

# Korpuspragmatik

## Paradigma zwischen Handlung, Gesellschaft und Kognition

*Ekkehard Felder, Marcus Müller und Friedemann Vogel*

### Inhalt

1. Begriffsbestimmung und Motivation
  2. Infrastrukturen für die Korpuspragmatik
  3. Probleme und Leitfragen der Korpuspragmatik
  4. Zur Anlage dieses Bandes (Erläuterung der Makrostruktur)
  5. Zu den einzelnen Beiträgen
- Literaturverzeichnis

### 1. Begriffsbestimmung und Motivation

Pragmatik als linguistische Teildisziplin erforscht das Handeln in und mit Sprache. CHARLES W. MORRIS (1946) prägte den Ausdruck für die sich sukzessive etablierende Linguistik und fasste darunter die Lehre der Zeichenverwendung, die er der Syntax als Lehre der Zeichen-Kombinatorik und der Semantik als Lehre der Zeichen-Bedeutung an die Seite stellte. Neben der Dimension der Zeichen-Verwendung berücksichtigt die Morrissche Konzeption auch das Verhältnis von Zeichen und Zeichen-Benutzer – es handelte sich also um eine semiotische Perspektive. Die Pragmatik zeichnete sich immer schon durch eine spezifische Ambivalenz aus. Einerseits handelt es sich im obigen Sinne um ein Forschungsprogramm der Linguistik. Auf der anderen Seite ist die Pragmatik aber auch eine Art linguistische Weltanschauung, welche die Verwendungszusammenhänge von Sprache auf interaktionaler, kognitiver und sozialer Ebene als Explicans für deren lexikalische und syntaktische Verfasstheit auffasst. Wir haben es dabei mit einer Fokussierung der Sprachbetrachtung auf „Sprache in Situationen“ und der Perspektive des Sprachhandelns bzw. kommunikativen Wirkens von sprachlich agierenden Akteuren zu tun.

Von der korpuslinguistischen Forschung der letzten Jahre lässt sich sagen, dass sie in diesem zweiten Sinne nahezu vollständig sprachpragmatisch motiviert ist, und zwar insofern sie mit Bezug auf J.R. FIRTH (1957) Sprache als im Gebrauch situierte, serialisierte und kulturell kontextualisierte Ausdruckskomplexe begreift (ausführlich dazu

TOGNINI-BONELLI 2001: 157ff.).<sup>1</sup> In diese Richtung weist die Begriffsbestimmung von ARCHER u. a. (2008: 620):

Corpus pragmatics is so named because, like corpus linguistics in general, it involves analysing actual patterns of language use, using a collection of natural texts. Increasingly, these texts tend to be in an electronic form, which means that researchers are able to make use of computers when analysing their data.

Vor diesem Hintergrund lassen sich nun aber Studien, die ein sprachstrukturimmanentes Interesse haben, von solchen unterscheiden, die Regelmäßigkeiten in den Phänomenbereichen ‚Kognition‘, ‚Handlung‘, ‚Interaktion‘, ‚Gesellschaft‘ oder ‚empirische Epistemik‘ auf der Basis authentisch verwendeter Sprache (gesprochen und geschrieben) erforschen. Diese wollen wir unter dem Etikett *Korpuspragmatik* zusammenfassen. Diesem Typus von korpuslinguistischer Forschung stellt sich die Aufgabe, linguistische Fragestellungen mit pragmatischer Stoßrichtung mittels Korpora zu untersuchen, indem aus den sprachlichen Befunden der Korpusanalyse auf solche Phänomene der Handlungstypik (Form-Funktions-Korrelation; FELDER 2006, DEPPERMANN 2007) zu schließen ist, die mit sprachlichen Ausdrucksweisen systematisch im Zusammenhang stehen.<sup>2</sup> Diese Art von Korpuslinguistik, die in unserem Band adressiert wird, unterscheidet sich von der sprachsystematisch ausgerichteten wie etwa der Korpuslexikologie und -lexikographie oder der Korpusgrammatik dadurch, dass das allgemein geltende, mehr oder weniger saliente Glaubensfundament des Kontextualismus Auswirkungen auf praktische Erkenntnisinteressen und Methodologie der Studien hat. Das führt uns zu der folgenden Definition:

Unter Korpuspragmatik verstehen wir einen linguistischen Untersuchungsansatz, der in digital aufbereiteten Korpora das Wechselver-

- 
- 1 In Deutschland sind in diesem Zusammenhang auch die Arbeiten Feilkes (z.B. 1996) zur Begründung korpuslinguistischer Studien herangezogen worden (z.B. STEYER/BRUNNER 2009: 4). Ein Beispiel für die Kontinuität zwischen der älteren Gebrauchstheorie der Bedeutung und der Korpuslinguistik ist die korpuslexikologische Arbeit HERINGERS (1999), die sich ausführlich auf die klassischen Orte der Handlungssemantik, angefangen bei der Spätphilosophie Wittgensteins, beruft.
  - 2 Auf eine Eingrenzung in diesem Sinne läuft auch die Begriffsbestimmung in ARCHER u.a. (2008: 620) hinaus. Der Terminus „corpus pragmatics“ wird dort als Abgrenzung zur „computational pragmatics“ verwendet. Während die „computational pragmatics“ sich die technisch-handwerkliche Aufgabe stelle, Computern das Verständnis für das Inferenzproblem der sprachlichen Kommunikation zu ermöglichen, und zwar hauptsächlich in den Bereichen „abduction“, „reference resolution“, „interpretation and generation of speech acts“, „interpretation and generation of discourse structure and coherence relations“, so versuchten Forscher der „corpus pragmatics“ die genannten Problembereiche *sich selbst* verständlich zu machen.

hältnis zwischen sprachlichen Mitteln einerseits und Kontextfaktoren andererseits erforscht und dabei eine Typik von Form-Funktions-Korrelationen herauszuarbeiten beabsichtigt. Solche Kontextfaktoren betreffen potenziell die Dimensionen *Handlung*, *Gesellschaft* und *Kognition*. Die Analyse bedient sich insbesondere einer Kombination qualitativer und quantitativer Verfahren.

Diejenigen Grundsätze der Korpuspragmatik, aus welchen axiomatisch die Verfahrensbedingungen der einzelnen Forschungsfelder abgeleitet werden, sollen hier über den Ordnungsterminus ‚Diskurs‘ eingeführt werden: ‚Diskurs‘ wollen wir hier vorerst in der weiten Definition von ‚Sprachgebrauch‘ (z.B. FASOLD 1990) verstehen.<sup>3</sup> Diese Definition wird nun in drei Punkten spezifiziert, die wir für Sprache als Gegenstand der Korpuspragmatik geltend machen wollen.

Als Erstes sollen der Handlungscharakter und die multimodale Verfasstheit von Sprache betont werden, wie sie in der Begriffsdefinition VAN DIJKS (2008a: 116) zum Ausdruck kommt:

[...] I shall simply use the term “discourse” for any form of language use manifested as (written) text or (spoken) talk-in-interaction, in a broad semiotic sense. This includes visual structures, such as layout, letter type or pictures for written or printed text, and gestures, facework and other semiotic signs for spoken interaction.

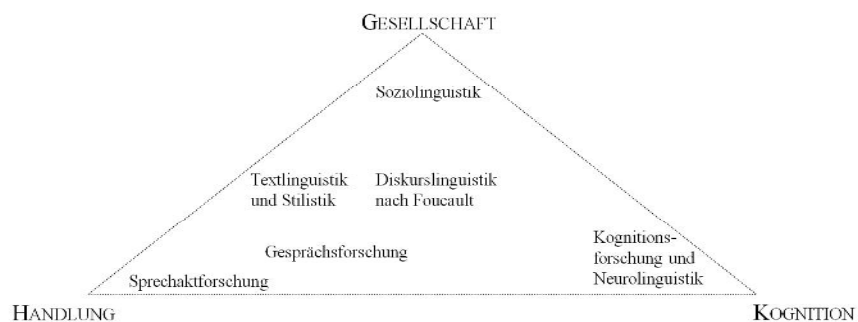
Insbesondere soll hier betont werden, dass für korpuspragmatische Studien in unserem Sinne prinzipiell schriftliche und mündliche Sprachdaten in gleicher Weise in Frage kommen, und sich die diesbezügliche Entscheidung über die Korpusbildung rein am jeweiligen Erkenntnisinteresse ausrichtet. Mit der Wahl der Modalität der Korpusdaten sind selbstverständlich auch grundlegende Entscheidungen zum Repräsentationsformat von Sprache im Korpus und ggf. zu Analyseverfahren verbunden. Eine besonders interessante Frage in diesem Zusammenhang ist die nach der Integration korpuslinguistischer Methoden in das Paradigma der Gesprächsforschung bzw. der interaktionalen Linguistik (vgl. dazu ARCHER u. a. 2008: 618; DEPPERMAN/HARTUNG in diesem Band).

Zweitens ist mit dem Terminus ‚Diskurs‘ das Theorem angesprochen, dass die Regelmäßigkeit von Sprache sich aus der Regelmäßigkeit von Sprachverwendungssituationen ableitet (FEILKE 1996, TOMASELLO 2006: 21). Diese Anschauung impliziert, dass sprachliche Muster nur unter Berücksichtigung der Varianz ihrer Verwendungskontexte be-

3 Diesen weiten Diskursbegriff legt VIRTANEN (2009) ihrem Überblicksaufsatz *Corpora and Discourse Analysis* zugrunde: „Discourse analysis [...] is here interpreted in a very wide sense, encompassing text linguistics, discourse analysis/studies, conversation analysis and pragmatics“ (VIRTANEN 2009: 1043).

schrieben werden können. Das sprachliche Verweisen auf Kontexte als verstehensrelevante Deutungsrahmen untersucht die Kontextualisierungsforschung (vgl. GUMPERZ 1982, AUER 1986, BUSSE 2007). Sie bildet einen wichtigen Bezugspunkt für die Korpuspragmatik. Bedenkt man in diesem Zusammenhang, dass ein großer Teil unseres individuellen Wissens auf der Rezeption von Sprachzeichen, wie sie in sprachlichen Äußerungen mündlicher und schriftlicher Art verbreitet werden, basiert, so ist evident, dass die auf diese Weise gewonnenen Wissensbestände und Erfahrungen in kommunikativen Formulierungsroutinen reproduziert und dadurch teilweise zu „kollektiven“ Wissensbeständen bzw. Erfahrungsmustern verdichtet werden (FELDER/MÜLLER 2009).

Drittens betont die zentrale Verwendung des Terminus ‚Diskurs‘ zur Explikation des programmatischen Etiketts ‚Korpuspragmatik‘, dass Sprache grundsätzlich in einem explikativen Feld zu verorten ist, das mit der Begriffstrias ‚Gesellschaft‘, ‚Handlung‘, ‚Kognition‘ umrissen ist. Dieses explikative Feld lässt sich nun zur Differenzierung der einzelnen Forschungsbereiche heranziehen. So wie auch in der traditionellen Pragmatik mehrere Pragmatiken unterschieden werden (vgl. ERNST 2002), so muss man also auch von unterschiedlichen Korpuspragmatiken sprechen.



Grafik 1: Das explikative Feld der Korpuspragmatik

In diesem Rahmen lassen sich nun folgende Korpuspragmatiken unterscheiden:

a) Textlinguistik und Stilistik

Korpusstudien, die strukturelle Untersuchungen auf Textebene vornehmen bzw. die transphrastische Sprachebene in den Blick nehmen, können dann ‚korpuspragmatisch‘ genannt werden, wenn sie transphrastische Phänomene als kontextsensitive Varianten modellieren. Ebenfalls zu dieser Kategorie rechnen wir solche variationslinguistischen Studien, welche die statistische Analyse mehrdimensionaler sprachstruktureller Muster auf Gattungs- und Medienzugehörigkeit beziehen, wie etwa der *multidimensional approach* im Rahmen der *register studies* (BIBER/CONRAD/REPPEN 1998, BIBER 2006, 2009). Insbesondere diese Studien sind in unserem explikatorischen Dreieck nicht leicht zu verorten, da zwar sowohl die interaktionale als auch die gesellschaftliche Dimension von Sprache in den Begründungszusammenhängen der Studien eine Rolle spielen (nämlich über den Situations-Begriff, vgl. dazu MÜLLER in diesem Band), es sich faktisch aber um vergleichende Analysen struktureller Variation handelt. Wegen der begrifflichen Grundlegung und dem variationslinguistischem Interesse ließen sich die *register studies* aber auch der Soziolinguistik zurechnen (s. u.). Die kognitive Ebene wird in den *register studies* explizit ausgeschlossen. In dieselbe Kategorie wollen wir korpusstilistische Arbeiten einordnen, die ebenfalls über mehrdimensional angelegte Strukturmerkmalsanalysen sprachliche Variation auf Textebene untersuchen, und zwar entweder bezogen auf kommunikative Gattungen, gesellschaftliche Gruppen oder einzelne Autoren (SEMINO/SHORT 2004, HOOVER 2007). Die Korpusstilistik gehört zu den ältesten Einsatzbereichen von Korpora überhaupt. Sie ist auch der Schlüssel zum Einsatz von Korpora zur Untersuchung literarischer Gattungen im Rahmen der Stylistics britischer Prägung (z.B. MAHLBERG 2007; Überblick bei BUSSE/MCINTYRE 2010, BUSSE 2010).

b) Sprechaktforschung

Studien, die von der Sprechakttheorie nach Searle ausgehen, nutzen Korpora, um über die Analyse von Sprachstrukturen Aussagen über Illokutionen und Handlungsmuster zu machen (Überblick in ARCHER u. a. 2008: 616f.). Dabei kann man zwei Typen unterscheiden: Studien des ersten Typs arbeiten qualitativ kategorienbildend im Sinne des *corpus-based approach* (z.B. KOHNEN in diesem Band), Studien des zweiten Typs berücksichtigen Häufigkeitsverteilungen und machen Aussagen über die quantitative Typizität von Illokutionen und Textfunktions-

typen. Beispielhaft für diese Forschung steht die *move-analysis*. (BIBER/CONNER/UPTON 2007). Dabei stellt sich die Aufgabe, einen Zusammenhang zwischen dem eruierten Strukturmuster und dem angenommenen Sprechakttyp plausibel zu machen (vgl. TAAVITSAINEN/JUCKER 2010: 16f.). Diese schwierigen Zuordnungsverfahren sind ja schon der sog. *armchair*-Variante der Sprechakttheorie zum Problem geworden und als solche bereits bei SEARLE (1971) unter dem Stichwort ‚indirekte Sprechakte‘ diskutiert worden.

### c) Soziolinguistik

ROMAINE (2008: 96) betont in ihrem Überblicksaufsatz *Corpus linguistics and sociolinguistics*, dass soziolinguistische Studien gewissermaßen schon immer Korpusstudien gewesen seien, insofern sie auf Sammlungen schriftlicher oder mündlicher Sprache basierten, welche unter einem bestimmten Aspekt (meist einer als ‚außersprachlich‘ gedachten Variable wie Region, Alter, Geschlecht, Klasse) zusammengestellt würden. Das gilt mindestens für die quantitative Richtung (LABOV 1966). Der Aspekt der strengen Empirie und Datenbasiertheit kann auch für die qualitativen Zugänge der interaktionalen Soziolinguistik (GUMPERZ 1982) geltend gemacht werden. Aus diesem klassischen Gegenstandsbereich der Soziolinguistik sind korpuslinguistische Verfahren im Rahmen der eher strukturell interessierten Variationsstudien nach der Art der *register studies* (s.o.) und neuerdings auch in dialektgeographischen Forschungen (Überblick in ANDERWALD/SZMRECSANYI 2009) angewandt worden (vgl. auch MAIR 2009). Um die oben genannten korrelativen Studien mit elektronischen Korpora und sprachstatistischen Verfahren durchführen zu können, bedarf es entweder Korpora, die mit den entsprechenden Metadaten ausgezeichnet sind (s. u.), z.B. das *British National Corpus (BNC)*, oder aber kleinere Spezialkorpora, deren Zusammenstellung auf einer soziolinguistisch relevanten Variablen beruht, z.B. das *Corpus of London Teenage Language (COLT)*. Viele variationslinguistische Studien beruhen auf Vergleichen von Korpora verschiedener Standardvarietäten, insbesondere für das Englische mittels der BROWN/LOB-Korpusgruppe (vgl. HUNDT 2008; BAKER 2010: 59ff.). Für das Deutsche sind solche Studien durch das Korpus C4-Projekt, einem Zusammenschluss des Digitalen Wörterbuchs der deutschen Sprache des 20. Jahrhunderts (DWDS), des Austrian Academy Corpus (AAC), des Korpus Südtirol und des Schweizer Text Korpus (CHTK), möglich.<sup>4</sup> Insbesondere auf der Grundlage von Daten des BNC wurde zu den Variablen Klasse

---

<sup>4</sup> <http://www.korpus-c4.org>

(ROMAINE 2008: 103) und Alter (ROMAINE 2008: 107f.), vor allem aber zur Gendervariable (SCHMID 2003; ROMAINE 2008: 104; BAKER 2010: 32ff.) gearbeitet. Eine Synthese aus quantitativen und qualitativen Zugängen bilden soziostilistisch gerahmte Corpus-driven-Studien, die mehr oder weniger explizit sozialkonstruktivistisch argumentieren (SCHARLOTH/BUBENHOFER 2010). ANDERSEN (2010) hebt in seinem Überblick zu korpuslinguistischen Studien im Bereich der Soziolinguistik die Bedeutung erstens der Korpora mündlicher Sprache und zweitens einer ausgebauten Metadatenstruktur hervor. BAKER (2010) führt in seinem Einführungsbuch neben den genannten auch diskurslinguistische Forschungen im engeren Sinne auf, die hier aber gesondert behandelt werden sollen (s. u.).

#### d) Kognitionsforschung und Neurolinguistik

Im florierenden Forschungsbereich der kognitiven Linguistik ist die Arbeit mit elektronischen Korpora schon lange einschlägig. Allerdings handelt es sich hier in vielen Fällen um Studien, in denen sprachstrukturelle Befunde kognitivistisch interpretiert werden (z.B. BYBEE 2006). Daneben gibt es Studien etwa aus dem Bereich der Wordnet-Projekte, in denen unter anderem ausführlich kognitivistische Interpretationen von Korpusbefunden mit experimentellen psycholinguistischen Studien verglichen werden (z.B. FELLBAUM u. a. 2005). Weniger gut ausgebaut ist der Rückgriff auf Korpora in der medizinischen Psycho- und Neurolinguistik (vgl. WENDELSTEIN/SATTLER in diesem Band). Hier stellt sich das grundlegende Problem der Vermittlung von Befunden, die überindividuell geprägte Ausdrucksmuster betreffen, und Aussagen über individuelle mentale Sachverhalte. Im neurolinguistischen Bereich geht das nur über Reihenuntersuchungen von Teilkorpora, die jeweils einzelnen Individuen zugeordnet sind (vgl. ebd.). Die Spracherwerbsforschung, die durch die Korpuslinguistik einen großen Schub erhalten hat, wäre hier anzuschließen –, da wir aber den unserem Band zugrundeliegenden Begriff von ‚Korpuspragmatik‘ nicht bis zur Unkenntlichkeit dehnen wollen, wird die Spracherwerbsforschung hier nicht weiter behandelt (vgl. den Überblick in DIESEL 2009).

#### e) Gesprochene-Sprache-Forschung und interaktionale Linguistik

Der korpuslinguistische Zugriff auf mündliche Sprache basiert bisher im Wesentlichen auf dem funktionalistischen Sprachvariationsmodell der *register studies*: Hier wird mündliche Sprache als Pool sprachlicher Register in gleicher Weise wie schriftliche Sprache behandelt. Untersucht werden klassisch textlinguistische Fragen wie die nach der Informationsstruktur (BIBER/CONRAD/REPPEN 1998: 108ff.), Fragen nach

sprachlichen Modellierungsmöglichkeiten interpersonalen Beziehungen z.B. mittels Diskurspartikeln (AIJMER 2002, 2008; BUTLER 2008) und solche nach der sprachstrukturellen Markierung verschiedener Situationstypen der Mündlichkeit (ADOLPHS 2008). Meist wird in diesen Studien vergleichend gearbeitet, entweder indem generell schriftliche mit mündlicher Sprache verglichen wird (BIBER/CONRAD/REPPEN 1998) oder indem verschiedene Register der Mündlichkeit verglichen werden (ADOLPHS 2008). In BIBER (2006) wird z.B. Variation in schriftlicher und mündlicher Sprache im akademischen Kontext zuerst bezogen auf Phänomene der Lexik und Grammatik und schließlich multidimensional untersucht. In jedem Fall steht die sprachstatistische Verteilung sprachlicher Ausdrücke auf Register und Situationstypen im Vordergrund.

Neben diesem breit ausgebauten Forschungsstrang wird der Einsatz von Korpora auch in der sequenzorientierten Gesprächsforschung diskutiert (ARCHER u. a. 2008: 618, s. DEPPERMAN/HARTUNG in diesem Band). Hier ist die Frage nach dem möglichen Einsatz von Korpora besonders interessant, aber auch besonders kontrovers, da sie stärker als in anderen Teildisziplinen den methodologischen Kern des Fachs berührt: Das grundlegende Methodologem, auf das sich viele Arbeiten aus der Gesprächsforschung berufen, besagt, dass sprachliche Befunde nur in ihrer Funktion für den interaktionalen, lokalen Aufbau von Sinn zu interpretieren sind (SPRANZ-FOGASY 1997, HAUSENDORF 2000, GÜNTNER 2000; DEPPERMAN 2007). Eine enggeführte Auslegung dieses Satzes schließt korpuslinguistische Studien per se aus, da ja hier mit Befunden gearbeitet wird, die mindestens im Bezug auf die Interaktionsstruktur der zugrunde liegenden Gespräche dekontextualisiert sind. Es ergibt sich zwar durch die Serialität der Befunde ein neuer Erklärungszusammenhang, der aber von den meisten GesprächsforscherInnen, insbesondere im Bezug auf die quantitativen Implikationen, kritisch beurteilt wird. Auf der anderen Seite ergibt sich vor allem durch Gesprächskorpora (WICHMANN 2008, ADOLPHS/KNIGHT 2010), für das Deutsche v.a. diejenigen des Mannheimer Instituts für deutsche Sprache (DEPPERMAN/HARTUNG, in diesem Band), eine wachsende Infrastruktur für serielle Untersuchung auch der Gesprächsfunktion sprachlicher Konstruktionen. Auch sind in der Gesprächsforschung Studien längst einschlägig, die mit einem klassischen *corpus-based approach* qualitativ typenbildend die Gesprächsfunktion einzelner Ausdrücke oder Ausdrucksmuster beschreiben, die Ergebnisse dann aber ausschließlich an einzelnen Sequenzanalysen plausibel machen (z.B. die Studie IMOS 2007 zur Gesprächsfunktion matrixfähiger Verben, die auf einem Korpus aus Transkripten der Sendung *Big Brother* basiert). Auch



die neu aufgelegte Transkriptionkonvention GAT 2 ist in Teilen für Korpusabfragen designed (SELTING u. a. 2009: 355). Ein Annotationsverfahren für die sequenzorientierte Korpusanalyse von Gesprächen schlägt ARCHER (2005) vor: Neben anderen soziopragmatischen Informationen werden zu jeder Sprecheräußerung sequenzbezogene annotiert. Archer setzt die Kategorien *initiation / response / response-initiation / report / follow up / follow up initiation* an (vgl. auch ARCHER/CULPEPER 2009: 6390f.).

Komplementär zu dieser Perspektive der Serialität sprachlicher Handlungen in Interaktionen scheint eine andere Entwicklung innerhalb der Gesprächsforschung zu stehen: die Analyse multimodaler Interaktion (SCHMITT 2007). Hier bieten sich aber erstens Verfahren der Datentriangulation an, zweitens besteht die Aufgabe darin, korpuslinguistisch verwertbare Repräsentationsformate für multimodale Interaktion zu entwickeln. Auf lange Sicht ist wahrscheinlich mit der Möglichkeit zu rechnen, multimodale Muster in den Bereichen der Mimik, Gestik und Proxemik computerunterstützt auszuwerten und ggf. mit klassisch korpuslinguistischen Befunden korrelieren zu können.

#### f) Diskurslinguistik nach Foucault

Studien, die auf einem engerem Diskursbegriff im Sinne von ›thematisch, funktional und intertextuell aufeinander bezogene Text- und Gesprächswelten‹ (BUSSE/TEUBERT 1994: 14) basieren, setzen einen engen Zusammenhang zwischen situational geprägter sprachlicher Interaktion, gesellschaftlicher Ausdrucksprägung und der Erkenntnisfunktion von Sprache an (KONERDING 2005, 2007). Insofern ist diese Art von Studien exakt in der Mitte unseres explikatorischen Dreiecks zu verorten. Begrifflich gehen wir von dem Sachverhalt aus, dass sich die meisten dieser Studien mehr oder weniger ausführlich auf die Arbeiten Foucaults beziehen, und greifen mit WARNKE (2007) auf die Etikettierung ‚Diskurslinguistik nach Foucault‘ zurück (vgl. den Überblick bei KONERDING 2009a). Während in WARNKE (2007) die Bezeichnung auch als Abgrenzung u. a. gegenüber der Critical Discourse Analysis (CDA) gemeint ist, sollen hier alle Strömungen der Diskursanalyse, deren Diskursbegriff die oben genannten Kriterien erfüllt, einbezogen werden. Hinsichtlich der methodologischen Positionierung innerhalb dieses Feldes stimmen wir aber mit WARNKE (2007) bzw. WARNKE/SPITZMÜLLER (2008) und SPITZMÜLLER/WARNKE 2011 überein. Diese Positionierung betrifft im Kern eine beschreibende Grundhaltung der Diskursanalyse gegenüber einer kritisierend-

wertenden.<sup>5</sup> In Diskursanalysen nach Foucault geht es meist darum, soziokognitive Gegebenheiten wie kollektive Einstellungen, Wissensformationen, und Mentalitäten oder gesellschaftliche Asymmetrien an der Analyse sprachlicher Muster auszuweisen. Dabei lassen sich – unabhängig von der methodologischen Verortung – vier Orientierungen hinsichtlich der Untersuchungsinteressen unterscheiden. Dabei ist zu bemerken, dass in vielen der Arbeiten, die im Folgenden genannt werden, zwei oder mehrere Forschungsorientierungen im Sinne von Mehrebenenmodellen (FELDER 2003, 2009, GARDT 2007, 2008, MÜLLER 2007, WARNKE/SPITZMÜLLER 2008, VOGEL 2009) zusammenwirken. Wir beschränken uns dabei auf sprachwissenschaftliche Diskursanalysen:

#### α) Handlungsorientierung

Handlungsorientierte Diskursstudien gehen entweder ausdrucksbezogen oder akteursbezogen vor. In ausdrucksbezogenen Studien wird die an Foucault orientierte Forschungsposition grundgelegt, um Systematiken der Konfiguration sprachlicher Ausdrucksweisen zu beschreiben. Das zentrale Erkenntnisziel ist in diesen Studien erreicht, wenn die entsprechenden Muster beschrieben und im Hinblick auf Funktions- und Wirkungspotentiale erörtert sind. Die Diskursfunktion dieser Ausdrücke dient dabei primär als Gliederungsaspekt. Eine typische Leitfrage solcher Studien unterliegt folgendem Muster: Mit welchen sprachlichen Mitteln werden bestimmte Begriffe, Themen und Perspektiven auf Sachverhalte in Diskursen konstituiert? (z.B. JUNG 1994, LIEBERT 2002, FELDER 2003, BUSCH 2004, KONERDING 2005, SPITZMÜLLER 2005, MÜLLER 2007, VOGEL 2009, LI 2011) Hier bietet sich die Korpuslinguistik als methodischer Zugriff unmittelbar an. Meist wird dabei so vorgegangen, dass signifikante sprachliche Muster in Teil- oder Vergleichskorpora festgestellt werden, welche nach diskurstheoretischen Gesichtspunkten zusammengestellt wurden. Die Ergebnisse werden dann entweder direkt diskurstheoretisch ausgedeutet (BUBENHOFER 2009) oder sie dienen als Vorstrukturierung für qualitative Analysen einzelner exemplarischer Textpassagen. Diese können eher auf die Diskursfunktion einzelner Ausdrucksmuster orientiert sein (MÜLLER 2009, MÜLLER im Druck), oder – im Sinne des hier unter γ) diskutierten Ansatzes – hermeneutisch auf Wissensformationen hinsichtlich einer oder

---

5 Vgl. zur Diskussion zwischen deskriptiver und kritischer Diskursanalyse auch den Band „Diskursanalyse zwischen Deskription und Kritik“, hg. von INGO WARNKE, MARTIN REISIGL und ULLA MEINHOF (in Vorb.)

mehrerer Sprachbenutzergruppen ausgedeutet werden (VOGEL 2009, LI 2011, FELDER in diesem Band).

Akteursbezogene Diskursstudien lassen sich grob in zwei Typen einteilen: Erstens deskriptiv orientierte Arbeiten, die im diskurslinguistischen Paradigma kategorienbildend nach Handlungstypen (LIEBERT 2002, FELDER 2003, LI 2011), Diskursstrategien (MÜLLER 2007) und Interaktionsmustern fragen (SPIEB 2011). Zweitens gesellschaftskritisch orientierte Arbeiten, welche Persuasionsstrategien einzelner dominanter Diskursakteure aufzeigen und hinterfragen (WODAK u. a. 1999). Besonders in diesem zweiten Ansatz, welcher eine Domäne der CDA ist, wird der Einsatz von Korpora schon relativ lange diskutiert und auch erprobt (PARTINGTON u. a. 2004, darin v. a. GARZONE/SANTULLI 2004; BAKER u. a. 2008; MAUTNER 2009; Überblick in BAKER 2010).

### β) Gesellschaftsorientierung

Außerdem sind Studien zu nennen, die in einer übergeordneten Perspektive die Gesellschaftlichkeit des Sprachgebrauchs in den Blick nehmen. Als Determinante sprachlicher Perspektivierungen wird dann weder der einzelne Akteur noch die Systematik des Sprechens gesehen, sondern vielmehr gesellschaftliche Verhältnisse, die entweder als ‚außersprachlich‘ oder als über das Sprachliche hinausgreifend gedacht sind. Untersuchungsgegenstände sind zum Beispiel gesellschaftliche Asymmetrien (Machtanalyse: FAIRCLOUGH 1989; VAN DIJK 2008b) oder die soziokommunikative Konstruktion bestimmter Wissensbereiche (KNOBLAUCH 1995, TEUBERT 2010). Dazu gehört eine unüberschaubare Zahl an Publikationen aus den verschiedenen Bereichen der qualitativen Sozialforschung, der Gesprächsforschung und der empirischen Linguistik, die sich schon durch ihre Titelformulierungen in die Nachfolge des soziologischen Klassikers von BERGER/LUCKMANN (2009) *Die gesellschaftliche Konstruktion der Wirklichkeit* stellen.<sup>6</sup> Korpuslinguistische Arbeiten in diesem Bereich sind vor allem im Bereich der kritischen Analyse von Mediendiskursen angesiedelt (WODAK/REISIGL 1999, BAKER 2005; BAKER/MCENERY 2005, MAUTNER 2009, TEUBERT 2010) und meist aus den großen Referenzkorpora des Englischen, v.a. aus dem *British National Corpus* heraus gewonnen.

---

6 Anstelle einer beliebigen Nennung einschlägiger Titel sei folgende Zahl genannt: Eine Anfrage beim Online-Bookseller Amazon mit dem Suchstring „the social construction of“ am 28.02.2011 bringt 15.956 Titel. Die Zahl mag die überaus große Beliebtheit dieser Formulierung und des damit verbundenen Forschungsprogramms illustrieren.

### γ) Wissensorientierung

Als dritten Typ nennen wir wissensorientierte Diskursanalysen. Hier lassen sich eher auf die individuelle Kognition eines angenommenen Diskursteilnehmers ausgerichtete Arbeiten (ZIEM 2008) von epistemologisch orientierten Analysen kollektiven Wissens unterscheiden. Als Untersuchungskategorien werden Modellbildungen (z.B. in der Techniksprache JAKOB 1991 und in der Wirtschaftssprache HUNDT 1995), Metaphern (in der Virologie LIEBERT 1995), Topoi (z.B. in der Politik WENGELER 2003 oder grundsätzlich KONERDING 2005, 2008), Sprachhandlungstypen (in der Klimatologie LIEBERT 2002 und im Recht FELDER 2003, LI 2011), Präsuppositionen (z.B. in Ideologien LOBENSTEIN-REICHMANN 2008) und Frames (KONERDING 1993, 2009, ZIEM 2008) herangezogen. In diesem Bereich ist das korpuslinguistische Vorgehen nicht sonderlich weit verbreitet. Das erklärt sich leicht dadurch, dass Wissensbestände nur interpretativ aus Texten plausibel gemacht werden können, ein qualitativer Zugang also unumgänglich erscheint. Allerdings haben sich hier korpuslinguistische Methoden zur Vorstrukturierung und Identifizierung zentraler Diskursmuster bewährt, in denen Wissenskonfigurationen ausgedrückt sind (BUBENHOFER 2008, HARDT-MAUTNER 1995, STUBBS 1996, WODAK/DE CILIA/REISIGL/LIEBHART 1999, VOGEL 2010). Auch können auf diese Weise umstrittene Wissensbereiche, „agonale Zentren“ (FELDER, STEGMEIER in diesem Band) identifiziert werden.

## 2. Infrastrukturen für die Korpuspragmatik

Wie sich bereits aus den im obigen Überblick eingestreuten Verweisen auf Korpora und Infrastruktur ergibt, sind korpuspragmatische Untersuchungen nicht auf einen bestimmten Korpus typ beschränkt. Prinzipiell sind alle Arten von Korpusdaten mehr oder wenig gut pragmatologisch ausdeutbar. Allerdings kristallisieren sich sehr wohl Infrastrukturen heraus, die sich für korpuspragmatische Untersuchungen besonders gut eignen.

Prinzipiell können korpuspragmatische Studien sowohl mit Referenz- als auch mit Spezialkorpora durchgeführt werden. Vor allem in den Bereichen Soziolinguistik und gesellschaftsorientierte Diskursanalyse nach Foucault wurde mit dem BNC korpuspragmatisch gearbeitet. Da das BNC zu ca. 10 % Transkripte gesprochener Sprache enthält, in die soziologische Sprecherinformationen zu Geschlecht, Alter und so-

zialen Status als markup im Kopf integriert sind<sup>7</sup>, eignet es sich insbesondere für korrelative soziopragmatische Studien (BAKER 2010: 32ff.). Familien von nationalen Referenzkorpora wie die BROWN/LOB-Gruppe für das Englische oder die C4-Korpusgruppe für das Deutsche sind für vergleichende Analysen von Standardvarietäten geeignet (s.o.). Auf Referenzkorpora beziehen sich in unserem Band Deppermann/Hartung, Becker, Kohnen und Mauthner. Teubert arbeitet im Sinne des *Web as Corpus*-Gedankens mit Internetdaten. Allerdings spricht für viele korpuspragmatische Anliegen doch einiges dafür, auf selbst anzulegende oder schon existierende Spezialkorpora zurückzugreifen. Damit kann erstens durch die Auswahl der Sprachdaten sichergestellt werden, dass der jeweilige Untersuchungsgegenstand auch angemessen repräsentiert ist, und zweitens können ggf. Informationen über kommunikative Gattungen, Situationen, Sprecherrollen u. ä. so kodiert werden, dass es dem Untersuchungsziel möglichst exakt dient. In der jüngsten Zeit hat es – parallel zum weiteren Ausbau der großen Referenzkorpora und zum *Web as Corpus*-Programm (HUNDT 2007) – verstärkte Bemühungen um den Aufbau von Spezialkorpora gegeben, die mit Text-, Situations- und Sprecherinformationen angereichert sind. ROMERO-TRILLO (2008: 6) spricht gar von einer Fülle neuer Korpora, die soziale, geographische und altersbezogene Informationen enthalten. Wir stimmen Romero-Trillo darin zu, dass solche Informationen für quantitative Analysen im Bereich der Korpuspragmatiken unerlässlich sind. Spezialkorpora liegen im vorliegenden Band den Beiträgen von Bär/von Consbruch, Felder, Frank, Müller, Scharloth/Bubenhofer, Senkbeil, Stegmeier, Vogel und Wendelstein/Sattler zu Grunde.

### Das Heidelberger Korpus

Die Beiträge Felder, Müller, Stegmeier, Vogel und Wendelstein beziehen sich jeweils auf Teilkorpora des ‚Heidelberger Korpus‘ (HeideKo). HeideKo wurde von den Herausgebern dieses Bandes als eine gemeinsame Infrastruktur für Spezialkorpora konzipiert. Die einzelnen Teilkorpora wurden unter thematischen Gesichtspunkten zusammengestellt, es handelt sich also um eine Forschungsumgebung für diskurslinguistische Untersuchungen (FELDER/MÜLLER/VOGEL 2010). Gemeinsames Erkenntnisinteresse der Teilprojekte ist die Analyse gesellschaftlicher Konflikte im Spiegel der Sprache. Methodisch besteht das gemeinsame Ziel in der Weiterentwicklung diskurslinguistischer Metho-

<sup>7</sup> Ausgestaltet nach den Maßgaben der *Text Encoding Initiative* (TEI): <http://www.tei-c.org/Guidelines/>

den durch eine Kombination aus quantitativen mit qualitativen Analyseverfahren (z.B. MÜLLER i.Dr.). Die zentrale Datenorganisation ermöglicht einzelnen Untersuchungsprojekten zum einen kontrastive Analysen durch Rückgriff auf andere themengebundene Referenzkorpora, zum anderen den Zugriff auf gemeinsame Analyseinstrumente. Das Heidelberger Korpus basiert auf einer hierarchisch strukturierten Access-Datenbank, deren Funktionalität durch VBA- und Python-Programmierung im Hinblick auf diskurslinguistische Fragestellungen (wie die Visualisierung von intertextuellen Verknüpfungen u. ä.) erweitert wird. Es handelt sich um eine multimediale Datenbank mit den Dateitypen Text, Transkript, Audiofile, Videofile, PDF, Bild. Für jeden Dateityp stehen separate Metadaten-Sets zur Verfügung. Die einzelnen Datensätze und -typen lassen sich untereinander in Beziehung setzen. Für die nähere Arbeit mit dem HeideKo wurden diverse Import-, Bearbeitungs- und Exportformulare entwickelt, so dass auf eine Arbeit mit Kommandozeilen weitestgehend verzichtet und HeideKo auch in die Lehre integriert werden kann. Die Exportformulare ermöglichen dabei Filterung, Vorsortierung auf Basis der Metadaten und einfachen Export der Datenbank-Daten in externe Analysesoftware (wie AntConc; vgl. L. ANTHONY 2005) und Formate (TXT, XML). Benutzerfreundlichkeit, umfassende Dokumentation und Tagsets der Korpora garantieren eine nachhaltige Korpusnutzung. Die Korpusprimärdaten werden umfassend digitalisiert und in Rohform in die Datenbank eingespeist. In individuellen Teilkorpora können beim Export automatisch Wortarten annotiert werden (POS-Tagging mit Hilfe des am IMS entwickelten Treetaggers<sup>8</sup>). Die mündlichen Sprachdaten werden orthographisch transkribiert.<sup>9</sup> Die am HeideKo beteiligten Teilkorpora werden mit gleichen Basismetadaten sowie projektspezifischen Metadaten ausgezeichnet. Neben den Basismetadaten zu Medium, Erscheinungsdatum usw. werden Text-, Gesprächs- bzw. Bildsorte, Oberthema, Redegegenstand, situatives Setting sowie akteursbezogene Daten wie Akteursberuf und -institution, Sprecherrolle und soziale Rolle vermerkt. Die sehr genaue und aufwändige Auszeichnung von Sprecherdaten erlaubt die Analyse von Korrelationen von sprachlichen und sozialen Mustern und ist damit die Kerninformation der korpuspragmatischen Arbeit. Dazu wurde jeder Diskursbeitrag, also auch jeder einzelne turn aus transkri-

---

8 <http://www.ims.uni-stuttgart.de/projekte/corplex/TreeTagger/>

9 Als Abweichung von der orthographischen Norm gilt die generelle Kleinschreibung (nach GAT). Markierte Betonungen werden mit Großschreibung gekennzeichnet. Markierte para- und nonverbale Performanzen werden nach GAT in filterbaren Klammerausdrücken annotiert.

bierten Gesprächen oder Wortlautprotokollen in ein eigenes Datenblatt importiert, so dass statistische Aussagen über Korrelationen von Sprecherrollen, Situationen, Gattungen, Medium und Sprachstrukturmustern ermöglicht werden. Der interaktionale Kontext der primärmündlichen Daten wird über eine im Hintergrund in XML implementierte Thread-Struktur dargestellt. Derzeit umfasst HeideKo 22396 Texte (12,46 Mio. Token) und 47 Transkripte (283.828 Token).

### 3. Probleme und Leitfragen der Korpuspragmatik

In dem folgenden Kapitel wollen wir in aller gebotenen Knappheit die aus unserer Sicht wichtigsten Fragen und Probleme aufwerfen, die sich aus der korpuspragmatischen Arbeit ergeben:

- a. Wie kann man von sprachstrukturellen Befunden auf sprachpragmatische schließen (Form-Funktions-Korrelation)? Und vice versa: Welche typisierten Sprachgebrauchs- und Handlungskonstellationen weisen bestimmte Spezifika im Sprachmaterial auf (Funktions-Form-Korrelation)?

Die klassische pragmatische Begriffsbildung, z.B. im Rahmen der Sprechakttheorie, beruht auf nicht empirischen, introspektiven Zugängen zur sprachlichen Praxis. Schon in deren Rahmen, besonders aber innerhalb der empirischen Pragmatik, stellt sich das Problem der Zuordnung sprachlicher Strukturmuster zu Sprachhandlungstypen und Kontextschemata. Ganz allgemein formuliert: Wie können Zusammenhänge zwischen Formen, Funktionen und Kontexten untersucht werden?<sup>10</sup> Im Bereich der Korpuspragmatik wirft diese Frage vor allem technische Probleme auf: Unter welchen Prämissen soll nach welchen Ausdrücken gesucht werden? Wie lassen sich Klassen funktionsäquivalenter Ausdrücke bilden und quantitativ plausibilisieren?<sup>11</sup> Wie können Korpora annotiert oder mit Metadaten ausgezeichnet werden, so dass sich Form-Funktions-Korrelationen errechnen lassen? Welche mehrdimensionalen Auswertungsverfahren gibt es, um Korrelationen zwischen Sprachmustern und mehreren Kontextdimensionen zu berechnen?

---

10 Zur Problematik aus empirisch-korpuslinguistischer Perspektive vgl. BELICA/KEIBEL/KUPIETZ/PERKUN (2010).

11 Einen originär induktiven Ansatz zur automatischen Exploration von Funktionsäquivalenz allein auf Basis von systematischen Kontextvergleichen beschreiben etwa VACHKOVÁ/BELICA (2009).

Die Frage nach der Auszeichnung von Korpusdaten mit Kontextinformationen wird gelöst, indem diese entweder in die Datenblätter selbst integriert (wie z.B. in Teilen des BNC) oder in korrelierten Metadateien kodiert werden (wie im HeideKo). ARCHER/CULPEPPER (2003) haben in einem aufwendigen Verfahren ein historisches Gesprächskorpus mit dynamischer Kontextinformation kodiert, indem jede einzelne Äußerung mit soziologischen Minimalinformationen zu Sprecher und Adressaten annotiert wird. Dieses überaus interessante Verfahren ist aber nur für sehr kleine Korpora praktikierbar und durch den hohen Anteil an parzellierter Interpretationsarbeit auch sehr fehleranfällig.

b. Welche theoretischen und methodologischen Implikationen hat der Einsatz von Korpora in der Sprachpragmatik? Welche Forschungsbereiche und Forschungsaspekte werden evtl. ausgeblendet, wenn verstärkt mit Korpora gearbeitet wird?

Der fortschreitende Einsatz von Sprachkorpora in den verschiedenen Richtungen der linguistischen Pragmatik hat selbstverständlich Konsequenzen erstens für Konjunkturen von Forschungsthemen und zweitens für die implizite oder explizite Sprachtheorie der Pragmatik: Den meisten korpuspragmatischen Studien liegt eine implizite Diskurstheorie zugrunde, die den Aspekt der Serialität und Musterhaftigkeit von Sprache betont. Es ergibt sich in der Gesamtschau das Bild einer Entsubjektivierung des Gebrauch-Begriffs. Hinsichtlich der Forschungsgegenstände überwiegen naturgemäß solche Themen, die ausdrucksbezogen angegangen werden können, die also etwa die Diskursfunktion bestimmter Partikel, Lexeme oder sprachlicher Muster betreffen. Außerdem werden solche Gattungen bevorzugt, die leicht über Korpora erschließbar sind bzw. von den öffentlich zugänglichen Korpora bereits abgedeckt sind. Das betrifft insbesondere einerseits die Presse- und Literatursprache, andererseits auch die mündlichen und schriftlichen Gattungen des akademischen Diskurses (FLOWERDEW 2002; BIBER 2009). Viele bisher unbeachtete Bereiche, auch und gerade in der historischen Pragmatik (JUCKER/TAAVITSAINEN 2010), können durch den Aufbau neuer Korpora erschlossen werden. Viele, insbesondere historische Varietäten der Mündlichkeit sind aber prinzipiell nicht durch Korpusbildung zu erschließen, sondern müssen durch indirekt schließende Verfahren erforscht werden. Hier scheint es wichtig, darauf zu achten, dass fehlende Korpora nicht dazu führen, dass ganze Forschungsbereiche keine Aufmerksamkeit mehr erfahren.



c. Welche Möglichkeiten der Methodentriangulation ergeben sich?

Um einen nachhaltigen *corpus bias* in den linguistischen Pragmatiken zu vermeiden und die Ergebnisse der Korpusstudien durch anderweitig erzielte Ergebnisse zu ergänzen, bietet sich für viele Bereiche Methodentriangulation an (BAKER 2006: 15ff.). Das betrifft zum einen die korpusinterne Verbindung von qualitativen mit quantitativen Ansätzen, zum anderen den komplementären Einsatz etwa von kommunikationsethnographischen und konversationsanalytischen Methoden in der Soziolinguistik, der Gesprächsforschung und der Diskurslinguistik und von experimentellen Methoden in der Sprechaktforschung und der kognitiven Linguistik. Solche auf Triangulation beruhende Studien sind in den meisten angesprochenen Bereichen aber noch weitgehend ein Desiderat. Am weitesten scheint hier die kognitive Linguistik fortgeschritten (z. B: FELLBAUM u. a. 2005). Die Durchführung solcher Studien setzt voraus, dass die Korpuslinguistik weder als revolutionäres Allheilmittel der Pragmalinguistik überhöht noch als technokratisches Teufelszeug verdammt wird und die verschiedenen Methodenkreise der empirischen Pragmatik sich nicht in Grabenkämpfen aufreiben, sondern als komplementäre Werkzeuge einer multiperspektivischen Sprachanalyse aufgefasst werden.

#### 4. Zur Anlage dieses Bandes (Erläuterung der Makrostruktur)

Im Mittelpunkt des in diesem Sammelband adressierten Forschungskontextes stehen Untersuchungen von Form-Funktions-Korrelationen an Korpusdaten. An Beispielen themengebundener Korpora (also inhaltlich zusammengehörender Texte und Gespräche als Ausschnitt eines Diskurses) sollen die spezifischen Erkenntnispotentiale eines Zusammenspiels von korpuslinguistischen Beschreibungsverfahren und sprachpragmatischen Fragestellungen demonstriert werden. Die Aufmerksamkeit richtet sich in erster Linie auf die sprachimmanente Perspektivität sprachlicher Performanz in thematisch zusammenhängenden Texten und Gesprächen. Dabei gilt das spezifische Untersuchungsinteresse bei klarer Explizierung einer inhaltlichen und methodischen Fragestellung primär den sprachlichen Mitteln und Mustern im Spiegel von kontextabstrahierten Verwendungssituationen. In diesem Sinne verbinden die Beiträge hermeneutische Aspekte mit methodischen Problemen der Operationalisierung von Sprachbeschreibungen.

Die Beiträge des Bandes kreisen um drei Zentren. In einem ersten Teil *Methoden und Zugänge* sind Beiträge versammelt, die schwerpunktmäßig methodische Fragen im Umfeld des oben dargelegten

Untersuchungsprogramms der Korpuspragmatik behandeln. Alle methodologischen Fragen der korpuslinguistischen Analysemöglichkeiten werden hinsichtlich ihres epistemologischen Potentials für Diskursanalysen erörtert.

In einem zweiten Teil *Studien und Ergebnisse* liegt der Fokus auf einzelnen Untersuchungen. In deren Mittelpunkt stehen Korpora aus den Themenbereichen rituelle Kommunikation, Recht, Religion und Sport. Die Autoren zeigen an konkreten Fragestellungen, wie mit Hilfe korpuslinguistischer Verfahren aufschlussreiche Untersuchungsergebnisse erzielt werden können.

Der dritte Teil ist mit *Korpora und Projekte* überschrieben. Dort werden vor allem konzeptionelle Fragen aktueller Forschungsprojekte in Form von Werkstattberichten vorgestellt.

## 5. Zu den einzelnen Beiträgen

Marcus Müller behandelt ‚Kontext‘ als zentralen heuristischen Terminus der Korpuspragmatik: Nach einer Bestimmung von Sprache als Spur sozialer Interaktion werden die Kontextbegriffe der systemic functional linguistics und der Kontextualisierungstheorie diskutiert. Daraus wird ein mehrschichtiges Modell der Kontextualisierung entwickelt, welches als Basis für einen methodologischen Grundlegungsvorschlag für die Korpuspragmatik dient. Der interaktionale Begriff ‚Kontextualisierungshinweis‘ wird dazu im Sinne von ›Serialisierung eines sprachlichen Merkmals relativ zu Kontexten‹ probabilistisch weiterentwickelt. Es folgen zwei Exemplifizierungen an einem Korpus zur gegenwärtigen Bioethikdebatte.

Gerlinde Mautner diskutiert Möglichkeiten und Grenzen, Korpuslinguistik (CL) und kritische Diskursanalyse (CDA) miteinander zu kombinieren. Sie sieht erhebliches Potential nicht nur für die Effizienz, mit der Erkenntnisse über die diskursive Konstitution sozialer Realität gewonnen werden können, sondern auch für eine neue Qualität von Erkenntnissen: zum Beispiel das Entdecken von kollokativen Bindungen, die ohne Computerunterstützung nicht offenbar würden.

Ekkehard Felder stellt eine Linguistische Mediendiskursanalyse im Paradigma der pragma-semiotischen Textarbeit vor. Hierbei entwickelt er semi-automatische Methoden zur Bestimmung von Diskursthemen, Subthemen und agonalen Zentren (= konfligierende Geltungsansprüche um Wahrheitsaussagen) und exemplifiziert diese anhand eines thematischen Mauer-Korpus auf Lexem-, Syntagmen-, Satz-, Text- und Text-Bild-Ebene. Ein Ziel des Beitrags besteht darin, sog. agonale Zentren (im Sinne diskursiver Wettkämpfe um Geltungsansprüche) als Kern-

und Kristallisationspunkte von Diskursstreitigkeiten induktiv aus dem sprachlichen Material heraus transparent zu machen.

Thomas Kohnen untersucht den Zusammenhang von korpuslinguistischen Analysen und Performativität. Dabei steht methodisch das Grundproblem der Identifizierung und Zuordnung von Sprechakten in Sprachkorpora im Vordergrund. Form und Funktion seien prinzipiell kontingent, was automatische Zugänge vor erhebliche Probleme stelle. Gleichwohl, so illustriert der Autor an religiöser Prosa, ermöglichen annotierte Korpora sowie corpus-based-Studien zu mehr oder weniger ‚festen‘ Form-Funktions-Korrelaten globale Aussagen über die Performativität eines Textes.

Joachim Scharloth und Noah Bubenhofer illustrieren an ausgewählten Beispielen die in der Forschergruppe semtracks entwickelte Methode der datengeleiteten Stilanalyse als Grundlage zu einer Rekonstruktion kommunikativer Stile. Hierfür berechnen sie auf Basis POS-annotierter Korpora einzelne „ideologische“ Schlagwörter sowie kontrastive Pentagramme als Hinweise für gruppenspezifische Stilmerkmale.

Wolfgang Teubert betont das konstruktivistische Paradigma, nach dem alle Wirklichkeit nur in Diskursen konstituiert werde. Sprache als das Konstitutivum schlechthin sei genauso wie die Wirklichkeit selbst kontingent. Korpuslinguistische Methoden könnten diese Kontingenz transparent machen und damit zu einer Diskursanalyse beitragen, die sich in der Extraktion, Aufbereitung und demokratischen Bereitstellung von Daten zugunsten unterschiedlicher Interpretationsvorgänge in der Diskursgemeinschaft verdient mache.

Nils Reiter, Oliver Hellwig und Anette Frank stellen ihre computerlinguistische Arbeit zur Analyse von Ritualdynamiken auf Basis von NLP-Tools und schrittweise annotierten Korpora vor. Ihre ersten framebasierten Ansätze zeigen das offene Potential computergestützter Analysen als Bereicherung in Sozial- und Geisteswissenschaften.

Friedemann Vogel prüft das Erkenntnispotential (semi)automatischer Analyseinstrumente beim Nachvollzug juristischer Text- und Normarbeit. Im Fokus stehen dabei rekurrente Sprachmuster in großen Textkorpora als Spuren sedimentierter Rechtsdogmatik. Nach Klärung der theoretischen und methodologischen Vorannahmen illustriert der Autor einen korpusbasierten Zugang zur Kontextualisierung der Würde des Menschen (Art. 1 Abs. 1 GG) in Entscheidungen des Bundesverfassungsgerichts.

Martin Becker zeigt die Möglichkeiten korpuslinguistischer Zugänge im Bereich der Sprachgeschichte auf. Am Beispiel einer Korpusstudie zur Verwendung epistemischer Prädikate in der *Institution de la*

*religion chrétienne* Jean Calvins kommt er zu dem Schluss, dass quantitative und grundlegende qualitative Analyse notwendig miteinander zu vermitteln seien. Besonders das Zusammenspiel der Semantik der Prädikate und der Pragmatisierung ihres semantischen Potentials im konkreten Diskurs seien zu berücksichtigen, weil es wichtige Einsichten über die Textintention und ihre sprachliche Realisierung aufzuzeigen vermöge.

Karsten Senkbeil exemplifiziert im Paradigma der Corpus-Assisted Discourse Studies einen Zugang zu Diskursen auf Basis systematischer Keyword-Analysen. Der Autor zeigt, wie mittels kontrastiver Auswertung von Korpuswortlisten versteckte oder offenbare Ideologiemuster und damit verbundene Diskurssäulen im amerikanischen Sportdiskurs transparent und einer qualitativen Detailanalyse zugänglich gemacht werden können.

Arnulf Deppermann und Martin Hartung diskutieren Kriterien und Probleme bei der Erstellung systematischer Gesprächskorpora. Am Beispiel des Aufbaus des „Forschungs- und Lehrkorpus Gesprochenes Deutsch“ (FOLK) am Institut für Deutsche Sprache (Mannheim) diskutieren sie das Korpusdesign im Hinblick auf Ausgewogenheit bzw. Repräsentativität, Parametrisierung, Nutzungspotentiale sowie praktische und forschungsethische Probleme.

Jochen Bär und Benita von Consbruch diskutieren an einem ausgewählten Beispiel – einem Wörterbuchprojekt zum literatur- und kunsttheoretischen Diskurs der Goethezeit – Kriterien und Probleme bei der Arbeit mit Korpora in der historischen Lexikographie. In diesem Kontext sei die Erstellung eines Korpus nicht lediglich mechanische Zusammenstellung von Datenmengen, sondern vielmehr eine im vollen Sinne philologische Arbeit in der Auseinandersetzung mit historischen Texten, Quellenbasen und Kontext-Fachwissen.

Britta Wendelstein und Christine Sattler stellen ihre korpuslinguistischen Untersuchungen im Kontext psychologisch-psychiatrischer Forschung vor. Auf Grundlage transkribierter und mittels Metadaten ausgezeichnete biographischer Interviews (ILSE) sollen Unterschiede in sprachlichen Ausdrucksfähigkeiten in Beziehung zu kognitiven Leistungsständen und sozioökonomischen Faktoren gesetzt werden, um als Frühsymptom oder Ausdruck der kognitiven Reserve eingeordnet werden zu können.

Jörn Stegmeier reflektiert im Rahmen von Diskursanalysen auf der Basis umfangreicher Textkorpora darüber, in welchem komplementären Verhältnis hermeneutisch orientierte Textlektüre und semi-automatisierte Verfahren der Korpuslinguistik bei der linguistischen Untersuchung stehen können. Die Überlegungen sind in die Konzeption

und Zielsetzung des Forschungsprojekts „Computergestützte Diskursanalyse – Eine E-Learning-Plattform“ eingebettet und werden anhand eines Textkorpus zum Themenbereich „Palliativmedizin/Sterbehilfe“ exemplifiziert.

### Literaturverzeichnis

- ÄDEL, ANNELIE / REPPEN, RANDI (Hgg.) (2008): *Corpora and discourse. The challenges of different settings*. Amsterdam, Philadelphia.
- ADOLPHS, SVENJA (2008): *Corpus and Context. Investigating pragmatic functions in spoken discourse*. Amsterdam.
- ADOLPHS, SVENJA / KNIGHT, DAWN (2010): Building a spoken corpus. What are the basics? In: O'KEEFE, ANNE / MCCARTHY, MICHAEL (Hgg.): *The Routledge Handbook of Corpus Linguistics*. London. 38–52.
- AIJMER, KARIN (2002): *English Discourse Particles: Evidence from a Corpus*. Amsterdam, Philadelphia.
- AIJMER, KARIN (2008): At the interface between grammar and discourse – a corpus based study of some pragmatic markers. In: ROMERO-TRILLO, JESÚS (Hg.): *Pragmatics and Corpus Linguistics. A mutualistic entente*. Berlin, New York. 11–36.
- ANDERSEN, GISLE (2010): How to use corpus linguistics in sociolinguistics. In: O'KEEFE, ANNE / MCCARTHY, MICHAEL (Hgg.): *The Routledge Handbook of Corpus Linguistics*. London. 547–562.
- ANDERWALD, LIESELOTTE / SZMRECSANYI, BENEDIKT (2009): Corpus linguistics and dialectology. In: LÜDELING, ANKE / KYTÖ, MERJA (Hgg.): *Corpus Linguistics. An International Handbook*. Volume 2. 1126–1140.
- ANTHONY, LAURENCE (2005). AntConc: design and development of a freeware corpus analysis toolkit for the technical writing classroom Professional Communication Conference. IPCC 2005. 729–737.
- ARCHER, DAWN (2005): Questions and answers in the English courtroom (1640–1760): A sociopragmatic analysis. Amsterdam.
- ARCHER, DAWN / CULPEPER, JONATHAN (2003): Sociopragmatic annotation: New directions and possibilities in historical corpus linguistics. In: WILSON, ANDREW / RAYSON, PAUL / MCENERY, TONY (Hgg.): *Corpus linguistics by the Lune: A Festschrift for Geoffrey Leech*. Frankfurt am Main.
- ARCHER, DAWN / CULPEPER, JONATHAN / DAVIES, MATTHEW (2008): Pragmatic Annotation. In: LÜDELING, ANKE / KYTÖ, MERJA (Hgg.): *Corpus Linguistics. An International Handbook*. Volume 1. 613–641.
- ARCHER, DAWN / CULPEPER, JONATHAN (2009): Identifying key sociophilological in plays and trial proceedings (1640–1760): An empirical approach via corpus annotation. In: *Journal of Historical Pragmatics* 10.2 (2009). 260–283.
- AUER, PETER (1986): Kontextualisierung. In: *Studium Linguistik* 19 (1986). 22–47.
- AUER, PETER / DI LUZIO, ALDO (Hgg.) (1992): *Contextualization of language*. Amsterdam.
- BAKER, PAUL (2005): *Public Discourses of Gay Men*. London.
- BAKER, PAUL (2006): *Using Corpora in discourse analysis*. London.
- BAKER, PAUL (2010): *Sociolinguistics and corpus linguistics*. Edinburgh.

- BAKER, PAUL / MCENERY, TONY (2005): A corpus-based approach to discourses of refugees and asylum seekers in UN and newspaper texts. In: *Journal of Language and Politics* 4.2 (2005). 197–226.
- BAKER, PAUL / GABRIELATOS, COSTAS / KHOSRAVINIK, MAJID / KRZYZANOWSKI, MICHAL / MCENERY, TONI / WODAK, RUTH (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 (2008). 273–306. [WWW-document]  
<http://das.sagepub.com/cgi/content/abstract/19/3/273> (13.03.2009)
- BELICA, CYRIL / KEIBEL, HOLGER / KUPIETZ, MARC / PERKUHN, RAINER (2010): An empiricist's view of the ontology of lexical-semantic relations. In: STORJOHANN, PETRA (Hg.): *Lexical-Semantic Relations. Theoretical and practical perspectives*. 115–144 (*Linguisticæ Investigationes Supplementa* 28).
- BERGER, PETER L. / LUCKMANN, THOMAS (<sup>22</sup>2009): *Die gesellschaftliche Konstruktion von Wirklichkeit*. Frankfurt am Main.
- BIBER, DOUGLAS (1994): An analytical framework for register studies. In: BIBER, DOUGLAS / FINEGAN, EDWARD (Hg.): *Sociolinguistic perspectives on register*. New York. 31–56.
- BIBER, DOUGLAS (2006): *University Language. A corpus-based study of spoken and written registers*. Amsterdam, Philadelphia.
- BIBER, DOUGLAS (2009): Multi-dimensional approaches. In: LÜDELING, ANKE / KYTÖ, MERJA (Hgg.): *Corpus Linguistics. An International Handbook. Volume 2*. Berlin. 822–855.
- BIBER, DOUGLAS / CONRAD, SUSAN / REPPEN, RANDI (1998): *Corpus linguistics. Investigating language structure and use*. Cambridge.
- BIBER, DOUGLAS / CONNOR, ULLA / UPTON, THOMAS A. (2007) *Discourse on the Move: Using Corpus Analysis to Describe Discourse Structure*. Amsterdam.
- BUBENHOFER, NOAH (2008): Diskurse berechnen? Wege zu einer korpuslinguistischen Diskursanalyse. In: WARNKE, INGO / SPITZMÜLLER, JÜRGEN (Hgg.): *Methoden der Diskurslinguistik. Sprachwissenschaftliche Zugänge zur transtextuellen Ebene*. Berlin, New York. 407–434 (*Linguistik – Impulse und Tendenzen* 31).
- BUBENHOFER, NOAH (2009): *Sprachgebrauchsmuster. Korpuslinguistik als Methode der Diskurs- und Kulturanalyse*. Berlin, New York.
- BUBENHOFER, NOAH / SCHARLOTH, JOACHIM (2010): Kontext korpuslinguistisch: Die induktive Berechnung von Sprachgebrauchsmustern in großen Textkorpora. In: KLOTZ, PETER / PORTMANN-TSELIKAS, PAUL R. / WEIDACHER, GEORG (Hgg.): *Kontexte und Texte. Studien zu soziokulturellen Konstellationen literalen Handelns*. Tübingen. 85–109.
- BUSCH, ALBERT (2004): *Diskurslexikologie und Sprachgeschichte der Computertechnologie*. Tübingen.
- BUSSE, BEATRIX / MCINTYRE, DAN (Hgg.) (2010): *Language and Style*. London.
- BUSSE, BEATRIX (2010): Recent Trends in New Historical Stylistics. In: BUSSE, BEATRIX / MCINTYRE, DAN (Hgg.): *Language and Style*. London. 32–54.
- BUSSE, DIETRICH / TEUBERT, WOLFGANG (1994): Ist Diskurs ein sprachwissenschaftliches Objekt? In: BUSSE, DIETRICH / HERMANN, FRITZ / TEUBERT, WOLFGANG (Hgg.): *Begriffsgeschichte und Diskursgeschichte*. 10–28.

- BUSSE, DIETRICH (2007): Diskurslinguistik als Kontextualisierung – Sprachwissenschaftliche Überlegungen zur Analyse gesellschaftlichen Wissens. In: WARNKE, INGO (Hg.): Diskurslinguistik nach Foucault. Theorie und Gegenstände. Berlin, New York. 81–105 (Linguistik – Impulse und Tendenzen 25).
- BUTLER, CHRISTOPHER S. (2008): The subjectivity of basically in British English – a corpus based study. In: ROMERO-TRILLO, JESÚS (Hg.) (2008): Pragmatics and Corpus Linguistics. A mutualistic entente. Berlin, New York. 37–64.
- BYBEE, JOAN (2006). From usage to grammar: The mind's response to repetition. In: Language 82. 4 (2006). 711–733.
- CURZAN, ANNE (2009): Historical corpus linguistics and evidence of language change. In: LÜDELING, ANKE / KYTÖ, MERJA (Hgg.): Corpus Linguistics. An International Handbook. Volume 2. Berlin. 1091–1108.
- DEPPERMANN, ARNULF (2007): Grammatik und Semantik aus gesprächsanalytischer Sicht. Berlin, New York.
- DIESSEL, HOLGER (2009): Corpus linguistics and first language acquisition. In: LÜDELING, ANKE / KYTÖ, MERJA (Hgg.): Corpus Linguistics. An International Handbook. Volume 2. Berlin. 1197–1212.
- ERNST, PETER (2002): Pragmalinguistik. Grundlagen, Anwendungen, Probleme. Berlin, New York.
- FAIRCLOUGH, NORMAN (1989): Language and power. London.
- FASOLD, RALPH (1990): Sociolinguistics of Language. Oxford.
- FELKE, HELMUTH (1996): Sprache als soziale Gestalt. Frankfurt am Main.
- FELDER, EKKEHARD (2003): Juristische Textarbeit im Spiegel der Öffentlichkeit. Berlin, New York.
- FELDER, EKKEHARD (2006): Form-Funktions-Analyse von Modalitätsaspekten zur Beschreibung von Geltungsansprüchen in politischen Reden. In: SCHERNER, MAXIMILIAN / ZIEGLER, ARNE (Hgg.): Angewandte Textlinguistik. Linguistische Perspektiven für den Deutsch- und Fremdsprachenunterricht. Tübingen. 157–178.
- FELDER, EKKEHARD (2009): Sprache – das Tor zur Welt!? Perspektiven und Tendenzen in sprachlichen Äußerungen. In: FELDER, EKKEHARD (Hg.): Sprache. Im Auftrag der Universitätsgesellschaft Heidelberg. Berlin u. a. 13–57
- FELDER, EKKEHARD / MÜLLER, MARCUS (Hgg.) (2009): Wissen durch Sprache. Theorie, Praxis und Erkenntnisinteresse des Forschungsnetzwerks „Sprache und Wissen“. Berlin, New York.
- FELDER, EKKEHARD / MÜLLER, MARCUS / VOGEL, FRIEDEMANN (Hgg.) (2010): Das Heidelberger Korpus: Gegenstand, Methoden und Zielsetzungen themenbezogener Korpora. In: Zeitschrift für Germanistische Linguistik (ZGL) 38 (2010). 314–319.
- FELLBAUM, CHRISTIANE / DELFS, LAUREN / WOLFF, SUSANNE / PALMER, MARTHA (2005): Word Meaning in Dictionaries, Corpora, and the Speaker's Mind. In: BARNBROOK, GEOFF / DANIELSSON, PERNILLA / MAHLBERG, MICHAELA (Hgg.): Meaningful Texts: The Extraction of Semantic Information from Monolingual and Multilingual Corpora. Birmingham. 31–38.
- FIRTH, JOHN RUPERT (1957): Papers in Linguistics (1934–1951). London, New York, Toronto.
- FLOWERDEW, JOHN (Hg.) (2002): Academic Discourse. Harlow, UK.

- GARDT, ANDREAS (2007): Linguistisches Interpretieren. Konstruktivistische Theorie und realistische Praxis. In: HERMANN, FRITZ / HOLLY, WERNER (Hgg.): Linguistische Hermeneutik. Theorie und Praxis des Verstehens und Interpretierens. Tübingen. 263–280 (Reihe Germanistische Linguistik 272).
- GARDT, ANDREAS (2008): Kunst und Sprache. Beobachtungen anlässlich der *documenta 12*. In: BARSCH, ACHIM / SCHEUER, HELMUT / SCHULZ, GEORG-MICHAEL (Hgg.): Literatur – Kunst – Medien. Festschrift für Peter Seibert zum 60. Geburtstag. München. 201–224.
- GARZONE, GIULIANA / SANTULLI, FRANCESCA (2004): What can Corpus Linguistics do for Critical Discourse Analysis? In: PARTINGTON, ALAN / MORLWEY, JOHN / HAARMAN, LOUANN (Hgg.): *Corpora and Discourse*. Bern. 351–368.
- GUMPERZ, JOHN J. (1982): *Discourse strategies*. Cambridge.
- GÜNTHER, SUSANNE (2000): Vorwurfsaktivitäten in der Alltagsinteraktion. Grammatische, prosodische, rhetorisch-stilistische und interaktive Verfahren bei der Konstitution kommunikativer Muster und Gattungen. Tübingen.
- JAKOB, KARLHEINZ (1991): Maschine, mentales Modell, Metapher. Studien zur Semantik und Geschichte der Techniksprache. Tübingen.
- HARDT-MAUTNER, GERLINDE (1995): „Only Connect“. *Critical Discourse Analysis and corpus linguistics*. University of Lancaster (<http://ucrel.lancs.ac.uk/papers/techpaper/vol6.pdf>).
- HAUSENDORF, HEIKO (2000): Zugehörigkeit durch Sprache. Eine linguistische Studie am Beispiel der deutschen Wiedervereinigung. Tübingen.
- HERINGER, HANS-JÜRGEN (1999): Das höchste der Gefühle: empirische Studien zur distributiven Semantik. Tübingen.
- HUNDT, MARIANNE (Hg.) (2007): *Corpus linguistics and the web*. Amsterdam.
- HUNDT, MARIANNE (2008): Text Corpora. In: LÜDELING, ANKE / KYTÖ, MERJA (Hgg.): *Corpus Linguistics. An International Handbook*. Volume 1. Berlin. 168–186.
- HUNDT, MARKUS (1995): Modellbildung in der Wirtschaftssprache. Zur Geschichte der Institutionen- und Theoriefachsprachen der Wirtschaft. Tübingen.
- HOOVER, DAVID L. (2007): Corpus Stylistics, Stylometry, and the Styles of Henry James. In: *Style* 41.2 (2007). 174–203.
- IMO, WOLFGANG (2007): *Construction Grammar und Gesprochene-Sprache-Forschung: Konstruktionen mit zehn matrixsatzfähigen Verben im gesprochenen Deutsch*. Tübingen.
- JUCKER, ANDREAS / TAAVITSAAINEN, IRMA (Hgg.) (2010): *Historical Pragmatics*. Berlin, New York.
- JUNG, MATTHIAS (1994): *Öffentlichkeit und Sprachwandel: Zur Geschichte des Diskurses über die Atomenergie*. Opladen.
- KNOBLAUCH, HUBERT (1995): *Kommunikationskultur. Die kommunikative Konstruktion kultureller Kontexte*. Berlin, New York.
- KONERDING, KLAUS-PETER (1993): *Frames und lexikalisches Bedeutungswissen. Untersuchungen zur linguistischen Grundlegung einer Frametheorie und zu ihrer Anwendung in der Lexikographie*. Tübingen.
- KONERDING, KLAUS-PETER (2005): Diskurse, Themen und soziale Topik. In: FRAAS, CLAUDIA / KLEMM, MICHAEL (Hgg.): *Mediendiskurse*. Frankfurt am Main. 9–38.