

## Das juristische Referenzkorpus (JuReko) - Computergestützte Rechtslinguistik als empirischer Beitrag zu Gesetzgebung und Justiz

,  
isabelle.gauer@medienkultur.uni-freiburg.de  
Albert-Ludwigs-Universität Freiburg, Deutschland

,  
hamann@coll.mpg.de  
Max-Planck-Institut für Gemeinschaftsgüter, Deutschland

,  
friedemann.vogel@medienkultur.uni-freiburg.de  
Albert-Ludwigs-Universität Freiburg, Deutschland

Sprachwissenschaftler\_innen und Jurist\_innen haben gemein, dass sie mit Texten arbeiten. Der juristische Umgang mit Texten ist allerdings geprägt und überformt von den Verfassungsgeboten der Rechtssicherheit und Vorhersehbarkeit der Interpretation von Normtexten, die eine disziplinäre Standardisierung erfordern: „Im Gegensatz zur grundsätzlich nicht normierbaren Alltagssprache oder zur Offenheit literaturwissenschaftlicher Interpretationen ist die Sprache des Rechts auf weitestgehende Verbindlichkeit, Deutlichkeit und Disziplin (zumindest) angelegt“ (Jeand’Heur 1998: 1287). Juristische Fachtexte lassen sich deshalb nur mit einem stark spezialisierten fachsprachlichen Sach- und (impliziten) Methodenwissen adäquat verstehen (vgl. hierzu Vogel 2012b: 34ff.).

Das spezialisierte Fach(sprach)wissen in der Jurisprudenz hat mindestens drei Funktionen: Erstens soll es juristische Entscheidungsarbeit valide und zuverlässig organisieren; zweitens soll es die Komplexität der Lebenswelt auf ‚rechtsrelevante‘ und verfahrenssichere, also in juristischen Kategorien verarbeitbare, Ausschnitte reduzieren; drittens stiftet es binnendisziplinäre Identität (Ingroup): Wer die Sprache und die ‚Denke‘ der Jurisprudenz nicht beherrscht, hat vor Gericht schlechte Karten.

All diese in der Regel für Laien nicht erkennbaren Funktionen stehen hinter sog. „ Subsumtionen“, also der juristischen Auslegungsmethode. Damit ist kein rechtspositivistisches ‚Anwenden‘ eines objektiv oder subjektiv vorgegebenen ‚Gesetzesinhalts‘ gemeint. Die juristische „Auslegung“ von Normen ist vielmehr ein komplexer Prozess der Ko(n)textualisierung von

Lebenswelt (zu beurteilender Sachverhalt, „Fall“) und Textwelt (inter- und intratextuelle Verknüpfung von Norm- und dogmatischen Texten). Lebens- und Textwelt sind dabei nicht lediglich ‚gegeben‘ und „im Sinne eines kybernetischen Informationsübertragungsmodells“ im Hinblick auf ‚die‘ Norm zu „decodieren“ (so noch Baden 1977: 14ff.; vgl. dazu kritisch Busse 2005). Sie ‚geben‘ dem hermeneutisch tätigen Rechtsarbeiter vielmehr sinnlich wahrnehmbare Hinweisreize (Gumperz 1982: 131f.), die gemeinsam mit bereits bestehendem, institutionalisiertem juristischen Norm(sprach)wissen in mentalen Modellen Sinn-voll gemacht werden können (Hörmann 1980). Rechtsnormen sind also keine absoluten Entitäten, sondern Ergebnis konstruktiver Textarbeit mit unterschiedlichen versprachlichten Eingangsdaten und Geltungsansprüchen (Müller et al. 1997; Felder 2003).

Seit rund 30 Jahren widmet sich die Rechtslinguistik als gemeinsame Teildisziplin von Rechts- und Sprachwissenschaft diesen Vertextungsverfahren im Recht (vgl. Vogel 2016). Juristische und linguistische Untersuchungen erfolgten dabei bislang ausschließlich mittels qualitativer Zugänge und auf Basis weniger hundert Texte. Die Ergebnisse geben wichtige Einblicke in die Mikroprozesse unseres sprachbasierten Rechtssystems, sei es vor Gericht, in der Verwaltung oder in der Gesetzgebung (Überblick bei Felder / Vogel 2016). In der frühen Rechtskybernetik und heutigen Rechtsinformatik hingegen wird das Recht meist als logisch operierendes Ontologiesystem zu formalisieren versucht, das die semantische Struktur seiner realen performativen Bearbeitung jedoch vernachlässigt (vgl. zur Kritik am „Subsumtionsautomaten 2.0“ Kotsoglou 2014; Vogel 2015). Erst neuere Ansätze einer „evidenzbasierten Jurisprudenz“ (Hamann 2014) und rechtstheoretisch fundierten Korpuslinguistik in den USA (Mouritsen 2010, 2011) sowie in Deutschland (Vogel 2012a; Vogel et al. 2015; Hamann 2015) versprechen praxisnahe Analysen und Einsichten in die ‚Makroökonomik‘ juristischer Fachsprache und -kommunikation.

An dieser Stelle setzt ein seit 2014 laufendes und von der Heidelberger Akademie der Wissenschaften finanziertes Projekt zur Konzeption und Auswertung eines „Juristischen Referenzkorpus“ (JuReko) an (Vogel / Hamann 2015). Ziel des Projektes, das den Kern der „International Research Group Computer Assisted Legal Linguistics“ ( CAL<sup>2</sup> 2014-2016 ) bildet, ist im ersten Schritt der Aufbau eines kontrollierten, zunächst statischen Fachtext-Korpus, das alle wichtigen Textsorten aus Judikative, Legislative und Rechtswissenschaft umfasst (v.a. Aufsätze aus juristischen Fachzeitschriften, Entscheidungstexte und Normtexte; Zielgröße: rund eine Milliarde fortlaufender Wortformen). Die Textdaten werden zunächst im html-Format gewonnen und anschließend in mehreren Konvertierungsschritten TEI-P5-konform kodiert. Dafür kommen xsl-Transformationen zum Einsatz, die auf die unterschiedlichen Webseitenstrukturen angepasst werden. Im Anschluss werden die Texte mit Part-of-Speech und weiteren Annotationen und Metadaten angereichert, wobei

die speziellen Anforderungen einer rechtslinguistischen Textanalyse und -verarbeitung im Vordergrund stehen.

Das Korpus bildet im zweiten Schritt die Grundlage für die Erprobung neuer computerlinguistischer Methoden zur Analyse insbesondere juristischer Semantik bzw. Dogmatik sowie zur Beschreibung von Wortschätzen und grammatischen Mustern in verschiedenen Rechtsbereichen auf Basis geeigneter Metriken. In Zusammenarbeit mit Praktikern aus Gesetzgebung und Rechtsprechung werden weitere Untersuchungsprojekte abgeleitet und vorbereitet. Hierzu zählt etwa die Entwicklung von Werkzeugen für die rechtslinguistisch wie korpusstatistisch-empirisch fundierte Optimierung der Gesetzesredaktion.

Der Vortrag stellt das Infrastrukturvorhaben „JuReko“ vor und diskutiert Möglichkeiten und Grenzen des durch die Projektgruppe entwickelten Ansatzes der „Computergestützten Rechtslinguistik“ als komplementären Beitrag zur qualitativen, juristischen Hermeneutik. Dabei wird anhand von Beispielen sowohl auf die textlinguistischen als auch technischen Details des Projektes eingegangen. Im Ausblick steht die Erweiterung des JuReko um Rechtstexte des britischen Case Law als Ausgangspunkt für ein Europäisches Rechtskorpus (European Law Corpus) und damit eine weltweit einzigartige Grundlage für rechts(sprach)kulturvergleichende Studien.

## Bibliographie

**Baden, Eberhard** (1977): *Gesetzgebung und Gesetzesanwendung im Kommunikationsprozess*. Studien zur jur. Hermeneutik u. zur Gesetzgebungslehre. Baden-Baden: Nomos-Verlagsgesellschaft.

**Busse, Dietrich** (2005): „Ist die Anwendung von Rechtstexten ein Fall von Kommunikation? Rechtslinguistische Überlegungen zur Institutionalität der Arbeit mit Texten im Recht“, in: Lerch, Kent D. (ed.): *Die Sprache des Rechts*. Recht Vermitteln: Strukturen, Formen und Medien der Kommunikation im Recht. 3 Bände. Berlin: Walter De Gruyter 23–54.

**CAL<sup>2</sup>** (2014-2016): *International Research Group: Computer Assisted Legal Linguistics*. University of Freiburg <http://www.cal2.eu/> [letzter Zugriff 08. Januar 2016].

**Felder, Ekkehard** (2003): *Juristische Textarbeit im Spiegel der Öffentlichkeit*. Berlin / Boston: De Gruyter.

**Felder, Ekkehard / Vogel, Friedemann** (eds.) (2016): *Handbuch Sprache im Recht*. Berlin / Boston: De Gruyter Mouton

**Gumperz, John Joseph** (1982): *Discourse strategies*. Cambridge: University Press.

**Hamann, Hanjo** (2014): *Evidenzbasierte Jurisprudenz. Methoden empirischer Forschung und ihr Erkenntniswert für das Recht am Beispiel des Gesellschaftsrechts*. Tübingen: Mohr Siebeck.

**Hamann, Hanjo** (2015): „Der "Sprachgebrauch" im Waffenarsenal der Jurisprudenz. Die Rechtspraxis im

Spiegel der quantitativ-empirischen Sprachforschung“, in: Vogel, Friedemann (ed.): *Zugänge zur Rechtssemantik*. Interdisziplinäre Ansätze im Zeitalter der Mediatisierung zwischen Introspektion und Automaten. Berlin / New York: Walter De Gruyter 184-204.

**Hörmann, Hans** (1980): „Der Vorgang des Verstehens“, in: Kühlwein, Wolfgang (ed.): *Sprache und Verstehen*. Tübingen: Narr 17–29.

**Jeand'Heur, Bernd** (1998): „Die neuere Fachsprache der juristischen Wissenschaft seit der Mitte des 19. Jahrhunderts unter besonderer Berücksichtigung von Verfassungsrecht und Rechtsmethodik“, in: Hoffmann, Lothar / Burkhardt, Armin / Ungeheuer, Gerold / Wiegand, Herbert Ernst / Steger, Hugo / Brinker, Klaus (eds.): *Fachsprachen: ein internationales Handbuch der Fachsprachenforschung und Terminologiewissenschaft*. (= Handbücher zur Sprach- und Kommunikationswissenschaft 14.1). Berlin: De Gruyter 1286–1295.

**Kotsoglou, Kyriakos N.** (2014): „Subsumtionsautomat 2.0. Über die (Un-)Möglichkeit einer Algorithmisierung der Rechtserzeugung“, in: *Juristenzeitung* 69, 9 451–457.

**Mouritsen, Stephen C.** (2010): „The Dictionary Is Not a Fortress: Definitional Fallacies and a Corpus-Based Approach to Plain Meaning“, in: *Brigham Young University Law Review* 1915–1980 <http://www.lawreview.byu.edu/archives/2010/5/10Mouritsen.pdf> [letzter Zugriff 07. November 2012].

**Mouritsen, Stephen C.** (2011): „Hard Cases and Hard Data: Assessing Corpus Linguistics as an Empirical Path to Plain Meaning“, in: *The Columbia Science and Technology Law Review* 8: 156–205 <http://www.stlr.org/cite.cgi?volume=13&article=4> [letzter Zugriff 07. November 2012].

**Müller, Friedrich / Christensen, Ralph / Sokolowski, Michael** (1997): *Rechtstext und Textarbeit* (= Schriften zur Rechts-theorie). Berlin: Duncker & Humblot.

**Vogel, Friedemann** (2012a): „Das Recht im Text. Rechtssprachlicher Usus in korpuslinguistischer Perspektive“, in: Felder, Ekkehard / Müller, Marcus / Vogel, Friedemann (eds.): *Korpuspragmatik*. Thematische Korpora als Basis diskurslinguistischer Analysen. Berlin / Boston: De Gruyter 314–353.

**Vogel, Friedemann** (2012b): *Linguistik rechtlicher Normgenese*. Theorie der Rechtsnormdiskursivität am Beispiel der Online-Durchsuchung (= Sprache und Wissen 9). Berlin / Boston: De Gruyter.

**Vogel, Friedemann** (2015): „Zwischen Willkür, Konvention und Automaten: Die interdisziplinäre Suche nach Bedeutungen in Recht und Gesetz“, in: Vogel, Friedemann (ed.): *Zugänge zur Rechtssemantik*. Interdisziplinäre Ansätze im Zeitalter der Mediatisierung zwischen Introspektion und Automaten. Berlin / New York: Walter De Gruyter.

**Vogel, Friedemann** (2016): „Rechtslinguistik: Zur Bestimmung einer Fachrichtung“, in: Felder, Ekkehard / Vogel, Friedemann (eds.): *Handbuch Sprache im Recht* (=

Handbücher Sprachwissen 12). Berlin / Boston: De Gruyter Mouton.

**Vogel, Friedemann / Christensen, Ralph / Pötters, Stephan** (2015): *Richterrecht der Arbeit – empirisch untersucht. Möglichkeiten und Grenzen computergestützter Textanalyse am Beispiel des Arbeitnehmerbegriffs*. Berlin: Duncker & Humblot.

**Vogel, Friedemann / Hamann, Hanjo** (2015): „Vom corpus iuris zu den corpora iurum – Konzeption und Erschließung eines juristischen Referenzkorpus (JuReko)“, in: *Jahrbuch der Heidelberger Akademie der Wissenschaften für 2014*. Heidelberg: Winter.