A Transnational Analysis of News and Tweets about *Nuclear Phase-Out* in the Aftermath of the Fukushima Incident

Philipp Heinrich[1], Christoph Adrian[2], Olena Kalashnikova[3], Fabian Schäfer[3], Stefan Evert[1]

Friedrich-Alexander-University Erlangen-Nuremberg

[1]Chair of Computational Corpus Linguistics, Bismarckstr. 6, 91054 Erlangen

[2]Chair of Communication Science, Findelgasse 7, 90402 Nuremberg

[3]Chair of Japanese Studies, Artilleriestr. 70, 91052 Erlangen

{philipp.heinrich, christoph.adrian, olena.kalashnikova, fabian.schaefer, stefan.evert}@fau.de

Abstract

Taking the impact of the Fukushima incident on the global discourse about nuclear energy as a case study, the present paper shows how to integrate computational linguistic methods into corpus-based discourse analysis (CDA). After an extensive literature review with regards to the related hermeneutic work, we present the corpus linguistic methods and point out methodological extensions. These extensions include visualization techniques that might help hermeneutic researchers explore large corpora, and second-order collocates, which help triangulate the semantics of lexical items. In our case study, we firstly give an in-depth analysis of the discourses that have formed around salient lexical items, in particular *nuclear phase-out* and *energy transition*, in the German *Frankfurter Allgemeine Zeitung* (FAZ) and the Japanese *Yomiuri* (both are conservative newspapers of the respective countries). We then provide preliminary results for the impact that the discourse had on German Twitter data. Last but not least, we show what effect the discourse has had on second-order collocates of the lexical item *Germany* in the FAZ corpus.

Keywords: Corpus-Based Discourse Analysis, Fukushima, Computer-Mediated Communication, Visualization, Collocations

1. Introduction

The world is currently witnessing a fundamental transformation of the political public sphere on a global scale. On the one hand, one can observe a powerful process of politicization in the guise of increasingly successful populist and oftentimes anti-European or even anti-democratic parties. On the other, some of the largest global protest movements against nuclear power, free-trade agreements, or social and economic inequality have occurred since the 1970s. In this regard, *Fukushima* represents a prototypical example for the intramedial and transmedial connections between national media discourses, namely in (1) the public spheres as represented in the mass media ("edited mass communication") and (2) the semi-public sphere in social media ("mass self-communication").

We consider the emergence of a global media discourse on renewable and nuclear energy in the aftermath of the Fukushima Daiichi disaster to be an ideal and politically highly relevant case study to study these multiple facets of this transformation of the transnational public sphere, particularly the shift in attitudes and opinions towards nuclear energy. This is not only due to the fact that it was a truly transnational global media event covered by the mass media worldwide (Hartwig et al., 2016; Hayashi, 2013; Ito, 2012; Koch et al., 2016), but also because it was an impetus for the formation of a transnationally networked anti-nuclear social movement (Kobayashi, 2011; Liscutin, 2011; Obinger, 2015; Sano et al., 2012; Slater et al., 2012; Thomson et al., 2012; Tsuda, 2011).

In our research, we analyze the tempo-spatial dissemination and framing of media discourses across different media and languages (i.e. their culture-specific linguistic realizations) in a complex connective network, which enables us to detect distinct distributional and linguistic patterns of the multiple forms of manipulation in the transnational algorithmic

public sphere. In the paper at hand we present preliminary results of our approach to study the multidirectionality and intermedia complexity of the transformation of the transnational public sphere through a transmedial and transnational analysis of the articulated (verbal) connectivities in the media discourse on nuclear energy in Japan and Germany in the mass media and a social media network (Twitter[1]) in the years 2011–2014.

In particular, we will take multiple levels into consideration: (a) the intra-media process of dissemination of statements, attitudes, and opinions framing nuclear energy, (b) the inter-media reciprocities and convergences between social media and the mass media (e.g. trending topics on Twitter becoming news, test-marketing of headlines online before ending up in the printed version of a paper, or the dissemination of journalistic news via Twitter), (c) both of these on a temporal (i.e. 2011–2014) and a transnational axis (i.e. Germany and Japan).[2]

The two countries have been chosen because they have comparable mass media systems (public and private broadcasting, newspapers with a broad spectrum of conservative and liberal papers). Moreover, the two countries were affected by the Fukushima disaster in significantly different ways, which makes them ideal for contrastively studying transnational processes. It is usually argued that the accident turned into a global watershed for domestic atomic en-

---

[1]Twitter has turned into one of the most important tools in election campaigns and political agitation, aside from Facebook (Conway et al., 2015; Davis et al., 2016; Jungherr et al., 2016). Twitter is considered a particularly rich source of social media data because it has several advantages over other social media networks: accessibility, metadata availability, sample size, and the briefness of tweets (Burghardt, 2015; Mejova et al., 2015).

[2]The present work is based on one newspaper from each country as well as German Twitter data. Future research will include more newspapers and Japanese Twitter data.

ergy policies and public attitudes towards the use of nuclear power and the use of renewable energy. However, while in Germany the so-called "Fukushima Effect" has contributed to an immediate phase-out plan, in Japan itself, nuclear energy still remains a core element in the domestic energy mix. Moreover, with Japanese and German, our study covers western and non-western languages, offering not only promising results from linguistic and intercultural comparisons but also poses certain difficulties when studying distant languages in comparison or relation.

2. Related hermeneutic work

The global impact of the second largest nuclear accident in human history is commonly described as the Fukushima Effect in recent publications focusing on the short-term and long-term economic, political, or sociocultural consequences of the disaster (Arlt and Wolling, 2016; Hartwig et al., 2016; Hindmarsh and Priestley, 2016; Wolling and Arlt, 2014; Zeh and Odén, 2014). However, existing studies looking into the so-called Fukushima Effect in the sense of a transnational media discourse (e.g. the coverage and framing of the event itself or detectable changes in opinions and attitudes towards nuclear energy; cf. Scheufele (1999)) remain one-dimensional in terms of the methods or theoretical approaches applied, the types of media taken into consideration, and with regard to geographical regions under study, with a particular bias on Germany and Japan. They analyze attitudes and public opinion either across different media but only for a single national context, or from a transnational perspective but then only with regard to a single medium (usually newspapers).

Moreover, data sets cover a relatively short time period (usually from 4 weeks to at most 12 months after March 2011). Wolling and Arlt (2014) for instance, drawing on an analysis of German mass media coverage and survey data, diagnose a moderate Fukushima effect on Germany's political decision to phase out nuclear energy in summer 2011, arriving at the conclusion that the effect of Fukushima was much more drastic than that of Chernobyl because Japan was considered a high-tech country with tight security standards. Nienierza (2014), as well as Seiffert and Fähnrich (2014) arrive at a similar conclusion based on a content analysis of newspaper articles, and emphasize that nuclear energy had already been framed rather negatively in Germany prior to 2011.

Kepplinger and Lemke (2015) carried out an international and multimedia study of TV news and newspapers (but not social media) in the immediate weeks after the tsunami and meltdown, arguing that news coverage of Japan was related to domestic nuclear energy policy significantly more often in Germany and Switzerland (which opted for a phase-out as well), than in France or the UK, where nuclear energy still remains a core element of the domestic energy mix. Similarly, Hayashi (2013) studied German TV news reporting and found that 40% of the reports about the events in Fukushima were connected to Germany's domestic nuclear policy (cf. Ito (2012) and Honma (2016) for a critical perspective). Abe (2015), based on a content analysis of newspaper editorials from Mar 2011–Dec 2012, divides the Japanese newspaper landscape into a denuclearization

camp (Asahi, Mainichi) and a pro-nuclear policy camp (Sankei, Yomiuri); Yoshino (2013) conducted a comparatively smaller study for Asahi and Yomiuri, but arrives at similar conclusions. Based on content and sentiment analysis of newspaper articles from Mar–Sept 2011, they found that while left-leaning Asahi frequently referred positively to Germany's phase-out, the conservative paper Sankei criticized Germany's dependency on its neighbors to sustain its energy demands in the transition period after the phase-out and emphasized Japan's economic vulnerability.

This bias in the reporting of many newspapers in Japan is also supported by a study by Satoh (2011) (see also Schäfer (2012); Schäfer (2017)), who analyzed the very limited reporting of anti-nuclear demonstrations in Tokyo (Mar–Sept 2011). Hartwig et al. (2016, 114), diagnosing a "gap in studies analyzing international news in Japan concerning Germany's energy policy shift after Fukushima", studied the Japanese mass media coverage (newspapers) of nuclear energy in Germany to understand the potential impact of what Gono'i (2015) has called a "boomerang [Fukushima] effect" on domestic nuclear energy policies in Japan. One might thus argue that the Fukushima Effect in the mass media and on social media worked in at least two directions, namely as a (re-)framing of nuclear energy inside and outside of Japan after the accident on the one hand, and as the reverse effect of this (re-)framing outside of Japan on the Japanese media discourse on the other.
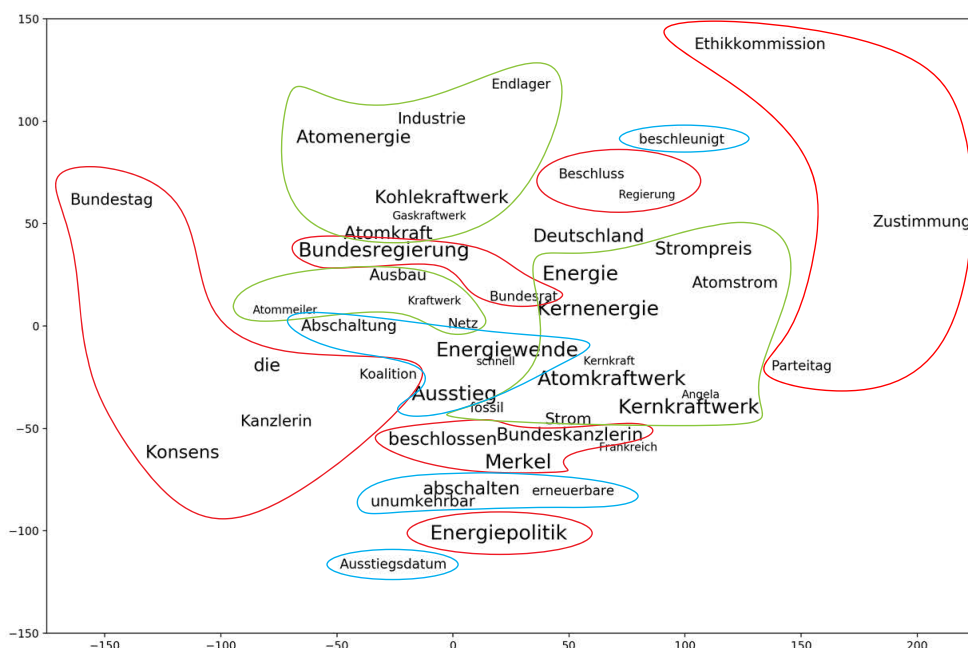
3. Methodology

We touch related methodological work in the present section. However, it should be noted that we are striving for a methodological revision of discourse analysis, which is why there is only a very limited selection of related work at hand.

3.1. Keywords, Collocations, and Discourse

One of our goals is to give an in-depth analysis of the language data produced by mass and social media after the Fukushima incident. We biuld on corpus-based discourse analysis (CDA) (see e. g. Baker (2006); Baker et al. (2008)), which in principal boils down to the aggregation and subsequent deconstruction of concordance lines. The categories in which the textual data is divided in a CDA have to be made up by the hermeneutic researcher while dealing with the data, and cannot be known a priori. This approach thus differs fundamentally from automatic topic modelling or classification into ad hoc categories.

We understand discourses to be formed around lexical items, which we call key words, key items, or simply (discourse) nodes. The collocations of these key items are interpreted as attitudes or stances that can be taken towards the nodes. An example for a discourse is given by Baker (2006, 86): "refugees as victims."[3]

---

[3]Note that key words in this context does not necessarily mean words that have been determined by a keyword analysis in a traditional (corpus) linguistic sense. In fact, keyword analyses and collocation analyses can be translated into one another: a classical keyword analysis, which compares frequency lists of two corpora against one another, is equal to a collocation analysis where the context equals the respective texts of the keyword analysis.

Figure 1: Paragraph-based collocates of *nuclear phase-out* (Atomausstieg) in the FAZ corpus of 2011, automatically arranged by semantic similarity using word embeddings and t-SNE. Size represents association strength to the node word (*nuclear phase-out*). Semantic fields are already loosely clustered together, which makes arrangement easier for hermenuitic interpretation. The post-hoc categories identified by the hermeneutic interpreter are *politics* (red), *actions and challenges of energy transition* (blue), and *economic capital and issues* (green).

In a collocation analysis, the window size (together with the association measure) influences the collocates that can be retrieved. A very conservative window size of, say, 3 will retrieve lexical items closely attributed to the node word, whereas wider window sizes might retrieve connotated items. Besides window-based collocations, we use generalized contexts such as tweets, paragraphs, and whole articles in the present paper. In our analyses, we use the standard log-likelihood measure as implemented in CQP-web (see Hardie (2012) for implementation details).

We do not use automatic Word Sense Disambiguation (WSD) techniques since we are interested in a more fine-grained triangulation of the semantic space: methodologically commited to distributional semantic models, we assume that the semantics of a word depends on the specific contexts at hand. It is thus reasonable and straightforward to look at second-order collocates when diving deeper into the semantics of a word, i.e. to retrieve items that collocate with the lexical node in the context of a discourse (see section 5.). Automatic WSD techniques, in contrast, can only categorize homographs into pre-defined (and very coarse) categories.

Although CDA reduces research bias – since the researcher is confronted with empirical data – this triangulation comes with an intense work load. We therefore actively de-

velop (and use) visualization techniques for semantic fields, which make the results of keyword and collocation analyses (which are essentially n-best lists which have to be sorted by semantic categories) easier to interpret. On that account, we project word vectors (Mikolov et al., 2013) produced by parsing Wikipedia data onto a two-dimensional space using t-distributed stochastic neighbour embedding (t-SNE) (van der Maaten and Hinton, 2008). For an outcome of this process, see figure 1, which displays collocates of the lexical node *nuclear phase-out* in a newspaper corpus. Semantically similar words are automatically grouped together; the hermeneutic interpreter thus can rely on pre-grouped words to form categories.[4]

3.2. Corpora and corpus preparation

We use several newspaper corpora and data from social media, particularly Twitter, in our research. In the present paper, we focus on the *Frankfurter Allgemeine Zeitung* (FAZ) for Germany, and the *Yomiuri* for Japan (we have acquired corpora of both newspapers for the years 2011–2014), both being very conservative newspapers in the respective country. Additionally, we use a sample of 10,266,835 German

---

[4]Note that we can only display words for which vectors have been created. To this end, words that do not appear on Wikipedia frequently enough are not displayed at all.

tweets collected in the years 2011–2014. We also have a sample of roughly 300,000,000 Japanese tweets which we will analyze in further reserach.

The newspaper corpora are pre-processed with standard tools (TreeTagger[5] for German and MeCab[6] for Japanese). The TreeTagger software lemmatizes German tokens; MeCab splits Japanese texts into short-unit morphemes.

Processing computer-mediated communication (CMC) is more difficult: we use specialized software for tokenization (Proisl and Uhrig, 2016) and POS-tagging (Proisl, 2018) of German tweets. Lemmatization of CMC data is work in progress; to this end, we use TreeTagger to lemmatize German CMC. Moreover, building on earlier work, we deduplicate the data, which helps getting rid of social bots and other unwanted noise (Schäfer et al., 2017).

3.3. Measuring Impact

In the paper at hand, we measure the impact of an event (the Fukushima incident) and therefore know a priori what discourses we are looking for (and what kind of reaction we can expect): the incident had an obvious adverse effect on the public stance about nuclear energy, and anti-nuclear movements have thus formed in social media and have been seized by the mass media on the one hand; on the other, social media was influenced by news reports in the mass media; and discourses have subsequently found their way into politics. We hope that once we have analyzed these particular discourses, we can proceed to use the characteristics to give more information to the hermeneutic interpreter.

Note that key words and their collocates are suitable linguistic items for measuring impact: the statistical significance of collocates indicates the salience of the discourse at hand; the propagation of the found discourses (consisting of node words and their respective collocates) through the social network gives insights into the reception among the agents in the network.

4. Case study *Fukushima Effect*

We focus on the key word *nuclear phase-out* (脱原発) for the Japanese Yomiuri in the years 2011 and 2014. We first extract article-based collocations (see section 3.1.) of this keyword in section 4.1. We then show in section 4.2. how the discourse in German mass media was influenced by the Fukushima incident. In particular, we look at paragraph-based collocates of *nuclear phase-out* (Atomausstieg) and the more prevalent discourse around the topic node *energy transition* (Energiewende) in the German corpus of the FAZ.

Section 4.3. concludes the case study with a temporal analysis of the propagation of the lexical items through the German Twitter network. In section 5., we then show how the *nuclear phase-out* debate has spilled over to the distributional semantics of the lexical item *Germany* in the FAZ corpus.

4.1. Analysis of Japanese mass media

In general, the change in discourses in the Yomiuri can be analyzed by performing a keyword analysis of the subcorpus of all the articles of 2014 compared to the subcorpus of 2011 (or vice versa). Such an analysis retrieves obvious items such as the lexical items *Tōhoku earthquake* and *tsunami*, *Fukushima accident*, *evacuation*, *radiation*, and *Khadafi* for 2011, compared to *Abe*, *Olympics* (*Sochi, Brazil*), *Ukraine*, *law-evading drug*, *Nankai megathrust earthquakes*, *consumption tax*, *reinterpretation of Japan's constitution*, *collective self-defense right*, *Mount Ontake eruption*, and *Islam* for 2014. We will leave it to the reader to come up with interpretations for this general keyword analysis.

An article-based collocation analysis of *nuclear phase-out* retrieves insights into the discourse surrounding the topic that we are interested in. In 2011, the top 50 collocates can roughly be categorized in three major semantic fields. Firstly, political actors and politics itself: *Kan* (菅), *Noda* (野田), *prime minister* (首相), the *democratic party* (民主党), *administration* (政権), as well as *policy* (政策), *transition* (転換), and *strategy* (方針). Secondly, the collocates are economic issues connect with the nuclear energy sector: *nuclear energy* (原子力), *power generation* (発電), *electricity* (電力), *operation* (稼働), and *reconstruction* (復興). Thirdly, collocates in the technological category that touch negative aspects of the nuclear energy sector: *safety* (安全), *fuel* (燃料), *shutdown* (停止), and *location* (立地).

In 2014, *nuclear phase-out* collocates on an article-level mostly with elections and politics (*point of dispute* (争点), *claim* (主張), *appeal* (訴える,訴え), *zero* (ゼロ), *policy* (政策), *street speech* (演説, as used in 街頭演説)) and the various political actors *Koizumi* (小泉), *Masuzoe* (舛添), *Hosokawa* (細川), *Abe* (安倍), the *Japan Restoration Party* (維新), and *councillor* (議員). There are fewer words regarding economics in comparison to 2011: *operation* (稼働), *energy* (エネルギー), *power generation* (発電), *electricity* (電力), and the economic policies advocated by Shinzō Abe (*Abenomics*, アベノミクス).

The differences in the *nuclear phase-out* discourse can thus be summarized in the following way: Firstly, the political actors changed. Secondly, not only was the Japanese administration discussed in 2011, but also political decisions of European leaders (Berlusconi, Sarkozy, Merkel, and Schröder). In 2014, political actors and organisations advocating phase-out and criticizing Abenomics appear, and discussions transfer to the street, which is suggested by the collocate *(street) speech* ((街頭) 演説); *nuclear phase-out* is thus used in political campaigns. Among the speakers cited in the Yomiuri are former prime ministers Koizumi and Hosokawa, who became active supporters of nuclear phase-out. Debates in the economic sector were concerned on the restart of nuclear power plants after the inspections and critics towards Abenomics.

4.2. Analysis of German mass media

Table 1 gives an overview of the FAZ in the years 2011 and 2014 as well as the the discourses about *nuclear phase-out* and *energy transition* on a token- and paragraph-level. The salience of the *energy transition* debate has arguably not

restriction	#occurrences	per mill.	#articles	#occ. in par.	#par.	#tokens in par.
nuclear phase-out	1,051	7.05	748	919	844	104,121
– in 2011	525	16.08	409	525	473	57,828
– in 2014	103	2.95	88	97	93	10,753
energy transition	7,166	48.07	3,808	6,354	5,422	621,178
– in 2011	1,474	38.11	840	1,272	1,133	130,319
– in 2014	1,308	41.90	805	1,308	1,129	131,880

Table 1: Sub-corpora restricted to lemma searches for two key words. Left hand side: descriptive figures for the whole corpus (the whole corpus consists of 306,580 articles divided into 1,598,208 paragraphs with 149,058,904 tokens running text). Right hand side: descriptive figures for paragraphs.

much changed between 2011 and 2014: the lexical item *energy transition* (Energiewende) appeared 7,166 in 3,808 different articles in the whole corpus (amounting to 48.07 occurrences per million words); in 2011, it appeared 38.11 times per million words, compared to 41.90 times per million in 2014.[7]

The right hand side of table 1 shows the distribution of the respective items in paragraphs; *nuclear phase-out* e. g. has appeared 919 times in 844 different paragraphs[8], and these paragraphs amount to 104,121 tokens. When looking at the lexical item *nuclear phase-out*, we are diving deeper into a discourse about *energy transition*, since the former is one of the collocates of the latter. Note that the discourse about *nuclear phase-out* had become very prominent in the aftermath of the Fukushima incident in 2011, and had subsided by 2014.

The top 50 paragraph-based collocates of *nuclear phase-out* both in 2011 and in 2014 can roughly be divided into three categories just like in Japan (cf. figure 1 keeping in mind footnote 4): Firstly, political actors (Merkel, *federal chancellor* (Bundeskanzlerin), *federal government* (Bundesregierung)) and political issues such as *ethics commission* (Ethikkommission), *lifetime extension* (Laufzeitverlängerung), and *electricity supply* (Stromversorgung). Secondly, economic actors (German energy producers such as RWE, Eon, and EnBW) and the *price of electricity* (Strompreis). Thirdly, technological issues, such as *nuclear*, *gas-fired*, and *coal power plant* (Gaskraftwerk, Kohlekraftwerk, Atomkraftwerk) as well as *atomic mile* (Atommeiler) and the *electricity grid* (Stromnetz).

The technolocial aspects, political issues and actors have remained the same both in 2011 and 2014, presumably because the issues of a *nuclear phase-out* are associated with specific aspects, actions (for example the shutdown of nuclear power plants), the concept of *energy transition* and the challenges it presents. However, a deeper analysis of the collocates and concordance lines reveals some differences. In 2011, the FAZ mainly covered political actors, the energy industry and issues related to energy policy (*security of supply* (Versorgungsicherheit), *nuclear fuel taxation* (Brennelementesteuer)). Political actors debated about issues relating to the previous decision to prolonging the lifetime of nuclear power plants and the imminent decision to phase out nuclear energy. In 2014, the discourse shifted slightly towards collocates such as *investor protection* (Investorenschutz) and potential *lawsuits* (Klage).[9]

4.3. Analysis of German social media

Figure 2 shows the impact that the Fukushima incident had on discussions in the German Twitter network. Two preliminary notes: firstly, the top panel shows the monthly signal strength, which varies substantially over time. Whereas we have rather good coverage in 2011, we only have roughly one tenth of the data available for subsequent years.

Secondly, whereas the dashed lines show the salience of the topics in terms of a relative proportion of the number of tweets containing the topic node compared to all available tweets, the solid line shows the results when only taking into account originals (i. e. no retweets), which have furthermore been cleansed from any duplicates[10]. The effect of retweets and duplicates (some of this effect presumably being produced by social bots) can be interpreted in terms of "echo chambers", amplifying the original signal. Note that the figure displays a decrease in relative frequency when moving from the dashed to the solid line and is not due to an decrease of data material.[11]

We will restrict ourselves to two conclusions in the scope of this paper: Firstly, the *nuclear phase-out* debate sparked immediately after the incident and subsided rapidly afterwards. Secondly, although the *energy transition* debate also received a stark impetus on March 11, 2011, the debate seems to be independently prominent: it is a sustained topic that is being sparked many times in the aftermath. These increases in relative frequency can be used to extract points in time that indicate external incidents or events.

---

[7]This means that the discourse about this topic was more prominent in the years in between.

[8]The remaining occurrences can be attributed to titles, subtitles, and figure captions.

[9]The story appears similar when looking at article-based collocates; however, as expected, the collocates are more general when retrieved on the basis of articles. They then e. g. comprise the *TTIP*, and the midterm elections in the US. The *nuclear phase-out* debate thus seems to have been covered in the context of other debates.

[10]The deduplication takes place on heavily normalized tweets (Schäfer, 2017).

[11]In the case of the deduplicated data, the number of tweets containing the topic node as well as the denominator ("available tweets") is decreased.
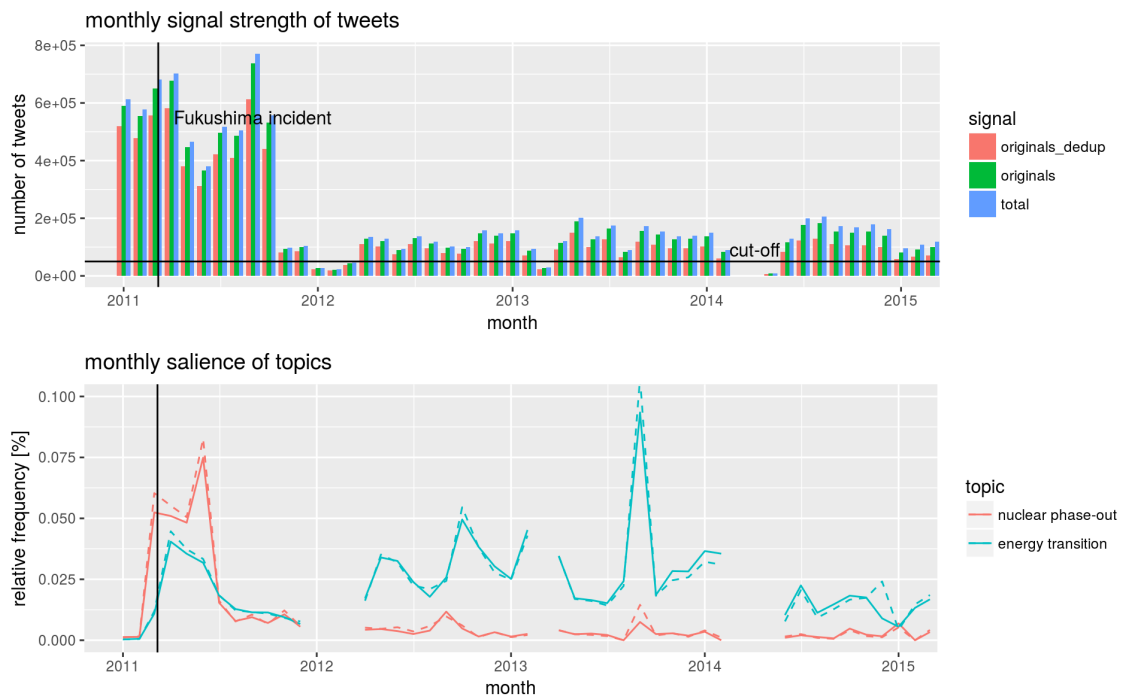
Figure 2: Top panel: Signal strength of our tweet data set (on a monthly basis). For our analyses, we only take into account months for which we have at least 50,000 tweets (as indicated by a horizontal line). Bottom panel: Temporal propagation of key words through the German Twitter network. The vertical line represents the time of the Fukushima incident.

5.   Second-order collocates

We conclude by looking at second-order collocates[12] of the very salient[13] lexical item *Germany* in the FAZ corpus. Figure 3 shows collocates of *Germany* in the whole corpus (left hand panel) and in the sub-corpus of all paragraphs containing the lexical item *energy transition*. Recall that the visual arrangements is due to breaking down word vectors into two dimensions; the salience of the collocate determines its size. Note that the arrangement is the same in both settings (we use the same projection in both cases), but the size and selection of collocates varies.

In the unrestricted case, the most prominent collocates of *Germany* are realted to the Europe and its countries (*federal republik* (Bundesrepublik), *Europe* (Europa), *Austria* (Österreich), *Great Britain* (Großbritannien), and *France* (Frankreich)), as well as *Kabel* and *Alternative* (two very frequent multi-word units are "Kabel Deutschland", the largest cable television provider in Germany, and "Alternative für Deutschland", a right-wing political party that has gained popularity in recent years and has even

made it into the German federal parliament in 2017) and *church* (Kirche), *evangelic* (evangelisch), *Jew* (Jude), *Muslim* (Muslim), and *Turk* (Türke). These collocates are relatively stable over time and can both be found in 2011 and 2014.

In the *energy transition* paragraphs, however, other collocates become salient, such as *future* (Zukunft) and *human being* (Mensch), alongside the more obvious items *energy transition* (Energiewende) *nature conservation* (Naturschutz), *(industrial or business) location* (Standort). What can be seen here is thus a spill-over effect of the energy transition debate into the very meaning of *Germany*, putting more weight on environmental issues and eliminating others such as religious and ethnic communities.

6.   Conclusion and future research

Taking the *Fukushima Effect* as a case study, the present paper has shown how to marry computational linguistic resources with corpus linguistic methods. It points out several improvements over existing corpus linguistic techniques, in particular visualization by means of breaking down word vectors into two dimensions, and the definition of second-order (discourse) collocates. From a hermeneutic stance, it has shed light on the discourse that has been sparked after the Fukushima incident and that has propagated through the social network.

We report work in progress. In a next step, we will analyze further discourse nodes and connect the results from

---

[12]For ease of implementation, we use window-based collocations in the present study. The window-size is set to 5, and log-likelihood has been chosen as association measure.

[13]Albeit we already have a very large corpus of newspapers, looking at second-order collocates is only possible if the number of joint occurrences of discourse node and discourse collocate is high. Infrequent items such as *nuclear phase-out* are thus not yet suitable to be inspected.
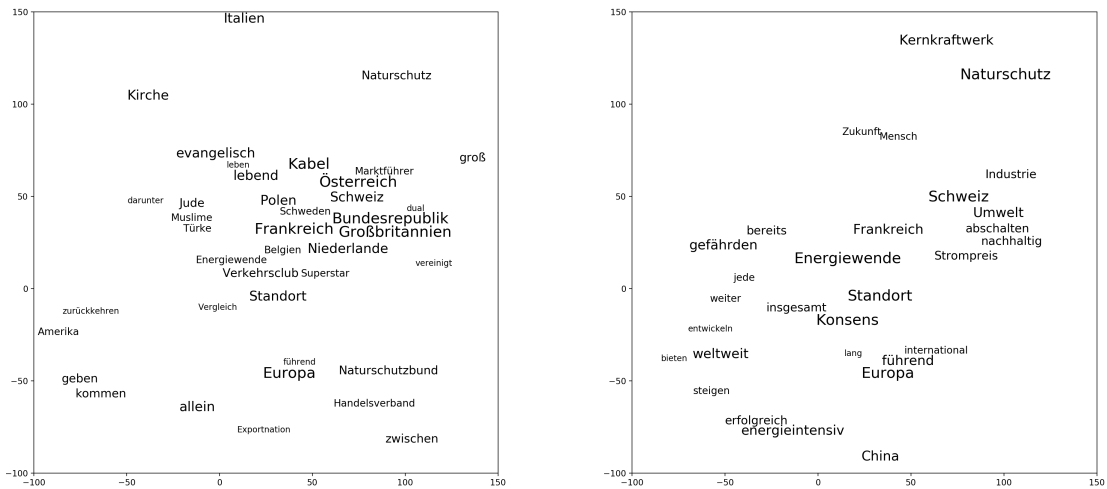
Figure 3: Collocates of the lexical item *Germany* (Deutschland) in the FAZ. Left hand panel: collocates when using the whole corpus as data basis. Right hand panel: second-order-collocates of *Germany* in energy-transition paragraphs.

across German and Japanese newspapers as well as the Twitter network. Further research is additionally necessary to make the methodology consistent (especially with regards to window-, paragraph-, or article-based collocates). Moreover, the Japanese Twitter data will be made usable by means of MeCab with a specialized CMC dictionary (Sato et al., 2017), and specialized word vectors will be created from Japanese Wikipedia and German and Japanese CMC data.

In the long run, we will also extend our data basis to more newspapers (particularly a left-wing newspaper from each country), and we will take into account behavioral operator-based connectivities for Twitter data (namely non-verbal actions like *retweeting* and *replying* to tweets).

7. Bibliographical References

Abe, Y. (2015). The nuclear power debate after Fukushima: a text-mining analysis of Japanese newspapers. *Contemporary Japan*, 27(2), January.

Arlt, D. and Wolling, J. (2016). Fukushima effects in Germany? Changes in media coverage and public opinion on nuclear power. *Public Understanding of Science*, 25(7):842–857, October.

Baker, P., Gabrielatos, C., KhosraviNik, M., Krzyżanowski, M., McEnery, T., and Wodak, R. (2008). A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse & Society*, 19(3):273–306.

Baker, P. (2006). *Using Corpora in Discourse Analysis*. Continuum, London.

Burghardt, M. (2015). Introduction to Tools and Methods for the Analysis of Twitter Data. *10plus1: Living Linguistics*, 1.

Conway, B. A., Kenski, K., and Wang, D. (2015). The Rise of Twitter in the Political Campaign: Searching for Intermedia Agenda-Setting Effects in the Presidential Primary. *Journal of Computer-Mediated Communication*, 20(4):363–380, July.

R. Davis, et al., editors. (2016). *Twitter and elections around the world: Campaigning in 140 characters or less*. Routledge, New York.

Gono'i, I. (2015). 2015-nen ANPO, Minshushugi wo futatabi hajimeru wakamono-tachi (ANPO in 2015. The Youth that is restarting Democracy), September.

Hardie, A. (2012). CQPweb – combining power, flexibility and usability in a corpus analysis tool. *International Journal of Corpus Linguistics*, 17(3):380–409.

Hartwig, M., Okura, S., Tkach-Kawasaki, L., and Kobashi, Y. (2016). Identifying the "Fukushima Effect": Assessing Japanese Mass Media Coverage of International Nuclear Power Decisions. *Journal of International and Advanced Japanese Studies*, 8:109–124.

Hayashi, K. (2013). Kiwadatsu doitsu no genpatsujikohōdō: Fukushima genpatsu jiko no kokusai hikakukenkyū yori: Tokushū 3.11 Fukushima dai'ichi genshiryoku hatsudensho jiko no meguru shakaijōohaō no kenshō: terebi jānarizumu, sōsharu media no tokusei to kadai (Highlighting Germany's nuclear accident news reporting: From an international comparative study on news reporting about the Fukushima accident (Special issue on the verification of social information environment covering the 3.11 Fukushima Dai'ichi Nuclear Power Plant Accident: Characteristics and challenges of TV journalism and social media). *Gakujutsu dōkō*, 18(1):50–55.

R. A. Hindmarsh et al., editors. (2016). *The Fukushima effect: a new geopolitical terrain*. Number 29 in Routledge studies in science, technology and society. Routledge, New York, NY.

Honma, R. (2016). *Genpatsu puropaganda (Nuclear Propaganda)*. Iwanami, Tōkyō.

Ito, H. (2012). Fukushima dai ichi genpatsu jiko ikō no genshiryoku hōdō – jiko-go 3-kagetsu-kan no shinbun shasetsu no ronchō kara miete kuru koto (The Atomic Energy News after the Accident at Fukushima No. 1 Nuclear Power Plant Viewed it from the Tone of Newspapers Editorials during the period Three Months after the Accident). *Pūru gakuin daigaku kenkyū kiyō*, 52:199–212, December.

Jungherr, A., Schoen, H., and Jürgens, P. (2016). The Mediation of Politics through Twitter: An Analysis of Messages posted during the Campaign for the German Federal Election 2013: The Mediation of Politics Through Twitter. *Journal of Computer-Mediated Communication*, 21(1):50–68, January.

Kepplinger, H. M. and Lemke, R. (2015). Instrumentalizing Fukushima: Comparing Media Coverage of Fukushima in Germany, France, the United Kingdom, and Switzerland. *Political Communication*, pages 1–23, June.

Kobayashi, A. (2011). *Saigai to social media: konran, soshite saisei e to michibiku hitobito no 'tsunagari' (The Catastrophe and Social Media: Confusion and the Connectivity that Led to Regeneration)*. Maikomi shinsho, Tōkyō.

Koch, M., Meyer, H., Nishiyama, T., and Zöllner, R. (2016). *Media-Contents und Katastrophen: Beiträge zur medialen Verarbeitung der Großen Ostjapanischen Erdbebenkatastrophe*. Iudicium Verlag.

Liscutin, N. (2011). Indignez-Vous! 'Fukushima,' New Media and Anti-Nuclear Activism in Japan. *Asia-Pacific Journal*, 9(47 (1)), November.

Yelena Mejova, et al., editors. (2015). *Twitter: A Digital Socioscope*. Cambridge University Press, New York.

Mikolov, T., Chen, K., Corrado, G., and Dean, J. (2013). Efficient estimation of word representations in vector space. *CoRR*, abs/1301.3781.

Nienierza, A. (2014). Die größte anzunehmende Umbewertung? Eine Frame-Analyse der deutschen Presseberichterstattung über Kernenergie nach den Reaktorunfällen von Tschernobyl (1986) und Fukushima (2011). In Jens Wolling et al., editors, *Fukushima und die Folgen – Medienberichterstattung, Öffentliche Meinung, Politische Konsequenzen*, pages 31–54. Univ.-Verl., Illmenau.

Obinger, J. (2015). *Alternative Lebensstile und Aktivismus in Japan: der Aufstand der Amateure in Tokyo*. Springer, Wiesbaden.

Proisl, T. and Uhrig, P. (2016). SoMaJo: State-of-the-art tokenization for German web and social media texts. In Paul Cook, et al., editors, *Proceedings of the 10th Web as Corpus Workshop (WAC-X) and the EmpiriST Shared Task*, pages 57–62, Berlin. Association for Computational Linguistics.

Proisl, T. (2018). SoMeWeTa: A Part-of-Speech Tagger for German Social Media and Web Texts. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC'18)*.

Sano, M., Varga, I., Kazama, J., and Torisawa, K. (2012). Requests in Tweets During a Crisis: A Systemic Func-

tional Analysis of Tweets on the Great East Japan Earthquake and the Fukushima Daiichi Nuclear Disaster. In *39th ISFC*, pages 135–140, Sydney, Australia, July.

Sato, T., Hashimoto, T., and Okumura, M. (2017). Implementation of a word segmentation dictionary called mecab-ipadic-NEologd and study on how to use it effectively for information retrieval (in Japanese). In *Proceedings of the Twenty-three Annual Meeting of the Association for Natural Language Processing*, pages NLP2017–B6–1. The Association for Natural Language Processing.

Satoh, K. (2011). What Should the Public Know?: Japanese Media Coverage on the Antinuclear Movement in Tokyo between March 11 and November 30, 2011.

Scheufele, D. (1999). Framing as a theory of media effects. *Journal of Communication*, 49(1):103–122, March.

Schäfer, F., Evert, S., and Heinrich, P. (2017). Japan's 2014 General Election: Political Bots, Right-Wing Internet Activism and PM Abe Shinzō's Hidden Nationalist Agenda. *Big Data*, 5:1 – 16.

Schäfer, F. (2012). Fukushima: Rumours, Emotions and Rousseau's General Will in the Digital Age.

Schäfer, F. (2017). *Medium als Vermittlung: Medien und Medientheorie in Japan. (Medium as Mediation: The Media and Media Theory in Japan*. Springer, Wiesbaden.

Seiffert, J. and Fähnrich, B. (2014). Vertrauensverlust in die Kernenergie. Eine historische Frameanalyse. In Wolling, Jens et al., editors, *Fukushima und die Folgen – Medienberichterstattung, Öffentliche Meinung, Politische Konsequenzen*, pages 55–74. Univ.-Verl., Illmenau.

Slater, D. H., Nishimura, K., and Kindstrand, L. (2012). Social Media, Information and Political Activism in Japan's 3.11 Crisis. *Asia-Pacific Journal*, 10(24 (1)), June.

Thomson, R., Ito, N., Suda, H., Lin, F., Liu, Y., Hayasaka, R., Isochi, R., and Wang, Z. (2012). Trusting Tweets: The Fukushima Disaster and Information Source Credibility on Twitter. In *9th Int. ISCRAM Conf.*, pages 1–10.

Tsuda, D. (2011). Sosharu media wa Tōhoku o saisei kanō ka? Lokaru komyuniti no jiritsu to fukkō (Can Social Media Revitalize the Tōhoku Region? Autonomy and Reconstruction of Local Communities). *Shisō chizu beta*, 2:52–73.

van der Maaten, L. and Hinton, G. (2008). Visualizing High-Dimensional Data Using t-SNE. *Journal of Machine Learning Research*, 9:2579–2605.

Jens Wolling et al., editors. (2014). *Fukushima und die Folgen - Medienberichterstattung, Öffentliche Meinung, Politische Konsequenzen*. Universitätsverlag TU Ilmenau, Ilmenau.

Yoshino, Y. (2013). Datsugenpatsu, hangenpatsuk ōdō ni kansuru shinbunkiji no sōi: Asahi shinbun to yomiuri shinbun (The Differences between Newspaper Articles on the Anti-Nuclear Power Movement. Asahi Shinbun and Yomiuri Shinbun). *Chikushi Jogakuen University*, 8:89–100.

Zeh, R. and Odén, T. (2014). Energieträger in der Berichterstattung. Die Nachwehen von Fukushima in

Schweden und Deutschland. In Jens Wolling et al., editors, *Fukushima und die Folgen – Medienberichterstattung, Öffentliche Meinung, Politische Konsequenzen*, pages 211–232. Universitätsverlag TU Ilmenau, Ilmenau.