

Das LDA-Toolkit¹

Korpuslinguistisches Analyseinstrument für kontrastive Diskurs- und Imageanalysen in Forschung und Lehre

FRIEDEMANN VOGEL

Abstract

This contribution discusses the function and use of a new corpus linguistic freeware (under Creative Commons Licence), called „LDA-Toolkit“. After a brief introduction, the tools and their functionality are presented, followed by two examples of analyses. The LDA-Toolkit has been especially developed for linguistic discourse and image analysis for use in research and teaching. The handling is easy (GUI) and does not require any programming skills. Its tools contain a token list, a concordance tool, keyword analysis, contrastive key-cluster, key-ngram and key-pos-gram analysis, contrastive co-occurrence analysis and easy auxiliary devices to visualize results as word nets or word clouds. All tools as well as all results of the analyses are concentrated and recorded in a virtual working platform called „LDA tree“ or „LDA explorer“, because the results are visualized in a hierarchic tree model of co(n)texts. Furthermore, the records of this main tool can easily be saved as status of work and reloaded for later analysis. The functionality of the LDA-toolkit is exemplified by two examples: the first example illustrates a corpus-driven method of linguistic discourse study using keyword and contrastive ngram analysis. It focuses on the legal debate about euthanasia and advanced health care directives in Germany. The second example outlines a corpus-assisted approach to analyzing public media images and stereotypes. It examines the German pirate party's image („Piratenpartei“) in German newspapers.

1. Einführung

Diskursanalysen interessieren sich seit den grundlegenden Arbeiten von M. Foucault (1973, 1974a, 1974b) vor allem für die Prämissen und Prozeduren gesellschaftlicher Episteme und finden sich mittlerweile in zahlreichen geistes- und sozialwissenschaftlichen Disziplinen². Ähnliches gilt für die Untersuchung von Images als – häufig strukturalistisch missver-

standene – sprachliche ‚Abbilder‘ einer Person oder eines Sachverhalts in ‚der‘ Öffentlichkeit. Auch LinguistInnen wenden sich seit nunmehr 2 Jahrzehnten der theoretischen Modellierung und Untersuchung von Diskursen (vgl. A. Gardt 2007; I. Warnke (Hrsg.) 2007, I. Warnke & J. Spitzmüller (Hrsg.) 2008) und Images (F. Vogel 2010a, 2010b) zu. Die Besonderheit des linguistischen Zugangs liegt in der Fokussierung auf die sprachlichen Konstitutionsbedingungen von akteurs- und gruppenbezogenen Wissensrahmen und Geltungsansprüchen in Diskursen und damit zugleich in dem Versuch, die einzelnen Interpretationsakte an den sprachlichen Strukturen im Ko(n)text plausibel zu machen. Damit einher geht in der Regel eine Kritik an nicht-linguistischen Analyseansätzen, ihre hermeneutische Arbeit häufig nicht transparent zu machen oder implizit für Ausdruck und Inhalt ein ontisches Abbildverhältnis anzunehmen. Die mittlerweile kaum mehr überschaubaren Studien, zu denen auch erste Lehrbücher zählen (vgl. J. Spitzmüller & I. Warnke 2011), arbeiten hierbei überwiegend qualitativ und auf unterschiedlichen Strukturebenen vom Morphem bis hin zu intertextuellen Verweisungsstrukturen (jedoch mit einem Überhang an lexikalisch orientierten Ansätzen).

Durch Anregungen aus der Korpuslinguistik wurden seit einigen Jahren³ auch erste quantitative Ansätze zur Analyse von Diskursen und Images auf Basis sehr großer Korpora entwickelt (vgl. hierzu im Überblick: E. Felder, M. Müller & F. Vogel (Hrsg.) 2012). Hierzu zählen insbesondere Arbeiten von W. Teubert⁴, der Corpus-Assisted Discourse Studies (vgl. A. Partington, J. Morley & L. Haarman (Hrsg.) 2004), der Critical Discourse Analysis (CDA; vgl. P. Baker 2006 und P. Baker et al 2008), der Mediendiskursanalyse im Paradigma der pragma-semiotischen Textarbeit (E. Felder 2012) sowie auch neuere Ansätze der Forschungsgruppe *semtracks*⁵.

Voraussetzung hierfür waren und sind zahlreiche neue computergestützte Hilfsmittel und Instrumente (Algorithmen, Software), die in erster Linie einem Zweck dienen: der automatischen Berechnung und Identifizierung jener sprachlichen Muster, die in einer einzeltextorientierten und intuitiv geleiteten Untersuchung gerade nicht oder nur zufällig aufgefunden werden könnten. Die besondere Leistungsfähigkeit der verschiedenen Hilfsmittel liegt dabei in der kontrastiven Visibilisierung von Ko(n)textstrukturen. Das methodische Set umfasst u. a. ausgefeilte Suchmaschinen und komplexe Suchanfragen, sortierte Konkordanzen (zeitweise Darstellung von KWICs⁶), n-Gramm- und Cluster-Analysen⁷ sowie Kollokations- bzw. Kookkurrenzanalysen.

Alle frühen korpuslinguistischen Zugänge zu sprachlichen Strukturen wenden sich implizit oder explizit gegen eine „Lehnstuhllinguistik“, gegen Untersuchungsansätze also, die ihre Interpretationsergebnisse allein auf die Introspektion des Analytikers stützen⁸. In der Hoffnung auf eine

Objektivierung der eigenen Arbeiten haben viele Diskurslinguisten diese Kritik mittlerweile aufgenommen und sich die Tools der Korpuslinguistik zu Eigen gemacht. – Die Vielzahl neuer technischer Instrumente spiegelt sich dabei auch in der Vielfalt ihrer (nicht immer systematischen) Anwendungen wieder. Mit anderen Worten: Der korpuslinguistisch inspirierten Diskursanalyse fehlt es bislang oft an einem ausgeglichenen methodologischen Verhältnis von quantitativ-Möglichem und qualitativ-Notwendigem.

Ein Grund hierfür könnte darin liegen, dass bislang kein Analyseinstrument speziell für diskurslinguistische Fragestellungen bereitsteht, sondern wenn, dann nur auf allgemeine Sprachanalysen hin ausgerichtete Software Verwendung findet. Diese Software ist jedoch häufig keine Freeware, in ihrer Funktionalität eingeschränkt, schlecht bis gar nicht dokumentiert oder erfordert zumindest rudimentäre Programmierkenntnisse. Diese Umstände tragen wohl dazu bei, dass entsprechende Instrumente zum Zwecke von Diskursanalysen nur punktuell – und in der Lehre noch immer selten⁹ – eingesetzt, geschweige denn in eine native Methodologie integriert werden.

Vor diesem Hintergrund wird am Germanistischen Seminar der Universität Heidelberg und mit Unterstützung der Heidelberger Computerlinguistik¹⁰ ein Windows-basiertes Programm entwickelt,

- das speziell für kontrastive diskurslinguistische Fragestellungen zugeschnitten ist,
- bisherige quantitative Diskursanalyseansätze der Korpuslinguistik aufgreift (wie Keyword-, Key-Cluster- und Key-N-Gramm-Analyse, vgl. hierzu etwa K. Senkbeil 2012, N. Bubenhofer 2009),
- in einer intuitiv bedienbaren Benutzeroberfläche auch für die Lehre eingesetzt werden können soll, und
- für nicht-kommerzielle Nutzung kostenlos zum Download¹¹ angeboten wird.

Im Folgenden wird zunächst die Funktionalität des Toolkit vorgestellt (2) sowie im Anschluss anhand der Bearbeitung zweier exemplarischer diskurslinguistischer Fragestellungen illustriert (3).

2. Das LDA-Toolkit: Korpuslinguistische Arbeitsumgebung „zum Klicken statt Tippen“¹²

Das LDA-Toolkit wird von Linguisten für Linguisten entwickelt und besteht aus mehreren, unabhängig von einander nutzbaren Teilkomponenten:

- LDA-Baum bzw. „Kotext-Explorer“: Auf die analytische Praxis hin organisierte Arbeitsumgebung durch Zusammenführung der nachfolgenden Analyse-Tools und Darstellung deren Ergebnisse in einer hierarchischen Baumstruktur, wahlweise unter Verwendung von statistischen Signifikanztests.
- Tokenlist: Frequenzanalyse aller Ausdrücke (Token vs. Lemma) in einem Korpus.
- Keyword-Analyse: Statistischer Vergleich von Ausdrucksfrequenzlisten im Hinblick auf signifikante Typik einzelner Ausdrücke für die jeweiligen Vergleichskorpora.
- Kontrastive (Key-) Cluster-, (Key-) N-Gramm- und (Key-) POS-Gramm-Analyse: Frequenzanalyse und wahlweise statistische Auswertung von Mehrworteinheiten bzw. annotierten Wortartfolgen mit und ohne Referenzausdruck.
- Kontrastive Kookkurrenzanalyse: Wahlweise einzelkorpusbezogene, korpusvergleichende oder profilvergleichende Frequenzanalyse und statistische Auswertung von Ausdrücken im Kotext eines Bezugswortes.
- KWIC: Einfache Konkordanzansicht zu einem Suchausdruck¹³.
- Visualisierung: Semiautomatische Code-Generierung (für den Export) und Visualisierung von Belegen und hierarchischen Kotextstrukturen als Wort-Wolke sowie als Wort-Netz.

Die einzelnen Tools erlauben spezifische Kotexteinschränkungen, komfortable Filter-, Sortier- und Exportmöglichkeiten sowie die Berücksichtigung von Annotationen (nach Token, Lemma und Part-of-Speech-Tags). Als Signifikanztests stehen derzeit einseitige T- sowie Chi-Square-Tests zur Verfügung¹⁴. Alle Einzeltools können über eine Schnittstelle in eigene Skripte oder Programme eingebunden werden.

Die LDA-Tools arbeiten unter zwei Voraussetzungen: Erstens können derzeit nur deutschsprachige und mit dem Stuttgarter TreeTagger¹⁵ Wortart-annotierte Korpora geladen werden. Zweitens werden die Texte bislang ausschließlich in Minuskeln verarbeitet (also keine Unterscheidung von Groß- und Kleinschreibung). Letzteres wird in zukünftigen Versionen als optionale Funktion (Normalschreibung oder Minuskelverarbeitung) zur Verfügung stehen.

2.1 *Der LDA-Baum („Kotext-Explorer“)*

Der LDA-Baum ist die zentrale Arbeitsumgebung des Toolkit, in der die einzelnen Analysetools in Registerkarten übersichtlich eingebunden und ihre Ergebnisse in einer Baumstruktur abgelegt werden. Die Darstellung

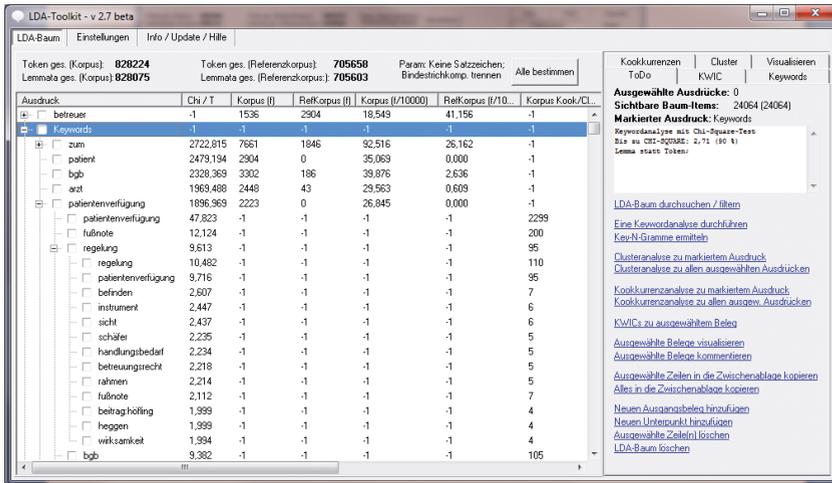


Abbildung 1. Screenshot – LDA-Baum mit Illustration hierarchischer Kotextabhängigkeiten.

der Ergebnisse in einer Baum-Struktur hat verschiedene Vorteile: Erstens können eine Vielzahl an unterschiedlichen Einzelergebnissen (Wort- und Frequenzlisten einschließlich Kommentierung und Dokumentation der jeweils zugrunde gelegten Analyseparameter) übersichtlich dargestellt und weiterverarbeitet werden; eine ‚Zusammenschau‘ verschiedener Teilergebnisse wird dadurch gefördert. Zweitens können hierarchische Kotextabhängigkeiten direkt in der Baumstruktur visualisiert werden: damit wird es zum Beispiel möglich, mittels mehrgliedriger Kookkurrenzanalysen systematisch Kotexte und Kotextüberschneidungen einzelner Ausdrücke nachzuvollziehen und Textbezüge transparent zu dokumentieren. Drittens schließlich lassen sich die Ergebnisse der LDA-Bäume in dieser Form komfortabel als „Arbeitsstände“ abspeichern und bei späterer Weiterarbeit erneut laden.

Zur Bearbeitung des LDA-Baums stehen verschiedene Hilfsfunktionen zur Verfügung, unter anderem:

- Automatische („Batch“-) Analysen zu ausgewählten Ausdrücken: Um wiederholte Wartezeiten in Folge aufwendiger Einzelanalysen zu vermeiden, ist es möglich, zum Beispiel Kookkurrenzanalysen zu den Ausgangsausdrücken vorzubereiten und anschließend den Analysevorgang automatisiert arbeiten zu lassen. Die Einzelergebnisse werden dann im Hintergrund im LDA-Baum hierarchisch zu den Aus-

Ausdruck	Chi / T	Kommentar
Keywords	-1	
patient	2479,194	Akteur
arzt	1969,488	Akteur
patientenverfügung	1896,969	Normtext
regelung	9,613	
schäfer	2,235	Akteur
heggen	1,999	Akteur
wille	8,951	
patient	5,087	Akteur
betroffene	4,537	Akteur
betreute	2,213	Akteur
arzt	7,109	Akteur
pflegepersonal	3,157	Akteur
patient	2,678	Akteur
angehörige	1,981	Akteur
betreuer	5,929	Akteur
arzt	4,400	Akteur
patient	3,201	Akteur
vormundschaftsgericht	2,423	Akteur
bosbach	3,925	Akteur / Politik
vertreter	3,663	Akteur

Abbildung 2. Screenshot – Filterung und Kommentierung von Datensätzen.

gangsausdrücken abgelegt und stehen nach Ende der Analyse en block zur Verfügung.

- Sortierung nach Ausdruck (alphabetisch), Korpusvorkommen (Frequenz) und Signifikanzwerten;
- Suche und Filterung von angelegten Datensätzen zur Unterstützung der qualitativen Auswertung der statistischen Analyseergebnisse;
- Kommentierung: Jeder Beleg im LDA-Baum kann manuell oder durch vorherige Filterung automatisch mit zusätzlichen Metainformationen kommentiert werden. Diese Kommentare selbst lassen sich anschließend wiederum filtern und erleichtern damit die Klassifizierung von einzelnen Belegen nach qualitativen Kriterien.
- Hinzufügen einzelner Ausgangs- und Subknoten: Im LDA-Baum werden üblicherweise automatisch generierte Analyseergebnisse abgelegt. Daneben lassen sich jedoch auch manuelle Datensätze und Knotenhierarchien anlegen, etwa zum Zwecke von Batch-Analysen (vgl. o.) oder zur anschließenden semiautomatischen Visualisierungen (ohne Programmierkenntnisse).
- Export: Schließlich lassen sich einzelne, mehrere oder alle Datensätze in Form von Tabellen (zur direkten Übertragung copy & paste in externe Tabellenkalkulationsprogramme) oder kommasspariert in den Arbeitsspeicher oder als Textdatei (*.csv) exportieren und abspeichern.

2.2 Tokenlist

Das Tool erlaubt einfache Ausdrucksfrequenzanalysen. Ausdrücke können sowohl als Token als auch als Lemmata (Grundformen nach der Lemmatisierung des TreeTagger¹⁶) und in Abhängigkeit von Wortarten gezählt und als Textdatei gespeichert werden. Das folgende Beispiel zeigt den Text-Output des Tools, einerseits die berücksichtigten Analyseparameter (u. a. Berücksichtigung nur von Substantiven und Eigennamen), andererseits die 10 häufigsten Belege der Ergebnisliste¹⁷:

```
<Metadata>
<Types>2558</Types>
<Token>9633</Token>
<Quellverzeichn.>C:\...\taz</Quellverzeichn.>
<LemmaStattToken>1</LemmaStattToken>
<KeineSatzzeichen>1</KeineSatzzeichen>
<BindestrichkompositaGetrennt>1</BindestrichkompositaGetrennt>
<POS-Filter>#NN#NE</POS-Filter>
</Metadata>
```

Lemma	Frequenz
<i>bka</i>	297
<i>gesetz</i>	269
<i>spd</i>	151
<i>computer</i>	148
<i>schäuble</i>	94
<i>online</i>	90
<i>durchsuchung</i>	78
<i>polizei</i>	75
<i>innenminister</i>	73
<i>bundesrat</i>	72

2.3 Keyword-Analyse

Unter „Keywords“ werden im Rahmen des LDA-Toolkit Ausdrücke verstanden, die in Abhängigkeit zu einem frei wählbaren Signifikanzniveau signifikant häufiger in einem Primär- als in einem Referenzkorpus vorkommen. Zu diesem Zweck werden Ausdrucksfrequenzlisten (Token vs. Lemma) sowohl zu Primär- als auch Referenzkorpus erstellt und kontrastiv mittels Chi-Square-Signifikanztest miteinander verglichen. Mittels vorheriger Angabe eines Signifikanzniveaus lassen sich die Ergebnisse sortieren und eingrenzen. Das Tool erlaubt ferner auch eine Filterung nach Wortarten, so dass etwa Verben oder Substantive ermittelt werden können, die ‚typisch‘ für das eine und ‚untypisch‘ für das Vergleichskorpus sind.

Das Prinzip der kontrastiven Keyword-Analyse bildet eine zentrale Säule für induktive, corpus driven orientierte Untersuchungsansätze. Sie ermöglicht die Herausarbeitung systematischer Unterschiede zwischen verschiedenen Korpora und damit eine datenbasierte anstelle einer deduktiv geleiteten Hypothesengenerierung (vgl. hierzu exemplarisch K. Senkbeil 2012).

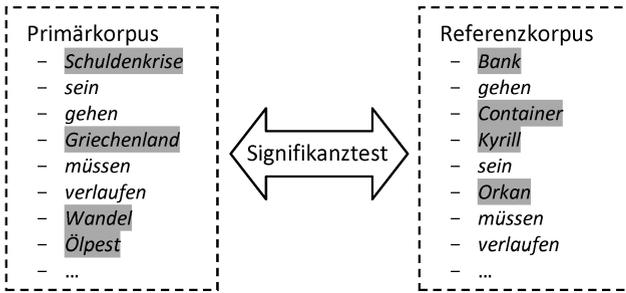


Abbildung 3. Veranschaulichung der Funktionsweise der Keywordanalyse.

2.4 (Key-) Cluster-, (Key-) N-Gramm- und (Key-) POS-Gramm-Analyse

Genauso wie Keywords lassen sich auch Key-Mehrworteinheiten als Indices für korpuspezifische, thematische Schlüsselwörter und Schlüssel-syntagmen (im Sinne F. Hermanns 1994) ermitteln. Die automatische Cluster-Analyse des LDA-Toolkit umfasst hierfür verschiedene Parameter:

- Cluster-Analyse im engeren Sinne: Gesucht und zählend gruppiert werden Mehrworteinheiten beliebiger Größe, in denen ein zuvor gewählter Ausgangsausdruck enthalten ist. Beispiel: Die 10 häufigsten Cluster der Größe 3 mit dem Basisausdruck *menschlich* in einem Korpus aus juristischen Fachtexten zur Fachdebatte um ‚Sterbehilfe‘ und ‚Patientenverfügung‘:

Cluster (Token)	Frequenz
<i>des menschlichen lebens</i>	38
<i>das menschliche leben</i>	22
<i>schutz des menschlichen</i>	17
<i>menschlichen lebens und</i>	14
<i>grenzbereich menschlichen lebens</i>	13
<i>schutz menschlichen lebens</i>	11
<i>im grenzbereich menschlichen</i>	10
<i>menschlichen lebens vorrang</i>	10
<i>schutzes menschlichen lebens</i>	7
<i>des schutzes menschlichen</i>	7

Cluster (Lemmata)	Frequenz
<i>d menschlich leben</i>	64
<i>schutz menschlich leben</i>	18
<i>menschlich leben und</i>	16
<i>grenzbereich menschlich leben</i>	13
<i>d schutz menschlich</i>	13
<i>im grenzbereich menschlich</i>	10
<i>menschlich leben vorrang</i>	10
<i>menschlich leben in</i>	9
<i>menschlich leben ein</i>	7
<i>d menschlich würde</i>	6

- N-Gramm-Analyse: Das Korpus wird in Mehrworteinheiten (MWE) beliebiger Größe aufgeteilt und anschließend gruppiert und gezählt. Beispiel: Die 10 häufigsten Quartgramme in einem Korpus aus juristischen Fachtexten zur Fachdebatte um ‚Sterbehilfe‘ und ‚Patientenverfügung‘:

4-Gramme (Token)	Frequenz
§ bgb § 1904	341
art . gg artikel	307
des § bgb §	182
bgb § 1901a absatz	168
§ bgb § 1901	152
bgb § 1904 absatz	148
§ bgb § 1901a	143
im hinblick auf die	109
2 gg artikel 2	97
die entwicklung des arztrechts	85

4-Gramme (Lemmata)	Frequenz
§ bgb § @card@	592
medr jahr @card@ seite	441
§ stgb § @card@	421
bgb § @card@ absatz	312
art . gg artikel	307
d § bgb §	188
im hinblick auf d	168
d mutmaßlich wille d	167
d § stgb §	147
nach § bgb §	139

- Als Ausgangsdaten lassen sich sowohl Token und Lemmata als auch Part-of-Speech-Tags (annotierte Wortarten) der Analyse zu Grunde legen. Das nachfolgende Beispiel zeigt die 5 häufigsten Tri-POS-Gramme eines juristischen Fachtextkorpus mit exemplarischer Auflö- sung in Token-Belege:

Tri-POS-Gramme	Frequenz	Exemplarische Übersetzung in Token
nn art nn	23843	bundesministerium[nn] der[art] justiz[nn]
nn card nn	22464	artikel[nn] 20[card] gg[nn]
art adja nn	18198	die[art] rechtzeitige[adja] errichtung[nn]
art nn art	15875	die[art] handreichungen[nn] des[art]
appr art nn	15574	durch[appr] das[art] vormundschaftsgericht[nn]

- Sowohl Cluster- als auch N-Gramm-Analysen können automatisch kontrastiv – also vergleichend zu einem Referenzkorpus – ermittelt werden. Die Logik des Algorithmus entspricht der von Keyword-Analysen (einschließlich Anwendung von Chi-Square-Signifikanz- test); das Ergebnis sind also korpuspezifische Key-Cluster bzw. Key- N-Gramme. Das folgende Schaubild illustriert diese Funktion am Beispiel einer kontrastiven Bigrammanalyse auf Grundlage zweier ju- ristischer Fachtextkorpora¹⁸. Die kontrastive Signifikanzprüfung tilgt jene Bigramme, die in beiden Korpora relativ ähnlich verteilt sind:

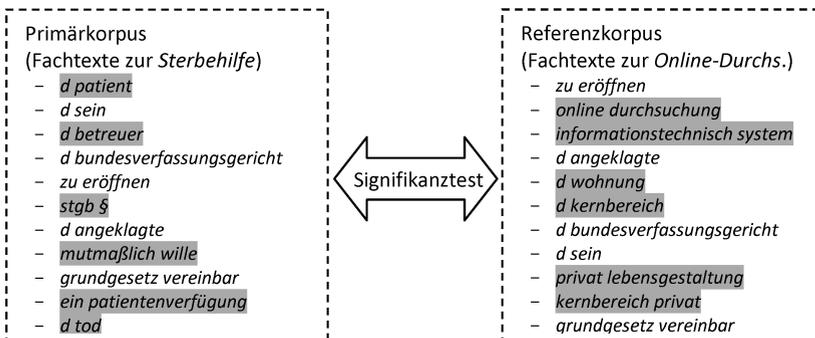


Abbildung 4. Veranschaulichung der Funktionsweise kontrastiver Key-N-Gramm-Analyse.

Cluster- und N-Gramm-Analysen können schließlich auch nach einer automatischen vorherigen Filterung von ausgewählten Wortarten bzw. POS-Tags erfolgen. Auf diesem Wege lassen sich etwa alle Synsemantika ausblenden und damit signifikante thematische Cluster auf Basis häufig (nicht unmittelbar, sondern über die Herausfilterung von Synsemantika entstandenen Leerstellen) aufeinander folgender Autosemantika ermitteln. Das folgende Beispiel zeigt etwa das Ergebnis einer Trigramm-Analyse unter alleiniger Berücksichtigung von Vollverben bzw. Substantiven im juristischen Fachkorpus zur ‚Sterbehilfe‘-Debatte (zu ihrer diskursanalytischen Deutung siehe unten, Kap. 3):

Trigramme (Lemmata) aus Verben	Frequenz	Trigramme (Lemmata) aus Nomen	Frequenz
<i>folgen befinden regeln</i>	20	<i>art artikel gg</i>	267
<i>bestehen sterben erleiden</i>	12	<i>artikel gg artikel</i>	231
<i>gelten folgen befinden</i>	12	<i>gg artikel absatz</i>	211
<i>lassen folgen befinden</i>	10	<i>bverfge bverfge jahr</i>	195
<i>ansehen folgen befinden</i>	8	<i>gesetz überprüfung lebensbeendigung</i>	30
<i>feststehen nehmen widersprechen</i>	8	<i>grundsatz bundesärztekammer sterbebegleitung</i>	30
<i>entscheiden folgen befinden</i>	8	<i>tötung verlangen stgb</i>	28
<i>verlangen respektieren verlieren</i>	7	<i>einwilligung betreuer behandlungsabbruch</i>	27
<i>rechtfertigen entsprechen dienen</i>	7	<i>ethik recht medizin</i>	27
<i>indizieren anbieten einwilligen</i>	7	<i>verlangen hilfe selbsttötung</i>	25

Statistische Analysen wie diese sind mit Hilfe des LDA-Toolkit mit wenigen Mausklicks und ohne Programmierkenntnisse realisierbar.

2.5 (Kontrastive) Kookkurrenzanalyse

Kookkurrenzen lassen sich mit L. Lemnitzer & H. Zinsmeister (2006: 197) beschreiben als „gemeinsame Vorkommen zweier oder mehrerer Wörter in einem Kontext von fest definierter Größe [...]. Das gemeinsame Vorkommen sollte höher sein, als bei einer Zufallsverteilung erwartbar wäre.“ Die Kookkurrenzanalyse des LDA-Toolkit ermöglicht die Berechnung von Kookkurrenzen in diesem Sinne, jedoch mit unterschiedlichen Modi: Zunächst können Kookkurrenzpartner (wahlweise gefiltert nach Wortarten bzw. POS-Tags) zu einem Ausgangswort und ausschließlich im Primärkorpus ausgezählt werden. Ferner lassen sich die Frequenzen der einzelnen Kookkurrenzpartner im Primärkorpus mittels Einstichproben-T-Test auf ihre Signifikanz prüfen (so auch möglich bei AntConc, vgl. L. Anthony 2005). Der Nachteil dieses Verfahrens ist jedoch, dass hierzu abstrakte Annahmen über wahrscheinliche Verteilungen der Ausdrücke in ein- und demselben Korpus getroffen werden müssen und folglich keine kontrastive Analyse (wie etwa bei Keywords) erfolgt.

Eine tatsächlich kontrastive Analyse lässt sich im LDA-Toolkit schließlich mittels Chi-Square-Test ebenso durchführen, indem zur Kon-

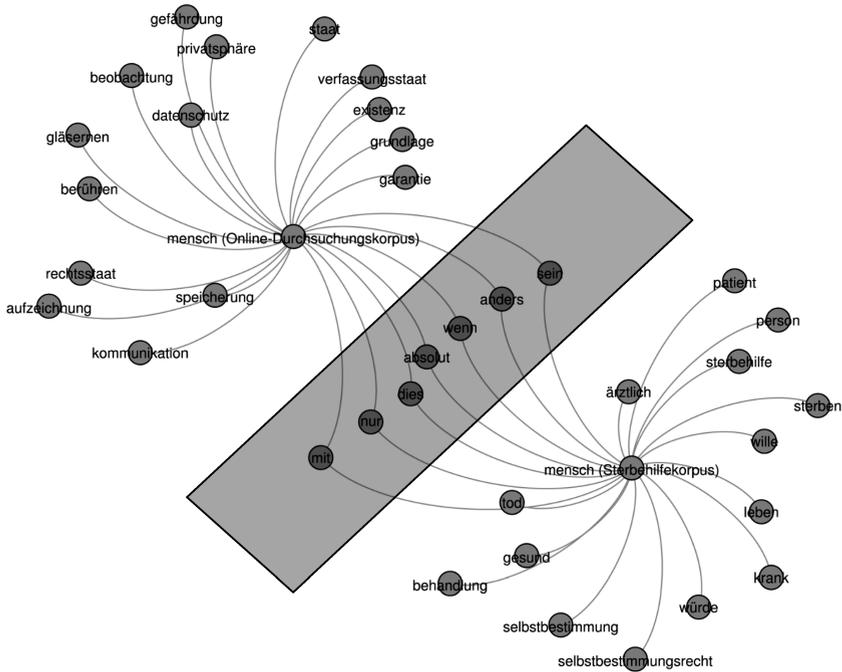


Abbildung 5. Veranschaulichung der funktionsweisen kontrastiven Kookkurrenzanalyse.

trastierung a) ein Referenzkorpus, oder b) ein anderer Ausdruck des Primärkorpus herangezogen wird. Im ersten Fall werden korpusbedingte Unterschiede der Kookkurrenzprofile deutlich, im zweiten Fall lassen sich die Wortfelder zweier unterschiedlicher Ausdrücke und deren Kontextprofile gegenseitig abgrenzen und konturieren. Das heißt konkret, dass potentiell genau diejenigen Kookkurrenzpartner signifikant hervorgehoben werden, die typisch für das Kotextprofil A, nicht aber für das Kotextprofil B sind. Oder umgekehrt formuliert: Diejenigen Kookkurrenzpartner, die in beiden Kotextprofilen gleichermaßen auftreten, fallen durch die Signifikanzprüfung heraus. Die Vorteile dieses Vorgehens sind jedoch an eine wichtige Bedingung geknüpft: Es bedarf in jedem Einzelfall sehr genauer Überlegungen im Hinblick auf die Auswahl der jeweiligen Kontrastwörter. Das obige Schaubild illustriert die kontrastive Kookkurrenzanalyse zum Ausdruck *mensch* (einschließlich Komposita) zweier juristischer Fachtextkorpora und dabei die korpusstypischen sowie korpusunspezifischen Kookkurrenzpartner (Abb. 5).

Dieses kontrastive Verfahren zur Ermittlung von ‚typischem Gebrauch‘ eines Ausdrucks über die Analyse ‚seiner typischen Kotexte‘ im

3. Exemplarische Analysen mit Hilfe des LDA-Toolkit

Im Folgenden wird die Funktionalität des LDA-Toolkit am Beispiel zweier exemplarischer Analysen illustriert und dabei insbesondere der Mehrwert von quantitativen Zugängen zur Unterstützung qualitativer Diskurs- und Imageanalysen deutlich gemacht. Der Fokus der Analysen liegt daher nicht auf einer umfassenden Beschreibung der Phänomene, sondern auf Fragen des Untersuchungsdesigns unter Rückgriff auf das LDA-Toolkit. Das erste Beispiel widmet sich der diskurslinguistischen Untersuchung des Rechtsdiskurses zur sog. *Sterbehilfe* (3.1); das zweite Beispiel illustriert den korpusgestützten Zugang linguistischer Stereotypen- bzw. Imageanalysen anhand des Medienbildes der *Piratenpartei* im Kontrast zur *CDU* (3.2).

3.1 Korpusgestützte Diskurslinguistik am Beispiel des juristischen Fachdiskurses zur ‚Sterbehilfe‘

Die nachfolgende exemplarische Analyse ist Teil einer größeren Untersuchung zum juristischen Binnen- und öffentlichen Mediendiskurs zu ‚Sterbehilfe‘ und ‚Palliativmedizin‘ unter Beteiligung von Prof. Dr. Ekkehard Felder und Janine Luth am Germanistischen Seminar der Universität Heidelberg²². Das leitende Erkenntnisinteresse dieses diskurslinguistischen Projektes lässt sich in folgenden drei Fragen zusammenfassen:

- Welche Sachverhalte und Referenzobjekte sind in der Debatte um ‚Sterbehilfe‘ und ‚Palliativmedizin‘ Gegenstand von semantischen Kämpfen (vgl. E. Felder 2006)?
- Welche Diskursakteure bzw. Akteursgruppen spielen eine leitende Rolle und wie spiegeln sich ihre Interessen und Wissensrahmen (kognitive Konzepte) in der sprachlichen Zubereitung der ‚Welt der ‚Sterbehilfe‘, wieder? Und schließlich
- mit welchen sprachlichen Mitteln werden Geltungsansprüche (auf allen Ebenen des Sprachsystems) in der Auseinandersetzung um eine rechtliche Regulierung von ‚Sterbehilfe‘ und ‚Palliativmedizin‘ im institutionalisierten Kontext juristischer Textarbeit konstituiert?

Unter Zuhilfenahme des LDA-Toolkit können diese Fragen nicht ‚automatisch‘ beantwortet, gar ‚errechnet‘ werden. Insbesondere im Kontext juristischer Textarbeit (auf die an dieser Stelle nicht näher eingegangen werden kann; vgl. hierzu grundlegend D. Busse 1992a) bedarf es zur Einordnung von Rechtsausdrücken eines differenzierten Hintergrundwissens über die Abläufe in der Institution Recht, zur juristischen Theorie und Methodik sowie nicht zuletzt zur Sonderstellung der juristischen Fachsprache und -kommunikation. Ein sorgfältiger Umgang mit den Al-

gorithmen des LDA-Toolkit ermöglicht jedoch eine corpus driven-orientierte Vorstrukturierung des komplexen Untersuchungsmaterials und damit sowohl die Generierung von (neuen) Hypothesen als auch eine datengeleitete Kontrastierung (und Kontrolle) introspektiver Annahmen. Dies soll im Folgenden an einigen Beispielen illustriert werden.

Der Untersuchung zugrunde liegen zwei juristische Fachtextkorpora vergleichbarer Größe, die manuell zusammengestellt und für kontrastive Analysen herangezogen wurden. Die Korpora wurden mit Hilfe des TreeTaggers lemmatisiert und wortartannotiert.

- Primärkorpus: Juristische Fachtexte zur Debatte um ‚Sterbehilfe‘ und ‚Palliativmedizin‘: 211 Fachtexte (0,9 Mio Wortformen), bestehend aus Gerichtsentscheidungen verschiedener Instanzen (BVerfG, BGH, OLG, LG), Aufsätzen aus einschlägigen juristischen Fachzeitschriften sowie Kurzmeldungen (Beck-link) zu juristischen Debatten. Die Auswahl der Texte erfolgte mittels Datenbankabfrage zum Ausdruck *Sterbehilfe* (einschl. Komposita usw.), der Rechtstextdatenbanken von Beck- und Juris-Online sowie nach einer manuellen Sichtung der Texte auf inhaltliche Relevanz²³. Zeitlich wurde keine Einschränkung vorgenommen, die älteste (belegte) Entscheidung datiert auf das Jahr 1964, der Großteil der Texte ist neueren Datums.
- Referenzkorpus: Juristische Fachtexte zur Debatte um die sog. ‚Online-Durchsuchung‘: 108 Fachtexte (0,8 Mio Wortformen), bestehend aus Gerichtsentscheidungen der Oberinstanzgerichte (BVerfG und BGH), juristischen Gutachten sowie Aufsätzen aus einschlägigen Fachzeitschriften. Die Auswahl der Texte erfolgte ebenfalls mittels Datenbankabfrage zum Ausdruck *Online-Durchsuchung* ODER *BKAG*²⁴ (einschl. Komposita usw.), der Rechtstextdatenbanken von Beck- und Juris-Online sowie nach manueller Relevanzsichtung (s. o.). Zeitlich wurde keine Einschränkung vorgenommen; die Mehrheit der Texte erschien nach 2005²⁵.

(a) Um einen **Überblick zu zentral verhandelten Sachverhalten und Referenzobjekten** im ‚Sterbehilfe‘-Diskurs zu gewinnen, bietet sich zunächst eine Keywordanalyse an, deren Ergebnisbelege (hochsignifikante Keywords) nach einem abstrakten Kategorienmodell²⁶ (EREIGNISSE, AKTEURE, ORTE, GEGENSTÄNDE, HANDLUNGEN usw.)²⁷ sowie mittels Kotextsichtung thematisch gruppiert werden. Zur Bestimmung von Keywords im juristischen ‚Sterbehilfe‘-Diskurs wird als Referenzkorpus das o.g. Rechtstextkorpus zur Debatte um ‚Online-Durchsuchungen‘ herangezogen, um möglichst diskursunspezifische Ausdrücke mittels Signifikanzprüfung herauszufiltern.

Nach Auswertung der ersten 200 hochsignifikanten ($P \geq 95\%$) Keywords lassen sich etwa folgende Sachverhalte und Referenzobjekte sortieren:

Wiederkehrende **AKTEURE** im juristischen Fachdiskurs zur ‚Sterbehilfe‘ sind **BETROFFENE** (*Patienten, Frau, Kranke, Kind, Suizidenten*), **DRITTE** (*Vormundschaftsgericht, [gesetzlicher] Betreuer/Vertreter, Bevollmächtigte und Angehörige*), **MEDIZINER** (*Ärzte, Taupitz [Gutachter], Pflegepersonal*), **VEREINIGUNGEN** (z. B. *Ethikkommission der Bundesärztekammer, Hospiz Stiftung*) sowie **KONFESSIONEN** (*Kirche*). Ferner sind **AKTEURE: PROZESSBETEILIGTE** (*Angeklagte, Kläger, Zeuge*), dogmatisch einschlägige **GERICHTE** (v. a. **OBERLANDESGERICHTE** Frankfurt und München (*OLG*), das **BUNDESVERFASSUNGSGERICHT** (*BVerfG*) sowie vereinzelte **LANDGERICHTE** (*LG*) sowie Akteure aus **EXEKUTIVE** und **LEGISLATIVE** (Wolfgang Bosbach (CDU), der 2008 zum *Verfahrensrecht der Patientenverfügung* einen Gesetzesentwurf vorgelegt hat²⁸, und die *Enquete-Kommission* des Bundestags).

Eine besondere Rolle spielen **MEDIZINER** und ihr Verhältnis zu anderen Akteuren. So konkurriert das professionelle Selbstbild der Mediziner mit Rechtsgrundsätzen und Ansprüchen von **BETROFFENEN** und **DRITTEN** (*Berufspflicht, Gewissensentscheidung, Schweigepflicht*, u. v. a.); im Fokus steht auch das umstrittene medizinische **HANDELN** im konkreten Fall (*Behandlung, künstliche [Ernährung], Behandlungsabbruch, lebensverlängernde [Maßnahme], Indikation, Medikament, [diverse medizinische] Hilfe, Tun²⁹, Sonde, Beratung, Aufklärung*) sowie die medizinische Expertise vor Gericht (*Gutachten, [gilt es, den Erkenntnisstand der medizinischen] Wissenschaft [zur Zeit der Behandlung zu ermitteln]*).

RECHTSORDNUNG und WERTE im Rechtsdiskurs zur ‚Sterbehilfe‘: Zahlreiche Keywords lassen sich als Spuren von semantisch (interpretativ) **umstrittenen Normtexten und Normtextstellen** gruppieren. Zu unterscheiden sind i) ‚privatrechtliche‘, ii) ‚strafrechtliche‘ und iii) ‚grundrechtliche‘ Rechtsgebiete.

- i) Als Indices für Semantische Kämpfe im Rahmen des **Privatrechts** spielen das *BGB*³⁰ respektive das *Betreuungsrecht* um die Paragraphen § 1901a *BGB* (*Patientenverfügung* und deren *Verbindlichkeit*), § 1991c i. V. m. § 1896 Abs. 2 *BGB* (*Willensbekundung* in Form einer *Patientenverfügung* konkurriert mit jener der Form *Vorsorgevollmacht*), § 1901 *BGB* (*Rechtsstellung und Aufgaben* von gesetzlichen *Betreuern* bei der Pflege von **BETROFFENEN**) und § 1904 *BGB* (konkurrierende Rechtsstellung des *Betreuers* gegenüber einem *Betreuungsgericht*) eine wichtige Rolle.

- ii) Rund um die Keywords *StGB*, *straflos*, *Strafrecht*, *lebenslang*, *Freiheitsstrafe*, *strafbar*, *Tatherrschaft* und *Tötungsdelikt* sowie damit verbundene Normtextparagrafen des **Strafrechts** (§ 323c *StGB* – ‚Unterlassene Hilfeleistung‘ sowie §§ 211 ff. *StGB* – ‚Mord/Totschlag‘) zeichnen sich Ausdrucks- und Konzeptualisierungskonkurrenzen ab zu der Frage nach der Abgrenzung und Abgrenzbarkeit von ‚strafbaren versus nicht strafbaren Handlungen, die den Tod eines BETROFFENEN nach sich ziehen‘.
- iii) Der Rechtsdiskurs um ‚Sterbehilfe‘ kreist insbesondere auch um die ‚Wahrung bzw. Verletzung‘ von **Menschen- und Grundrechten**. Genauer geht es um divergierende Versuche der Bedeutungsfestschreibung einzelner Normtextausdrücke (*Menschenwürde*, *Leben*, *Selbstbestimmung* usw.) mit Bezug auf den jeweiligen Fall.

Neben Keywords, die auf die RECHTSORDNUNG im Besonderen rekurrieren, finden sich zahlreiche signifikante Ausdrücke, die **ETHISCHE WERTE** akzentuieren und als Topoi in der Auseinandersetzung prozedieren. Um die Deutungshoheit über verschiedene Sachverhalte konkurrieren im Wesentlichen ‚individuelle‘ (BETROFFENE und ANGEHÖRIGE), ‚medizinische‘ (Hippokratischer Eid), ‚religiöse‘ und ‚juristisch-menschenrechtliche‘ Wertekataloge. Ihre Funktion scheint zweierlei: Zum einen dienen diese (meist impliziten) Wertekataloge als Orientierungsmaßstab sowie Legitimation für das eigene Handeln in einem konkreten Fall. Zum anderen dienen sie als ‚Lückenfüller‘ zur Konkretisierung abstrakter Rechtssätze, wobei dieser Einfluss ETHISCHER WERTE auf die Rechtsdogmatik selbst umstritten bleibt. Deutlich wird der Topos-Charakter ETHISCHER WERTE bereits an folgenden signifikanten Keywords nebst ihrer attributiven Kollokationen:

- [*ethisch X*] mit X = *rechtlich, juristisch, Grund, Aspekt, Hinsicht, beeinflusst, Bedenken, Diskussion, Frage, Problem, Verantwortung*
- [*X Wertvorstellung*] mit X = *allgemeine, persönliche, religiöse*
- [*X Überzeugung*] mit X = *ärztliche*

Zentral für die Debatte um ‚Sterbehilfe‘ ist offenbar auch die sprachliche Konstitution der **LEBENSWELT und der FORMEN von ‚Sterbehilfe‘**. Ohne hierauf im Einzelnen eingehen zu können, so vermittelt die folgende Liste an Keywords bereits einen Eindruck zu der Divergenz der Perspektiven:

- [[*sicher / absehbar / ohnehin / alsbald*] sterben werden]: Topos der STERBEHILFE-Befürworter, nach dem ‚das Handeln für das OB des Sterbeprozesses keine Rolle spielt‘ und demnach der Handelnde sich nicht strafbar mache.

- [*Sterben zu lassen*]: Topos der STERBEHILFE-Kritiker mit implizitem Vorwurf der ‚unterlassenen Hilfeleistung (oder gar Mord)‘.

*alle drei beteiligten Ärzte diskutierten gegen 2.00 Uhr des 4. 6. 1987, ob ein Transport in ein Krankenhaus sinnvoll sei; sie gingen zu diesem Zeitpunkt aus, daß die Patientin **ohnehin sterben würde** [= Perspektive der Befürworter]. Zur Überzeugung der StrK [= Strafkammer, FV] beschließen sie sodann gemeinsam [...], Frau V nicht mehr in ein Krankenhaus zu bringen und sie mit einer schnell verabreichten Überdosis Dolantin zu **töten** [= Perspektive der Kritiker] (BGHSt 42, 301)*

- Weitere signifikante Formen divergierender (Lebenswelt- wie Rechts-) Perspektiven: *Hilfe beim Sterben* versus *Hilfe zum Sterben*, *Behandlungsabbruch*, *in Würde sterben*, *aktive Sterbehilfe* versus *passive Sterbehilfe*, *direkte Sterbehilfe* versus *indirekte Sterbehilfe*, *Suizid*, *freiverantwortlicher Suizid*, *Suizidbeihilfe*, *assistierter Suizid*, *töten*, *Tötung*, *Selbsttötung*, *Selbstmord*, *Tötung auf Verlangen*, *Beihilfe zum Mord*, *Fremdtötung*, *Euthanasie* usw.

Während es bei den o. g. Keywords mehr um die Handlungskontexte der ‚Sterbehilfe‘ geht, finden sich zahlreiche Keywords zur Konkretisierung der **ZUSTÄNDE des BETROFFENEN**. Von dessen sprachlicher Konstitution hängt zentral die Legitimation der verschiedenen **HANDLUNGEN** ab:

- *Leben / Sterben / Tod*: Wo ist die Grenze anzusetzen zwischen einem ‚guten‘ und ‚leidenden‘ Leben? Ist *Leben* nur auf die ‚physische‘ oder auch auf die ‚psychisch-kognitiv-emotionale‘ Qualitätsebene hin zu bewerten? Was ist ein ‚natürlicher‘ oder ‚unnatürlicher‘ Sterbeprozess? Diese Fragen hängen eng zusammen auch mit folgenden Keywords:

- *Leiden*, *Krankheit*, *irreversibel*, *Verlauf* [tendenziell: *tödlich*], *Schmerz*, *Erkrankung*, *unheilbar*, *Eintritt* (eines best. Ereignisses), *Entscheidungsunfähigkeit*, *Demenz*, *Bewusstlosigkeit*, *Koma*, *Einwilligungs[un]fähigkeit*, *Hirntod* u. v. a.: Grundsätzliche Auseinandersetzung um die ‚physischen‘ versus ‚psychisch-kognitiven‘ Fähigkeiten sowohl im Hinblick auf die ‚gesundheitliche Verlaufsprognose‘ als auch auf die ‚Voraussetzungen für eine Willensäußerung‘.

Neben diesen lassen sich weitere Sachverhalte und Referenzobjekte als Zentren der semantischen Kämpfe im ‚Sterbehilfe‘-Diskurs skizzieren (insbesondere auch ORTE des STERBENS, die PATIENTENVERFÜGUNG im Besonderen usw.). An dieser Stelle soll jedoch nur deutlich werden, wie durch eine Systematisierung kontrastiv ermittelter Keywords die Daten effektiv strukturiert und ein induktiver Überblick zu globalen Diskurssedimenten (‚in‘ Korpora) gewonnen werden kann.

(b) Eine andere Möglichkeit, diskursive Sedimente kontrastiv zu ermitteln, besteht in der Berechnung von Key-N-Grammen allein unter Berücksichtigung von Autosemantika (insb. Substantive, Eigennamen, Vollverben) bzw. nach vorheriger Tilgung aller Synsemantika im Korpus. Die daraus resultierenden, signifikanten N-Gramme lassen sich interpretieren als wiederkehrende thematische Muster in (hier: Rechts-) Texten. Im Folgenden werden exemplarisch hochsignifikante³¹ Trigramme aus lemmatisierten Autosemantika und damit thematische Rekurrenz in der juristischen Debatte um die ‚Sterbehilfe‘ – analog zur Auswertung von Keywords (s. o.) – interpretativ gruppiert. Dabei zeigt sich auch, dass zahlreiche der oben bereits genannten Keywords erneut auftauchen, nun aber in thematischer Relation zueinander.

- Einschlägige Texte der Dogmatik (Gerichtsentscheidungen, rechtswissenschaftliche Literatur), auf die in der Debatte als Geltungsansprüche (sog. *herrschende Meinung*) häufig rekurriert wird:

Trigramm	χ^2	f / 10.000 Primärkorpus	f / 10.000 Referenzkorpus
<i>bgbst jahr seite</i>	655,527	24,959	0,000
<i>njw jahr seite</i>	520,277	74,205	30,158
<i>famrz jahr seite</i>	471,026	17,941	0,000
<i>zrp jahr seite</i>	258,053	12,723	1,144
<i>bgb njw njw</i>	206,462	11,381	1,449
<i>btprax jahr seite</i>	168,116	6,407	0,000
<i>mdr jahr seite</i>	159,307	6,072	0,000
<i>arztr arztrecht jahr</i>	48,021	1,831	0,000
<i>olg frankfurt a.m.</i>	47,330	2,105	0,114
<i>fußnote olg köln</i>	39,216	1,495	0,000
<i>olg münchen njw</i>	34,414	1,312	0,000

- Umstrittene Normtext(teile) – sowohl bestehender (de lege lata) als auch geplanter (de lege ferenda) Rechtslage:

<i>art gg artikel</i>	97,380	9,398	2,821
<i>gg artikel absatz</i>	90,418	7,475	1,906
<i>1901a absatz bgb</i>	55,225	2,105	0,000
<i>gesetz änderung betreuungsrecht</i>	30,412	1,159	0,000
<i>de lege lata</i>	27,211	1,037	0,000
<i>gesetz überprüfung lebensbeendigung</i>	24,009	0,915	0,000
<i>bgb 1901b bgb</i>	22,409	0,854	0,000
<i>ausdruck geltung verschaffen</i>	15,206	0,580	0,000
<i>entwurf gesetz änderung</i>	13,605	0,519	0,000
<i>stgb tötung verlangen</i>	13,605	0,519	0,000
<i>verbindlichkeit grenze patientenverfügung</i>	15,206	0,580	0,000

- Die zentrale Auseinandersetzung um die Konkretisierung des WILLENS, des gesundheitlichen ZUSTANDS des BETROFFENEN sowie die Rolle von DRITTEN bei der ‚Bestimmung‘ dieser beiden Sachverhalte:

<i>wille patient entsprechen</i>	28,811	1,098	0,000
<i>einwilligung betreuer behandlungsabbruch</i>	21,608	0,824	0,000
<i>wille betroffene entsprechen</i>	20,808	0,793	0,000
<i>patientenautonomie ende leben</i>	19,207	0,732	0,000
<i>tod zeit eintreten</i>	18,407	0,702	0,000
<i>einwilligung betreuer behandlung</i>	16,806	0,641	0,000
<i>verbindlichkeit grenze patientenverfügung</i>	15,206	0,580	0,000
<i>peg sonde ernähren</i>	13,605	0,519	0,000
<i>arzt patienten verhältnis</i>	13,605	0,519	0,000
<i>gesundheitszustand heilbehandlung eingriff</i>	13,605	0,519	0,000
<i>einwilligung betreuer maßnahme</i>	13,605	0,519	0,000
<i>maßnahme wille patient</i>	12,805	0,488	0,000
<i>verbindlichkeit grenze patientenverfügung</i>	15,206	0,580	0,000

- Die Rolle von Wertungsmaßstäben und damit verbundene Akteure in der Debatte:

<i>ethik recht medizin</i>	21,608	0,824	0,000
<i>grundsatz bundesärztekammer sterbebegleitung</i>	20,808	0,793	0,000
<i>arbeitsgruppe patientenautonomie lebensende</i>	16,806	0,641	0,000
<i>kommission ethik recht</i>	16,806	0,641	0,000
<i>bioethik kommission land</i>	13,605	0,519	0,000
<i>kommission land rheinland</i>	13,605	0,519	0,000
<i>überzeugung verlauf annehmen</i>	12,805	0,488	0,000
<i>zurückgreifen wertvorstellung entsprechen</i>	12,805	0,488	0,000
<i>lebensentscheidung wertvorstellung überzeugung</i>	12,805	0,488	0,000
<i>überzeugung wertvorstellung lebenserwartung</i>	12,805	0,488	0,000
<i>kriterium zurückgreifen wertvorstellung</i>	12,805	0,488	0,000

(c) Kontrastive N-Gramm-Analysen geben auch Hinweise auf wiederkehrende argumentative Routinen bzw. Topoi. Von den für juristische Fachdiskurse ebenfalls aufschlussreichen Textverweiselementen abgesehen³² zeigt eine Auswertung hochsignifikanter 4-Gramme (ohne jegliche Filterung) zum Beispiel folgende vier Topoi:

- Zentraler Topos ‚eine Handlung X muss der ‚vermuteten‘ INTENTION des BETROFFENEN gerecht werden‘: Damit ‚legale‘ Formen der Sterbehilfe‘ auch legitim erfolgen können, haben sich die Handlungen, die insb. MEDIZINER oder DRITTE vornehmen (möchten), an die INTENTION des BETROFFENEN zu halten. Juristen respektive Richter stehen dann vor der Frage, wie eine INTENTION zu ‚ermitteln‘ sei. Die Mehrworteinheit *wirklicher oder mutmaßlicher Wille* stellt sich diesbezüglich als juristische usuelle Wortverbindung heraus, mit der genau auf dieses Problem abgehoben wird. Mit anderen Worten: Im Fokus der Argumentation steht in der Regel nicht die (häufig fehlende oder umstrittene) ‚Willens-Selbstäußerung‘ des BETROFFENEN, sondern häufig ‚Art, Form, Zeit usw. der ‚externen‘ Willens-Zuschreibung‘ von Seiten DRITTER.

Wille d Betroffene entsprechen; wirklich oder mutmaßlich Wille; d mutmaßlich Wille d; mutmaßlich Wille d Patient; d Wille d Patient; mutmaßlich Wille d Betroffene; d Selbstbestimmungsrecht d Patient; d Einwilligung d Betreuer; d Widerruf d Einwilligung; Ermittlung d mutmaßlich Wille; ein Einwilligung d Betreuer; Wille d Patient entsprechen; d erklärt oder mutmaßlich; d Selbstbestimmungsrecht d Betroffene u. v. a.

- TOPOS ‚Sterbehilfe‘ bezogene Handlungen sind nur dann/auch dann nicht zulässig, wenn der ‚Prozess des Sterbens‘ bereits eingesetzt hat‘: Diese Argumentation ist Teil einer ganzen Topoi-Gruppe, in der die ‚Strafbarkeit‘ bzw. ‚Straflosigkeit‘ des Handelns von MEDIZINERN oder DRITTEN von der sprachlichen Konstitution des ZUSTANDS des BETROFFENEN abhängig gemacht wird.

irreversibel tödlich verlauf annehmen; sterbevorgang noch nicht einsetzen; ein tödlich verlauf annehmen; tödlich verlauf nehmen haben u. a.

*Die vom XII. Zivilsenat angenommene Begrenzung – keine Geltung des Verbots lebenserhaltender Maßnahmen, bevor die Grunderkrankung einen **irreversiblen tödlichen Verlauf angenommen** hat – ist mit dem auch durch Art. 1 und Art. 2 GG und Art. 8 EMRK garantierten Selbstbestimmungsrecht eines Patienten nicht vereinbar (Kutzer 2004: Probleme der Sterbehilfe – Entwicklung und Stand der Diskussion)*

- Topos ‚DRITTE (vor allem MEDIZINER und RICHTER) dürfen nicht ‚über den Kopf der BETROFFENEN hinweg‘ über derart existentielle Fragen wie die ‚Sterbehilfe‘ entscheiden‘: Dieser Topos – aufgerufen über die signifikante Mehrworteinheit *[[nicht] Richter] über Leben und Tod* (häufig in Distanzmarkern) – wird von verschiedenen Akteuren jeweils zur Delegitimierung einer anderen Akteursgruppe bzw. deren Argumentation im Diskurs erhoben: Juristen gegenüber der Autorität medizinischer Handlungen, Angehörige gegenüber staatlicher (exekutiver und judikativer) Akteure usw.

Auch wenn die Vormundschaftsrichter(innen) nicht „Richter über Leben und Tod“ sind und werden sollen: Eine konsistente und praxistaugliche Regelung des Sterbehilferechts, die den verfassungsrechtlichen Postulaten nach Patientenautonomie und Integritätsschutz gleichermaßen gerecht wird, setzt voraus, dass der Gesetzgeber verstärkt auch rechtstatsächliche Daten und Erkenntnisse berücksichtigt. (Schäfer 2005:

Regelung der Patientenverfügungen aus Sicht der Rechtspraxis)

- Topos ‚Eine ‚gesetzliche Lücke‘ kann oder soll (nicht) über Richterrecht geschlossen werden‘: Signifikant häufig tritt im ‚Sterbehilfe‘-Korpus das 4-Gramm *im Wege der Rechtsfortbildung* auf. Die Mehrworteinheit ist Hinweis auf eine noch sehr junge bzw. im Wandel begriffene Rechtspraxis im Bereich ‚Sterbehilfe‘, denn sie expliziert methodische Auseinandersetzungen um eine ‚normtextbezogene‘ (‚im Sinne des Gesetzes‘) versus ‚mangels gesetzlicher Regelung auf Richterrecht bezogene‘ Konstitution juristischer Normen. Mit anderen Worten: die Mehrworteinheit dient zur Legitimation einer ‚vom (expliziten) Gesetzestext entfernten‘ Sachverhalts- und Bedeutungskonstitution durch den Richter³³.

*Richtig ist auch, dass der Gesetzgeber ein Verhalten des Betreuers, das auf Durchsetzung eines solchen Wunsches gerichtet ist, keinem Genehmigungsvorbehalt unterworfen hat. Daraus lässt sich jedoch nicht auf ein „beredtes Schweigen“ des Gesetzes schließen, das es verbieten könnte, **im Wege der Rechtsfortbildung** die unterlassene Einwilligung des Betreuers in lebensverlängernde oder -erhaltende Maßnahmen einer vormundschaftsgerichtlichen Kontrolle zu unterwerfen.*

(BGH XII ZB 2/03 vom 17. 03. 2003)

(e) Um die **rechtssprachlichen Besonderheiten auch auf grammatischer Ebene** herauszuarbeiten, kann schließlich auch ein Vergleich von POS-Grammen, also rekurrenten Mustern aus Part-of-Speech-Tags aufschlussreich sein. Zu diesem Zweck werden im Folgenden mit Hilfe des LDA-Toolkit signifikante Key-N-POS-Gramme (der Größe 1 sowie der

Größe 4) ermittelt und aufgelöst. Als Referenzkorpus dient ein Korpus aus 4.723 Medientexten (2,7 Mio Wortformen) aus überregionalen Zeitschriften und Zeitungen zum Thema ‚Sterbehilfe‘ und ‚Palliativmedizin‘.

Key-N-POS-Gramme der Größe 1 zeigen die Verteilungen der einzelnen Wortart-Vorkommen im Rechtstext- versus Medientextkorpus zur ‚Sterbehilfe‘.

POS	χ^2	f / 10.000 Rechtskorpus	f / 10.000 Medienkorpus
CARD	59781,666	712,667	183,009
XY	7001,851	86,031	22,147
APPRART	5890,426	312,601	175,130
NN	5735,577	2918,141	2500,394
ADJA	1480,139	764,597	643,462
APPO	338,353	5,639	1,828
VVIZU	252,210	23,565	15,285
PDAT	92,843	50,399	42,356
ART	65,488	1145,607	1113,491
FM	45,512	9,540	7,182
PTKNEG	40,118	96,155	88,615
KOUS	20,351	140,817	134,250

Die relative Häufigkeit der Tags CARD und XY ist wenig überraschend, da sich im Rechtstextkorpus erwartungsgemäß mehr Zahlen (Normtextparagrafen) sowie das Paragraphenzeichen (XY hier: §) belegt sind als in einem nicht-juristischen Korpus³⁴. Die Mehrzahl an Präpositionen (APPRART, APPO) im Rechtstextkorpus erklärt sich vor allem in Textverweisen, d. h. in der intertextuellen Verweisungskultur juristischer Argumentation (*zur Fußnote, im Beitrag, nach Artikel* usw.). Substantive (NN), Adjektive (ADJA) und Artikelgebrauch (ART) deuten auf vermehrten Nominalstil in der juristischen Fachsprache hin (siehe dazu auch im Folgenden). Die Signifikanz des Tags VVIZU bestätigt frühere grammatische Studien (F. Vogel 2012: 423 ff.) in der Hypothese, über modale Infinitive in der Rechtssprache häufiger als in der Mediensprache Geltungsanspruch zu realisieren (*bei Entscheidungsfähigkeit aufzuzeigen, weitere Behandlungen abzulehnen, Aussagekraft Willensbekundungen beizumessen, darin sicher zuzustimmen* u. a.). Hinter PDAT verbergen sich attribuerende Demonstrativpronomen (*dieser Beschluss, diese Entscheidung*), die in ihrer deiktischen Funktion ein wichtiges Element der intra- und intertextuellen Verweisungen darstellen. Die signifikante Häufung von Negationspartikeln (PTKNEG) verweist auf die argumentative Textstruktur juristischer (v. a. auch richterlicher) Auseinandersetzungen, in der Fremdmeinungen in die eigenen Darlegungen integriert und (zu begründend) verworfen werden. Auf den Sprachhandlungstyp des Argu-

mentierens hin verweist schließlich auch die wiederkehrende Satzverknüpfung mittels unterordnender Konjunktionen (KOUS).

Die nachfolgenden Key-N-POS-Gramme der Größe 4 zeigen für die Rechtssprache typische grammatische Muster, das heißt vor allem komplexer, hypotaktischer Satzbau und Dominanz an (genitiv-)attributreichen Nominal- und Präpositionalphrasen (im Vergleich zur Mediensprache)³⁵:

POS-Gramm	χ^2	f / 10.000 Rechtsk.	f / 10.000 Medienk.	Exemplarische Belege
ART NN ART NN	1269,535	132,886	88,506	<i>das Bundesministerium der Justiz; die Handreichungen des Ministeriums; eine Einwilligung des Bevollmächtigten</i>
NN APPRART ADJA ADJA	1130,956	5,741	0,362	<i>Äußerungen zur eigenen bisherigen; Rat zur freiwilligen regelmäßigen; Augenblick am ernstlich unbedingten</i>
ADJA NN ART NN	814,172	70,639	45,047	<i>rechtzeitige Errichtung einer Vorsorgevollmacht; allgemeinen Beginn des Sterbeprozesses</i>
KOUS ART NN ART	662,321	13,245	4,813	<i>wenn der Fall der; daß der Einwilligung des; ob die Grundrechte des</i>
NN ART NN APPRART	644,311	17,429	7,444	<i>Verwendung der Vollmacht im; Begründung des Beschlusses zum; Grundsätze der Bundesärztekammer zur</i>
ADJA NN APPRART NN	634,265	29,615	15,800	<i>fertige Patientenverfügung zum Ankreuzen; beigefügte Erläuterungen zur Patientenverfügung; ergangenen Gerichtsentscheidungen zur frage</i>
ADJA ADJA NN ADJD	468,677	4,245	0,797	<i>gewandelten allgemeinen Wertvorstellung dahingehend; lebensgefährliche ärztliche Eingriffe erforderlich</i>
NN ART NN PTKNEG	456,118	6,543	1,901	<i>Informationsbedürfnis der Bürger nicht; Willen des Patienten nicht</i>
ART ADJA NN ART	379,983	46,983	32,335	<i>die rechtzeitige Errichtung einer; einem vorzeitigen Missbrauch der; einer analogen Heranziehung des</i>

(f) Berücksichtigt man neben diesen weitere semiautomatische Auswertungsroutinen (wie kontrastive Kookkurrenzanalysen zu Ausdrücken unterschiedlicher FORMEN der STERBEHILFE oder zu grammatischen Markern für konzeptuelle Divergenz wie Adversativ- und Konzessivkonnektoren), lassen sich folgende **Sedimente zentraler semantischer Kämpfe im ‚Sterbehilfe‘-Diskurs** nachzeichnen:

- Divergenz verfassungsrechtlicher Grundsätze in ihrer Konkretisierung auf die Lebenswelt (Fälle) der Sterbehilfe: ‚Patientenautonomie‘ versus ‚Lebensschutz‘;
- divergente Ko(n)textualisierung des ‚(Rechts-)Verhältnisses von ‚Betroffenem (-Willen)‘ und ‚Betreuer / Vormundschaftsgericht (-Entscheidung)‘ im Falle einer Sterbehilfe‘;
- Kollision ‚unterschiedlicher Wertmaßstäbe‘ (insbesondere Kirche versus (Verfassungs-)Recht versus individuelle Ethik von Betroffenen versus allgemeine Ethik einer Kultur);

- Kollision ‚unterschiedlicher Rollenverständnisse‘ (Arzt versus Jurist versus Betroffene versus Dritte); so soll etwa *die Verbindlichkeit der Patientenverfügung gegen den ärztlichen Paternalismus* [= sign. Keyword, FV] *befestig[t]* werden³⁶;
- Kollision der Beurteilungsmaßstäbe zur (medizinischen) Bewertung eines Sterbehilfe (nicht) legitimierenden Gesundheits- bzw. Krankheits- bzw. ‚Sterbe‘-Zustands (Kontextvariable); insb.: ‚biologisches‘ versus ‚kognitives‘ Leben‘;
- Divergenz von ‚(juristischer) Theorie‘ versus ‚(medizinischer oder Alltags-) Praxis‘;
- Besonderheit der deutschen Rechtsgeschichte: Verhältnis von ‚Sterbehilfe (heute)‘ und ‚(historischer) Euthanasie‘.

3.2 *Linguistische Imageanalyse am Beispiel der PIRATENPARTEI in deutschsprachigen Medien*

Der Ausdruck „Medienimage“ meint im Kontext einer linguistischen Imageanalyse (vgl. dazu F. Vogel 2010a, 2010b) weder kognitive „mentale Bilder“ (vgl. G. H. Mead ¹³1965: 332), noch unterstellt er ein kohärentes Bild an Sachverhaltsbeschreibungen einzelner Medientexte oder gar pauschal ‚der‘ Medien. Mit „Medienimages“ bezeichne ich vielmehr ein heuristisches Konstrukt, das sowohl der Kontingenz und Diversifizität, als auch temporal gehäufte Gleichförmigkeit medialer Sachverhaltszuschreibungen Rechnung zu tragen sucht. Das heißt, mediale Prädikationen unterschiedlichster Variationen (ganze Nominalsätze, Konnexionen, adjektivische oder adverbiale Attribuierungen, Komposita usw., bis hin zu bloßer kotextueller Nähe verschiedener Ausdrücke) im Hinblick auf bestimmte Sachverhalte, Objekte oder Akteure (Gruppen und Gruppenmitglieder) können – ohne dass es von konkreten Produzenten intendiert wäre – schlichtweg gehäuft auftreten. Diese Häufungen oder rekurrenten Sprachmuster sind Teil unseres allgegenwärtigen ‚medialen Hintergrundrauschens‘, das fortwährend bewusst zu verarbeiten wir schlichtweg überfordert sind. Gleichwohl bietet dieses Hintergrundrauschen fortwährender Medienpräsenz allein durch seine Informationsdichte prototypische Schemata³⁷ an. Dieses Wissen koorientierende mediale Angebot nenne ich Public-Images (PI). Öffentliche Images sind

„Teil global-diskursiver Ereignisse, die sich in Form typischer, d. h. wiederkehrender, ko(n)textsensitiver Sprachmuster manifestieren und damit ausdrucksseitige Vorlagen für spezifische Standardwerte in Frames bzw. Stereotypen bereitstellen.“ (F. Vogel 2010a: 350)

Eine „linguistische Imageanalyse“ (LIma) versucht in Anlehnung an diskurs- und korpuslinguistische Ansätze diesen Zusammenhang in einem Set aus qualitativen und quantitativen Methoden effektiv zu operationalisieren und insistiert insbesondere auf folgende drei Prämissen:

„Erstens abstrahiert sie weitestgehend von konkreten Akteuren, insb. von einzelnen Textproduzenten oder -rezipienten; zweitens wendet sie sich damit (heuristisch) zugleich ab von dem Paradigma der geschlossenen Texteinheit als Basis für Ganztextanalysen und fokussiert vielmehr auf großkorpusbasierte, textübergreifende rekurrente Sprachmuster als Indikatoren für mediale Zuschreibungsmuster³⁸, die schließlich drittens (vorsichtige) Rückschlüsse zulassen auf mögliches prototypisch verankertes Wissen in der Bevölkerung.“
(F. Vogel 2010b: 92)

Inwiefern der Funktionsumfang des LDA-Toolkit für eine Linguistische Imageanalyse fruchtbar gemacht werden kann, soll im Folgenden an einer exemplarischen Analyse zum Medienimage der PIRATENPARTEI illustriert werden. Die zentrale, hier nur andeutungsweise zu beantwortende Fragestellung lautet: Welche sprachlichen Prädikate werden der PIRATENPARTEI zugeschrieben, auf welchen sprachlichen Ebenen werden diese realisiert und inwiefern lassen die damit verbundenen Muster rückschließen auf potentielle Stereotype bei den Medienrezipienten?

Der Untersuchung zugrunde liegen zwei lemmatisierte und POS-annotierte Korpora:

- Primärkorpus: 3.931 Texte (2,61 Mio Token) aus überregionalen Zeitschriften und Zeitungen³⁹ aus dem Zeitraum vom 01. 01. 2011 bis einschl. 19. 05. 2012. Jeder Text sollte mindestens einmal die Zeichenfolge *piraten*, nicht jedoch die Ausdrücke *Somalia* oder *Afrika* enthalten (= Minimalhypothese).
- Referenzkorpus: 9.897 Texte (5,32 Mio Token) aus dem Zeitraum vom 01.01. bis einschl. 19. 05. 2012. Die Minimalhypothese bestand im mindestens einmaligen Vorkommen (je Text) der Zeichenkette *CDU* oder *Christdemokraten*.

Das LDA-Toolkit erlaubt die Durchführung kontrastiver Kookkurrenzanalysen, d. h. die kontrastive Kotext-Sedimentierung eines Ausdrucks gegenüber (dem Kotext) eines anderen Vergleichsausdrucks (vgl. o. 2.5). Auf die vorliegende Fragestellung angewandt lassen sich damit jene Kotextmuster ermitteln, die als potentielle Prädikationen typisch für die *Piratenpartei*, jedoch ‚untypisch‘ für die *CDU* wären. Im Folgenden wurde eine solche Kookkurrenzanalyse durchgeführt mit einem berück-

sichtigten Kotextintervall von acht Wörtern links bis acht Wörtern rechts des Suchausdrucks mit Berücksichtigung von Substantiven, Eigennamen und Adjektiven. Die sich ergebenden hochsignifikanten Kookkurrenzpartner wurden via weiterer Kookkurrenzanalysen, KWICs und Textdurchsicht zu thematischen Wortfeldern geclustert. Thematische Cluster der ersten 100 hochsignifikanten Substantive und Eigennamen sind:

- AKTEURE:
 - PERSONEN: *Sebastian Nerz* (Parteivorsitzender), *Marina Weisband* (Geschäftsführerin), *Bernd Schlömer* (Bundesvorsitzender), *Gerhard Anger* (Vorsitzender in Berlin), *Matthias Bahner* (umstrittenes *Ex-NPD-Mitglied*), *Jörg Tauss* (umstrittenes *Ex-SPD-Mitglied*), *Rick Falkvinge* (Piratenpartei-Gründer in Schweden).
 - GRUPPEN: *Piratenfraktion*, *Piratenwähler*, [*neue*] *Bewegung*, *Anhänger*, *Wikileaks*, *Musiker*, *Erstwähler*, *Internetaktivisten*, *Protestwähler*, *Newcomer*, *Nichtwähler*, *Hacker*, *Sozialpiraten*, *Netzaktivisten*, *Freibeuter*, *Schwarm*, *Datenschutzverfechter*, *Nerds*.
 - FUNKTIONSTRÄGER: *Geschäftsführerin*, *Mitglied* [der Piratenpartei], *Bundesschiedsrichterin*, *Bundesschiedsgericht*.
 - INSTITUTIONEN: *FORSA*, *Bundesschiedsgericht*.
- THEMEN und DOMÄNEN: *Politik* [v. a.: negativ: ‚gegen‘ *etablierte / klassische / bisherige Politik*; positiv: *transparente / offene / echte Politik*; *Protest*] *Internet*, *Urheberrecht*, *Künstler*, *Eigentum*, *Transparenz*, *Grundeinkommen*, *Netzpolitik*, *Bürgerrecht*, *liquid (democracy)*, *Demokratie*, *Kernthema*, *NPD* (Ausschluss bzw. Distanzierung), [*in die rechte*] *Ecke* [*stellen*], *Nahverkehr*, *filesharing*, *Geschlecht* (Piratenpartei lehnt Erfassung von Geschlechtsmerkmalen bei Behörden ab).
- ORTE bzw. LOKALISIERUNGEN: *Abgeordnetenhaus*, *Berliner*, *Landesparlament*, *Piratenpad* (Wikiplattform der Piratenpartei), *Saar*, *Parlament*, *Netz*, *Web*, [*in die rechte*] *Ecke* [*stellen*].
- OBJEKTE: *Spionagesoftware*, *Server*, *Inhalt* (der Piratenpartei sei (zu) gering), *Programm*, *IT*, *Latzhose*⁴⁰.
- EREIGNISSE bzw. ZUSTÄNDE: *Erfolg*, *Aufstieg*, *Wahlerfolg*, *Umfrage*, *Höhenflug*, *Umfragehoch*, *Triumph*, (*schärfere*) *Auseinandersetzung* (mit der Piratenpartei kündigen *Grüne / CSU* usw. an), *Tanzverbot* (in Hessen von der Piratenpartei kritisiert), *Abgeordnetenhauswahl*, *Parteitag*, *NPD*, *Wahlrend*, *Piratenparteitag*, *Überraschung*, *Tendenz*, *Bundesparteitag*, *Mitgliederboom*, *Zulauf* (an Mitgliedern).
- ABSTRAKTA: *Phänomen* (der Piratenpartei), *Fünfprozenthürde* (bei Landtagswahlen zu überwinden), *Klischee*.
- HANDLUNGEN: *Umgang* (mit X = brisantes Thema), (umstrittene) *Äußerung*, *Sprung / Einzug* (in Parlamente), *Entern* (als gebrauchssähnliche,

PIRATENPARTEI-bezogene Variante zu ‚einen (politischen, sozialen usw.) Ort erobern‘).

Thematische Cluster der ersten 100 hochsignifikanten Adjektive lauten:

- LOKATIV: *berliner, schwedisch, bayerisch, bundesweit, tunesisch, finnisch* [*Piraten(partei)*].
- TEMPORAL: *rasant* (Mitgliederzulauf, Wahlerfolge; „*wie die NSDAP*“), *vorübergehend* [-e *Erscheinung, -es Phänomen*], *erstmalig* (Einzug in einen Landtag), *anfänglich* [-er *Erfolg, noch sehr dürftige Piraten-Programm*], *zeitgemäß* [-e *Reformierung von Urheberrecht und hessischem Tanzverbot*].
- MODALITÄT: *potentiell* [-e *Wähler / Kandidaten / neue politische Kraft / [Koalitions-]Partner*], *augenscheinlich* [-er *Erfolg bei Wahlen*].
- FORM: *orangefarben, knallorange* (div. Utensilien der Selbstdarstellung).
- HABITUELL: *naiv, romantisch, modern, aufstrebend, charismatisch, unschuldig, hektisch, locker, typisch* (wenig soziale Verantwortung [tragen]), *vermieftes T-Shirt, Sexismus, Transparenz, Plagiatsoftware-Einsatz*, *ernst* ([*nehmen*]: eigene Ziele der Piratenpartei versus von anderen Parteien), *diszipliniert* [-er *Piraten-Parteitag*], *männlich, rechtsextrem / rechtsradikal* [-e *Äußerung / Vergangenheit, Gedankengut, Tendenz, Strömung, Einzelmeinung* [zurückgewiesen]], *antisemitisch* [-e *Äußerungen*], *professionell* (-e *Handeln* [beabsichtigen]), *anarchistisch, blöd* (im Hinblick auf Talkshow-Verhalten, *da stehen*), *irre* (v. a. ironisch und in Glossen⁴¹), *schillernd, juvenil*, (seltener:) *frauenfeindlich, frivol* [-er *Umgang mit dem Schutz des geistigen Eigentums*], *unerfahren, linksliberal, dynamisch, kleinlaut, nachdenklich, jugendlich, grundoptimistisch, hoffnungsvoll, braun, wild, albern, revolutionär* u. v. a.
- SOZIAL/ORGANISATORISCH: *basisdemokratisch, gegründet, hierarchiefrei, nichtkommerziell, verwaltend* [-e *Funktion des Parteivorstands*], *irrelevant* (*Meinung eines Einzelmitglieds gegenüber der Meinungsbildung der Basis*).
- WIRKUNG und FOLGEN: *sensationell, etabliert, gewählt, erfolgreich, beneidenswert, zweifellos* [*Farbe ins Geschehen bringen, ein Erfolg*], *unglaublich* (im Landtag), (seltener:) *attraktiv, fundamental* (-e *Veränderung der Parteienlandschaft*), *frisch* [-er *Wind, -es Gesicht, gewählt*], *antisemitisch* (-e *Hass Mails / Beschimpfungen / Anfeindungen* gegenüber der scheidenden Geschäftsführerin Weisband), *normal* [-e *Konkurrenten* aus Sicht anderer Parteien], *kinderpornografisch* (-e *Links auf Piratenpad* von Unbekannten eingestellt).
- THEMENBEZOGEN: *netzpolitisch, geistig* [-es *Eigentum*], *kulturpolitisch, kostenlos* [-er *Nahverkehr, -e Downloads* usw.], *netzaffiner, inhaltlich* (*durchaus etwas vorweisen, -e Leere, schwach, -e Ziele fehlen, -e Gründe* (für Wählerzuspruch fehle), *zu den Grünen keine Alternative anbieten, nichts mit Na-*

zis zu tun, -er kurs, -e Kernpunkte: Freiheit und Bürgerbeteiligung und Technikoffenheit, Auseinandersetzung [als CDU/Grüne/SPD] mit der [Piraten-] Partei), bedingungslos [-es Grundeinkommen], digital [-e Welt / Gesellschaft / Traum / Demokratie / Generation, -es Wissen, Zeitalter, -er Zeitgeist + (KKP:) Nerds, authentisch, Idee, Vertreter, jung, Politik, besser [verstehen] u. a.], transparent [-e Politik, Verwaltung, Staat, Meinungsbildungsprozess, Parteitag], frei [entscheiden, -er Zugang /-e Wähler], maximal / größtmöglich [-e Transparenz].

Bereits diese Liste an gruppierten, hochsignifikanten, rekurrenten Kontextmustern konturiert ein stereotypisiertes Medienimage der PIRATENPARTEI und lässt Hypothesen zu einer damit geprägten, generalisierenden Frame-Struktur in der Bevölkerung (bzw. bei Medienrezipienten) formulieren:

- (a) Der ‚typische‘ PIRAT⁴² ist tendenziell ‚männlich, jung, blass (*lichtscheu*), Computerspezialist oder ‚Hacker‘, unorganisiert, mit einem offenen (bzw. *naiven*), zuversichtlichen und etwas idealistischen Gemüt, experimentierfreudig bzw. verschroben oder komisch und verstößt gelegentlich (gerne) gegen ‚etablierte‘ Verhaltens- bzw. Erscheinungsregeln‘. IHN eint mit anderen Gleichgesinnten in der PIRATENPARTEI eine ‚Grundskepsis (bzw. Verdrossenheit) gegenüber ‚etablierter‘ Politikformen, eine informationstechnologische Sozialisierung sowie das Bedürfnis nach bzw. die Hoffnung auf ‚nachvollziehbare und basisgebundene‘ Organisation des gesellschaftlich-politischen Lebens mit Hilfe von Informationstechnik‘.
- (b) Wahrgenommen wird die PIRATENPARTEI insbesondere durch einzelne FUNKTIONSTRÄGER, die sich jedoch ‚selbst an das Votum ihrer Mitglieder halten möchten‘, sich also ‚als ein (kleines) Rädchen im großen Lager‘ verstehen. Allen voran (im Zeitintervall des Korpus) und ‚entgegen dem [damit wiederum bestätigten] Stereotyp‘ erhielt die Geschäftsführerin *Marina Weisband* große Beachtung in der (ver-)öffentlich(t)en Wahrnehmung. Die PIRATENPARTEI ist ‚kein regionales, unverständliches *Phänomen* (mehr), sondern Teil einer mittlerweile internationalen *Bewegung*‘.
- (c) Die PIRATENPARTEI ist sehr ‚erfolgreich und gewinnt zunehmend an (wiederum einst politikverdrossenen) Mitgliedern und Fürsprechern‘. Sie ist ‚anders‘ als die ‚etablierten‘ Parteien, bringt *neuen Wind* und ist insbesondere ‚für Jung-, Erst- und Protestwähler eine Alternative zur Nicht- oder Linkswahl‘. Mit ihrem *Triumph* ‚verschiebt die neue PIRATENPARTEI die bisherige Parteienlandschaft‘

und wird (daher) von politischen Konkurrenten *ernst genommen*, man ‚muss sich mit ihr *auseinandersetzen*‘.

- (d) Das POLITISCHE KERNTHEMA sowie die ‚Kompetenz‘ der PIRATENPARTEI ‚liegt‘ in erster Linie im Bereich ‚informationstechnologischer, datenkommunikativer und basisdemokratischer Organisation des gesellschaftlichen Lebens‘ sowie – mit großem Abstand – in ‚kulturpolitischen Fragestellungen‘. Im Kontext kulturpolitischer Fragen erscheint die PIRATENPARTEI als ‚Befürworter von („doch urheberrechtsgeschützten oder illegalen“) kostenlosen Inhalts- und Güterangeboten für alle Gesellschaftsmitglieder‘ (Auseinandersetzung mit Künstlern usw.). Die PIRATENPARTEI wird tendenziell als ‚inhaltsarm‘ wahrgenommen (dominantes Bindestrichkompositum: *Ein-Themen-Partei*) und von anderen Parteien zunehmend kritisiert.
- (e) Die PIRATENPARTEI ‚hat ein Problem mit antisemitischen und sexistischen Äußerungen bzw. Äußerungsereignissen und Inhalten‘ und erscheint – ‚wie insgesamt im politischen Alltag‘ – im Umgang damit ‚unerfahren‘: *Beneidenswert sind die Vertreter der Piratenpartei nicht: direkt vom Kindergarten in die Politik* (Welt, 27. 04. 2012). Die PARTEI ‚ist‘ ferner ‚sehr heterogen und lässt sich (noch nicht?) einer bestimmten ‚politischen Farbe‘ (rot/links versus schwarz/konservativ) zuordnen‘. Häufig wird sie mit ‚der Fraktion DIE GRÜNEN in ihrer Anfangszeit‘ assoziiert.
- (f) In konservativen Medien erscheint das Image der PIRATENPARTEI tendenziell negativer konnotiert als in linksliberalen Medien.

Die stereotypen Prädikationsketten im Medienimage PIRATENPARTEI werden stark von zwei Quellkonzepten genährt, nämlich aus Wissensrahmen zur ‚Märchenwelt der (See-) Piraten‘ einerseits sowie aus Konzepten der ‚räuberischen (Daten-) Piraterie der ‚realen‘ Welt‘ andererseits. Die Quellsemantik ‚märchenhafter Piraten‘ erscheint in tendenziell positiver konnotierter, metaphorischer sowie ironischer Perspektive (eher in linksliberalen Medien) und realisiert sich auf allen Ebenen des Sprachsystems. AKTEURE der PIRATENPARTEI *erobern*, *entern* oder *setzen Segel* (Verben); FUNKTIONSTRÄGER sind *Piratenkapitäne* oder *Piratenbräute* (Komposita). Die Quellsemantik ‚räuberischer Piraterie‘, also das ‚illegale Entwenden und Verbreiten von Gütern und vor allem Daten ohne Autorisierung und damit verbundene „Wertvernichtung“‘ erscheint tendenziell eher in konservativen Medien. Der Ausdruck *Pirat* wird dabei oft in den Kontext des ‚illegalen Filesharings‘ gestellt und die PIRATENPARTEI zu ‚Fürsprechern einer (implizit abzulehnenden) anarchistischen, d. h. das Privateigentum nicht wahrenen Subkultur‘:

geht es nach der **Piratenpartei**, dann soll das private Filesharing „bald völlig legal sein. Stärkung der Privatkopie“ nennen sie das. Was wird passieren, wenn alle meine Romane als E-book online frei verfügbar sind?

(Welt, 27. 04. 2012)

Doppelinterview [im Spiegel, FV] mit Jan Delay und einem Vertreter der **Piratenpartei**. Die Idee ist, einen Popstar und einen Netzpolitiker über die Zulässigkeit von Musikraubkopien diskutieren zu lassen

(Welt, 19. 04. 2012)

Für die hier formulierten Hypothesen zu Medienimage bzw. Stereotypen lassen sich auch auf einer Metaebene, nämlich in Prädikationen mit der expliziten Qualifizierung als *Vorurteil* oder *Klischee* (letzteres auch signifikant) bestätigen⁴³.

Piraten-Klischee: jung, männlich, mit Nerd-Ausstrahlung

(Spiegel, 12. 03. 2012)

Klischee, dass die Piraten keine Inhalte zu bieten hätten

(Spiegel, 07. 05. 2012)

beide passen nicht ins Bild der Klischee-Piraten. Sie haben wenig Ahnung von IT, ihre Facebook-Konten gelöscht und kämpfen mit den Tücken der Technik

(FR, 11. 04. 2012)

Klischee über die Piratenpartei als leicht verschrobener Internet-Clique

(FR, 20. 09. 2012)

Landesliste, auf der zwischen 14 Männern nur eine einzige Frau steht, zuständig für – Achtung Klischee! – Jugend

(taz, 17. 09. 2011)

ist auch nie jemand fern, der Augenklappe oder Freibeuter-Hut aufgesetzt hat. Viele Piraten wissen: Klischees stimmen zwar nicht unbedingt. Sie führen jedoch dazu, dass über die Piraten berichtet wird

(WaS, 29. 04. 2012)

*Wer in den noch jungen Piraten eine chaotische, unorganisierte oder ineffiziente Partei sieht, lebt von **Vorurteilen**. die Saar-Piraten jedenfalls agieren durchweg professionell und ausgesprochen pragmatisch*

(WaS, 11. 03. 2012)

*die acht vom Arbeitskreis Kommunalpolitik sind weit weg von all den **Vorurteilen**, die gegen die Piraten im Umlauf sind: Spaßpartei, Chaoten, Ein-Punkt-Partei*

(WaS, 25. 09. 2012)

4. Ausblick

Die LDA-Tools werden in ersten Lehrveranstaltungen und Qualifikationsarbeiten erprobt und kontinuierlich weiterentwickelt. Dabei gilt es, die quantitativ-automatischen Analyseinstrumente als Hilfsmittel zur qualitativen Diskurs- und Imageanalyse zu modellieren und das Methodenset praxisnah zu erweitern, ohne leichtfertig etwaigen ‚Statistik-Bias‘ zu unterliegen. Hierfür sollen für die Zukunft auch weitere Tools in die ‚digitale Arbeitsumgebung‘ des LDA-Toolkit eingebunden werden, wie insbesondere eine semiautomatische Erstellung von „Diskurskarten“ (vergleichende Kotextprofile)⁴⁴, die nicht lediglich zur Visualisierung von Ergebnissen, sondern vielmehr zur induktiven Analyse von Kotextstrukturen beitragen können.

In kommenden Folgeversionen werden schließlich auch englisch-, französisch- und spanischsprachige Korpora als Eingangsdaten verarbeitet sowie zum Zwecke eines komfortableren Korpusmanagements Schnittstellen zur IMS Corpus Workbench⁴⁵ geschaffen.

Anmerkungen

1. „LDA-Toolkit“ = Linguistisches Diskursanalyse-Toolkit.
2. Vgl. im Überblick etwa B. Kerchner (2006) für die Politikwissenschaft, R. Keller (2008: 21–96) für die Wissenssoziologie, P. Reuber & C. Pfaffenbach (2005: 198 ff.) für die (Human-) Geographie sowie H. Kudlich & R. Christensen (2009) exemplarisch für die Rechtswissenschaften (unter Ausblendung der vielfältigen juristischen Arbeiten zur Rhetorik im Recht).
3. M. Pêcheux (1969) entwickelte bereits in den siebziger Jahren einen theoretischen Rahmen für eine „Automatische Diskursanalyse“, die jedoch erst später auch computertechnisch fruchtbar gemacht werden konnte (vgl. N. Helsloot & T. Hak 2007).
4. Vgl. grundlegend D. Busse & W. Teubert (1994) zum Diskurs als „virtuelle Textkorpora“ sowie zuletzt W. Teubert (2006), W. Teubert et al (2007), W. Teubert & A. Cermáková (2007: 77 ff.) und W. Teubert (2012).
5. <http://www.semtracks.org/web/> (07. 05. 2012).
6. KWIC = Keywords in Context.
7. Cluster- und N-Gramm-Analyse zählen beide in einem bestimmten Wortintervall Mehrworteinheiten (MWE), mit (Cluster) und ohne (NGrams) Bezugswort, und geben diese nach Häufigkeit bzw. Signifikanz aus. Trigramme des vorangegangenen Satzes wären etwa: {[Cluster- und N-], [und N-Gramm], [N-Gramm-Analyse], [Gramm-Analyse zählen]} usw.
8. Vgl. hierzu etwa die Diskussion in W. Kallmeyer & G. Zifonun (Hrsg.) (2006).
9. Eine insbesondere für Einführungsseminare zur Korpuslinguistik geeignete, sehr übersichtliche und mittlerweile sehr ausführliche Plattform bietet der Online-Kurs von N. Bubenhofer (2006–2011).
10. Für wichtige Hinweise danke ich insbesondere Prof. Dr. Sebastian Padó (Institut für Computerlinguistik, Universität Heidelberg).
11. Derzeit steht eine 2.7 beta-Version zur Verfügung unter: <http://lda-toolkit.friedemann-vogel.de> (08. 05. 2012).

12. Diesen prägnanten Titel verdanke ich einer Rezension von N. Bubenhofer (2012) zu dem vorliegenden LDA-Toolkit.
13. Das KWIC-Tool erlaubt lediglich einfache Such-Abfragen, jedoch mit Berücksichtigung von Kotexteinschränkungen und der Möglichkeit, die KWICs als Token, Lemmata, POS-Tags oder daraus folgenden Kombinationen auszugeben (z. B. können POS-Gramme in Token[POS] aufgelöst werden). Für komplexere Suchanfragen und Sortiermodi stehen bereits zahlreiche einfach zu bedienende Freeware-Tools zur Verfügung (wie AntConc, ConcApp u. ä.).
14. Mittelfristig wird auch ein Signifikanztest mittels Log-Likelihood-Ratio (LLR) eingebunden werden.
15. <http://www.ims.uni-stuttgart.de/projekte/corplex/TreeTagger/> (01. 07. 2011).
16. <http://www.ims.uni-stuttgart.de/projekte/corplex/TreeTagger/> (01. 07. 2011).
17. Zugrunde lag ein Medientextkorpus zur Debatte um die sogenannte *Online-Durchsuchung*.
18. Es handelt sich um Fachtextkorpora vergleichbarer Größe zur juristischen Debatte zum einen um die Verrechtlichung von *Sterbehilfe* und *Patientenverfügung*, zum anderen um die Rechtmäßigkeit von *Online-Durchsuchungen*, d. h. den heimlichen Zugriff der Sicherheitsbehörden auf private Computer von Verdächtigen.
19. <http://www.wordle.net/> (05. 09. 2011).
20. Graph Visualization Software; <http://www.graphviz.org/> (05. 09. 2011).
21. <http://gephi.org/> (05. 09. 2011).
22. E. Felder, J. Luth & F. Vogel (in Vorb.).
23. Die Relevanzprüfung orientierte sich an den Kriterien des Umfangs (nicht nur einmalige Nennung des Ausdrucks im Rahmen eines thematisch fremden Falles, sondern auch Verhandlung des Themas in mehr als einem Absatz), der Quellenrelevanz (Einschlägigkeit der Zeitschriften im juristischen Fachdiskurs) sowie an Beratungen mit verschiedenen Fachjuristen.
24. *BKAG-E* = Entwurf zur Novellierung des BKA-Gesetzes einschl. einer darin vorgesehenen Regelung zur sog. *Online-Durchsuchung* (2006).
25. Dieses Korpus ist Teil bzw. Gegenstand einer umfassenden Studie zur Normgenese der ‚Online-Durchsuchung‘ unter besonderer Berücksichtigung von Judikative, Exekutive, Legislative, Medien und Rechts- bzw. IT-Wissenschaft (vgl. F. Vogel 2012).
26. Grundlage dieses allgemeinen Kategorienmodells sind die Arbeiten von K. P. Konerding (1993, 1997, 2008). Zur Operationalisierung Frame-basierter Diskursanalysemodelle zeigt Konerding in Anlehnung an die Frametheorie von L. Barsalou ein Verfahren, mit dem „sich Beschreibungsrahmen für allgemeinste semantisch-sortale Typen [von] Konzepte[n] konstruieren lassen, und zwar unter Rückgriff auf semantische Typisierungen von Substantiven und Verben, verbale Argumentationsstrukturen sowie Topik-Prädikations-Strukturen der deutschen Standardsprache“ (K. P. Konerding 2005: 15 f.). Dabei reduziert Konerding auf Basis von Wörterbuchlemmata systematisch Hyperonymtypen zu abstrakten Oberkategorien (Substantivklassen wie GEGENSTAND, ORGANISMUS, PERSON/AKTANT, EREIGNIS u. a.) und generiert auf der Basis von Thesaurus-Prädikationsschemata methodische Fragestellungen, die Aufschluss über Frame-Slot(-Value)-Strukturen in Texten geben können. Die von Konerding induktiv geronnenen Oberkategorien lassen sich als Orientierungsrahmen im Sinne eines allgemeinen Suchrasters für eine diskursspezifische Sortierung möglicher Sachverhalte und Referenzobjekte (z. B. MEDIZINISCHE vs. JURISTISCHE AKTEURE) fruchtbar machen (vgl. F. Vogel 2012: 41 f.).
27. Zum besseren Nachvollzug der hermeneutischen Arbeit werden die Ergebnisse in folgender Notation dokumentiert: Belege werden kursiv (Fokus auf Ausdrucks-

- ebene) gesetzt; Konzepte bzw. Attribute als Hypothesenkonzentrate der Interpretationsarbeit werden in spitze Klammern bzw. einfachen Anführungszeichen gesetzt („Intention des Betroffenen“: ‚selbstbestimmt‘ vs. ‚fremdbestimmt‘); Sachverhalte bzw. Referenzobjekte als induktiv-deduktiv gewonnene heuristische Analysekategorien werden in Majuskeln gesetzt (AKTEURE, EREIGNISSE usw.).
28. Vgl. BT-Drucksache 16/11360.
 29. Das Vollverb *Tun* ist sehr dominant bei der wertneutralen Hervorhebung ärztlichen Handelns: *Erlaubnis zur Beihilfe an einem Suizid durch positives Tun* [kann] ein strafbares Untätigbleiben folgen / verbotene Tötung auf Verlangen (durch **Tun**), wenn der Arzt die künstliche Beatmung auf Wunsch [abstellt].
 30. Bürgerliches Gesetzbuch.
 31. Exemplarisch ausgewertet werden hier allein hochsignifikante Trigramme mit $P \geq 99,95\%$ ($X^2 \geq 12,12$).
 32. In den bisherigen Beispielen zeigen sich wiederholt Verweise auf Normtexte sowie insbesondere Texte der juristischen Dogmatik. Diese Verweise sind Elemente eines umfassenden intertextuellen Geflechts, mittels dessen Juristen in der Diskussion Geltungsansprüche („Normativität“) zur Unterstützung ihrer Prädikationen realisieren. Ihnen genauer nachzugehen sprengte hier jedoch den Rahmen.
 33. Sprachwissenschaftlich betrachtet wird natürlich jede Form der juristischen Sachverhalts- und Bedeutungskonstitution erst durch den Richter als Textproduzent bzw. -rezipient manifest. „Die“ Norm gibt es nicht; Normen liegen nicht in Gesetztexten verborgen (traditionelle rechtstheoretische Fehlannahme, von D. Busse (1992b: 14) zutreffend als „Topftheorie“ der Kommunikation“ bezeichnet), sondern werden erst durch den deklarativen Sprechakt formelhaft (*im Namen des Volkes* ...) durch den Richter konstituiert und können jederzeit durch andere juristische Akteure neu ein- und umgeschrieben werden. Im juristischen Fachdiskurs werden jedoch unterschiedliche Methoden – und damit Geltungsanspruch konstituierende Textroutinen – differenziert, mittels derer eine Bedeutungszuschreibung legitimiert werden kann. Diesbezüglich konkurrieren bzw. ergänzen sich u. a. die vier juristischen Canones, nach denen eine Norm bestimmt werden kann: am „Wortlaut“ des Gesetzes (Grammatische Auslegung), durch intertextuelle Vergleiche von Normtextbezügen und Regelungen (Systematische Auslegung), durch Überlegungen zum „Sinn und Zweck“ einer Regelung (Teleologische Auslegung) sowie durch Berücksichtigung der Normtextgenese (Historische bzw. Genetische Auslegung). Vgl. dazu etwa H. Kudlich & R. Christensen (2004).
 34. Der Tag FM steht im Tagset des TreeTaggers für fremdsprachliche Ausdrücke; im Rechtstextkorpus zur ‚Sterbehilfe‘ steht es allerdings für nicht erkannte Normtext-Bezeichnungen.
 35. POS-Gramme, die Zahlen enthalten (v. a. Normtextverweise, andere Textverweise, Datumsangaben usw.), werden hierbei ausgeklammert.
 36. H. Holzhauser. 2004. Von Verfassungen wegen: Straffreiheit für passive Sterbehilfe (ZRP, 41).
 37. Prototypikalität orientiert sich dabei an der sprachlichen Präsentation von Sachverhalten: Diejenigen Informationen und Kontexte, die wiederholt versprachlicht werden (rekurrente Sprachmuster), prägen unsere Wissensrahmen, und umgekehrt: prototypisches Wissen zeigt sich häufig auch in sprachlichen Stereotypen (Redewendungen, Formelhafte Mehrworteinheiten usw.; vgl. H. Putnam ²1990, H. Feilke 1989, M. Heinemann (Hrsg.) 1998, K. P. Konerding 2001).
 38. Einzelne kommunikative Äußerungen (Einzeltexte) werden nur insofern berücksichtigt, als sie transtextuell rekurrent auftretende Sprachmuster zu interpretieren, also zu disambiguieren helfen.

39. Es handelt sich um die Frankfurter Rundschau, den Spiegel, Spiegel Online, Die Welt, Die Welt Online, Die Welt am Sonntag, Focus Magazin, Der Stern sowie die tageszeitung (taz).
40. *durch die Piratenpartei feiert die Latzhose ein Comeback in der Politik* (FR, 06. 10. 2012).
41. Z. B.: *Apocalypse now; Irre! Ehemalige Spitzenpiraten treiben unser Universum an den Rand des Untergangs* [= Titel] (taz, 20. 02. 2012); „*Wie irre ist denn der Pirat?*“ [= Interviewfrage] (Welt, 10. 04. 2012).
42. Häufig auch über die signifikante Kookkurrenz *Nerd* und wiederum deren Kotext koloriert: eine *Partei der Computer-Nerds, der Freaks* (Spiegel, 23. 04. 2012); *unrasierter, zottelhaarer Brillen-Nerd* (taz, 30. 07. 2011); *Nerdinchen. Das ist wohl ein weiblicher Nerd, ein Computerfreak in der Version eher nett* (mit Bezug auf eine Piratenpartei-Vorsitzende; FR, 04. 02. 2012) u. a.
43. Negationen widersprechen den Hypothesen nicht, im Gegenteil: sie präsupponieren vielmehr die Existenz von Pauschalurteilen (Stereotypen), indem sie diese negieren.
44. Vgl. hierzu die Ansätze zur induktiven Kotextexploration, online unter: <http://friedemann-vogel.de/projekte/diskurskartographie> (01.07.2011); eine Verbesserung der dort zu sehenden Ergebnisse verspricht die Anwendung vektorbasierter Ähnlichkeitsmaße (vgl. L. Lee 1999).
45. <http://www.ims.uni-stuttgart.de/projekte/CorpusWorkbench/> (24. 05. 2012).

Literatur

- Anthony, L. 2005. AntConc: design and development of a freeware corpus analysis toolkit for the technical writing classroom Professional Communication Conference. In: IPCC, 729–737.
- Baker, P. 2006. *Using Corpora in Discourse Analysis*. London: Continuum.
- Baker, P., Gabrielatos, C., Khosravini, M., Krzyzanowski, M., McNery, T. & Wodak, R. 2008. A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. In: *Discourse & Society*, 19/3, 273–306.
- Bubenhof, N. 2006–2011. Einführung in die Korpuslinguistik: Praktische Grundlagen und Werkzeuge. [Online-document] <http://www.bubenhof.com/korpuslinguistik/> (08. 05. 2012).
- Bubenhof, N. 2009. Sprachgebrauchsmuster. Korpuslinguistik als Methode der Diskurs- und Kulturanalyse. Berlin/New York: Walter de Gruyter. (= Sprache und Wissen 4).
- Bubenhof, N. 2012. LDA-Toolkit: Korpusanalyse zum Klicken statt Tippen. In Sprechtafel. Linguistische Notizen. [Online-document] <http://www.bubenhof.com/sprechtafel/2012/05/05/lda-toolkit-korpusanalyse-zum-klicken-statt-tippen/> (08. 05. 2012).
- Busse, D. & Teubert, W. 1994. Ist Diskurs ein sprachwissenschaftliches Objekt? Zur Methodenfrage der historischen Semantik. In: *Begriffsgeschichte und Diskursgeschichte*, 10–28. Hrsg. von D. Busse, F. Hermanns und W. Teubert. Darmstadt: Westdeutscher Verlag.
- Busse, D. 1992a. *Recht als Text. Linguistische Untersuchungen zur Arbeit mit Sprache in einer gesellschaftlichen Institution*. Tübingen: Max Niemeyer Verlag.
- Busse, D. 1992b. *Textinterpretation. Sprachtheoretische Grundlagen einer explikativen Semantik*. Opladen: Westdeutscher Verlag.
- Feilke, H. 1989. Funktionen verbaler Stereotype für die Alltagssprachliche Wissensorganisation. In: *Kognition und Kommunikation. Beiträge zur Psychologie der Zeichenverwendung*, 71–84. Hrsg. von C. Knobloch. Münster: Nodus.

- Felder, E. 2006. Semantische Kämpfe in Wissensdomänen. Eine Einführung in Benennungs-, Bedeutungs- und Sachverhaltsfixierungs-Konkurrenzen. In: *Semantische Kämpfe. Macht und Sprache in den Wissenschaften*, S. 13–46. Hrsg. von E. Felder. Berlin/New York: Walter de Gruyter.
- Felder, E. 2012. Pragma-semiotische Textarbeit und der hermeneutische Nutzen von Korpusanalysen für die linguistische Mediendiskursanalyse. In: *Korpuspragmatik. Thematische Korpora als Basis diskurslinguistischer Analysen*, 115–174. Hrsg. von E. Felder, M. Müller und F. Vogel. Berlin/New York: Walter de Gruyter. (= *Linguistik – Impulse und Tendenzen* 44).
- Felder, E., Luth, J. & Vogel, F. (in Vorb.). „Patientenautonomie“ und „Lebensschutz“. Der Rechtsdiskurs zur „Sterbehilfe“ aus linguistischer Perspektive. [Online-document] <http://www.gs.uni-heidelberg.de/sprache02/felder/> (24. 05. 2012).
- Felder, E., Müller, M. & Vogel, F. (Hrsg.). 2012. *Korpuspragmatik. Thematische Korpora als Basis diskurslinguistischer Analysen*. Berlin/New York: Walter de Gruyter. (= *Linguistik – Impulse und Tendenzen* 44).
- Foucault, M. 1973. *Archäologie des Wissens*. Frankfurt am Main: Suhrkamp Verlag.
- Foucault, M. 1974a. *Die Ordnung der Dinge: eine Archäologie der Humanwissenschaften*. Frankfurt am Main: Suhrkamp Verlag.
- Foucault, M. 1974b. *Die Ordnung der Diskurse*. Frankfurt am Main: Suhrkamp Verlag.
- Gardt, A. 2007. Diskursanalyse – Aktueller theoretischer Ort und methodische Möglichkeiten. In: *Diskurslinguistik nach Foucault. Theorie und Gegenstände*, 27–52. Hrsg. von I. Warnke. Berlin/New York: Walter de Gruyter.
- Heinemann, M. (Hg.). 1998. *Sprachliche und soziale Stereotype*. Frankfurt am Main: Peter Lang Verlag.
- Helsloot, N. & Hak, T. 2007. Pêcheux's Contribution to Discourse Analysis. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 8 (2), Art. 1. [Online-document] <http://nbn-resolving.de/urn:nbn:de:0114-fqs070218> (07. 05. 2012).
- Hermans, F. 1994. Schlüssel-, Schlag- und Fahnenwörter. Zu Begrifflichkeit und Theorie der lexikalischen „politischen Semantik“. Heidelberg: Univ. (Arbeiten aus dem Sonderforschungsbereich 245, „Sprache und Situation“, Heidelberg, Mannheim; 81).
- Kallmeyer, W. & Zifonun, G. 2006. *Sprachkorpora – Datenmengen und Erkenntnisfortschritt*. Berlin/New York: Walter de Gruyter.
- Keller, R. 2008. *Wissenssoziologische Diskursanalyse. Grundlegung eines Forschungsprogramms*. Tübingen: Verlag für Sozialwissenschaften.
- Kerchner, B. 2006. *Diskursanalyse in der Politikwissenschaft. Ein Forschungsüberblick*. In: *Foucault: Diskursanalyse der Politik. Eine Einführung*, 33–67. Hrsg. von B. Kerchner und S. Schneider. Tübingen: Verlag für Sozialwissenschaften.
- Konerding, K. P. 1993. *Frames und lexikalisches Bedeutungswissen: Untersuchungen zur linguistischen Grundlegung einer Frametheorie und zu ihrer Anwendung in der Lexikographie*. Tübingen: Max Niemeyer Verlag.
- Konerding, K. P. 1997. *Grundlagen einer linguistischen Schematheorie und ihr Einsatz in der Semantik*. In: *Methodologische Aspekte der Semantikforschung*, 57–84. Hrsg. von I. Pohl. Frankfurt am Main: Peter Lang Verlag.
- Konerding, K. P. 2001. *Sprache im Alltag und kognitive Linguistik: Stereotype und schematisiertes Wissen*. In: *Sprache im Alltag. Beiträge zu neuen Perspektiven in der Linguistik. Herbert Wiegand zum 65. Geburtstag gewidmet*, 151–172. Hrsg. von A. Lehr, M. Kammerer, K. P. Konerding, A. Storrer, C. Thimm und W. Wolski. Berlin/New York: Walter de Gruyter.

- Konerding, K. P. 2005. Diskurse, Themen und soziale Topik. In: *Mediendiskurse. Bestandsaufnahme und Perspektiven*, 9–38. Hrsg. von Claudia Fraas und Michael Klemm. Frankfurt am Main: Peter Lang Verlag.
- Konerding, K. P. 2008. Diskurse, Topik, Deutungsmuster – Zur Komplementarität, Konvergenz und Explikation sprach-, kultur- und sozialwissenschaftlicher Zugänge zur Diskursanalyse auf der Grundlage kollektiven Wissens. In: *Methoden der Diskurslinguistik. Sprachwissenschaftliche Zugänge zur transtextuellen Ebene*, 117–150. Hrsg. von I. H. Warnke und J. Spitzmüller. Berlin/New York: Walter de Gruyter.
- Kudlich, H. & Christensen, R. 2004. Die Kanones der Auslegung als Hilfsmittel für die Entscheidung von Bedeutungskonflikten. In: *Juristische Arbeitsblätter*, H.36, 74–83.
- Kudlich, H. & Christensen, R. 2009. *Die Methodik des BGH in Strafsachen*. Köln: Heymanns.
- Lee, L. 1999. Measures of Distributional Similarity. In: *37th Annual Meeting of the ACL*, 25–32.
- Lemnitzer, L. & Zinsmeister, H. 2006. *Korpuslinguistik. Eine Einführung*. Tübingen: Gunter Narr Verlag.
- Mead, G. H. ¹³1965. *Mind, Self & Society. From the standpoint of a social behaviorist*. Chicago, London: The University of Chicago Press.
- Partington, A., Morley, J. & Haarman L. (Hrsg.). 2004. *Corpora and Discourse*. Bern: Peter Lang Verlag.
- Pêcheux, M. 1969. *Analyse automatique du discours*. Paris: Dunod.
- Putnam, Hilary ²1990: Die Bedeutung von „Bedeutung“. Herausgegeben und übersetzt von Wolfgang Spohn. Zweite Auflage. Frankfurt am Main: Klostermann.
- Reuber, P. & Pfaffenbach, C. 2005. *Methoden der empirischen Humangeographie*. Braunschweig: Wissenschaftliche Buchgesellschaft.
- Scharloth, J. & Bubenhofer, N. 2012. Datengeleitete Korpuspragmatik: Korpusvergleich als Methode der Stilanalyse. In: *Korpuspragmatik. Thematische Korpora als Basis diskurslinguistischer Analysen*, 195–230. Hrsg. von E. Felder, M. Müller und F. Vogel. Berlin/New York: Walter de Gruyter. (= *Linguistik – Impulse und Tendenzen* 44)
- Senkbeil, K. 2012. The Language of American Sports. A corpusassisted discourse study. Methodologische Überlegungen. In: *Korpuspragmatik. Thematische Korpora als Basis diskurslinguistischer Analysen*, 384–413. Hrsg. von E. Felder, M. Müller und F. Vogel. Berlin/New York: Walter de Gruyter. (= *Linguistik – Impulse und Tendenzen* 44).
- Spitzmüller, J. & Warnke, I. 2011. *Diskurslinguistik. Eine Einführung in Theorien und Methoden der transtextuellen Sprachanalyse*. Berlin/New York: Walter de Gruyter.
- Teubert, W. & Čermáková, A. 2007. *Corpus linguistics. A short introduction*. London: Continuum.
- Teubert, W. 2006. Korpuslinguistik, Hermeneutik und die soziale Konstruktion der Wirklichkeit. In *Linguistik online*, 28. Hrsg. von E. Hentschel. [Online-document] http://www.linguistik-online.de/28_06/teubert.htm (19. 03. 2008).
- Teubert, W. 2012. Von der Korpuslinguistik zur Analyse thematischer Diskurse. In: *Korpuspragmatik. Thematische Korpora als Basis diskurslinguistischer Analysen*, 231–278. Hrsg. von E. Felder, M. Müller und F. Vogel. Berlin/New York: Walter de Gruyter. (= *Linguistik – Impulse und Tendenzen* 44).
- Teubert, W., Hoey, M., Stubbs, M. & Mahlberg, M. 2007. *Text, Discourse and Corpora*. London: Continuum.
- Vogel, F. 2010a. *Linguistische Imageanalyse (LIma). Grundlegende Überlegungen und exemplifizierende Studie zum Öffentlichen Image von Türken und Türkei in*

- deutschsprachigen Medien. In: Deutsche Sprache (DS). Zeitschrift für Theorie, Praxis, Dokumentation (04/2010), 345–377.
- Vogel, F. 2010b. Ungarn – das Tor zum Westen. Das Bild Ungarns in deutschsprachigen Medien 1999 – 2009. Eine linguistische Imageanalyse. In: Jahrbuch der ungarischen Germanistik, 87–124. Hrsg. von A. Masát und E. Tichy. Bonn/Buda-pest: DAAD.
- Vogel, F. 2011. Gefechtsspuren im gesetzgebenden Diskurs: Die Debatte um Normierung von „Online-Durchsuchungen“ aus rechtslinguistischer Perspektive. In: Sprachreport 3/2011, 7–14.
- Vogel, F. 2012. Linguistik rechtlicher Normgenese. Theorie der Rechtsnormdiskursivität am Beispiel der Online-Durchsuchung. Berlin/New York: Walter de Gruyter. (= Sprache und Wissen 9)
- Warnke, I. & Spitzmüller, J. 2008. Methoden und Methodologie der Diskurslinguistik – Grundlagen und Verfahren einer Sprachwissenschaft jenseits textueller Grenzen. In: Methoden der Diskurslinguistik. Sprachwissenschaftliche Zugänge zur transtextuellen Ebene, 3–54. Hrsg. von I. Warnke und J. Spitzmüller. Berlin/New York: Walter de Gruyter.
- Warnke, I. 2007. Diskurslinguistik nach Foucault – Dimensionen einer Sprachwissenschaft jenseits textueller Grenzen. In: Diskurslinguistik nach Foucault. Theorie und Gegenstände, 3–24. Hrsg. von I. Warnke. Berlin/New York: Walter de Gruyter.

Anschrift des Verfassers:

Jun.-Prof. Dr. Friedemann Vogel
Universität Freiburg
Institut für Medienkulturwissenschaft
Werthmannstraße 16
79085 Freiburg im Breisgau
E-Mail: friedemann.vogel@medienkultur.uni-freiburg.de
Web: www.mkw.friedemann-vogel.de

