

Big Data – Big Problems? Zur Kombination qualitativer und quantitativer Methoden bei der Erforschung politischer Social-Media-Kommunikation

Michael Klemm und Sascha Michel

Zusammenfassung

Social Media stellen die Medienwissenschaft vor neue theoretische und methodische Herausforderungen, geht es doch um die Analyse großer, dynamischer und vielfältig vernetzter Datenmengen (Big Data) über raumzeitliche Grenzen hinweg. Gefragt ist daher eine systematische Kombination qualitativer mit quantitativen Verfahren, da klassische mikroanalytische Fallanalysen aufgrund des Umfangs, der Vielfalt und Eigendynamik der Daten in vernetzten Online-Medien allein an ihre Grenzen stoßen, automatisierte Tracking-Verfahren hingegen blind für die Feinheiten kommunikativen Handelns bleiben. Den Ertrag einer solchen Methodenkombination demonstrieren wir im Beitrag am Beispiel der politischen Twitter-Kommunikation in Deutschland.

Ausgangspunkte: Aktuelle methodologische Herausforderungen der Medienlinguistik bei der Analyse politischer Online-Kommunikation

Soziale Netzwerke wie *Twitter*, *Facebook*, *YouTube* und Co. haben nicht nur quasi im Sturm die private wie teils auch berufliche Kommunikation erobert und verändert, sie stellen auch die medienwissenschaftliche Forschung vor neue, insbesondere methodologische Herausforderungen bei der Analyse der darin stattfindenden Kommunikation – vor allem angesichts enormer Datenmengen (Stichwort: Big Data), neuer Verbreitungs- und Vernetzungstechnologien (Stichwort: Viralität) und neuer Interaktionsformen (Stichwort: „Liken“ und „Teilen“). Die qualitative Text- und Diskursanalyse hat ihre Stärken in der reichhaltigen und kontextsensiblen Mikroanalyse und Interpretation kommunikativer Daten. Sie hat jedoch ihre Grenzen in der Bearbeitung und Auswertung solch großer dynamischer Korpora. Es reicht nicht aus, Einzelfallanalysen durchzuführen, wenn Soziale Medien gerade durch die zeit- und raumübergreifende Vernetzung der Nutzerkommunikation geprägt sind und dadurch weitgreifende dynamische Diskursnetze (z.B. „Blogosphere“ oder „Twittersphere“) entstehen. Mit Methoden der sog. „Digital Humanities“ wiederum können große Datenkorpora bewältigt werden (vgl. z.B. Kumar et al. 2013), diese konzentrieren sich aber meist auf rein quantitative Auswertungen. Eine systematische Kombination von qualitativer Diskursanalyse und informationstechnischen Auswertungsverfahren wäre daher äußerst wünschenswert, steht aber bislang noch weitgehend aus (vgl. zumindest Ansätze in Thimm et al. 2011, 2012; Klemm/Michel 2013, 2014a). Dies gilt umso mehr, als Soziale Medien quasi nebenbei eine Fülle an hochinteressanten Daten erzeugen, die man mit Tools des *Social Media Monitoring* auslesen und in die Analysen integrieren kann (vgl. Abschnitt „Quali trifft Quanti“).

In einer transdisziplinären Kooperation könnten neue Methoden der quantitativen wie qualitativen medienwissenschaftlichen Forschung entstehen, indem die informationstechnologische Auswertung großer Datenmengen (Makroebene) mit der detaillierten Analyse charakteristischer Fallbeispiele und (virtueller) Feldstudien (Mikroebene) verbunden wird (z.B. in Form von Textanalyse und Ethnografie). Mit einem solchen Methodenmix sollten dann auch neue Fragestellungen jenseits der überkommenen Trennlinie zwischen quantitativer und qualitativer Forschung in den Blick genommen werden können.

Unser aktueller Fokus liegt dabei auf politischen Diskursen in Online-Medien (vgl. Klemm / Michel 2013; 2014a). Hier können mit einer Kombination von quantitativen mit qualitativen Methoden eine Fülle von Fragen in Angriff genommen werden:

- Wie und wozu verwenden Politikerinnen bzw. Politiker und Bürgerinnen bzw. Bürger das Social Web – in politischen Diskursen, aber auch (inszeniert) privat?
- Wie verändert sich das Social-Media-Handeln von Politikerinnen und Politikern über einen längeren Zeitraum hinweg? Lassen sich zum Beispiel differenzierte Twitterstile identifizieren?
- (Wie) Verändern Politikerinnen und Politiker z.B. ihr Twitterverhalten und ihre Kommunikation, wenn sie eine andere politische Position übernehmen oder sich die Koalitionen ändern?
- (Wie) Reagieren Politikerinnen und Politiker in Twitter auf öffentlich diskutierte politische Krisen bzw. Imagekrisen?
- Wie beeinflussen Wahlkämpfe das Twitterverhalten, quantitativ wie qualitativ?
- Welche Gemeinsamkeiten und Unterschiede zwischen verschiedenen Sozialen Netzwerken (Facebook, Twitter etc.) können festgestellt werden?
- Wie lassen sich kommunikative Strukturen, spezifische Sprachhandlungen, politische Kommunikationsstile (z.B. interaktiv-bürgerorientierte oder rein distributive Nutzung von Twitter, vgl. Thimm et al. 2011) oder Wissen und Einstellungen in komplexen Diskursen informationstechnisch auslesen und dann qualitativ interpretieren?
- Welche Schlagwörter/Themen/Frames sind in einem Social-Media-Diskurs besonders verbreitet?
- (Wie) Initiieren Politikerinnen und Politiker oder auch Bürgerinnen und Bürger Diskurse in Sozialen Medien? Welche Prinzipien führen zur verstärkten Beachtung dieser Diskursimpulse?
- Verändern die neuen Onlinemedien wirklich das Verhältnis zwischen Politikerinnen und Politikern/Parteien/Parlamenten und Bürgerinnen und Bürgern/Wählerinnen und Wählern/Interessengruppen (ePartizipation)?

- Was unterscheidet neue Parteien (z.B. Piraten), neue soziale Bewegungen (z.B. *Occupy Wallstreet*) oder neue politische Akteure (z.B. *Wikileaks*) mit den von ihnen propagierten alternativen Formen politischer Willensbildung in medienstruktureller und kommunikativer Hinsicht von etablierten Formen?
- Wie verbreiten sich Diskurse (z.B. identifizierbar über Twitter-Hashtags und Keywords) im zeitlichen Ablauf und regionaler Differenzierung (Stadt – Land/lokal – global)?
- Wie national oder transkulturell sind aktuelle politische Diskurse?
- Wie inszenieren sich Politikerinnen und Politiker bzw. Parteien multimodal mittels Fotos und Videos in Sozialen Netzwerken, etwa im Hinblick auf Faktoren wie Glaubwürdigkeit, Kompetenz, Stärke, Volksnähe?

Einige dieser Fragen wollen wir im Beitrag zumindest exemplarisch anschneiden.

Quali trifft Quanti: Zum Einsatz von Social-Media-Monitoring-Tools in qualitativen Analysen

Um solche Fragen erforschen zu können, benötigt man ein umfangreiches Datenkorpus, das auch quantitativ ausgewertet werden muss. Dabei ist der Umgang mit großen Datenmengen im Rahmen der Diskurs- und Korpuslinguistik nichts Neues (vgl. u.a. Bubenhofer 2013), er erhält aber angesichts der Big Data in Sozialen Medien mit teils in die Millionen gehenden Daten/Token, die sich zudem laufend ändern können, eine neue Dimension.¹ Zudem waren und sind die Korpora der etablierten Korpuslinguistik meist nicht ohne Weiteres der Öffentlichkeit zugänglich und in der Regel auf Sprachdaten aus Printmedien beschränkt.

Die Sozialen Medien bieten hingegen geradezu eine Flut an allgemein verfügbaren Daten unterschiedlichster Zeichenqualitäten. Zudem lassen sich zum Beispiel für die Twitterkommunikation über die eigentlichen sprachlichen Äußerungen hinaus zahlreiche Metadaten mittels der sog. „API“ (*Application Programming Interface*) auslesen, der Programmierschnittstelle von Twitter, mit der die Integration von Twitter auf anderen Webseiten ermöglicht wird (vgl. z.B. Gaffney/Puschmann 2013; Einspänner et al. 2013). Für eine professionelle und anwenderspezifische Auswertung sowie dauerhafte Speicherung dieser Korpora und Metadaten benötigt man Programmierkenntnisse,² aber für viele Forschungsfragen bieten

¹ Vgl. zu den Herausforderungen einer „Online-Diskursanalyse“ z.B. Fraas, Meier & Pentzold (2013).

² “In the current state, the ability of individual users to effectively interact with ‘their’ Twitter data hinges on their ability to use the API, and on their understanding of its technical constraints” (Puschmann & Burgess 2013, S. 11).

inzwischen Websites ihre Daten und Auswertungen öffentlich an.³ Zudem steht eine ganze Reihe an frei verfügbaren „Social-Media-Monitoring-Tools“ zur Verfügung, die teils kostenfrei genutzt werden können, teils kommerziell angeboten werden, um über individuelle Suchanfragen authentische Twitter-Metadaten etwa im Hinblick auf persönliche Nutzerdaten (Follower, Followees, Anzahl der Tweets und Retweets, zeitliche und geografische Verteilung etc.), behandelte Themen/Schlagwörter und einzelne Hashtags, die Nutzung der @-Funktion und vieles mehr auszuwerten. Solche einfach nutzbaren Tools sind für Twitter zum Beispiel *Twitonomy*, *Topsy* oder *Hootsuite*, für Facebook *Fanpage Karma* oder *Socialbakers*.

Fallbeispiel: Twitterstile von Spitzenpolitikern

Wie solche Social-Media-Monitoring-Tools gezielt zur Unterstützung bei qualitativen Forschungsfragen eingesetzt werden können, möchten wir im Folgenden an Fallbeispielen aus der politischen Twitterkommunikation zeigen. Welche Rolle Twitter für Politikerinnen und Politiker und Bürgerinnen und Bürger bzw. Wählerinnen und Wähler inzwischen spielt und welche strukturellen Besonderheiten der Kommunikationsform Tweet man dabei beachten sollte, haben wir an anderer Stelle ausführlich erläutert (vgl. Klemm/Michel 2013; 2014a). Hier soll es daher speziell darum gehen, das Ineinandergreifen von quantitativer Auswertung und qualitativen Methoden darzulegen und zu zeigen, welchen Erkenntnisgewinn eine solche gezielte Kombination zum Beispiel für die Medien- und Politolinguistik bringen kann.

Ein Ziel der medienlinguistischen Erforschung politischer Online-Kommunikation kann es z.B. sein, die „Twitter-Profile“ oder Twitterstile einzelner Politikerinnen und Politiker herauszuarbeiten und miteinander zu vergleichen, um typische Muster heutiger Politik(er)kommunikation zu bestimmen.

Betrachten wir dazu zum Beispiel den SPD-Vorsitzenden Sigmar Gabriel (vgl. dazu auch ausführlich Klemm/Michel 2013), der inzwischen zum Vizekanzler aufgestiegen ist. Hier lohnt zunächst eine quantitative Auswertung seiner Tweets mittels des Monitoring-Tools *Twitonomy* (<http://www.twitonomy.com>), um einen ersten Eindruck zu gewinnen (vgl. Abb. 1).

Gabriel hat – nachdem er nach eigener Aussage lange nichts von Sozialen Medien gehalten hat (vgl. Doemens 2012) – im Zeitraum vom 3. Mai 2012 bis 29. Oktober 2013 insgesamt 1.861 Tweets verfasst. So plötzlich, wie er anfing (nach einem Gespräch mit einer Journalistin), hörte er auch wieder auf, kurz bevor er zum deutschen Vizekanzler ernannt wurde.⁴

³ Ein Beispiel für die politische Kommunikation ist „Bundestwitter“ (www.bundestwitter.de), bei dem alle Tweets der Bundestagsabgeordneten in Echtzeit erscheinen und zudem vielfältige Informationen zu Politiker-Accounts bereitgestellt werden (vgl. Puppe 2013). Eine ähnliche Seite ist Pluragraph.de.

⁴ Die Änderung der Twitterkommunikation (in diesem Falle gar deren Einstellen) nach einer Veränderung der politischen Funktion/Position lässt sich – schon quantitativ – auch bei vielen anderen deutschen Spitzenpolitikerinnen und -politikern nach der Bundestagswahl beobachten, was nicht nur mit

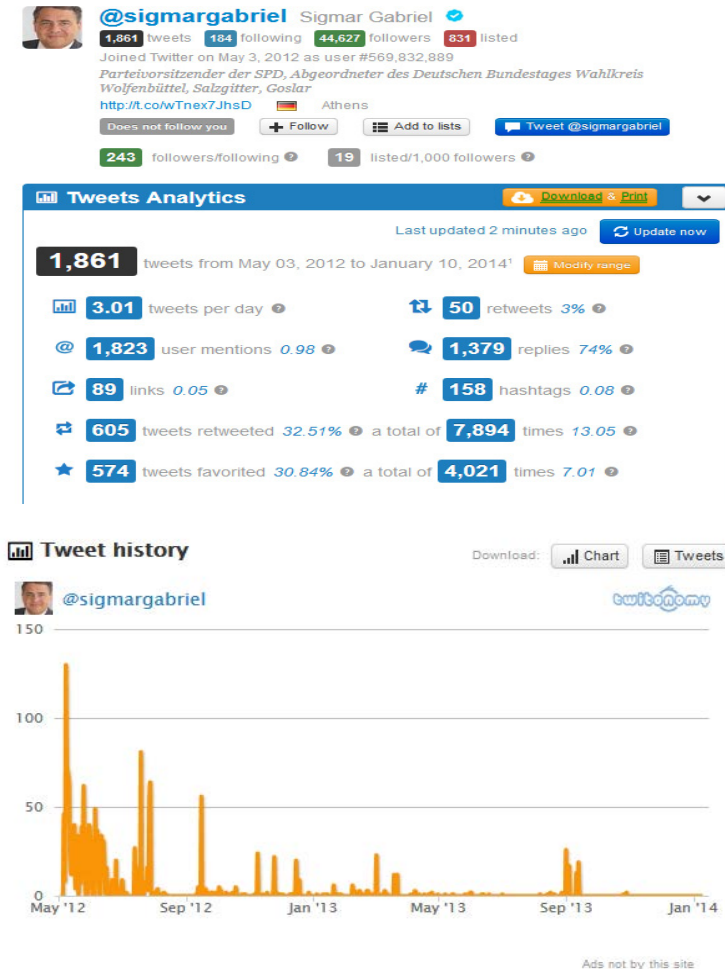


Abbildung 1: Twitonomy-Auswertung der Tweets von Sigmar Gabriel

Die zeitliche Verteilung seiner Tweets ist diskontinuierlich: Nach einer besonders aktiven Anfangsphase mit bis zu 130 Tweets am Tag ging das Aufkommen innerhalb weniger Wochen zurück (ohne dass ein Tag ohne Tweets verging), erreichte aber immer wieder Spitzen von bis zu 80 Tweets am Tag. Erst die qualitative Analyse zeigt, dass diese Spitzen mit bestimmten

dem Ende des Wahlkampfes erklärt werden kann – wie auch Nachfragen bei diesen Politikerinnen und Politikern ergeben haben.

Ereignissen verbunden waren: mit zwei „Twiterviews“ (sozusagen Bürgersprechstunden via Twitter), zudem mit TV-Duellen im Bundestagswahlkampf 2013, die Gabriel per Twitter im Social TV kommentierte. Aussagekräftiger noch als die Verlaufsanalyse sind die Interaktionsdaten. So wird deutlich, dass immerhin ein Drittel seiner Tweets weitergeleitet wurden, er somit als Multiplikator wirkte. Gabriel selbst leitete hingegen nur in 3 Prozent seiner Tweets die Äußerungen anderer weiter – und dann bevorzugt Tweets des damaligen SPD-PR-Chefs Vehlewald. Asymmetrisch ist auch das Verhältnis von Followern und Followees. Während er (Stand: Januar 2014) seine Followerzahl kontinuierlich auf fast 45.000 Personen gesteigert hat (ein Spitzenwert für deutsche Politikerinnen und Politiker), folgte Gabriel selbst nur 184 Personen – und diese stammten fast ausschließlich aus dem Umfeld der SPD, waren demnach Teil der parteipolitischen PR.

Besonders auffällig ist, dass er in 74 Prozent seiner Tweets auf andere antwortete bzw. Personen in seinen Tweets adressierte, darunter viele „Normalbürger“ (so weit sich das überprüfen lässt). Gabriel kann man somit zum Typus „interaktiv-persönlicher Twitterer“ zählen (vgl. Thimm et al. 2012, S. 303), der mit Bürgerinnen und Bürgern sowie anderen Politikerinnen und Politikern in unmittelbaren Kontakt tritt – im Gegensatz zu vornehmlich „thematisch-informativen“ Politikerinnen und Politikern, die insbesondere über Hyperlinks Informationen streuen wie bei klassischen Pressemitteilungen. In Gabriels Tweets kamen Links so gut wie nicht vor. Auch Hashtags, mit denen er sich aktiv an aktuellen Twitter-Diskursen hätte beteiligen können, verwendete er nur in 50 Tweets – und diese vor allem im Zusammenhang mit dem selbst gewählten Hashtag *#fragsigmar* zur Organisation der Twiterviews oder als Zuschauer bei *#tvduell*. Inhaltliche Hashtags fehlen fast komplett.

Solche Daten liefern somit relevante Informationen, aber sie müssen stets interpretiert und an einzelnen Tweets konkretisiert werden. Was bedeutet es, dass Gabriel häufig in Interaktion mit anderen Twitterern trat? Welche kommunikativen Funktionen und Sprachhandlungen waren damit verbunden? Über welche Themen tauschte er sich aus? Welchen sozialen Stil pflegte er in seinen Tweets, wie privat wurde er bei seinen Äußerungen? All dies lässt sich nur durch die detaillierte Analyse einzelner Tweets herausfinden, gehört aber zwingend dazu, um einen politischen Twitter-Stil umfassend zu rekonstruieren.⁵ Nur detaillierte Textanalysen können zeigen, dass sein Sprachstil häufig umgangssprachlich, impulsiv, direkt und pointiert war, was ihm einerseits Respekt in der Twittersphäre eingebracht (vgl. Kriwoj 2012), aber auch „Shitstorms“ ausgelöst hat (u.a. zum Thema Vorratsdatenspeicherung).

Allerdings kann die quantitative Herangehensweise auch auf relevante Korpusstellen und Fragestellungen aufmerksam machen. So kann man z.B. automatisch erheben lassen, welche Tweets von Gabriel am meisten weitergeleitet wurden (vgl. Abb. 2).

⁵ Ein Beispiel für die Grenzen rein quantitativer oder automatisierter Auswertung von Tweets diskutieren wir in Klemm/Michel (2014a, S. 29): Ein Zuschauer kommentiert eine Äußerung des SPD-Politikers von Dohnanyi während einer Talksendung per Tweet mit „Klaus von Dohnanyi (CDU)“. Solch feine Ironie und diese twittertypische Art von politischem Protest eines Zuschauers kann man nur textanalytisch-interpretativ aufspüren (vgl. dazu auch Abschnitt 4).

🔄 Tweets most retweeted



+ More on @sigmargabriel's retweeted tweets

Abbildung 2: Retweet-Statistik als Orientierung für qualitative Forschung

An dritter Stelle seiner am meisten weitergeleiteten Tweets findet man (in Abbildung 2 hervorgehoben) eine Replik von Gabriel auf eine abfällige Bemerkung eines Bürgers über sein Körpergewicht, den er im Gegenzug ebenfalls beleidigt: „Lieber dick als doof. Ganz schlimm, wenn man beides ist: also passen Sie schön auf: bloß nicht zunehmen.“ Dieser Tweet vom 6. Mai 2012 (also kurz nachdem Gabriel mit dem Twittern begonnen hat) wurde insgesamt 117 Mal weitergeleitet und 111 Mal favorisiert. Unklar bleibt aber, wie die Twitterer Gabriels Äußerung kommentiert haben: etwa negativ als unangemessene Beleidigung eines Bürgers durch einen Politiker oder anerkennend als schlagfertige Reaktion? Hier ist die automatische Auswertung an ihrer Grenze angelangt und muss man sich „klassisch“ qualitativ-interpretativ mit den 117 Retweets auseinandersetzen. Über Twitonomy kann man die Tweets immerhin anzeigen und speichern lassen.

Wer retweetet den Beitrag? Meist sind es Vieltwitterer mit vielen Followern, meist zudem (so weit dies erkennbar ist) Sympathisanten oder SPD-Genossen, aber auch etliche „Piraten“ und Netzaktivisten, ein BILD-Vize-Chefredakteur und andere Journalistinnen und Journalisten, zudem viele Bürgerinnen und Bürger ohne erkennbare Parteipräferenz oder sonstigen Politikfokus beim Twittern, aber mit einem offenkundigen Faible für Sprüche, wie man an deren sonstigen Tweets erkennen kann. Wie wird der Retweet kommentiert? Meist gar nicht, er wird schlicht weitergeleitet. In keinem der 117 Fälle wird er explizit negativ kommentiert, nur in einem Fall etwas spöttisch („Ich bin für ein Gewichtslimit von max. 120 kg für

#SigmarGabriel“). Es gibt auch explizites Lob: „Haha. Gut gekontert.“ oder „@sigmargabriel ist sowas von gut auf Twitter“. Es wird sogar der Hashtag *#twitternwiesigmargabriel* erfunden. Insgesamt wurde somit die Bürgerschelte eher anerkennend kommentiert und als schlagfertig goutiert. Der damalige Neu-Twitterer Gabriel wurde von der Twittersphere wohlwollend aufgenommen. Der Fall zeigt zumindest symptomatisch, dass es für Politikerinnen und Politiker nicht schädlich sein muss, sich mit Bürgerinnen und Bürgern in der Twitter-Öffentlichkeit anzulegen.

Wie hilfreich Social-Media-Monitoring für qualitative Fragestellungen sein kann, zeigt sich z.B. auch, wenn man mit dem Tool „Mentionmap“ in Zeitabständen untersucht, mit welchen Themen bzw. Hashtags ein Politiker in der Twittersphere in Verbindung gebracht wird. So treten bei Gabriel im Mai 2013 politische Schlagwörter wie „Mindestlohn“ oder „Armutsbericht“ in den Vordergrund, die zum Kerngeschäft der SPD gehören, zudem „SPD-Bürgerkonvent“. Hier scheint das Agenda-Setting somit funktioniert zu haben – alles Weitere muss freilich interpretativ erforscht werden. Im Januar 2014 fehlen solche Schlagwörter, stattdessen wird Gabriel vor allem in Verbindung mit politischen Sendungen im Fernsehen gebracht. Dies könnte mit dem abrupten Ende seiner Twitteraktivität zusammenhängen: Er setzt keine Themen mehr selbst, sondern wird nur noch in massenmedialen Kontexten erwähnt.

Solche Analysen kann man vergleichend für Politikerinnen und Politiker aller Parteien und Ebenen und in unterschiedlichen Plattformen durchführen und so in der Zusammenschau mit dieser Kombination aus quantitativen und qualitativen Methoden zu einer Typologie des politischen Kommunizierens und Handelns in Sozialen Netzwerken gelangen (vgl. dazu auch Thimm et al. 2012; Siri/Seeßler 2013). So erkennt man z.B. bereits anhand der Twitonomy-Daten, dass die als in Sozialen Medien überaus aktiv geltende CDU-Politikerin Julia Klöckner ganz im Gegensatz zu Gabriel dem thematisch-informativen Stil folgt (erkennbar an zahlreichen Tweets mit Links und Hashtags). Die anschließende inhaltliche Analyse ihrer Tweets zeigt, dass sie Twitter in erster Linie als politisches (Wahlkampf-)Instrument und zum strategischen Agenda-Setting (meist aus ihrer Rolle als Oppositionspolitikerin heraus) verwendet. Im Gegensatz zu Gabriel nutzt sie kaum die direkten Dialogmöglichkeiten des Mediums und trennt zudem klar zwischen öffentlich und privat – sie agiert „strictly to the role“ (Siri/Seeßler 2013).

Nochmals anders ist der Twitterstil des Bonner SPD-Politikers und heutigen Parlamentarischen Staatssekretärs Ulrich Kelber: Er betrieb bis zur Bundestagswahl 2013 via Twitter aktive Oppositionsarbeit und begab sich dazu – vor den Augen der Bürgerinnen und Bürger bzw. Leserinnen und Leser (ähnlich wie Plenum des Bundestags oder in einer Polit-Talkshow) – in den offenen Schlagabtausch mit den politischen Kontrahenten, u.a. mit dem CDU-Politiker Peter Altmaier. Seine Strategie des „(Regierungs-)Politiker STELLEN“ lässt sich aber nicht rein quantitativ nachweisen (auch wenn sich über die Auswertung der @-Adressierungen von Kelber sozusagen als Suchanweisung zeigt, wer seine Hauptadressaten waren und dass diese im Gegensatz zu Gabriel nicht aus der eigenen Partei stammten), sondern über Textanalysen und

vor allem ethnografische Beobachtungen sowie qualitative Interviews,⁶ in denen Kelber seine Motive offenlegt (zum Beispiel: „Das ist eine Möglichkeit, wo ich sticheln kann. Über Social Media kann man jemanden stellen.“). Kelber inszenierte sich als Experte, als Gegner auf Augenhöhe, als Politiker mit Ministerformat – dass er nach der Wahl immerhin Staatssekretär wurde, lässt auf den Erfolg dieser Kommunikationsstrategie schließen. Kelber strebte via Twitter und vorbei an den Massenmedien, zu denen er weniger Zugang hatte, Diskursmacht an: Er versucht, Diskurse zu initiieren, zu lenken, zu dominieren oder auch Gegendiskurse anzustoßen. Er nutzte Twitter auch zur „Prozesstransparenz“ (Siri/Seeßler 2013) und durchaus auch manchmal „publicly private“ (ebd.), wenn er z.B. mit Altmaier in aller Öffentlichkeit duzend und frotzelnd ins Gericht ging („Red nicht ständig um den heißen Brei rum“ oder „Sorry, aber könntest du bitte einmal deinen Job machen? Und die CDU-Vorsitzende könnte auch mal ran.“).

Die drei vorgestellten Twitter-Stile von zwei Politikern und einer Politikerin decken bei weitem nicht das ganze Spektrum ab, wie und wozu Twitter politisch genutzt werden kann (vgl. dazu auch Parmelee/Bichard 2011; Siri/Seeßler 2013). Sie zeigen aber exemplarisch, wie quantitative und qualitative Methoden ineinander greifen (müssen), um Kommunikationskonstellationen wie Netzwerkmedien und komplexe politische Diskurse – im Sinne einer „Medienkulturlinguistik“ (Klemm/Michel 2014b) – holistisch erfassen zu können.

Fallbeispiel: Hashtag #GroKo

Das nächste Fallbeispiel bezieht sich darauf, wie die Kombination quantitativer und qualitativer Auswertungen von Daten in Sozialen Netzwerken dazu beitragen kann, semantische Wort- und Begriffsanalysen durchzuführen. Dies soll am Beispiel eines Wortes, das zunächst in den Sozialen Netzwerken als Hashtag aufkam, aber dann auch von den traditionellen Medien aufgegriffen und schließlich zum Wort des Jahres 2013 in Deutschland gewählt wurde, exemplarisch gezeigt werden. Das Kurzwort *GroKo* (< *Große Koalition*) entstand im Zuge der Koalitionsverhandlungen zwischen CDU/CSU einerseits und SPD andererseits nach der Bundestagswahl 2013. Die Prominenz dieses Wortes – und das geht z.T. auch aus der öffentlichen Erklärung der Gesellschaft für deutsche Sprache hervor – lässt sich auf zwei Aspekte zurückführen: 1. Die Homophonie zu *Kroko* (< *Krokodil*) ermöglicht semantische Assoziationen, die gelegentlich auch visuell als Karikatur umgesetzt wurden, in der ein unterdimensionierter SPD-Parteichef Sigmar Gabriel im überdimensionalen Maul der als Krokodil dargestellten Bundeskanzlerin Angela Merkel herausschaute. 2. Das Wort steht symptomatisch für die sprachlichen Prägkräfte, welche die technischen Beschränkungen bestimmter digitaler Kommunikationsplattformen mit sich bringen und dadurch für die sprachlich-kommunikativen Innovationen, die

⁶ Sascha Michel untersucht aktuell in seiner Dissertation „Der Politiker als Kommunikator“ u.a. das individuelle Twitterhandeln von Politikerinnen und Politikern unterschiedlicher Parteien und Ebenen. Dazu kombiniert er die stilpragmatische Analyse von Politiker-Tweets mit ausführlichen Begleitstudien/teilnehmenden Beobachtungen und Tiefeninterviews, um das alltägliche Kommunizieren von Politikerinnen und Politikern insbesondere in Sozialen Medien holistisch herauszuarbeiten.

sich aus der digitalen in die analoge Welt distribuieren. Dieser Hashtag wurde aufgrund der Zeichenbeschränkung auf 140 Zeichen bei Twitter gebildet und konnte alternative Hashtags wie *#GroKoa* oder *#GroKoal* sowohl frequenziell als auch hinsichtlich der Bedeutung für die Nutzer deutlich überholen. Hierdurch und durch die Tatsache, dass *GroKo* auch nach den Koalitionsverhandlungen als Hashtag funktional eingesetzt wurde/wird, avancierte es zu einem „öffentlichen Wort“ (Dang-Anh et al. 2013: 152) und unterscheidet sich von eher temporalen hochfrequenziellen Hashtags wie z.B. *#tvduell*.

Der Gebrauch dieses in Sozialen Netzwerken neu gebildeten und primär dort verwendeten Begriffs wirft jedoch einige Fragen auf: Wie hat sich die Verwendung frequenziell entwickelt, d.h. gibt es Hoch- oder Niedrigphasen der Verwendung? In welchen Kontexten tritt das Wort auf und ergeben sich dadurch semantische Rückschlüsse? Worauf bezieht sich das Wort in den jeweiligen Kontexten und lassen sich dadurch referenzielle Muster ableiten?

Die quantifizierenden Tools www.tweetbinder.com und www.hashtagify.me liefern zunächst einmal statistische Aussagen über die Entwicklung von *#GroKo*. Eine Suchanfrage bei Tweetbinder etwa zeigt, dass zwischen dem 05.01.2014 und dem 10.01.2014 in 2.000 Tweets von insgesamt 1.406 Nutzern der Hashtag vorkam. Von diesen 2.000 Tweets wurden mit 43,26 % die Mehrheit aller Tweets (873) retweetet, also von Nutzerinnen und Nutzern weitergeleitet. Neue Tweets machten 24,67 % (498) aller Tweets aus, 23,33 % (472) enthielten Verlinkungen und Fotos und nur 8,72 % (176) wurden im Rahmen von Antworten auf andere Tweets verwendet.⁷ Damit wird deutlich, dass in diesem Zeitraum, der kurz nach der Regierungsbildung liegt, weniger der Formulierung neuer Tweets mit diesem Hashtag eine bedeutende Funktion zukommt als vielmehr der Weiterleitung von Tweets.

Die dadurch naheliegende Vermutung, dass die Bildung neuer Tweets mit dem Hashtag *#GroKo* zu diesem Zeitpunkt allmählich stagniert oder sogar rückläufig ist, wird durch eine Suchanfrage bei Hashtagify.me bestätigt. Demnach war der „Popularity Trend“ im Januar 2014 sowohl im Wochen- (-2) als auch im Monatstrend (-4) rückläufig.⁸ Was Hashtagify.me auch liefert, ist eine frequenzielle Korrelationsanalyse, die grafisch demonstriert, mit welchen anderen Hashtags *#GroKo* wie oft erscheint. Die Ergebnisse dieser Art von „Frameanalyse“ sind dabei keineswegs überraschend, stellen doch *#SPD*, *#CDU*, *#CSU*, *#Gabriel* und *#Merkel* die wichtigsten Protagonisten der großen Koalition dar. Mit *#VDS*, *#NSA* und *#Energiewende*

⁷ Eine Anfrage bei Tweetbinder für den Zeitraum vom 19.-25.07.2014 ergab folgende Verteilung: 790 Tweets insgesamt, 40,02 % Retweets, 13,46 % neue Tweets, 27,27 % mit Verlinkungen und Fotos und 19,22 % als Antwort. Somit zeigt sich, dass sich die Frequenz dieses Hashtags zwar deutlich reduziert (mehr als halbiert) hat, es aber in Antworten offenbar eine stärkere prozentuale Verwendung findet.

⁸ Auch hier zeigt eine Anfrage vom 26.07.2014, dass *#GroKo* mit Höchstwerten von 30 die Höchstwerte vom Dezember 2013 und Januar 2014 (teilweise über 50) nicht mehr erreicht, sich aber erstaunlich stabil erweist (der Wochentrend steigt sogar um 5 Punkte, der Monatstrend um 1 Punkt). Dies alles spricht für eine allmähliche Lexikalisierung des Wortes.

werden zudem zentrale und zwischen den beteiligten Parteien strittige Punkte des *#Koalitionsvertrags* abgedeckt, die zudem in Twitter eigene Diskurse gebildet haben.⁹

Mit diesen quantitativen Aussagen zu Entwicklung und Wortkorrelationen bekommen wir ein Indiz, einen Hinweis darauf, in welchen Begriffsfeldern sich *#GroKo* befinden könnte, welche Begriffsassoziationen bestehen können, welche evtl. stärker oder schwächer ausgeprägt sind. Mehr erfährt man durch solche Makroanalysen indes nicht, sodass man etwa schnell an die empirischen Grenzen stößt, wenn es um die kontextuellen Bedeutung(en) und Referenzbezüge des Wortes geht. Diese lassen sich nur durch qualitative Mikroanalysen von Fallbeispielen mit dem Ziel der fallübergreifenden Mustererkennung bestimmen. Die makroanalytischen Ergebnisse zeigen uns, wo wir mikroanalytisch ansetzen und „suchen“ müssen, weshalb ihnen jenseits der visuellen Präsentation statistischer und allgemeiner Tendenzen und Fakten auch eine ganz konkrete, unmittelbare Funktion im Zusammenspiel mit qualitativen Untersuchungen zukommt. Das soll im Folgenden durch den Referenzbezug von *#GroKo* verdeutlicht werden.

Die quantitative Korrelationsanalyse zeigt, dass der Hashtag in irgendeiner Korrelation zu *#SPD*, *#CDU*, *#CSU*, *#Gabriel*, *#Merkel*, *#VDS*, *#NSA* und dem *#Koalitionsvertrag* steht. Somit empfiehlt es sich, in einem ersten Schritt Tweets mit diesen Korrelationen für eine qualitative Analyse heranzuziehen, um anschließend nach weiteren, evtl. peripheren Korrelationen zu suchen. Die Analyse zeigt, dass diese Hashtags attributiv, also ergänzend zu *#GroKo* erscheinen können, indem sie entweder *#GroKo* näher spezifizieren oder umgekehrt. Viele Beispiele machen aber auch deutlich, dass *#GroKo* auf die korrelierten Hashtags referiert, so dass sich drei Referenzbezüge dieses Hashtags herausarbeiten lassen:

1. Referenz auf die Koalition (generisch):



#GroKo wird hier mit den an der großen Koalition beteiligten Parteien – in Hashtags aufgelistet – gleichgesetzt.

2. Referenz auf die Bundesregierung:

⁹ Interessanterweise hat sich die Korrelation im Zeitraum 01.2014-07.2014 kaum geändert, es tauchen nach wie vor die gleichen Protagonisten auf. Bei den Themen wiederum wurde *#Energiewende* durch *#Mindestlohn* ersetzt. Solche Hashtagkorrelationen erweisen sich demnach als äußerst dynamisch und situationsabhängig.



In diesem Tweet nimmt ein Twitteruser mit *#GroKo* auf die Bundesregierung Bezug, mehr noch, er setzt die *#GroKo* (lästernd) mit der Bundeskanzlerin gleich.

3. Referenz auf Themen des Koalitionsvertrags:



Mindestlohn, Vorratsdatenspeicherung (VDS) und PKW-Maut sind Themen, die hier per Hashtagverkettung mit der *#GroKo* – freilich gerahmt in einem Lästermodus – identifiziert werden. Die (junge) GroKo wird hier vom FDP-Politiker Wissing als zerstritten dargestellt.

Gerade das letzte Beispiel illustriert den zuvor bereits angedeuteten Umstand erneut eindrücklich, dass quantitative Erhebungen meist nur top-down die qualitative Validierung und Spezifizierung bereits ermittelter Themen etc. erlauben, bei (twittertypischen) sprachlichen Phänomenen wie Lästern, Ironie, Wortwitz, Metaphern etc. jedoch, die es nur interpretativ und bottom-up zu ermitteln gilt, auf qualitative Analysen angewiesen sind.

Zwischen Makro und Mikro: Einige (vorläufige) Forschungsergebnisse

Mit einer solchen Kombination von quantitativen wie qualitativen Methoden haben wir bislang vor allem die politische Kommunikation in Twitter und Facebook untersucht: sowohl Politikerbeiträge wie Bürgeräußerungen und deren interaktive Vernetzung. Die bisherigen Erkennt-

nisse (vgl. Klemm/Michel 2013; 2014) legen nahe, dass Politik einen hohen Stellenwert in der Twitterkommunikation hat. Dabei sind die Tweets meist auf „persönliche Öffentlichkeiten“ (Schmidt 2012) orientiert, die den jeweiligen Twitterern „folgen“, potenziell kann aber jeder über twitter.com oder auch in seiner „Timeline“ die eingehenden Tweets verfolgen, favorisieren, kommentieren, weiterleiten. Tweets sind somit Bausteine eines unendlichen flüchtigen öffentlichen Kommunikationsflusses, auch über politische Themen jedweder Art, und liefern dabei häufig (über die Nutzung von Operatoren wie # und @) „Diskurshäppchen“, die schnell in der Twittersphäre (über RT) verbreitet und via Massenmedien popularisiert werden können.

Das deliberative Potenzial dieser neuen und direkten Kommunikationskanäle wird aber noch bei Weitem nicht ausgeschöpft. Den Politikerinnen bzw. Politikern und Parteien dienen Social Media – wie quantitative Auswertungen z.B. der Adressatenstruktur und Hyperlink-Frequenz zeigen – (noch) eher der Information und Mobilisierung ihrer Parteigänger/Wähler als der Interaktion und Kommunikation, sie sind mehr ein „Branchenmedium“ (Christoph Bieber) untereinander oder für die gezielte Information von Journalistinnen und Journalisten denn ein Draht zum Bürger, auch wenn der Twitterstil von Gabriel Ansätze dazu zeigte. Die auf Politikerinnen und Politiker bezogenen Bürgeräußerungen dienen laut unseren bisherigen quantifizierenden wie interpretativen Studien weniger der Politik(er)aneignung im engeren Sinne, sprich der ernsthaften rationalen Reflexion oder der Integration politischer Themen in die eigene Lebenswelt, als der Selbstdarstellung und dem Spaß der Community – oder dem öffentlichen Abbau von Politikfrust.

In Twitter findet man eine Fülle an politischen Deutungs- und Bewertungsangeboten zu aktuellen Diskursen (insbesondere, wenn sie Hashtags beinhalten wie etwa bei der Sexismus-Debatte 2013 rund um *#aufschrei*), sie können politischen Protest manifestieren und der Koordinierung und Mobilisierung von Aktivisten dienen (wie bei Stuttgart 21, vgl. Thimm/Bürger 2012). Durch die Kommunikationsform bedingt (vgl. Klemm/Michel 2013, S. 13f.) gibt es aber kaum echte dialogische Aushandlungen, eher Meinungsbekundungen als Dialoge, eher Evaluation als Argumentation, eher Amüsement (mit der Ironie als typischer Interaktionsmodalität) als Raisonement. Immerhin: Politisches Handeln wird in Twitter als (zumindest inszenierte) Echtzeitkommunikation öffentlich sichtbar, auswertbar, kritisierbar; Bürger können sich unmittelbar an Diskursen beteiligen oder diese mittels Hashtags initiieren und Politikerinnen und Politiker dabei direkt adressieren. Möglich machen dies neuartige Vernetzungsstrukturen über technische Operatoren wie RT, # oder @ – die zugleich für die Forschung wichtige Parameter für die quantitative Auswertung wie qualitative Interpretation bereitstellen.

Fazit und Ausblick: Auf dem Weg zur „Roboter-Medienforschung“?

Die spielend leichte automatisierte Erhebung und Auswertung großer Datenmengen ist verführerisch. Sie liefert eine Unmenge an Metadaten, die bereits für sich betrachtet Interpretationen ermöglichen. Damit geht freilich die Gefahr einher, es mit dieser rein quantitativen Auswertung bewenden zu lassen, um Big Data vergleichbar und generalisierbar erfassen zu können.

Wir hoffen gezeigt zu haben, dass eine derartige Reduktion, bei aller Begeisterung über die Reichhaltigkeit und Aussagekraft der mitgelieferten Metadaten, ein Fehler wäre.

In den Massenmedien ist bereits der „Roboter-Journalismus“ zunehmende Praxis, das heißt die automatische Generierung von Texten aus Datenbanken durch Unternehmen wie „Narrative Science“, etwa in der Sport- und Börsenberichterstattung.

„Niemand verfolgt Twitter aufmerksamer als die Roboter. [...] Kein Journalist hat die Zeit, Millionen Tweets aufzufinden, zu lesen und zu analysieren, aber Narrative Science schafft das mühelos und im Handumdrehen. Beabsichtigt ist nicht nur, über Statistiken zu berichten, sondern dem Leser zu vermitteln, was die ganzen Zahlen bedeuten.“ (Morozov 2012)

Sind wir vielleicht auch schon auf dem Weg zu einer „Roboter-Medienforschung“, die sich im Sog der „Digital Humanities“ auf automatisierte und statistisch vergleichbare Big-Data-Analysen beschränkt und qualitative stilpragmatische Fallanalysen als obsolet deklariert? Bislang bleiben jedenfalls die meisten Twitter-Untersuchungen bei quantitativen Beschreibungen stehen. Einerseits sollte deutlich geworden sein, dass gerade die Social-Media-Kommunikation eine Unmenge an neuen und spannenden (Meta-)Daten generiert, die relativ einfach und erkenntnisreich mit den Methoden der praktischen Informatik analysiert werden können – und (insbesondere bei Diskursanalysen) auch sollten. Andererseits ist die rein quantitative Auswertung eher eine Heuristik, eine Suchhilfe als ein fertiges Ergebnis: Alle Daten bedürfen der kontextgesättigten Interpretation, alle Erkenntnisse müssen an konkreten Kommunikaten überprüft und belegt werden. Oder anders formuliert: Tracking-Verfahren weisen aus der Makroperspektive auf interessante Befunde hin, die auf der Mikroebene mit qualitativen Methoden interpretiert oder zumindest ergänzt werden können – und müssen.

Hier wollen wir ansetzen und so weitere Muster, Stile und Funktionen politischer Twitter-Kommunikation identifizieren. Hierfür sind holistische Analysen erforderlich, die sowohl Produktionsaspekte (Motive, Praktiken und Stile twitternder Politiker bzw. Bürger/Zuschauer) wie die Kommunikate wie die Aneignung dieser Tweets in Beziehung zueinander setzen (vgl. Klemm/Michel 2014b), auch indem wir die quantitativ ausgewertete Datenbasis durch stilpragmatische Textanalysen, ethnografische Begleitstudien und Interviews qualitativ anreichern – was im vorliegenden Rahmen nur angedeutet werden konnte.

Literatur

- Bubenhof, Noah (2013): Quantitativ informierte qualitative Diskursanalyse. Korpuslinguistische Zugänge zu Einzeltexten und Serien. In: Roth, Kersten Sven & Spiegel, Carmen (Hrsg.): *Angewandte Diskurslinguistik*. Berlin: Akademie Verlag, S. 109–134.
- Dang-Anh, Mark; Einspänner, Jessica & Thimm, Caja (2013): Kontextualisierung durch Hashtags. Die Mediatisierung des politischen Sprachgebrauchs im Internet. In: Diekmannshenke, Hajo & Niehr, Thomas (Hrsg.): *Öffentliche Wörter. Analysen zum öffentlich-medialen Sprachgebrauch*. Stuttgart: ibidem, S. 137–159.

- Doemens, Karl (2012): Sigmar Gabriel entdeckt Facebook für sich. *Frankfurter Rundschau online*, 25.01.2012. Online: <http://url9.de/Vt0> [Stand vom 20-07-2014].
- Einspänner, Jessica; Dang-Anh, Mark & Thimm, Caja (2013): Computer-assisted content analysis of Twitter data. In: Weller, Katrin et al. (Hrsg.): *Twitter and Society*. New York: Peter Lang, S. 97–108.
- Fraas, Claudia; Meier, Stefan & Pentzold, Christian (Hrsg.) (2013): *Online-Diskurse. Theorien und Methoden transmedialer Online-Diskursforschung*. Köln: von Halem.
- Gaffney, Devin & Puschmann, Cornelius (2013): Data collection on Twitter. In: Weller, Katrin et al. (Hrsg.): *Twitter and Society*. New York: Peter Lang, S. 55–67.
- Klemm, Michael & Michel, Sascha (2013): Der Bürger hat das Wort. Politiker im Spiegel von Userkommentaren in Twitter und Facebook. In: Diekmannshenke, Hajo & Niehr, Thomas (Hrsg.): *Öffentliche Wörter. Analysen zum öffentlich-medialen Sprachgebrauch*. Stuttgart: ibidem, S. 113–136.
- Klemm, Michael & Michel, Sascha (2014a): Social TV und Politikaneignung. Wie Zuschauer die Inhalte politischer Diskussionssendungen via Twitter kommentieren. *Zeitschrift für Angewandte Linguistik (ZfAL)* 60 (1), S. 3–35.
- Klemm, Michael & Michel, Sascha (2014b): Medienkulturlinguistik. Plädoyer für eine holistische Analyse von (multimodaler) Medienkommunikation. In: Benitt, Nora et al. (Hrsg.): *Korpus – Kommunikation – Kultur: Ansätze und Konzepte einer kulturwissenschaftlichen Linguistik*. Trier: WVT, S. 183–215.
- Kriwoj, Sachar (2012): So twittern Politiker: Sigmar Gabriel. *UdL digital*, 30.07.2012. Online: <http://www.udldigital.de/so-twittern-poliker-sigmar-gabriel> [Stand vom 20-07-2014]
- Kumar, Shamanth; Morstatter, Fred & Liu, Huan (2013): *Twitter Data Analytics*. New York: Springer.
- Morozov, Evgeny (2012): Texte in null Komma nichts: Roboterjournalismus. *Frankfurter Allgemeine Zeitung* vom 04.04.2012. Online: <http://www.faz.net/-hbj-6yw8g> [Stand vom 20-07-2014].
- Parmelee, John H. & Bichard, Shannon L. (2011): *Politics and the Twitter Revolution: How Tweets Influence the Relationship Between Political Leaders and the Public*. Lanham u.a.: Lexington.
- Puppe, Thomas (2013): *Messung und Visualisierung von Twitter-Einfluss und Tweet-Inhalten der MdB*. Masterarbeit, HTW Berlin.
- Puschmann, Cornelius & Burgess, Jean (2013): *The Politics of Twitter Data*. Online: <http://ssrn.com/abstract=2206225> [Stand vom 20-07-2014].
- Schmidt, Jan Henrik (2012): Das demokratische Netz? *Politik und Zeitgeschichte* 3/2012, S. 3–8.

- Siri, Jasmin & Seßler, Katharina (2013): *Twitterpolitik. Politische Inszenierung in einem neuen Medium*. München: IfM. Online: <http://medienpolitik.eu/cms/media/pdf/Twitterpolitik.pdf> [Stand vom 20-07-2014].
- Thimm, Caja & Bürger, Tobias (2012): *Digitale Citoyens: Politische Partizipation in Zeiten von Social Media. Fallanalysen zur politischen Beteiligung in Deutschland, Ägypten und China*. Online: http://www.bapp-bonn.de/medien/downloads/publikation_forschungsergebnisse/publikation_forschungsergebnisse_kurz.pdf [Stand vom 20-07-2014].
- Thimm, Caja; Dang-Anh, Mark & Einspänner, Jessica (2011): Diskurssystem Twitter: Semiotische und handlungstheoretische Perspektiven. In: Anastasiadis, Mario & Thimm, Caja (Hrsg.): *Social Media. Theorie und Praxis digitaler Sozialität*. Frankfurt am Main: Lang, S. 265–285.
- Thimm, Caja; Einspänner, Jessica & Dang-Anh, Mark (2012): Twitter als Wahlkampfmedium. Modellierung und Analyse politischer Social-Media-Nutzung. *Publizistik* 57 (3), S. 293–313.