

# „SO ETWAS WIE EINE BOTSCHAFT“: KORPUSLINGUISTISCHE ANALYSEN DER BUNDESTAGSWAHL 2009

von Noah Bubenhofer, Tobias Dussa, Sarah Ebling, Martin Klimke, Klaus  
Rothenhäusler, Joachim Scharloth, Suarès Tamekue, Saskia Vola  
(Forschergruppe semtracks)

„Yes, wie gäh“ sei die Formel, mit der man den Wahlkampf der Bundestagswahl 2009 zusammenfassen müsse – meinte zumindest die Bild-Zeitung nach dem TV-Duell zwischen Angela Merkel und Frank-Walter Steinmeier. Doch auch wenn der Wahlkampf lau war, wurde er dennoch von einer großen Zahl von Politologen, Rhetorikanalystikerinnen, Kommunikationsprofis und Journalistinnen begleitet und kommentiert. Der Hunger nach Expertise ist nach der Wahl nun wohl fürs Erste gestillt. Doch spätestens beim nächsten Wahlkampf werden die Analysen wieder gefragt sein.

Doch wer gehört zur Zunft der Analytiker? Für die semtracks-Forschergruppe war die Bundestagswahl eine weitere Gelegenheit, mit einem linguistischen Blick die Wahlen zu begleiten. Bisher sind Sprachwissenschaftlerinnen und -wissenschaftler eher selten in den medialen Expertenrunden vertreten, obwohl in der Linguistik in vielen Subdisziplinen eine lange Tradition der politischen Analyse existiert. Doch werden diese Arbeiten hauptsächlich im wissenschaftlichen Kontext rezipiert und im medialen Alltag kaum berücksichtigt.

Das liegt einerseits an den unterschiedlichen Diskursregeln, die in Wissenschaft und Medienwelt herrschen: Das Leitmotiv der Wissenschaft ist Differenzierung, jenes der Medien Pointierung. Es liegt aber auch an den unterschiedlichen Formaten: Ein 200-seitiges Buch, das die Früchte einer mehrjährigen Forschungsarbeit über rhetorische Muster in Wahlkampfreden darstellt, mag wissenschaftlich relevant, wertvoll und angemessen sein, es ist für mediale Bedürfnisse aber schlicht zu dick und bereits veraltet. Gefragt sind Analysen zur gestrigen Rede und zum Parteiprogramm, das am letzten Samstag verabschiedet wurde.

## Der semtracks-Political-Tracker

Im Rahmen eines Forschungsprojekts am Heidelberg Center for American Studies, gefördert durch den In-

novationsfonds FRONTIER der Universität Heidelberg, hat die Forschergruppe semtracks bereits bei den US-Wahlen 2008 (Bubenhofer et al. 2008a, b) und nun bei den vergangenen Wahlen in Deutschland (Bubenhofer et al. 2009) linguistische Methoden für aktuelle Analysen zu nutzen versucht. Das Ziel dieses „Political Trackers“ war, aktuelle Kurzanalysen zu den Parteien und politischen Akteuren der Bundestagswahl 2009 online zu publizieren und damit auch die Medien anzusprechen. Die Analysen sollten Aussagen über

stilistische und rhetorische Besonderheiten ermöglichen. Die Basis hierfür waren die für die einzelnen Politiker/innen und Parteien typischen linguistischen Profile, die stilistische, semantische und syntaktische Phänomene umfassten. Um die geforderte Aktualität zu gewährleisten, boten sich korpuslinguistische und maschinelle Methoden der Sprachanalyse an.

## Korpus

Die Untersuchungskorpora müssen primäre Äußerungen der politischen Akteure umfassen, um die in den Korpora festgelegten linguistischen Profile der politischen Akteure berechnen zu können. Dafür bieten sich folgende Quellen an:

- Redemanuskripte bzw. selbst erstellte Transkripte von Reden, bei denen keine Manuskripte verfügbar sind
- Pressemitteilungen der Parteien
- Parteiprogramme

Diese Quellen können über die Websites der Parteien oder Kandidat/innen leicht beschafft werden. Bei größeren Mengen (z. B. Pressemitteilungen) gelangen automatisierte Verfahren des Downloads zum Einsatz.

Kern der Analyse sind stets Vergleiche zwischen verschiedenen Teilkorpora: Ein Teilkorpus kann bestehen aus:

- allen bereits gehaltenen Reden eines Akteurs
- einer bestimmten Rede
- allen Pressemitteilungen einer Partei
- einem oder mehreren Programmen einer Partei

Das Korpus umfasst damit auch Dokumente, deren Entstehung bereits einige Jahre zurückliegt, um Vergleiche sowohl auf synchroner als auch auf diachroner Ebene anstellen zu können.

## Methoden der semtracks-Matrix-Analyse

Für die maschinelle Korpusanalyse verwenden wir ein Bündel von Methoden, die wir unter dem Titel „semtracks-Matrix-Analysen“ zusammenfassen. Die Methoden fußen auf theoretischen Überlegungen zur Sprachgebrauchsanalyse, die in Bubenhofer (2009) dargelegt sind. Das Korpus ist mit Wortarten-Informationen annotiert und lemmatisiert.

## Einzellexeme

Der einfachste Zugang zu linguistischen Spezifika der Teilkorpora bietet die Berechnung von typischen Lexemen. Es werden dabei für zwei Teilkorpora (z. B. Korpus A: alle Reden von Angela Merkel, Korpus B: alle Reden von Frank-Walter Steinmeier) je Frequenzlisten aller vorkommenden Lexeme berechnet, gesondert nach Wortarten. Anschließend werden pro Lexem die Frequenzen in den beiden Teilkorpora verglichen und die Signifikanz des Frequenzunterschieds mittels eines Log-Likelihood-Tests (Manning/Schütze 2002, S. 174) berechnet. Resultat sind nach Signifikanzwert geordnete Listen von Lexemen, die jeweils für das eine Teilkorpus im Vergleich zum anderen Teilkorpus typisch sind.

In einem weiteren Schritt werden die typischen Lexeme in sog. Wortwolken visualisiert, wobei die Typik des Lexems durch die Größe des abgebildeten Lexems dargestellt wird. Das erlaubt einen raschen Überblick über die lexikalische Spezifik der Teilkorpora. Wortwolken der typischen Nomen zeigen dabei meist inhaltliche Präferenzen der Sprecher und können verwendet werden, um die typischen Schlagworte von Parteiprogrammen oder Reden herauszuarbeiten.

Interessanter sind oft aber auch Wortwolken anderer Wortarten. Abbildung 1 (auf Seite 4) zeigt z. B. die typischsten Adjektive von Angela Merkel und Frank-Walter Steinmeier in jenen Reden, die sie als Kanzlerin bzw. Außenminister gehalten haben (Merkel: 278 Reden, 768.549 Wortformen; Steinmeier: 190 Reden, 330.084 Wortformen). Während Steinmeiers Reden gespickt sind mit Adjektiven, die entsprechend seiner Funktion als Außenminister geopolitisch motiviert sind (*afghanisch, afrikanisch, auswärtig, außenpolitisch, bilateral, deutsch-französisch, deutsch-polnisch, europäisch, humanitär, international, israelisch, palästinensisch, militärisch, nuklear, regional, politisch, sicher*), entstammen die Adjektive, die typisch für die Reden Angela Merkels sind, anderen Wortfeldern. Es dominieren wertende (*gut, interessant, richtig, spannend, vernünftig, wunderbar, schön*) und intensivierende Adjektive (*außerordentlich, allergrößt, bestimmt, vollkommen, völlig, wirklich, riesig, unendlich, dramatisch*). Ein genauerer Blick in die Wortlisten zeigt zudem, dass bei Merkel insbesondere die intensivierenden Adjektive *vollkommen, unglaublich, allergrößt* sehr viel häufiger vorkommen als bei Steinmeier. Interessant ist auch, dass insbesondere die umgangssprachlichen Ausdrücke *toll* und *spannend* und weitere positive Gefühle zum Ausdruck bringende Adjektive wie *wunderschön* und *wunderbar* fast ausschließlich bei Merkel vorkommen.

Wortwolke: Steinmeier

afrikanisch aktiv aktuell amerikanisch  
arabisch asiatisch **auswärtig**  
außenpolitisch bilateral dauerhaft  
demokratisch **deutsch** direkt dringend eng  
entscheidend erfolgreich ernst europäisch  
französisch friedlich geehrt gegenseitig  
**gemeinsam** genau global **heutig**  
häufig international kalt klug knapp  
**kommend konkret** konstruktiv kritisch  
**kulturell** kurz künftig langfristig militärisch  
multilateral mutig nachhaltig nah **neu** offen  
**palästinensisch politisch** polnisch  
positiv recht **regional** russisch schwierig  
**sicher** strategisch technologisch tief  
transatlantisch täglich unmittelbar  
verantwortlich **verehrt** vereint **vergangen**  
wachsend weit weltweit westlich wirtschaftlich  
zahlreich zentral **zivil** zunehmend öffentlich

Abbildung 1: Typische Adjektive in den Reden von Frank-Walter Steinmeier und Angela Merkel

Es finden sich weitere Anzeichen dafür, dass die Sprache Angela Merkels emotionaler ist als die von Steinmeier. So benutzt sie bei Adressierungen häufiger die Form *liebe/r Herr/Frau* und in Dankesformeln das Adjektiv *herzlich*. Steinmeier hingegen benutzt bei Adressierungen bevorzugt die distanzierenden Vokabeln *gehrt* oder *verehrt*, letzteres rund 10 mal häufiger als Angela Merkel. Bezieht man auch Verben in die Analyse mit ein, dann zeigt sich, dass Angela Merkel häufiger das Verb *fühlen* benutzt. Steinmeier hingegen benutzt Gefühlswortschatz meist nur desemantisiert im Kontext ritualisierter Phrasen wie *ich freue mich, dass*. Eine weitere Besonderheit Merkel'scher Reden sind Heckenausdrücke mit einschränkender Funktion, insbesondere Lexeme wie *relativ*, *ungefähr* oder *wahrscheinlich*.

## Komplexe n-Gramme

Der Fokus auf Einzellexeme ist vergleichsweise primitiv und lässt die Kontextspezifik der Semantik von Lexemen außer Acht. Das gleiche Verfahren der Berechnung von für ein Teilkorpus typischen Lexemen kann nun auf die Berechnung von typischen Mehr-

Wortwolke: Merkel

**allergrößt** ander anschließend ausreichend  
außerordentlich **bestimmt** christlich  
dankbar deutlich dramatisch ehrlich  
**einfach** einzeln entsprechend erheblich  
fest froh **ganz** geistig gering **gesamt**  
gleich **gut** hart **herzlich** hoch  
**interessant** jeweilig jung klein lieb  
mittelständisch plötzlich privat **relativ**  
**richtig** riesig römisch schnell schwer schön  
sozial spannend stolz technisch  
**unglaublich** unterschiedlich vereinigt  
**vernünftig** verschieden  
**vollkommen** vorhanden **völlig**  
wahrscheinlich wesentlich wichtig **wirklich**  
wunderbar zufrieden

worteinheiten (n-Grammen) ausgeweitet werden (vgl. Bubenhofer 2009, S. 149ff.). Dabei werden alle im Korpus vorkommenden n-Gramme berechnet, wobei bestimmte Parameter wie Länge der Mehrworteinheit (aus zwei, drei oder mehr Wörtern bestehend) oder Spannweite (sind Lücken zwischen den Wörtern erlaubt?) festgelegt werden.

Normalerweise werden n-Gramme als Folge von Wortformen verstanden. Interessanter ist es, nicht nur die Wortformen, sondern auch das Lemma, die Wortart oder andere annotierte Elemente als weitere Dimensionen mit einzubeziehen. Ein solches komplexes n-Gramm kann dann aus einer beliebigen Kombination aller dieser Elemente bestehen. In der Wortfolge "ich glaube, dass" mit den entsprechenden Wortart- und Lemmainformationen „ICH, Personalpronomen (PPER) / GLAUBEN, finites Verb (VV-FIN) / [Komma] / DASS, unterordnende Konjunktion (KOUS)“ stecken demnach Kombinationen der Art:

- Ich glaube , dass
- ICH GLAUBEN , DASS
- PPER glaube , dass

- PPER GLAUBEN, dass
- Ich VVFIN, dass
- Ich glaube, KOUS
- PPER VVFIN, dass
- ...

Es ergeben sich also bei drei Dimensionen  $3^4=81$  Kombinationsmöglichkeiten; bezieht man weitere Annotationen mit ein, entstehen sehr große Datenmengen, die sich jedoch stark reduzieren, wenn nur die für das jeweilige Teilkorpus signifikanten n-Gramme berücksichtigt werden. Zusätzlich kann es sinnvoll sein, auf einer Dimension nur gewisse Elemente zu berücksichtigen; bei den folgenden Analysen wurden auf der Dimension der Wortformen nur Funktionswörter berücksichtigt (auf der Ebene der Lemmata gab es hingegen keine Restriktionen).

Bei der Analyse der Reden von Angela Merkel ergeben sich im Kontrast zu Frank-Walter Steinmeier typische Muster der Art “PPER VVFIN, dass“, realisiert durch Phrasen wie:

Beispiele (1):

*Wir sehen, dass*  
*Sie wissen, dass*  
*Wir wissen, dass*  
*Ich weiß, dass*  
*Ich sage, dass*

Beispiele (2):

*Ich glaube, dass*  
*Ich hoffe, dass*  
*Ich vermute, dass*  
*Ich denke, dass*

Es handelt sich also um Hauptsätze in der ersten Person Singular oder Plural (ich/wir), denen sich ein Komplementsatz anschließt, der mit „dass“ eingeleitet wird. Die Verben des Hauptsatzes haben dabei (1) teilweise konstativen bzw. faktiven Charakter (2), teilweise subjektivierenden bzw. relativierenden Charakter. Weitere Varianten des Musters haben ebenfalls diese Wirklichkeitskonstituierende Funktion, die z. B. auch den Zuhörer mit einem kollektiven „wir“ mit einbezieht:

*Wir erleben, dass*  
*Wir wissen, dass*  
*Wir spüren, dass*  
*Wir sehen, dass*

Ebenfalls auffällig sind adversative Satzanschlüsse (*Aber ich finde/glaube/vermute; Aber wir wissen/glauben etc.*), die auf die Auseinandersetzung mit einer

im vorangehenden Satz genannten Gegenposition reagiert, die entweder relativiert oder korrigiert wird. In Verbindung mit faktiven Verben konstruieren sie eine andere Wirklichkeit als die in der Vorrede entworfene, in Verbindung mit nicht-faktiven stellen sie sie zumindest in Aussicht.

Typisch für argumentative Kontexte sind auch konsekutive Satzanschlüsse, die sich in Mustern zeigen wie *Deshalb werden/haben wir jetzt/hier, Deshalb ist NOMINALPHRASE, Deshalb ist es so/leben/lauch/ganz etc.* Die genannten Muster kommen signifikant häufiger bei Merkel als bei Steinmeier vor. Die Formulierung konsekutiver Relationen und damit die Inszenierung eines argumentativen Politikstils sind also ein sehr ausgeprägtes Merkmal der Merkel’schen Reden. Dies sagt freilich noch nichts über die Qualität der vorgebrachten Argumente und damit der Gültigkeit ihrer Folgerungen aus, denn nicht immer korrespondiert die sprachliche Inszenierung mit den Inhalten.

Signifikant unterscheiden sich die Reden von Merkel und Steinmeier auch in der Verwendung von deiktischen Ausdrücken. Angela Merkel verwendet zur Betonung einzelner Passagen oder Punkte signifikant häufiger deiktische Ausdrücke in Kombination mit Modalverben. Dies hat die Funktion, einzelne Aspekte der Rede hervorzuheben:

*Ich möchte an diesem Tage*  
*Ich möchte an dieser Stelle*  
*Ich möchte an diesem Abend*  
*Wir müssen an dieser Stelle*  
*Ich darf an dieser Stelle*

Schließlich findet sich auch ein Formulierungsmuster signifikant häufiger bei der Kanzlerin, das Vagheit indiziert, nämlich die Wendung *so etwas wie*:

*so etwas wie eine Botschaft*  
*so etwas wie ein Wahrzeichen*  
*so etwas wie eine Richtschnur*  
*so etwas wie ein Gemeinschaftsgefühl*  
*so etwas wie eine Brücke*  
*so etwas wie eine Krise*

Frank-Walter Steinmeier tendiert in seinen Reden hingegen zum Nominalstil, z. B. mit Aneinanderreihungen von Nominalphrasen im Genitiv, die sich bei ihm sieben mal häufiger als bei Angela Merkel finden:

*Folge des Zerfalls der Sowjetunion*  
*Tagen der Entscheidung der EU*  
*Autorität des Gouverneursrates der IAEO*  
*Erklärung des Botschafters des Irans*  
*Klärung des Status des Kosovo*



Auch die Aneinanderreihung von Präpositionalgruppen ist sonst ein typisches Phänomen von fachsprachlichen und wissenschaftlichen Texten:

*angesichts der Tätergruppen von New York  
wegen des Informationsaustausches mit Großbritan-  
nien  
in den Erklärungen von Berlin  
aufgrund der Gespräche über Atomwaffen  
im Kampf gegen den internationalen Terrorismus  
in den letzten Jahren seit der deutschen*

Besonders häufig (über fünfzig mal häufiger als Merkel) benutzt Steinmeier Phrasen mit der Präposition *für* in Kombination mit Genitivattributen:

*für die Arbeit der Sicherheitsbehörden  
für die Wiederaufnahme der Verhandlungen  
für die Bekämpfung der Intoleranz  
für die Sicherheit der Menschen  
für den Wiederbeginn der Gespräche*

Steinmeier gibt sich in seinen Reden zudem als Meister der Aufzählung zu erkennen. Er benutzt Paar- und Drillingsformen in allen Varianten:

– als Paarformen von Nomen (drei mal häufiger als Merkel):

*ein enger Freund und Verbündeter  
den Vereinigten Staaten und Russland  
die spezifischen Interessen und Anliegen  
den amerikanischen Freunden und Partnern  
den kritischen Fragen und Themen*

– als Paarformen von Adjektiven (drei mal häufiger als Merkel):

*wirtschaftliches und kulturelles Potential  
sicherheitspolitischer und wirtschaftlicher Hin-  
sicht  
konstruktiven und zukunftsgerichteten Kommuni-  
kation  
gesellschaftlichem und kulturellem Austausch  
technologischen und kulturellen Leistungen*

– als Paarformen von Adjektiv-Nomen-Verbindungen (3,5 mal häufiger als Merkel):

*sozialer Marktwirtschaft und nationaler Selbstbe-  
stimmung  
andere Nachbarn und regionale Großmächte  
kulturelle Identität und kulturelle Differenz  
absolute Sicherheit und objektive Garantien  
bedeutende Verlage und etablierte Schriftsteller  
gemeinsame Gegenwart und gemeinsame Zukunft*

– als Paarformen von Verben, die im Skopus eines Modalverbs in Letztstellung vorkommen (zwei mal häufiger als Merkel)

*EU-Ratspräsidentschaft suchen und ausbauen  
wollen  
Flughafens helfen und verhindern sollen  
Austausch ausbauen und intensivieren wollen  
Kontinent zeigen und teilen wollten  
Welt kämpfen und werben möchten  
Dinge herstellen und beschreiben kann*

Wo Paarformen sind, sind auch Drillingsformen nicht weit: Auch die Aufzählungen dreier Nomen (1,8 mal häufiger) oder Adjektive (drei mal häufiger) ist ein signifikantes Merkmal der Steinmeier'schen Rhetorik.

*Sicherheit, Stabilität und Stärkung  
Verhalten, Augenmaß und Vernunft  
Gewicht, Stimme und Handlungsfähigkeit  
Fragen, Frieden und Integration  
Demokratie, Menschenrechte und Rechtsstaatlich-  
keit  
politischer, sicherheitspolitischer und wirtschaft-  
licher  
wirtschaftlichem, gesellschaftlichem und kultu-  
rellem  
diplomatischer, nachrichtendienstlicher und bun-  
despolizeilicher  
offenen, konstruktiven und zukunftsgerichteten  
wissenschaftlichen, technologischen und kultu-  
rellen*

Ein weiteres Charakteristikum in Steinmeiers Reden sind Satzanschlüsse mit *und* in Varianten mit Personalpronomen, finitem Hilfsverb und Adjektiv (*Und ich bin sicher/froh/zuversichtlich*), mit Sprachhandlungs-  
verben (*Und ich wünsche Ihnen/sage es/verspreche Ihnen*) oder Modalverben (*Und wir müssen/sollten/wollen etc.*).

Während bei Angela Merkel sprachliche Mittel der Betonung und Emphase mittels deiktischer Ausdrücke (*Ich möchte an dieser Stelle ...*) signifikant häufiger auftreten, zeichnen sich Frank-Walter Steinmeiers Reden durch eine höhere Frequenz von redestrukturierenden Elementen aus. Insbesondere Formulierungen, in denen er sich über den folgenden Inhalt der Rede äußert, wie *Lassen Sie mich ...* und *Erlauben Sie mir*, finden sich deutlich häufiger als bei Angela Merkel (neun mal häufiger).

Auch die Kombination von Modalverb und Sprachhandlungsverb in Verbindung mit *ich* findet sich deutlich (acht mal) häufiger in den Reden Steinmeiers:

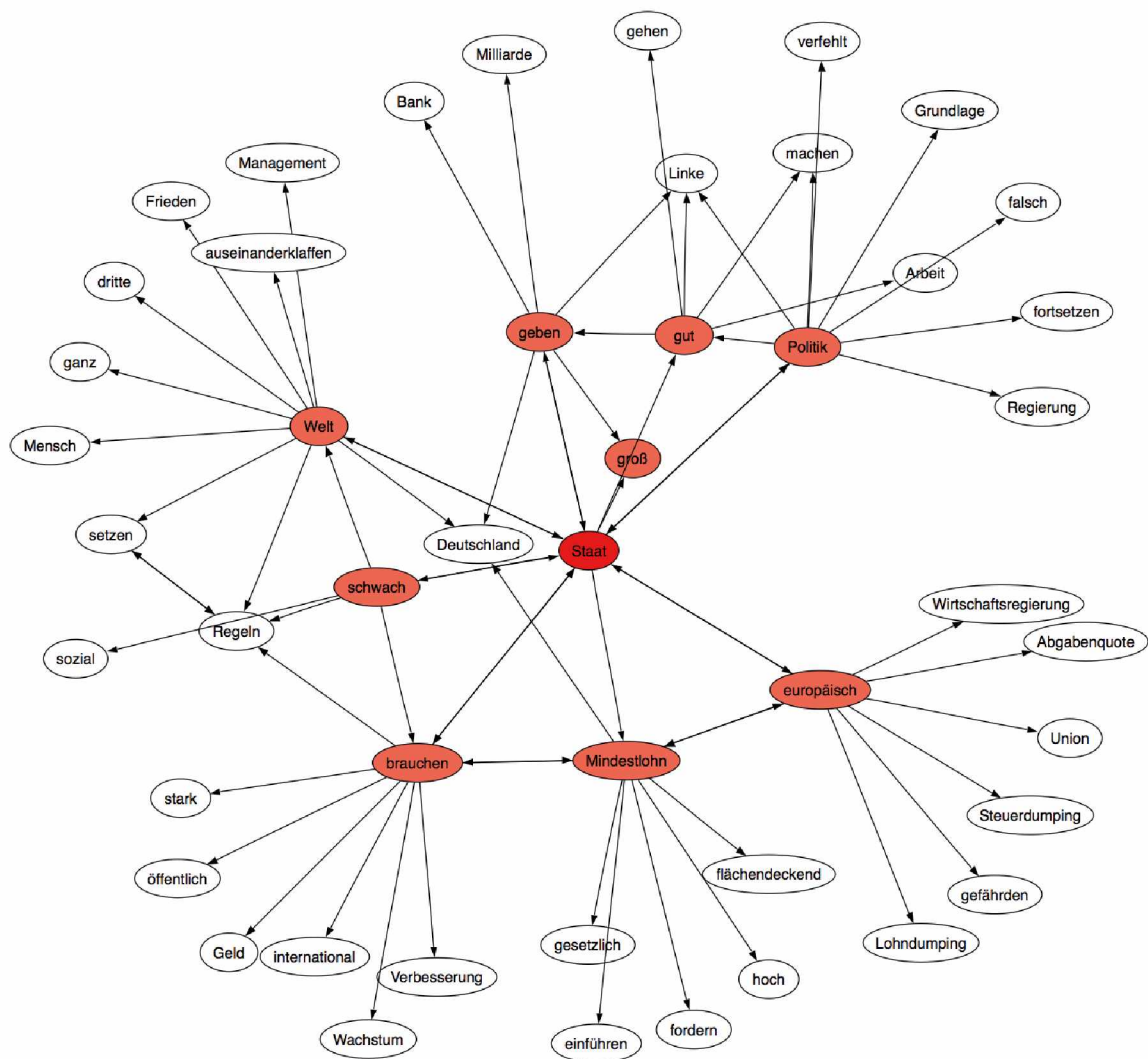


Abbildung 2: Kollokatoren zu „Staat“ bei Oskar Lafontaine

*wesentliche Gefahrenpunkte will ich erwähnen  
 zweiten Aspekt möchte ich hervorheben  
 gesamten Hause möchte ich danken  
 zweiten Aspekt möchte ich betonen*

Insgesamt zeigen die Analysen wesentliche Unterschiede in der politischen Rhetorik der beiden Kandidaten. Angela Merkel spricht mit mehr emotionaler Emphase, aber zugleich argumentativer als Frank-Walter Steinmeier. Sie ist zudem stärker darum bemüht, ihre persönliche Weltsicht zur Geltung zu bringen. Frank-Walter Steinmeier hingegen erweckt in seinen von Nominalstil geprägten, stärker strukturierten Reden eher den Eindruck des kompetenten Experten, der auf zahlreiche feststehende Wendungen (vor allem Paar- und Drillingsformen) zurückgreift.

## Kollokationen

Eine klassische Möglichkeit, typische Kontexte von bestimmten Lexemen systematisch auszuwerten, ist

die Berechnung von Kollokatoren. Mit statistischen Mitteln (hier: Log-Likelihood-Test) werden die Frequenzen ausgewertet, mit denen Lexeme in der definierten Umgebung eines spezifischen Lexems erscheinen. Resultat ist eine Liste von Kollokatoren, die signifikant häufig zusammen mit dem definierten Lexem auftreten (Manning/Schütze 2002, S. 172ff., Bubenhofer 2009, S. 131ff.; vgl. für eine elaborierte Art der Berechnung von Kollokationen auch Belica 2001).

Für die semtracks-Analysen eignet sich diese Methode, um die unterschiedliche Verwendung von bestimmten Begriffen darzustellen. Die Abbildungen 2 und 3 zeigen die primären und sekundären Kollokatoren zu *Staat* in den Reden von Oskar Lafontaine und Guido Westerwelle, die während des Wahlkampfes als Oppositionsführer der Parteien „Die Linke“ bzw. FDP agierten.

Wie die Kollokationsanalyse zum Lemma *Staat* in den Reden Oskar Lafontaines zeigt, referiert der Par-

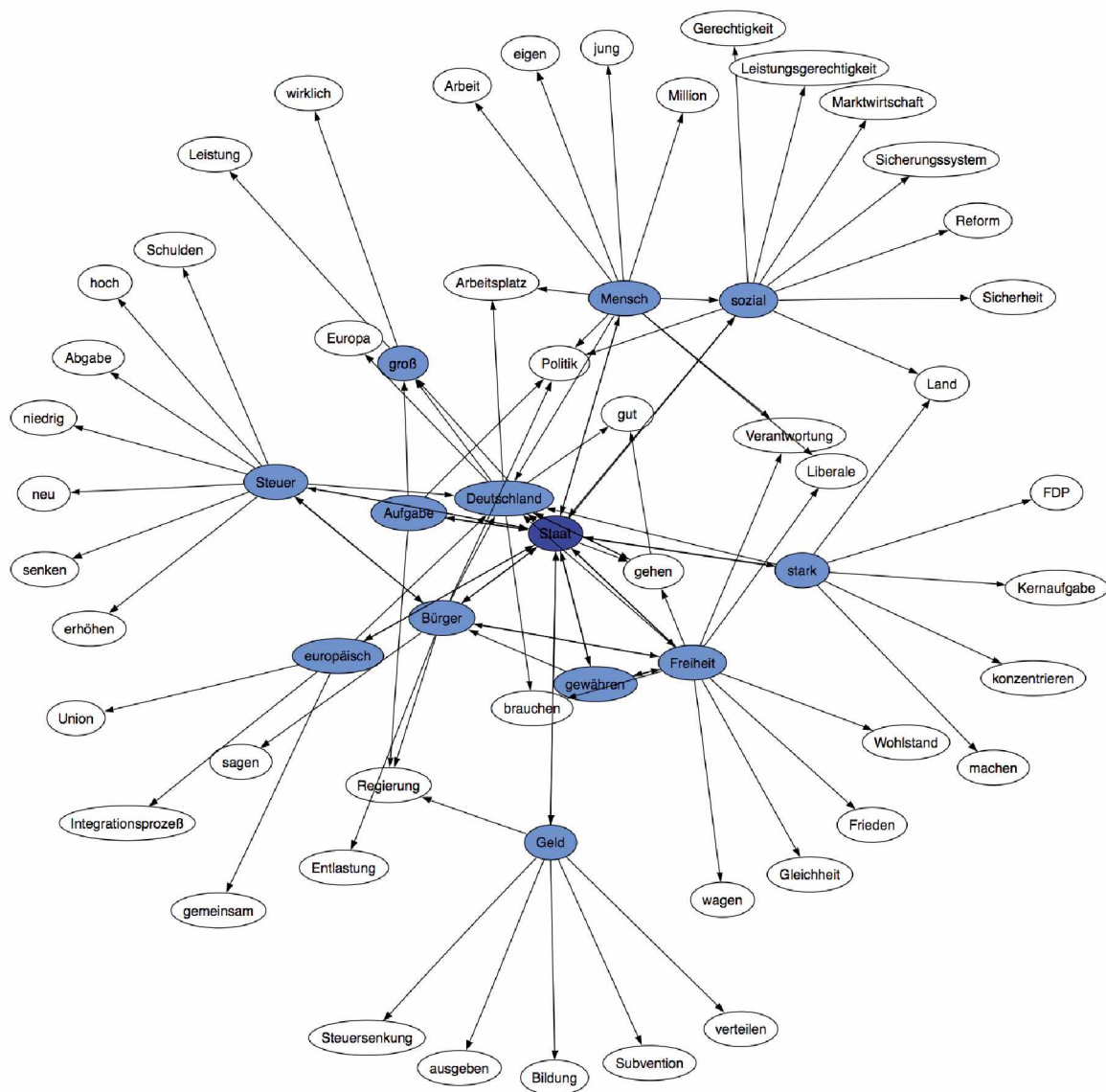


Abbildung 3: Kollokatoren zu „Staat“ bei Guido Westerwelle

teivorsitzende der Partei „Die Linke“ häufig auf die *europäischen Staaten*, aber auch auf die *schwachen Staaten* der Welt. In den Staaten der *europäischen Union* macht er vor allem die Gefahr von *Steuer- und Lohndumping* aus, hofft jedoch zugleich auf die Einführung eines *europäischen Mindestlohns*. Angesichts des *Auseinanderklaffens* der sozialen Verhältnisse brauchen die *schwachen Staaten*, aber auch die *sozial Schwachen Regeln*, die den Mächtigen Grenzen setzen. Nicht weniger als elf mal verwendet Lafontaine in seinen Reden ein Zitat von Montesquieu in verschiedenen Abwandlungen, das diese Lesart des Kollokationendiagramms stützt: „Zwischen dem Starken und dem Schwachen befreit das Gesetz, während die Freiheit unterdrückt“. Als Vorsitzender einer Oppositionspartei, aber auch weil der Staat den *Banken Milliarden gibt*, hält Lafontaine die Politik der Bundesregierung für *falsch* und *verfehlt* und nicht ganz überraschend die der eigenen Partei für *gut*.

Während der Staat für Lafontaine also Regeln setzen soll, um die Schwachen zu schützen, ist für den Vorsitzenden der FDP *Freiheit* eine entscheidende Kategorie, wenn es um die Aufgaben des Staates geht.

Wenn Guido Westerwelle über den Staat spricht, dann philosophiert er häufiger als Oskar Lafontaine über den Staat als abstrakte Ordnungsmacht: Der Staat müsse die *Freiheit* der *Bürger* *gewähren*. Freiheit ist ohne Verantwortung der *Menschen* nicht denkbar. Der *starke Staat* müsse sich auf seine *Kernaufgaben* konzentrieren. Zugleich steht der Staat für Westerwelle aber auch für *soziale Sicherheit*, sichert die *soziale Marktwirtschaft* und achtet auf *soziale Gerechtigkeit*, die freilich als *Leistungsgerechtigkeit* bestimmt wird. Dass das Wort *Staat* bei Westerwelle häufig mit dem Wort *Steuer* und *Geld* verknüpft ist, belegt, wie zentral wirtschafts- und finanzpolitische Kategorien für die FDP sind.



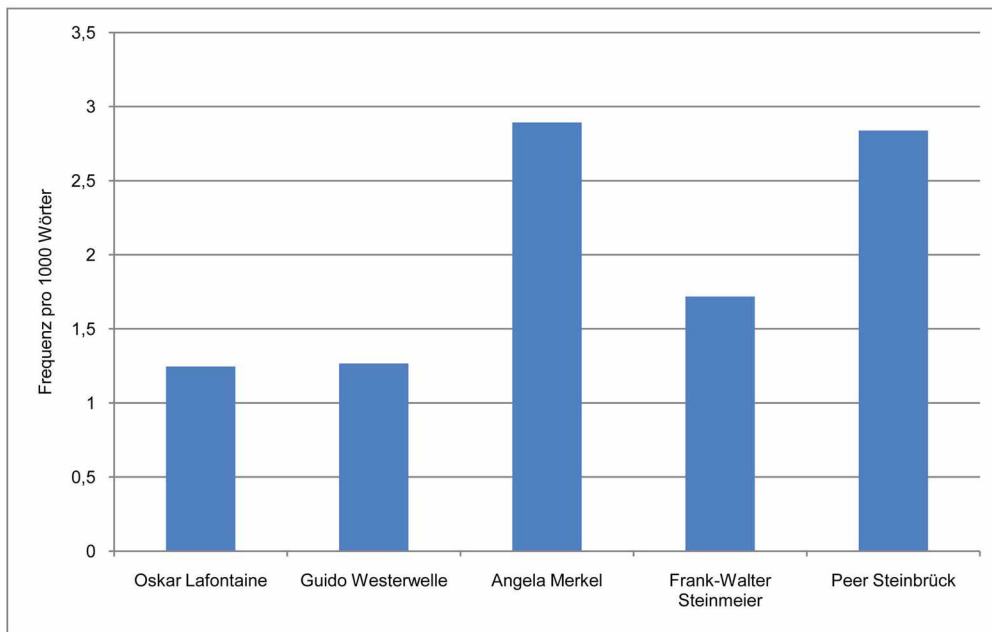


Abbildung 4: Gebrauch von Intensivierern aus dem absoluten und extrem hohen Intensivierungsbereich

### Semantische Subklassen

Als vierte methodische Herangehensweise, um rhetorisch-stilistische Unterschiede aufzudecken, bewährt sich die Suche nach zu semantischen Kategorien zusammengefassten Lexemen. Ein Beispiel ist die Kategorie der Intensivierer (*völlig, möglichst, kaum* etc.), die wiederum in unterschiedliche Klassen eingeteilt werden (vgl. Os 1989, Biedermann 1969, Bierwisch 1987). Besonders interessant sind der „absolute“ und der „extrem hohe Intensivierungsbereich“. Zum absoluten Intensivierungsbereich zählen unter anderem Ausdrücke wie:

*absolut, gänzlich, grundlegend, gründlich, im geringsten, komplett, längst, rein, restlos, schlechterdings, total, überhaupt, unbedingt, voll, vollständig, von Grund auf, durchweg, fundamental, grundsätzlich* etc.

Zum extrem hohen Intensivierungsbereich zählen dann vor allem Wörter, die die Superlativform aufweisen, und Adjektive, die von sich aus einen hohen Grad der Intensivierung beinhalten:

*höchst, äußerst, zutiefst, aufs höchste, (nicht) im geringsten, im höchsten Maße, allerbest, möglichst, schärfstens, sehnlichst, hervorragend, großartig, wunderbar, ungemain, irrsinnig, idiotisch, unheimlich, riesig, kolossal, sagenhaft, fabelhaft, traumhaft, wunderbar, zauberhaft, schrecklich, unvorstellbar, unsäglich, unbeschreiblich* etc.

Intensivierer kodieren Emotionen und den Grad von Überzeugungen, bzw. der Rigorosität, mit der sie vertreten werden.

Analysiert man das Auftreten von Intensivierern in den Reden wichtiger Akteure des Wahlkampfs 2009, ergibt sich folgendes Bild (vgl. Abbildung 4): Angela Merkel verwendet mit Abstand am meisten Intensivierer aus dem extrem hohen bzw. absoluten Bereich. Das steht im Einklang mit den für sie typischen Mustern, die Emphase und Emotionalität ausdrücken (z. B. die Verwendung intensivierender Adjektive). Guido Westerwelle und Oskar Lafontaine, die bis zu den

Wahlen beide Oppositionsführer waren, zeigen keinen auffälligen Gebrauch dieser Intensivierer.

Neben der Frequenz von Intensivierern ist die Untersuchung einer Vielzahl weiterer Wortklassen nützlich, um rhetorisch-stilistische Eigenheiten aufzudecken. So z. B. Phänomene der Sprachthematization (*so genannte X*, Benutzung von Anführungszeichen) oder skandalisierende Vokabeln wie *Dreistigkeit, Katastrophe, Irrwitz, absurd* etc. Solche Wortklassen können teilweise auch über datengeleitete Verfahren gewonnen werden. Für den Fall der skandalisierenden Vokabeln geht das beispielsweise über Abfragen von „Ich finde das + Nominalgruppe/Adjektiv“ in großen Korpora, um zu bewertenden Nomina und Adjektiven zu gelangen.

### Fazit

Die Methoden der semtracks-Forscherguppe bewegen sich zwischen korpusbasierten (corpus-based) und korpusgeleiteten (corpus-driven) Verfahren (Bubenhof 2009, S. 99ff.). Die Berechnung von typischen Lexemen oder komplexen n-Grammen legt sich nicht vorgängig auf bestimmte Lexeme fest, sondern bezieht alle möglichen Lexeme oder n-Gramme mit in die Analyse ein. Erst der Signifikanztest, der die unterschiedlichen Frequenzen der Lexeme bzw. n-Gramme in jeweils zwei Korpora vergleicht, führt zu nach Typik geordneten Listen. Dadurch gelangt man in einem induktiven Verfahren zu den sprachlichen Phänomenen, die für das jeweilige Untersuchungskorpus typisch



sind und die im Anschluss kategorisiert und damit ge-  
deutet werden.

Genauso wichtig ist aber das hypothesengeleitete, korpusbasierte Verfahren, bei dem z. B. von einer Gruppe von Lexemen angenommen wird, dass sie für emotionalen Stil stehen. Deren Auftretensfrequenz in den Korpora hilft dann, die Korpora zu charakterisieren und zu vergleichen – und damit auch die Hypothese zu überprüfen.

Im semtracks-Blog<sup>2</sup> haben wir bis zur Bundestagswahl 14 Analysen veröffentlicht, die die bedeutendsten Akteure des Wahlkampfes mit einbezogen. Diese Art der Publikation, verbunden mit geeigneten Visualisierungen war wahrscheinlich für das Presseecho mitverantwortlich, das die Analysen auslösten.<sup>3</sup> Damit bot sich die Möglichkeit, eine interessierte Öffentlichkeit auf sprachliche Aspekte des Wahlkampfes aufmerksam zu machen und so auch die Linguistik als Disziplin in den Fokus der Öffentlichkeit zu rücken.

#### Anmerkungen

- <sup>1</sup> Unter Verwendung des TreeTaggers (Schmid 1994).
- <sup>2</sup> Vgl. <<http://semtracks.com/politicaltracker/bundestagswahl/>>
- <sup>3</sup> Vgl. <<http://semtracks.com/index.php?id=Press>>

#### Literatur

Belica, Cyril: (2001): Kookkurrenzdatenbank CCDB. Eine korpuslinguistische Denk- und Experimentierplattform für die Erforschung und theoretische Begründung von systemisch-strukturellen Eigenschaften von Kohäsionsrelationen zwischen den Konstituenten des Sprachgebrauchs, Institut für Deutsche Sprache, Mannheim: <<http://corpora.ids-mannheim.de/ccdb/>>.

Biedermann, Reinhard (1969): Die deutschen Gradadverbien in synchronischer und diachronischer Hinsicht. Heidelberg: Dissertation.

Bierwisch, Manfred (1987): Semantik der Graduierung. Grammatische und konzeptuelle Aspekte von Dimensionsadjektiven (=Studia Grammatica, Bd. 25/27). Berlin: Akademie Verlag, S. 91-286.

Bubenhofer, Noah (2009): Sprachgebrauchsmuster. Korpuslinguistik als Methode der Diskurs- und Kulturanalyse. Berlin / New York: de Gruyter.

Bubenhofer, Noah / Klimke, Martin / Scharloth, Joachim (2008a): political tracker – U.S. Presidential Campaign '08: A Semantic Matrix Analysis: <<http://semtracks.com/politicaltracker/>>.

Bubenhofer, Noah / Klimke, Martin / Scharloth, Joachim (2008b): The Word War: Yes, He Did. How Obama won the (rhetorical) battle for the White House, International Relations and Security Network, ISN ETH Zurich: <[www.isn.ethz.ch/Current-Affairs/Special-Reports/The-Word-War-Yes-He-Did/Analysis](http://www.isn.ethz.ch/Current-Affairs/Special-Reports/The-Word-War-Yes-He-Did/Analysis)>.

Bubenhofer, Noah / Klimke, Martin / Scharloth, Joachim (2009): political tracker – Bundestagswahl '09. Eine Semantische Matrixanalyse: <<http://semtracks.com/politicaltracker/bundestagswahl/>>.

Manning, Christopher D. / Schütze, Hinrich (2002): Foundations of Statistical Natural Language Processing. Cambridge, Massachusetts: The MIT Press.

Os, Charles van (1989): Aspekte der Intensivierung im Deutschen. Tübingen: Narr.

Schmid, Helmut (1994): Probabilistic Part-of-Speech Tagging Using Decision Trees, Institut für maschinelle Sprachverarbeitung. Stuttgart: Universität Stuttgart.

Noah Bubenhofer ist wissenschaftlicher Mitarbeiter am Institut für Deutsche Sprache. Alle Autoren sind Mitarbeiter der Forschergruppe semtracks.