

Beitrag aus:

Sonderband 3 der ZfdG: Wie Digitalität die Geisteswissenschaften verändert: Neue Forschungsgegenstände und Methoden. Hg. von Martin Huber und Sybille Krämer. 2018. DOI: [10.17175/sb03](https://doi.org/10.17175/sb03)

Titel:

Digitale Methoden und Werkzeuge für Diskursanalysen am Beispiel der Wikipedia

Autor/in:

Eva Gredel

Kontakt:

eva.gredel@phil.uni-mannheim.de

Institution:

Universität Mannheim, Philosophische Fakultät

GND:

[105697995X](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63869-p0050-9)

ORCID:

[0000-0002-3689-9834](https://orcid.org/0000-0002-3689-9834)

DOI des Artikels:

[10.17175/sb003_005](https://doi.org/10.17175/sb003_005)


Nachweis im OPAC der Herzog August Bibliothek:

[1007371234](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63869-p0050-9)

Erstveröffentlichung:

27.06.2018

Lizenz:

Sofern nicht anders angegeben 

Medienlizenzen:

Medienrechte liegen bei den Autoren

Letzte Überprüfung aller Verweise:

22.06.2018

GND-Verschlagwortung:

[Diskursanalyse](#) | [Elektronische Enzyklopädie](#) | [Digitalisierung](#) | [Digital Humanities](#) |

Zitierweise:

Eva Gredel: Digitale Methoden und Werkzeuge für Diskursanalysen am Beispiel der Wikipedia. In: Wie Digitalität die Geisteswissenschaften verändert: Neue Forschungsgegenstände und Methoden. Hg. von Martin Huber und Sybille Krämer. 2018 (= Sonderband der Zeitschrift für digitale Geisteswissenschaften, 3). PDF Format ohne Paginierung. Als text/html abrufbar unter DOI: [10.17175/sb003_005](https://doi.org/10.17175/sb003_005).

Eva Gredel

Digitale Methoden und Werkzeuge für Diskursanalysen am Beispiel der Wikipedia

Abstracts

Die Diskurslinguistik in der Tradition Foucaults als relativ neue Disziplin der Linguistik geht der Frage nach, wie soziale Wirklichkeiten in transtextuellen Einheiten konstruiert werden. Bisher werden dabei noch kaum Texte aus digitalen Medien berücksichtigt. Ziel des Beitrags ist es, am Beispiel der Wikipedia das Methodeninventar der Diskurslinguistik in zwei Richtungen zu erweitern: Zum einen sollen spezifische Analysekatégorien für Diskurse in digitalen Medien systematisch dargestellt werden. Zum anderen sollen bereits vorhandene Methoden der Korpuslinguistik und Digital Methods im Hinblick auf die Anforderungen digitaler Diskursanalysen beschrieben werden, um diese in einer Methodologie für digitale Diskursanalysen zu bündeln.

As a relatively new discipline in linguistics, discourse linguistics, in the tradition of Foucault, explores how social realities are constructed in transtextual units. Until now, texts from digital media sources have been barely acknowledged. The goal of this paper is to expand the inventory of methods for discourse analysis in two directions using the example of Wikipedia. First, specific categories for the analysis of discourse in digital media will be described systematically. Second, existing approaches in the field of corpus linguistics and digital methods will be described, keeping in mind the requirements of digital discourse analysis, in order to include them in a methodology for digital discourse analysis.

1. Einleitung

Die Einschätzung des Internets und der dort angebotenen Plattformen hat sich in den letzten Jahren grundlegend gewandelt: Die Auffassung des Internets als virtuelle Welt wurde in den 2000er Jahren durch das Bild des Resonanzraums sozialer Realitäten abgelöst.¹ Zuletzt hat sich die Haltung verbreitet, dass im Social Web durch Interaktion verschiedenster Akteure soziale Wirklichkeiten konstruiert und verändert werden.² Die gesellschaftliche Relevanz und Brisanz digitaler Diskurse zeigt sich aktuell beispielsweise auch in politischen Kontexten: So ist vor Wahlen die Angst vor Fake News, die im Social Web die Runde machen und dort Meinungen beeinflussen, bei politisch Verantwortlichen, Journalisten und Journalistinnen und ganz generell bei Internetnutzern und Internetnutzerinnen größer denn je. Dies ist dadurch begründet, dass diskursive Dynamiken auf digitalen Plattformen wie Twitter oder Facebook gezeigt haben, dass die veränderten Kommunikationsverhältnisse und die somit gewachsene Rolle digitaler Medien in z.T. wahlentscheidenden Diskurssträngen nicht zu unterschätzen sind. Internet-Phänomene und neue digitale Praktiken wie Edit Wars, Trolling und Astroturfing oder die viralen evolutionstheoretischen Verbreitungsmuster von Ideen wie Memes (z.B. [#lovewins](#)) sind Ausgangspunkte für digitale Diskurse, die hohe Reichweiten erlangen und auch in nicht-

¹ Rogers 2013, S. 19f.

² Rogers 2013, S. 21.

digitalen Medien verhandelt werden. Diese Entwicklungen machen digitale Interaktion zu einem besonders interessanten Gegenstandsbereich für Diskursanalysen, der aber noch kaum systematisch in diesem Forschungsfeld bearbeitet wird.

Für das Forschungsfeld der Diskurslinguistik bedeutet dies, dass bisherige Foki, Methoden und Forschungspraktiken unter neuen Vorzeichen diskutiert werden müssen: Berücksichtigten linguistische Diskursanalysen bisher v.a. nicht-digitale Texte (»Newspaper Bias« der Diskurslinguistik³) sind nun digitale Texte aus Twitter, Facebook und Wikipedia verstärkt zu berücksichtigen, um digitale Diskurse systematisch beschreiben und analysieren zu können.

Ziel des vorliegenden Beitrags ist es, am Beispiel der Wikipedia diesen oben beschriebenen Gegenstandsbereich in den Fokus diskursanalytischer Arbeiten zu rücken und Charakteristika von Diskursen in digitalen Medien systematisch zu beschreiben. Dabei soll eine diskurslinguistische Methodologie skizziert werden, die in ihren Beschreibungskategorien und Analysewerkzeugen die Spezifika digitaler Diskurse berücksichtigt. Um sich diesen Herausforderungen zu stellen, sind die Unterschiede zwischen digitalisierten Methoden und digitalen Methoden herauszuarbeiten.

2. Diskurslinguistik

Der Begriff *Diskurs* wurde in den letzten Jahren in der Linguistik in der Schnittmenge der Linguistik und der Soziologie profiliert und somit zu einem interdisziplinär beachteten Diskussionsgegenstand entwickelt. Die Diskurslinguistik als Teildisziplin mit »Interesse an textübergreifenden, also transtextuellen Sprachstrukturen«⁴ untersucht dabei, wie sich einzelne Lexeme, Mehrworteinheiten und komplexere Sprachgebrauchsmuster (z.B. Narrative⁵) in großen thematisch zentrierten Textmengen etablieren und durchsetzen. Diskurse können im Sinne der Diskurslinguistik als große Mengen an Texten oder Äußerungen verstanden werden, die zeitlich aufeinander folgen, bestätigend oder kritisch aufeinander bezogen werden, differenzierend fortgeführt werden und (soziale) Netzwerke und Wissensgemeinschaften kennzeichnen. In Diskursen werden einerseits sprachliche Traditionen etabliert (z.B. anhand bestimmter »Protometaphern«⁶), zugleich haben Diskursakteure die Möglichkeit, permanent Neuschöpfung von diskursiv konstituiertem Sinn einzubringen (z.B. innovative Metaphern zu setzen⁷).

Dem ersten Entwurf einer linguistischen Diskursanalyse in der Germanistischen Linguistik⁸ folgte bereits Anfang der 1990er Jahre die Idee, diskurs- und korpuslinguistische Ansätze zu kombinieren⁹, um die für Diskurse konstitutiven großen Mengen an Texten und Äußerungen

³ Warnke 2013, S. 191.

⁴ Spitzmüller / Warnke 2011, S. 22.

⁵ Vgl. Bubenhofer et al. 2013, passim.

⁶ Vgl. Liebert 1996, S. 808.

⁷ Vgl. Gredel 2014, passim.

⁸ Vgl. Busse 1987, passim.

⁹ Vgl. Busse / Teubert 1994, passim.

zu handhaben. Seit den 2000er Jahren erlebt die Diskurslinguistik in der germanistischen Linguistik und ihre Einführung mit der Korpuslinguistik die Phase der Konsolidierung als Teildisziplin durch synoptische Sammelbände¹⁰ sowie durch Lehrbücher.¹¹ Dies geht mit der Ausdifferenzierung diskurslinguistischer Erkenntnisinteressen einher, die sich in erweiternden Ansätzen wie der Diskurslexikographie¹², der Diskursgrammatik¹³ oder Diskurssemiotik¹⁴ und je eigenen methodischen Zugängen zu Diskursen widerspiegeln. Bisher bleiben diese Bemühungen allerdings oftmals beschränkt auf nicht-digitale bzw. digitalisierte Untersuchungsgegenstände wie z.B. Zeitungstexte. Die korpuslinguistische Beschäftigung mit digitalen Daten in der Diskursanalyse steht momentan noch am Anfang, auch wenn beispielsweise Korpora internetbasierter Kommunikation nach und nach in bestehende Forschungsinfrastruktur integriert werden.

3. Digitale Sprachressourcen in der Korpuslinguistik zur Analyse digitaler Diskurse

Mittlerweile sind in der Computerlinguistik eine Vielzahl von Methoden des Natural Language Processings¹⁵ vorhanden und in der Korpuslinguistik stehen Verfahren für die Textaufbereitung und -analyse großer Textmengen zur Verfügung.¹⁶ Die Entwicklung neuer korpuslinguistischer Verfahren geht auch mit der Erarbeitung neuer digitaler Sprachressourcen einher, die zum Ziel haben, die Nachnutzbarkeit der erstellten Korpora für eine große wissenschaftliche Öffentlichkeit zu ermöglichen.

In diesen digitalen Sprachressourcen sind häufig Korpora, also Textsammlungen, verzeichnet, die ursprünglich nicht digital vorlagen (z.B. das Marx-Engels-Korpus im Deutschen Referenzkorpus¹⁷), dann aber digitalisiert wurden, um sie korpuslinguistischen Untersuchungen zugänglich zu machen. Die Anreicherung dieser digitalisierten Texte mit linguistischen Metadaten (z.B. POS-Tagging) ermöglicht es dann auch, komplexe sprachwissenschaftliche Analysen durchzuführen.¹⁸ Das Spektrum digitalisierter Texte bzw. Textsorten beispielsweise in **DeReKo** ist entsprechend groß: So finden sich dort neben Zeitungstexten auch belletristische Texte oder Plenarprotokolle des Deutschen Bundestages. Diese Korpora eignen sich dazu, die für diskursanalytische Studien wichtige Serialität von Sprachgebrauchsmustern über Suchanfragen zu rekonstruieren und quantitativ deren Verwendungskontexte z.B. in Kookkurrenzanalysen zu explorieren, wie dies auch Ingo H. Warnke beschreibt:

¹⁰ Vgl. Warnke / Spitzmüller 2008, *passim*; Felder et al. 2012, *passim*; Busse / Teubert 2013, *passim*.

¹¹ Vgl. Spitzmüller / Warnke 2011, *passim*; Niehr 2014, *passim*; Bendel Larcher 2015, *passim*.

¹² Vgl. Kämper 2006, *passim*.

¹³ Vgl. Warnke et al. 2014, *passim*.

¹⁴ Vgl. Hess-Lüttich et al. 2017, *passim*.

¹⁵ Vgl. Lobin 2010, *passim*.

¹⁶ Vgl. Lemnitzer / Zinsmeister 2006, *passim*; Perkuhn et al. 2012, *passim*.

¹⁷ Vgl. Marx-Engels-Korpus in DeReKo 2017, *passim*.

¹⁸ Perkuhn et al. 2012, S. 57f.

»Die Korpusorientierung der Diskurslinguistik unterstreicht [...] ein Interesse an seriellen Sprachgebrauchsdaten. Das Primat der Serialität ist durchaus auch kompatibel mit einem Begriff von Diskurs als Aussageninformation.«¹⁹

Zunehmend finden jedoch auch genuin digital verfasste Texte Eingang in die beschriebene Forschungsinfrastruktur. Die besonderen methodischen Herausforderungen dieser Erweiterung der Forschungsinfrastruktur beschreiben Beißwenger et al. im Zusammenhang mit dem Referenzkorpus zur internetbasierten Kommunikation (DeRiK²⁰), das zur linguistischen Analyse internetbasierter Kommunikation seit den 1990er Jahren in den Fokus der Linguistik gerückt ist.²¹ In DeReKo (Das deutsche Referenzkorpus) sind mittlerweile drei Korpora integriert, die Texte internetbasierter Kommunikation verzeichnen: Das Usenet News Corpus, das Dortmunder Chat-Korpus und die Wikipedia-Korpora²², wobei zu letzteren sogar multilinguale Korpora über das Corpus Search, Management and Analysis System (COSMAS II) zur Verfügung stehen. Diese Ressourcen werden bisher noch kaum diskurslinguistisch genutzt, sind aber für die Analyse digitaler Untersuchungsgegenstände sehr relevant. Die Aufbereitung, Anreicherung der Wikipedia-Daten mit linguistischen Metadaten und ihre Integration in die Korpusinfrastruktur in DeReKo ist gut beschrieben.²³ Die in Korpusinfrastrukturen implementierten Algorithmen nehmen, um die oben bereits erwähnte und für Diskursanalysen wichtige Serialität sichtbar zu machen, eine »diagrammatische Umformung des Textes [...] in Konkordanzen«²⁴ vor. Diese diagrammatischen Operationen auf der Grundlage der abgefragten Muster der Textoberfläche (beispielsweise in Konkordanzzeilen) ermöglichen dem Diskursanalytiker eine neue Sicht auf die Daten, wie Noah Bubenhofer und Klaus Rothenhäusler beschreiben:

»Die Einheit des Einzeltextes wird aufgelöst und der neue locus, also das Ensemble einzelner Textpartien, ermöglicht eine neue Sicht auf die Texte. [...] Um andere Ordnungen zu ermöglichen, dienen diagrammatische Methoden: Immer geht es darum, eine Art Diagramm zu entwickeln, das Texte oder Textteile repräsentiert, jedoch kraft seiner Zeichenhaftigkeit Operationen erlaubt, die mit den Texten selber nicht möglich wären.«²⁵

Für diese genannten diagrammatischen Operationen ist bisher bei korpuslinguistischen Verfahren allein die textuelle Verfasstheit der Daten auf der Textoberfläche ausschlaggebend. Die korpuslinguistischen Verfahren stoßen jedoch bei Analysen an ihre Grenzen, bei denen die Datenauswahl und -abfrage nicht allein über sprachliche Muster bzw. Suchstrings auf der Textoberfläche möglich oder sinnvoll ist. Dies ist beispielsweise bei multimodalen Analysen oder aber bei sprachvergleichenden Analysen verschiedener Namensräume oder verschiedener Sprachversionen der Fall: So liegen in DeReKo die Wikipedia Artikelseiten getrennt von den Diskussionsseiten vor (in je eigenen Archiven bzw. Korpora), obwohl diese in der digitalen Plattform hypertextuell verlinkt sind. Die hypertextuelle Verlinkung muss

¹⁹ Warnke 2015, S. 233.

²⁰ Vgl. Beißwenger et al. 2013, *passim*.

²¹ Vgl. Beißwenger / Storrer 2008, *passim*.

²² Lungen / Kupietz 2017, S. 20.

²³ Vgl. Bubenhofer et al. 2011, *passim*, Margaretha / Lungen 2014, *passim*.

²⁴ Bubenhofer / Rothenhäusler 2016, S. 63.

²⁵ Bubenhofer / Rothenhäusler 2016, S. 62.

nach der Korpusanfrage über Suchstrings auf der Textoberfläche manuell rekonstruiert werden. Die komplexe Struktur der Wikipedia und die daraus resultierende Komplexität der Diskursfragmente in Wikipedia sind bisher für korpuslinguistische Verfahren nicht nachzubilden.

4. Diskurse – digital am Beispiel der Wikipedia

Für die in diesem Beitrag folgende linguistische Betrachtung digitaler Diskurse am Beispiel der Wikipedia ist die zentrale Unterscheidung von Text und Hypertext notwendig, wie sie Angelika Storrer²⁶ am Beispiel der Wikipedia vorgenommen hat. Da die Wikipedia eine digitale Plattform mit hypertextuellen Eigenschaften ist, stellt sie eine komplexe, aber auch wertvolle linguistische Ressource dar, die jedoch methodische Herausforderungen an linguistische Diskursanalysen mit sich bringt. Die der Wikipedia zugrundeliegende Wiki-Software MediaWiki bedingt die hypertextuellen Merkmale der Online-Enzyklopädie wie Nicht-Linearität, Interaktivität, Multimodalität, Adaptivität und Dynamik.²⁷

Die wichtigste Eigenschaft ist die der Nicht-Linearität: Storrer definiert Hypertexte als »nicht-linear organisierte Texte, die durch Computertechnik verwaltet werden«. ²⁸ Die Besonderheit ist dabei, dass »[d]er Autor eines Hypertextes [...] seine Daten auf mehrere Module – im WWW werden solche Module üblicherweise als Seiten bezeichnet – [verteilt].« ²⁹ Hypertexte befördern mit dieser Struktur die selektive Rezeption der dargebotenen Inhalte, da »jedes Modul mehrere Links enthalten kann, sodass die Nutzer je nach Vorlieben und Interessen selbst entscheiden können, welche Module sie in welcher Reihenfolge abrufen möchten.« ³⁰ Bei der Wikipedia finden sich eine ganze Reihe verschiedener Linkarten, die die nicht-lineare Natur der Wikipedia bedingen und je nach Blickwinkel technische Affordanzen des Systems oder auch Grenzen der individuellen Beiträge in der kollaborativen Wissenskonstruktion darstellen:

»Bei der Produktion von Hypertexten ist der Spielraum von Hypertextautoren wesentlich determiniert von der Funktionalität des Hypertextsystems, insbesondere von den vom jeweiligen System unterstützten Strukturierungskonzepten und von den Navigations- und Orientierungswerkzeugen, die dem Nutzer für die Rezeption angeboten werden.« ³¹

Die Wikipedia ist derart strukturiert, dass sich neben den – häufig isoliert – rezipierten Artikelseiten mit dem zentralen Lemma auch Diskussionsseiten, Benutzer- und Benutzerdiskussionsseiten, die Versionsgeschichte sowie Metaseiten finden. Diese verschiedenen Arten bzw. Typen von Seiten sind über Hyperlinks miteinander verknüpft, folgen allerdings je eigenen Regeln. Über die sogenannten Interwiki Links sind die Sprachversionen

²⁶ Storrer 2008, S. 318ff.

²⁷ Storrer 2008, S. 318ff.

²⁸ Storrer 2008, S. 318.

²⁹ Storrer 2008, S. 319.

³⁰ Storrer 2008, S. 319.

³¹ Storrer 2008, S. 320.

hypertextuell verbunden. Seiten, die gleichen Regeln folgen und ähnlich aufgebaut sind, werden deshalb zu sogenannten Namensräumen zusammengefasst.

Während die Artikelseiten mit dem Lemma (Abbildung 1, A) den enzyklopädischen Kern der Online-Enzyklopädie und somit das Produkt diskursiver Aushandlung darstellen und sich durch textorientiertes Schreiben auszeichnen³², stellen die korrespondierenden Diskussionsseiten (Abbildung 1, B) den Raum diskursiver Aushandlung dar: Diskursakteure versuchen dort, in Interaktion mit anderen ihr für die Artikelseiten präferiertes Vokabular durchzusetzen, um Sachverhalte perspektivisch zu konstituieren.

Herring bezeichnet diese situative Kopräsenz von Artikel- und Diskussionsseiten als Text-Text-Konvergenz: »Less prototypically (because it involves the convergence of text with text rather than the convergence of text with another mode), CMC is also illustrated by [...] ›talk‹ pages associated with Wikipedia articles.«³³

Grundlegend für (diskurs-)linguistische Untersuchungen ist auch, dass sich die sprachlichen Phänomene in den verschiedenen Namensräumen – überwiegend aufgrund präskriptiver Aussagen der Diskursgemeinschaft in der Wikipedia – maßgeblich unterscheiden: Während die Artikelseiten durch textorientiertes Schreiben charakterisiert sind, zeichnen sich die Diskussionsseiten durch netztypische Besonderheiten des interaktionsorientierten Schreibens aus³⁴, deren Analyse interaktionsanalytischer Kategorien bedarf.³⁵ Im Fall der Wikipedia findet sich u. a. dort das oben bereits genannte Merkmal der Interaktivität, das sich auf die Interaktion von Diskursakteur zu Diskursakteur bei der kollaborativen Erarbeitung von Wikipedia-Artikeln bezieht. Die diskursanalytische Relevanz der so beschriebenen Wikipedia-Struktur machen Birte Arendt und Philipp Dreesen deutlich, auch wenn sie in ihrer Analyse der Wikipedia lediglich die Artikelseiten und nicht die Diskussionsseiten in die konkrete empirische Analyse einbeziehen:

»Ingo Warnke (2013:108) macht auf das wechselseitige Verhältnis der Auffassung vom ›Diskurs als Praxis‹ (›Handlungsvollzug‹) und vom ›Diskurs als Arrangement‹ (›Handlungsprodukt‹) aufmerksam. Für die Untersuchung bedeutet das, dass der Diskurs in den Artikeln ex post und – in der Dokumentation der Diskussionen und Artikelversionen – ansatzweise in actu zugänglich ist.«³⁶

Neben der bereits genannten Text-Text-Konvergenz ermöglicht die Wikipedia zudem die Integration unterschiedlicher semiotischer Ressourcen (mediale Konvergenz³⁷), was die zunehmend multimodale Verfasstheit (Multimodalität) der Wikipedia bedingt: Die Autoren können Bild-, Ton- und Videomaterial einbringen, um Artikelseiten auszugestalten. Versteht

³² Storrer 2012, S. 277.

³³ Herring 2013, S. 5.

³⁴ Storrer 2012, S. 277.

³⁵ Vgl. Beißwenger 2016, *passim*.

³⁶ Arendt / Dreesen 2015, S. 433.

³⁷ Vgl. Bucher 2013, *passim*.

man Diskurse als »semiotisches Kohärenzphänomen«³⁸ rückt die Frage in den Vordergrund, inwiefern sich Bildinventare zu ausgewählten Fallbeispielen in verschiedenen Sprachversionen der Wikipedia unterscheiden und welche Muster der Text-Bild-Kombination diskursspezifisch sind.³⁹ Die bereits erwähnte Adaptivität stellt für multimodale Analysen ein noch ungelöstes Problem dar: »Adaptivität meint in einem einfachen Sinne, dass Daten aus ein und derselben digitalen Datenbasis für verschiedene Zugriffsmedien, also Computerbildschirm, Smartphone, Tablet-Computer, in jeweils spezifischer Weise präsentiert werden.«⁴⁰ Je nach Zugriffsmedium können so z.B. Text-Bild-Konstellationen sehr unterschiedlich ausfallen. Die Dynamik von Hypertexten im Allgemeinen und im Besonderen im Beispiel der Wikipedia ist gleichzeitig Gewinn und Herausforderung für diskursanalytische Studien: »Die Wikipedia besteht nicht aus einer festen Anzahl von Artikeln [...], sondern wird kontinuierlich um neue Artikel erweitert. Auch der Inhalt der Artikel ist nicht statisch, sondern kann bei Bedarf jederzeit verändert und aktualisiert werden.«⁴¹ Im dynamischen Hypertext der Wikipedia ist es vor allem in der Versionsgeschichte (Abbildung 1, C) möglich die »diskursive Historizität«⁴² der Artikel und der dort relevant gesetzten Termini zu rekonstruieren.

Es zeigt sich somit, dass aus einer diskursanalytischen Perspektive zahlreiche Elemente der Wikipedia diskursanalytisch relevant sind und eine umfassende Diskursanalyse mehrere, wenn nicht sogar alle Namensräume und deren hypertextuelle Verlinkung berücksichtigen muss. Ein Aspekt der bereits erwähnten Non-Linearität und der damit einhergehenden Verknüpfung verschiedener Module ist besonders relevant und ermöglicht eine zusätzliche Erweiterung diskursanalytischer Erkenntnisinteresse:

Eine zentrale Rolle bei der hypertextuellen Struktur der Wikipedia spielt ein besonderer Typ hypertextueller Verweise: Die rund 290 Sprachversionen der Wikipedia sind über die sogenannten Interwiki-Links miteinander verknüpft. Dies kann mit Sandrini als Phänomen »interlingualen Kohärenz«⁴³ gedeutet werden kann. Zur deutschen Artikelseite mit dem Lemma *Krimkrise* ist in der englischen Sprachversion der Artikel mit dem Lemma *Annexation of Crimea by the Russian Federation* verknüpft. Dass sich in der englischen Wikipedia das Lemma *Annexation of Crimea by the Russian Federation* durchgesetzt hat, ist beachtenswert, da der Begriff *Crimean Crisis* ebenfalls auf den Diskussionsseiten der englischen Sprachversion vorkommt. Da es sich bei den Artikeln verschiedener Sprachversionen der Wikipedia oftmals nicht um bloße Übersetzungen handelt, qualifiziert sich die Wikipedia als Ressource zum Sprach- und Kulturvergleich.⁴⁴

³⁸ Vgl. Hess-Lüttich et al. 2017, *passim*.

³⁹ Gredel 2017b, S. 145ff.

⁴⁰ Storrer 2012, S. 287.

⁴¹ Storrer 2008, S. 321.

⁴² Vgl. Kämper et al. 2016, *passim*.

⁴³ Sandrini 2012, S. 246.

⁴⁴ Vgl. Gredel 2016, *passim*.



Abb. 1: Eigenschaften und Strukturen der Wikipedia als Hypertext. Quelle: © Wikipedia, 2017.

Bei der Analyse der Diskussionsseiten und der Versionsgeschichte ist zudem auffällig, dass die jeweiligen Lemmata in beiden Sprachen »agonale Zentren« darstellen und sich die Akteure an sogenannten Edit Wars beteiligen, d.h. konkurrierende Termini wie *annexation*, *incorporation* und *accession* durch Überarbeitungen des Artikels immer wieder austauschen. Erst eine von einem Administrator herbeigeführte Abstimmung auf der Diskussionsseite sorgt für die Festlegung auf das oben genannten Lemma.

Wie das Beispiel gezeigt hat, stellt die Wikipedia einen einzigartigen Untersuchungsgegenstand dar, der jedoch methodische Herausforderungen mit sich bringt: Digitale Daten strukturieren die Datenquelle und machen sie zu einer komplexen Ressource. Abbildung 2 zeigt die schematische Darstellung eines Diskursfragments in der Wikipedia:

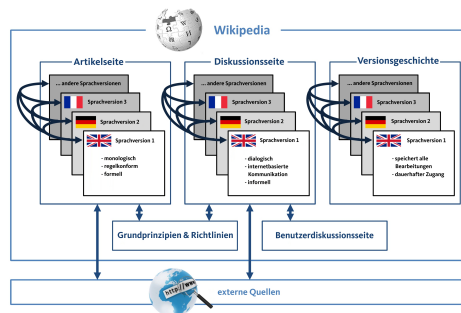


Abb. 2: Schematische Darstellung eines Diskursfragments in der Wikipedia. © Eigene Darstellung, 2017.

Um sich diesen methodischen Herausforderungen, die durch die Eigenschaften wie Nicht-Linearität, Multimodalität, Adaptivität, Offenheit, Interaktivität und Multilingualität entstehen, zu stellen, sind die Unterschiede zwischen digitalisierten Methoden und digitalen Methoden herauszuarbeiten und bei der Entwicklung neuer Methoden und digitaler Sprachressourcen zu beachten. Wie das Paradigma der Digital Methods die beschriebenen Charakteristika digitaler Diskurse berücksichtigt, ist Gegenstand des folgenden Kapitels.

5. Methoden und Tools – digital: Das Beispiel Wikipedia

Wie bereits in Kapitel 5 beschrieben, stellen die in digitalen Diskursen vorgefundenen genuin digitalen Daten bzw. Objekte (z.B. Hyperlinks, Hashtags, Timestamps, IPs) und die sich daraus ergebenden Diskursstrukturen und Diskursfragmente bisherige diskursanalytische Verfahren vor methodische Hürden. Hilfestellung bei der Auswahl und Erhebung von Daten für diskursanalytische Studien können jedoch Methoden und Tools geben, die einem Paradigma zuzurechnen sind, das sich explizit der Analyse digitaler Daten verschreibt: Die unter dem Begriff *Digital Methods* gefassten Methoden und Tools, die um Richard Rogers von der Digital Methods Initiative an der Universität Amsterdam entwickelt wurden, werden klar gegenüber Digitized Methods abgegrenzt. Digitalisiert sind Methoden dann, wenn sie aus nicht-digitalen Medien in digitale Medien übertragen werden (z.B. Online-Umfragen) und beim Transfer die Berücksichtigung genuin digitaler Daten und Methoden vernachlässigt wird. Digital Methods können im Gegensatz dazu folgendermaßen beschrieben werden: »Digital methods not only think with online devices. They also take stock of the availability and exploitability of digital objects so as to recombine them fruitfully«⁴⁵ Unter dem Diktum »follow the medium«⁴⁶ empfehlen die Vertreter der Digital Methods Initiative, technische Affordanzen der jeweiligen digitalen Plattform als Forschungsaffordanzen⁴⁷ zu verstehen:

»I introduce the term ›research affordances‹ of digital media, focusing specifically on the analytical affordances of platforms and engines as devices in digital research; they deal with the relation between objective, medium and method, and are specific to the actors and contexts of use.«⁴⁸

Wichtig ist an dieser Stelle der Hinweis von Konstanze Marx und Georg Weidacher darauf, dass »die einzelnen Plattformen und das Internet generell [...] verschiedene Formen von Kommunikation«⁴⁹ ermöglichen. Die jeweiligen Affordanzen und auch die Constraints sind für jede Plattform einzeln zu beschreiben und bei der Entwicklung von Methoden und Tools zu berücksichtigen. Dies spiegelt sich auch in der Sammlung relevanter Tools zur Analyse digitaler Daten wider, wie sie beispielsweise die Digital Methods Initiative in Amsterdam entwickelt. So finden sich dort für digitale Plattformen wie Twitter, Facebook oder Wikipedia je spezifische Tools, die jeweils nur für eine Plattform eingesetzt werden können. Die Zielsetzung lautet dann: »Follow the methods of the medium as they evolve, learn from how the dominant devices treat natively digital objects, and think along with those object treatments and devices so as to recombine or build on top of them«⁵⁰

Die so verfassten und frei verfügbaren Tools ermöglichen je einzelne Aspekte oder Stadien des Umgangs mit digitalen Daten. Bisher wurden diese Tools fast nur in medien- und sozialwissenschaftlichen Studien genutzt. In diskursanalytischen bzw. noch konkreter diskurslinguistischen Studien finden diese Tools noch keinen Einsatz. Ziel der folgenden

⁴⁵ Rogers 2013, S. 1.

⁴⁶ Rogers 2013, S. 24.

⁴⁷ Weltverde 2016, S. 13.

⁴⁸ Weltverde 2016, S. 13.

⁴⁹ Marx / Weidacher 2014, S. 83.

⁵⁰ Rogers 2013, S. 5.

Wie Abbildung 3 am Beispiel des Diskursfragments zum Lemma *Burn-Out* zeigt, extrahiert das Tool – ausgehend von einer Artikelseite (hier die deutsche Seite unter dem Lemma *Burn-Out*) – die Bilder und Grafiken aller hypertextuell verlinkten Sprachversionen. Das Bildmaterial der Artikelseite einer Sprachversion wird in jeweils eine Zeile überführt. Das Tool nutzt somit die bereits in Kapitel 4 erwähnte interlinguale Kohärenz, die über die Interwiki-Links hergestellt wird. Die Interwiki-Links als technische Affordanz zur Verlinkung der Sprachversionen wird somit als »research affordance« (im Folgenden Forschungsaffordanz) genutzt, um Daten mithilfe einer diagrammatischen Operation neu zu ordnen. Das so erstellte Diagramm kann dann Ausgangspunkt für multimodale und kontrastive Diskursanalysen sein. Bereits dieses Beispiel zeigt plakativ, dass in den verschiedenen Sprachversionen ganz unterschiedliche Zahlen an Bildern und Grafiken vorhanden sind. Auffällig ist zudem, dass in der arabischen und in der französischen Wikipedia eine niedergebrannte Kerze zur Illustration der Protometapher *Burn-Out* herangezogen wird. Das Tool bietet somit den Ausgangspunkt dafür, kulturelle Konvergenzen und Divergenzen zu rekonstruieren.

5.2 Digitale Tools – kontrastiv

Wie das Beispiel oben bereits gezeigt hat, ist die Wikipedia mit ihren rund 290 Sprachversionen Hinweis darauf, dass Diskurse im Web 2.0 nicht auf einzelne Sprachen oder Kulturen begrenzt sind: Sie sind multilingual und multikulturell konstituiert. Die über Links verknüpften Sprachversionen der Wikipedia können diskursanalytisch als hypertextuell und thematisch zusammenhängende Korpora genutzt werden, die kulturelle Differenzen und Gemeinsamkeiten sichtbar machen. Das oben skizzierte Beispiel liefert Evidenz für kulturelle Differenzen bei multimodaler Gestaltung bestimmter Themen und legitimiert zugleich kontrastiv konzipierte Diskursanalysen. Noch vor wenigen Jahren, waren kontrastiv konzipierte Diskursanalysen in der Linguistik wenig verbreitet: »Da aber die Diskursforschung erst in den Anfängen steckt, will ich nicht vom interdiskursiven Vergleich sprechen und ordne deshalb den Vergleich von zwei Diskursen zwei[er] Sprach- und Kommunikationskulturen der kontrastiven Textologie zu.«⁵⁴ Zwischenzeitlich hat sich jedoch auch diese Erweiterung diskursanalytischer Interessen etabliert und konsolidiert: »Die kontrastive Diskurslinguistik versteht sich als eine Forschungspraxis, die einerseits aus den Methoden der Diskurslinguistik und andererseits aus den Methoden der kontrastiven bzw. interkulturellen Linguistik schöpft.«⁵⁵ Das so skizzierte Forschungsprogramm kontrastiver Diskursanalysen wird zwischenzeitlich methodisch und terminologisch weiter ausdifferenziert. Eine für Wikipedia-Analysen wichtige begriffliche Differenzierung ist die zwischen Paralleldiskurse⁵⁶ und »transnationale Diskurse«⁵⁷. Rogers beschreibt seine Methode als »web content analysis« und beschreibt die für dieses Verfahren zentralen Elemente der Wikipedia:

»The approach taken in the comparative study is relatively straightforward. The comparisons across language versions of Wikipedia are based on a form of web content analysis that

⁵⁴ Bilut-Homplewicz 2008, S. 488.

⁵⁵ Czachur 2013, S. 333.

⁵⁶ Czachur 2013, *passim*.

⁵⁷ Gür-Şeker 2012, S. 306.

focuses on basic elements that comprise an article: its title, authors (or editors), table of contents, certain content details, images, and references.«⁵⁸

Das Tool **Manypedia** macht ebenfalls von der technischen Affordanz der Interwiki-Links Gebrauch und bietet die Möglichkeit, zwei hypertextuell durch Interwiki-Links verknüpfte Seiten in zwei parallelen Fenstern anzeigen und in eine gewünschte Zielsprache übersetzen zu lassen. Grundlage für die Übersetzung ist die Google Translate API.

5.3 Akteursorientierte Perspektive

In den letzten Jahren ist die akteursorientierte Perspektive in den Fokus der Diskursanalyse gerückt. Heidrun Kämper beschreibt die Funktion von Akteuren aus einer diskurslinguistischen Perspektive folgendermaßen: »Die besondere Funktion der Akteure im Diskurs besteht darin, Sinn zu schaffen und gleichzeitig selbst Sinnträger zu sein.«⁵⁹ Die Möglichkeiten der einzelnen Akteure, Diskurse zu beeinflussen und sprachlichen Mustern zu Dominanz im Diskurs zu verhelfen, ergibt sich dabei über deren jeweilige Rolle und Position in einem spezifischen Kontext: »Diskurse, verstanden als soziale Praktiken, sind mehr oder weniger symmetrische oder asymmetrische Strukturen, innerhalb derer Akteure je entsprechende Rollen und Positionen einnehmen.«⁶⁰

Gerade am Beispiel der Wikipedia mit dem großen Spektrum verschiedener Rollen ergibt sich ein ausdifferenziertes Beziehungsgefüge: Der unangemeldete Autor (»IP-ler«⁶¹) kann durch Anlegen eines Benutzerkontos aus der Quasi-Anonymität heraustreten und weitere Rechte erwerben: »Wenn ein Neuling einen Artikel bearbeitet hat, dann wird dies für die Allgemeinheit erst sichtbar, wenn ein Benutzer mit Sichterstatus sein Okay gegeben hat [...]. Man erhält den Status [eines Sichters] nach sechzig Tagen, außerdem muss man mindestens dreihundertmal bearbeitet haben«⁶². Wahlämter wie das des Administrators gehen mit Rechten wie Sperren und Löschen von Seiten einher.⁶³

Den ungeübten Nutzern der Online-Enzyklopädie bleiben die oben beschriebenen Akteurskonstellationen intransparent. In medialen Diskursen wird diese angebliche Intransparenz der Autorschaft in der kollaborativen Textproduktion häufig als Nachteil bzw. Constraint der Wikipedia genannt. Allerdings bietet die Versionsgeschichte bereits die Möglichkeit, einzelne Bearbeitungen nachzuvollziehen. Die Benutzerdiskussionsseiten geben Aufschluss darüber, in welcher Position (Sichter, Administrator, etc.) sich ein Wikipedia-Autor befindet. Aggregierte Daten sind ebenfalls in der Wikipedia hinterlegt bzw. verlinkt: Das Tool **WikiHistory** zeigt Statistiken zu Artikeln an, die die prozentuale Beteiligung eines Autors an der

⁵⁸ Rogers 2013, S. 166.

⁵⁹ Kämper 2017, S. 259.

⁶⁰ Kämper 2017, S. 263.

⁶¹ Eine IP ist eine Internetadresse in Form einer mehrziffrigen Zahl, die Beiträge in der Wikipedia über Internetzugänge rekonstruierbar macht. IP-ler sind Internetnutzer, die keinen Account in der Wikipedia anlegen.

⁶² van Dijk 2010, S. 34.

⁶³ van Dijk 2010, S. 35.

Entstehung eines Artikels enthalten. Das Tool **Wikiwatchdog** bietet eine andere Sicht auf die akteursorientierte Analyse: Im Zentrum steht die Nachverfolgung der bereits erwähnten IP'ler, die als nicht angemeldete Nutzer die Wikipedia bearbeiten. In das Suchfeld des Tools können IP-Adressen oder Domains eingegeben werden. **Wikiwatchdog** zeigt dann alle von dieser IP-Adresse oder Domain getätigten Bearbeitungen der Wikipedia-Artikel. Das Tool dient somit zur Deanonymisierung von Bearbeitungen und kann als Hilfsmittel herangezogen werden, um alle Diskursäußerungen einzelner Akteure zu aggregieren.

Auch das Tool **whoCOLOR** befördert die akteursorientierte Perspektive diskursanalytischer Arbeiten, indem es auf der textuellen Oberfläche der Artikelseite des analysierten Diskursfragments die jeweilige Autorschaft visualisiert:

»WhoColor is a [...] userscript for the Tamper-/Greasemonkey browser extensions for Chrome and Firefox. when you open an (english) wikipedia article it creates a color-markup on the text, showing the original authors of the content, an author list ordered by percentages of the article written and (soon) additional provenance information.«⁶⁴

Der prozentuale Anteil an der Erstellung des jeweiligen Artikels ist nach Autoren sortiert am rechten Rand der Seite gelistet. Die so auf die Artikelseite projizierten Informationen geben Aufschluss über die Rolle einzelner Akteure bei spezifischen diskursiven Aushandlungen.



Abb. 4: Funktionsweise des Tools whoColor. © f-squared.org.

Weitere Tools wie **Whovis** und **Wikiwho** zur Visualisierung von Autoren-Beteiligung und Interaktion in der Online-Enzyklopädie Wikipedia befinden sich in der Entwicklung.

5.4 Agonalität

Berücksichtigt man in Diskursanalysen nicht nur die Artikelseiten, sondern auch die Diskussionsseiten, wird deutlich, dass viele Elemente des enzyklopädischen Kerns der Wikipedia Produkt intensiver diskursiver Aushandlungsprozesse sind. Die auf den Diskussionsseiten ausgetragenen Konflikte und Kontroversen sind somit diskursanalytisch sehr relevant. Eine von Ekkehard Felder im Feld der Diskurslinguistik etablierte Kategorie ist die des agonalen Zentrums:

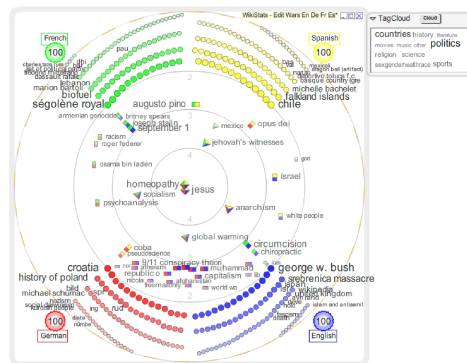
⁶⁴ **whovisual: interactive visualizations to explore wiki article development.**

Unter agonalen Zentren verstehe ich einen sich in Sprachspielen manifestierenden Wettkampf um strittige Akzeptanz von Ereignisdeutungen, Handlungsoptionen, Geltungsansprüchen, Orientierungswissen und Werten in Gesellschaften.⁶⁵

Aus einer linguistischen Perspektive ist der genannte Wettkampf häufig ein »semantische[r] Kampf«, d. h. ein »Kampf um angemessene Benennungen.«⁶⁶ In der Wikipedia zeigen sich solche semantischen Kämpfe einerseits auf den Diskussionsseiten der Wikipedia und andererseits in der Versionsgeschichte anhand bestimmter digitaler Praktiken (wie beispielsweise Edit Wars). Das Tool **Search Cristal** detektiert und bewertet die Intensität von Edit Wars in mehreren Sprachversionen der Wikipedia folgendermaßen: »The controversy M of an article is defined by summing the weights of all mutually reverting editor pairs, excluding the topmost pair, and multiplying this number by the total number of editors E involved in the article.«⁶⁷ Berücksichtigt werden bei der Messung der Intensität auch das bisherige Engagement eines Wikipedia-Autors (»senior editor« versus »junior editor«⁶⁸). Die zugehörige Visualisierung dient dazu, Themen ausfindig zu machen, die in mehreren Sprachversionen der Wikipedia kontrovers verhandelt werden und Gegenstand von Edit Wars sind:

»The searchCrystal visualization toolset will be used to compare, visualize and identify Wikipedia pages that are highly contested in multiple languages. Similar to a bullseye display, searchCrystal uses a radial mapping so that the Wikipedia pages contained in all the language lists that are being compared are mapped to the center of the display and the number of lists that contain the same page decreases toward the periphery of the display.«⁶⁹

Abbildung 5 zeigt die Visualisierung in Search Cristal für die deutsche, englische, französische und spanische Sprachversion.



⁶⁵ Felder 2015, S. 108.

⁶⁶ Felder 2006, S. 1.

⁶⁷ Yasseri et al. 2014, S. 27.

⁶⁸ Yasseri et al. 2014, S. 27.

⁶⁹ Yasseri et al. 2014, S. 28.

Abb. 5: Das Tool Search Cristal visualisiert Edit Wars. © [searchCrystal](#).

Kontroversen in Wikipedia, die diskursanalytisch als agonale Zentren gedeutet werden können, sind auch Gegenstand des Tools [Contropedia](#). Das angewendete Verfahren kann folgendermaßen beschrieben werden:

»A score is calculated for each wiki link according to the volume of activity around it. More precisely, the score for each wiki link is based on the number of substantial disagreeing edits to sentences containing the wiki link.«⁷⁰

Auch bei diesem Tool werden also Informationen aus dem Namensraum der Versionsgeschichte auf die Artikelseiten projiziert und visualisiert: Die kontrovers verhandelten Begriffe eines Wikipedia-Artikels werden je nach Kontroversität unterschiedlich intensiv eingefärbt. In der sogenannten Dashboard-Funktion kann zudem der zeitliche Verlauf der Kontroversen nachverfolgt werden.

6. Diskussion und Fazit

Die Erweiterung des Gegenstandsbereichs der Diskurslinguistik auf den Bereich digitaler Medien bringt methodische und methodologische Herausforderungen mit sich, denen mithilfe zweier Paradigmen begegnet werden kann: Zum einen stehen zunehmend Korpora internetbasierter Kommunikation in der etablierten Korpusinfrastruktur wie beispielsweise die Wikipedia-Korpora in DeReKo zur Verfügung. Zum anderen haben Verfahren und Tools, die in den Sozial- und Medienwissenschaften entwickelt wurden, das Potential Diskurse in digitalen Medien Diskursanalysen zugänglich zu machen. Die in diesem Beitrag exemplarisch verhandelte digitale Plattform Wikipedia wurde als komplexe, aber auch vielversprechende Ressource für digitale Diskursanalysen vorgestellt: Explorative Studien haben gezeigt, dass alle Namensräume der Wikipedia als Datenmaterial diskursanalytisch genutzt werden sollten.⁷¹ Neben den Artikel- und Diskussionsseiten sind auch die korrespondierenden und hypertextuell verknüpften Versionsgeschichten in den Fokus empirischer Studien zu rücken, um den Prozess der kollaborativen Wissenskonstruktion vollständig zu erfassen. Die Komplexität der so konstituierten Diskursfragmente wird zusätzlich dadurch erhöht, dass den jeweiligen Namensräumen ein je eigener Stil und je eigene digitale Praktiken eigen sind.⁷² Sowohl bei korpuslinguistischen als auch bei digitalen Methoden werden die Daten zur Analyse in neue Konstellationen bzw. neue Ansichten überführt, was häufig mit »diagrammatischen Operationen«⁷³ einhergeht. Der Unterschied zwischen korpuslinguistischen Methoden gegenüber digitalen Methoden kann folgendermaßen beschrieben werden: Während digitale Methoden aus den Sozial- und Medienwissenschaften bei der Umsetzung diagrammatischer Methoden dem Medium und seinen technischen Affordanzen folgen (»follow the medium«⁷⁴)

⁷⁰ Pentzold et al. 2017, S. 6.

⁷¹ Vgl. Gredel 2017a, *passim*.

⁷² Vgl. Gredel 2017a, *passim*.

⁷³ Bubenhofer / Rothenhäusler 2016, S. 62.

⁷⁴ Rogers 2013, S. 24.

folgen die korpuslinguistischen Methoden gemäß ihrer disziplinären Herkunft der Verfasstheit der Textoberfläche (»follow the language«). Beide Ansätze haben Vor- und Nachteile: Bei korpuslinguistischen Untersuchungen anhand digitaler Korpora sind die sprachlichen Daten (z.B. aus der Wikipedia) häufig durch linguistische Metadaten angereichert, was komplexe linguistisch informierte Suchanfragen ermöglicht. Allerdings sind die Strukturen der Diskursfragmente (also beispielsweise das Zusammenspiel der Namensräume in der Wikipedia oder die hypertextuelle Verknüpfung verschiedener Sprachversionen der Online-Enzyklopädie) in bisherigen Korpora nicht nachvollziehbar: So sind die Artikel- und Diskussionsseiten in je eigene DeReKo-Korpora überführt. Diagrammatische Operationen im Bereich der Digital Methods bleiben zwar häufig »näher« an der Struktur der Diskursfragmente in digitalen Medien, ermöglichen jedoch aufgrund fehlender linguistischer Metadaten keine komplexen linguistisch informierten Suchabfragen.

Um neue digitale Untersuchungsgegenstände adäquat beschreiben zu können, erscheint deshalb eine Integration beider methodischer Paradigmen im Sinne eines »mixed methods«-Ansatzes aussichtsreich. Die Integration etablierter korpuslinguistischer Methoden mit den in den Sozial- und Medienwissenschaften verbreiteten Digital Methods stellt allerdings noch ein Desiderat dar. Für den Untersuchungsgegenstand der digitalen Plattform Wikipedia wurde in diesem Beitrag der Versuch unternommen, das Potential einer Auswahl digitaler Tools und den damit verbundenen Methoden für diskursanalytische Studien aufzuzeigen. Deutlich wurde, dass die Idee der Anwendung dieser Tools und Methoden ein lohnenswertes Desiderat im Bereich linguistischer Diskursanalysen ist und die Bündelung zu einer Methodologie digitaler Diskursanalysen ein aussichtsreiches Projekt darstellt.

Bibliographische Angaben

Birte Arendt / Philipp Dreesen: Kontrastive Diskurslinguistik. Werkstattbericht zur Analyse von deutschen und polnischen Wikipedia-Artikeln. Standortbestimmung, eine reflektierende Vorbemerkung. [\[Nachweis im GBV\]](#) In: Diskurs interdisziplinär. Zugänge, Gegenstände, Perspektiven. Hg. von Heidrun Kämper / Ingo H. Warnke. Berlin u.a. 2015, S. 427–445. (= Diskursmuster, 6). [\[Nachweis im GBV\]](#)

Michael Reißwenger: Praktiken in der internetbasierten Kommunikation. [\[Nachweis im GBV\]](#) In: Sprachliche und kommunikative Praktiken. Hg. von Arnulf Deppermann / Helmuth Feilke / Angelika Linke. Berlin u.a. 2016, S. 279–310. (= Jahrbuch 2015 des Instituts für Deutsche Sprache). [\[Nachweis im GBV\]](#)

Michael Reißwenger / Angelika Storrer: Corpora of Computer-Mediated Communication. In: Corpus Linguistics. An International Handbook. Hg. von Anke Lüdeling / Merja Kytö. Volume 1. Berlin u.a. 2008, S. 292–308. (= Handbücher zur Sprache und Kommunikationswissenschaft, 29.1). [\[Nachweis im GBV\]](#)

Michael Reißwenger / Maria Ermakova / Alexander Geyken / Lothar Lemnitzer / Angelika Storrer: DeRiK: A German Reference Corpus of Computer-Mediated Communication. In: LLC the journal of digital scholarship in the humanities 28 (2013), H. 4, S. 531–537. [\[Nachweis im GBV\]](#) Siehe auch Preprint PDF: [\[online\]](#)

Sylvia Bendel Larcher: Linguistische Diskursanalyse. Ein Lehr- und Arbeitsbuch. Tübingen 2015. [\[Nachweis im GBV\]](#)

Zofia Bilut-Homplewicz: Prinzip Kontrastivität. Einige Anmerkungen zum interlingualen, intertextuellen und interlinguistischen Vergleich. In: Vom Wort zum Text. Studien zur deutschen Sprache und Kultur. Hg. von Czachur Waldemar / Marta Czyżewska. Warszawa 2008, S. 483–492. [\[Nachweis im GBV\]](#)

Noah Bubenhofer / Stefanie Haupt / Horst Schwinn: A Comparable Corpus of the Wikipedia: From Wiki Syntax to POS Tagged XML. In: Arbeiten zur Mehrsprachigkeit. Serie B 96 (2011) S. 141–144. URN: [urn:nbn:de:bsz:mh39-51897](#)

Noah Bubenhofer / Nicole Müller / Joachim Scharloth: Narrative Muster und Diskursanalyse: Ein datengeleiteter Ansatz. In: Zeitschrift für Semiotik 35 (2013), S. 419–444. [\[Nachweis im GBV\]](#)

Noah Bubenhofer / Klaus Rothenhäusler: »Korporatheken«: Die digitale und verdatete Bibliothek. [\[online\]](#) In: 027.7 Zeitschrift für Bibliothekskultur 4 (2016), H. 2, S. 60–71. [\[online\]](#)

Hans-Jürgen Bucher: Online-Diskurs als multimodale Netzwerk-Kommunikation. Plädoyer für eine Paradigmenweiterung. In: Online-Diskurse. Theorien und Methoden transmedialer Online-Diskursforschung. Hg. von Claudia Fraas / Stefan Meier / Christian Pentzold. Köln 2013, S. 57–101. [\[Nachweis im GBV\]](#)

Dietrich Busse: Historische Semantik. Analyse eines Programms. Stuttgart 1987. (= Sprache und Geschichte, 13). [\[Nachweis im GBV\]](#)

Dietrich Busse / Wolfgang Teubert: Ist Diskurs ein sprachwissenschaftliches Objekt? Zur Methodenfrage der historischen Semantik. In: Begriffsgeschichte und Diskursgeschichte. Methodenfragen und Forschungsergebnisse der historischen Semantik. Hg. von Dietrich Busse / Fritz Hermanns / Wolfgang Teubert. Opladen 1994, S. 10–28. [\[Nachweis im GBV\]](#)

Dietrich Busse / Wolfgang Teubert: Linguistische Diskursanalyse. Neue Perspektiven. Wiesbaden 2013. [\[Nachweis im GBV\]](#)

Waldemar Czachur: Kontrastive Diskurslinguistik. Sprach- und kulturkritisch durch Vergleich. In: Diskurslinguistik im Spannungsfeld von Deskription und Kritik. Hg. von Ulrike Hanna Meinhof / Martin Reisigl / Ingo H. Warnke. Berlin 2013, S. 325–350. [\[Nachweis im GBV\]](#)

Deutsches Referenzkorpus (DeReKo). Ausbau und Pflege der Korpora geschriebener Gegenwartssprache. Hg. vom Institut für Deutsche Sprache. Mannheim 2017. [\[online\]](#)

Ziko van Dijk: Wikipedia. Wie Sie zur freien Enzyklopädie beitragen. München 2010. [\[Nachweis im GBV\]](#)

Ekkehard Felder: Semantische Kämpfe in Wissensdomänen. Eine Einführung in Benennungs-, Bedeutungs- und Sachverhaltsfixierungs-Konkurrenzen. In: Semantische Kämpfe. Macht und Sprache in den Wissenschaften. Hg. von Ekkehard Felder. Berlin u.a. 2006, S. 13–46. [\[Nachweis im GBV\]](#)

Ekkehard Felder: Pragma-semiotische Textarbeit und der hermeneutische Nutzen von Korpusanalysen für die linguistische Mediendiskursanalyse. In: Korpuspragmatik. Thematische Korpora als Basis diskurslinguistischer Analysen. Hg. von Ekkehard Felder / Ekkehard / Marcus Müller / Friedemann Vogel. Berlin u.a. 2012, S. 115–174. [\[Nachweis im GBV\]](#)

Ekkehard Felder: Lexik und Grammatik der Agonalität in der linguistischen Diskursanalyse. [\[Nachweis im GBV\]](#) In: Diskurs - interdisziplinär. Zugänge, Gegenstände, Perspektiven. Hg. von Heidrun Kämper / Ingo H. Warnke. Berlin u.a. 2015, S. 87–121. (= Diskursmuster, 6). [\[Nachweis im GBV\]](#)

Korpuspragmatik. Thematische Korpora als Basis diskurslinguistischer Analysen. Hg. von Ekkehard Felder / Marcus Müller / Friedemann Vogel. Berlin u.a. 2012. [\[Nachweis im GBV\]](#)

Eva Gredel: Diskursdynamiken: Metaphorische Muster zum Diskursobjekt Virus. Berlin u.a. 2014 (= Sprache und Wissen, 17). [\[Nachweis im GBV\]](#)

- Eva Gredel: Digitale Diskursanalysen: Kollaborative Konstruktion von Wissensbeständen am Beispiel der Wikipedia. [\[Nachweis im GBV\]](#) In: Wissensformate in den Medien. Hg. von Sylvia Jaki / Anette Sabban. Berlin 2016, S. 317–339. (= Kultur, Kommunikation, Kontakte, 25). [\[Nachweis im GBV\]](#)
- Eva Gredel (2017a): Digital discourse analysis and Wikipedia: Bridging the gap between Foucauldian discourse analysis and digital conversation analysis. In: Journal of Pragmatics 115 (2017), S. 99–114. [\[Nachweis im GBV\]](#)
- Eva Gredel (2017b): Diskursensitivität von Bildern: Semiotische Strategien in Zeitungsartikeln zu den Olympischen Winterspielen in Sotschi. In: Diskurs - semiotisch. Aspekte multiformaler Diskurskodierung. Hg. von Ernest W.B. Hess-Lüttich / Ingo H. Warnke / Martin Reisigl / Heidrun Kämper. Berlin u.a. 2017, S. 145–163. (= Diskursmuster, 14). [\[Nachweis im GBV\]](#)
- Derya Gür-Şeker: Transnationale Diskurslinguistik. Theorie und Methodik am Beispiel des sicherheitspolitischen Diskurses über die EU-Verfassung in Deutschland, Großbritannien und der Türkei. Bremen 2012 (= Sprache, Politik, Gesellschaft, 6). [\[Nachweis im GBV\]](#)
- Susan Herring: Discourse in Web 2.0: familiar, reconfigured, and emergent. In: Discourse 2.0. Language and New Media. Hg. von Deborah Tannen. Washington, DC. 2013, S. 1–25. [\[Nachweis im GBV\]](#)
- Diskurs – semiotisch. Aspekte multiformaler Diskurskodierung. Hg. von Ernest W. B. Hess-Lüttich / Ingo H. Warnke / Martin Reisigl / Heidrun Kämper. Berlin u.a. 2017. (= Diskursmuster, 14). [\[Nachweis im GBV\]](#)
- Heidrun Kämper: Diskurs und Diskurslexikographie. Zur Konzeption eines Wörterbuchs des Nachkriegsdiskurses. URN: [urn:nbn:de:bsz:mh39-36215](#) In: Deutsche Sprache 34 (2006), H. 4, S. 334–353. [\[Nachweis im GBV\]](#)
- Heidrun Kämper: Personen als Akteure. In: Handbuch Sprache in Politik und Gesellschaft. Hg. von Kersten Roth / Martin Wengeler / Alexander Ziem. Berlin u.a. 2017, S. 259–279. (= Handbücher Sprachwissen, 19). [\[Nachweis im GBV\]](#)
- Heidrun Kämper / Ingo H. Warnke / Daniel Schmidt-Brücken: Diskursive Historizität. In: Textuelle Historizität. Interdisziplinäre Perspektiven auf das historische Apriori. Hg. von Heidrun Kämper / Ingo H. Warnke / Daniel Schmidt-Brücken. Berlin u.a. 2016, S. 1–8. (= Diskursmuster, 12). [\[Nachweis im GBV\]](#)
- Konstanze Marx / Georg Weidacher: Internetlinguistik. Ein Lehr- und Arbeitsbuch. Tübingen 2014. [\[Nachweis im GBV\]](#)
- Lothar Lemnitzer / Heike Zinsmeister: Korpuslinguistik. Eine Einführung. Tübingen 2006. [\[Nachweis im GBV\]](#)
- Eliza Margaretha / Harald Lungen: Building linguistic corpora from Wikipedia articles and discussions. URN: In: Journal for Language Technology and Computational Linguistics 2 (2014), S. 59–82. PDF. [\[online\]](#)
- Wolf-Andreas Liebert: Die transdiskursive Vorstellungswelt zum Aids-Virus. Heterogenität und Einheiten von Textsorten im Übergang von Fachlichkeit und Nichtfachlichkeit. In: Fachliche Textsorten. Komponenten, Relationen, Strategien. Hg. von Hartwig Kalverkämper / Klaus-Dieter Baumann. Tübingen 1996, S. 789–811. [\[Nachweis im GBV\]](#)
- Henning Lobin: Computerlinguistik und Texttechnologie. München 2010. (= UTB, 3282). [\[Nachweis im GBV\]](#)
- Thomas Niehr: Einführung in die linguistische Diskursanalyse. Darmstadt 2014. [\[Nachweis im GBV\]](#)
- Harald Lungen / Marc Kupitz: CMC Corpora in DeReKo. URN: [urn:nbn:de:bsz:mh39-62592](#) In: Proceedings of the Workshop on Challenges in the Management of Large Corpora and Big Data and Natural Language Processing. (CMLC: 5, Cfp: 1, BigNLP 2017, Birmingham, 24.07.2017) Birmingham 2017, S. 20–24. URN: [urn:nbn:de:bsz:mh39-62434](#)
- Rainer Perkuhn / Holger Keibel / Marc Kupietz: Korpuslinguistik. Paderborn 2012. [\[Nachweis im GBV\]](#)
- Christian Pentzold / Esther Weltevrede / Michele Mauri / David Laniado / Andreas Kaltenbrunner / Erik Borra: Digging Wikipedia. The online encyclopedia as digital cultural heritage gateway and site. PDF. [\[online\]](#) In: ACM Journal on Computing and Cultural Heritage 10 (2017), H. 1, Artikel 5. [\[Nachweis im GBV\]](#)
- Richard Rogers: Digital Methods. Cambridge 2013. [\[Nachweis im GBV\]](#)
- Peter Sandrini: Kohärenz in mehrsprachigen Webauftritten. In: »Es geht sich aus ...« zwischen Philologie und Translationswissenschaft. Festschrift für Wolfgang Pöckl. Hg. von Peter Holzer / Carolin Feyrer / Vanessa Gampert. Frankfurt/Main 2012, S. 243–252. [\[Nachweis im GBV\]](#)
- Jürgen Spitzmüller / Ingo H. Warnke: Diskurslinguistik. Eine Einführung in Theorien und Methoden der transtextuellen Sprachanalyse. Berlin u.a. 2011. [\[Nachweis im GBV\]](#)
- Angelika Storrer: Hypertextlinguistik. In: Textlinguistik. 15 Einführungen. Hg. von Nina Janich. Tübingen 2008, S. 315–331. [\[Nachweis im GBV\]](#)
- Angelika Storrer: Neue Text- und Schreibformen im Internet: Das Beispiel Wikipedia. Siehe auch PDF-Preprint: [\[online\]](#) In: Textkompetenzen für die Sekundarstufe II. Hg. von Helmuth Feilke / Juliane Köster. Freiburg 2012, S. 277–304. [\[Nachweis im GBV\]](#)
- Ingo H. Warnke: Urbaner Diskurs und maskierter Protest – Intersektionale Feldperspektiven auf Gentrifizierungsdynamiken in Berlin Kreuzberg. [\[Nachweis im GBV\]](#) In: Angewandte Diskurslinguistik. Felder, Probleme, Perspektiven. Hg. von Roth Kersten / Carmen Spiegel. Berlin 2013, S. 189–221. [\[Nachweis im GBV\]](#)
- Ingo H. Warnke: Diskurs. [\[Nachweis im GBV\]](#) In: Handbuch Sprache und Wissen. Hg. von Ekkehard Felder / Andreas Gardt. Berlin u.a. 2015, S. 221–241. [\[Nachweis im GBV\]](#)

Ingo H. Warnke/ Janina Wildfeuer/ Daniel Schmidt-Brücken / Wolfram Karg: Diskursgrammatik als wissensanalytische Sprachwissenschaft. [\[Nachweis im GBV\]](#) In: Kommunikation Korpus Kultur. Ansätze und Konzepte einer kulturwissenschaftlichen Linguistik. Hg. von Nora Benitt / Christopher Koch / Katharina Müller / Sven Saage / Lisa Schüler. Trier 2014, S. 47–67. [\[Nachweis im GBV\]](#)

Ingo H. Warnke / Jürgen Spitzmüller: Methoden der Diskurslinguistik. Sprachwissenschaftlicher Zugänge zur transtextuellen Ebene. Berlin u.a. 2008. [\[Nachweis im GBV\]](#)

Esther Weltverde: Repurposing digital methods. The research affordances of platforms and engines. Amsterdam 2016. [\[online\]](#)

Taha Yasseri / Anselm Spoerri / Mark Graham / Janos Kertesz: The Most Controversial Topics in Wikipedia: A Multilingual and Geographical Analysis. In: Global Wikipedia: International and Cross-Cultural Issues in Online Collaboration. Hg. von Pnina Fichman / Noriko Hara. Lanham, MD. u.a. 2014. [\[Nachweis im GBV\]](#)

Abbildungsnachweise und -legenden

Abb. 1: Eigenschaften und Strukturen der Wikipedia als Hypertext. Quelle: © Wikipedia, 2017.

Abb. 2: Schematische Darstellung eines Diskursfragments in der Wikipedia. © Eigene Darstellung, 2017.

Abb. 3: Tool »Wikipedia Cross-lingual Image Analysis«. © Digital Methods Initiative.

Abb. 4: Funktionsweise des Tools whoColor. © f-squared.org.

Abb. 5: Das Tool Search Cristal visualisiert Edit Wars. © [searchCrystal](https://searchcrystal.com/).