When politicians talk: assessing online conversational practices of political parties on Twitter

Lietz, Haiko; Wagner, Claudia; Bleier, Arnim; Strohmaier, Markus

Veröffentlichungsversion / Published Version
Konferenzbeitrag / conference paper

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
GESIS - Leibniz-Institut für Sozialwissenschaften

When Politicians Talk:
Assessing Online Conversational Practices of Political Parties on Twitter

Haiko Lietz
GESIS & University of Duisburg-Essen
Cologne and Duisburg, Germany
haiko.lietz@gesis.org

Claudia Wagner
GESIS & University of Koblenz
Cologne and Koblenz, Germany
claudia.wagner@gesis.org

Arnim Bleier
GESIS - Leibniz Institute for the Social Sciences
Cologne, Germany
arnim.bleier@gesis.org

Markus Strohmaier
GESIS & University of Koblenz
Cologne and Koblenz, Germany
markus.strohmaier@gesis.org

Abstract

Assessing political conversations in social media requires a deeper understanding of the underlying practices and styles that drive these conversations. In this paper, we present a computational approach for assessing online conversational practices of political parties. Following a deductive approach, we devise a number of quantitative measures from a discussion of theoretical constructs in sociological theory. The resulting measures make different – mostly qualitative – aspects of online conversational practices amenable to computation. We evaluate our computational approach by applying it in a case study. In particular, we study online conversational practices of German politicians on Twitter during the German federal election 2013. We find that political parties share some interesting patterns of behavior, but also exhibit some unique and interesting idiosyncrasies. Our work sheds light on (i) how complex cultural phenomena such as online conversational practices are amenable to quantification and (ii) the way social media such as Twitter are utilized by political parties.

Introduction

In recent years, Twitter has established itself as a popular medium for public mass-personal communication (Wu et al. 2011). Our community has studied different aspects of Twitter communication, such as network structures (e.g., Jürgens, Jungherr, and Schoen 2011; Larsson and Moe 2012), conversational practices (e.g., Honeycutt and Herring 2009; Huang, Thornton, and Efthimiadis 2010), or analyses of dynamics (e.g., Becker, Naaman, and Gravano 2011; Bruns and Burgess 2011). While this work has advanced our understanding about communication on Twitter, we know little about how these different perspectives can be integrated and quantified, and how they relate to conversational practices in a political context. Future computational social scientists would certainly benefit from computational tools and instruments that translate theoretical constructs from sociology to quantifiable measures that are amenable to computation and therefore can be applied on larger scales.

For example, it can be expected that in online conversations, political parties attempt to define themselves by adopting unique topical foci and conversational styles that are inherently difficult to assess. Having computational methods available to track and assess these styles would greatly increase our sociological understanding of such processes. In Figure 1, we present preliminary evidence for this idea. The three networks depict the static outcome of communication among politicians on Twitter dissected according to politicians' following, retweeting, and mentioning practices. Table 1 provides accompanying statistics about these networks. Interesting differences are present. For example, homophily effects are weakest for mentioning practices and the variance of homophily is also largest for this practice (cf. Table 1). This indicates that all parties tend to follow and retweet members of their own party but may use mentioning for inter-party debate. This finding is in line with previous results from a study on Twitter use before the U.S. midterm election of 2010 (Conover et al. 2011). Sound computational methods that would enable such qualitative insights into online conversational practices on a wide scale would represent a useful addition to the arsenal of social science methods.

Research questions: These interesting yet preliminary differences in the static outcomes of conversational practices lead us to expect to see differences in the socio-cultural processes that produce these outcomes as well. While our observations above suggest that following, retweeting, and mentioning exhibit unique traits, we know little about the particularities and idiosyncrasies of these practices and their party-specific adoption.

For example, what are the different purposes that tagging, retweeting, and mentioning serve in online political conversations? How do they differ from each other? How consistently are different practices used across different parties? Moreover, one would also expect that these different practices are effected to varying extents by external events or factors. For example, how would a (TV) debate or the day of the election itself effect the online conversational practices of parties in general or individual parties specifically?

Approach: In this paper, we set out to answer these and related questions by presenting and applying a computational approach to assessing the socio-cultural dynam-
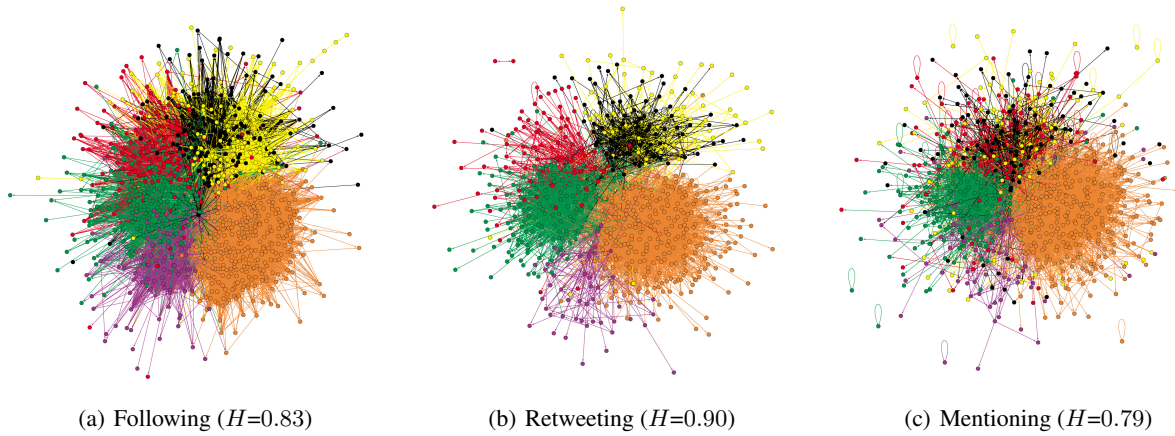
(a) Following (H=0.83) (b) Retweeting (H=0.90) (c) Mentioning (H=0.79)

Figure 1: **Examples of online conversational practices on Twitter**: Structures of the aggregate following, retweeting, and mentioning networks of German politicians from 9 weeks before to 4 weeks after the federal election 2013. The vertices in the networks correspond to user handles and are color-coded by party affiliation (colors given in Table 1). Arcs correspond to following/retweeting/mentioning relationships and are colored by sender. Structural differences between different practices can be observed: For example, homophily H effects are lower in the mentioning network (0.79) than in the following (0.83) and retweeting (0.90) networks. CDU/CSU and FDP, which formed the last government coalition in Germany, are tightly knit in the follow and retweet networks. The Pirates are largely decoupled from a relatively pluralistic mentioning space where all other parties transact. The networks were laid out using the Kamada-Kawai algorithm.

ics of online conversational practices *over time*. Our work is rooted in relational sociology, specifically in theoretical work that considers episodes of stability and change in practice (Mohr and White 2008; White 2008; Fuhse 2009; Padgett and Powell 2012). We are interested in making different aspects of online conversational practices, in particular cultural focus, - similarity, and - reproduction as well as institutions and punctuations, amenable to quantitative measurements. In doing so, we follow a *deductive style of research*, deriving measures from a theoretical discussion of sociological constructs. While this enables us to root our measures in theory, it makes validation a challenging endeavour. To evaluate our approach nonetheless, we choose to apply it to a particular case, i.e., to conversational practices of political parties on Twitter before, during, and after the German federal election of September 22nd, 2013. This enables us to generate insights into the practical utility of our deductive approach in a real world scenario, as well as into the conversational practices of the case itself.

Contributions: The contributions of our work are threefold: First, we present and discuss several sociological constructs related to conversational practices on a theoretical level. Second, we present a computational approach that deduces measures for each of the sociological constructs of interest. While the constructs are grounded in sociological theory, the proposed measures stem from computer science, social science, information science, and related fields. Third, to demonstrate the utility of our computational approach, we conduct a case study on the German federal election 2013 and present empirical insights into the conversational practices of German politicians during the course of this event.

The paper is structured as follows: After related work we

Table 1: **Statistics and dataset for the German federal election 2013 on Twitter – parties differ in several interesting ways:** Consistently across all conversational practices, the Pirate party exhibits the most homophilic behavior. Mentioning is most strongly used by the Pirates and the Greens ($D = 0.07$), the two largest parties in terms of microblogging politicians (312 and 178), but not in terms of how many votes they actually received (2.2% and 8.4%). D denotes network density, $\bar{k}$ denotes average in/out-degree of nodes, $\bar{w}$ denotes average in/out-weight of nodes, and H denotes homophily, i.e., the tendency of a political party to communicate within party boundaries, computed on the individual level.

Party	Election Result	Politicians	Following				Retweeting				Mentioning			
			D	$\bar{k}_{out}$	$\bar{k}_{in}$	H	D	$\bar{w}_{out}$	$\bar{w}_{in}$	H	D	$\bar{w}_{out}$	$\bar{w}_{in}$	H
CDU/CSU	41.5%	158	0.15	32	38	0.78	0.08	14	14	0.86	0.04	16	22	0.64
SPD	25.7%	143	0.18	35	41	0.80	0.05	7	9	0.84	0.04	11	17	0.72
FDP	4.8%	143	0.17	35	35	0.78	0.05	7	9	0.84	0.03	6	9	0.55
Greens	8.4%	178	0.21	50	51	0.82	0.08	24	24	0.89	0.07	31	29	0.77
Left	8.6%	97	0.23	30	32	0.79	0.07	8	13	0.91	0.04	13	14	0.68
Pirates	2.2%	312	0.16	57	52	0.89	0.06	40	38	0.93	0.07	73	69	0.92
Total	**91.2%**	**1,031**	**0.05**	**44**	**44**	**0.83**	**0.02**	**25**	**25**	**0.90**	**0.02**	**39**	**39**	**0.79**

introduce the sociological background and constructs which form the theoretical foundation of our computational approach which we present subsequently. Finally, we describe our empirical study which demonstrates the utility of our approach, discuss our empirical insights, and conclude our work.

Related Work

Previous work which is relevant for our research focused either on analyzing the hashtagging, mentioning, and retweeting behavior of Twitter users in general or on analyzing the role of Twitter in a political context. To our best knowledge, ours is the first work which coherently studies the tagging, mentioning, and retweeting behavior of politicians as socio-cultural processes, and presents comparative analysis that yield unique insights into online conversational practices of politicians and political parties on Twitter.

Hashtags, Mentions, and Retweets: Twitter is used for many purposes, including the reporting of daily activities, communicating with other users, sharing information, and reporting, or commenting on, news (Java et al. 2007). As such it enables several *conversational practices*. Hashtags are primarily used to describe news or communications with others and to find other users' tweets about certain topics. Since tagging behavior is inspired by the observed use of hashtags in a users' network (Huang, Thornton, and Efthimiadis 2010), coherent semantic structures emerge from hashtag streams (Wagner and Strohmaier 2010). Retweeting is the forwarding of other users' tweets. By 2010, conventions as to how, why, and what users retweet had emerged, but the practice had not yet stabilized (Boyd, Golder, and Lotan 2010). The retweetability of a Twitter message is related to its informational content and value and the embeddedness of its sender in following networks (Suh et al. 2010). Mentions, sometimes called @mentions or replies, emerged to be Twitter's convention for the interactive use of addressing others, although it is also being used for other purposes like referencing. In 2009, most tweets without a mention reported daily activities while tweets with @ signs exhibited much higher variance in terms of topics and types of content (Honeycutt and Herring 2009).

Elections on Twitter: With the rise of social media many researchers got interested in exploring the role of Twitter for politics. Scientists from different backgrounds started exploring to what extent Twitter can predict election results with contradicting conclusions. Some research suggests that election results can be predicted by analyzing Twitter (e.g., Livne et al. 2011), while Jungherr (2013) shows that, at least for the German multi-party system, election result can not be predicted using Twitter. Metaxas, Mustafaraj, and Gayo-Avello (2011) conduct a meta study and conclude that electoral predictions using the published research methods on Twitter data are not better than chance.

Political discourse on Twitter: Besides the interest in the voting behavior of persons and its reflection on Twitter, researchers also got interested in studying political discourse on Twitter. For example, Conover et al. (2011) analyze the retweet and mention networks from 6 weeks leading up to the 2010 U.S. midterm election. Interestingly, the authors find extremely limited connectivity between right- and left-leaning users in the retweet network, but not in the mention network. This indicates that retweets and mentions are different conversational practices. The work of Jürgens, Jungherr, and Schoen (2011) shows that on Twitter new gatekeepers and ordinary users tend to filter political content based on their personal preferences. Therefore, political communication on Twitter might be highly dependent on a small number of users, critically positioned in the structure of the network.

Politicians on Twitter: In previous research Thamm and Bleier (2013) found that retweets are motivated by professional uses while replies are used mainly for personal communication. Schweitzer (2011) explored political e-campaigns in Germany and concluded that they increasingly reflect those patterns of traditional election coverage that have been held accountable for rising political alienation among the public, i.e., strategic news and extensive negativism.

Following the suggestion of Murthy (2012) who emphasizes the importance of a sociological understanding of Twitter, we now present an approach based on sociological constructs that allows to assess political online conversations.

Theoretical Constructs

In this section, we elaborate the theoretical sociological constructs behind conversational practices which build the basis of our computational approach. In doing so we start from sociological network theory known as relational sociology (Carley 1991; Mohr and White 2008; White 2008; Fuhse 2009; Padgett and Powell 2012). The central premise is that social life is complex and stochastic and that *identities*, which can be persons, groups, or higher-level agents, try to gain *control* over the ensuing uncertainty through regularities. In many contexts, control is gained by collectively forming a densely clustered and culturally coherent community, triadic closure and homophily being the mechanisms (McPherson, Smith-Lovin, and Cook 2001; Kossinets and Watts 2009).

We first discuss an accessible understanding of culture and then offer an idealtype of how culture is reproduced by styles of practice. Because of the stochastic nature of social formations, we finally deal with system perturbation.

Cultural Facts and Homophily: Gaining control is a project in the socio-cultural space-time of simultaneous emergence of culture and feedback on structure: As agents act, they are not only engaged in transactions with other agents, they also make reference to pieces of information or *facts* such as ideas, beliefs, concepts, symbols, knowledge, etc. From this practice, a culture emerges which can be understood as a distribution of referenced facts (Carley 1991). But as agents in transactions are confronted with facts, existing culture positively or negatively feeds back onto agent behavior. Computer simulations are able to produce the densely clustered and culturally coherent communities which are ubiquitously found if cultural similarity breeds social connection (Carley 1991; Axelrod 1997).

Styles, Institutions, and Reproduction: If an identity has gained control and maintains it over time it has a *style*. Styles are inert mechanisms of reproduction. It is due to the existence of such dynamical regularity that identities are predictable (Kosinski, Stillwell, and Graepel 2013; Schoen et al. 2013). Cycles of reproduction (styles) simultaneously determine practices of individuals and groups (Mohr and White 2008; White 2008).

Distributions of facts like words, religious beliefs, surnames, and citations are fat-tailed (Clauset, Shalizi, and Newman 2009). *Institutions* are cultural facts that are stably reproduced over time. Thereby, they become popular and are consequently found in the fat tail of probability distributions. Mechanisms of cumulative advantage (Simon 1955; Dellschaft and Staab 2008; Papadopoulos et al. 2012) are theoretically compatible and have been shown to be capable of producing these skewed distributions. However, not all facts in the tail are necessarily institutions, since facts can also become popular due to short activity bursts rather than prolonged activity.

Punctuations: Continuity and normality is just one side of social life. Only in equilibrium persons and groups can reproduce uninterruptedly. In reality, social life is *punctuated* by minor and major events that interrupt the normal flow of reproduction. Such punctuations may reflect scientific (Kuhn 1962), political (Brunk 2001), or economic (Padgett and Powell 2012, Ch. 6) innovations or generally perturbations originating from inside or outside the observed system. Any account of socio-cultural processes must account for both normality and change (Padgett and Powell 2012). It is both reproduction and punctuation that we study in this paper.

Operationalization of Constructs

Next, we describe how we operationalize the sociological constructs described in the previous section for assessing online conversational practices of politicians and political parties on Twitter.

Approach: The central idea behind conversational practices as socio-cultural processes is that a transaction – a user using a hashtag, a user following another user, a user retweeting another user, or a user mentioning another user – involves a social subject (the referencing user) and a cultural object (a referenced hashtag, followee, retweetee, or mentionee). In the case of following, retweeting, and mentioning, subjects and objects are of the same type, exemplifying the duality of the social and the cultural. In the empirical study which we will present later we are interested in how politicians practice online conversations on Twitter, concretely we study their mentioning, retweeting, and hashtagging practices. We use aggregations of users (parties) as *objects of study* and referenced facts (i.e., user handles and hashtags) as *units of analysis*. For analyzing dynamics, tweets are batched into bins of one week each. Though our operationalization is specific to the context of our empirical study (politics on Twitter), our computational approach for assessing online conversational practices is general and can be applied to study the practices of other groups of users on

Twitter (or even on other stream-based social systems which provide support for conversations and allow to observe the referencing of facts over time).

Cultural Focus and Similarities

In the following, we interpret the observable cultural objects (i.e., hashtags and user handles) of conversational practices (i.e., retweeting, hashtagging, and mentioning) of Twitter users as cultural *facts*. The culture of a group of users (which corresponds to a party i in our case) is represented by a vector $\sigma_i = (a_1, a_2, ..., a_n)$ where the elements are the frequencies of (a) the hashtags that members of party i used, (b) the users that members of party i retweeted, or (c) the users which members of party i mentioned within the observation period.

Cultural Focus: To quantify the extent to which a party i reveals a *cultural focus F* on few selected hashtags or users (facts), the normalized Shannon entropy (Shannon 2001) – a measure of disorder – of the party's fact vector is used:

$$F(\sigma_i) = 1 - \frac{-\sum_{j=1}^{n} p(a_j) * \log_2 p(a_j)}{\log_2(n)} \quad (1)$$

Here, $p(a_j)$ corresponds to the frequency of a cultural fact a_j for party i divided by the frequency of all other facts of that party. Because the denominator holds the maximum entropy and normalized entropy is subtracted from 1, F falls in the range $[0, 1]$ where 0 means that there is no focus (highest entropy) and 1 that the focus is highest (no entropy).

Cultural Similarity: Culturally similar agents are more probable to interact with each other. To reveal potential homophily effects, cultural similarities of parties are studied over time. We propose to measure the *cultural similarity S* of two parties i and j using the cosine similarity (Baeza-Yates and Ribeiro-Neto 1999) of their fact vectors:

$$S(\sigma_i, \sigma_j) = \frac{\sigma_i \cdot \sigma_j}{\parallel \sigma_i \parallel \parallel \sigma_j \parallel} \quad (2)$$

Because facts cannot have negative frequencies, similarities are in the range $[0, 1]$ where 0 indicates no similarity and 1 highest similarity. Cosine measures the similarity for individual party pairs. To obtain a score for a single party i, we compute its average cosine similarity to all other parties j. We consciously do not weight scores by user size or tweet volume because we operate on the level of parties as objects of study.

Styles, Institutions and Reproduction

Cultural Reproduction: Styles are mechanisms of reproduction of focus. To study and quantify styles, we proceed in two steps. First, we measure the reproduction of cultural focus from week to week. In a second step, the stability of reproduction over time is studied to generate a quantitative judgement about styles. Reproduction is operationalized by an extended version of the Rank Biased Overlap (RBO) metric (Wagner et al. 2014) which allows to measure stability in social streams. The *cultural reproduction R* of a party i is

defined as follows:

$$R_i(\sigma_i, p) = (1 - p) \sum_{d=1}^{\infty} \frac{2 \cdot \sigma_i^{1:d}(t1) \cap \sigma_i^{1:d}(t2)}{|\sigma_i^{1:d}(t1) + \sigma_i^{1:d}(t2)|} p^{(d-1)} \quad (3)$$

The cultural fact vectors $\sigma_i(t1)$ and $\sigma_i(t2)$ of party i hold the frequencies of the facts which have been observed at $t1$ and $t2$, respectively. These vectors are not necessarily conjoint since new facts can be introduced at any point in time. Let $\sigma_i^{1:d}(t1)$ and $\sigma_i^{1:d}(t2)$ be the ranked vectors at depth d. The scores fall in the range $[0, 1]$ where 0 means that the culture of a party i at $t1$ is completely different from its culture at $t2$ (culture is unstable, turnover is maximum), and 1 means that it is identical (i.e., does not change over time and is therefore perfectly stable). The parameter p ($0 \leq p < 1$) determines how steep the decline in weights is. The smaller p is, the more tail-weighted the metric is, the tail being the high frequency part of the distribution. We chose $p = 0.9$ which means that facts in the fat tail are most relevant for assessing the cultural stability of a party compared to changes in the head of the distribution (86% of the weight is given to the 10 highest ranked cultural facts if $p = 0.9$). This makes intuitively sense because, if a party changes a cultural fact which is core to them, they will become very unstable according to our measures, while, if they change a fact which is just one among many with low frequencies, this will only have a small impact on their stability.

Mathematically, a style is present when $\delta R_i / \delta t \geq 0$. If R decreases over time, the party will dissolve into chaos as less and less facts are reproduced. If R increases, the party will eventually freeze as reproduction reaches 1 and cultural facts will forever be the same. From a systems perspective, a constant turnover of facts is desired.

Institutionness: Cultural reproduction is strongly related to the construct of stable facts called institutions which are expected to be found in the fat tail of the probability distribution of cultural facts. Two notions resonate with the idea of stable facts: that they are (a) highly referenced and (b) continuously highly referenced. It seems straightforward to divert the Hirsch index (Hirsch 2005) from its intended use. Originally proposed to quantify an individual's scientific research output, a scientist has an index h if h of his papers have at least h citations each, and all but h papers have no more than h citations each. Applied to our temporal case, a fact has an index h if for h weeks it is referenced at least h times, and in all but h weeks no more than h times. A normalization – one part of which is a division of a paper's citation rate by the citation rate of an average paper in the subject area – has been proposed to make the Hirsch index comparable across scientific disciplines with different publication and citation characteristics (Radicchi and Fortunato 2008). In analogy, a fact a has index h if for h weeks it is referenced at least h/h_0 times, and in all but h weeks no more than h/h_0 times, where h_0 is the week-specific reference rate of an average fact. Let's call this index the *institutionness I* of fact a. Since data ranges over 13 weeks, I is in the range $[0, 13]$.

Punctuations

Burstiness: Having a style means being predictable. Punctuations interrupting the normal flow of reproduction leave traces in Twitter in the form of activity bursts. To operationalize, we refer to Kleinberg's (2003) observation that social streams are "punctuated by the sharp and sudden onset of particular episodes". A fact is considered to burst if it leaves a period where its popularity relative to other facts is small and enters a period in which its relative popularity is reasonably large.

Suppose there are n batches of transactions made by party i (in our case $n = 13$ since we use weeks as batches). The t-th batch contains r_t transactions with references to fact a out of a total of d_t transactions. Let $R = \sum_{t=1}^{n} r_t$ and $D = \sum_{t=1}^{n} d_t$ and $p_s = \frac{R}{D} 2^s$. To reveal the burstiness of a fact a, we compute for each week t the party-specific cost of that fact to be transformed in a burst state s:

$$\gamma(s, r_t, d_t) = -\ln\left(\binom{d_t}{r_t} p_s^{r_t} (1 - p_s)^{d_t - r_t} \right) \quad (4)$$

The weight of a fact's burst – its *burstiness B* – in the period $[t1, t2]$ is the improvement in cost by being in state $s = 1$ for the period $[t1, t2]$:

$$B(a, t1, t2) = \sum_{t=t1}^{t2} \left(\gamma(0, r_t, d_t) - \gamma(1, r_t, d_t) \right) \quad (5)$$

In words, for each bursting fact a its burstiness in a period $t1$ (onset) to $t2$ is obtained. Facts can burst multiple times. There is no upper bound for a fact's burst weight. The final

Table 2: **Operationalization of sociological constructs**.

Theoretical Construct	Measure	Description
Cultural Focus F	Shannon Entropy (Shannon 2001)	*How strongly does a party focus on cultural facts?*
Cultural Similarity S	Cosine Similarity (Baeza-Yates and Ribeiro-Neto 1999)	*How similar are parties in terms of their culture vectors?*
Cultural Reproduction R	Rank Biased Overlap (Wagner et al. 2014)	*How stable is a party's culture vector over time?*
Institutionness I	Hirsch Index (Hirsch 2005)	*How many weeks is a fact referenced that many times?*
Burstiness B	Kleinberg's Burst Weight (Kleinberg 2003)	*How popular is a fact in a given period relative to other facts?*
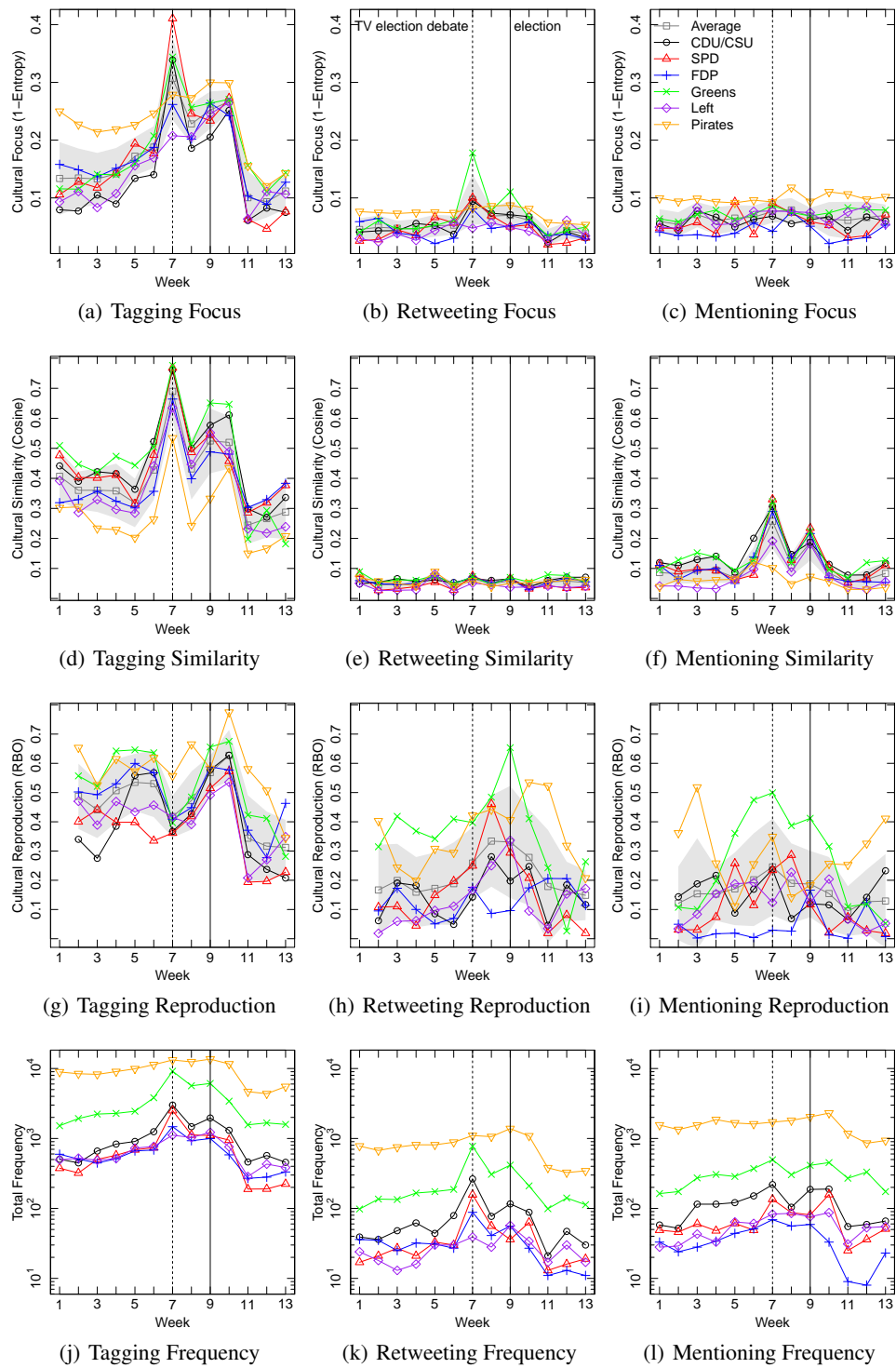
Figure 2: **Online conversational practices of the six main political parties during the German federal election 2013:** Tagging, retweeting, and mentioning practices are shown over a period of 13 weeks. The three conversational practices (columns) are assessed using measures (rows) summed up in Table 2. For the first three rows, curves for an average party are shown, obtained by averaging the six party scores. Shaded regions show standard deviations. The weeks of the election (straight line) and the TV election debate (broken line) clearly mark episodes of irregular activity. The last row shows the total number of references made to facts and is included as reference.

burstiness measure is normalized by the weight of a party's strongest burst and is, therefore, also in the range $[0, 1]$.

The German Federal Election 2013

In this section we describe the dataset which we used to conduct our empirical study of three conversational practices of politicians before, during, and after the German federal election 2013. Then, we present our results. In particular, we present our findings on hashtagging, mentioning, and retweeting practices of German politicians during the 2013 elections. While our approach would also be applicable to analyze following practices (as introduced in Figure 1(a)), we could not perform this analysis since the fine-grained temporal data necessary to study the evolution of follower networks on Twitter is impossible to collect via the public Twitter API.

Dataset

The dataset consists of German politicians and their communicative transactions on the microblogging platform Twitter between July 20th and October 18th, 2013. This corresponds to about 9 weeks before and 4 weeks after the federal election of September 22nd. A preliminary version of the "Twitter corpus of candidates" described by Kaczmirek et al. (2013) is used. Our dataset contains 1,031 user account handles, 98% of which were the "most relevant" candidates for the German parliament in early 2013, and their tweets. 983 politicians were active in the sense that they either tagged a tweet or followed, retweeted, or mentioned another politician during our window of observation. In total, there are 123,819 tweets containing at least one hashtag, 16,292 tweets retweeting at least one other politician (identified by "RT username"), and 25,778 tweets mentioning at least one other politician (identified by "@ username").

Each politician belongs to one political party. The Social Democratic party SPD is the oldest (150yrs), followed by the Conservatives CDU/CSU (68yrs), the Liberals FDP (65yrs), the Greens (Die Grünen, 34yrs), the Left (Die Linke, 24yrs), and the new, internet-affine Pirates (Piratenpartei, 7yrs).

Results

Before going into the details of the dynamics that created the networks shown in Figure 1(b) and 1(c), we first address the two episodes that left spikes in the curves of Figure 2 which visualizes the main dynamics results. Both relate to offline events. The first is the election of September 22nd in week 10. Message volumes, and average tagging focus and retweeting reproduction rate increase towards the events and decrease afterwards. Though the election day was early in week 10, we can see that most of the communication occurred in week 9. Therefore, the horizontal line marking the event in Figure 2 is in week 9. The second is the TV election debate between the two candidates for chancellor, A. Merkel (CDU) and P. Steinbrück (SPD), of September 1st, early in week 7. One can tell because the corresponding hashtag #tvduell is the strongest burst for the CDU/CSU and FDP, the second strongest for the SPD and the Greens, the third strongest for the Left, and the fourth strongest for the

Pirates, restricted to week 7 in each case. See Table 3 for the burstiness of selected facts. Despite the election being the main event in this stream, the debate marks the maximum in terms of total message frequency (cf. Figure 2(j)) and the average tagging focus and similarity (cf. Figure 2(a) and 2(d)). It also left spikes in the average retweeting focus and mentioning similarity (cf. Figure 2(b) and 2(f)). In the following, we will discuss these events, average party dynamics, as well as the peculiarities of single parties.

Cultural Focus: The two major events in the stream have different impacts on the three practices. Figure 2(a) shows that all parties almost monotonously concentrate their focus on hashtags towards the election (average increases from 0.13 in week 1 to 0.27 in week 10) and drastically lose focus afterwards (down to 0.11). The TV debate increased the tagging focus of the CDU/CSU, SPD, FDP, and the Greens, but did not gain enough attention by the Left and the Pirates to do so for them. The latter were the two parties which could not hope for being a junior partner in the coming government coalition.

The two events hardly caused changes of focus in the retweeting practice (cf. Figure 2(b)). One can further see that the reweeting as well as the mentioning practice have a very low focus (high entropy) compared to the tagging practice and don't seem to be impacted by external events. For the retweeting practice we observe a slight increase of focus during the TV debate. A closer look into our data reveals that the debate caused the SPD and Greens to retweet what their most vocal members had to say. The SPD's chairman S. Gabriel and the Green's parliamentary managing director V. Beck were even subject of their parties' strongest retweet bursts. These two parties' activities contributed most to slightly increasing the retweet focus on average.

The absence of any change in the mentioning practice (cf. Figure 2(c)) reveals that, even if real world events like the debate or the election itself caused the average party to sharpen its thematic focus, it did so without changing its focus of who to talk to.

Cultural Similarity: Even though parties increase their individual thematic focuses towards the election, they don't become more similar to each other until the election debate causes them all to increasingly focus on overlapping hashtags (cf. Figure 2(d)). This general tagging behavior suggests that, absent punctuations, parties try to maintain distinctiveness by finding a topical niche and focusing on it. This time, the election debate also impacts the Left and Pirates and causes them to become more similar to other parties, though they remain least similar to other parties over time. This suggests that all parties included topics which emerged during the TV debate into their tagging practice.

Interestingly, neither of the events had an impact on the similarities between parties according to their retweeting practices, but both impacted their mentioning similarities (cf. Figure 2(e) and 2(f)).

One possible explanation for this observation is that the debate narrowed down the pool of politicians to those which were subject of the debate. Evidence for this explanation can be found by taking a closer look at the data: both par-

ties whose candidates were debating (CDU/CSU and SPD) experienced their strongest bursts as their members started mentioning the SPD's candidate P. Steinbrück who also bursted strongly for the Greens and received most mentions that week from the FDP. The CDU/CSU candidate A. Merkel, who did not maintain a Twitter account, was tagged instead, her hashtag #merkel bursting in week 7 for all parties (see Table 3 for details). This indicates that the TV debate caused changes in the mentioning and tagging practices of most parties, while neither the debate nor the election caused any party to notably start retweeting another party's voices.

Cultural Reproduction: So far, we have shown that parties increase their thematic focus towards the election and become more similar according to their tagging practices as the election comes closer and the debate bursts. However, it remains unclear how stably the parties reproduce their focuses. Recall that a style is absent if cultural reproduction decreases.

Average tagging reproduction – the stability of practice – increases from 0.49 to 0.63 in election week 9 where it peaks and subsequently drops (cf. Figure 2(g)). The average party's stability increases, focuses are increasingly reproduced, and turnover is reduced. But the style is strongly perturbed by the TV debate. The sudden drop in reproduction means that the punctuation keeps parties, on average, from increasingly narrowing their focuses. They lose stability. Again, the Left and the Pirates are not affected by this event. But least effected is the SPD. This is because the SPD had the largest turnover of hashtags in the two weeks before the debate (i.e., the SPD was least stable in weeks 5 and 6, as one can see in Figure 2(g)). Consequently, the event did not punctuate their style. The CDU/CSU, on the other hand, had a strong reproduction of focus and was consequently negatively affected by the event. The debate does not have an impact on reproduction in retweeting (cf. Figure 2(h)) and mentioning (cf. Figure 2(i)). This is expected

because focus and similarity were not, or only weakly, affected by the event. There is a slight trend in the retweet practice for all parties, except the FDP (cf. Figure 2(h)). Parties increasingly reproducing their retweeting focus, while not becoming more similar to each other (cf. Figure 2(e)), suggests that they maintain distinctiveness in terms of who their spokespersons are – the homophily principle at work.

But all growth of reproduction must come to an end and actually does after the election, as one can see in Figures 2(g), 2(h), and 2(i). Election styles terminate and post-election styles start, from lower levels of reproduction. Average mentioning reproduction is noisy but rather constant over time, indicating that the debate about which politician is worth commenting on is progressing with smooth turnover.

Institutionness: Institutionness I captures how long a fact is getting referenced often. Since election styles had a duration of up to 10 weeks, any fact with $I > 10$ can be regarded to represent a defining factor for a party's identity. Indeed, among the facts with $I > 10$, we find environment-related hashtags only for the Greens and internet and surveillance-related tags mostly for the Pirates. Unsurprisingly, each party's name hashtag (#cdu, #csu, #spd, etc) has a score $I = 13$. Table 3 shows that the election hashtag #btw13 has the same score $I = 13$ for all parties but the FDP ($I = 10$). This points at the fundamental difference between the two events discussed above. The election is a long-term topic that endures and is repeatedly reproduced by long-term styles. The TV debate is a bursty punctuation interrupting these styles (see low institutionness and high burstiness scores for all parties in Table 3). Politicians getting retweeted or mentioned for more than 10 weeks are all top-runners and authorities with functions like minister, vice chairman, or political managing director.

Burstiness: So far, there is no evidence that hashtags that were important online in the TV debate episode were actually new. But in fact, the average proportion of hashtags not referenced before week 7 is 65% across parties, a percent-

Table 3: **Institutionness I and burstiness B of selected facts:** The election received continuous attention by all parties (high I scores) but bursted only for some parties (it was the Left's strongest burst). The short-lived TV debate, on the other hand, bursted strongly for all parties. In the process, the one debater's hashtag #merkel and the other debater's username @peersteinbrueck also bursted for almost all parties. Square brackets give onset and end of burst episodes.

Fact	Description	CDU/CSU		SPD		FDP		Greens		Left		Pirates	
		I	B	I	B	I	B	I	B	I	B	I	B
#btw13	Federal election (Sep 22)	13	0	11	0	10	0.69 [8,10]	13	0.18 [10,10]	11	1.00 [8,10]	13	0.46 [10,10]
#tvduell	TV election debate (Sep 1)	3	1.00 [7,7]	3	0.94 [7,7]	1	1.00 [7,7]	3	0.66 [7,7]	1	0.75 [7,7]	2	0.85 [7,7]
#merkel	CDU/CSU, chancellor	11	0.10 [7,7]	11	0.25 [7,7]	4	0.25 [7,7]	13	0.18 [7,7]	8	0.39 [6,7]	13	0.10 [7,7]
RT sigmargabriel	SPD, chairman	1	0	3	1.00 [7,8]	0	0	2	0.26 [5,8]	0	0	0	0
RT Volker_Beck	Greens, parliamentary managing director	2	0	3	0	1	0	8	1.00 [7,7]	1	0	1	0.26 [6,7]
@peersteinbrueck	SPD, candidate for chancellor	5	1.00 [7,8]	8	1.00 [6,8]	5	0	6	0.78 [7,8]	2	0	2	0.24 [7,7]

age not reached in any other week. The increase of focus and similarity, and the decrease in reproduction in tagging are likely due to this renewal.

Of all hashtags having its burst onset in week 7, due to the selection of the sample, only a handful used by the Pirates is devoid of political relevance or meaning. Table 3 shows examples of selected facts and their burstiness scores.

Discussion

In our empirical study, we found that the focus on hashtags is more concentrated and similarities between parties through common hashtag usage are much more pronounced than foci on, and similarities through, retweeted or mentioned politicians. Tagging is also more prone to perturbation by offline events than retweeting or mentioning. Retweet practices are much more stable because the users who enjoy attention tend to remain the same even as bursts of topical activity propagate through the system. Even though studying each practice is diagnostic of socio-cultural process, studying tagging is most indicative of culture as the use of language is the ultimate form on symbolic communication (Padgett and Powell 2012, Ch. 4). The notion of "social" is stronger in retweeting and mentioning, where the facts and the users referencing the facts are identical.

Furthermore, we used our approach to study the effects of a major offline event, the TV debate of the two top candidates three weeks prior to the election. It left a distinct mark in the tagging practice of politicians but only impacted some aspects of their retweeting an mentioning practice. In addition to revealing differences between distinct practices we also found interesting differences between individual parties. For example, while most parties increased their tagging focused during the week of the TV election debate, the focus of the Pirates and the Left was not impacted by this event. This points at differences in what parties were trying to control: All but the latter were campaigning for becoming a part of the new government – the latter never had a chance to be more than part of the opposition. The results of our empirical study can be seen as anecdotal evidence for the fact that our approach is suitable for detecting differences in the conversational practices of one or several groups of users.

Regarding the objects of study, a typical dynamic for an average party was described: In the run-up to the election a party has a style of practice geared toward controlling uncertainties in communication. Certain topics are focused on and reproduced. Members of other parties are not retweeted but mentioned if an event – in out case the TV election debate – creates a need to do so. After the election, attention and control is directed in another direction. Style is altered but certain institutions – well-established (self-)references in topics and users – remain at the core of how parties define themselves.

An interesting and plausible extension of this work would be the study of inter-party similarities rather than average similarities over longer periods of time. Finally, our computational approach in general can also be applied to understand how non-politicians participate in political conversations as well.

Conclusions

We have presented a computational approach to assessing online conversational practices of political parties on Twitter. Our approach is rooted in, and informed by, theoretical constructs from relational sociology, in particular cultural focus, - similarity, and - reproduction of agents as well as institutions and punctuations. We devised measures for each of these dimensions to enable the computational assessment of socio-cultural structure and dynamics of online conversational practices. We presented our approach and demonstrated its usefulness in a study on the German federal election 2013.

Overall, we find that several online conversational practices differ significantly. We highlight these differences along with interesting commonalities among political parties by studying political communication on Twitter.

While our computational approach was applied to a single case study, the approach is not limited to a single case. Rather it is general enough that it can be applied to other contexts in which different groups of agents communicate and to other social media systems similar to Twitter. We hope that our work equips future computational social scientists with an improved instrument to assess and study online conversational practices in social media over time.

References

Axelrod, R. 1997. The dissemination of culture: A model with local convergence and global polarization. *Journal of Conflict Resolution* 41:203–226.

Baeza-Yates, R. A., and Ribeiro-Neto, B. 1999. *Modern Information Retrieval*. Boston, MA: Addison-Wesley Longman Publishing Co., Inc.

Becker, H.; Naaman, M.; and Gravano, L. 2011. Beyond trending topics: Real-world event identification on twitter. In *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, 438–441. Menlo Park, CA: AAAI Press.

Boyd, D.; Golder, S.; and Lotan, G. 2010. Tweet, tweet, retweet: Conversational aspects of retweeting on twitter. In *Proceedings of the Forty-third Hawaii International Conference on System Sciences*. IEEE Computer Society Press.

Brunk, G. G. 2001. Self-organized criticality: A new theory of political behaviour and some of its implications. *British Journal of Political Science* 31:427–445.

Bruns, A., and Burgess, J. 2011. #ausvotes: How twitter covered the 2010 australian federal election. *Communication, Politics and Culture* 44:37–56.

Carley, K. 1991. A theory of group stability. *American Sociological Review* 56:331–354.

Clauset, A.; Shalizi, C. R.; and Newman, M. E. J. 2009. Power-law distributions in empirical data. *SIAM Review* 51:661–703.

Conover, M.; Ratkiewicz, J.; Francisco, M.; Gonçalves, B.; Flammini, A.; and Menczer, F. 2011. Political polarization on twitter. In *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*. Menlo Park, CA: AAAI Press.

Dellschaft, K., and Staab, S. 2008. An epistemic dynamic model for tagging systems. In *Proceedings of the Eighteenth ACM Conference on Hypertext and Hypermedia*, 71–80. New York, NY: ACM Press.

Fuhse, J. A. 2009. The meaning structure of social networks. *Sociological Theory* 27:51–73.

Hirsch, J. E. 2005. An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences* 102:16569—-16572.

Honeycutt, C., and Herring, S. C. 2009. Beyond microblogging: Conversation and collaboration via twitter. In *Proceedings of the Forty-Second Hawaii International Conference on System Sciences*. Los Alamitos, CA: IEEE Computer Society Press.

Huang, J.; Thornton, K. M.; and Efthimiadis, E. N. 2010. Conversational tagging in twitter. In *Proceedings of the Twenty-First ACM Conference on Hypertext and Hypermedia*, 13–18. New York, NY: ACM Press.

Java, A.; Song, X.; Finin, T.; and Tseng, B. 2007. Why we twitter: Understanding microblogging usage and communities. In *Proceedings of the Ninth WebKDD and First SNA-KDD Workshop on Web Mining and Social Network Analysis*, 56–65. New York, NY: ACM Press.

Jürgens, P.; Jungherr, A.; and Schoen, H. 2011. Small worlds with a difference: New gatekeepers and the filtering of political information on twitter. In *Proceedings of the Third International Web Science Conference*. New York, NY: ACM Press.

Jungherr, A. 2013. Tweets and votes, a special relationship: The 2009 federal election in germany. In *Proceedings of the Second Workshop on Politics, Elections and Data*, 5–14. New York, NY: ACM Press.

Kaczmirek, L.; Mayr, P.; Vatrapu, R.; Bleier, A.; Blumenberg, M.; Gummer, T.; Hussain, A.; Kinder-Kurlanda, K.; Manshaei, K.; Thamm, M.; Weller, K.; Wenz, A.; and Wolf, C. 2013. Social media monitoring of the campaigns for the 2013 german bundestag elections on facebook and twitter. *GESIS-Working Papers* (29).

Kleinberg, J. 2003. Bursty and hierarchical structure in streams. *Data Mining and Knowledge Discovery* 7:373–397.

Kosinski, M.; Stillwell, D.; and Graepel, T. 2013. Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences* 110:5802–5805.

Kossinets, G., and Watts, D. J. 2009. Origins of homophily in an evolving social network. *American Journal of Sociology* 115:405–450.

Kuhn, T. S. 1962. *The Structure of Scientific Revolutions*. Chicago, IL: University of Chicago Press.

Larsson, A. O., and Moe, H. 2012. Studying political microblogging: Twitter users in the 2010 swedish election campaign. *New Media & Society* 14:729–747.

Livne, A.; Simmons, M. P.; Adar, E.; and Adamic, L. A. 2011. The party is over here: Structure and content in the 2010 election. In *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, 201–208. Menlo Park, CA: AAAI Press.

McPherson, M.; Smith-Lovin, L.; and Cook, J. M. 2001. Birds of a feather: Homophily in social networks. *Annual Review of Sociology* 27:415–444.

Metaxas, P. T.; Mustafaraj, E.; and Gayo-Avello, D. 2011. How (not) to predict elections. In *Third International Conference on Privacy, Security, Risk and Trust and Third International Conference on Social Computing*, 165–171. Los Alamitos, CA: IEEE Computer Society Press.

Mohr, J. W., and White, H. C. 2008. How to model an institution. *Theory and Society* 37:485–512.

Murthy, D. 2012. Towards a sociological understanding of social media: Theorizing twitter. *Sociology* 46:1059–1073.

Padgett, J. F., and Powell, W. W. 2012. *The Emergence of Organizations and Markets*. Princeton, NJ: Princeton University Press.

Papadopoulos, F.; Kitsak, M.; Serrano, M. A.; Boguñá, M.; and Krioukov, D. 2012. Popularity versus similarity in growing networks. *Nature* 489:537–540.

Radicchi, F., and Fortunato, C. S. 2008. Universality of citation distributions: towards an objective measure of scientific impact. *Proceedings of the National Academy of Sciences* 105:17268–17272.

Schoen, H.; Gayo-Avello, D.; Metaxas, P. T.; Mustafaraj, E.; Strohmaier, M.; and Gloor, P. 2013. The power of prediction with social media. *Internet Research* 23:528–543.

Schweitzer, E. J. 2011. Normalization 2.0: A longitudinal analysis of german online campaigns in the national elections 2002–9. *European Journal of Communication* 26:310–327.

Shannon, C. E. 2001. A mathematical theory of communication. *Mobile Computing and Communications Review* 5:3–55.

Simon, H. A. 1955. On a class of skew distribution functions. *Biometrika* 42:425–440.

Suh, B.; Hong, L.; Pirolli, P.; and Chi, E. H. 2010. Want to be retweeted? large scale analytics on factors impacting retweet in twitter network. In *Proceedings of the Second International Conference on Social Computing*, 177–184. Los Alamitos, CA: IEEE Computer Society Press.

Thamm, M., and Bleier, A. 2013. When politicians tweet: A study on the members of the german federal diet. arxiv:1305.1734.

Wagner, C., and Strohmaier, M. 2010. The wisdom in tweetonomies: Acquiring latent conceptual structures from social awareness streams. In *Proceedings of the Third International Semantic Search Workshop*. New York, NY: ACM Press.

Wagner, C.; Singer, P.; Strohmaier, M.; and Huberman, B. 2014. Semantic stability in social tagging streams. In *Proceedings of the Twenty-Third International World Wide Web Conference*.

White, H. C. 2008. *Identity and Control: How Social Formations Emerge*. Princeton, NJ: Princeton University Press.

Wu, S.; Hofman, J. M.; Mason, W. A.; and Watts, D. J. 2011. Who says what to whom on twitter. In *Proceedings of the Twentieth International Conference on World Wide Web*, 705–714. New York, NY: ACM Press.