – von der intratextuellen Ebene, die lexikogrammatistische, propositionsorientierte und textstrukturelle Aspekte abdeckt, über die akteursbezogene Ebene (Interaktionsrollen, Diskurspositionen, Medialität) bis hin zur transtextuellen Ebene (z.B. Frames, Topoi, diskurssemantische Grundfiguren usw.).

Indikatoren können – durch Annotationsschemata und -richtlinien systematisiert – erfasst, diskurslinguistisch relevanten Phänomenen zugeordnet und auch für quantifizierende und automatisierte Analyseverfahren zugänglich gemacht werden. Letztlich zielen diskurslinguistische Ansätze auf die Erforschung der interessegeleiteten Aushandlungsprozesse im Zuge der Konstruktion von Wissen in Diskursen (vgl. Spitzmüller/Warnke 2011: 41–48). Dieses diskursiv konstruierte Wissen kann ebenfalls durch Annotation erschlossen werden. Auch in den hier beschriebenen Beispiel-Projekten sind Wissenskonstrukte und die damit verbundenen agonalen Zentren (vgl. Felder 2013: 13–28) Gegenstand der Untersuchung bzw. Annotation: Im Poetiken-Projekt (vgl. Abschnitt 3.2) wird die Aushandlung der definitiven Merkmale des Metaphernbegriffs bzw. dessen terminologische Einordnung und Dynamik

---

4 Diskurslinguistische Mehr-Ebenen-Analyse

erfasst. Im Empathie-Projekt (vgl. Abschnitt 3.3) geht es unter anderem um die metakommunikativ-diskursive Konstruktion von Kriterien guter Kommunikation im Experten-Laien-Spektrum zu wissenschaftlichen Themen. Im Projekt zu wissenschaftlichen Textroutinen (vgl. Abschnitt 3.1) steht nicht das konstruierte Wissen selbst im Vordergrund, sondern die diskursiven Konstruktionsprozesse in Form von heuristischen Praktiken, wodurch der diskurslinguistische Bezug auf einer anderen Ebene liegt.

In diesem Abschnitt stehen jedoch nicht die Diskurs-Aspekte der untersuchten Gegenstände im Vordergrund, sondern die Diskursivität des Annotationsprozesses selbst wird reflektiert. Es wird also die diskursive Wissensaushandlung in Annotationsprojekten in den Blick genommen, wobei hier Überschneidungen mit denen der untersuchten Diskurse auftreten können, was im Weiteren gezeigt wird. Allerdings ist auf der Ebene der Annotation bei Aushandlungsprozessen nicht die Agonalität im Sinne eines semantischen Kampfes und der Durchsetzung bestimmter Bedeutungen das zentrale diskursive Phänomen. Vielmehr wird eine kollaborative Form der Aushandlung von Kategorien vollzogen, die Teil eines Systems sind, durch das für die Fragestellung bzw. das Erkenntnisinteresse des Forschungsprojekts relevante Aspekte gemeinsam operationalisiert werden. Dafür muss in kollaborativen Projekten eine Vielfalt an Fragen im diskursiven Austausch und sukzessive im Annotationsprozess geklärt werden:

Naheliegend ist die Frage nach der Benennung und dem Bedeutungsumfang von Kategorien und der Granularität ihrer Ausdifferenzierung in Unterkategorien, also dem Aufbau der taxonomischen oder ontologischen Struktur des Annotationsschemas. Dabei spielt die Trennschärfe der Kategorien eine zentrale Rolle, damit auch digital-quantifizierende Analyseprozesse auf Annotationen zugreifen und zu aussagekräftigen Ergebnissen führen können. Aber auch Entscheidungen über die Einheitengrößen der Text-Segmente, die annotiert werden, die Bestimmung von Indikatoren für Kategorien, die jeweilige Interpretationstiefe und den Umgang mit Zweifelsfällen müssen verhandelt und in Richtlinien eindeutig und transparent formuliert und dokumentiert werden. Diese Aushandlungen können über unterschiedliche Kanäle umgesetzt werden. Eine Möglichkeit stellen Wiki-Plattformen dar, auf denen Guidelines in Foren diskutiert, weiterentwickelt und mittels Versionierung auch in ihren unterschiedlichen Stadien dokumentiert werden

können. Einige kollaborative Annotationswerkzeuge bieten auch entsprechende Möglichkeiten direkt in der Annotations-Ansicht des Tools an, z.B. die Anzeige von zuvor festgelegten Guidelines oder die Kommentierung von Zweifelsfällen. Aber auch mündliche Diskussionen – synchron zum Annotationsprozess oder gesondert im Rahmen von Projekt-Treffen, sei es face-to-face oder per Online-Konferenz – spielen in kollaborativen Projekten eine wichtige Rolle. Solche Aushandlungsmöglichkeiten müssen allerdings methodisch angemessen integriert werden. In Phasen der Ermittlung des Inter-Annotator-Agreements kann es bspw. sinnvoll sein, Absprachen im Prozess zunächst zu vermeiden, um Abweichungen und somit den Bedarf für Zweifelsfallklärungen und Differenzierungen ermitteln zu können.

Diese Aushandlungsprozesse als wissenschaftlichen Metadiskurs über Aspekte der Untersuchungsgegenstände anzusehen, erscheint zunächst als adäquate Perspektive. Diskurse im Rahmen des Annotationsprozesses haben die untersuchten Diskurse als Gegenstand und sind – als wissenschaftliche Methode – in der Regel kein einflussnehmender Teil davon, ihre Akteure (die Annotator/inn/en) sind normalerweise nicht mit denen der untersuchten Diskurse identisch, die Kategorien dienen der Systematisierung bzw. Abstraktion und somit der wissenschaftlich-distanzierten und objektivierenden bzw. explizierenden Analyse vor dem Hintergrund einer bestimmten Fragestellung.

Annotieren kann insofern als eine Art Rekontextualisierung der untersuchten Diskurse angesehen werden. Annotationsgegenstände (nicht nur Texte, auch z.B. Bilder, Videos, 3D-Objekte, vgl. Lordick et al. 2016: 190–199) werden aus ihrem ursprünglichen Kontext herausgelöst, ggf. digitalisiert, als Forschungsgegenstand aufbereitet in ein Korpus integriert und insofern schon dadurch in einen neuen Kontext eingebettet. Oder sie werden im ursprünglichen textuellen Umfeld belassen und dort annotiert, bspw. mittels Web-Annotation, also dem Anbringen von Annotationen (mit unterschiedlichem Sichtbarkeits- bzw. Veröffentlichungsstatus) direkt an Webseiten (vgl. Lordick et al. 2016: 194–199). Als eine Form des HerauslöSENS kann aber natürlich auch die Selektion der einzelnen Segmente für die Annotation betrachtet werden. Diese Elemente werden über die Einordnung in ein durch bestimmte Erkenntnisinteressen geprägtes abstrahierendes Schema bzw. die Weiterverarbeitung in einem Analyseverfahren neu kontextualisiert. D.h. es wird an ein gesamtes Korpus, an die einzelnen Texte (bzw. allgemeiner: digitalen

Objekte), aber auch an die segmentierten und mit Tags versehenen Stellen darin, ein neuer Gedankenhorizont perspektivierend angeschlossen.

Dies ist gut beschreibbar mit den Rekontextualisierungs-Dimensionen nach Bauman und Briggs (1990: 75–76): Durch die meta-pragmatische Kennzeichnung des rekontextualisierten Materials als Forschungsgegenstand ist die Dimension des *framing* gegeben. Darunter verstehen Bauman und Briggs „the metacommunicative management of the recontextualized text“ (Bauman/Briggs 1990: 75). Die ausdrucksseitige Transformation (*form*-Dimension) ist durch das Tagging, also die Markierung und Auszeichnung der annotierten Stellen, gegeben, die dann als Repräsentation der annotierten Gegenstands-Segmente in den weiteren Analyseprozess eingehen, woraus zugleich die *function* des rekontextualisierten Materials im neuen Kontext hervorgeht. Hinzu kommen in diesem Funktionszusammenhang die textuellen Ebenen des Kategoriensystems und der ausformulierten Annotationsrichtlinien. Die digitale Erschließung im Zuge des Korpusaufbaus und die Weiterverarbeitung mit digitalen Annotationstools bedingen die mediale Transformation der Gegenstände (*translation* -Dimension) und ihre Neustrukturierung (*emergent structure*).

Hier können auch die vier Grundoperationen des Rekontextualisierens nach Muntigl, Weiss und Wodak (2000: 77) ansetzen: Umordnung, Ergänzung, Tilgung und Ersetzung. Das annotierte Material wird durch Selektion, Tagging, schematische Kategorien-Zuweisung und anschließende Auswertung umgeordnet. Durch Anreicherung mit Zusatzinformationen aus dem Tagset und den verknüpften Guidelines wird es ergänzt und perspektiviert. Annotierte Gegenstandssegmente werden aber auch zur Aufbereitung für den weiteren Analyseprozess in ihrer ursprünglichen Form getilgt und durch Tags ersetzt. Die Systematisierung des Kontexts in Form eines Annotationsschemas und die ausführliche Erklärung der Kategorisierungsbedingungen und des Bezugsgeflechts in Annotationsrichtlinien machen die Wissens- und Relevanzstrukturen des neuen Kontexts und auch den damit verbundenen Erkenntnisweg explizit.

Ein Aspekt des Rekontextualisierungs-Paradigmas besteht allerdings auch darin, dass sinnkonstituierende Rekontextualisierungspraktiken nicht nur von Diskursforscher/inne/n im Zuge der Hervorbringung ihrer Untersuchungsgegenstände vollzogen werden, sondern auch von Diskursakteur/inn/en. Diskurse werden im Grunde ständig neu kontex-

tualisiert – bspw. in unterschiedlichen Akteurskonstellationen, institutionellen Rahmungen und medialen Umgebungen (gerade in digitalen Medien, die vielfältige Umordnungs- und Verknüpfungsmöglichkeiten bieten, etwa durch das schon erwähnte Social-Tagging per Hashtags). Die Differenzierung zwischen Kontextualisierung und Rekontextualisierung bezieht sich auf den Aspekt des Herauslösens und den Ausprägungsgrad des Bruchs, der damit in der Sinnkonstitution des bisherigen Diskursverlaufs bewirkt wird. Eine trennscharfe Abgrenzung nach diesen Kriterien wird hier jedoch als problematisch angesehen. Jedenfalls lässt sich zeigen, dass Rekontextualisierungspraktiken sowohl ein Forschungsprozess als auch ein Diskursphänomen darstellen, ebenso wie die Praktik des Annotierens.

Daher lässt sich – gerade vor dem Hintergrund der beschriebenen Beispiel-Projekte – die Frage anschließen, inwiefern die Perspektive der Diskursforschung je nach Fragestellung auch sinnvoll als Weiterführung von Diskursen, also als eine Form des Diskursphänomens der Rekontextualisierung angesehen werden kann – und nicht nur als reiner Metadiskurs. Beispielsweise wird im Poetiken-Projekt der in den Poetiken geführte Diskurs um die terminologische Bestimmung des Metaphernbegriffs und seiner Definitionskomponenten auf der Ebene der Annotation fortgeführt. Es geht zwar auch darum, den Diskurs in den Gegenständen nachzuvollziehen, zu beschreiben und zu analysieren, aber dieser Forschungsprozess ist ohne eine Weiterführung des Diskurses um die Definition des Metaphernbegriffs und seiner Komponenten nicht zu haben, weil dies Voraussetzung für die sukzessive Entwicklung des Kategoriensystems ist.

Auch im Projekt zu Sprachhandlungen mit Empathie-Bezug in wissenschaftlichen Blogs spielt dieser Aspekt eine Rolle. In den untersuchten Diskursen werden sowohl Kriterien für adäquates soziales Rollenverhalten im Experten-Laien-Spektrum (bzw. in der Wissenschaftskommunikation generell, auch zwischen Wissenschaftler/inne/n unterschiedlicher Disziplinen z.B.) als auch für empathische Kommunikation und ihr Stellenwert im Diskurs metakommunikativ verhandelt. In der Diskursforschungsperspektive des Annotationsprojekts müssen diese Aspekte ebenfalls kollaborativ ausgehandelt werden, um sie in ein adäquates Kategoriensystem überführen zu können. Man befasst sich also auf beiden Ebenen mit denselben Fragestellungen. Hinzu kommt der Anspruch des Projekts, Möglichkeiten der konstruktiven Nutzung von Em-

pathie als Ressource für eine erfolgreiche Wissenschaftskommunikation bzw. als kognitiven und emotiven „Türöffner“ für eine erfolgreichere sachorientierte Verständigung zu erarbeiten (vgl. Bender/Janich 2020, Janich 2020).

An diesen Beispielen zeigt sich, dass die Methode des diskurslinguistischen, kollaborativen Annotierens neben einer metadiskursiven Ebene auch Formen der diskursiven Involviertheit aufweist. Zum einen ist Involviertheit ein zentraler Aspekt des epistemologischen Paradigmas des Spurenlesens (vgl. Kogge 2007: 182–184), das beim (diskurs-)linguistischen Annotieren insofern eine wichtige Rolle spielt, als zur induktiven Kategorienbildung zu allererst explorativ Wissen über Kategorisierungskriterien und die relevante Indikatorik gewonnen werden muss und Tagging als ständige Suche nach Indizien für bestimmte Kategorien zu verstehen ist. Involviertheit steht auf diese Weise in Wechselwirkung mit der konzeptionellen Abgelöstheit bzw. Abstraktion der Metaebene des Annotationsschemas. Die Methode der kollaborativen, diskursiv-hermeneutischen Annotation mit sukzessiver Kategoriendifferenzierung hebt sich – ganz analog zu Kogges Wissensgenerierungskonzept des Spurenlesens – also nicht kategorisch von ihren Gegenständen ab (vgl. Kogge 2007: 183). Wissen kommt dabei „durch ‚Involviert‘, durch Verwicklung der Forschung und der Forschenden ins erforschte Material zustande“ (Kogge 2007: 184). Die damit korrespondierenden kulturtechnischen, schematisierenden Praktiken des Annotierens haben ihre Funktion darin, „als Werkzeuge der systematischen Möglichkeitserkundung zur materialen Forschung beizutragen“ (Kogge 2007: 184).

Involviertheit bedeutet in den beschriebenen Projekt-Szenarien aber unter Umständen auch Verwicklung im Sinne einer diskurskritischen Perspektive, wie schon oben angesprochen. Diskurslinguistische Annotation ist insofern nicht unbedingt als rein deskriptive Methode zu begreifen, wobei die Abgrenzung zwischen Deskription und Kritik in der Diskurslinguistik ohnehin nicht aufrechterhalten wird (Reisigl/Warke 2013: 26–27). Das Explizieren von Diskurs-Phänomenen und ihrer funktionalen Einbettung und Wirkung im Diskurs kann auch dem konstruktiven Herausarbeiten von Verbesserungsmöglichkeiten bzw. ‚Best Practices‘ dienen – zum Beispiel im Sinne erfolgreicher Wissenschaftskommunikation wie oben beschrieben – und somit einen Ansatz zur kritischen Diskursanalyse im Sinne der angewandten Linguistik darstellen. Annotation

ist also in mehrerlei Hinsicht im Spannungsfeld zwischen Metadiskursivität und diskursiver Involviertheit zu verorten.

### 5. Fazit: Annotation als linguistische Methode in digitalisierungsbedingten Spannungsfeldern

Annotation wurde im vorliegenden Aufsatz aus diskurslinguistischer Perspektive als Praktik beschrieben, indem mehrere Schnittstellen bzw. Spannungsfelder – vor allem vor der Folie des Digitalen – dargestellt wurden, an/in denen das Annotieren verortet und über die es in unterschiedlichen Ausprägungen charakterisiert werden kann. Aus diskurslinguistischer Sicht kann Annotation als Diskursphänomen und somit als Untersuchungsgegenstand, aber auch als Methode gesehen werden. Es wurde gezeigt, inwiefern methodisches Annotieren eine Schnittstelle zwischen hermeneutisch-qualitativer und automatisiert-quantifizierender Analyse darstellt und wissenschaftliche Metadiskursivität mit diskursiver Involviertheit verbindet. Diese drei Dimensionen bedingen ein viertes Spannungsfeld im Bereich der Methoden-Einordnung (siehe Abb. 5) – nämlich das zwischen Rogers dichotomisch konzeptualisierten Kategorien ‚digitized‘ und ‚digital‘.

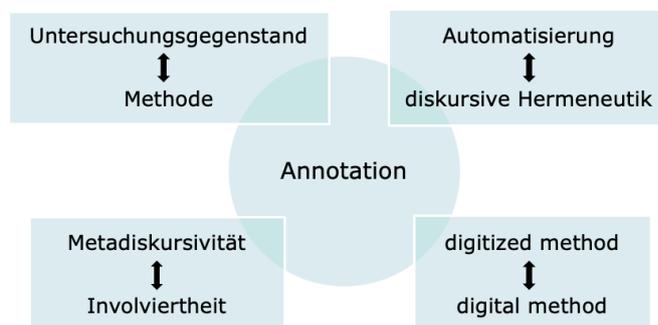


Abb. 5: Verortung von digitaler Annotation in vier Spannungsfeldern

Annotation geht als digitale Methode ganz klar über eine rein mimetische Übertragung ihrer ursprünglich analogen Form ins Digitale hinaus. Sie ist durch ihre Entwicklung hin zur Einbindung algorithmischer Prozesse und ihre Etablierung als diskursive alltagsweltliche Medienpraktik in das

Digitale bzw. digitale Medien eingelassen und somit auch als ‚method of the medium‘ anzusehen. Diese Digitalität wird aber auch rückgebunden an schon im Analogen entwickelte, systematisch kategorisierende Interpretationsformen. Annotation verbindet somit analog und digital geprägte Praktiken und Denkweisen, hermeneutische und algorithmisch-quantifizierende Analyseparadigmen und ist insofern prototypisch für digitale linguistische – oder allgemeiner – geisteswissenschaftliche Forschung. Zusammenfassend erscheint also eine Verortung in allen vier Spannungsfeldern instruktiv und relevant.

## 6. Literatur

- Agosti, Maristella; Bonfiglio-Dosio, Giorgetta; Ferro, Nicola (2007): A Historical and Contemporary Study on Annotations to Derive Key Features for Systems Design. In: *International Journal on Digital Libraries*, 8 (1) 1–19. DOI:10.1007/s00799-007-0010-0.
- Aijmer, Karin (2015): Pragmatic Markers. In: Aijmer, Karin; Rühlemann, Christoph (Hrsg.): *Corpus Pragmatics. A Handbook*. Cambridge: Cambridge University Press, 195–218.
- 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*. 2016. DOI: 10.17175/2016\_004
- Alscher, Stefan; Bender, Michael; Rapp, Andrea; Richter, Sandra et al. (2016): ePoetics – 20 Poetiken (1770–1960), TEI-konformes XML, HTML-Lesefassung, synoptische Faksimile-Ansicht, inkl. Annotationen des BMBF-Projekts epoetics, publiziert im Deutschen Textarchiv DTA, URL: <http://www.deutschestextarchiv.de/search/metadata?corpus=epoetics>.
- Archer, Dawn; Culpeper, Jonathan; Davies, Matthew (2008): Pragmatic Annotation. In: Lüdeling, Anke; Kytö, Merja (Ed.): *Corpus Linguistics. An International Handbook*. Berlin: de Gruyter, 613–641.
- Arn, Christof (2016): *Agile Hochschuldidaktik*. Weinheim/Basel: Beltz.
- Artstein, Ron; Poesio, Massimo (2008): Inter-Coder-Agreement for Computational Linguistics. In: *Computational Linguistics* 34 (4), 555–596. DOI: <https://doi.org/10.1162/col.07-034-R2>
- Auer, Peter (1986): Kontextualisierung. In: *Studium Linguistik* 19, 22–47.
- Baumann, Richard; Briggs, Charles L. (1990): Poetics and Performance as Critical Perspectives on Language and Social Life. In: *Annual Review of Anthropology* 19, 59–88.
- Becker, Maria; Bender, Michael; Müller, Marcus (in Vorb.): Classifying Heuristic Textual Practices in Academic Discourse. A Deep Learning Approach to Pragmatics. In: *International Journal of Corpus Linguistics* (zur zweiten Review-Phase angenommen).

- Bender, Michael; Müller, Marcus (2020): Heuristische Textpraktiken. Eine kollaborative Annotationsstudie zum akademischen Diskurs. In: Zeitschrift für Germanistische Linguistik 48 (1): 1-46. DOI: <https://doi.org/10.1515/zgl-2020-0001>
- Bender, Michael; Janich, Nina (2020): Empathie in der Wissenschaftskommunikation. Eine Forschungsskizze. In: Jacob, Katharina; Konecny, Klaus Peter; Liebert, Andreas: Sprache und Empathie. Reihe: Sprache und Wissen (SuW). Berlin, Boston: de Gruyter, 425–449. DOI: <https://doi.org/10.1515/9783110679618-014>
- Bender, Michael (2016a): Forschungsumgebungen in den Digital Humanities: Nutzerbedarf, Wissenstransfer, Textualität. Reihe: Sprache und Wissen (SuW) 22. Berlin, Boston: de Gruyter. 2016.
- Bender, Michael (2016b): Digitale Methoden und Kulturtechniken. In: Friese, Heidrun; Rebane, Gala; Nolden, Marcus; Schreiter, Miriam: Handbuch soziale Praktiken und digitale Alltagswelten. Heidelberg: Springer. DOI: [https://doi.org/10.1007/978-3-658-08460-8\\_48-1](https://doi.org/10.1007/978-3-658-08460-8_48-1)
- Bender, Michael; Kollatz, Thomas; Rapp, Andrea (2018): Objekte im digitalen Diskurs – epistemologische Zugänge zu Objekten durch Digitalisierung und diskursive Einbindung in virtuelle Forschungsumgebungen und -infrastrukturen. In: Hilgert, Markus; Hofmann, Kerstin P.; Simon, Henrike (Hrsg.): Objekt epistemologien. Zur Vermessung eines transdisziplinären Forschungsraums. Berlin Studies of the Ancient World 59. Berlin: Edition Topoi, 107–132. (URL: [edition-topoi.org/articles/details/1444](http://edition-topoi.org/articles/details/1444))
- Bernard, Andreas (2018): Das Diktat des #hashtags. Über ein Prinzip der aktuellen Debattenbildung. Frankfurt a.M.: Fischer.
- Biemann, Chris (2012): "Creating a system for lexical substitutions from scratch using crowdsourcing". Language Resources and Evaluation 47 (1), 97–122.
- Bradley, John (2012): Towards a richer sense of digital annotation: Moving beyond a „Media“ orientation of the annotation of digital objects. Digital Humanities Quarterly 6 (2). URL: <http://www.digitalhumanities.org/dhq/vol/6/2/000121/000121.html>. (Abrufdatum: 15.8.2019)
- Braun, Steffen (2016): Einflussfaktoren auf den Wechsel des Abschlussprüfers. Eine empirische Analyse bei kapitalmarktorientierten Unternehmen. Diss. Darmstadt. URL: <http://tuprints.ulb.tu-darmstadt.de/5377>
- Breyer, Thimo (2013): Empathie und ihre Grenzen: Diskursive Vielfalt – phänomenale Einheit? In: Ders. (Hg.): Grenzen der Empathie. Philosophische, psychologische und anthropologische Perspektiven. München: Fink (Übergänge 63), 13–44.
- Bubenhof, Noah (2013): Quantitativ informierte qualitative Diskursanalyse. Korpuslinguistische Zugänge zu Einzeltexten und Serien. In: Roth, Kersten Sven; Spiegel, Carmen (Hrsg.): Angewandte Diskurslinguistik. Felder, Probleme, Perspektiven. Berlin: Akademie, 109–134.
- Dang-Anh, Mark; Einspänner, Jessica; Thimm, Caja (2013): Kontextualisierung durch Hashtags: die Mediatisierung des politischen Sprachgebrauchs im Internet. In: Diekmannshenke, Hajo; Niehr, Thomas (Hrsg.): Öffentliche Wörter: Analysen zum öffentlich-medialen Sprachgebrauch. Stuttgart: ibidem (Perspektiven Germanistischer Linguistik 9), 137–159. URL: <http://nbn-resolving.de/urn:nbn:de:0168-ss0ar-55975-5>

- Dang-Anh, Mark; Einspänner, Jessica; Thimm, Jaja (2012): Mediatisierung und Medialität in Social Media: Das Diskursystem "Twitter". In: Marx, Konstanze; Schwarz-Friesel, Monika (Hrsg.): Sprache und Kommunikation im technischen Zeitalter. Wieviel Internet (v)erträgt unsere Gesellschaft? Berlin, 68–94.
- Feilke, Helmuth (1994): Common sense-Kompetenz. Überlegungen zu einer Theorie „sympathischen“ und „natürlichen“ Meinens und Verstehens. Frankfurt am Main: Suhrkamp.
- Felder, Ekkehard (2013): Faktizitätsherstellung mittels handlungsleitender Konzepte und agonaler Zentren. Der diskursive Wettkampf um Geltungsansprüche. In: Ders. (Hg.): Faktizitätsherstellung in Diskursen. Die Macht des Deklarativen. Berlin/Boston: de Gruyter (Sprache und Wissen 13), 13–28.
- Giles, Jim (2005): Internet encyclopaedias go head to head. *Nature* 438, 900–901.
- 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>.
- Golder, Scott; Huberman, Bernardo (2006): Usage Patterns of Collaborative Tagging Systems. *Journal of Information Science*, 32 (2), 198–208.
- Gredel, Eva (2018): Digitale Methoden und Werkzeuge für Diskursanalysen am Beispiel der Wikipedia. In: Huber, Martin; Krämer, Sybille (Hg.): Wie Digitalität die Geisteswissenschaften verändert: Neue Forschungsgegenstände und Methoden. Sonderband der Zeitschrift für digitale Geisteswissenschaften, 3. DOI: 10.17175/sb003\_005
- Gruber, Helmut (2017): Quoting and retweeting as communicative practices in computer mediated discourse. In: *Discourse, Context and Media* 20, 1–9.
- Gumperz, John J. (1982): *Discourse Strategies*. Cambridge.
- Ide, Nancy (2017): Introduction: The Handbook of Linguistic Annotation. In: Ide, Nancy; Pustejovsky, James (Hg.): *Handbook of Linguistic Annotation*. Vol. I, Dordrecht: Springer, 1–18.
- Janich, Nina (2020): What Do You Expect? Linguistic Reflections on Empathy in Science Communication. In: *Media and Communication* 2020, 8 (1).
- Juska-Bacher, Britta; Biemann, Chris; Quasthoff, Uwe (2013): Webbasierte linguistische Forschung: Möglichkeiten und Begrenzungen beim Umgang mit Massendaten. In: *Linguistik online* 61, 4/2013, URL: [https://www.linguistik-online.net/61\\_13/juskabacherBiemannQuasthoff.html](https://www.linguistik-online.net/61_13/juskabacherBiemannQuasthoff.html)
- Kogge, Werner (2007): Spurenlesen als epistemologischer Grundbegriff: Das Beispiel der Molekularbiologie. In: Krämer, Sybille; Kogge, Werner; Grube, Gernot (Hg.): *Spur. Spurenlesen als Orientierungstechnik und Wissenskunst*. Frankfurt am Main: Suhrkamp, 182–221.
- Lordick, Harald; Becker, Rainer; Bender, Michael; Borek, Luise; Hastik, Canan; Kollatz, Thomas; Mache, Beata; Rapp, Andrea; Reiche, Ruth; Walkowski, Niels-Oliver (2016): Digitale Annotationen in der geisteswissenschaftlichen Praxis. In: *Bibliothek Forschung und Praxis*. Hrsg. Bonte, Achim et al. Band 40, Heft 2 (Juli 2016). Berlin, Boston: de Gruyter, 186–199. (URL: <https://www.degruyter.com/downloadpdf/j/bfup.2016.40.issue-2/bfp-2016-0042/bfp-2016-0042.pdf>)
- Marx, Konstanze (2019): Von #Gänsehaut bis #esreicht – Wie lässt sich ein Territorium neuer Sagbarkeit konturieren? Ein phänomenologischer Zu-

- gang. In: Eichinger, Ludwig; Plewnia, Albrecht (Hg.): Neues vom heutigen Deutsch, 245–264. Berlin, Boston: De Gruyter. DOI: <https://doi.org/10.1515/9783110622591-012>
- McCarty, Willard (2004): Modeling – A study in words and meanings. In: Schreibman, Susan/Siemens, Ray/Unsworth, John (Hg.): A companion to digital humanities. Oxford, 254–271.
- Müller, Marcus (2015): Sprachliches Rollenverhalten. Korpuspragmatische Studien zu divergenten Kontextualisierungen in Mündlichkeit und Schriftlichkeit. Berlin, Boston.
- Müller, Marcus; Stegmeier, Jörn (2016): Twittern als #Alltagspraxis des Kunstpublikums. In: Zeitschrift für Literaturwissenschaft und Linguistik (LiLi) (4/2016), 499–522.
- Müller, Marcus; Mell, Ruth (im Druck): Zwischen Fach und Wort. Fragen, Methoden und Erkenntnisse der Terminologiedynamik. In: Dominika Bopp, Stefaniya Ptashnyk, Kerstin Roth & Tina Theobald (Hg.): Wörter - Zeichen der Veränderung (Gedenkschrift für Jörg Riecke). Berlin, Boston: De Gruyter. (vorauss. 06/2020)
- Müller-Prove, Matthias (2008): Modell und Anwendungsperspektive des Social Tagging. In: Gaiser, Birgit; Hampel, Thorsten; Panke, Stefanie (Hrsg.): Good Tags - Bad Tags. Social Tagging in der Wissensorganisation. Münster u.a.: Waxmann, 15–22.
- Muntigl, Peter; Weiss, Gilbert; Wodak, Ruth (2000): European Union Discourses on Un/employment: An interdisciplinary approach to employment policy-making and organizational change. Amsterdam (= Dialogues on Work and Innovation).
- Rapp, Andrea (2017): Manuelle und automatische Annotation. In: Jannidis, Fotis/Kohle, Hubertus/Rehbein, Malte (Hg.): Digital Humanities. Eine Einführung. Stuttgart: Springer, 253–267.
- Reisigl, Martin; Warnke, Ingo (2013): Diskurslinguistik im Spannungsfeld von Deskription, Präskriptiv und Kritik. Eine Einleitung. In: Meinhof, Ulrike; Reisigl, Martin; Warnke, Ingo (Hg.): Diskurslinguistik im Spannungsfeld von Deskription und Kritik. Berlin: Akademie Verlag, 7–36.
- Rogers, Richard (2013): Digital Methods. Cambridge: MIT Press.
- Schiller, Anne; Teufel, Simone; Stöckert, Christine; Thielen, Christine (1999): Guidelines für das Tagging deutscher Textkorpora mit STTS (Kleines und großes Tagset). Stuttgart/Tübingen. URL: <http://www.sfs.uni-tuebingen.de/resources/stts-1999.pdf>
- Searle, John (1969): Speech Acts. Cambridge: Cambridge Univ. Press.
- Spitzmüller, Jürgen; Warnke, Ingo (2011): Diskurslinguistik. Eine Einführung in Theorien und Methoden der transtextuellen Sprachanalyse. Berlin/New York: de Gruyter.
- Steffek, Jens; Müller, Marcus; Behr, Hartmut (2020): Terminological Entrepreneurs and Discursive Shifts in International Relations: How a Discipline Invented the “International Regime”. In: International Studies Review viaa003/2020.
- Stricker, Stefanie (2009): Definitiorische Vorklärungen. In: Rolf Bergmann, Stefanie Stricker (Hg.): Die althochdeutsche und altsächsische Glossographie. Ein Handbuch. 2 Bände. Berlin/ New York 2009, 20–32.

- Vander Wal, Thomas (2005): Folksonomy. Presentation at the Online Information Conference. London. URL: <http://vanderwal.net/essays/051130/folksonomy.pdf>
- Weisser, Martin (2015): Speech Act Annotation. In: Ajmer, Karin; Rühlemann, Christoph (Ed.): *Corpus Pragmatics. A Handbook*. Cambridge: Cambridge University Press, 84–116.
- Weisser, Martin (2018): *How to Do Corpus Pragmatics on Pragmatically Annotated Data*. Amsterdam / Philadelphia: John Benjamins.
- Zappavigna, Michelle (2011): Ambient affiliation: A linguistic perspective on Twitter. In: *New Media & Society* 13 (5), 788–806.
- Zappavigna, Michelle (2015). Searchable talk: the linguistic functions of hashtags. *Social Semiotics* 25 (3), 274–291. <http://dx.doi.org/10.1080/10350330.2014.996948>.