



Zeitschrift für Diskursforschung

Journal for Discourse Studies

Herausgegeben von Reiner Keller | Werner Schneider | Willy Viehöver

- **Thomas Niehr / Eva Dickmeis / Bianka Trevisan / Eva-Maria Jakobs**
Neue Wege der linguistischen Diskursforschung
- **Marlon Barbehön**
Städtische Europadiskurse und die Konstitution lokalpolitischer
Möglichkeitsräume: Das Beispiel Feinstaubpolitik in Frankfurt am Main
- **Simon-Philipp Pfersdorf**
Die diskursive Konstruktion von Regulierungswissen am Beispiel
des gesellschaftlichen Umgangs mit Nanotechnologie
- **Reiner Keller / Achim Landwehr / Wolf-Andreas Liebert /
Werner Schneider / Jürgen Spitzmüller / Willy Viehöver**
Diskurse untersuchen – ein Gespräch zwischen den Disziplinen (Teil 2)

Inhaltsverzeichnis

Reiner Keller / Willy Viehöver / Werner Schneider

Editorial 110

Themenbeiträge

Thomas Niehr / Eva Dickmeis / Bianka Trevisan / Eva-Maria Jakobs

Neue Wege der linguistischen Diskursforschung 113

Marlon Barbehön

Städtische Europadiskurse und die Konstitution lokalpolitischer
Möglichkeitsräume: Das Beispiel Feinstaubpolitik in Frankfurt am Main 137

Simon-Philipp Pfersdorf

Die diskursive Konstruktion von Regulierungswissen am Beispiel
des gesellschaftlichen Umgangs mit Nanotechnologie 159

*Reiner Keller / Achim Landwehr / Wolf-Andreas Liebert / Werner Schneider /
Jürgen Spitzmüller / Willy Viehöver*

Diskurse untersuchen – ein Gespräch zwischen den Disziplinen (Teil 2) 183

Review

Annette Knaut

Rezension zu Michael Kauppert / Irene Leser (Hrsg.):
Hillarys Hand. Zur politischen Ikonographie der Gegenwart. 208

Berichte

Maya Halatcheva-Trapp / Wolf J. Schünemann

Die Diskursive Konstruktion von Wirklichkeit II –
Interdisziplinäre Perspektiven einer wissensoziologischen Diskursforschung 216

Thomas Niehr / Eva Dickmeis / Bianka Trevisan / Eva-Maria Jakobs

Neue Wege der linguistischen Diskursforschung

Computerbasierte Verfahren der Argumentanalyse

Zusammenfassung: Der vorliegende Beitrag beschreibt und diskutiert eine neuartige Verbindung quantitativer und qualitativer Verfahren für die Analyse von Big Data in der linguistischen Diskursforschung. Der vorgestellte Ansatz kombiniert Methoden der diskurslinguistischen Argumentationsanalyse mit Methoden des Linguistischen Text Mining. Das Ziel der Methodenentwicklung ist ein computergestütztes Verfahren für die semi-automatisierte Identifizierung und Analyse von Argumenten in großen Textkorpora. Erprobt wird das Verfahren an einem Diskurs über Infrastrukturmaßnahmen. Im Beitrag werden sprachliche Mittel vorgestellt, die im Korpus gemeinsam auftreten und damit als Merkmale von Argumentmustern betrachtet werden können. Solche Argumentmuster können das Vorkommen von Argumenten und ihren Verwendungsweisen in Texten indizieren.

Schlagwörter: Linguistische Diskursanalyse, Argumentationsanalyse, Text Mining, Argumentation Mining, Argument

Abstract: The following paper presents and discusses an innovative approach for the automated detection and analysis of arguments in big data corpora. The approach links qualitative methods from discourse linguistics with quantitative text mining methods. The aim is to identify linguistic means that indicate the occurrence of an argument and its various production modes in large text corpora. The public discourse about the introduction of transport infrastructure is used as an application example. As a result, the paper presents various linguistic means found in the corpus that can be used as features of argument patterns.

Keywords: Discourse Linguistics, Argumentation Analysis, Text Mining, Argumentation Mining, Argument

1 Einleitung

Der Beitrag diskutiert am Beispiel eines Diskurses über Infrastrukturmaßnahmen das Potential der Verbindung qualitativer und quantitativer Vorgehensweisen für die Weiterentwicklung linguistisch orientierter Diskursanalysen. Der Ansatz nutzt diskurslinguistische Methoden der Argumentationsanalyse und kombiniert diese mit neuen Entwicklungen des »Argumentation Mining«. Der Fokus richtet sich auf computergestützte Verfahren der Identifizierung und Analyse von Argumenten. Im Folgenden werden zunächst Ansätze der linguistischen Diskursanalyse (Kapitel 2), zu Argumentation (Kapitel 3) und Verwendungsweisen von Argumenten (Kapitel 4) sowie Herausforderungen ihrer computergestützten Identifikation (Kapitel 5) dargestellt und diskutiert. Kapitel 6 bietet einen

Überblick zu neueren Ansätzen des Text und Argumentation Mining. In Kapitel 7 wird fallbeispielbasiert ein Ansatz zur Verbindung diskurslinguistischer und computergestützter Verfahren vorgestellt. Der Beitrag schließt mit einem Ausblick auf offene Forschungsfragen (Kapitel 8).

2 Linguistische Diskursanalyse

In der linguistischen Diskursanalyse wird seit einiger Zeit eine Diskussion darüber geführt, inwieweit quantitative Methoden das Paradigma einer hermeneutisch ausgerichteten Diskursanalyse ablösen können oder – stärker noch – sollen (Bubenhofner 2008, 2013; Scharloth/Eugster/Bubenhofner 2013; Niehr 2015). Letzteres wird damit begründet, dass der Umgang mit »Big Data« und unüberschaubar großen Textkorpora andere Methoden verlange als die Einzelsichtung von Belegen und deren Interpretation mittels philologischer Analysen. Die Lösung wird in rein quantitativ orientierten Verfahren gesehen, da nur sie in der Lage seien, große Textmengen zu bearbeiten.

Das Ziel linguistischer Diskursanalysen besteht üblicherweise darin, anhand sprachlicher Muster Regelmäßigkeiten und typische Verteilungen festzustellen, um daraus auf Denkmuster, ihre Verbreitung und ihren Wandel zu schließen. Die in Diskursen erfolgende Argumentation bietet einen möglichen Zugang, um solche Denkmuster zu erforschen. Erste Ansätze zur systematischen Analyse von Argumentationen in öffentlichen Diskursen stammen von Wengeler (2003) und Niehr (2004). Am Beispiel der Argumentation pro und contra Einwanderung entwickeln sie ein methodisches Instrumentarium, das themenunabhängig für die Analyse diskursbasierter Argumentationen eingesetzt werden kann. Die Analysen gehen zwar quantifizierend vor, sie verstehen sich jedoch als hermeneutisch, da sie eine interpretierende Analyse aller im Textkorpus vorhandenen Belege voraussetzen. Basierend auf der interpretierenden Analyse werden Belege bzw. Textstellen kategorisiert und anschließend, aufgrund ihrer Frequenz als für den jeweiligen Diskurs typische Muster identifiziert. Der Fokus liegt eindeutig auf der linguistischen Interpretation von Belegen, die für die anschließende Kategorienbildung unerlässlich ist. Wengeler (2003, S. 297 ff.) ordnet nach diesem Verfahren Einzelbelege aufgrund von semantischen Kriterien Topoi zu; Niehr (2004, S. 129 ff.) nutzt pragmatische Kategorisierungskriterien, anhand derer er verschiedene Argument-Verwendungen unterscheidet und auf diese Weise die Umstrittenheit von Argumenten aufzeigt.

3 Argumentation

In einer der Rationalität verpflichteten Gemeinschaft müssen strittige Geltungsansprüche durch Argumente verteidigt bzw. die jeweiligen Opponenten durch Argumente überzeugt werden (Kopperschmidt 1989, S. 25 ff.; Apel 1999, S. 424 ff.; Habermas 1981, S. 37 ff.). In diesem Sinne ist Argumentieren eine komplexe sprachliche Handlung, mit der SprecherInnen/SchreiberInnen das Ziel verfolgen, einen in Zweifel gezogenen Geltungs-

anspruch zu verteidigen (Klein 1980, S. 19). Argumentation wird erst dann nötig, wenn Geltungsansprüche tatsächlich bestritten und mithin problematisch werden. Ähnlich fasst Kopperschmidt (2005, S. 59) Argumentation als »eine geregelte Abfolge (Sequenz) von Sprechhandlungen [...], die zusammen ein mehr oder weniger komplexes, kohärentes und intentionales Beziehungsnetz zwischen Aussagen bilden, das der methodischen Einlösung von problematisierten Geltungsansprüchen dient«. Das argumentative Verfahren besteht darin »durch Rede und Gegenrede bzw. durch Nachfragen und Angeben von Gründen« problematisierte Geltungsansprüche zu verteidigen (ebd., S. 75). Als Mittel der Argumentation werden üblicherweise Argumente herangezogen, die bezogen auf ihre Plausibilität und Relevanz bewertbar sind. Insofern lassen sich Argumente als sprachliche Äußerungen definieren, die die Funktion haben, in einer Argumentation Geltungsansprüche zu verteidigen.

Dass dieses komplexe sprachliche Verfahren nicht adäquat durch einen klassischen Syllogismus aus Oberprämisse, Unterprämisse und daraus folgender Konklusion modelliert werden kann, hat bereits Toulmin (1996\1958) gut begründet nachgewiesen. Er verweist zu Recht darauf, dass durch einen Syllogismus keine Erkenntniserweiterung stattfindet, da die Konklusion lediglich durch Umstellung bzw. Kombination der Prämissen entstehe: Akzeptiert man die Prämissen eines Syllogismus, so muss man gemäß den Regeln der formalen Logik auch seiner Konklusion zustimmen, da diese implizit in den Prämissen enthalten ist.

In realen Argumentationssituationen, d.h. bei »substantiellen« Argumentationen (Toulmin 1996, S. 112 f.), hilft die formale Logik nicht weiter, und zwar aus mehreren Gründen. Ihr Augenmerk ist auf die formale Gültigkeit von Schlüssen gerichtet, ohne deren Inhalte zu beachten. Im Zentrum der formalen Logik steht mithin die Art der Verknüpfung von Aussagen beliebigen Inhalts. Dass dies zu absurden Konsequenzen für die Argumentationsanalyse führt, zeigt beispielsweise die sogenannte materiale Implikation des Typs $p \rightarrow q$ [lies: wenn p, dann q]. Dem Schluss kommt der Wahrheitswert »falsch« lediglich unter der Bedingung zu, dass p der Wahrheitswert »wahr« und q der Wahrheitswert »falsch« zugeschrieben wird. Alle anderen Kombinationen von »wahr« und »falsch« führen zu einem »wahren« Konditional. Ein Beispiel dafür ist der Satz *Wenn Frankreich in Australien liegt, ist die See süß* (Quine 1974, S. 39 ff.).

Der zweite Grund dafür, dass die formale Logik zur Analyse realer Argumentationen nicht geeignet ist, ist darin zu sehen, dass letztere meist durch einen hohen Grad an Implizitheit gekennzeichnet sind. Insbesondere die von Toulmin als »Schlussregel« bezeichnete Komponente, die den Übergang von Daten zur Konklusion gewährleisten soll (im klassischen Syllogismus die Oberprämisse), findet sich eher selten in realen Kommunikationssituationen (eine Ausnahme sind wissenschaftliche Texte). Gerade im öffentlichen Diskurs, der in den Massenmedien stattfindet, bleibt es häufig bei thesenartigen Behauptungen, die zunächst nicht weiter argumentativ gerechtfertigt werden (Niehr i.Vb. sowie die Beispiele in Niehr 2014, S. 107 ff.).

Mit Toulmin (1996, S. 148 ff.) lässt sich zusammenfassend feststellen, dass die Kriterien der formalen Logik nicht auf Argumentationen des täglichen Lebens anzuwenden sind. Hier gelten vielmehr die Standards der »Alltagslogik« (Kienpointner 1992), für die

als zentrales Kriterium das der Plausibilität angesetzt werden muss. Plausibilität aber ist nicht aufgrund formaler Merkmale beschreibbar, da Plausibilität eine inhaltlich bestimmte Kategorie ist. Als Bedeutungsparaphrase zu *plausibel* finden sich im Duden Universalwörterbuch (2011, S. 1348) die Adjektive *einleuchtend*, *verständlich*, *begreiflich*. Damit wird eine weitere Besonderheit dieser Kategorie deutlich: Ob ein Argument plausibel erscheint, kann immer nur diskursiv ermittelt werden und setzt die Ratifizierung durch die jeweiligen Kommunikationsbeteiligten voraus.

4 Verwendungsweisen von Argumenten

Wenn Argumentation dazu dient, strittige Geltungsansprüche zu rechtfertigen, liegt es nahe, Argumentation als eine geordnete Abfolge von Pro- und Contra-Argumenten aufzufassen. Betrachtet man reale Argumentationen des öffentlichen Diskurses, so zeigt sich, dass diese Auffassung zu korrigieren ist. Dies liegt einerseits an der bereits beschriebenen Implizitheit von Argumenten, die es nicht immer ermöglicht, im Text vorkommende sprachliche Äußerungen eindeutig als Pro- oder Contra-Argumente zu klassifizieren (Niehr 2004, S. 115 f., 129). Ein zweiter Grund ist, dass die an einer Argumentation Beteiligten häufig Argumente zitieren. Diese »uneigentliche« Verwendung von Argumenten lässt sich linguistisch als »Erwähnung« (im Gegensatz zum »Vollzug«) von Argumenten kategorisieren (Niehr i.Vb.); sie kann unterschiedlichen Zwecken dienen und mithin unterschiedliche argumentative Funktionen übernehmen.

Mit Niehr (2004, S. 130 ff.) lassen sich vier Verwendungsweisen von Argumenten unterscheiden. Argumente werden argumentativ, neutral zitierend, positiv zitierend oder negativ zitierend verwendet.

Die argumentative Verwendung ist die »eigentliche« Verwendung eines Arguments als Argument. Sie lässt sich auch als Vollzug bzw. Gebrauch eines Arguments beschreiben. Dies zeigt das folgende Beispiel:

(1) »Der Ausbau der Elektromobilität und der Bau der Campusbahn sind auch aus Umweltgesichtspunkten für unsere Stadt wichtig. Die Bahn kann in der ersten Ausbaustufe 1,6 Millionen Buskilometer ersetzen. Der Betrieb der Bahn wird dazu beitragen, die Feinstaubbelastung der Luft im Aachener Kessel zu senken.« (Aachener Nachrichten, 31.02.2013)

In Beispiel (1) wird als Argument für den Bau der Campusbahn angeführt, dass der Ausbau der Elektromobilität die Feinstaubbelastung in Aachen verringern werde. Die Verringerung der Feinstaubbelastung in Aachen wird mithin als Argument verwendet, um den Bau der Campusbahn argumentativ zu stützen.

Die Verwendungsweisen von Argumenten in den Beispielen (2) bis (4) unterscheiden sich insofern von der eigentlichen Verwendungsweise in Beispiel (1), als der/die Argumentierende ein Argument lediglich zitierend aufgreift (»erwähnt«), um es funktional für seine Argumentation zu nutzen.

In Beispiel (2) wird ein Argument neutral, d. h. ohne Bewertung zitiert. Neutrale Argument-Zitate sind u. a. charakteristisch für Nachrichtentexte, in denen die Argumentation einer Person oder Institution referiert wird. In Beispiel (2) wird das Argument des finanziellen Risikos, das für den Haushalt der Stadt mit dem Bau der Campusbahn entsteht, neutral zitiert:

(2) »Slawinski erläuterte sein Hauptargument, das Prestigeprojekt Campusbahn sei gefährlich, weil es den Stadthaushalt zusätzlich tief in die roten Zahlen treibe und drohte damit, dass zahlreiche freiwilligen [sic!] Leistungen in einem dann genehmigungsbedürftigen Nothaushalt nicht mehr gezahlt werden könnten.« (Aachener Nachrichten, 03.03.2013)

Die neutral zitierende Argument-Verwendung stellt Argumentationsanalysen, die sich an Oberflächenmerkmalen orientieren, vor besondere Herausforderungen. Denn das Zitieren von Argumenten schließt auch argumentative Verknüpfungen durch Konjunktionen ein. So kann an der Textoberfläche eine Argument-Verwendung signalisiert werden, wiewohl lediglich ein Argument zitiert wird.

Positiv zitierende Argument-Verwendungen werden üblicherweise realisiert, indem ein Argument zitiert und gleichzeitig positiv bewertet und damit bekräftigt wird (Beispiel 3). Die Bewertung signalisiert Zustimmung zu einer im Diskurs vertretenen Position oder wertet die eigene Argumentation implizit auf, indem sie als übereinstimmend mit einer anderen positiv bewerteten Argumentation dargestellt wird. Letzteres wird durch sogenannte Autoritätsargumente bewirkt, die ihre argumentative Kraft zu mehr oder minder großen Teilen aus der Berufung auf die Argumente von Autoritäten beziehen. In Beispiel (3) erfolgt dies durch positives Zitieren eines Arguments der Stadtkämmerin:

(3) »Es kann nur eine Entscheidung geben: nämlich nein! Denn die Stadtkämmerin als oberste Kassenhüterin von Aachen hat dieses Vorhaben für die Stadt als absolut nicht stemmbar erklärt. Jeder weiß doch, dass öffentliche Bauvorhaben zunächst minimal kalkuliert werden, um dann nach Genehmigung ins Uferlose zu steigen.« (Aachener Zeitung, 26.02.2013)

Negativ zitierende Argument-Verwendungen zeichnen sich dadurch aus, dass ein gegnerisches Argument zunächst zitiert und dann entkräftet oder widerlegt wird. Die Besonderheit dieser Art der Argument-Verwendung besteht darin, dass sie häufig wie eine neutrale Argument-Zitation aussieht und erst im weiteren Textverlauf deutlich wird, dass eine negative Bezugnahme auf das jeweilige Argument vorliegt. In Beispiel (4) wird das Argument der Finanzierbarkeit durch eine Campusbahn-Abgabe negativ zitiert:

(4) »Das tollkühnste Argument zum Thema Finanzierbarkeit, welches ich aus dem Munde eines Aachener Kommunalpolitikers vernahm, war, bald entfiere ja der Solidarbeitrag zur Einkommensteuer, und den könne man dann ja in Aachen als Campus-

bahn-Abgabe erheben. Ein überraschendes Verständnis von der Finanzverfassung unseres Staates, wenn nicht totale Ahnungslosigkeit.« (Webkommentar, Aachener Zeitung, 15.01.2013)

Für die Analyse von Alltagsargumentationen, beispielsweise im Kontext von Infrastrukturentscheidungen, dürfte insbesondere die positiv bzw. negativ zitierende Verwendung von Argumenten von Bedeutung sein. Das gehäufte positive Zitieren eines Arguments ist nicht nur Ausdruck seiner Verbreitung, sondern auch Indiz für seine Akzeptanz. Umgekehrt kann das häufige negative Zitieren eines Arguments als Hinweis auf seine Umstrittenheit gedeutet werden.

Zusammenfassend ist festzustellen, dass Argumentationsanalysen, die die in Diskursen vorkommenden Argumentationen adäquat analysieren wollen, verschiedene Verwendungsweisen von Argumenten berücksichtigen müssen. Niehr (2004) entwickelt dafür eine Vorgehensweise: Ziel der Diskursanalyse ist demnach zu rekonstruieren, mit welchen Argumenten in einem Diskurs für oder gegen einen strittigen Sachverhalt argumentiert wird. Hierzu werden in den Texten des Korpus konkret vorkommende, diskursrelevante Einzelargumente identifiziert. Die sprachlich sehr unterschiedlich realisierten Argumente werden auf ihren gemeinsamen inhaltlichen Kern untersucht und zu prototypisch formulierten Argumentklassen zusammengefasst. Die Diskursrelevanz eines Arguments wird quantitativ über seine Vorkommenshäufigkeit im Korpus bestimmt (Niehr 2004, S. 111 f.).

5 Herausforderungen der computergestützten Identifikation von Argumenten

Im Hinblick auf große Korpora wäre es sehr hilfreich, Argumente computergestützt identifizieren zu können. Ansätze, die dies leisten wollen, müssen sich verschiedenen Herausforderungen stellen. Eine Herausforderung besteht darin, dass – wie in Kapitel 3 erläutert – Argumente nicht anhand formaler Kriterien erfassbar sind, da es sich um funktionsbezogene Kategorien handelt (Kopperschmidt 1989, S. 94 ff.).

Eine zweite Herausforderung ergibt sich mit der Kontexteinbettung. Da *Argument* eine funktionale, kontextsensitive Kategorie ist, lassen sich Argumente immer nur in einem sprachlichen Kontext als Argumente identifizieren. Oder anders formuliert: Eine Äußerung ist nicht per se Argument, sondern kann lediglich in einem spezifischen sprachlichen Kontext eine argumentative Funktion übernehmen. So kann die Äußerung

(5) Die staatliche Förderung von Elektromobilität reduziert die Umweltbelastung.

als Argument dienen, wenn es darum geht, den Geltungsanspruch zu untermauern, dass die staatliche Förderung der Elektromobilität ein politisch sinnvolles Instrument ist. Als prototypisches Mittel der Argumentation gilt die kausale Verknüpfung, da mit ihr Gründe angegeben werden können. Dementsprechend lässt sich Beispiel (5) in ein ausformuliertes Argument (Beispiel 6) überführen, dem ein kausaler Schluss zugrunde liegt:

(6) *Die staatliche Förderung von Elektromobilität ist ein politisch sinnvolles Instrument, weil durch sie die verkehrsbedingte Umweltbelastung reduziert werden kann.*

Allerdings stellt es eine unzulässige Verkürzung dar, Argumentation auf die kausale Verknüpfung von Aussagen zu reduzieren (vgl. dazu Niehr i. Vb.). Dies ist insbesondere im Hinblick auf das Auffinden von Argumenten in Texten von Bedeutung: Die Vorstellung, eine gezielte Suche nach kausalen Konjunktionen wie *denn*, *da*, *weil* fördere sämtliche im Text enthaltenen Argumente zutage, wird der komplexen Sprachhandlung Argumentieren nicht gerecht, da auch zahlreiche andere logische Verknüpfungen von Äußerungen funktional als Argument eingesetzt werden können (z.B. finale oder konzessive). Zu denken ist dabei an Äußerungen wie (7) oder (8):

(7) *Wir sollten die staatliche Förderung von Elektromobilität verstärken, damit die verkehrsbedingte Umweltbelastung reduziert werden kann.*

(8) *Obwohl die staatliche Förderung von Elektromobilität eine zusätzliche finanzielle Belastung darstellt, sollten wir sie verstärken, damit die verkehrsbedingte Umweltbelastung reduziert werden kann.*

Die Vielfalt der sprachlichen Realisierungsvarianten für die Formulierung von Argumenten ist eine der zahlreichen Herausforderungen, die eine computergestützte Suche nach Argumenten bewältigen muss.

Eine weitere Schwierigkeit kommt hinzu, wenn man bedenkt, dass Argumentation kein Phänomen ist, das ausschließlich anhand von Oberflächenmerkmalen in Texten lokalisiert werden kann. Argumentation bedarf keineswegs – wie die Beispiele (6) bis (8) suggerieren – der Verknüpfung von Äußerungseinheiten durch Konjunktionen, die die Logik der Verknüpfung verdeutlichen. Dies zeigt der folgende Textausschnitt:

(9) »In den letzten Wochen ist in Aachen und Umgebung von einigen Leuten bezüglich der Wiedereinführung einer Stadtbahn viel Unruhe erzeugt worden. Man könnte meinen, die Welt geht unter, wenn die Stadtbahn kommt. Großbaustellen wie Berlin und Stuttgart 21 müssen dann erhalten, um diese Maßnahmen mit der Aachener Stadtbahn in Verbindung zu bringen. Dies ist doch völlig absurd. Da rechnen einige Leute aus, mit wie viel Euro jeder Einzelne von uns pro Jahr mehr belastet wird, wenn die Bahn kommt. Es wird aber nicht darüber nachgedacht, wie viel jeder von uns pro Jahr zahlen muss für die Milliarden Euro, die zur Stützung der Banken und nach Griechenland geflossen sind. Diese Gelder sind verloren. Wenn die Stadtbahn einmal fährt, kann sie uns keiner mehr wegnehmen und es ist eine gute Investition in die Zukunft.« (Leserbrief, Aachener Zeitung, 02.03.2013)

Bei Beispiel (9) besteht eine Leistung der Text-RezipientInnen darin, durch Inferenzziehung die Passagen des Textes zu erkennen, die die mit dem Text erhobenen Geltungsansprüche argumentativ absichern sollen. In realen Texten bzw. realen Argumentationen

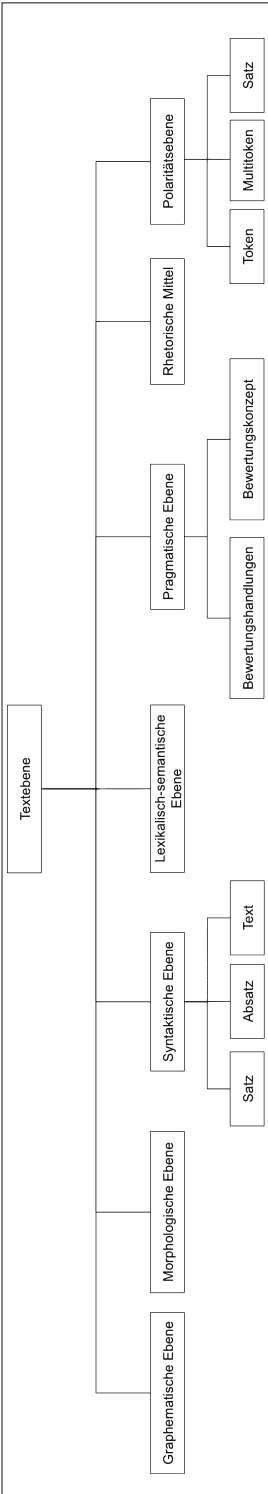


Abb. 1: Mehrebenen-Annotationsmodell von Trevisan (2014, S. 134)

bleibt vieles implizit, das auf gemeinsamem Weltwissen von SchreiberInnen/SprecherInnen und RezipientInnen beruht. Das als geteilt vorausgesetzte Weltwissen muss nicht immer wieder vollständig expliziert werden. Hohe Redundanz wäre sogar kontraproduktiv, da das Wiederholen nicht problematisierter Geltungsansprüche – die ständige Wiederholung allgemein akzeptierter Sachverhalte oder Bewertungen – zu Abwehrreaktionen der RezipientInnen führen kann.

So wird in den Beispielen (5) bis (8) stillschweigend vorausgesetzt, dass es sinnvoll und wünschenswert ist, die Umweltbelastung zu reduzieren. Dieser Geltungsanspruch ist in unserer Gesellschaft unumstritten und muss daher nicht ausformuliert werden. Dennoch ist er zweifelsohne impliziter Bestandteil der Argumentation.

Die Implizitheit von Begründungen erschwert das computergestützte Erkennen von Argumenten wesentlich. Wenn Äußerungen wie im Beispiel (9) keine oder nur rudimentäre Oberflächenmerkmale für Argumentationen enthalten und auf das gemeinsame Weltwissen von AutorInnen und RezipientInnen rekurrieren, stellt dies die computergestützte Argumentsuche vor ganz neue Herausforderungen.

6 Computergestützte Argumentanalyse

Disziplinen wie Informatik und Computerlinguistik bieten verschiedene Ansätze, die sich für die computergestützte Identifizierung von Argumenten nutzen lassen. Der hier beschriebene Ansatz präferiert Methoden des Linguistischen Text Mining, speziell des Argumentation Mining.

6.1 Text Mining

Text Mining ist eine Methodik, die aus der Informatik stammt. Sie wurde erstmals von Feldman und Dagan (1995) als *Knowledge Discovery in Textual Database (KDD)* eingeführt. Gegenstand der Methodik ist die (semi-)automatische Textanalyse schwach strukturierter Inhalte (E-Mails, Zeitungsartikel, Blogkommentare). Für

die Analyse großer Textkorpora werden Methoden des *Information Retrieval (IR)*, der *Information Extraction (IE)* und des *Natural Language Processing (NLP)* genutzt (Mehler/Wolff 2005, S. 1–10; Trevisan/Jakobs 2015, S. 168). Ziele ihrer Anwendung sind u.a. das Identifizieren von Schlüsselthemen, Trends und Diskursverläufen.

Neuere Ansätze des Text Mining verbinden Ansätze der Informatik mit linguistischen Theorien (*Linguistisches Text Mining*). In Trevisan (2014) z.B. wird ein theoretisch-methodischer Ansatz für die computergestützte Analyse sprachlichen Bewertens in Blogkommentaren entwickelt, der textmusterspezifische Phänomene berücksichtigt. Kern des Ansatzes ist ein linguistisches Mehrebenen-Annotationsmodell, das Merkmale sprachlichen Bewertens auf verschiedenen linguistischen Ebenen erfasst und annotiert. Das Modell umfasst sechs textbezogene Annotationsebenen (graphematische, morphologische, lexikalisch-semantische, syntaktische, pragmatische Ebene; Polarität) und 114 Annotationskategorien (vgl. Abbildung 1).

Auf der *graphematischen Ebene* werden graphostilistische bzw. typographische Mittel annotiert, die Bewertungshandlungen und Polaritäten anzeigen (z.B. Emoticon, Anführungszeichen, Iteration). Auf der *morphologischen Ebene* werden Tags zur Auszeichnung von Flexion und Wortbildung vergeben (z.B. Abkürzung, Akronym, Netzjargon-Ausdrücke, Lexeme mit e-Tilgung). Auf der *syntaktischen Ebene* werden textstrukturbezogene Phänomene annotiert (Anfang und Ende eines Satzes, Absatzes oder Textes). Auf der *lexikalisch-semantischen Ebene* werden themenbezogene Lexeme annotiert, die u.a. für die Erstellung themenspezifischer Lexika verwendet werden.

Auf der *pragmatischen Ebene* werden Bewertungshandlungen und Komponenten des Bewertungskonzepts annotiert. Bewertungshandlungen sind Sprechakte wie LOBEN, IRONISIEREN oder BEHAUPTEN. Komponenten des Bewertens erfassen nach Sandig (1979, S. 139-140; Sandig 1991, S. 226-228), z.B. Bewertungssubjekt, Bewertungsgegenstand oder Bewertungsaspekt. Auf der Ebene *rhetorische Mittel* werden metaphorische Formeln annotiert (z.B. Metapher, Allegorie oder Oxymoron). Sandig (1991) zufolge dienen rhetorische Mittel u.a. als Ausdruck der Maßstabserfüllung. Auf der Ebene *Polarität* werden Bewertungsrichtung-anzeigende Token (Lexeme), Mehrworteinheiten (Phrasen) und Sätze annotiert.

6.2 Argumentation Mining

Ein junger Bereich des Text Mining ist der anspruchsvolle Versuch, Methoden des Natural Language Processing für Zwecke der Argumentanalyse zu nutzen (Trevisan/Jakobs 2015, S. 181). Dieser Bereich wird als *Argumentation Mining* bezeichnet. Argumentation Mining-Verfahren wurden bislang primär für englischsprachige Texte entwickelt, zum Teil finden sich auch Ansätze für deutsche Texte.

Die meisten Forschungsansätze des Argumentation Mining ermitteln Argumente bezogen auf ausgewählte Textsorten, wie z.B. Rechtsschrift (Mochales-Palau/Moens 2009; Bach et al. 2013; Ashley/Walker 2013; Wyner et al. 2010), Online-Debatte (Cabrio/Villata

2012), Produktbewertung (Garcia-Villalba/Saint-Dizier 2012; Wyner et al. 2012), Zeitungsartikel und Mitschrift von Gerichtsverfahren (Feng/Hirst 2011), wissenschaftlicher Beitrag und Abstract (u.a. Liakata et al. 2012; Teufel 2010; Mizuta et al. 2005). Bislang gibt es nur wenige aufbereitete, d.h. annotierte Ressourcen, etwa Webkorpora oder heterogene Korpora (bestehend aus on- und offline verfügbaren Diskursfragmenten) (Reed et al. 2008).

Die (semi-)automatische Identifikation von Argumenten in Texten erfolgt musterbasiert, d.h. über die Ermittlung von Argumentmustern. *Argumentmuster* beschreiben Kombinationen häufig gemeinsam auftretender Mittel, die repräsentativ sind für das zu identifizierende Argument. Die Identifikation von Mustern erfolgt auf zwei Weisen: mit Machine Learning-Algorithmen oder mit Verfahren der Pattern Recognition. Eine häufig genutzte Methode ist das Training von *Machine Learning-Algorithmen* anhand von Merkmalen wie Signalwörtern (z.B. *weil*), Satzposition oder Mehrworteinheiten (z.B. *ich komme zum Schluss, dass*). Mochales-Palau und Moens (2009, S. 103) analysieren mit dieser Methode Argumente in juristischen Texten mit einer Genauigkeit von bis zu 80%. Die Argument-Identifizierung mit *Verfahren der Pattern Recognition* (Mustererkennung) basiert meist auf grammatik- und regelbasierten (z.B. in der Programmiersprache Prolog) oder hybriden Verfahren. Mochales-Palau und Moens (2009) beispielsweise kombinieren den Machine Learning-Ansatz mit einer kontextfreien Grammatik, um nicht nur die Position von Argumenten, sondern auch ihre Relationen zueinander zu analysieren. Sie erreichen dabei eine Genauigkeit von etwa 60%. Kang und Saint-Dizier (2014) entwickeln eine Prolog-basierte Sprache (*DISLOG*), die speziell der Analyse von Diskursstrukturen (hier: in Texten) dient.

Im deutschsprachigen Raum werden Verfahren der computergestützten Argumentanalyse erst seit kurzem entwickelt und dafür Ressourcen (Korpora, Lexika) aufgebaut. Ein umfangreiches annotiertes Korpus ist z.B. das *Potsdam Commentary Corpus* (Stede/Neumann 2014). Das Korpus ist auf unterschiedlichen linguistischen Ebenen annotiert und dient insbesondere der Erprobung bestehender Annotations- bzw. Analyseschemata und dem Training von Machine Learning-Algorithmen. Dumm und Lemke (2013, S. 3) erarbeiten ein Lexikon mit harten und weichen Argumentmarkern (»Signalwörter«), das für Zwecke der computergestützten Argumentanalyse genutzt werden kann. Harte Argumentmarker (bestimmte Konjunktionen, Präpositionen und Adverbien) ermöglichen es, in einer Vielzahl der Fälle argumentative Zusammenhänge zu erkennen (Grund – Folge). Weiche Argumentmarker sind dagegen Floskeln oder Mehrworteinheiten (z.B. *wie jeder weiß, Tatsache ist*).

7 Verschränkte Methodenentwicklung

Die in diesem Beitrag beschriebene Methodik kombiniert den Ansatz von Niehr (2004) mit Methoden des Linguistischen Text Mining für Zwecke der computergestützten Argumentanalyse. Die Methodenentwicklung wird im Folgenden beschrieben bezogen auf Schritte der Datenerhebung und -aufbereitung sowie der Datenanalyse.

7.1 Datenerhebung und -aufbereitung

Die Methodik wurde an einem themenzentrierten Korpus entwickelt und erprobt (Texte zu Infrastrukturentscheidungen). Als Fallbeispiel diente die öffentliche Kontroverse um die Einführung einer Straßenbahn (»Campusbahn«) in Aachen. In das Korpus gingen Artikel lokaler Printmedien, Online-Beiträge in Foren und sozialen Medien sowie mündliche und schriftliche Diskursbeiträge politischer Parteien, Experten und Bürgerinitiativen ein, in denen für bzw. gegen die Einführung der Campusbahn argumentiert wird. Die Texte wurden schlagwortbasiert erhoben (z.B. *Campusbahn*, *Elektrobahn*, *Straßenbahn*). Die Vorgehensweise kombinierte traditionelle und computergestützte Verfahren bei der Erhebung schriftlicher und mündlicher on- und offline zugänglicher Daten.

Die *offline zugänglichen Daten* umfassen schriftliche und mündliche Daten, die als Print vorliegen und für Zwecke des Projekts eingescannt wurden, archivierte digitalisierte Textdokumente sowie Audio-Mitschnitte (z.B. von Podiumsdiskussionen), die uns zur Verfügung gestellt wurden. Die Dateien wurden zum Teil aufbereitet: Die als pdf-Dateien vorliegenden Dokumente wurden für das Einlesen in MaxQDA und EXMARaLDA in txt-Dateien umgewandelt und manuell um Fehler bereinigt; die Audio-Dateien wurden partiell als Leseversion transkribiert.

Die *online zugänglichen Daten* umfassen schriftliche und mündliche Äußerungen im Internet. Sie wurden computergestützt recherchiert und extrahiert. Herausforderungen bei der Extraktion betrafen u.a. den unterschiedlichen Aufbau von Websites (z.B. Abfolge Newsartikel und -kommentar) und damit verbundene methodische Anpassungen. Parallel wurden online produzierte oder publizierte schriftliche Diskursbeiträge mit der Screenshotsoftware *FireShot* aufgezeichnet, um sie dauerhaft verfügbar zu machen.

Die erhobenen Daten entstanden im Zeitraum September 2010 bis August 2013; das Korpus umfasst mehr als 4.000 Diskursbeiträge, eine Gesamtübersicht bietet Tabelle 1. Die Speicherung der Daten erfolgte auf einem Server. Die Daten wurden manuell mit Metadaten (Quelle, Textsorte, Veröffentlichungsdatum, Dateiformat) ausgezeichnet und anhand dieser geordnet.

Für die Methodenentwicklung wurde aus dem Gesamtkorpus ein Trainingskorpus extrahiert. Es besteht aus 19 Webartikeln der Online-Ausgaben zweier lokaler Tageszeitungen in Aachen (Aachener Zeitung, AZ und Aachener Nachrichten, AN) sowie 333 Webkommentaren. Die Webkommentare beziehen sich auf die Webartikel und wurden – meist unter Verwendung eines Pseudonyms – von Autoren verfasst, die sich als Nutzer auf der Webseite der jeweiligen Lokalzeitung registrieren mussten.

Das Trainingskorpus wurde für die Analyse in EXMARaLDA (vgl. Kapitel 7.3) mit Methoden des *Natural Language Processing* aufbereitet:

1. Die Daten wurden automatisch tokenisiert (vgl. Abbildung 4), d.h. die Texte in durch Leerzeichen getrennte Einheiten bzw. Zeichenketten segmentiert.
2. Die Token wurden automatisch mit morphosyntaktischen Kategorien versehen (Part-of-Speech-Tagging mittels WebTagger, Genauigkeit: 94,8%; Neunerdt/Reyer/Mathar 2014) und lemmatisiert, d.h. flektierte Begriffe (z.B. *geht*) auf ihre lexikalische Grundform zurückgeführt (z.B. *gehen*). Der WebTagger – eine Weiterentwicklung des Stan-

| | Online zugänglich | | Offline zugänglich | | Σ |
|-------------------|-------------------|----------|--------------------|----------|--------|
| | schriftlich | mündlich | schriftlich | mündlich | |
| Anzeige | | | X | | 6 |
| Blogbeitrag | X | | | | 15 |
| Flyer | X | | X | | 6 |
| Forumsbeitrag | X | | | | 380 |
| Interview | | X | | | 4 |
| Leserbrief | | | X | | 524 |
| Newsgroup-eintrag | X | | | | 1 |
| Plakat | X | | | | 8 |
| Podiumsdiskussion | | | | X | 1 |
| Positionspapier | X | | X | | 30 |
| Sitzungsprotokoll | X | | | | 4 |
| Webartikel | X | | | | 99 |
| Webkommentar | X | | | | 2913 |
| Zitat | | | X | | 130 |
| Zeitungsartikel | | | X | | 323 |
| | | | | | Σ 4444 |

Tab. 1: Zusammensetzung des Gesamtkorpus

dard-Taggers *TreeTagger* (Schmid 1995) für das Tagging von Webtexten – taggt die tokenisierten Primärdaten (Texte) nach dem *Stuttgart-Tübingen-TagSet* (STTS) (Schiller/Teufel/Stöckert 1999). Das Tagset bzw. Kategoriensystem umfasst 54 Kategorien: 48 Tags für die Hauptwortarten (Nomen, Verb, Artikel, Adjektiv, Pronomen, Kardinalzahl, Adverb, Konjunktion, Adposition, Interjektion, Partikel) und ihre Subklassen (z.B. finites Verb, adverbiales Adjektiv) sowie sechs Tags zur Auszeichnung von Nichtwörtern, fremdsprachlichem Material und Satzzeichen.

- Die tokenisierten, getaggtten und lemmatisierten Daten wurden manuell korrigiert. Die Korrektur betrifft falsch vergebene Tags (z.B. finites Vollverb/VVFIN statt normales Nomen/NN) und die Ergänzung fehlender Lemmata, die in den Daten als »unknown« klassifiziert wurden. Häufig handelt es sich bei diesen Lemmata um themenspezifische Begriffe, die in den Tagger-eigenen, themenunspezifischen Lexika nicht enthalten und damit unbekannt sind (z.B. *Campusbahn*).

Das oben beschriebene Verfahren der Datenaufbereitung beschreibt ein Standardverfahren der computergestützten Textanalyse und gilt als Vorbedingung für weitergehende Analysen wie die linguistische Mehrebenen-Annotation (Trevisan/Neunerdt/Jakobs 2012, S. 182).

7.2 Das Kategoriensystem

Für die Argumentanalyse des Trainingskorpus wurde das diskurslinguistische Analyseverfahren nach Niehr (2004) mit dem in Trevisan (2014) entwickelten linguistischen Ansatz des Text Minings zusammengeführt. Der Ansatz basiert auf dem Kategoriensystem von Niehr (2004), das durch die linguistischen Ebenen des Annotationsansatzes von Trevisan (2014) erweitert wurde. Das Ergebnis wurde durch die Kategorie *Argumentträger* ergänzt (vgl. Abbildung 2); sie ermöglicht die Zuordnung der im Korpus genannten Argumente zu ihren jeweiligen TrägerInnen.

Das Grundgerüst des Kategoriensystems bilden drei Kategorien: Argumentträger, Verwendungsweise und linguistische Ebene. Jede Kategorie umfasst Unterkategorien ersten und zweiten Grades, anhand derer das im Text identifizierte Argument näher beschrieben wird. Die Ausprägung von Kategorien variiert textsortenspezifisch (Beispiel graphematische Ebene: die Verwendung von Emoticons ist typisch für Blogkommentare, nicht jedoch für Zeitungsartikel).

Die Kategorie *Argumentträger* dient der Identifizierung der am Diskurs beteiligten SprecherInnen. Es werden zwei Typen unterschieden: Argumentträger, die im Diskurs ein Argument vorbringen (Subj. = Subjekt) versus Argumentträger, denen im Diskurs ein Argument zugeschrieben wird (Obj. = Objekt). Für jeden identifizierten Argumentträger bzw. jede identifizierte Argumentträgerin werden textbasiert Argumentträgereigenschaften ermittelt, z.B. anhand der Nennung ihres Geschlechts oder Alters in dem jeweils untersuchten Textexemplar. Von besonderem Interesse ist im Fallbeispiel die Bestimmung seiner bzw. ihrer gesellschaftlichen Rolle (z.B. PolitikerIn, UmweltaktivistIn). Auf diese Weise lassen sich rollen- bzw. akteurspezifische Argumentationsweisen identifizieren und vergleichen.

Die Kategorie *Verwendungsweise* beschreibt nach Niehr (2004) die Art und Weise, in der ein Argument vorgebracht wird (vgl. Kapitel 5): argumentativ (das Argument wird vom Argumentträger produziert) oder zitierend (ein bereits in den Diskurs eingebrachtes Argument wird von einem Sprecher bzw. einer Sprecherin reproduziert). Anhand der Art der zitierenden Reproduktion (positiv, neutral und negativ) können Bewertungen des Arguments durch den Sprecher bzw. die Sprecherin rekonstruiert werden.

Die Kategorie *linguistische Ebene* erlaubt die Erfassung und Beschreibung der sprachlichen Mittel, mit denen ein Argument realisiert wird. Aus Trevisan (2014) wurden die graphematische und die lexikalisch-semantische Ebene übernommen. Einige Ebenen ihres ursprünglichen Mehrebenen-Annotationssystems wurden für Zwecke des Fallbeispiels weiterentwickelt bzw. präzisiert und umbenannt: die Analyseebenen Sprechakt (*bisher*: pragmatische Ebene), Position (*bisher*: syntaktische Ebene) und Bewertungsrich-

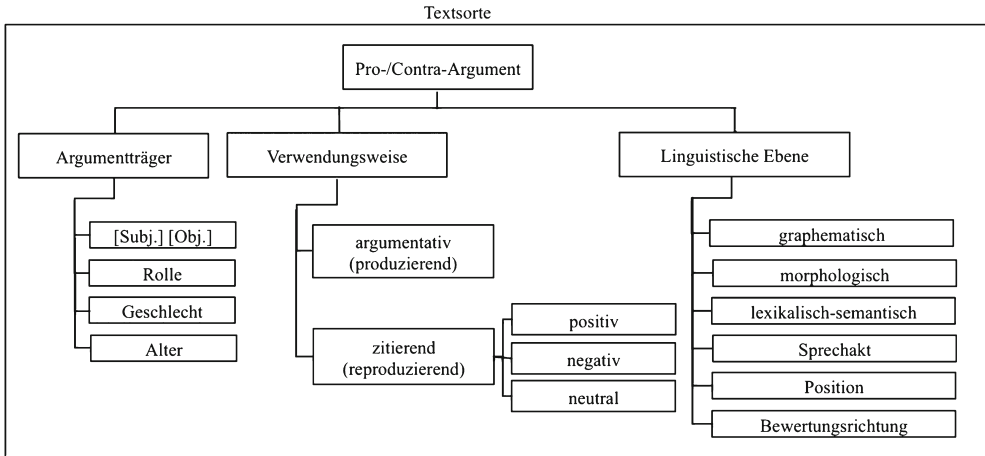


Abb. 2: Integriertes Kategoriensystem

tion (*bisher*: Polarität). Die morphologische Ebene wurde durch folgende Unterkategorien ergänzt: Argumentmarker nach Dumm und Lemke (2013) sowie Kommunikationsverben nach Harras (2007). Die Idee, auf Harras’ Aufstellung von Kommunikationsverben zurückzugreifen, gründete auf dem Problem, Indikatoren für die Identifizierung von Argument-Verwendungsweisen zu finden, die sich für eine computergestützte Analyse eignen. Kommunikationsverben wie *sagen*, *erläutern* und *nennen* leiten häufig die Wiedergabe von Argumenten in Texten ein und eignen sich daher für die computergestützte Bestimmung analyserelevanter Textpassagen. Kommunikationsverben können darüber hinaus die Art der Verwendungsweise eines Arguments durch den Sprecher bzw. die Sprecherin indizieren (z.B. positiv: *zustimmen*, negativ: *bestreiten* oder *anzweifeln*).

Wichtig für den hier beschriebenen Ansatz ist die Berücksichtigung der Textsortenspezifisch erhobener Textdaten. Die sprachliche Realisierung eines Arguments kann textsortenabhängig (z.B. Zeitungskommentar vs. News-Kommentar) variieren.

7.3 Integrative Analysemethodik

Die Methodenentwicklung richtet sich auf das Ziel, sprachliche Muster zu identifizieren, die bestimmte Verwendungsweisen von Argumenten indizieren. Die dafür entwickelte Analysemethodik kombiniert vier Schritte: (1) getrennte Datenanalyse (diskurslinguistische und computergestützte Argumentanalyse), (2) Aggregation, (3) Metaanalyse (Korpus-Statistiken und Mustererkennung) und (4) Musterdefinition. Die Methodik integriert manuelle, automatische und semi-automatische Methoden.

(1) *Datenanalyse*: Die Datenanalyse erfolgt parallel mit zwei Methoden: manuell-qualitativ durch eine diskurslinguistische Argumentanalyse, semi-automatisch-quantitativ durch eine computergestützte Argumentanalyse.

Diskurslinguistische Argumentanalyse: Die manuelle diskurslinguistische Argumentanalyse folgt dem in Niehr (2004) entwickelten Verfahren. Die Auszeichnung der Argumente erfolgte in MaxQDA. In der Analyse wurden diskursrelevante Argumente zum Thema »Für und wider der Einführung einer Straßenbahn« identifiziert, basierend auf der Annotation Argumentklassen gebildet und die Verwendungsweise der Argumente bestimmt. Die Codierung erfolgte durch zwei unabhängige CodiererInnen; für die finale Codierung wurden die Ergebnisse verglichen (*Intercoder-Agreement*).

Computergestützte Argumentanalyse: Parallel zur qualitativen Argumentanalyse wurden die Texte des Trainingskorpus mit computergestützten Methoden semi-automatisch analysiert. Die Analyse fokussiert die in Kapitel 7.2 beschriebenen linguistischen Ebenen. Sie erfolgte in EXMARaLDA, da das Tool eine Daten-Annotation über mehrere Ebenen erlaubt (*Mehrebenen-Annotation*).

Für die Mehrebenen-Annotation wurden die aufbereiteten Textdaten in EXMARaLDA importiert (Importformat: .txt). Die Annotation der graphematischen, morphologischen und lexikalisch-semantischen Ebene sowie der Analyseebenen Position und Bewertungsrichtung erfolgte automatisch mittels AutoAnnotator (Trevisan et al. 2014b, Trevisan/Neunerdt 2015/in Vorbereitung, Genauigkeit: 95,1%); die Annotationsebenen wurden in EXMARaLDA ebenfalls automatisch durch den AutoAnnotator angelegt. In einem nachfolgenden Schritt wurden Annotationsfehler manuell identifiziert und korrigiert. Die Ebene Sprechakt wurden ausschließlich manuell annotiert, da sie interpretativ ist und mit computergestützten Verfahren nur eingeschränkt annotiert werden kann. Der Ansatz erlaubt die Annotation und Betrachtung sprachlicher argument-indizierender Mittel. Durch die Mehrebenen-Annotation wird der Tatsache Rechnung getragen, dass jedes Token bzw. jede Texteinheit spezifische Funktionen erfüllt, die auch für die Konstitution evaluativer Aussagen und Argumente relevant sind (z.B. *nur* als abschwächendes Element evaluativer Aussagen, *aber* als Argumentmarker).

(2) *Aggregation:* Die Ergebnisse der manuellen und semi-automatischen Analysen (Argumente und Argument-Verwendungsweisen, Argumentklassen, Mehrebenen-Annotation) wurden in EXMARaLDA zusammengeführt (aggregiert). Dies erforderte u.a., die Annotationsergebnisse der manuellen diskurslinguistischen Argumentanalyse manuell zu übertragen. Für die Übertragung wurden die Analyseebenen Verwendungsweise sowie Pro-/Contra-Argument manuell angelegt. Abbildung 3 zeigt die aus dem Kategoriensystem (Abbildung 2) übernommenen Ebenen (graphematische Ebene, lexikalisch-semantische Ebene, Sprechakt, Bewertungsrichtung) sowie Unterebenen (Partikel, Kommunikationsverben, Argumentmarker, Position_Satz), die nicht im Kategoriensystem dargestellt werden (vgl. Abbildung 3). EXMARaLDA bietet für die Anlage von Kategorien und Unterkategorien nur eine Spalte, sie erscheinen daher zusammen »gelistet«.

(3) *Metaanalyse:* Die Analyseergebnisse des ersten Schritts – Argumente, Verwendungsweisen und linguistische Annotation – bilden die Ausgangsdaten für den zweiten Schritt, die Metaanalyse, in der eine Korpus-Statistik für Zwecke der Mustererkennung erstellt und genutzt wird.

Korpus-Statistik: Das Mehrebenen-annotierte Trainingskorpus wurde quantitativ untersucht (relatives versus absolutes Vorkommen eines Mittels). Ermittelt wurden häufig

| | 16 | 317 | 318 | 319 | 320 | 321 | 322 | 323 | 324 | 325 | 326 | 327 | 328 | 329 | 330 | 331 | 332 |
|-------------------------------|----|-------------------|-------|------|-----|------------|-----|------------|-----|------|----------|-------|------|-------------|-----------|--------|-------|
| X [txt] | | Ich | bin | für | die | Einführung | der | Campusbahn | , | weil | die | Stadt | vom | Autoverkehr | entlastet | wird | |
| X [pos] | | PPER | VAFIN | APPR | ART | NN | ART | NN | \$, | KOUS | ART | NN | APPR | ART | NN | VVPP | VAFIN |
| X [lemma] | | ich | sein | für | d | Einführung | d | Campusbahn | , | weil | d | Stadt | vom | Autoverkehr | entlasten | werden | |
| ProArgument [Nachhaltigkeit] | | Nachhaltigkeit | | | | | | | | | | | | | | | |
| Verwendungsweise | | neutral zitierend | | | | | | | | | | | | | | | |
| Graphematische Ebene | | MARK | | | | | | | | | | | | | | | |
| Partikel | | | | | | | | | | | | | | | | | |
| Kommunikationsverben | | | | | | | | | | | | | | | | | |
| Argumentmarker | | | | | | | | | | | | | | | | | |
| Lexikalisch-semantische Ebene | | | | | | | | | | | | | | | | | |
| Sprechakt | | | | | | | | | | | | | | | | | |
| Position_Satz | E | SA | | | | | | | | | | | | | | | |
| Bewertungsrichtung | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | KONNEKTP | | | | | | |
| | | | | | | | | | | | AM | | | | | | |
| | | | | | | | | | VEH | | | | | | EL | | |
| | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | + |

Abb. 3: Mehrebenen-annotierter Korpusausschnitt
 Legende der Tag-Bezeichnungen: X [txt]: Originaltext; X [pos]: morphosyntaktische Kategorie; X [lemma]: Grundform; MARK: Markierung; KONNEKTP: Konnektivpartikel; AM: Argumentmarker; VEH: Transportmittel; EL: themenspezifisches Einzelexem; SA: Satzanzfang; +: positiv bewertendes Token).

auf tretende Mittel der linguistischen Ebenen (z.B. Interjektion, Argumentmarker, Hervorhebung durch Anführungszeichen), die mit dem Auftreten produktiver und re produktiver Verwendungsweisen von Argumenten kookkurrieren.

Mustererkennung: Das Trainingskorpus wurde – neben der Erstellung der Korpus-Statistiken – nach häufig auftretenden Tag-Kombinationen oder Mustern durchsucht, die mit dem Auftreten produktiver und reproduktiver Verwendungsweisen von Argumenten einhergehen, z.B. Tagkombinationen wie [+ , ^] [EL, VEH] (+: positiv bewertendes Token; ^: Intensivierer; EL: themenspezifisches Einzelexem; VEH: Transportmittel). Die Grundidee ist, dass das Zusammenspiel und die Kombination verschiedener sprachlicher Mittel ein Argument sowie seine Verwendungsweise indiziert.

In einem ersten Schritt wurde das Trainingskorpus explorativ nach Mustern durchsucht, d.h. die Suche basiert nicht auf vordefinierten Suchparametern (z.B. häufig auftretende Tag-Kombinationen), sondern verläuft ungesteuert. Ziel und Zweck der Suche ist die (ungezielte) Ermittlung von Tagkombinationen innerhalb von Sätzen, die Muster für Verwendungsweisen und Argumente repräsentieren. Diese Muster können mehr oder weniger komplex ausfallen.

In einem zweiten Schritt wurden alle Muster, die sich einer Argument-Verwendungsweise zuordnen lassen (z.B. negativ zitierend)

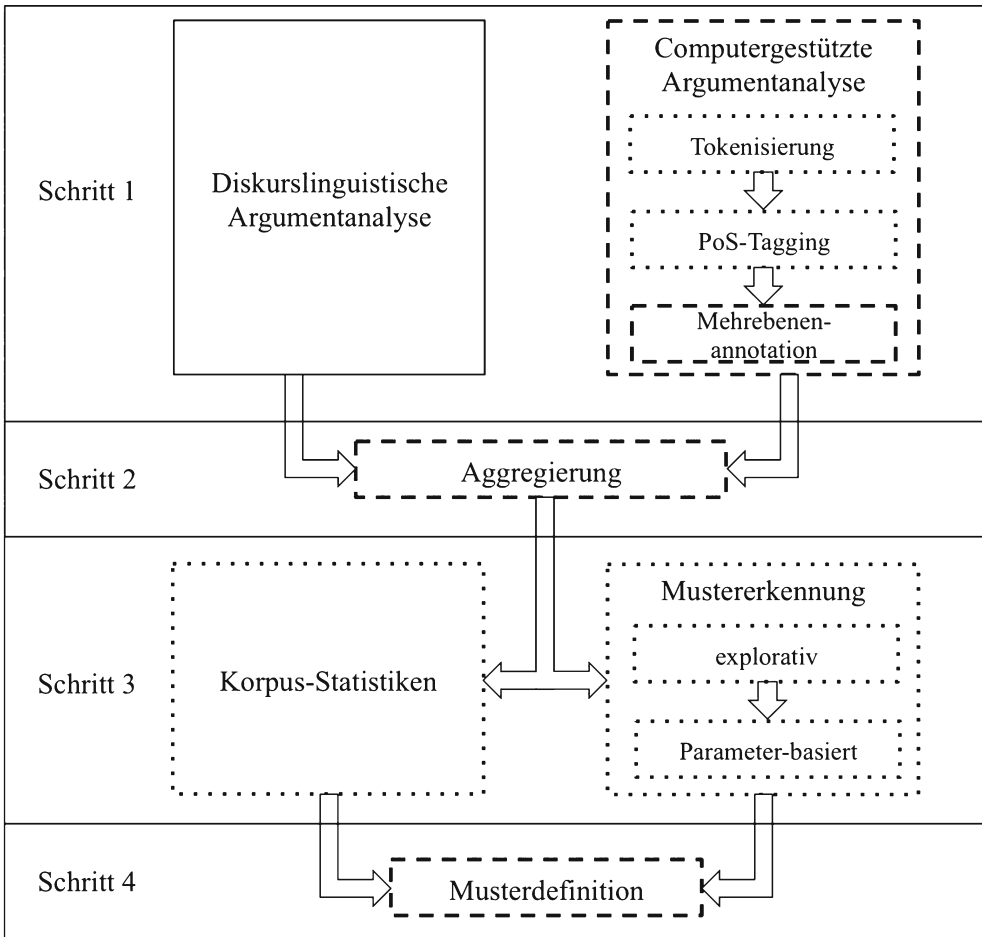


Abb. 4: Integrative Analysemethodik.

Legende: manuelle Methoden: durchgezogene Linie; automatische Methoden: gepunktete Linie; semi-automatische Methoden: gestrichelte Linie.

auf Gemeinsamkeiten und Unterschiede untersucht und solcherart die »Basiskomponenten« von Mustern einer Verwendungsweise ermittelt.

(4) *Musterdefinition*: Die Basiskomponenten wurden hinsichtlich ihrer »Repräsentativität« für Verwendungsweisen geprüft. Prüfkriterium war die Auftretenshäufigkeit in der Korpus-Statistik. Die identifizierten Muster wurden auf relevante Muster beschränkt und nur diese definiert. Das Ergebnis ist die Festlegung relevanter, argument-indizierender Tagkombinationen. Abbildung 4 zeigt die integrative Methodik im Überblick.

7.4 Ergebnisse

Die oben beschriebene Methodik wurde auf das Trainingskorpus angewandt mit dem Ziel, Muster zu identifizieren, die bestimmte Verwendungsweisen von Argumenten indizieren. Die statistische Auswertung liefert Aussagen zu folgenden Punkten: Auftretenshäufigkeit von Tags und Argumenten (Korpus-Statistiken, Abbildung 5), Anzahl von Mustern je Argument-Verwendungsweise (Tabelle 2), häufig auftretende Tagkombinationen bzw. Muster je Verwendungsweise (Tabelle 3). Die Ergebnisse der statistischen Auswertung werden im Folgenden vorgestellt.

Befunde zur Auftretenshäufigkeit von Tags und Argumenten

Die Auftretenshäufigkeit liefert Hinweise darauf, welche Kategorien bei der automatischen Mustererkennung relevant sein könnten. Die Darstellung der Korpus-Statistik in Abbildung 5 zeigt die absolute Auftretenshäufigkeit von Tags per aggregierter Annotationsebene. Mit Ausnahme der Statistik Pro-/Contra-Argumente wurden für alle anderen Ebenen die fünf am häufigsten auftretenden Tags ermittelt (Top 5). Im Falle der lexikalisch-semantischen Ebene war dies nicht möglich – die Auswertung ergab lediglich zwei häufig auftretende Tags.

Die Verteilung von Tags variiert je nach Annotationsebene. Einige Ebenen zeigen tendenziell eine ähnliche Verteilung häufig auftretender Tags, z.B. bei Pro-/Contra-Argumenten und Argumentmarkern. Die Verteilung auf anderen Ebenen indiziert dagegen starke Abweichungen, z.B. im Falle der Pos-Ebene oder der Unterebene musterindizierender Kommunikationsverben.

Anzahl von Mustern je Argument-Verwendungsweise

Tabelle 2 bietet einen Überblick, in wie vielen Mustern eine bestimmte Argument-Verwendungsweise im Trainingskorpus realisiert wird.

| Verwendungsweise | n |
|--------------------------|-----|
| argumentativ (ARG) | 512 |
| negativ zitierend (NEGZ) | 478 |
| neutral zitierend (NEUZ) | 354 |
| positiv zitierend (POSZ) | 95 |

Tab. 2: Anzahl identifizierter Muster je Verwendungsweise

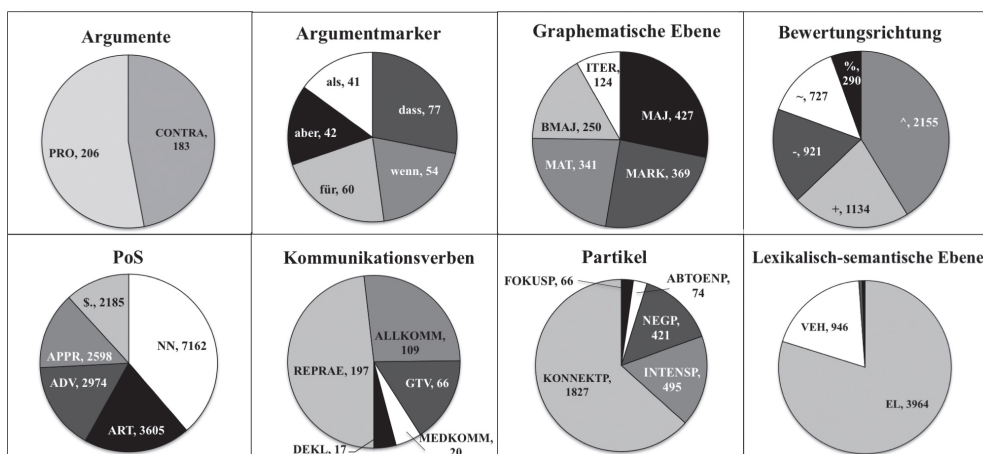


Abb. 5: Korpus-Statistiken.

Legende der Abkürzungen geordnet nach Ebenen: Argument-Ebene – PRO: Pro-Argument, CONTRA: Contra-Argument; graphematische Ebene – ITER: Iteration, MAJ: Majuskelschreibung, MARK: Markierung, MAT: mathematisches Zeichen, BMAJ: Binnenmajuskel; Ebene der Bewertungsrichtung – ^: Intensivierer, +: positiv bewertendes Token, -: negativ bewertendes Token, : Umkehrer, %: Verminderer; PoS-Ebene – NN: normales Nomen, ART: Artikel, ADV: Adverb, APPR: Präposition, \$: satzbeendende Interpunktion; Ebene der Kommunikationsverben – ALLKOMM: allgemeine Kommunikationsverben, GTV: Gesprächs- und themenstrukturierende Verben, MEDKOMM: mediale Kommunikationsverben, DEKL: Deklarative, REPRAE: Repräsentative; Partikel-Ebene – ABTOENP: Abtönungspartikel, INTENSP: Intensivierungspartikel, KONNEKT: Konnektivpartikel, FOKUSP: Fokuspartikel, NEGP: Negationspartikel; lexikalisch-semantische Ebene – VEH: Transportmittel, EL: themenspezifisches Einzellexem.

Die meisten Muster sind Repräsentationsformen argumentativer Verwendungsweisen (n=512). An zweiter Stelle folgen Muster für negativ zitierende (n=478), an dritter Stelle neutral zitierende Verwendungsweisen (n=354). Am geringsten ist der Anteil von Mustern für positiv zitierende Verwendungsweisen (n=95).

Die identifizierten Muster sind mehr oder weniger komplex; sie umfassen partiell mehr als zehn Beschreibungskategorien bzw. Tags. Je komplexer ein Muster ist, desto seltener tritt es auf. Die Komplexität des Musters scheint damit ein Indikator für die Repräsentativität eines Musters. Für die Charakterisierung von Korpora interessieren insbesondere relevante, d.h. repräsentative Muster und ihre Beschaffenheit.

Häufig auftretende Tagkombinationen bzw. Muster je Verwendungsweise

Es konnten acht Muster ermittelt werden, die in unterschiedlichem Maße (mehr oder weniger typisch) Argument-Verwendungsweisen des Trainingskorpus charakterisieren (Tabelle 3).

| Muster | Verwendungsweise | | | |
|---------------------|------------------|------|------|------|
| | ARG | NEGZ | POSZ | NEUZ |
| [+, ^, ~] [EL, VEH] | 20 | 11 | 8 | 0 |
| [+, ^, ~] | 71 | 136 | 20 | 0 |
| [EL, VEH] | 133 | 116 | 23 | 74 |
| [+, ^] [EL, VEH] | 32 | 13 | 8 | 12 |
| [+] [EL, VEH] | 42 | 18 | 8 | 23 |
| [LOC, VEH] | 95 | 47 | 23 | 6 |
| [ITER, MARK] | 21 | 10 | 0 | 0 |
| [ADJA, NN] | 282 | 239 | 26 | 179 |

Tab. 3: Häufig auftretende Muster je Verwendungsweise

Die Ergebnisse zeigen, dass die Anzahl von Mustervorkommen je Verwendungsweise stark variiert. Das Muster $[+, ^, \sim]$ (+: positiv bewertendes Token, ^: Intensivierer, ~: Umkehrer; vgl. Beispiel 10) tritt beispielsweise häufig bei negativen zitierenden Verwendungsweisen auf ($n=136$), wo hingegen das Muster $[EL, VEH]$ (EL: themenspezifisches Einzellexem, VEH: Transportmittel; vgl. Beispiel 11) insbesondere im Zusammenhang mit argumentativer Verwendungsweise zu finden ist ($n=133$).

(10) »Sie hält Elektrobusse für wesentlich_[^] flexibler_[+], auch glaubt sie nicht_[~], dass jede_[^] Straße für eine Bahntrasse geeignet_[+] sei.« (Webkommentar, Aachener Nachrichten, 04.03.2013)

(11) »Die Fahrzeiten_[EL] belegen, dass die Fahrtzeit_[EL] aus dem Südraum durch die Campusbahn_[VEH] schneller werden - trotz umsteigen - und die Fahrt_[EL] komfortabler.« (Webkommentar, Aachener Zeitung, 14.01.2013)

Fünf der acht ermittelten Muster treten bei allen Argument-Verwendungsweisen auf; die restlichen drei Muster werden dagegen nur für bestimmte Verwendungsweisen genutzt – sie können im Umkehrschluss bezogen auf das Trainingskorpus als Ausschlusskriterien für spezifische Argument-Verwendungsweisen interpretiert werden.

8 Fazit und Ausblick

Die oben dargestellte integrative Methodik hat erhebliches Potential für die Analyse großer Korpora. Sie erlaubt in der Verbindung manueller und computergestützter Verfahren die Identifizierung von Argumenten und Verwendungsweisen von Argumenten, die Analyse und statistische Auswertung von Mitteln für die Realisierung von Argumenten sowie die Ermittlung mehr oder weniger typischer Argumentmuster. Die Leistungsfähigkeit bzw. Performanz der Methodik ist an weiteren Korpora zu testen. Die definierten Muster sind anhand von Validierungskorpora auf ihre Gültigkeit zu überprüfen. Es gibt eine ganze Reihe bisher unbeantworteter Fragen, die erheblichen Forschungsbedarf andeuten, wie zum Beispiel die Fragen, welche Parameter sich – über die bisher genannten hinaus – auf die Nutzung argument-indizierender Mittel auswirken oder wie groß die Varianz von Argumentmustern ist. Insgesamt ist der Ansatz in vielfältiger Weise ausbaubar. Was bislang u. a. fehlt, sind Verfahren für die Zuordnung von Argumenten zu Argumentträgern als Voraussetzung für die computergestützte Ermittlung akteurspezifischer Argumentationsweisen. Der Ansatz stößt u. a. dann an Grenzen, wenn es um die Ermittlung impliziter Argumente geht. Er macht aber auch deutlich, dass die Beschränkung auf entweder qualitative oder aber quantitative Methoden nicht zielführend ist. Im Gegenteil: Ihre Verbindung und partiell gegenseitige Modifikation ermöglicht vielfältige Ansatzpunkte für eine veränderte Qualität diskurslinguistischer Studien und neue Fragestellungen.

Literatur

- Apel, K.-O. (1999): Transformation der Philosophie. Band 2: Das Apriori der Kommunikationsgemeinschaft. Frankfurt am Main: Suhrkamp.
- Ashley, K. D./Walker, V. R. (2013): Toward Constructing Evidence-based Legal Arguments Using Legal Decision Documents and Machine Learning. In: Proceedings of the Fourteenth International Conference on Artificial Intelligence and Law (ICAIL ,13, ACM), S. 176–180.
- Bach, N. X./Le Minh, N./Oanh, T. T./Shimazu, A. (2013): A Two-Phase Framework for Learning Logical Structures of Paragraphs in Legal Articles. In: ACM Transactions on Asian Language Information Processing (TALIP) 1 (12), S. 1–32.
- Bubenhof, N. (2008): Diskurse berechnen? Wege zu einer korpuslinguistischen Diskursanalyse. In: Warnke, I./Spitzmüller, J. (Hrsg.): Methoden der Diskurslinguistik. Sprachwissenschaftliche Zugänge zur transtextuellen Ebene. Berlin und New York: de Gruyter, S. 407–434.
- Bubenhof, N. (2013): Quantitativ informierte qualitative Diskursanalyse. Korpuslinguistische Zugänge zu Einzeltexten und Serien. In: Roth, K. S./Spiegel, C. (Hrsg.): Angewandte Diskurslinguistik. Felder, Probleme, Perspektiven. Berlin: Akademie-Verlag, S. 109–134.
- Bubenhof, N./Scharloth, J. (2013): Korpuslinguistische Diskursanalyse. Der Nutzen empirisch-quantitativer Verfahren. In: Meinhof, U./Reisigl, M./Warnke, I. H. (Hrsg.): Diskurslinguistik im Spannungsfeld von Deskription und Kritik. Berlin: Akademie Verlag, S. 147–169.
- Cabrio, E./Villata, S. (2012): Natural Language Arguments: A Combined Approach. In: De Raedt, L./Bessiere, C./Dubois, D./Doherty, P./Frasconi, P./Heintz, F./Lucas, P. J. F. (Hrsg.): ECAI, Frontiers in Artificial Intelligence and Applications 242. IOS Press, S. 205–210.

- Dumm, S./Lemke, M. (2013): Argumentmarker. Definition, Generierung und Anwendung im Rahmen eines semi-automatischen Dokument-Retrieval-Verfahrens. Schriftenreihe des Verbundprojektes »Postdemokratie und Neoliberalismus«, Diskussion-Paper 3
- Feldman, R./Dagan, I. (1995): Knowledge Discovery in Textual Databases (KDT). In: Proceedings of the First International Conference on Knowledge Discovery and Data Mining (KDD-95), S. 112–117.
- Feng, V. W./Hirst, G. (2011): Classifying Arguments by Scheme. In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies 1 (HLT ,11, ACL), S. 987–996.
- García-Villalba, M. P./Saint-Dizier, P. (2012): A Framework to Extract Arguments in Opinion Texts. In: International Journal of Cognitive Informatics and Natural Intelligence 3 (6), S. 62–87.
- Habermas, J. (1981): Theorie des kommunikativen Handelns. Band 1: Handlungsrationalität und gesellschaftliche Rationalisierung. Frankfurt am Main: Suhrkamp.
- Kang, J./Saint-Dizier, P. (2014): Argument Mining in TextCoop. In: Workshop on Argument Mining, Dundee.
- Kienpointner, M. (1992): Alltagslogik. Struktur und Funktion von Argumentationsmustern. Stuttgart-Bad Cannstatt: Frommann-Holzboog.
- Klein, W. (1980): Argumentation und Argument. In: Zeitschrift für Literaturwissenschaft und Linguistik (38-39), S. 9–57.
- Kopperschmidt, J. (1989): Methodik der Argumentationsanalyse. Stuttgart-Bad Cannstatt: Frommann-Holzboog.
- Kopperschmidt, J. (2005): Argumentationstheorie zur Einführung. Hamburg: Junius.
- Liakata, M./Saha, S./Dobnik, S./Batchelor, C. R./Rebholz-Schuhmann, D. (2012): Automatic Recognition of Conceptualization Zones in Scientific Articles and TwoLife Science Applications. In: Bioinformatics 7 (28), S. 991–1000.
- Mehler, A./Wolff, C. (2005): Einleitung: Perspektiven und Positionen des Text Mining. In: Zeitschrift für Computerlinguistik und Sprachtechnologie: Themenschwerpunkt Text Mining 1 (20), S. 1–18.
- Mizuta, Y./Collier, N. (2005): Zone Analysis in Biology Articles as a Basis for Information Extraction. In: Proceedings of the International Joint Workshop on Natural Language Processing in Biomedicine and its Applications (JNLPBA ,04, ACL), S. 29–35.
- Mochales Palau, R./Moens, M.-F. (2009): Argumentation Mining: The Detection, Classification and Structure of Arguments in Text. In: Twelfth International Conference on Artificial Intelligence and Law (ICAIL '09, ACM), S. 98–109.
- Neunerdt, M./Reyer, M./Mathar, R. (2014): Efficient Training Data Enrichment and Unknown Token Handling for POS Tagging of Non-standardized Texts. In: Proceedings of the Twelfth Conference on Natural Language Processing (KONVENS), 08.-10.10.2014, Hildesheim (D).
- Niehr, Th. (2004): Der Streit um Migration in der Bundesrepublik Deutschland, der Schweiz und Österreich. Eine vergleichende diskursgeschichtliche Untersuchung. Heidelberg: Winter.
- Niehr, Th. (2014): Einführung in die linguistische Diskursanalyse. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Niehr, Th. (2015): Die Universalität im öffentlichen Sprachgebrauch. Ein Plädoyer für das Zusammenwirken von quantitativen und qualitativen Methoden der Diskursforschung. In: Roth, K. S./Spitzmüller, J./Arendt, B./Kiesendahl, J. (Hrsg.): Sprache, Universität, Öffentlichkeit. Bremen: Hempen, S. 134–146.
- Niehr, Th. (i.Vb.): Argumentation in Texten. Erscheint in: Roth, K. S./Wengeler, M./Ziem, A. (Hrsg.): Handbuch Sprache in Politik und Gesellschaft. Berlin und Boston: de Gruyter.
- Quine, W. V. O. (1974): Grundzüge der Logik. Frankfurt am Main: Suhrkamp.
- Reed, C./Mochales Palau, R./Rowe, G./Moens, M.-F. (2008): Language Resources for Studying Argument. In: Proceedings of the 6th Conference on Language Resources and Evaluation (LREC '08), S. 91–100.
- Sandig, B. (1979): Ausdrucksmöglichkeiten des Bewertens. Ein Beschreibungsrahmen im Zusammenhang eines fiktionalen Textes. In: Deutsche Sprache 7, S. 137–159.

- Sandig, B. (1991): Formeln des Bewertens. In: Palm, C. (Hrsg.): »EUROPHRAS 90«. Akten der internationalen Tagung zur germanistischen Phraseologieforschung. Aske, Schweden 12.-15. Juni 1990. Stockholm: Almqvist & Wiksell international, S. 227–252.
- Scharloth, J./Eugster, D./Bubenhöfer, N. (2013): Das Wuchern der Rhizome. Linguistische Diskursanalyse und Data-driven Turn. In: Busse, D./Teubert, W. (Hrsg.): Linguistische Diskursanalyse: neue Perspektiven. Wiesbaden: VS, S. 345–381.
- Schiller, A./Teufel, S./Stöckert, C. (1999): Guidelines für das Tagging deutscher Textkorpora mit STTS. Technischer Bericht. Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart und Seminar für Sprachwissenschaft, Universität Tübingen.
- Schmid, H. (1995): Improvements in Part-of-Speech Tagging with an Application to German. In: Proceedings of the EACL SIGDAT Workshop, Dublin (IRL).
- Stede, M./Neumann, A. (2014): Potsdam Commentary Corpus 2.0: Annotation for Discourse Research. In: Proceedings of the Language Resources and Evaluation Conference (LREC), Reykjavik.
- Teufel, S. (2010): Recognising Structure in Scientific Discourse: Applications to Summarization and Citation Indexing. CSLI Publications.
- Toulmin, St. (1996/1958): Der Gebrauch von Argumenten. Weinheim.
- Trevisan, B. (2014): Bewerten in Blogkommentaren. Mehrebenenannotation sprachlichen Bewertens. Dissertation. RWTH Aachen University.
- Trevisan, B./Neunerdt, M./Jakobs, E.-M. (2012): A Multi-level Annotation Model for Fine-grained Opinion Detection in German Blog Comments. In: Proceedings of KONVENS 2012, Wien, S. 179–188.
- Trevisan, B./Dickmeis, E./Hemig, T./Jakobs, E.-M./Niehr, Th. (2014a): Indicators of Argument-conclusion Relationships. An Approach for Argumentation Mining in German Discourses. In: Proceedings of the First Workshop on Argumentation Mining (ACL), 26.06.2014. Baltimore, S. 104–105.
- Trevisan, B./Neunerdt, M./Hemig, T./Jakobs, E.-M./Mathar, R. (2014b): Detecting Ironic Speech Acts in Multilevel Annotated German Web Comments. In: Proceedings of KONVENS, Workshop »NLP 4 CMC: Natural Language Processing for Computer-Mediated Communication / Social Media«.
- Trevisan, B./Jakobs, E.-M. (2015): Linguistisches Text Mining. In: Keller, B./Klein, H.-W./Tuschl, S. (Hrsg.): Zukunft der Marktforschung. Heidelberg: Springer Gabler, S. 167–185.
- Trevisan, B./Neunerdt, M. (2015/in Vorbereitung): IRONISIEREN: Verarbeitung und Annotation internetbasierter Kommunikation. Am Beispiel von Webkommentaren. In: Imo, W./Spieß, C. (Hrsg.): Empirische Erforschung internetbasierter Kommunikation. Berlin: Akademie-Verlag.
- Wengeler, M. (2003): Topos und Diskurs. Begründung einer argumentationsanalytischen Methode und ihre Anwendung auf den Migrationsdiskurs (1960–1985). Tübingen: Niemeyer.
- Wyner, A./Mochales-Palau, R./Moens, M.-F./Milward, D. (2010): Approaches to Text Mining Arguments from Legal Cases. In: Francesconi, E./Montemagni, S./Peters, W./Tiscornia, D./Wyner, A./Mochales-Palau, R./Moens, M.-F./Milward, D. (Hrsg.): Semantic Processing of Legal Texts. Berlin/Heidelberg: Springer [LNCS; 6036], S. 60–79.
- Wyner, A./Schneider, J./Atkinson, K./Bench-Capon, T. (2012): Semi-automated Argumentative Analysis of Online Product Reviews. In: Fourth International Conference on Computational Models of Argument (COMMA '12).

Anschriften:

Prof. Dr. Thomas Niehr

RWTH Aachen University

Institut für Sprach- und Kommunikationswissenschaft

Lehr- und Forschungsgebiet Germanistische Sprachwissenschaft

Eilfschornsteinstr. 15

52062 Aachen

t.niehr@isk.rwth-aachen.de

Eva Dickmeis, M.A.
RWTH Aachen University
Institut für Sprach- und Kommunikationswissenschaft
Lehr- und Forschungsgebiet Germanistische Sprachwissenschaft
Eilfschornsteinstr. 15
52062 Aachen
e.dickmeis@isk.rwth-aachen.de

Dr. Bianka Trevisan
RWTH Aachen University
Institut für Sprach- und Kommunikationswissenschaft / HCI-Center
Textlinguistik und Technikkommunikation
Campus-Boulevard 57
52074 Aachen
b.trevisan@tk.rwth-aachen.de

Prof. Dr. Eva-Maria Jakobs
RWTH Aachen University
Institut für Sprach- und Kommunikationswissenschaft / HCI-Center
Textlinguistik und Technikkommunikation
Campus-Boulevard 57
52074 Aachen
e.m.jakobs@tk.rwth-aachen.de