

# **Fukushima – Eine quantitativ informierte Diskursanalyse anhand von Sprachgebrauchsmustern**

In der germanistischen Linguistik beginnt man sich in den 90er Jahren zunehmend stärker für den Diskursbegriff von Michel Foucault als eine sprachwissenschaftliche Analysemöglichkeit zu interessieren. Seitdem hat sich viel getan in der Forschung und die Diskurslinguistik ein reiches Arsenal unterschiedlicher Analysemethoden vorgestellt (vgl. Einleitungskapitel im vorliegenden Band). Die meisten Arbeiten, welche im diskurslinguistischen Rahmen zu verorten sind und dabei Korpora als Hilfsmittel zur Erkenntnisgewinnung verwenden, gehen qualitativ bzw. ‚corpus-based‘ vor, d.h., sie haben bestimmte Vorannahmen bzw. Ausgangshypothesen und suchen dann deduktiv bspw. in einer Textsammlung nach relevanten Belegen hierfür. Eine andere Möglichkeit, gemeinsame Strukturen vieler Texte zu finden, ist der ‚corpus-driven‘-Ansatz, bei dem induktiv häufig frequentierte (resp. statistisch signifikante) sprachliche Muster ermittelt werden. Derartige Kollokationen, Ein- bzw. Mehrwort-Einheiten, Keywords oder Cluster etc. bilden dann den Ausgangspunkt für eine qualitativ-interpretative Analyse – die Auswertung schlägt um in eine corpus-based-Perspektive, diesmal jedoch mit einer empirischen Fundierung. An dieser Stelle wird für eine weitere Nutzbarmachung solcher quantitativ-korpuslinguistischer Methoden zur Beantwortung diskurslinguistischer Fragestellungen plädiert.

Die in den Arbeiten von Spitzmüller/Warnke (2011) zusammengefasste Diskurslinguistische Mehr-Ebenen-Analyse (DIMEAN) enthält neben vielen anderen Ansätzen u.a. die Untersuchungskategorien der ‚Einwort‘- wie ‚Mehrwort-Einheiten‘ einer ‚wortorientierten Analyse‘, worunter auch ‚Schlag‘- bzw. ‚Schlüsselwörter‘ aufgelistet werden. Dass von Noah Bubenhofer (2009) dargelegte Konzept der ‚Sprachgebrauchsmuster‘ basiert auf einem frequenzorientierten Kollokationen-Begriff, der in seinen Anfängen zurückgeht auf den britischen Kontextualismus von John R. Firth. Derartige induktiv ermittelte, für einen bestimmten Sprachausschnitt ‚typische‘ Muster stellen eine sehr fruchtbare Operationalisierung der genannten DIMEAN-Untersuchungskategorien dar. Dadurch, dass sie textübergreifend per definitionem eine gleichbleibende Form von Wörtern darstellen, beschränken sie sich jedoch nicht nur auf eine wortorientierte Ebene, sondern werden zugleich einer von der Diskurslinguistik oft geforderten trans- bzw. intertextuellen Betrachtungsweise gerecht. In Sprachgebrauchsmustern kondensieren sich wiederkehrende Sprachhandlungen der Teilnehmer eines Diskurses, deren ganz spezifische Erscheinungsform eine pragmatische Aussagekraft besitzt.

Der vorliegende Aufsatz fasst wesentliche Erkenntnisse meiner im Oktober 2012 eingereichten Magisterarbeit mit dem Originaltitel ‚Fukushima – Eine korpuslinguistische Diskursanalyse anhand von Sprachgebrauchsmustern‘ zusammen. *Fukushima* ist eine verkürzende Bezeichnung (wenn auch gerechtfertigt, wie der spätere Ergebnisteil noch zeigen wird) für den Untersuchungsgegenstand; den deutschsprachigen, printmedialen Diskurs um die desaströsen Ereignisse, welche in Japan mit dem großen Tohoku-Erdbeben vom 11. März 2011 und dem sich anschließenden Tsunami ihren Anfang nahmen. Analysiert wird hier ein Diskurs, welcher als ein nach inhaltlich-semantischen Gesichtspunkten kompiliertes Korpus definiert wird. Es ist hier mithin ein zentrales Anliegen, zu zeigen, dass für einen derartigen, im Vorhinein thematisch festgelegten Diskurs induktiv typische Ein- bzw. Mehrwort-Einheiten gebildet werden können, die sich anschließend als Sprachgebrauchsmuster sinnvoll interpretieren lassen.

## Inhalt

1	Begründung der Methode.....	2
2	Muster in der Sprache .....	7
2.1	Kollokationen, N-Gramme, Mehrwort-Einheiten, ... ..	7
2.2	Schlüssel- und Schlagwörter – über die Typik von Ein- und Mehrwort-Einheiten ..	10
2.3	Log-Likelihood.....	13
3	Korpus und Diskurs .....	14
4	Korpuskompilierung .....	16
5	Ergebnisteil .....	19
5.1	Corpus-driven: Listen der berechneten typischen Ein- und Mehrwort-Einheiten.....	19
5.2	Corpus-based: qualitative Analyse ausgewählter Ergebnisse .....	26
6	Zusammenfassung der Ergebnisse – Hypothesenbildung.....	40
7	Fazit.....	41
	Quellenverzeichnis .....	43

## 1 Begründung der Methode

“[...] by searching for frequent collocations, we can glimpse the recurrent wordings which circulate in the social world, and glimpse how linguistic categories become social categories.” (Stubbs 1996, S. 194)

Die Diskurslinguistik hat es nicht leicht, die Theorie Michel Foucaults für sich nutzbar zu machen.<sup>1</sup> Dies hat damit zu tun, dass Foucault sich in seinen Werken immer wieder einer eindeutigen Festlegung entzieht. Angesichts einer zum Teil verwirrenden Vielfalt dieses Begriffes proklamieren Spitzmüller/Warnke (2011, S. 5) nicht zu Unrecht „die ‚Unordnung‘ des Diskurses“. Ganz allgemein formuliert, bezeichnet der Diskurs das komplexe, interdependente Verhältnis von der Konstitution der Welt *durch*, sowie vom Aufscheinen der Welt(Sicht) *in* Sprache und Wissen, von sozialem Handeln, Machtstrukturen, diachronen u. synchronen sozialen Wertesystemen. Ein solch globaler Ansatz macht es auf der einen Seite schwer, die Theorie in adäquate diskurslinguistische Methoden zu überführen; besteht doch die „Gefahr von Missverständnissen“ – aus diesem Grund bleibt deshalb „stets zu klären, welche Lesart von ‚Diskurs‘ in einem wissenschaftlichen Text gemeint ist“ (Spitzmüller/Warnke 2011, S. 7). Auf der anderen Seite jedoch kann es als eine Chance gesehen werden, ständig neue Methoden für diesen Zweig der Sprachwissenschaft fruchtbar zu machen; lehnte doch Foucault selbst schon das Wort „Lehre“ ab und sah seine Bücher eher als „Einladungen, öffentliche Gesten“ (Foucault 1996, S. 33). So können hier (möglichst gut begründet) folglich jene Aspekte des Diskursbegriffes herausgesucht werden, welche sich im Zusammenhang mit hiesiger linguistischer Untersuchung für eine empirisch-methodische Umsetzung eignen.<sup>2</sup>

---

<sup>1</sup> Zum Interesse am Diskursbegriff in unterschiedlichen Disziplinen der Geistes- und Sozialwissenschaften allgemein vgl. Keller et al. (2011).

<sup>2</sup> In meiner Arbeit demonstriere ich eine Diskursskizzierung mithilfe datengeleiteter korpuslinguistischer Methoden. Dennoch sollten letztere nicht ausschließlich als eine Operationalisierung traditioneller Diskursanalyse missverstanden werden (vgl. Bubenhofer/Scharloth 2013, S. 4).

So bezeichnet Foucault (1973, S. 170) den Diskurs u.a. als eine „Menge an Aussagen“, wobei er wiederum offen lässt, was genau man unter einer ‚Aussage‘ versteht.<sup>3</sup> Vage definiert er diese:

„als ein letztes, unzerlegbares Element, das in sich selbst isoliert werden kann und in ein Spiel von Beziehungen von anderen ihm ähnlichen Elementen eintreten kann. Ein Punkt ohne Oberfläche, der aber in Verteilungsplänen und spezifischen Formen von Gruppierungen ausgemacht werden kann. Ein Korn, das an der Oberfläche eines Gewebes auftaucht, dessen konstitutives Element es ist. Ein Atom des Diskurses.“ (Foucault 1973, S. 117)

Soll der Diskurs untersucht werden, spielen einige grundlegende Leitprinzipien eine wichtige Rolle: „Vier Begriffe müssen demnach der Analyse als regulative Praxis dienen: die Begriffe des Ereignisses, der Serie, der Regelmäßigkeit, der Möglichkeitsbedingung“ (Foucault 1982, S. 37). Das Begriffspaar ‚Ereignis‘|‚Serie‘ wird von Spitzmüller/Warnke (2011, S. 125ff.) in Beziehung mit dem Paar ‚Aussage‘|‚Diskurs‘ gesetzt: Einzelereignisse wie eine konkrete parole, eine sprachliche Erscheinung bestimmen eine ganze Diskursserie, welche selbst wiederum ein Ereignis/einer Aussage Sinn verleiht. Hiermit ist auf die holistische Einheit von Einzelphänomenen und Gesamtzeichensystem hingewiesen (siehe unten). Mit der ‚Regelmäßigkeit‘ ist ein weiterer bedeutsamer Terminus in der geplanten Umsetzung einer Diskursanalyse angeschnitten. Man könnte auch von ‚Musterhaftigkeit‘ sprechen, welche auf die eine oder andere (zu operationalisierende) Art und Weise in einem Diskurs in Erscheinung tritt und beschrieben werden soll:

„In einem weiteren Schritt gilt es, ‚Regelmäßigkeiten‘ (also Muster) zwischen den und innerhalb der Serien zu beschreiben und damit die diskursive Formation zu rekonstruieren. Diese ‚Regelmäßigkeiten‘ sucht Foucault allein in der ‚Positivität‘, also dem materiellen Auftreten von Ereignissen und Serien im Diskurs, nicht ‚hinter‘ dem Diskurs.“<sup>4</sup> (Spitzmüller/Warnke 2011, S. 71)

Dies macht den Musterbegriff (vgl. Bubenhofer 2008, S. 409) generell für den Diskurs sehr bedeutend<sup>5</sup> – und für die wissenschaftlichen Disziplinen im Allgemeinen, wie dieses Zitat nahelegt: „Eine fundamentale Aufgabe jeder Wissenschaft ist die Schaffung einer Ordnung, das Finden von Mustern in der Menge mannigfaltiger, unübersichtlicher Daten.“ (Köhler 2005, S. 4).

Was außerdem wichtig für eine Diskursanalyse ist und eine solche überhaupt erst für die Sprachwissenschaft interessant macht, ist nicht zuletzt die Berücksichtigung der ‚Transtextualität‘; also die Erweiterung der klassischen linguistischen Kategorien Wort – Satz – Text.<sup>6</sup> Die Sprache bildet hierbei die Positivität, in welcher der Diskurs in Erscheinung

---

<sup>3</sup> So erfährt man bei ihm nicht konkret, was eine ‚Aussage‘ ist, sondern was es alles *nicht* ist. Er lehnt es ab, sein Konzept von bekannten linguistischen Kategorien vereinnahmen zu lassen: „[...] die Aussage ist keine Einheit derselben Art wie der Satz, die Proposition oder der Sprechakt; sie gehorcht nicht den gleichen Kriterien; aber sie ist ebenfalls keine Einheit, wie ein materieller Gegenstand es sein könnte, der seine Grenzen und seine Unabhängigkeit besitzt“ (Foucault 1973, S. 126). Vgl. hierzu weiter Foucault (1973, S. 148/149), Spitzmüller/Warnke (2011, S. 70; Fußnote 5) und Jung/Wengeler (1999, S. 146f.).

<sup>4</sup> Martin Wengeler (2003) konzentriert sich in seiner Untersuchung auf die Herausarbeitung von ‚Topoi‘ bzw. ‚Argumentationsmuster‘. So heißt es bei ihm: „Die Analyse von Argumentationen und die Herausarbeitung wiederkehrender Redeweisen und Aussagen gehören also zu den Aufgaben und Möglichkeiten einer diskurs- und mentalitätsgeschichtlich interessierten Sprachgeschichtsschreibung.“ (Wengeler 2003, S. 175). Unabhängig von seiner konkreten Methode, wird hier erneut die Wichtigkeit von ‚Muster‘- bzw. ‚Regelmäßigkeiten‘ deutlich.

<sup>5</sup> Als gemeinsames Ganzes bei all den diversen Bedeutungen des Diskursbegriffes in den verschiedenen geisteswissenschaftlichen Disziplinen heißt es: „Die Begriffsverwendung richtet sich immer auf die Analyse von Sprachgebrauch bzw. von mündlichen oder schriftlichen Texten und untersucht diese im Hinblick auf (formale) Regelstrukturen oder inhaltliche Strukturierungen.“ (Keller et al. 2011, S. 10)

<sup>6</sup> Vgl. hierzu die von Fix et al. (2002) herausgegebene Sammelschrift ‚Brauchen wir einen neuen Textbegriff?‘. Auf diese Preisfrage antwortet Ingo Warnke (2002, S. 125) im gleichen Band, dass man tatsächlich einen neuen Begriff benötige (ohne sich dabei gänzlich vom alten zu verabschieden), nämlich den der ‚Diskursivität‘: „Denn sprachliche Äußerungen erfolgen zwar in Texten, doch sind Texte eben keine isolierten Größen, sondern stehen im Verbund koexistierender Texte, sie sind Teil von Diskursen.“ (Warnke 2002, S. 131). Diskursanalyse ist erst

tritt<sup>7</sup>: „Deshalb sprechen wir auch von ‚transtextueller Sprachanalyse‘.“ (Spitzmüller/Warnke 2011, S. 14). Die einzelnen linguistischen Kategorien sind in diesem Zusammenhang wechselseitig abhängig: sie determinieren die Bedeutung im Diskurs von der sprachlich kleineren zur nächsthöheren Einheit, ‚Aszendenz‘, und von der höheren zu den kleineren Bestandteilen, ‚Deszendenz‘, wie Spitzmüller/Warnke (2011, S. 23f.) es nennen. Das Ganze steht also für seine Teile – und umgekehrt. Darüberhinaus spielt bei einer Diskursanalyse die Berücksichtigung des Kontextes, in dem sich sprachliche Zeichen, z.B. ein ganz bestimmter Begriff bewegt, eine wesentliche Rolle für die Konstitution seiner Bedeutung. Schon in der Darlegung ihres Konzeptes einer Historischen Semantik bzw. Diskursgeschichte stellt für Busse/Teubert (1994, S. 18) die Untersuchung von begrifflicher Ähnlichkeit zwischen Wörtern in unterschiedlich Texten von einem Korpus ein „ergiebige Zugriffsobjekt“ dar:

„Deshalb ist die zweite, wichtigere Grundlage der Diskursanalyse (neben der wortsemantischen Analyse) die textanalytische Erschließung des Sinns, der sich in syntagmatischen Verknüpfungen der Wörter ausdrückt.“ (Busse/Teubert 1994, S. 22)

Alle eben genannten Ausführungen bezüglich des Diskurses und seiner Untersuchung sind auf der einen Seite zwar äußerst punktuell bzw. selektiv und werden der kompletten Bandbreite dieses Begriffes nicht gerecht. Jedoch bilden sie eine Schnittmenge an Vorstellungen, Begriffen und Ansätzen davon, was einen Diskurs ausmacht, welche sich zugleich für eine empirische korpuslinguistische Operationalisierung eignen. Zusammengefasst handelt es sich bei den wesentlichen Charakteristika des Diskurses um eine Menge von ‚Aussagen‘, die regelmäßig, struktur- oder ‚musterhaft‘ auftreten. Sie manifestieren sich als ‚sprachliche Zeichen‘ (dem genuin linguistischen Untersuchungsgegenstand) in einem ganz speziellen, zu berücksichtigenden ‚Kontext‘ und weisen Querverweise bzw. verbindende Elemente zu unterschiedlichen Texten auf. Folglich sind solche Aussagen also auch durch ‚Transtextualität‘ gekennzeichnet.

All jene Eigenschaften werden durch eine korpuslinguistische Lesart sehr gut abgedeckt, welche im Gegensatz zum ‚horizontalen‘ Lesen eines Textes als ein individuelles, abgeschlossenes Einzelereignis stattdessen ‚vertikal‘ nach ‚rekurrenten‘, textübergreifend gleichbleibenden, formalen Mustern Ausschau hält (vgl. Tognini-Bonelli 2001, S. 3).<sup>8</sup> Ein Plädoyer für den stärkeren Einsatz korpuslinguistischer Methoden geschieht nicht zuletzt durch die Einsicht, dass viele diskurslinguistische Untersuchungen nicht als „korpuslinguistisch“ im engeren Sinne bezeichnet werden können, wenn sie auf eine Sammlung von nur wenigen dutzend Texten zurückgreifen<sup>9</sup>, Fragen zur Repräsentativität überhaupt nicht beantworten und die Textsammlung lediglich als einen „Zettelkasten“<sup>10</sup>

---

dann transtextuell, „wenn sie nicht nur einzelne bzw. vereinzelte Texte untersucht – was auch eine traditionelle Aufgabe der Rhetorik und Stilistik oder Textlinguistik sein könnte –, sondern wenn sie eine Mehrzahl, besser: Vielzahl von Texten bzw. Aussagen in verschiedenen Texten, verschiedenen Medien, von verschiedenen Akteuren usw. analysiert, und zwar eine Vielzahl, die strukturelle Übereinstimmungen und Handlungsbezüge aufweist.“ (Spitzmüller/Warnke 2011, S. 187f.).

<sup>7</sup> „Sprache konstituiert sich im Diskurs, sie ist der Diskurs schlechthin.“ (Teubert 2012, S. 258). Vgl. auch Warnke (2007, S. 14) sowie Kap. 3.

<sup>8</sup> Mit der vertikalen Anzeigenform sind konkret ‚Keyword in Context‘-Listen (kurz: ‚KWIC‘, vgl. Lemnitzer/Zinsmeister 2006, S. 89) gemeint, in der alle Vorkommen von einer gesetzten sprachlichen Einheit (z.B. ein bestimmtes Suchwort) in einem Korpus aufgeführt werden. Diese klassischste Darstellungsform der Korpuslinguistik macht sprachliche Beziehungen über die Einheit des Textes hinaus sichtbar (vgl. Bubenhofer/Scharloth 2013, S. 147 sowie Bubenhofer 2009, S. 38).

<sup>9</sup> Vgl. auch Perkuhn (2007, S. 471), der fordert, dass Daten möglichst umfangreich sein sollen – je mehr desto besser, sodass sie „wegen ihrer schieren Masse mit dem Intellekt nicht einmal gesichtet werden könnten“. (Köhler 2005, S. 5)

<sup>10</sup> „Sind Korpora nur Belegsammlungen mit Zettelkästen in elektronischer Form? Mitnichten! In entsprechender Größe [...] und mit den entsprechenden Analysemethoden eröffnen sie eine eigene Perspektive in der linguistischen Forschung – die korpuslinguistische Perspektive.“ (Perkuhn/Belica 2006, S. 2)

verwenden ohne automatische Verfahren zu benutzen (vgl. Spitzmüller/Warnke 2011, S. 35/37). So stellen viele diskurslinguistische Ansätze deduktive Verfahren dar, die von vorgelagerten Überlegungen ausgehen und in entsprechenden Texten nach Belegen hierfür suchen. Bei dieser ‚corpus-based‘-Perspektive respektive „korpusgestützten Ansätzen“ (Lemmitzer/Zinsmeister 2006, S. 37) besteht jedoch die Gefahr, dass – bedingt durch die Suche nach Vordefiniertem – unbekannte Phänomene übersehen werden können (vgl. Bubenhofer 2008, S. 411 sowie Bubenhofer 2009, S. 17/61f./101). Um einer zu starken „hermeneutischen Befangenheit“ entgegenzuwirken, sollte der Diskurs selbst die Richtung vorgeben:

“It is the discourse itself, and not a language-external taxonomy of linguistic entities, which will have to provide the categories and classifications that are needed to answer a given research question. This is the corpus-driven approach.“ (vgl. Teubert 2005, S. 4)

Das sog. corpus-driven-Paradigma betont die Relevanz quantitativer Daten für die (Korpus-)Linguistik, doch in den Sozial- und Kulturwissenschaften allgemein zeichnet sich ein data-driven Turn ab.<sup>11</sup> Die wesentlichen Leitprinzipien eines solch induktiven Ansatzes werden von Tognini-Bonelli (2001, S. 84) zusammengefasst: die ermittelten Daten und Evidenzen werden in ihrem natürlichen Zusammenhang im Korpus als Ganzes betrachtet (nicht als bloße Einzelbelegsammlung), sie geben die Richtung vor und sollen nicht etwa zuvor aufgestellte Hypothesen bestätigen. Rekurrente Muster bzw. häufige Vorkommen von sprachlichen Zeichen bilden dabei die wichtigste Grundlage für eine linguistische Hypothesen- und Kategorienbildung (welche einer quantitativen Analyse nachgelagert zu erfolgen hat; vgl. Steyer/Lauer 2007, S. 494), wobei sowohl die An- als auch die Abwesenheit von Mustern eine Aussagekraft haben. Eine zentrale (diskutierenswerte) Prämisse des corpus-driven-Ansatzes ist, dass eine Relation zwischen Frequenz und Salienz besteht: je häufiger etwas in einer Textsammlung vorkommt, als desto bedeutsamer wird es für den Diskurs angesehen. Solche sich wiederholenden Erscheinungen sind auf der äußersten „sprachlichen Oberfläche“ zu verorten (Bubenhofer/Scharloth 2013, S. 148), was einem Verständnis vom Diskursbegriff aber nicht widerspricht; geht es doch bei der Aussage um „ein Korn, das an der Oberfläche eines Gewebes auftaucht“ (siehe Foucault-Zitat oben).<sup>12</sup> Sprachgebrauchsmuster, nach denen man hier fahndet, sind ein wichtiges Indiz bzw. „Symptome“ für kulturelle oder soziale Phänomene (vgl. Scharloth/Bubenhofer 2012, S. 196). Ganz allgemein bezeichnen sie:

„[...] Einheiten bestehend aus mehreren linguistischen Entitäten wie Wörtern, Wortarten, Lemmata, Tempus- oder Modusinformationen etc., die in einem Korpus musterhaft, also rekurrent, verwendet werden und pragmatisch gedeutet werden können.“ (Bubenhofer/Scharloth 2013, S. 153)

---

<sup>11</sup> Vgl. hierzu den Aufsatz von Scharloth et al. (2013). Sie sehen die Digitale Revolution als einen Auslöser für eine radikale Neuorientierung der Linguistik. Die Analyse und Auswertung riesiger quantitativer Daten erlaubt eine völlig neue Perspektive auf Sprache, der sich eine Diskurslinguistik auf keinen Fall entziehen sollte. Datengeleitete Untersuchungen stellen also nicht nur eine empirische Methode dar, sondern einen neuen Denkstil (vgl. Scharloth et al. 2013, S. 3; Seitenzahl im PDF-Dokument). In einem anderen Aufsatz wurden keine großen Worte gescheut, die neuen digitalen Möglichkeiten der Linguistik mit anderen wissenschaftlichen Untersuchungsinstrumenten zu vergleichen: “Such research is nevertheless an excellent illustration of how the technological tools available to the researcher shape, dictate even, the object of study, as the radio-telescope in astronomy, the particle accelerator in physics and, of course, the concordancer in linguistics [...]” (Partington/Morley 2004, S. 8). Doch ganz gleich, wie man es bezeichnet – das immense Potential solcher sprachlicher Daten muss unbedingt genutzt werden, den andernfalls drohe etwa der Diskurslinguistik eine mögliche Marginalisierung: „[...] dann wird Spitzenforschung im Bereich Sprachanalyse als Gesellschaftsanalyse nicht mehr an Universitäten stattfinden, sondern in privatwirtschaftlichen Unternehmen“ (Scharloth et al. 2013, S. 2)

<sup>12</sup> „Die Korpuspragmatik deutet signifikant häufig auftretende sprachliche Muster in Korpora als Ergebnis rekurrenter Sprachhandlungen der Autorinnen und Autoren der im Korpus enthaltenen Texte bzw. der sie autorisierenden Institutionen und Gruppen. [...] Sie sucht nach pragmatischen Spuren an der (inzwischen rehabilitierten) sprachlichen Oberfläche, nach Mustern, in die sich ein Gebrauchswert eingeschrieben hat.“ (Scharloth/Bubenhofer 2012, S. 196)

Solche hochfrequenten Entitäten; dabei kann es sich um einzelne Wörter (‚Einwort-Einheiten‘, kurz: ‚EWE‘) oder aber auch um mehrere zusammenhängende Wörter (‚Mehrwort-Einheiten‘, kurz: ‚MWE‘) handeln; werden von linguistischen Programmen in großer Zahl berechnet; nicht alle davon können oder lassen sich sinnvoll untersuchen. Erst diejenigen, welche im Nachhinein vom Forscher selektiert, eingegrenzt, systematisiert, ausgewertet, pragmatisch gedeutet, ergo interpretiert werden, erhalten den Status eines Sprachgebrauchsmusters im engeren Sinne (vgl. Bubenhofer 2009, S. 49ff., 122). So ein Interpretationsschritt markiert bereits das Umschlagen von einer corpus-driven-Berechnung hin zu einer corpus-based-Betrachtungsweise. Zahlen alleine haben noch keine Aussagekraft, deswegen sollte bei einer quantitativen korpuslinguistischen Untersuchung stets eine solch „gemischte“ Verfahrensweise praktiziert werden, bei der Induktion und Deduktion zusammenspielen (vgl. Bubenhofer 2009, S. 102ff.). Die Befunde, also die gewonnenen Sprachgebrauchsmuster erlauben eine evidenzbasierte Diskursskizzierung:

„Sie sind zwar weniger Indikatoren für die Wirkkräfte des Diskurses auf das *inhaltlich* Sagbare, dafür zeigen sie an, *wie* etwas gesagt werden muss. Solche Muster sind also Indikatoren für ‚Sprechweise‘ oder die ‚diskursive Praxis‘ (Foucault [1973, S. 276]), die in einem Diskurs gilt.“<sup>13</sup> (Bubenhofer 2008, S. 410)

Der Gedanke, dass durch korpuslinguistische Methoden erhaltene Muster Rückschlüsse auf den Diskurs erlauben, ist aber keineswegs neu. Schon Michael Stubbs sagt:

“Vocabulary and grammar provide us with the potential and resources to say different things. But often this potential is used in regular ways, in large numbers of text, whose patterns therefore embody particular social values and views of the world. Such discourse patterns tell us which meanings are repeatedly expressed in a discourse community.”<sup>14</sup> (Stubbs 1996, S.158)

Spitzmüller/Warneke (2011, S. 37f.) heben die Leistung einer fruchtbaren Annäherung von Diskurs- und quantitativer Korpuslinguistik ausdrücklich hervor, jedoch kritisieren sie die Loslösung von Aussagen bzw. Sprachgebrauchsmustern von der Einheit des Textes und sehen dadurch sowohl die ‚Kontextorientierung‘ als auch die strukturelle Aszendenz vom Text zum Diskurs (siehe oben) als gefährdet an. Diese Befürchtungen können relativiert werden (vgl. Scharloth et al. 2013, S. 5): Mehrworteinheiten erscheinen evtl. fragmentiert, jedoch kann deswegen keinesfalls von fehlender Kontextorientierung gesprochen werden. Nicht nur die heutige Korpuslinguistik verschreibt sich einer kontextsensitiven Analyse von sie interessierenden Begriffen in Form von Konkordanzzeilen. Auch speziell die (in dieser Arbeit) ermittelten Sprachgebrauchsmuster gehen auf einen Kollokationenbegriff zurück, wie ihn schon der ‚britische Kontextualismus‘ (siehe Kap. 2.1) theoretisch begründet hat, und für welchen der syntagmatische Kontext ein eminent wichtiges Kriterium darstellt. Was die Aszendenz speziell vom Text zum Diskurs anbelangt, haben Spitzmüller und Warnke nicht ganz unrecht, dennoch ist hier in jedem Fall eine holistische „Herleitung des Diskurses aus kleineren sprachlichen Einheiten“ (Spitzmüller/Warneke 2011, S. 23) gegeben – sie erfolgt eben direkt von der Wort- bzw. Aussagenebene (den frequenzbasierten Mustern) hin zum Diskurs. Denn auch jener Aspekt trifft eine grundlegende Auffassung der vorliegenden Arbeit, mit der die Intertextualität und deren intensive Reklame für eine Diskursanalyse begründet wird: es liegt in der Natur der berechneten Wortverbindungen, dass sie ihre hohe Auftretenshäufigkeit nur deshalb erhalten, indem sie als zusammenhängende Buchstaben- oder Wortketten textübergreifend gleich bleiben und an den verschiedensten Stellen im Diskurs (bzw. Korpus) in immer dieser gleichen Form reproduziert werden – sie stellen das verbindende Element, die auf äußerster Sprachoberfläche fass- und operationalisierbare intertextuelle Verknüpfung dar.

---

<sup>13</sup> Im späteren Ergebnisteil der hiesigen Arbeit zeigen sich an den gewonnenen Sprachgebrauchsmustern aber nicht nur das *Wie* des untersuchten Diskurses, sondern auch wesentliche *inhaltliche* Themenkreise.

<sup>14</sup> Siehe auch das Eingangszitat von Stubbs, welches diesem Kapitel als Motto vorangestellt wurde.

## 2 Muster in der Sprache

### 2.1 Kollokationen, N-Gramme, Mehrwort-Einheiten, ...

Grundlegend theoretisch mit Sprachgebrauchsmustern verbunden ist der Kollokationenbegriff, wie ihn der ‚britische Kontextualismus‘<sup>15</sup> maßgeblich geprägt hat und mit ihm sein Begründer, der Sprachwissenschaftler John R. Firth und in deren weiterer Tradition sein Schüler M.A.K. Halliday und John Sinclair (vgl. Lehr 1996, S.7). Lange bevor es die Computertechnik und große, maschinenlesbare Korpora sowie den Diskursbegriff von Michel Foucault überhaupt gab, hat Firth bereits 1935 Worte geschrieben, welche das Prinzip heutiger, induktiv arbeitender Korpuslinguistik treffen:

„We must take our facts from speech sequences, verbally complete themselves and operating in contexts of situation which are typical, recurrent, and repeatedly observable. Such contexts of situation should themselves be placed in categories of some sort, sociological and linguistic, within the wider context of culture.“ (Tognini-Bonelli 2001, S. 157 zitiert nach Firth 1957 [1935]<sup>16</sup>, S. 35)

Michael Stubbs (1996, S. 22ff.) charakterisiert die Genese des britischen Kontextualismus als eine von ihren Methoden und Ansätzen her zur Linguistik in der Tradition Noam Chomskys beinahe in direkter Opposition stehende Forschungsperspektive. Interessant im Kontext der verwendeten Methodik in der vorliegenden Arbeit ist ein Zitat von Chomsky, welches Stubbs (1996, S. 6) einer Äußerung von Sinclair gegenüberstellt:

„Like most facts of interest and importance [...] information about the speaker-hearer’s competence [...] is neither presented for direct observation nor extractable from data by inductive procedures of any known sort.“ (Chomsky 1965; S. 18)<sup>17</sup>

Sinclair ist dagegen der Ansicht, dass eine evidenzbasierte, korpuslinguistische Perspektive einen ganz anderen Zugang zur Sprache ermöglicht: „The ability to examine large text corpora in a systematic manner allows access to a quality of evidence that has not been available before“ (Sinclair 1991, S. 4). Zentrale Prinzipien des britischen Kontextualismus sind auch für das Selbstverständnis der vorliegenden Untersuchung bedeutsam.<sup>18</sup> So konstatiert Firth (1957a) in seinem wichtigen Aufsatz ‚Modes of Meaning‘, dass man linguistisch gesehen nie von *der* Bedeutung (z.B. eines Wortes) sprechen kann. Stattdessen

---

<sup>15</sup> Diese Bezeichnung firmiert hauptsächlich in deutschsprachigen Publikationen zum Thema (daneben ist auch gelegentlich vom ‚Firthianismus‘ die Rede). Im angelsächsischen Raum sind hierfür Begriffe wie ‚London School‘, ‚Functionalism‘ oder ‚Firthian Linguistics‘ gebräuchlich. Bei neueren Forschungsperspektiven, welche in deren Tradition stehen, spricht man von ‚Neofirthianismus‘ oder ‚Neo-Firthian Linguistics‘. Vgl. hierzu Lehr (1996, S. 6), die sich in ihre Monografie sehr umfassend mit dem britischen Kontextualismus und dem dort angesiedelten Kollokationenbegriff auseinandersetzt.

<sup>16</sup> Firth, J. R. (1957 [1935]): The Use and Distribution of Certain English Sounds. In: English Studies XVII, 1.

<sup>17</sup> Chomsky, Noam. (1965): Aspects of the Theory of Syntax. Cambridge, MA: MIT Press.

<sup>18</sup> Vgl. Stubbs (1996, S. 23): 1. Linguistik wird als eine ‚Sozialwissenschaft‘ (oder auch ‚Angewandte Wissenschaft‘) verstanden. 2. Sprache sollte stets vorrangig anhand von ‚authentischen Daten‘, also ‚echtem‘ Sprachgebrauch (Texte, Gespräche usw.) studiert werden; nicht mittels erfundener (introspektiv gebildeter) Sätze. 3. Sprache sollte ferner untersucht werden mit Blick auf die Texte als Ganzes (nicht mittels isolierter Fragmente ohne jegliche Kontextualisierung) und darüber hinaus auch textübergreifend in Form von Korpora. 3. Sprachwissenschaft sieht Form und Bedeutung von Sprache als untrennbar bzw. als zwei Seiten einer Medaille an 4. Sprachgebrauch ist sowohl charakterisiert als ein Rückgriff auf Routinen, auf Bekanntes wie auch als ein stetes Kreieren von Neuem. Im Sprachgebrauch kommen kulturell gebundene Eigenheiten und Praktiken zum Ausdruck: „language in use transmits culture“ (Stubbs 1996, S. 23). 5. Die Saussursche Vorstellung von einem streng vom grundlegenden Sprachsystem (‚langue‘) abzutrennenden Sprachgebrauch (‚parole‘), bzw. adäquat hierzu das Chomskysche Begriffspaar ‚Kompetenz‘ – ‚Perfomanz‘, muss revidiert werden. Vgl. hier auch Lehr (1996, S. 10).

verhält es sich mit ihr, wie er sagt, ähnlich dem sichtbaren Licht, welches bei näherer Analyse in ein ganzes Farbspektrum mit je unterschiedlichen Wellenlängen zerfällt.<sup>19</sup> Firth zeigt, dass Wörter nicht zufällig gewählt werden, sondern in ihrer Nähe eine Erwartbarkeit von anderen, ganz bestimmten Worten existiert. Er spricht hier von ‚Kollokationen‘ respektive der ‚kollokativen Bedeutung‘.<sup>20</sup> Auch wenn der (situative) Kontext bei Sprache eine sehr wichtige Rolle spielt, grenzt er hiervon die kollokative Bedeutung ab:

„It must be pointed out that meaning by collocation is not all the same thing as contextual meaning, which is the functional relation of the sentence to the processes of a context of situation in the context of culture.” (Firth 1957a, S. 195)

Das bedeutet: „Kollokationen sind Phänomene des syntagmatischen Kontextes” (Lehr 1996, S. 22). Die Eigenschaft, dass ganz bestimmte Wörter in Gegenwart anderer Wörter auftauchen und damit ihre (kontextuelle) Bedeutung determinieren, fasst Firth mit diesem vielzitierten Satz zusammen: „You shall know a word by the company it keeps“ (Firth 1957b, S. 11). Er blieb allerdings, was eine konkrete Operationalisierung seines Kollokationenbegriffes anbelangt, etwas unklar (vgl. Bubenhofer 2009, S. 112 und Lehr 1996, S. 28). Eine konkretere methodische Weiterentwicklung von Firths Konzept leistet Halliday mit folgender Definition:

„Collocation ist the syntagmatic association of lexical items, quantifiable, textually, as the probability that there will occur, at n removes (a distance of n lexical items) from an item x, the items a, b, c ... Any given item thus enters into a range of collocation, the items with which it is collocated being ranged from more to less probable; and delicacy is increased by the raising of the value of n and by the taking account of an item not only with one other but with two, three or more other items.” (Halliday 2002 [1961], S. 61)

In dieser Begriffsbestimmung sind einige wichtige Punkte enthalten. Zunächst wird auch hier die Kollokation als eine syntagmatische Beziehung von Wörtern bzw. lexikalischen Einheiten verstanden. Dieses Miteinandervorkommen wird jedoch eindeutig als quantifizierbar beschrieben (als ‚rekurrent‘ wie Firth es in seinem frühen Zitat ausdrückte; siehe oben): „damit ist der Grundstein gelegt, um Kollokationen über Auftretensfrequenzen definieren zu können, wie das in der Korpuslinguistik heute üblich ist“ (Bubenhofer 2009, S. 113).<sup>21</sup> Durch

---

<sup>19</sup> Sprachwissenschaft muss dieses Bedeutungsspektrum als Ganzes betrachten und dabei alle beteiligten Ebenen, die einzelnen Bedeutungsarten bzw. -modi abdecken; angefangen mit dem sozialen Kontext, über Syntax und Vokabular bzw. Lexik bis hin zu Phonologie und sogar Phonetik: hier wird die Untrennbarkeit von Form und Bedeutung deutlich (vgl. Firth 1957a, S. 192).

<sup>20</sup> Die deutsche Übersetzung dieses Begriffes von Firth stammt aus Lehr (1996, S. 21). Er erklärt, dass es im Englischen am Beispiel von *ass* nur ganz bestimmte, begrenzte Möglichkeiten von Kollokationen der vorangehenden Adjektive gibt: *silly ass*, *stupid ass*, *awful ass* etc. und dabei in unmittelbarer Umgebung meist Phrasen von persönlicher Adressierung verwendet verwenden wie *you*, *he*, *don't be* (vgl. Firth 1957a, S. 195).

<sup>21</sup> Zum Kollokationenbegriff in der Phraseologie vgl. auch Burger (2010, S. 38), der sie als ‚feste Wortverbindungen‘ definiert (*Zähne putzen* ist fest, weil geläufiger als die paradigmatische Konstellation *Zähne waschen*), deren Sinn ohne Probleme aus ihrem Wortlaut heraus erschlossen werden kann im Gegensatz zu ‚Idiomen‘ und ‚Teil-Idiomen‘ (*Öl ins Feuer gießen*). Für Hausmann (2004, S. 316) sind Kollokationen „grundsätzlich binäre Einheiten“, die in einem Determinationsverhältnis stehen, wobei die Komponenten definiert sind als ‚Basis‘ (*Bett, Zähne, Tisch*) + ‚Kollokator‘ (*machen, putzen, decken*). Den richtigen Kollokator zu einer Basis zu finden stellt z.B. eine Schwierigkeit für DaF-Lerner dar. Deswegen sieht Hausmann den Kollokationenbegriff mit Ausrichtung auf das Fremdsprachenlernen im Kontrast zum „computerlinguistische[n] Kollokationsbegriff, der [...] jede Art von Clusterbildung meint“ (Hausmann 2004, S. 320f.). „Zusammengestellt wurden diese Cluster von dem fleißigen Vollidioten Computer, der selbstredend nicht unterscheidet zwischen Lexem und Morphem und noch weniger zwischen Basis und Kollokator“ (Hausmann 2004, S. 320). Er spricht bei der Besetzung dieses Begriffes gar von einem „Terminologiekrieg“. Vielleicht schrumpft dieser Widerspruch aber, wenn man bedenkt, dass Kollokationen wie *Tisch + decken* etc. auch computer- bzw. korpuslinguistisch ermittelt werden können, weil sie sich eben durch eine höhere Frequenz auszeichnen als andere vergleichbare Wortkombinationen. Ganz allgemein für den Sprachgebrauch und der linguistischen Beschäftigung damit bilden Kollokationen ein vielversprechendes Konzept. So hält auch Burger fest: „Dieser phraseologische Bereich ist weitaus größer und für die Sprachverwendung wohl auch wichtiger, als man es früher angenommen hatte“ (Burger 2010, S. 52).

die Frequenzorientierung sind jene sprachlichen Einheiten nicht kongruent mit konventionellen linguistischen Kategorien.<sup>22</sup> Zum zweiten sagt Halliday, dass ausgehend von einem bestimmten Wort *X* manche Items stärker damit ‚kollokieren‘<sup>23</sup> als andere; es gibt gewisse, auf Wahrscheinlichkeit gründende, graduelle (probabilistische) Unterschiede. Manche Kollokationen sind dann lediglich frequenter als andere. Zum dritten spricht Halliday an, dass mit Kollokationen nicht nur direkt aufeinanderfolgende Token gemeint sind – etwas, das Bubenhofer (2009, S. 117) als ‚kontinuierliche Wortgruppen‘ bezeichnet – sondern sich zwischen den Items ‚eine Distanz von *n* lexikalischen Einheiten‘, ein ‚collocational span‘ oder ‚Kollokationsbereich‘ (Lehr 1996, S. 38) befindet und man so auch ‚diskontinuierliche Wortgruppen‘ (Bubenhofer 2009, S. 117) erhält.<sup>24</sup> Der britische Linguist John Sinclair, welcher wiederum ein Schüler Hallidays war, (vgl. ausführlicher Stubbs 1996, S. 22ff. sowie Sinclair 2004, S. vii) definiert: „Collocation (at present) ist the co-occurrence of words with no more than four intervening words“ (Sinclair 2004, S. 141). Hier wird also eine maximale Distanz der einzelnen lexikalischen Einheiten in einer diskontinuierlichen Wortgruppe festgesetzt.<sup>25</sup> Besteht die Länge solcher Wortgruppen aus einer vorher festgelegten Zahl von Wörtern, so spricht man in diesem Zusammenhang von ‚n-grams‘ bzw. ‚n-Grammen‘.<sup>26</sup> Diskontinuierliche Wortgruppen werden in der englischsprachigen Literatur auch als ‚skipgrams‘ bezeichnet (vgl. Cheng et al. 2006, S. 412f.).<sup>27</sup> Was die englische Terminologie bezüglich frequenzbasierter Wortverbindungen angeht, so gibt es viele Überlappungen weiterer einzelner Begriffe: ‚n-grams‘ werden auch als ‚bundles‘, ‚lexical bundles‘ oder ‚word cluster‘ bezeichnet (vgl. Cheng et al. 2006, S. 412). Biber et al. (2004) konstatieren selbst wiederum eine Begriffsvielfalt:

<sup>22</sup> Vgl. Partington/Morley (2004, S. 1; zitierte Seitenzahl basierend auf dem Online-Dokument). Selten hat man es mit kompletten Sätzen, Teilsätzen, Propositionen, festen Idiomen, inhaltlich abgeschlossen Aussagen oder dergleichen zu tun, was allerdings auch Foucaults Verständnis von einer Aussage entgegenkommt (vgl. Fußnote 2). Kollokationen, Mehrwort-Einheiten – bzw. ihr Pendant in der englischsprachigen Linguistik: ‚lexical bundles‘ (siehe unten) – bilden eine eigene Kategorie: „[...] lexical bundles are usually not complete structural units, but rather fragmented phrases or clauses with new fragments embedded“. [...] The fixedness, however, is different from that of other word combinations, because it is determined by frequency“ (Cortes 2004, S. 400)

<sup>23</sup> Der deutsche Begriff wurde von Lehr (1996, S. 38) übernommen. Sie führt weitere Übersetzungen ein für die einzelnen Bestandteile einer Kollokation. In so einer Wortverbindung geht man meist von einem bestimmten Wort/einer bestimmten lexikalischen Einheit von Interesse aus: dieses wird als ‚nodal item‘ oder kurz ‚node‘ bezeichnet respektive als ‚Kollokant‘. Die anderen Bestandteile, welche zusammen mit dem Kollokanten auftauchen (damit ‚kollokieren‘) werden als ‚Kollokat(e)‘ bezeichnet (vgl. Lehr 1996, S. 37f.).

<sup>24</sup> Ein Beispiel bildet das Wort *sowohl*, welches sehr stark mit *als auch* kollokiert, jedoch zumeist mit ein oder mehreren intervenierenden Wörtern.

<sup>25</sup> Hier taucht der Terminus der ‚co-occurrence‘ oder ‚Kookkurrenz‘ auf, welcher mit dem Begriff der Kollokation gleichgesetzt wird. Obzwar häufig synonym gebraucht, gibt es zwischen beiden von Autor zu Autor mehr oder weniger große Unterschiede (vgl. hierzu Bubenhofer 2009, S. 113ff. sowie auch Lemnitzer/Zinsmeister 2006, S. 196f.). Zentral damit verbunden ist in der deutschsprachigen Linguistik die Kookkurrenzdatenbank von Belica (2001ff.) sowie die ‚statistische Kookkurrenzanalyse‘: „Es geht um die Erfassung von Zeichenketten, die im Vergleich mit ihrem Gesamtvorkommen statistisch überproportional häufig in der Umgebung anderer Zeichenkettenkonfigurationen vorkommen. Die KA erfasst sie und ordnet sie in hierarchischen Kookkurrenzclustern an.“ (Belica/Steyer 2005, S. 8)

<sup>26</sup> Verbindungen aus zwei zusammenhängenden Wörtern würde man bspw. als Bigramme, aus drei Wörtern als Trigramme bezeichnen.

<sup>27</sup> Die Auffassung von Wortverbindungen, welche zwischen den einzelnen Bestandteilen Lücken zulassen und bei denen des Weiteren sogar noch die Reihenfolge der einzelnen Elemente beliebig sein kann, ist die freieste Definition und technische Umsetzung von frequenzbasierten Kollokationen. Für Wendungen, bei denen bspw. zwischen zwei Wortbestandteilen *A* und *B* intervenierende Token *C* auftauchen können, sodass nicht nur Kollokationen *AB*, sondern auch der Form *ACB* möglich sind und ferner durch Vertauschung der Position auch etwa *BA* oder *BCA* gefunden wird, dafür führen Cheng et al. den Terminus ‚congrams‘ (deutsch evtl.: ‚Konkgramme‘) ein (vgl. Cheng et al. 2006, S. 413). Eigens für das Konzept entwickelte diese Arbeitsgruppe (unter Mitwirkung von John Sinclair) das linguistische Programm ‚Concgram 1.0‘<sup>©</sup> (vgl. Greaves 2009, S. 1ff.).

“Multi-word sequences have been studied under many rubrics, including ‘lexical phrases’, ‘formulas’, ‘routines’, ‘fixed expressions’, ‘prefabricated patterns’ (or ‘prefabs’), and ‘lexical bundles’.” (Biber et al. 2004, S. 372)

In diesem Zitat kommt ein allgemeinerer Begriff zur Bezeichnung rekurrenter Wortverbindungen zur Sprache, und zwar ‚multi-word sequences‘; an anderer Stelle wird auch gesprochen von den ‚multi-word units‘; ‚multi-word expressions‘, ‚multi-word collocations‘ usw.<sup>28</sup> Hierfür könnte man den deutschen Terminus ‚Mehrwort-Einheiten‘ verwenden. Wo viele zuvor erwähnte Begriffe wie ‚n-Gramme‘, ‚kontinuierliche Wortgruppen‘ etc. Wortverbindungen mit bereits ganz bestimmten voreingestellten Parametern bezeichnen, hat man mit ‚Mehrworteinheit‘ „einen flexibleren Kollokationen-Begriff zur Verfügung“ (Bubenhofer 2009, S. 122).

Aus diesem Grund spreche ich auch in meiner Arbeit größtenteils von ‚Mehrwort-Einheiten‘; und um die gewonnenen frequenten Wortverbindungen in eine begriffliche Nähe zu den diskurslinguistischen Analysekatoren des DIMEAN-Modells von Spitzmüller/Warneke (2011) zu rücken. Synonym wird in der Ergebnisauswertung auch von ‚Cluster‘ gesprochen, da das Programm WordSmith Tools, mit welchem hier gearbeitet wurde (vgl. Scott 2008), diesen Ausdruck verwendet (vgl. Scott/Tribble 2006, S. 11f.). Es wurden in meiner Arbeit ausschließlich ‚kontinuierliche Wortgruppen‘ berechnet, ohne einen collocational span. In den allermeisten Fällen bestehen diese Verbindungen aus drei zusammenhängenden Wörtern, wobei dann von ‚Dreiwort-Einheit‘ geredet wird.

Es gibt eine riesige Fülle an berechneten MWE; diejenigen davon, welche für eine Interpretation ausgesucht, in ihrem inhaltlichen Kontext und hinsichtlich möglicher pragmatischer Funktionen beschrieben werden, erhalten den Status ‚Sprachgebrauchsmuster‘ (kurz: ‚SGM‘). Bei ‚diskontinuierlichen Mehrworteinheiten‘ ist dieser interpretatorische Schritt schwieriger zu vollziehen, da viele syntaktisch variable und inhaltlich folglich uneinheitliche MWE errechnet werden (vgl. Bubenhofer 2009, S. 204). Bei kontinuierlichen MWE wie in dieser Arbeit lassen sich die meisten davon jedoch sinnvoll interpretieren. Die Gesamtheit aller SGM und ihrer beschriebenen Phänomene wird als eine Diskursskizzierung von ‚Fukushima‘ verstanden. Ein Kriterium, das dabei die Bildung von Mehrwort-Einheiten auszeichnet, ist, wie dargelegt wurde, ihre hohe Frequenz; ein weiteres betrifft ihre ‚Typik‘. Darunter versteht man, dass Sprachgebrauchsmuster „häufig in Bezug auf bestimmte Sprachausschnitte“ sind (vgl. Bubenhofer 2008, S. 408), welche wiederum durch Vergleichskorpora repräsentiert werden. Im Folgenden wird daher theoretisch-methodisch dargelegt, wie solche typischen Mehrwort- und auch Einwort-Einheiten definiert bzw. ermittelt werden und welchen Sinn sie haben.

## **2.2 Schlüssel- und Schlagwörter – über die Typik von Ein- und Mehrwort-Einheiten**

In ihrer komplexen Diskurslinguistischen Mehrebenen-Analyse (DIMEAN) listen Spitzmüller/Warneke (2011, S. 140ff.) unter der intratextuellen Ebene bei den Ein- wie Mehrworteinheiten u.a. auch die sog. ‚Schlagwörter‘ und ‚Schlüsselwörter‘ auf. Dies ist ein wichtiger Aspekt, denn für eine wortorientierte Analyse des Fukushima-Diskurses sollten seine ihm eigenen, spezifischen sprachlichen Muster gesucht werden. Doch bleibt zu fragen, wann man von einem ‚Schlüsselwort‘ (oder ‚Schlüsselwortverbindung‘) spricht. Bei der intuitiven Verwendung dieses Begriffs gibt es meist keine Probleme; es scheint klar, dass es sich um ein wichtiges Wort, um eines mit zentraler Bedeutung handeln muss, welches besonders kennzeichnend oder charakteristisch für einen bestimmten Sachverhalt ist. Wissenschaftlich exakter versucht die Sprachwissenschaft auf diesen Terminus zu schauen.

---

<sup>28</sup> Vgl. etwa Oakes (1998, S. 184) oder Tognini-Bonelli (2001, S. 19).

Entscheidend für eine (diskurs-)linguistische Konzeptionierung des Begriffs ‚Schlüsselwort‘ sind die Arbeiten von Liebert (1994) und Hermanns (1994), auf die sich Spitzmüller/Warnke (2011, S. 143) in ihren Ausführungen zum Gegenstand ebenfalls beziehen. Fritz Hermanns führt bei einer Rekapitulation verschiedener Autoren, welche sich mit dem Konzept ‚Schlüssel- oder ‚Schlagwort‘ in einem stark gesellschaftlichen, insbesondere politischen Kontext befassen, einige von deren Eigenschaften auf. U.a. sind sie Elemente der Öffentlichkeitssprache, in ihnen sind bestimmte (politisch-ideologische) Programme verdichtet, sie weisen eine starke Affektbesetzung auf, d.h., sie haben entweder eine stark positive oder negative Wertung<sup>29</sup> und sie zeichnen sich durch temporäre Gebrauchshäufigkeit aus (vgl. Hermanns 1994, S. 11). Während Schlagwörter durch den Gebrauch sozialer Akteure gekennzeichnet sind, grenzt Hermanns hiervon das ‚Schlüsselwort‘ ab, welcher „ein Terminus aus der Betrachterperspektive ist“ (Hermanns 1994, S. 43) und Bedeutung für die Betrachter der Verwendung hat, um das Verständnis „aufzuschließen“ – ergo also für Historiker oder Linguisten. Die Überschneidungen mit dem Konzept ‚Schlagwort‘ sind aber sehr eng. Mit Bezug auf Firths „focal words“ (im Sinne von ‚fokal‘ bzw. ‚Fokus‘) arbeitet Liebert heraus, dass Schlüsselwörter immer auch sog. ‚Brennpunktörter‘ sind, „in denen die Kommunikationsaktivitäten immer wieder zusammentreffen, und die damit zu den lexikalischen Brennpunkten eines Kommunikationsprozesses werden“ (Liebert 1994, S. 10). Spitzmüller/Warnke (2011, S. 143) fassen nach Liebert<sup>30</sup> mehrere Merkmale von Schlüsselwörtern zusammen; u.a. sollen sie diskursbestimmend sein sowie das Selbstverständnis oder Ideale einer Gruppe oder ganzen Epoche zum Ausdruck bringen. Es ist klar, dass Schlag- und/oder Schlüsselwörter – gleichwohl schwer voneinander zu trennen (vgl. Spitzmüller/Warnke 2011, S. 144) – sehr eng mit einer diskurslinguistischen Analyse verbunden sind; sie „haben in Diskursen – und nur in Diskursen – ihren Ort und ihren Sinn“ (Hermanns 1994, S. 53). All die oben angesprochenen Eigenheiten derartiger Begriffe zielen oft auf die Pragmatik ab – und es wird im Laufe der Ergebnisanalyse dieser Arbeit sicher noch deutlich, dass sich in solchen Wörtern tatsächlich Programme verdichten und sie oft eine z.T. überwiegend positive oder aber vorwiegend negative konnotative Bedeutung haben. Alle eben skizzierten Charakteristika von Schlüssel- bzw. Schlagwörtern deuten zugleich ihre in hiesiger Arbeit erfolgte methodische Operationalisierung an. Letztere kondensiert sich als Schnittmenge obiger Aspekte wie etwa „temporäre Gebrauchshäufigkeit“, „diskursbestimmend“, „Brennpunktörter“. In (Schlüssel-)Wörtern können nur „Kommunikationsaktivität[en] immer wieder zusammentreffen“, wenn sie sich durch eine weitere Gemeinsamkeit auszeichnen, und zwar die der hohen Auftretensfrequenz.

Doch das allein reicht nicht aus. Man kann diesen Umstand sehr gut an sog. ‚Wortformenlisten‘ respektive ‚wordlists‘ erkennen, wie sie von korpuslinguistischer Software ausgegeben werden können. Darunter versteht man eine Liste von sämtlichen, in einem Text (in diesem Fall einem Korpus) vorkommenden ‚Types‘ und ‚Tokens‘<sup>31</sup>; man

---

<sup>29</sup> Positiv konnotierte Schlagwörter nennt Hermanns z.B. ‚Hochwertwörter‘ oder ‚Fahnenwörter‘ (die sich bestimmte Parteien oder Gruppierungen „auf ihre Fahne geschrieben haben“) und Schlagwörter mit negativer Bedeutung bezeichnet er etwa als ‚Unwertwörter‘ oder ‚Stigmawörter‘.

<sup>30</sup> Liebert, Wolf-Andreas (2003): Zu einem dynamischen Konzept von Schlüsselwörtern. Eine exemplarische Analyse am Beispiel Globalisierung. In: Zeitschrift für Angewandte Linguistik 38, S. 57–83.

<sup>31</sup> Vgl. zur Begrifflichkeit von ‚Type‘ und ‚Token‘ auch Lemnitzer/Zinsmeister (2006, S. 198). Unter einem ‚Worttype‘ oder ‚Type‘ wird eine nach festgelegten Kriterien bezüglich Ähnlichkeit oder Gleichheit zusammengefasste ‚Wortform‘ verstanden. Mit ‚Token‘ ist dann das konkrete Einzelvorkommen dieses Types an einer bestimmten Stelle im Korpus gemeint. Was man genau unter Types versteht, ist eine Sache der vorher festgelegten Definition, welche bei korpuslinguistischen, datengetriebenen Auswertungen eine nicht unerhebliche Rolle spielt. So stellen sich bspw. die Fragen, ob und aus wievielen Types Eigennamen und Bezeichnungen bestehen wie etwa *Peter*, *Kölner Dom*, *SPD*, *Hamburg*, *Angela Merkel* usw.; ob Groß- und Kleinschreibung beachtet wird; wie Zahlen behandelt werden wie *2012*, *365*, *100 Millionen* etc., welche auch sehr häufig in Texten vorkommen. Vgl. zu den Problemen von computerlinguistischer Type-Klassifikation

könnte hier von frequenzbasierten Einworteinheiten bzw. ‚Unigrammen‘ sprechen (in Anlehnung an den ‚n-grams‘-Begriff). Im hier untersuchten Fukushima-Korpus (siehe Kap. 4) kommen Wortformen wie *die*, *der*, *und*, *in* etc. am häufigsten vor<sup>32</sup>, denen jedoch nicht zwangsläufig eine sehr wichtige Schlüsselwortfunktion zugesprochen werden kann. Entscheidend ist, dass ein Type ein quantitativ großes Vorkommen hat im Vergleich zu einer anderen Entität. Denn ein Schlüsselwort hat immer nur eine gewisse Gebrauchshäufigkeit für ganz bestimmte Gesellschaften, Gruppen, Zeiten, Epochen, geografische Räume, Themen, Texte, Textsorten u.v.m., kurz: für einen ganz spezifischen Sprachausschnitt. Um generell die Eigenschaft von etwas zu definieren, kommt man kaum umhin, diese durch den Vergleich mit etwas Andersartigem zu gewinnen. Übertragen auf die Korpuslinguistik formuliert Alan Partington dies so:

“In some sense, all work with corpora is properly comparative. Even when a single corpus is employed, it is used to test the data it contains against another body of data. [...] Corpus-assisted studies of register, genre or discourse type are of course by definition comparative: it is only possible to both uncover and evaluate the particular features of a discourse type by comparing it with others.” (Partington 2006, S. 269)

Schlüsselwörter, so wie sie in hiesiger Arbeit verstanden werden, erhält man, indem die zu analysierende Textsammlung, hier also das ‚Spezialkorpus‘ zum Fukushima-Diskurs, einer zum Vergleich herangezogenen Textsammlung, einem ‚Referenzkorpus‘ (kurz: ‚RK‘) gegenübergestellt wird.<sup>33</sup> Für jede Wortform wird jeweils ihre Frequenz im hier bezeichneten ‚Fukushima-Korpus‘ (kurz: ‚FK‘) und im RK berechnet. Nicht die absolute Auftretenshäufigkeit eines bestimmten Types ist hierbei ausschlaggebend, sondern sein Frequenzunterschied im Kontrast mit dem Vergleichskorpus. Diese ermittelten Wörter bezeichnen Scott/Tribble nun als ‚keywords‘ – sie stellen die sog. ‚aboutness‘ (vgl. Scott/Tribble 2006, S. 58 und Baker 2004, S. 347) eines Textes oder einer Textsammlung dar. Im Folgenden wird hier für ‚keywords‘ die deutsche Übersetzung ‚Kenn-Wort‘ (kurz: ‚KW‘) verwendet.<sup>34</sup> Synonym wird auch von ‚typischen Einwort-Einheiten‘ gesprochen, wobei in hiesiger Untersuchung gleichermaßen typische Mehrwort-Einheiten gebildet werden, welche auch als ‚key cluster‘ (vgl. Partington/Morley 2004, S. 2f.) bzw. ‚KeyWords Cluster‘<sup>35</sup> (vgl. Scott 2008, WordSmith Tools Manual) bezeichnet werden.

---

Scott/Tribble (2006, 14f.). Der Wortformenbegriff, auf den die vorliegende Arbeit fußt, wird von WordSmith Tools auf eine sehr pragmatische Art und Weise definiert: Types sind in einer Textdatei grafisch gleiche Worteinheiten, welche nach links und rechts von u.a. Leerezeichen, Zeilenumbrüchen, Interpunktion (wie Komma, Satzpunkt, Ausrufezeichen, Klammern usw.) begrenzt werden. Um einmal bei den oben genannten Beispielen zu bleiben: das Programm zählt *suchte*, *gesucht* und *suche* (keine Lemmata), als drei verschiedene Types, *SPD* wird als eine Wortform gezählt; *Angela Merkel* dagegen besteht aus den zwei separat aufgeführten Wortformen *Angela* und *Merkel* und alle Zahlen bzw. Ziffern (*2012*, *365*, *100*, ...) werden unter dem gemeinsamen Type # (Hashtag) subsumiert.

<sup>32</sup> Es fällt auf, dass die frequentesten Types keine speziellen lexikalischen Einheiten sind, sondern grammatische Wörter (Präpositionen, Konjunktionen, Pronomen), die als eine Art Kitt zwischen den Elementen eines Textes fungieren. So ein Muster ist charakteristisch für genügend große Korpora (vgl. Scott/Tribble 2006, S. 23f.).

<sup>33</sup> Vgl. für ein einfaches Beispiel Scott/Tribble (2006, S. 58): häufig vorkommende Types in einem Kuchenrezept sind sehr wahrscheinlich grammatische Wörter wie *die*, *der*, *in*. Schlüsselwörter wie *Zucker*, *Mehl*, *Eier* erhält man durch die Kontrastierung mit einem anderen Text, welcher kein Kuchenrezept darstellt.

<sup>34</sup> Die Bezeichnung ‚Kenn-Wort‘ orientiert sich vom Wortlaut her nah am englischen Terminus ‚keyword‘, welcher auf ein spezielles technisches Verfahren verweist, das als eine methodische Operationalisierung des (linguistischen) Begriffs ‚Schlüsselwort‘ verstanden wird. Ein Kenn-Wort ist also im wörtlichen Sinne ein Type, welcher für einen bestimmten Sprachausschnitt *kennzeichnend* ist, wobei das Basismorphem *kenn* hier auf die ‚keyness‘ verweist (vgl. Kap. 2.3). Die Übersetzung in Form eines Bindestrich-Kompositums geschieht auch in Abgrenzung zum umgangssprachlich verwendeten Ausdruck *Kennwort* (im Sinne von Passwort).

<sup>35</sup> Als eine dem englischen Wortlaut nahe Übersetzung schlage ich ‚Kenn-Wortverbindung‘ vor (vgl. Fußnote 34).

Diese Definition erlaubt eine objektivere Sicht auf das Konzept von Schlüsselwörtern, die weniger anfällig ist für eine (kulturell bedingte) subjektive Einschätzung bezüglich der „Wichtigkeit“ bestimmter Begriffe (vgl. Baker 2004, S. 347). Zumeist wird diese Methode gebraucht, um bestimmte Textsorten, Genre oder Register auf ihre sprachlichen Eigenheiten hin zu untersuchen (vgl. Baker 2004, S. 347; Tribble 2000). Unabhängig von dem Einsatz zur Stilbeschreibung bestimmter sprachlicher Register kann jedoch darüber hinaus das Verfahren auch für diskursanalytische Untersuchungen angewendet werden, wie Baker explizit betont:

“Keywords can also be useful in helping to spot traces of discourse within language. While the term *discourse* has multiple meanings, I use it here to refer to a ‘system of statements which constructs an object’ (Parker 1992, 5)<sup>36</sup>.” (Baker 2004, S. 347)

Die eben dargestellten Überlegungen zum Konzept von Kenn-Wörtern machen eine für vorliegende Untersuchung wichtige Forschungsprämisse deutlich, die wiederum Baker auf den Punkt bringt, wenn er sagt:

„Therefore, a keyword analysis will focus only on lexical differences, not lexical similarities. Such a feature of WordSmith is not necessarily problematic, depending on the researcher’s focus, but it may result in the researcher making claims about differences while neglecting similarities to the point that differences are overemphasized.” (Baker 2004, S. 349)

Dem Unterschied wird größere Bedeutsamkeit zuteil als der Ähnlichkeit bestimmter sprachlicher Phänomene. Diese mögliche Gefahr einer Überbewertung solcher auf Unterschiede beruhenden Ergebnisse kann nicht stark genug betont werden, denn natürlich interessiert sich eine Diskursanalyse auch für Gemeinsamkeiten im Gebrauch spezieller Wörter in unterschiedlichen Diskursen.<sup>37</sup>

### 2.3 Log-Likelihood

Ein Kenn-Wort konkret berechnet wird, indem das relative Vorkommen eines bestimmten Wortes zur jeweiligen Gesamt-Korpusgröße einmal im Fukushima-Korpus und zum anderen im Referenzkorpus zueinander ins Verhältnis gesetzt wird. Die Beantwortung der Frage, ob es einen signifikanten Unterschied zwischen der Frequenz einer Wortform in den beiden Korpora gibt, ob also die Nullhypothese verworfen und von der Richtigkeit der Alternativhypothese ausgegangen wird<sup>38</sup>, ist bei Sprachdaten jedoch zweitrangig. Denn ganz gleich, welchen Type man in zwei Textsammlungen auf seine Häufigkeit hin analysiert – es existiert beinahe immer ein statistischer Unterschied.<sup>39</sup>

---

<sup>36</sup> Parker, Ian (1992): *Discourse Dynamics: Critical Analysis for Social and Individual Psychology*. London: Routledge.

<sup>37</sup> Das gilt sowohl für die Diskurs- wie auch für die Korpuslinguistik: „Two salient questions are ‘how similar are two corpora’, and ‘in what ways do two corpora differ?’“ (Kilgarriff 2001, S. 233)

<sup>38</sup> Vgl. allgemein Oakes (1998, S. 9). Man stellt die Nullhypothese  $H_0$  auf, welche davon ausgeht, dass es keinen Unterschied zwischen den beiden Token-Werten eines bestimmten Types in den beiden Korpora (FK + RK) gibt, respektive dass diese Unterschiede so klein sind, dass sie sich im Rahmen zufälliger Schwankungen ergeben. Zumeist geht das Untersuchungsdesign des Forschers davon aus, diese Behauptung zu verwerfen und die Richtigkeit der sog. Alternativhypothese  $H_1$  anzunehmen. Letztere besagt, dass doch ein signifikanter Unterschied zwischen diesen beiden Werten existiert, die Zugehörigkeit zu einem Korpus folglich sehr wohl Einfluss auf die Frequenz des Types hat. Bei der in meiner Untersuchung angewandten Methode werden hierbei alle im Fukushima-Korpus auftretenden Wortformen auf ihr Vorkommen im Referenzkorpus überprüft.

<sup>39</sup> Vgl. Kilgarriff (2001, S. 236), der ein offensichtliches Faktum von Sprache enthüllt; nämlich dass Wörter niemals zufällig verteilt sind, sondern in „Rudeln“ auftreten, sich zu „Klumpen“ (Bubenhofner 2009, S. 134) zusammenballen. Oder wie Church/Gale (1995, S.8) es mit einem schönen Vergleich deutlich machen: anders als ein Blitz, „schlägt“ ein Wort oft mehrmals an der gleichen Stelle ein. So kommt beispielsweise das Wort

Für das Testen von Unterschieden bei den Frequenzen können viele „klassische“ statistische Testverfahren nicht angewendet werden, weil sie üblicherweise eine Normalverteilung voraussetzen, welche bei Sprachdaten meist nicht gegeben ist (vgl. Köhler 2005, S. 10ff.). Ted Dunning (1993) schlägt daher den Log-Likelihood- bzw.  $G^2$ -Test vor.<sup>40</sup> Dieser Wert wird für jeden Type im FK berechnet<sup>41</sup>, wobei im Anschluss die einzelnen Wortformen in einer Liste („Typikprofil“ vgl. Bubenhofer 2009, S. 152), beginnend mit den höchsten  $G^2$ -Werten, sortiert werden. Diese Art Rangfolge der berechneten Ergebnisse interessiert hier mehr als das Testen von Hypothesen: die Wortformen mit den höchsten  $G^2$ -Werten sind hierbei jene, welche am meisten destinktiv bzw. „charakteristisch“ sind im Vergleich zu ihrem Vorkommen im Referenzkorpus (vgl. Kilgarriff 2001, S. 237; Rayson/Garside 2000, S. 3). Lediglich diese Types werden in meiner Untersuchung eingehender betrachtet, da sie, aufgefasst als die „dominantesten“ Schlag- bzw. Schlüsselwörter, die wichtigsten Inhalte des Fukushima-Diskurses repräsentieren. Auch werden hier typische Mehrwort-Einheiten errechnet, indem exakt das gleiche statistische Verfahren eingesetzt wird wie bei der Bildung der Kenn-Wörter.<sup>42</sup>

### 3 Korpus und Diskurs

Der ‚Diskurs‘ ist eine unerschöpfliche Größe. Schon Foucault ließ die Definition des Diskursbegriffes möglichst offen (siehe Kap. 1), versuchte einer Vorfestlegung auf ganz bestimmte Kategorien aus dem Weg zu gehen und wollte, dass man „die Instanz des Aussagenereignisses isoliert“ betrachten soll. Dies, so sagte er, geschieht nicht,

„um eine zahllose Menge von Fakten zu streuen. Es geschieht, um sicher zu sein, sie nicht auf Verfahren der Synthese zu beziehen, die rein psychologischer Natur wären (die Absicht des Autors, die Form seines Geistes, die Strenge seines Denkens, die ihn beschäftigenden Themen, das Vorhaben, das seine Existenz durchläuft und ihr Bedeutung gibt und um andere Formen der Regelmäßigkeit, andere Typen der Beziehung erfassen zu können“ (Foucault 1973, S. 44).

Wann immer ein Text entsteht, eine Aussage getätigt oder ein Satz ausgesprochen wird, sind diese sprachlichen Artefakte in einem praktisch unendlichen Geflecht mit anderen Äußerungen verbunden, sowohl synchron wie auch diachron. Speziell in der Sprachwissenschaft wird die Untersuchung des Diskurses wie erläutert sehr global angesetzt: „Diskurslinguistik nach Foucault analysiert in unserem Verständnis Sprache, Wissen und Akteure“ (Spitzmüller/Warnke 2011, S. 17). Nicht lediglich Sprache allein, ob gesprochen

---

*Diskurs* in vielen diversen Texten möglicherweise überhaupt nicht vor; um dann, wenn es sich etwa um einen diskurslinguistischen Text handelt, plötzlich sehr gehäuft aufzutreten.

<sup>40</sup> Rayson/Garside (2000, S.2f.) sehen  $G^2$  als eine gute Möglichkeit, um (zwei) Korpora miteinander zu vergleichen; im hier verwendeten WordSmith Tools wird diese Methode als Auswahlmöglichkeit für die Ermittlung von ‚Keyness‘ angeboten (vgl. Scott 2008, WordSmith Tools Manual); Bubenhofer (2009, S. 146f.) verwendet (neben dem  $\chi^2$ -Test) den Log-Likelihood-Test für die Erstellung seiner Ranglisten („Typikprofile“) und so wird auch in der vorliegenden Arbeit auf dieses Verfahren zurückgegriffen. Für einen konzisen Überblick über die Anwendbarkeit dieser und anderer Verfahren im Hinblick auf Sprachdaten vgl. Bubenhofer (2009, ab S. 134).

<sup>41</sup> Die Berechnung von  $G^2$  für meine Untersuchung geht von einer Vierfelder-Kontingenztafel aus. Vgl. dazu Rayson/Garside (2000, S. 3), Dunning (1993, S. 70); Kilgarriff (2001, S. 234) oder Bubenhofer (2009, S. 136). Sehr hilfreich, um die Berechnung einmal manuell nachvollziehen zu können ist zudem die Website UCREL (o.J.).

<sup>42</sup> Die linguistische Software sucht nach allen kontinuierlichen Mehrwort-Einheiten, welche im Spezialkorpus vorkommen, wobei in diesem Fall lediglich Dreiwort-Einheiten gewählt werden. Daraufhin wird die Frequenz einer Wortverbindung der Auftretenshäufigkeit von exakt der gleichen Dreiwort-Einheit im Referenzkorpus gegenübergestellt, um so die Log-Likelihood zu bestimmen.

oder geschrieben, sondern ferner auch ganz bestimmte Voraussetzungen, unter denen (sprachliche) ‚Aussagen‘ getätigt werden, sog. ‚Existenzbedingungen‘ bzw. ‚Möglichkeitsbedingungen‘ sind bei einer Analyse von Interesse (vgl. Spitzmüller/Warneke 2011, S. 123). Ferner „schreibt“ sich ein Diskurs nicht nur in Texten ein, sondern tritt auch zutage in Form von Bildern, Ton- wie Videoaufnahmen usw.; er ist multimodal<sup>43</sup> und darf bei seiner Analyse nicht auf Sprache allein beschränkt bleiben (vgl. Spitzmüller/Warneke 2011, S. 39). Des Weiteren beinhaltet eine Diskursanalyse „die Ebene der Akteure als handlungsbezogenen Aspekt der Diskurspraxis“ (Warnke/Spitzmüller 2008, S. 23). Im Besonderen hierzu heißt es: „Akteure können Individuen, Gruppen von Individuen, Netzwerke von Individuen, aber auch nicht-personale Handlungsinstanzen wie Institutionen, Parteien, Medien etc. sein.“ (Spitzmüller/Warneke 2011, S. 172).

Diese voneinander ganz verschiedenen Dimensionen bzw. Entitäten, bei denen eine Diskursanalyse ansetzen kann, spricht: der Unendlichkeit des Diskurses steht die forschungspraktisch und zeitökonomisch motivierte Unumgänglichkeit einer Beschränkung gegenüber. So ist eine wichtige Eingrenzung die thematische. Auch wenn dies wiederum ein kontingentes Konstrukt darstellt und es folglich nie eine „richtige“, sondern immer nur eine verhandelte Definition gibt, „ist das kein Grund, nicht mit thematischen Diskursen zu arbeiten, solange nur die Karten auf den Tisch gelegt werden“ (Teubert 2012, S. 261). Eine Notwendigkeit dergestalt kommt im folgenden Zitat zum Ausdruck:

“The discourse itself is unfathomable, inexhaustible, and as a whole, unexplorable. Perhaps we can approach the conundrum of representativeness more easily if we approach it from the other end, from the corpus.”<sup>44</sup> (Teubert/Čermáková 2004, S. 115)

U.a. im Korpus tritt der Diskurs also in Erscheinung – hier manifestiert er sich zeichenhaft.<sup>45</sup> Corpus-driven ermittelte Muster in einem Korpus haben eine pragmatische Funktion. Weil der Gesamtdiskurs schlicht nicht fassbar ist, steht außer Frage, dass er eine Eingrenzung erfahren muss, will man ihn untersuchen. Stets muss sich der Diskurslinguist der Tatsache bewusst sein, dass er selbst Teilhaber am Diskurs ist, dass er seinen Forschungsgegenstand konstituiert und er daher die Kompilierung seines Korpus‘ genau darzulegen hat.<sup>46</sup> Was zu diesem jeweiligen Ausschnitt an Texten hinzugefügt wird und was nicht, ist „möglichst genau mit kontrollierbaren Parametern wie beispielsweise temporalen, arealen, sprachlichen, textsortalen und natürlich thematischen Spezifikationen zu bestimmen“ (Teubert 2012, S. 259). Von einem solchen, thematisch-inhaltlich festgelegten Diskurs auszugehen; nach Kriterien, welche bestimmte Forscher vielleicht als relevant erachten, scheint Foucaults Vorstellungen (siehe Zitat oben) zu widersprechen. Auch Bubenhofer sieht die Chance in einem corpus-driven-Ansatz eher darin, „maximal“ induktiv zu arbeiten. So berechnet er aus einer Zufallsstichprobe von allen Zeitungstexten mehrerer Jahrgänge der ‚Neuen Zürcher Zeitung‘ frequente Wortverbindungen – und zwar quer durch alle Ressorts und Themengebiete – um letzten Endes Diskursgrenzen oder Themenkreise daraus abzuleiten.

---

<sup>43</sup> Vgl. zur multimodalen Diskursanalyse den Aufsatz von Meier (2011).

<sup>44</sup> Auch Jung/Wengeler (1999) kommen in ihrem Aufsatz zum gleichen Schluss, wenn sie sich mit der Frage auseinandersetzen, „ob ein bestimmter thematisch definierter Diskurs, etwa die europapolitische Diskussion, die Globalisierungsdebatte oder eben der ‚Migrationsdiskurs‘ [...] linguistisch überhaupt greifbar ist. Aus unserer Sicht kann die Antwort auf diese berechtigten Fragen nur eine korpuslinguistische sein: Ein bestimmter Diskurs bleibt in seiner Totalität als Menge aller Äußerungen zum gleichen Thema zwar eine abstrakte Entität, wird aber ausschnitthaft in Textkorpora zugänglich, und zwar in Korpora, die methodisch reflektiert zusammenzustellen und systematisch zu analysieren sind [...]“ (Jung/Wengeler 1999, S. 146f.)

<sup>45</sup> Vgl. auch Kap. 1 und Teubert (2012, S. 238): „Für mich ist ein Diskurs die Gesamtheit aller (gesprochenen oder geschriebenen) Texte, die den Kriterien entsprechen, durch die dieser Diskurs definiert ist.“

<sup>46</sup> „Gleichzeitig mit der Konstitution der Einheit des als Untersuchungsgegenstand gewählten konkreten Textkorpus ist der semantische Zusammenhang des Diskurses zu rekonstruieren, der den Grund für diese Wahl abgibt. Er muß durch den Aufweis von inhaltlichen und strukturellen Eigenschaften plausibel gemacht werden, die den zugehörigen Texten gemeinsam sind.“ (Busse/Teubert 1994, S. 17). Vgl. auch Teubert (2012, S. 259).

„Relevanzempfinden ist ein Ergebnis diskursiver Formationen. Dieses als Kriterium zu wählen, bedeutet, die offensichtlichen Strukturen des Diskurses zu replizieren. [...] Eine induktive Analyse, bei der Sprachgebrauchsmuster im Zentrum stehen, bietet aber zusätzlich die Chance, Diskurse nicht nur thematisch zu definieren, sondern als Cluster von ähnlichem Sprachgebrauch.“ (Bubenhofers 2008, S. 413)

Dass in vorliegender Arbeit mit dem Themenkreis um die Unglücksserie in Japan vom März 2011 eindeutig eine Vorfestlegung getroffen wird (siehe Kap. 4), lässt sich durch folgende Punkte legitimieren: 1. Es wird manchmal einfach vorausgesetzt oder hingenommen, dass die Untersuchungsgegenstände des Forschers schlicht „die ihn beschäftigenden Themen“ (siehe Foucault-Zitat) abbilden, die gelegentlich auch sehr persönlich motiviert sind. 2. Man kann für jedes Korpus, ganz gleich wie oder aus was es sich zusammensetzt, typische Ein- oder Mehrwort-Einheiten aufspüren, welche eine linguistische Aussagekraft besitzen. Denn zu den „Mythen“ der Korpuslinguistik heißt es:

„Mythos 7: Korpora enthalten manchmal ‚Quatsch‘. Eine technisch einwandfreie und authentische Abbildung vorausgesetzt, enthalten Korpora nur Fakten über den Sprachgebrauch.“ (Perkuhn/Belica 2006, S. 5)

Daraus folgt 3. aufs Neue die Frage, wie man die Grenzen eines Diskurses/Korpus‘ genau definiert, da jede begrenzte Textsammlung durch zuvor größtenteils bewusst gesetzte Restriktionen gekennzeichnet ist. So könnte man Bubenhofers Stichprobe auch (durch die Wahl der Korpustexte) als den vordefinierten ‚NZZ-Diskurs‘ bezeichnen. Denn letzterer bildet wie jeder Diskurs per se kein geschlossenes, sondern ein offenes System, das sich mit einer Vielzahl anderer (medialer) Diskurse überlappt (vgl. Jung/Wengeler 1999, S. 147ff.). Da die Natur eines Korpus‘ seine Beschränkung und gezielte Textauswahl ist, soll diese für vorliegende Arbeit nun genauer erklärt werden.

## 4 Korpuskompilierung

In meiner Untersuchung spreche ich verkürzend vom Fukushima-Diskurs. Gemeint ist aber, wie schon angesprochen, der gesamte Diskurs um die desaströsen Vorgänge in Japan vom März 2011. Was sich hier ereignet hat, ist aus der intensiven Berichterstattung in sämtlichen wichtigen Medien hierzulande noch zum großen Teil präsent – daher lediglich folgende kurze Rekapitulation: die Vorfälle nahmen ihren Anfang am 11. März 2011 um 14.46 Uhr Ortszeit: mit dem starken Erdbeben dessen Epizentrum vor der Küste der japanischen Insel Honshu, 373 km nordöstlich von Tokyo lag (vgl. USGS 2011).<sup>47</sup> Es bekam den offiziellen Namen ‚Tohoku-Erdbeben‘ und ist mit einer Magnitude von 9.0 „the largest earthquake recorded in Japan“ (JMA 2011). Durch diese exorbitanten Erdverschiebungen unter Wasser entstand eine Schockwelle, die nach nur wenigen Minuten das Festland erreichte und an bestimmten Küstenabschnitten Japans eine Höhe von teilweise mehr als 20 Metern erreichte (vgl. Fujii/Satake 2011, Fig. 2). Unzählige Helikopter- oder Amateuraufnahmen zeigten die Folgen des ausgelösten Tsunamis: immense Wassermassen, die sich landeinwärts bewegen, welche große Boote, ganze Gebäude und wegführende Autos mit sich reißen.<sup>48</sup> Mit einer Höhe von

---

<sup>47</sup> Im Meer östlich von der Küste „brach dabei innerhalb von weniger als drei Minuten ein Erdsegment von geschätzten 400 bis 500 km Länge und 200 km Breite ab, das bis in 25 km Tiefe reichte“ (Coulmas/Stalpers 2011, S. 19).

<sup>48</sup> Die kleine Hafenstadt Minamisanriku traf es besonders schwer: hier wurden 69 Prozent aller Häuser komplett zerstört (vgl. Coulmas/Stalpers 2011, S. 26f.). Aber auch an vielen anderen Orten der Ostküste Japans waren die Schäden flächendeckend und verheerend. Insgesamt 15.883 Menschen wurden durch das Unglück getötet, 6.150 wurden verletzt, 2.643 gelten sogar noch als vermisst; rund 130.000 Gebäude sind komplett eingestürzt, tausende Straßen und über einhundert Brücken wurden beschädigt (vgl. National Police Agency of Japan 2013).

14 Metern überschwemmt die Welle auch die Anlagen des Kernkraftwerkes Fukushima Daiichi (vgl. Coulmas/Stalpers 2011, S. 50ff.). In der Folge fallen der Strom, das Hauptkühlsystem, die Dieselaggregate aus, wobei die Batterien der Notstromversorgung nur wenige Stunden Energie lieferten. Letzten Endes schlagen sämtliche Kühlversuche fehl: die Reaktorblöcke 1 bis 4 werden durch Explosionen infolge der Bildung von Wasserstoffgas schwer beschädigt. Es kommt zu insgesamt drei Kernschmelzen in den Reaktoren 1 bis 3 (vgl. IAEA 2011, S. 29ff.). Am 12. April 2011 stufte die japanische Atomaufsichtsbehörde ‚NISA‘ (‚Nuclear and Industrial Safety Agency‘) die Vorfälle in Fukushima auf INES-Level 7<sup>49</sup> ein (vgl. Japanese Government 2011, IX-10). Bis zum heutigen Tag dauern die Folgen dieses Desasters an und wird u.a. gegen die radioaktive Belastung angekämpft.<sup>50</sup>

Ähnlich zu dem hier gegebenen, sehr kleinen Überblick über die Ereignisse in Fernost vom März 2011, so spielen in den Printmedien (wie natürlich auch im Fernsehen etc.), also im öffentlichen Diskurs, auch ein paar wiederkehrende Begriffe zur Bezeichnung dieses Themenkomplexes eine wichtige Rolle – es sind dies: *Japan*, *Erdbeben* (oder *Beben*), *Tsunami* und *Fukushima*. Sie erscheinen mir daher als der „kleinste gemeinsame Nenner“ der Berichterstattung und aus diesem Grund geeignet, um als semantisches Kriterium für die Kompilierung eines Korpus‘ verwendet werden zu können. Eine ähnliche Vorgehensweise verfolgt auch Storjohann (2007, S. 574f.) in ihrer Untersuchung zum Begriff *Globalisierung*, die unter Bezugnahme auf Busse/Teubert (1994) ihr virtuelles Korpus nach inhaltlichen bzw. semantischen Kriterien aufbaut, indem sie schlicht auf der Basis der schriftsprachlichen Korpora des Instituts für Deutsche Sprache Mannheim nach sämtlichen Wortformen des Lexems *Globalisierung* sucht.

Die Gewinnung der Texte erfolgte mithilfe der elektronischen Internetdatenbank ‚LexisNexis® Wirtschaft‘ (siehe Literaturverzeichnis).<sup>51</sup> Die Basis oder die Grundgesamtheit<sup>52</sup> aller Äußerungen zur Katastrophenserie in Japan bildet hier der öffentliche Diskurs zu Fukushima in der deutschen Presse, welche durch einen Querschnitt verschiedener printmedialer Vertreter repräsentiert werden soll.<sup>53</sup> Es wurden hierbei Texte in das Korpus

---

<sup>49</sup> Für die Einschätzung solcher atomaren Vorfälle wurde die siebenstufige Skala ‚INES‘ (‚The International Nuclear And Radiological Event Scale‘) eingeführt, bei der die Tschernobyl-Katastrophe von 1986 die oberste Messlatte bildet mit der höchsten Gefahrenstufe 7 (vgl. IAEA 2009, S. 1ff.).

<sup>50</sup> Es wurden auch schon Bücher zu den desaströsen Vorfällen in Japan geschrieben: z.B. vom Journalisten Johannes Hano (2011), der sich zur Zeit der Ereignisse im März 2011 in Tokio aufhielt. Über die gesellschaftlichen Folgen des Fukushima-Desasters sprechen der ehemalige Umweltminister Klaus Töpfer und der Wissenschaftsjournalist Ranga Yogeshwar (Töpfer/Yogeshwar 2011). Ein weiteres Buch wurde gemeinsam von der Japankorrespondentin der Wochenzeitschrift ‚Elsevier‘, Judith Stalpers und dem Direktor des Deutschen Instituts für Japanstudien in Tokio, Florian Coulmas verfasst (Coulmas/Stalpers 2011). Letztgenanntes ist besonders informativ: es enthält zahlreiche sehr gut recherchierte (hier genutzte) Daten und Fakten über die Unglücksserie in Japan, alles in einem eigenen Literaturverzeichnis belegt.

<sup>51</sup> Dass man LexisNexis sehr gut für die Kompilierung eines eigenen Korpus verwenden kann, zeigt Bubenhofer (2006-2011) sehr anschaulich mit einigen Beispielen auf seiner Website zur ‚Einführung in die Korpuslinguistik‘. Der Zugriff auf jenen Teil der Datenbank ist nur Inhabern einer Lizenz gestattet.

<sup>52</sup> Unter ‚Grundgesamtheit‘ versteht man dabei im linguistischen Zusammenhang einen definierten sprachlichen Bereich, über welchen man in seiner Ganzheit Aussagen treffen möchte. Dies ist allerdings oft ein hypothetisches Konstrukt, das nicht präzise bestimmt werden kann, da beispielsweise mündliche Aussagen kaum berücksichtigt werden können (vgl. Atkins et al. 1991, S. 6f. sowie Lemnitzer/Zinsmeister 2006, S. 50f.).

<sup>53</sup> Es wurden alle typischen, journalistischen Textsorten berücksichtigt, d.h. (Leit-) Artikel, Berichte, Meldungen, Reportagen, Interviews, Kommentare, Features, Leserbriefe, Glossen, Kritiken, etc. – unabhängig von ihrer Länge. Hierbei werden alle in derartigen Medien gängigen Ressorts aufgegriffen, sprich: Politik, Wirtschaft, Ausland, Sport, Lokales, Kultur, Feuilleton, Wissenschaft, Technik, Bildung, Service, Reisen, Ratgeber etc. pp. Dies umfasst alle deutschen Publikationsorgane (insgesamt 215 verschiedene), welche LexisNexis zur Verfügung stellt und unter dem Titel ‚Deutsche Presse (German Language News)‘ subsumiert. Darunter befinden sich zahlreiche überregionale wie regionale (Abonnement-) Zeitungen, Wochenmagazine, Boulevardblätter, Wirtschaftsblätter, speziellere Automobilzeitungen usw. Anzumerken ist ferner, dass zur Auswahl ein paar schweizerische und österreichische Blätter hinzugehören sowie auch einige Onlinepräsenzen

aufgenommen aus dem Zeitraum vom 11. März 2011 (Beginn des Erdbebens) bis zum 31. Mai 2012, was wiederum als eine synchrone Dimension verstanden wird. Die Textzusammenstellung mittels der oben genannten Suchbegriffe wurde folgendermaßen vollzogen: es wurden all jene Artikel in das Spezialkorpus aufgenommen, welche das Wort *Fukushima* enthalten sowie zusätzlich Texte, in denen sowohl der Begriff *Japan* wie gleichzeitig auch mindestens einer von den drei Begriffen *Erdbeben*, *Beben* oder *Tsunami* erscheint.<sup>54</sup> So wurde über diesen Zeitraum verteilt ein thematisch definiertes Korpus mit rund 7.000 Texten (exakt: 7.428) kompiliert.<sup>55</sup> Es umfasst insgesamt über 4 Mio. Token (exakt 4.358.938), wobei über 160.000 verschiedene Types gezählt wurden.

Wie in Kap. 2.2 dargestellt, ist es an dieser Stelle wichtig zu klären, wie das Referenzkorpus (RK) gebildet wurde. Denn die (semantischen) Kriterien seiner Zusammenstellung beeinflussen die Ergebnisse für die eigentlich zu analysierende Textsammlung. Scott (2006) hat sich mit dem Einfluss unterschiedlicher Vergleichskorpora auf das Spezialkorpus auseinandergesetzt. U.a. kommt er anhand einiger Beispielrechnungen mit jeweils gänzlich unterschiedlichen Textsorten von Referenzkorpora zu der Erkenntnis:

„[...] genre-specific RCs [reference corpus, M.T.] identify rather different keywords, which itself led to the conclusion that the aboutness of a text may not be one thing but numerous different ones.“ (Scott 2006, S. 11).<sup>56</sup>

Die Bedeutung eines Textes oder einer Textsammlung ist relativ und hängt davon ab, mit welchem Auge man darauf schaut. So könnte man (konkret bezogen auf hiesige Untersuchung) RK verwenden, welche diverse sprachliche Register und Genres, gewissermaßen einen „Querschnitt“ der deutschen Sprache enthalten mit gemischt schriftlichen/mündlichen (transkribierten) Texten und diversen Textsorten aus den unterschiedlichen Bereichen Wissenschaft, Belletristik, Gebrauchstexte, Gespräche etc. pp. Aber dann erhielte man möglicherweise typische Ein- und Mehrwort-Einheiten, die allgemeinere Sprechweisen eines ‚Mediendiskurses‘ an sich (Zeitungen/Zeitschriften/Magazine) widerspiegeln. Zweifelsohne wären solche Muster eine eigene Untersuchung wert (einmal mehr sei hier auf Bubenhofer 2008 verwiesen) – und höchstwahrscheinlich würden ggf. auch hier durch die Statistik induktiv weitere Muster gebildet, mit denen man im wahrsten Sinne des Wortes zuvor nicht „gerechnet“ hat. Dessen ungeachtet wird ein derart gelagertes Ziel in dieser Arbeit jedoch nicht verfolgt. Weil wie dargelegt ein Hauptaugenmerk auf die thematische Definition des Fukushima-Diskurses liegt, sollen daher typische Ein- und Mehrworteinheiten und in der Folge interpretatorisch

---

deutscher Periodika wie ‚Zeit-Online‘, ‚Spiegel-Online‘ etc. Jedoch lässt sich damit keinesfalls sagen, dass auch der ‚Online-Diskurs‘ erfasst wurde, weil dieser nicht einmal im Ansatz repräsentiert ist. Vgl. die Quellenbeschreibung unter:

<http://www.lexisnexis.com/de/business/auth/bridge.do?rand=0.5414886031342303> [Stand: 08.10.2012; nur mittels Lizenzrechte abrufbar].

<sup>54</sup> Es zeigt sich, dass die (Mindest-)Kombination von *Japan* und *Erdbeben* (bzw. *Beben*, *Tsunami*) im hier interessierenden Zeitraum vom 11. März 2011 bis zum 31. Mai 2012 in fast allen Fällen tatsächlich auf das große Tohoku-Erdbeben referiert. Dies muss natürlich nicht sein, denn Japan ist ein Land in einer geologisch aktiven Zone, welches in seiner Geschichte schon vielfach von Erderschütterungen oder Tsunamis heimgesucht wurde, was auch allgemeinerer Gegenstand von Zeitungsberichten sein kann. Sucht man aber nach Artikeln mit diesen Begriffen vor dem 11. März 2011, so erhält man beispielsweise für das ganze Jahr 2010 für die gesamte deutsche Presse bei LexisNexis lediglich ein paar hundert Treffer. Mit exakt den gleichen Suchwörtern steigt die Artikelzahl allein für die ersten Tage nach dem Stichdatum auf mehrere tausend an.

<sup>55</sup> Ein Großteil der Beiträge (70%) entspringt dem klassischen Publikationstyp ‚Zeitung‘, 14% sind ‚Webbasierte Publikationen‘, 8% Artikel aus ‚Magazine & Zeitschriften‘, 5% ‚Branchenfachpresse‘ sowie 3% ‚Agentur- & Pressemeldungen‘. In meiner Arbeit habe ich noch etwas detaillierter die Anteile ganz bestimmter Publikationsorgane am FK aufgelistet: so stammen die meisten Artikel zum Thema aus ‚Rheinische Post Düsseldorf‘ (5,5%), 5,0% aus ‚Berliner Morgenpost Online‘, 3,5% aus ‚Frankfurter Rundschau‘, 3,3% aus ‚Die Welt‘ u.s.f.

<sup>56</sup> Vgl. dazu auch Baker (2004, S. 349).

gewonnene Sprachgebrauchsmuster ermittelt werden, die themenspezifisch sind – nicht medienspezifisch (auch wenn die gewonnenen Muster natürlich stets als ein Teil des öffentlich-printmedialen Diskurses begriffen werden). Damit ist das zentrale, mithin sogar einzige Kriterium benannt, wonach sich das RK vom FK hier unterscheidet: das Thema. Für das RK wurden (ebenfalls mithilfe von LexisNexis) lediglich Texte ausgewählt, die sich nicht mit den Vorfällen in Japan vom März 2011 befassen, sondern komplett gemischte Themen aus der gesamten Bandbreite journalistischer Berichterstattung abdecken – von Tagespolitik über Wirtschaft, Kultur, Sport, Inlands-, Auslandsnachrichten u.v.m.<sup>57</sup> Da die Größe des Referenzkorpus‘ nur wenig Einfluss auf die Ergebnisse typischer EWE hat (vgl. Scott 2006, S. 8), wurde hier schlicht entschieden, ein RK von ungefähr ähnlichem Umfang wie das FK zu verwenden. Somit besteht das gebildete RK aus über 7000 Artikeln (genau 7773) und rund 4,5 Millionen Token (4.496.446).<sup>58</sup> Ferner muss noch angemerkt werden, dass bei der Bildung der Korpora und der berechneten EWE wie MWE keine Lemmatisierung oder linguistische Annotation erfolgte. Einerseits geschah dies aus zeitlichen Gründen nicht<sup>59</sup>, vielmehr aber aus der Überzeugung, dass solche Informationen für die Zwecke meiner diskurslinguistischen Analyse keinen zusätzlichen Nutzen erbringen.<sup>60</sup> Eine systematisierende Zusammenfügung der Klassen würde jeglicher (grammatischer) Form letztlich die Bedeutung absprechen und stünde zu sehr im klassischen Verständnis einer Saussurschen ‚langue‘-Auffassung.

## 5 Ergebnisteil

### 5.1 Corpus-driven: Listen der berechneten typischen Ein- und Mehrwort-Einheiten

---

<sup>57</sup> Ansonsten sind alle anderen Parameter exakt dieselben wie beim Fukushimakorpus: es handelt sich beim RK ebenfalls um Texte aus exakt den gleichen Publikationsorganen (Quelle ‚Deutsche Presse‘ bei LexisNexis). Auch wurden alle journalistischen Ressorts bzw. Textsorten berücksichtigt und es wurde ein identischer Zeitraum gewählt (11. März 2011 bis zum 31. Mai 2012). Um sicherzustellen, dass die Texte für das RK thematisch-inhaltlich sehr divers sind, wurden für die Korpus-Generierung die wenig inhaltsbezogenen (grammatischen) Begriffe *außer*, *solche*, *hinter*, *daher*, *hin*, *darüber*, *aufs*, *hinzu* und *diesmal* für die Suche abwechselnd gebraucht. Ausgeschlossen wurden dabei nur jene Artikel, welche die Wörter *Fukushima* und gemeinsam auftretende Begriffspaare wie *Japan* und *Erdbeben* etc. beinhalten.

<sup>58</sup> 80% aller Artikel im RK entfallen auf den Publikationstyp ‚Zeitungen‘, 7% auf ‚Webbasierte Publikationen‘, 6% auf ‚Magazine & Zeitschriften‘, 4% auf ‚Agentur- & Pressemeldungen‘ sowie 3% auf ‚Branchenfachpresse‘. 9,4% der Artikel stammen aus ‚Rheinische Post Düsseldorf‘, 4,9% aus ‚Mitteldeutsche Zeitung‘, 4,1% aus ‚Kölnische Rundschau‘, 3,7% aus ‚Hamburger Abendblatt Online‘ usw.

<sup>59</sup> Eine automatische Annotation eines größeren Korpus ist durchaus fehleranfällig (vgl. Baker 2004, S. 355) und eine manuelle angesichts eines Umfangs von 4 Mio. Token des FK nicht umsetzbar.

<sup>60</sup> Vgl. Perkuhn/Belica (2006, S. 3), die es als einen „Mythos“ bezeichnen, dass Korpora durch Annotationen an Wert gewinnen. Stubbs (2001, S. 27) zeigt anhand von englischsprachigen Beispielen, dass unterschiedliche Wortformen ein und desselben Lemmas mit teilweise ganz unterschiedlichen Wörtern kookkurieren. Folglich stellt eine Lemmatisierung, also Zusammenfassung zu Wortklassen, keinen Informationsgewinn, sondern einen Informationsverlust dar. Abgesehen von einem kurzen, gegebenen Ausblick verzichtet auch Bubenhofer (2009, S. 127) für seine Beispielsanalysen auf ein mit Wortarten oder diversen Lemmata annotiertes Korpus.

N	Key word	Freq.	%	RC. Freq.	RC. %	Keyness	P
1	FUKUSHIMA	10063	0,23	0		14277,06	0,00000000000000000000
2	JAPAN	10998	0,25	341		13004,74	0,00000000000000000000
3	ERDBEBEN	4377	0,10	35		5846,16	0,00000000000000000000
4	TSUNAMI	3417	0,08	22		4608,94	0,00000000000000000001
5	KATASTROPHE	3863	0,09	130		4507,85	0,00000000000000000001
6	TOKIO	3330	0,08	35		4380,05	0,00000000000000000001
7	JAPANISCHEN	3388	0,08	58		4293,90	0,00000000000000000001
8	AKW	2873	0,07	78		3458,69	0,00000000000000000002
9	TEPCO	2427	0,06	3		3399,07	0,00000000000000000002
10	JAPANISCHE	2555	0,06	46		3222,64	0,00000000000000000002
11	REAKTOREN	2124	0,05	40		2667,17	0,00000000000000000004
12	JAPANS	1959	0,04	19		2588,92	0,00000000000000000005
13	#	197052	4,52	173971	3,87	2341,19	0,00000000000000000007
14	ENERGIEWENDE	2052	0,05	99		2240,61	0,00000000000000000007
15	ATOMKRAFT	1929	0,04	69		2228,40	0,00000000000000000008
16	BEBEN	1743	0,04	41		2136,34	0,00000000000000000009
17	REAKTOR	1479	0,03	16		1941,54	0,00000000000000000012
18	ATOMKRAFTWERK	1478	0,03	21		1902,96	0,00000000000000000012
19	PROZENT	8770	0,20	4100	0,09	1884,68	0,00000000000000000013
20	STROM	2105	0,05	255		1713,86	0,00000000000000000017
21	ENERGIEN	1864	0,04	168		1711,16	0,00000000000000000017
22	JAPANER	1385	0,03	33		1694,89	0,00000000000000000017
23	UHR	6721	0,15	2929	0,07	1652,87	0,00000000000000000019
24	ATOMAUSSTIEG	1583	0,04	93		1651,80	0,00000000000000000019
25	ATOMKATASTROPHE	1259	0,03	19		1613,04	0,00000000000000000020
26	REGIERUNG	3375	0,08	903	0,02	1599,48	0,00000000000000000021
27	YEN	1213	0,03	21		1535,45	0,00000000000000000024
28	TSCHERNOBYL	1165	0,03	21		1469,07	0,00000000000000000027
29	TEL	1957	0,04	280		1466,97	0,00000000000000000027
30	AUSSTIEG	1450	0,03	105		1429,30	0,00000000000000000029
31	STRAHLUNG	1068	0,02	21		1335,15	0,00000000000000000036
32	ATOMKRAFTWERKE	1259	0,03	72		1322,43	0,00000000000000000037
33	RADIOAKTIVITÄT	989	0,02	10		1303,68	0,00000000000000000039
34	KERNENERGIE	992	0,02	18		1250,10	0,00000000000000000044
35	ENERGIE	1904	0,04	371		1178,71	0,00000000000000000053
36	ATOMENERGIE	1011	0,02	40		1147,57	0,00000000000000000058
37	ATOM	1098	0,03	78		1088,41	0,00000000000000000068
38	KAN	773	0,02	2		1070,75	0,00000000000000000072
39	ERNEUERBAREN	1002	0,02	58		1049,34	0,00000000000000000076
40	RWE	1079	0,02	87		1028,68	0,00000000000000000081
41	GRÜNEN	3319	0,08	1283	0,03	997,57	0,00000000000000000089
42	WIRTSCHAFT	2620	0,06	863	0,02	984,14	0,00000000000000000093
43	RADIOAKTIVE	771	0,02	14		971,51	0,00000000000000000097
44	MRD	1674	0,04	359		964,40	0,00000000000000000099
45	DEUTSCHLAND	5539	0,13	2885	0,06	935,90	0,00000000000000000109
46	KERNSCHMELZE	753	0,02	17		927,28	0,00000000000000000112
47	MEILER	864	0,02	46		923,00	0,00000000000000000114
48	GAU	953	0,02	75		915,65	0,00000000000000000117
49	DONNERSTAG	2077	0,05	597	0,01	914,01	0,00000000000000000117
50	KERNKRAFTWERKE	703	0,02	13		884,26	0,00000000000000000130

Abbildung 1: Liste der typischen Einwort-Einheiten (Kenn-Wörter) – sortiert nach G<sup>2</sup>-Wert (,Keyness‘)

N	Key word	Freq.	%	RC. Freq.	RC. %	Keyness	P
51	TOYOTA	930	0,02	80		867,69	0,00000000000000000138
52	KERNKRAFT	653	0,01	7		857,65	0,00000000000000000143
53	KONZERN	1389	0,03	294		808,42	0,00000000000000000172
54	MERKEL	2234	0,05	761	0,02	803,65	0,00000000000000000175
55	IN	78678	1,80	70151	1,56	803,04	0,00000000000000000176
56	MENSCHEN	5451	0,13	3000	0,07	800,08	0,00000000000000000178
57	MILLIARDEN	2372	0,05	853	0,02	792,82	0,00000000000000000183
58	RÖTTGEN	824	0,02	74		757,32	0,00000000000000000211
59	EON	657	0,02	24		756,21	0,00000000000000000212
60	KRAFTWERK	672	0,02	28		755,50	0,00000000000000000213
61	ERDBEBENS	537	0,01	2		737,61	0,00000000000000000230
62	QUARTAL	980	0,02	151		705,02	0,00000000000000000265
63	ERNEUERBARE	731	0,02	63		681,57	0,00000000000000000295
64	SIEMENS	772	0,02	78		678,94	0,00000000000000000298
65	REAKTORKATASTROPHE	513	0,01	5		677,68	0,00000000000000000300
66	BETREIBER	962	0,02	154		676,84	0,00000000000000000301
67	HAVARIERTEN	482	0,01	5		634,35	0,00000000000000000370
68	TSUNAMIS	469	0,01	5		616,19	0,00000000000000000405
69	STÄRKE	818	0,02	119		607,62	0,00000000000000000424
70	ÖKOSTROM	510	0,01	15		607,11	0,00000000000000000425
71	RADIOAKTIVEN	499	0,01	15		592,16	0,00000000000000000460
72	NAOTO	416		0		589,74	0,00000000000000000466
73	KRAFTWERKE	562	0,01	35		577,70	0,00000000000000000498
74	NETZ	1275	0,03	357		576,79	0,00000000000000000500
75	ENBW	560	0,01	36		570,82	0,00000000000000000517
76	POLITIK	2932	0,07	1436	0,03	570,68	0,00000000000000000518
77	ANLAGEN	1060	0,02	247		570,41	0,00000000000000000518
78	WELT	3723	0,09	2017	0,04	569,52	0,00000000000000000521
79	SENDAI	409		1		567,14	0,00000000000000000528
80	WASSER	1821	0,04	692	0,02	561,66	0,00000000000000000545
81	ARBEITER	651	0,01	70		558,30	0,00000000000000000555
82	BRENNSTÄBE	427		5		557,58	0,00000000000000000557
83	ATOMSTROM	493	0,01	20		557,02	0,00000000000000000559
84	ENERGIEPOLITIK	567	0,01	42		555,08	0,00000000000000000565
85	PRÄFEKTUR	399		1		553,01	0,00000000000000000572
86	NATURKATASTROPHEN	475	0,01	17		548,60	0,00000000000000000587
87	KATASTROPHEN	539	0,01	35		547,92	0,00000000000000000589
88	SCHÄDEN	895	0,02	180		541,41	0,00000000000000000612
89	KERNKRAFTWERK	422		7		536,23	0,00000000000000000632
90	KILOMETER	1526	0,04	538	0,01	524,19	0,00000000000000000679
91	NATURKATASTROPHE	403		5		524,13	0,00000000000000000679
92	RADIOAKTIV	385		2		523,45	0,00000000000000000682
93	STR	556	0,01	53		500,00	0,00000000000000000791
94	DAX	626	0,01	79		499,91	0,00000000000000000791
95	IAEA	380		4		499,65	0,00000000000000000792
96	UNGLÜCK	568	0,01	58		497,44	0,00000000000000000804
97	BIBLIS	439	0,01	18		495,03	0,00000000000000000816
98	AREVA	372		4		488,48	0,00000000000000000852
99	ABGESCHALTET	485	0,01	33		486,71	0,00000000000000000862
100	BUNDESREGIERUNG	1171	0,03	356		484,19	0,00000000000000000877

Fortsetzung von voriger Seite: Kenn-Wort-Liste mit den Items 51 bis 100



Man sieht in Abbildung 1, dass Types wie *Fukushima*, *Japan*, *Erdbeben*, *Tsunami* und *Beben* sehr hohe  $G^2$ -Werte haben, folglich sehr charakteristisch bzw. ‚typisch‘ für das FK sind.<sup>61</sup> Gerade den genannten Wörtern ist aber hinsichtlich der Bedeutung ihrer Frequenz hinzuzufügen, dass sie bekanntlich das Suchkriterium für die Zusammenstellung der Texte zur Unglücksserie in Japan und zugleich das (thematische) Ausschlusskriterium für die Kompilierung des Vergleichskorpus bildeten (deswegen kommt z.B. die Wortform *Fukushima* 0 Mal im RK vor).

Zwecks weiterer Visualisierung der Ergebnisse wurde die Abbildung 2 erstellt. Sie führt die bloßen Wortformen aus der KW-Liste in Form einer Wortwolke auf. In der Mitte der Seite befinden sich die typischen EWE mit der höchsten Log-Likelihood-Ausprägung. Zentral von der Mitte hin zur Peripherie der Seite nimmt dieser Wert für die jeweilige Wortform immer weiter ab. Zusätzlich wird dieser Sachverhalt über die Schriftgröße angezeigt: je größer also der  $G^2$ -Wert, desto größer auch das abgebildete Wort. Im Zentrum der Darstellung befinden sich demzufolge u.a. die Wortformen des Suchstrings. Interessant hierbei ist jedoch, dass *Fukushima* mit Abstand die höchste ‚Keyness‘ hat – es hätte gleichsam auch eines der Wörter *Erdbeben*, *Tsunami* oder *Beben* sein können. Allein dies rechtfertigt auch etwas die generalisierende Bezeichnung „Fukushima-Korpus“ oder „Fukushima-Diskurs“ für die Unglücksserie im Allgemeinen vom März 2011 in Japan. Blickt man auf die gewonnenen Schlagwörter, so findet man schnell weitere Indizien, welche dafür sprechen, dass den Vorfällen in und um das havarierte Kernkraftwerk Japans eine größere Aufmerksamkeit im öffentlichen Diskurs hierzulande zuteil wird, als den tragischen Auswirkungen der Naturkatastrophen. Hierzu zählen etwa KW wie *Atomkraftwerk*, *AKW*, *Reaktoren*, *Atomkraft*, *Reaktor*, *Atomkraftwerk*, *Atomkatastrophe* (höhere Log-likelihood als etwa *Naturkatastrophe*), *Tschernobyl*, *Strahlung*, *Atomkraftwerke*, *Radioaktivität*, *Kernschmelze*, *Gau* u.v.m. Doch diese Schlüsselwörter beziehen sich inhaltlich nicht nur auf die Ereignisse in *Japan*.<sup>62</sup> Stattdessen betreffen sie im beträchtlichen Ausmaß auch die *Energiepolitik* in *Deutschland*, ersichtlich etwa an den hochsignifikanten Types *Energiewende*, *Strom*, *Energien*, *Atomausstieg*, *Ausstieg*, *Erneuerbare*, *Ökostrom*, *Biblis*, *Abgeschaltet* usw. Dass die Ereignisse Debatten auf hoher bis höchster politischer Ebene sowohl hierzulande als auch in Fernost anstoßen, legen folgende soziale Akteure nahe, welche sich als KW wiederfinden: *Regierung*, *Kan* (damaliger japanischer Ministerpräsident), *Grünen*, *Merkel*, *Röttgen* (damaliger Umweltminister), *Naoto* (Vorname Kans), *Bundesregierung* etc. Verquickt mit den politischen sind die wirtschaftlichen Auswirkungen, was sich wiederum an diesen Akteuren bzw. KW ablesen lässt: *Tepco*, *RWE*, *Toyota*, *Konzern*, *EON*, *Siemens*, *Betreiber*, *ENBW*. Darüberhinaus erkennt man den Einfluss, den die Ereignisse auf die Ökonomie ausüben auch an den EWE: *Yen*, *Wirtschaft*, *Quartal*, *Milliarden*, *Schäden*, *DAX*. Nachfolgende Abbildung 3 führt die einhundert „dominantesten“ typischen Mehrwort-Einheiten auf.

---

<sup>61</sup> Die Liste liest sich folgendermaßen: ‚N‘ steht in diesem Fall für die Position; ‚key word‘ ist das jeweilige von WordSmith Tools berechnete Kenn-Wort bzw. typische Einworteinheit; ‚Freq.‘ steht für die absolute Auftrenshäufigkeit dieses Items im Fukushima-Korpus; das Prozentzeichen ‚%‘ stellt die relative Frequenz der entsprechenden Wortform im FK im Verhältnis zu seiner Gesamt-Token-Zahl dar; ‚RC. Freq.‘ ist die absolute Frequenz des Kenn-Wortes im Referenzkorpus; ‚RC. %‘ dementsprechend die relative Frequenz des Items im Verhältnis zur Gesamt-Token-Zahl des Referenzkorpus. Mit ‚Keyness‘ ist hier stets der berechnete  $G^2$ -Wert (Log-Likelihood) gemeint – die Liste ist absteigend nach der Höhe dieses Wertes sortiert. ‚P‘ gibt den P-Wert an (nicht Signifikanzlevel). Letzterer wurde der Vollständigkeit aufgeführt, um durch seine sehr kleinen Ausprägungen anzudeuten wie ‚hoch signifikant‘ (bzw. weit darüber hinaus) die Items sind. Wenn in einer Spalte (insbesondere in RC. %) Werte fehlen, ist die relative Frequenz so klein ist, dass sie von WordSmith Tools gar nicht erst angezeigt wird.

<sup>62</sup> Daneben verorten weitere typische Einworteinheiten wie *Tokio*, *Japanischen*, *Japanische*, *Japaner*, *Sendai* und *Präfektur* die Geschehnisse geografisch.

N	Key word	Freq.	%	RC. Freq.	RC. %	Keyness
1	NACH DER KATASTROPHE	537	0,01	7		695,92
2	ERDBEBEN IN JAPAN	482	0,01	0		683,31
3	KATASTROPHE VON FUKUSHIMA	457	0,01	0		647,86
4	NACH DEM ERDBEBEN	430		0		609,59
5	AUSSTIEG AUS DER	516	0,01	27		553,40
6	KATASTROPHE IN JAPAN	437	0,01	20		482,35
7	ATOMKATASTROPHE VON FUKUSHIMA	322		0		456,48
8	DER KATASTROPHE VON	322		0		456,48
9	ERDBEBEN UND TSUNAMI	312		0		442,30
10	DIE JAPANISCHE REGIERUNG	289		0		409,69
11	IM JAPANISCHEN FUKUSHIMA	281		0		398,35
12	DER KATASTROPHE IN	299		7		366,63
13	DER ERNEUERBAREN ENERGIEN	317		12		362,63
14	IM ATOMKRAFTWERK FUKUSHIMA	232		0		328,89
15	IN JAPAN UND	299		14		328,48
16	DER ATOMKATASTROPHE VON	228		0		323,22
17	NACH DEM BEBEN	224		0		317,55
18	AUS DER KERNENERGIE	214		0		303,37
19	NACH DER ATOMKATASTROPHE	232		5		287,19
20	ERDBEBEN DER STÄRKE	192		0		272,18
21	TSUNAMI IN JAPAN	192		0		272,18
22	KATASTROPHE IN FUKUSHIMA	186		0		263,68
23	DER ATOMKATASTROPHE IN	210		5		256,98
24	NACH DER REAKTORKATASTROPHE	178		0		252,34
25	IM ERSTEN QUARTAL	361		60		248,29
26	IN DER PRÄFEKTUR	175		0		248,08
27	REAKTORKATASTROPHE VON FUKUSHIMA	174		0		246,66
28	ATOMKATASTROPHE IN JAPAN	222		10		245,82
29	DAS ERDBEBEN IN	172		0		243,83
30	AUS DER ATOMENERGIE	227		12		242,88
31	AUS DER ATOMKRAFT	211		8		241,30
32	MINISTERPRÄSIDENT NAOTO KAN	170		0		240,99
33	DER REAKTORKATASTROPHE VON	165		0		233,91
34	ATOMKATASTROPHE IN FUKUSHIMA	164		0		232,49
35	NACH DEM TSUNAMI	160		0		226,82
36	NACH DEM VERHEERENDEN	153		0		216,89
37	AUSBAU DER ERNEUERBAREN	152		0		215,48
38	DEM ERDBEBEN UND	150		0		212,64
39	DER ERDBEBEN UND	150		0		212,64
40	ERDBEBEN TSUNAMI UND	149		0		211,22
41	DIE KATASTROPHE VON	146		0		206,97
42	DEM VERHEERENDEN ERDBEBEN	145		0		205,55
43	DIE ERNEUERBAREN ENERGIEN	145		0		205,55
44	ERDBEBEN UND DEM	141		0		199,88
45	DEM ERDBEBEN IN	140		0		198,47
46	IN JAPAN DIE	177		8		195,86
47	IN DER JAPANISCHEN	138		0		195,63
48	BEBEN DER STÄRKE	133		0		188,54
49	DIE NACHRICHTENAGENTUR KYODO	132		0		187,12
50	DIE KATASTROPHE IN	160		6		183,33

Abbildung 3: Liste der typischen Mehrwort-Einheiten – sortiert nach G<sup>2</sup>-Wert („Keyness“)

N	Key Word	Freq.	%	RC. Freq.	RC. %	Keyness
51	IN BADEN WÜRTTEMBERG	347		84		180,60
52	DAS ERDBEBEN UND	126		0		178,62
53	VON FUKUSHIMA UND	123		0		174,37
54	UND DEM TSUNAMI	121		0		171,53
55	REGIERUNGSSPRECHER YUKIO EDANO	120		0		170,11
56	TOKYO ELECTRIC POWER	119		0		168,70
57	IM ZWEITEN QUARTAL	199		22		168,68
58	ATOMKRAFTWERK FUKUSHIMA DAIICHI	118		0		167,28
59	FÜR DIE OPFER	178		15		167,23
60	EIN ERDBEBEN DER	116		0		164,44
61	DER JAPANISCHEN REGIERUNG	115		0		163,02
62	RUBRIK POLITIK S	568	0,01	228		160,77
63	MENSCHEN IN JAPAN	142		6		159,21
64	IN JAPAN ZU	112		0		158,77
65	DER PRÄFEKTUR FUKUSHIMA	111		0		157,35
66	ERDBEBEN UND DER	111		0		157,35
67	REGIERUNG IN TOKIO	111		0		157,35
68	IN JAPAN IST	110		0		155,94
69	CODE ZEI COPYRIGHT	108		0		153,10
70	DER FUKUSHIMA KATASTROPHE	108		0		153,10
71	ZEITUNGS CODE ZEI	108		0		153,10
72	DES ERDBEBENS UND	107		0		151,68
73	IM NORDOSTEN JAPANS	107		0		151,68
74	IN JAPAN HAT	132		5		150,98
75	DEM SCHWEREN ERDBEBEN	106		0		150,27
76	FÜR DEN WIEDERAUFBAU	106		0		150,27
77	ERDBEBENS IN JAPAN	105		0		148,85
78	DIE ATOMKATASTROPHE IN	104		0		147,43
79	RUBRIK WIRTSCHAFT S	366		115		145,60
80	DIE JAPANISCHE WIRTSCHAFT	101		0		143,18
81	JAPANISCHEN ATOMKRAFTWERK FUKUSHIMA	101		0		143,18
82	IM ERSTEN HALBJAHR	185		25		142,84
83	DEN AUSSTIEG AUS	136		8		141,85
84	VOR DER KATASTROPHE	100		0		141,76
85	NACH ANGABEN DER	303		83		140,18
86	DER KATASTROPHE IM	98		0		138,93
87	AUF ERNEUERBARE ENERGIE	97		0		137,51
88	DIE ATOMKATASTROPHE VON	96		0		136,09
89	BUNDESKANZLERIN ANGELA MERKEL	270		68		135,55
90	WIEDER ANS NETZ	120		5		134,90
91	DURCH DAS ERDBEBEN	93		0		131,84
92	IM AKW FUKUSHIMA	93		0		131,84
93	IN JAPAN SIND	93		0		131,84
94	DIE ZEIT INKLUSIVE	114		5		126,90
95	INKLUSIVE ZEIT MAGAZIN	114		5		126,90
96	ZEIT INKLUSIVE ZEIT	114		5		126,90
97	ERDBEBENKATASTROPHE IN JAPAN	89		0		126,17
98	EIN JAHR NACH	155		19		125,48
99	DEN AUSBAU DER	144		15		125,10
100	DES ATOMKRAFTWERKS FUKUSHIMA	88		0		124,75

Fortsetzung von voriger Seite: Liste typischer Mehrwort-Einheiten mit den Items 51 bis 100

## 5.2 Corpus-based: qualitative Analyse ausgewählter Ergebnisse

In diesem Kapitel werden nun einige von den dominantesten typischen Mehrworteinheiten ausgesucht und als Sprachgebrauchsmuster umfassend interpretiert, was mit zahlreichen Zitaten aus den (Zeitung-) Artikeln qualitativ gestützt wird. Sofern ein gewonnenes SGM pragmatisch gedeutet wird, geschieht dies nicht aufgrund weniger singulärer Belege. Stattdessen zeichnen sich in den Konkordanzen zum jeweiligen Muster inhaltliche Ähnlichkeiten und Übereinstimmungen ab. Nichtsdestotrotz bleibt eine beschriebene pragmatische Funktion ein vom Forscher abhängiger, konstruktivistisch-interpretativer Akt.<sup>63</sup>

### **Die SGM *Erdbeben in Japan, Katastrophe in Japan, Tsunami in Japan, ...***

Es erfolgt zunächst einmal die Analyse von SGM, welche eine Benennungsfunktion der Vorfälle in Fernost inne haben (siehe unten). Dazu gehören: die mit einem hohen Vorkommen von 482 Belegen im Fukushima-Korpus zweittypischste Mehrwort-Einheit *Erdbeben in Japan* ( $G^2 = 683,31$ ); die mit einer Log-Likelihood von 647,86 auf dritter Position rangierende Formel *Katastrophe von Fukushima* (Freq. = 457); das SGM *Katastrophe in Japan* mit einer Frequenz von 437 im FK ( $G^2 = 482,35$ ) und ferner die „erst“ an 21. Position folgende Kennwortverbindung *Tsunami in Japan* (Freq. = 192;  $G^2 = 272,18$ ).

Die MWE *Erdbeben in Japan* verweist nicht auf frühere Beben dieses Landes, sondern praktisch ausschließlich auf das Hauptbeben vom 11. März 2011 (siehe das Erscheinungsdatum des Artikels):

- (1) „Hunderte Tote und Sachschäden von mehreren Milliarden Dollar haben das gewaltige *Erdbeben in Japan* und der dadurch ausgelöste Tsunami gekostet. Dennoch ist das Land gemessen an historischen Erdbeben noch relativ glimpflich davon gekommen.“

Berliner Morgenpost Online, Freitag 11. März 2011 5:44 PM MET (GMT+1), RUBRIK: WISSEN

Dieses Muster hat eine „Etikettierungsfunktion“, weil es in kürzester Form in einer Dreiwort-Einheit zwei wichtige (nachrichtenrelevante) Informationen zusammenträgt, indem es 1. das katastrophale Ereignis (*Erdbeben*) und 2. den Ort auf der Erde, wo es stattgefunden hat (*in Japan*), benennt und somit auf eine ganz spezielle, aktuelle Naturkatastrophe in der Welt referiert. Um die außerordentlich große Magnitude dieses Bebens und seine Folgen zu charakterisieren, finden sich in den Konkordanzen sehr oft entsprechende Attributbezeichnungen, welche als Kollokate in der ersten Wortposition links von diesem Muster auftauchen: *verheerenden Erdbeben in Japan, schweren Erdbeben in Japan, katastrophale Erdbeben in Japan, gewaltigen Erdbeben in Japan*. Inhaltlich beschäftigen sich die Kontexte hier zu einem Teil mit den unmittelbaren Folgeschäden vor Ort und zum anderen Teil mit längerfristigen Schäden für verschiedene Wirtschaftsbranchen:

- (2) „Bei dem verheerenden *Erdbeben in Japan* und dem anschließenden Tsunami sind offenbar mehr als tausend Menschen ums Leben gekommen.“

Berliner Morgenpost Online, Freitag 11. März 2011 10:32 PM MET (GMT+1), Rubrik:VERMISCHTES

- (3) „Das *Erdbeben in Japan* trifft die Wirtschaft hart. Internationale Konzerne beklagen Schäden, zudem steigen wohl Versicherungsprämien. Doch vor allem ist das Beben eine Katastrophe für den hochverschuldeten Inselstaat: Es könnte Japan den finanziellen Todesstoß versetzen.“

manager magazin online, 11. März 2011 Freitag 5:30 PM GMT+1, RUBRIK: ERDBEBEN IN JAPAN

---

<sup>63</sup> Sämtliche Ein- wie Mehrwort-Einheiten werden in konkreten Textbelegen hier stets kursiv gedruckt, um ihr Auffinden zu erleichtern, wobei unter jedem Konkordanzbeispiel die von LexisNexis ausgegebenen Meta-Informationen bezüglich der Quelle aufgeführt werden.

- (4) „Kurserschütterung. Das *Erdbeben in Japan* hat auch den Kupferpreis erschüttert.“

Rheinische Post Duesseldorf, 14. März 2011 Montag

Interessant ist die hier gebrauchte bildliche Übertragung von seismischen Erdbebenwellen auf schwankende Börsenpreise. Tatsächlich hat die Katastrophenserie in Japan von 2011 darüber hinaus manche „Erschütterungen“ in der Wirtschaftswelt ausgelöst (siehe weiter unten). Diese beiden Auswirkungen, also Opferzahlen und (ökonomische) Einbußen dominieren die inhaltlichen Kontexte der MWE. Weniger, aber dennoch in Erscheinung tretend ist der Gebrauch dieser Formel als Ursachefaktor für das sich anschließende Atomunglück:

- (5) „Das *Erdbeben in Japan* hat ein Atomkraftwerk im Katastrophengebiet in eine kritische Lage gebracht. Nach dem Ausfall der Hauptkühlsysteme und der Stromversorgung in mindestens zwei Reaktoren des Kraftwerks Fukushima sprechen Experten von einem ‚beängstigenden Rennen gegen die Zeit‘.“

Berliner Morgenpost Online, Freitag 11. März 2011 11:02 PM MET (GMT+1), RUBRIK: VERMISCHTES

Die MWE *Tsunami in Japan* hat ebenfalls wie *Erdbeben in Japan* eine ähnlich gelagerte, semantisch-pragmatische Referenzfunktion, indem sie ganz spezifisch auf eine aktuelles desaströses Ereignis verweist, welches sich in einer bestimmten Gegend in der Welt ereignet hat. Diese enge kausale Verknüpfung beider Naturkatastrophen ist sehr wahrscheinlich auch ein Grund dafür, warum das SGM *Tsunami in Japan* häufigerer Bestandteil des Musters *Erdbeben und Tsunami in Japan* (Freq. = 35) ist. Auch hier sind die unmittelbaren inhaltlichen Kontexte der Konkordanzen (ähnlich wie bei den oben genannten äquivalenten SGM) zum Teil durch angerichtete Zerstörungen, Opferzahlen und direkte Auswirkungen auf Umwelt und Menschen in Japan gekennzeichnet:

- (6) „Die Bilder, die am Freitag aus Japan um die Welt gehen, sind schockierend. Die apokalyptischen Szenen erinnern an den Tsunami 2004 im Indischen Ozean. Damals starben mehr als 200.000 Menschen. Auch der *Tsunami in Japan* kostet unzählige Menschen das Leben.“

Spiegel Online, 11. März 2011 Freitag 8:54 PM GMT+1, RUBRIK: ERDBEBEN UND TSUNAMI

Ebenso wird hier noch mehr als für Berichte über Tote und Verletzte das Muster als Ursache für wirtschaftlich-finanzielle Folgen, Kosten, Gewinneinbrüche diverser Unternehmen (oft der japanischen Automobilbranche) und allgemeine Entwicklungen auf den Finanz-Märkten gebraucht:

- (7) „MONEY: Wie reagieren die Finanzmärkte?

Roselieb: Zunächst mit einem massiven Vertrauensverlust. Diese unmittelbare Panik haben wir sowohl nach dem 11. September 2001 als auch nach dem jüngsten *Tsunami in Japan* erlebt.“

Focus-Money, 11. Mai 2011, RUBRIK: S. 500x Ausg. 20

- (8) „Trotz der Krise in vielen Ländern und Naturkatastrophen wie dem *Tsunami in Japan* hat die Reisebranche 2011 neue Rekorde erzielt.“

Die Welt, Dienstag 17. Januar 2012, RUBRIK: WIRTSCHAFT; S. 11 Ausg. 14

Das Zitat (8) ist eines der seltenen Gegenbeispiele, bei dem der *Tsunami in Japan* nicht als Auslöser negativer wirtschaftlicher Entwicklungen fungiert. In dem Artikel selbst deutet sich diese Ungewöhnlichkeit in dem Wort „trotz“ an. Dass den monetären Konsequenzen mehr Aufmerksamkeit geschenkt wird als den Schilderungen über die Opfer, erklärt sich sehr wahrscheinlich damit, dass Japan ein Land ist, welches in der Wahrnehmung vieler Leser und Textproduzenten hierzulande weit entfernt liegt und die Menschen in Deutschland kaum persönlich von der Katastrophe betroffen sind. Stattdessen legt der Diskurs andere Bereiche offen, auf die hierzulande die Ereignisse in Japan großen Einfluss ausüben. Einen Bereich davon bilden wie erwähnt die Zweige Ökonomie und Industrie (siehe auch weiter unten die

SGM zu den ökonomischen Auswirkungen). Ein anderer Komplex, welcher gleichfalls maßgeblich die Inhalte von Konkordanz zu diesem SGM prägt, ist der in der Folge ausgelöste Vorfall im AKW Fukushima Daiichi:

- (9) „Das Begriff ‚Restrisiko‘ wurde auf erschreckende Weise Realität, als nach dem *Tsunami in Japan* im März das Atomkraftwerk Fukushima überflutet wurde und es daraufhin zu Kernschmelzen in mehreren Reaktorblöcken kam.“

Die Welt, Dienstag 27. Dezember 2011, RUBRIK: WIRTSCHAFT; S. 10 Ausg. 302

Während diese Muster auf einen speziellen Teil der Unglücksserie in Japan referieren, stellt die hochfrequente typische Formel *Katastrophe in Japan* (Freq. = 437;  $G^2 = 482,35$ ) eine allgemeinere Bezeichnung für alle tragischen Vorfälle in Japan mit Beginn vom 11. März 2011 dar. In den Konkordanz zeigt sich bereits in den sprachlichen Wortbildungen, inwiefern das Desaster in Japan „untergliedert“ wird. So fungiert das Wort *Katastrophe* in diesem SGM häufig als Bestandteil eines Bindestrich-Kompositums:

- (10) „Die Schüler des Thomas-Eßer-Berufskolleg zeigten sich mit den Opfern der *Dreifach-Katastrophe in Japan* solidarisch. Bei einem Sponsorenlauf sammelten sie Spendengelder.“

Kölnische Rundschau, Samstag 21. Mai 2011, RUBRIK: Lokales; S. 51

Der Begriff *Dreifach-Katastrophe* begegnet einem in diesem Diskurs mehrere Male, und konkreter in gleicher Weise diese Wendungen:

- (11) „Im japanischen Kulturzentrum in Niederkassel gedachten Japaner und Deutsche der Opfer der *Erdbeben-Katastrophe in Japan*.“

Rheinische Post Duesseldorf, 12. März 2012 Montag

- (12) „In den USA und Kanada mehren sich Befürchtungen, dass Müll von der *Tsunami-Katastrophe in Japan* quer über den Pazifik an die Westküste des amerikanischen Kontinents gespült werden könnte.“

Der Tagesspiegel, Freitag 17. Februar 2012, RUBRIK: WELTSPIEGEL; S. 28 Ausg. 21248

- (13) „Die *Atom-Katastrophe in Japan* könnte nach Ansicht des Münchener Strahlenbiologen und -mediziners Prof. Edmund Lengfelder noch schlimmere Folgen als die von Tschernobyl haben.“

Aachener Zeitung, Montag 14. März 2011, RUBRIK: SPEZIAL; S. 5

Das letzte Beispiel ist nur eine sprachliche Realisierungsform des Desasters. In den Konkordanz tauchen mehrmals die Kombinationen *Reaktor-Katastrophe in Japan*, *nukleare Katastrophe in Japan* und *atomare Katastrophe in Japan* auf, welche sich alle auf das havarierte Kernkraftwerk von Fukushima beziehen. Die Summierung dieser einzelnen Unglücke sorgt in der Presseberichterstattung teilweise für stark emotionalisierende, attributive Bezeichnungen; es ist an einigen Stellen die Rede von der sogenannten *schrecklichen Katastrophe in Japan*, *verheerenden Katastrophe in Japan*, *schweren Katastrophe in Japan*, *dreifachen Katastrophe in Japan* oder *Mega-Katastrophe in Japan*. Des Weiteren tritt folgende Formel mehrmals auf: das *Ausmaß der Katastrophe in Japan*:

- (14) „Um das *Ausmaß der Katastrophe in Japan* zu begreifen, gibt Claus Böttcher vom THW Pinneberg einen anschaulichen Vergleich. In Asien seien rund 80 000 Zivilschutzhelfer im Einsatz. Das sind doppelt so viele wie vor neun Jahren, als das Elbhochwasser bei Dresden wochenlang bekämpft werden musste.“

Hamburger Abendblatt, Freitag 18. März 2011, RUBRIK: REGIONAL; S. 1 Ausg. 65

Die Formel wird verwendet, um zu signalisieren, dass der Schaden respektive die Dimensionen eines Unfalls nicht bestimmbar oder noch nicht „ausgemessen“ sind. Gemessen werden bei einer Katastrophe zumeist die Zahlen der Opfer oder Zahlen zum wirtschaftlichen, finanziellen Schaden. Allgemein jedoch ist in den Konkordanz zum SGM *Katastrophe in*

*Japan* weniger von konkreten Opferzahlen, sondern mehr von (symbolischen) Hilfen für die Opfer die Rede (siehe Beispiel (10) sowie das SGM *Menschen in Japan* weiter unten). Den größten Raum bei den Kontexten zu diesem Muster nehmen aber aufs Neue die Auswirkungen und Folgen ein, welche die Vorfälle in Fernost haben. Dazu gehören die politischen, noch häufiger und entsprechend auffälliger die wirtschaftlichen Folgen für andere Länder und für Deutschland (insbesondere für die Automobilindustrie):

- (15) „In den USA muss die erste Autofabrik wegen der *Katastrophe in Japan* ihren Betrieb einstellen. Einem Pick-up-Werk von General Motors gehen langsam bestimmte Teile aus [aufgrund von Produktionsausfällen dieser Teile in Japan, M. T.]“

Autoflotte Online, 18. März 2011 Freitag

- (16) „Die *Katastrophe in Japan* wird laut Finanzmarktexperte Wolfgang Gerke vor allem die Börsenwerte von Rückversicherern und Energieversorgern treffen. Die übrigen Branchen in Deutschland würden voraussichtlich nicht allzu sehr darunter leiden, sagte Gerke gestern.“

General-Anzeiger (Bonn), Montag 14. März 2011, RUBRIK: POL; S. 4

Auf politischer Ebene betrifft das SGM praktisch nur ein Thema, nämlich die Zukunft der Kernenergie, welche bereits nur wenige Tage nach Beginn des Unglücks von deutschen Politikern infrage gestellt wird. Stellvertretend für die zahlreichen Beispiele sei hier nur eines aufgeführt:

- (17) „Gabriel [SPD, M. T.] sagt, die *Katastrophe in Japan* zeige, dass ‚der Super-GAU keine rein theoretische Rechengröße‘ sei: ‚Was in Japan ein Erdbeben ist, kann in Deutschland auch ein Flugzeugabsturz sein‘.“

Berliner Zeitung, Montag 14. März 2011, RUBRIK: POLITIK; S. 4 Ausg. 61

Als letztes sei an dieser Stelle das sich an dritter Position in der Liste typischer Dreiwort-Einheiten befindende SGM *Katastrophe von Fukushima* ( $G2 = 647,86$ ; Freq. = 457) erwähnt. Dieses SGM bildet quantitativ eine starke Sammelbezeichnung für die Vorfälle im japanischen AKW in der Stadt Fukushima und deren Folgen. Im Diskurs taucht die Kennwortverbindung als Erweiterung der metonymisch verwendeten EWE *Fukushima* auf:

- (18) „Die Landtagswahl, so betont er [Rainer Brüderle, M. T.], stand ganz im Zeichen extremer Entwicklungen: *Fukushima*, Euro-Krise, Libyen. [...] Diese Wahl stand unter ganz besonderen, extremen Vorzeichen: der Krieg in Libyen, die Krise des Euro, die *Katastrophe von Fukushima*.“

Kreis Anzeiger (Germany), Donnerstag 31. März 2011

Der in diesem Zitat dokumentierte Gebrauch des Einzelwortes *Fukushima* ist an vielen Stellen im Korpus belegt. Dabei kommt dem Begriff wie gesagt die Funktion einer Metonymie zu, denn es ist nicht die (gesamte) japanische Großstadt an sich angesprochen, sondern der Ausdruck nimmt meistens auf das havarierte Kernkraftwerk dieser Stadt Bezug (folglich nur auf einen Teil des Ganzen). Die Verweiskfunktion auf das Desaster realisiert in der deutschen Presse im unmittelbaren Zeitraum der Geschehnisse manchmal schon allein das Wort *Japan*, wie diese Stelle andeutet:

- (19) „Lange für seine Wirtschaftskraft und Hochtechnologie bewundert, ist *Japan* zum Synonym für die *Katastrophe von Fukushima* geworden.“

Financial Times Deutschland, Donnerstag 23. Juni 2011, RUBRIK: POLITIK; S. 11

Zusammenfassend kann konstatiert werden, dass mit den wichtigen SGM *Erdbeben in Japan*, *Tsunami in Japan*, *Katastrophe in Japan*, und *Katastrophe von Fukushima* eine Etikettierungsfunktion durch die Printmedien verbunden ist, mit der allgemein auf sämtliche Unglücksvorfälle in Japan verwiesen wird, also auf das starke Erdbeben, dem anschließenden Tsunami und dem sich daraus ableitenden Atomunglück des Kernkraftwerkes in Fukushima.

Damit ist gemeint, dass diese Formeln generalisierende Bezeichnungen sind. Generalisierend deshalb, weil es in der Geschichte Japans schon viele (Natur-)Katastrophen gab, wie z.B. Erdbeben (etwa von Kōbe im Jahr 1995), Tsunamis oder die Atombombenabwürfe auf Hiroshima und Nagasaki. Es liegt aber an der Kurzfristigkeit der printmedialen Agenda, dass mit der Bezeichnung *Katastrophe in Japan* jeweils nur die gegenwärtige gemeint ist. Allein die geografische Information reicht beim Leser aus, um den „richtigen“, den aktuellen Kontext mental zu aktivieren. Derartige verkürzende Bezeichnungen kommen in der Nachrichtenberichterstattung häufiger vor, wenn z.B. die Rede ist von der: *Katastrophe in Haiti* (Frühjahr 2010), *Katastrophe in Südostasien* (der Tsunami im Indischen Ozean von 2004) oder die (*Flut-*) *Katastrophe in Russland* vom Sommer 2012 uvm. – alles Wortverbindungen, welche ihre Referenzfunktion nur temporär und vor dem Hintergrund der aktuell präsenten Geschehnisse erfüllen können.

Die Kontexte zum SGM *Katastrophe von Fukushima* behandelt der folgende Abschnitt, denn es stellt sich heraus, dass besagte Wendung zumeist Komponente eines größeren Musters ist.

### **Das SGM nach (der Katastrophe von) Fukushima**

*Katastrophe von Fukushima* fungiert wie erläutert praktisch als ein Eigenname für ein besonderes Ereignis, welches tiefgreifende negative Folgen hat und Veränderungen nach sich zieht. Es sind auch diese unterschiedlichen Konsequenzen des Desasters, welche den Großteil der unmittelbaren Konkordanz ausmachen und direkt zu einem der prominentesten SGM führen. Die Kontexte des Clusters zeigen, dass es in vielen Fällen Bestandteil der Mehrworteinheit *nach der Katastrophe von Fukushima* ist. In diversen MWE begegnet einem dieses Sprachgebrauchsmuster immer wieder. Es lohnt sich, in diesem Fall einmal die zehn häufigsten Fünfwort-Einheiten zum Type *Fukushima* zu berechnen<sup>64</sup>:

N	Cluster	Freq.	Length
1	NACH DER KATASTROPHE VON FUKUSHIMA	150	5
2	NACH DER ATOMKATASTROPHE VON FUKUSHIMA	98	5
3	NACH DER REAKTORKATASTROPHE VON FUKUSHIMA	83	5
4	NACH DER ATOMKATASTROPHE IN FUKUSHIMA	52	5
5	DER REAKTORKATASTROPHE IM JAPANISCHEN FUKUSHIMA	52	5
6	NACH DER KATASTROPHE IN FUKUSHIMA	44	5
7	DER ATOMKATASTROPHE IM JAPANISCHEN FUKUSHIMA	42	5
8	NACH DER REAKTORKATASTROPHE IN FUKUSHIMA	33	5
9	SUPER GAU IM ATOMKRAFTWERK FUKUSHIMA	26	5
10	DER ATOMKATASTROPHE VON FUKUSHIMA UND	23	5

Abbildung 4: Die zehn häufigsten 5-Wort-Einheiten zum KW *Fukushima*

Mit einer überragenden Auftretenshäufigkeit von 150 im FK steht das SGM *nach der Katastrophe von Fukushima* hier an erster Stelle. Man erkennt ferner deutlich die zu einem Sprachgebrauchsmuster zusammenkommenden, miteinander verwandten MWE-Variationen: *nach der Atomkatastrophe von Fukushima* (Freq. = 98), *nach der Reaktorkatastrophe von Fukushima* (Freq. = 83), *nach der Atomkatastrophe in Fukushima* (Freq. = 52), *nach der Katastrophe in Fukushima* (Freq. = 44) u.e.m. Auch in der Liste mit den hundert typischsten

<sup>64</sup> Die linguistische Software zählt in diesem Zusammenhang die häufigsten kontinuierlichen Wortverbindungen in einem Bereich von fünf Wörtern links und fünf Wörtern rechts der Wortform *Fukushima* im Zentrum.

Dreiwort-Einheiten des FK (siehe oben) findet man viele Bestandteile dieses SGM. Dies ist sogleich ein gutes Beispiel dafür, wie eine corpus-driven-Untersuchung typische kontinuierliche Wortgruppen hervorbringt, mit denen man nicht unbedingt „rechnet“. Denn allein durch seine sprachliche Form her wirkt das Muster unscheinbar, eventuell wäre es beim manuellen Durchsuchen des Korpus, mit der Zielvorgabe nach prominenten Mustern Ausschau zu halten, übersehen worden. Ungeachtet dessen aber gehört es quantitativ zu den wichtigsten Sprachgebrauchsmustern im gesamten untersuchten Diskurs. Überblicksartig kann man dieses umfängliche SGM nun wie folgt zusammenfassen:

[Temporale Präposition]	[Artikel]	[desaströses Ereignis]	[Lokale Präposition]	[Ort]
<i>NACH</i>	<i>DER</i>	<b>KATASTROPHE</b> <i>ATOMKATASTROPHE</i> <i>REAKTORKATASTROPHE</i>	<i>IN</i> <i>VON</i>	<i>FUKUSHIMA</i>

Abbildung 5: Schematische Zusammenfassung des SGM *nach (der Katastrophe von) Fukushima*

Die drei mittleren Elemente wurden deshalb mit einer gestrichelten Linie umrahmt, um anzuzeigen, dass sie auch weggelassen werden können, sprich: zu diesem SGM gehört auch die Zweiwort-Einheit und kürzere Bezeichnung *nach Fukushima* (Freq. = 680). Fett gedruckt wurden die Wörter, welche im Zusammenhang mit der Wortverbindung öfter auftreten. Die Untersuchung der Kontexte der Cluster zeigen, dass oft Auswirkungen des atomaren Desasters auf menschliche Handlungen, Bewertungen oder Einstellungen erfolgen:

- (20) „Mit seiner Meinung hielt Rachel [CDU, M.T.] nicht lange hinterm Berg – und sparte dabei nicht mit Kritik, was die Vorgänge *nach der Katastrophe von Fukushima* betrifft: ‚Ich bin erschüttert, dass das Unglück stattgefunden hat. Aber noch mehr hat mich erschüttert, wie in Japan mit diesem Unglück umgegangen wurde‘.“

Aachener Zeitung, Mittwoch 22. Juni 2011, RUBRIK: LOKALES; S. 19

- (21) „Der Fernsehjournalist Takashi Uesugi ist einer von jenen, die davon erzählen, wie empfindlich der Stromriese Tepco reagiert, wenn dennoch über Unliebsames berichtet wird. [...] *Nach der Katastrophe in Fukushima* kampierte auch Uesugi in der Lobby von Tepco, weil er wissen wollte, was im Reaktor passiert.“

Der Spiegel, 23. Mai 2011, RUBRIK: WISSENSCHAFT+TECHNIK; NUKLEARTECHNIK; S. 114  
Ausg. 21

- (22) „*Nach der Reaktorkatastrophe von Fukushima* trat der Physiker häufig als Experte auf. Gemeinsam mit dem ehemaligen Umweltminister und früheren Unter-Generalsekretär der Vereinten Nationen Klaus Töpfer hat Ranga Yogeshwar [...] nun das Buch ‚Unsere Zukunft – Ein Gespräch über die Welt *nach Fukushima*‘ geschrieben.“

Rheinische Post Duesseldorf, 25. Januar 2012 Mittwoch

Die Konkordanz zeigen klar und deutlich, dass die diskutierten Auswirkungen des Desasters zumeist die politische, insbesondere die energiepolitische Ebene betreffen:

- (23) „Droht nun der große Blackout? Am Donnerstag konnten die Tokioter durch Stromsparen den Kollaps zunächst vermeiden. Doch die Frage bleibt bestehen: Wie will Japan seine Stromversorgung *nach der Katastrophe von Fukushima* sicherstellen?“

Aachener Zeitung, Dienstag 27. Dezember 2011, RUBRIK: TITELSEITE; S. 1

Die meisten politischen Konsequenzen in den Kontexten betreffen dabei die Debatte um die Zukunft der Atomenergie (welche im Zuge der Geschehnisse nun einhellig negativer bewertet wird) in den verschiedensten Ländern:

- (24) „Fast ein Jahr *nach der Katastrophe von Fukushima* sind bereits 52 der 54 Atomkraftwerke abgeschaltet worden. Die Reaktorexpllosion hat das Vertrauen der Japaner in die Kernkraft erschüttert. Doch die Atomlobby wehrt sich.“

Der Spiegel, 27. Februar 2012, RUBRIK: AUSLAND; JAPAN; S. 94 Ausg. 9

- (25) „Ein Jahr *nach der Katastrophe von Fukushima* ist Japan nach wie vor mit Aufräumarbeiten beschäftigt, während weltweit – auch in der Schweiz – über die zukünftige Energiepolitik debattiert wird. Die Mehrheit der Bevölkerung will den Ausstieg aus der Atomenergie.“

SonntagsZeitung, 11. März 2012, RUBRIK: FOKUS; NaN; S. 23

In den häufigsten Fällen jedoch beschäftigen sich die Kontexte dieses SGM mit Deutschland und den radikalen Konsequenzen für die hiesige Atomenergiedebatte:

- (26) „*Nach der Katastrophe von Fukushima* hat die Bundesregierung eine Kehrtwende in der Energiewirtschaft gemacht.“

Welt am Sonntag, 10. Juli 2011, RUBRIK: NRW Ausg. 28

- (27) „Deutschland soll aussteigen, da sind sich die Deutschen *nach der Katastrophe von Fukushima* einig wie selten: 71 Prozent sehen das nach einer Blitzumfrage der ARD inzwischen so, im Herbst vergangenen Jahres waren es nur 51 Prozent. Noch mehr, nämlich 80 Prozent, stimmen dem von der Regierung durchgesetzten Moratorium für alte Atommeiler zu.“

Der Spiegel, 4. April 2011, RUBRIK: TITELGESCHICHTE; S. 62 Ausg. 14

Allen genannten Beispielen ist resümierend folgende pragmatische Funktion gemeinsam, welche in der Deutschen Presse mit diesem SGM realisiert werden soll: Die Vorfälle im japanischen Atomkraftwerk Fukushima Daiichi werden als eine Zäsur bewertet, die einen Wendepunkt in vielen (politischen) Bereichen bedeutet. Die allgemeine Formel *nach (der Katastrophe von) Fukushima* markiert und benennt einerseits den Ausgangspunkt respektive Auslöser für eine Wende und fragt andererseits nach den Konsequenzen, Taten, Lehren und Handlungen, die dieses Ereignis nach sich ziehen (soll), um künftigen, ähnlich drohenden Schaden fortan abzuwenden.

### **Die SGM *Ausstieg aus der (Atomkraft) und für die Energiewende***

An fünfter Position in der Liste typischer MWE befindet sich die Wortgruppe *Ausstieg aus der* mit einer Frequenz von 516 im FK und 27 im RK. Es zeigt sich beim Blick in die Konkordanzen, dass dieses Cluster in den meisten Fällen auch Bestandteil ist der Vier-Wort-Einheiten *Ausstieg aus der Atomenergie* (Freq. = 171), *Ausstieg aus der Kernenergie* (Freq. = 136), *Ausstieg aus der Atomkraft* (Freq. = 111) und *Ausstieg aus der Kernkraft* (Freq. = 31). Hinter diesem SGM verbirgt sich inhaltlich die womöglich radikalste Konsequenz, welche die Rezeption der Katastrophenserie in Japan und insbesondere die Vorkommnisse im AKW Fukushima Daiichi für die deutsche Politik zur Folge hatte: nämlich ein beschleunigter Wechsel von atomar produziertem Strom hin zu anderen Energielieferanten:

- (28) „Die Koalition aus Union und FDP hatte als Konsequenz nach der Atomkatastrophe in Japan vor zwei Wochen beschlossen, dass die sieben ältesten deutschen Atomreaktoren für zunächst drei Monate kurzfristig vom Netz genommen werden müssen. Dies könnte der Auftakt zum endgültigen *Ausstieg aus der Atomkraft* in Deutschland sein.“

abendblatt.de - Hamburger Abendblatt Online, Sonntag 3. April 2011 6:21 PM GMT, RUBRIK: POLITIK

Das Fukushima-Desaster und die antizipierten Folgen eines Atomunfalls, der sich auf deutschem Boden ereignen könnte, löst eine Angst aus, welche die *Atomkraft* im Diskurs fast unisono negativ darstellt<sup>65</sup>, und dessen Nutzungsende nun möglichst schnell herbeigeführt werden soll. Liest man die Konkordanzzeilen vertikal, so fallen die folgenden, unmittelbar links des SGM stehenden Attribute häufig auf; respektive es ist die Rede von einem: *schnellen, baldigen, eingeleiteten, wirklichen, endgültigen, schnellstmöglichen, raschen Ausstieg aus der Atomkraft*, womit der erhöhte politische Druck auf diesen Prozess signalisiert wird:

- (29) „Der Leiter des Fraunhofer-Instituts für Solare Energiesysteme, Eicke Weber, plädierte für einen schnellen *Ausstieg aus der Atomkraft*. Ein Atomunfall gefährde Millionen von Menschen, sagte er. Die Kernkraft sei daher keine Technologie, die Menschen benutzen sollten.“

abendblatt.de - Hamburger Abendblatt Online, Donnerstag 28. April 2011 2:51 PM GMT, RUBRIK: POLITIK

Interessant an der MWE *Ausstieg aus der Atomkraft* ist das metaphorische Bild, welches hier benutzt wird und dem Quellbereich ‚Mobilität‘ entstammt. Für gewöhnlich wird mit *Aussteigen* das Verlassen eines Zuges oder Fahrzeuges assoziiert: jemand möchte nicht mehr mitfahren. Bei einigen Konkordanzen findet man in diesem Kontext auch weitere Mobilitätsmetaphern:

- (30) „spiegel ONLINE: Im ersten Atomkonsens unter Rot-Grün sollten die Meiler länger laufen als 2022. Ströbele: Damals hatten wir einen Koalitionspartner, und wir hatten auch keine atomare Zeitenwende wie Fukushima. Der *Ausstieg aus der Atomkraft* und der Umstieg auf erneuerbare Energien geht viel schneller, als wir 2001 geahnt haben.“

Spiegel Online, 25. Juni 2011 Samstag 7:53 AM GMT+1, RUBRIK: GRÜNEN-POLITIKER STRÖBELE

Hier wird also von einem *Umstieg* gesprochen, an anderen Stellen im Diskurs ist von einem *Einstieg* in alternative Energielieferanten die Rede. Das Wort *Aussteigen* suggeriert die Option, überhaupt nicht mehr involviert zu sein. Dass dieser übertragene Gebrauch auch Probleme mit sich bringen kann, deutet folgendes Beispiel an:

- (31) „Die zigtausend Tonnen Atommüll strahlen auch nach dem so genannten *Ausstieg aus der Atomkraft*“, sagt Lagatom-Sprecher Dirk Werner, „eine wirkliche Lösung des Endlagerdebakels ist nicht im Ansatz in Sicht“.

Hamburger Abendblatt Online, Montag 27. Februar 2012 3:02 AM GMT, RUBRIK: HAMBURG

Entscheidend ist hier die metasprachliche Markierung *so genannten*. Steigt man aus etwas aus, so ist es in vielen Fällen ein harmloses Unterbrechen einer Reise; steigt man hingegen *aus der Atomkraft* aus, so sagt es der Lagatom-Sprecher, ist die Reise bildlich gesehen noch lange nicht vorbei. Oder anders ausgedrückt, und um die Mobilitätsmetapher noch einmal zu bemühen: der gefährliche Zug mit der nuklearen Fracht (als Sinnbild für die noch Jahrtausende bis Millionen Jahre andauernde Radioaktivität), aus dem jemand ausgestiegen ist, fährt auch ohne seine Passagiere weiter.

Stellt das Wort *Atomkraft* in diesem Diskurs ein ‚Unwertwort‘ nach Hausmann (1994) dar, dessen Kontexte negativ sind, was sich allein an der Oberfläche des besagten SGM *Ausstieg aus der Atomkraft* ablesen lässt, so stehen dem gegenüber Types wie *Energiewende* (Position 14 in der EWE-Liste; 2052 Nennungen im FK) oder *Energien* (Position 21; Freq. = 1864) und in Verbindung hiermit SGM wie *der erneuerbaren Energien* (317), *die*

<sup>65</sup> Unter den häufigsten Dreiwort-Einheiten in Verbindung mit der auf Rang 15 positionierten typischen EWE *Atomkraft* sind bspw.: *aus der Atomkraft* (Freq.=211), *Ausstieg aus der* (118), *gegen die Atomkraft* (74), *von der Atomkraft* (65), *Nutzung der Atomkraft* (49), *Atomkraft nein danke* (47). Nicht nur in den Konkordanzen dieser einzelnen MWE, sondern bereits an deren Form erkennt man die negative Konnotation.

*erneuerbaren Energien* (145), für die *Energiewende* (Freq. = 115; mit großem Abstand die häufigste Dreiwort-Verbindung, welche den Type *Energiewende* als Bestandteil enthält).

(32) „Nach Einschätzung des Energieexperten Uwe Leprich von der Hochschule für Technik und Wirtschaft des Saarlandes müssen sich Eon und RWE neu erfinden. Die Konzerne hätten lange auf Großkraftwerke gesetzt, mit denen das meiste Geld zu verdienen gewesen sei. Mit dem Aufwind *der erneuerbaren Energien* hätten jetzt aber mittelgroße Unternehmen gute Chancen.“

Kölnische Rundschau, Donnerstag 11. August 2011, RUBRIK: Wirtschaft; S. 8

(33) „Jürgen Großmann hat einiges bewegt in seinen bisherigen vier Jahren als Vorstandsvorsitzender von RWE. Er hat mehrere Milliarden Euro in die Hand genommen und den von seinen Vorgängern verschlafenen Einstieg in *die erneuerbaren Energien* in Angriff genommen.“

Börsen-Zeitung, Dienstag 9. August 2011, RUBRIK: KOMMENTAR - RWE; S. 1 Ausg. 151

So wie in diesen beiden Exempeln weisen die Kontexte *die erneuerbaren Energien*<sup>66</sup> als positiv aus und beschäftigen sich oft mit den (politischen wie wirtschaftlichen) Prozessen, die einen Wandel hin zu diesen alternativen Energielieferanten herbeiführen können. Dieser, dem Diskurs nach wünschenswerte Prozess findet sich auch in der sprachlichen Struktur etwa des sehr frequenten Musters *für die Energiewende* wieder. Wäre dies nicht so oder eine derartige Entwicklung allgemein negativ besetzt, so würde hier eventuell das Muster *gegen die Energiewende* häufiger vorkommen. Dieses taucht aber nur ganze 11 Mal im kompletten Korpus auf.<sup>67</sup> Und schon allein dessen Konkordanzzeilen zeigen, dass sich hinter der Formel meistens nicht wirklich ernsthafte Einwände und Gegenmeinungen verbergen:

N	Concordance
1	sbau gefällt werden. Ich bin nicht <i>gegen die Energiewende</i> , aber ich trete für das Machbare
2	rde richte sich ausdrücklich nicht <i>gegen die Energiewende</i> , betonte der E.on-Sprecher. Aber
3	s ist aber natürlich kein Argument <i>gegen die Energiewende</i> , sondern eher ein mahndes Pläd
4	zent festgestellt haben, als Beleg <i>gegen die Energiewende</i> interpretieren? Überhaupt nicht.
5	uhe richte sich ausdrücklich nicht <i>gegen die Energiewende</i> , hatte das Unternehmen betont. D

Abbildung 6: Konkordanzzeilen zum SGM *gegen die Energiewende*

Stattdessen wird aufgezeigt, und das ist die eigentliche pragmatische Funktion des SGM *für die Energiewende*, mit welcher diversen politischen Bestrebungen, der Vollzug einer *Energiewende* günstig beeinflusst werden kann:

(34) „RWE und Eon wollen in dieser Woche die Weichen *für die Energiewende* stellen.“

Kölnische Rundschau, Montag 8. August 2011, RUBRIK: Wirtschaft; S. 6

(35) „Eine eigens eingesetzte Ethikkommission soll bis Ende des Monats politische Empfehlungen *für die Energiewende* vorlegen.“

Die Welt, Dienstag 10. Mai 2011, RUBRIK: POLITIK; S. 5 Ausg. 108

<sup>66</sup> Der Begriff *erneuerbare Energien* ist positiv konnotiert, weil er sein Augenmerk auf die künftige Verfügbarkeit der Energielieferanten (Sonne, Wind, Wasserkraft, Erdwärme usw.) richtet, die im Gegensatz zu den sog. *fossilen Brennstoffen* (Kohle, Öl,...) kein begrenztes Reservoir haben. Ob solche künftigen Energieformen daher, wie der Begriff suggerieren mag, bedingungslos positiv sind und nicht vielleicht andere ökologische, politische, wirtschaftliche oder soziale Nachteile mit sich bringen, bleibt abzuwarten.

<sup>67</sup> Darunter befinden sich jedoch auch Mehrfachnennungen, welche auf von LexisNexis generierte Duplikate zurückgehen.

In den Positionen unmittelbar links vom Cluster *für die Energiewende* befinden sich häufig diverse Pläne/Handlungen/Entitäten, welche nach Meinung unterschiedlicher sozialer Akteure für eine gelingende Umsetzung dieses Prozesses vonnöten sind. Neben den *Weichen* oder *Empfehlungen für die Energiewende* illustrieren dies ein paar weitere kurze, im Diskurs auftauchende Wörter: *Rohstoffe, Geld, Verantwortung, Drehbuch, Beauftragten* oder *Masterplan für die Energiewende*.

Das Wort *Energiewende* enthält eine Metapher, welche im gleichen Maße dem Quellbereich ‚Mobilität‘ entspringt, wobei mit *Wende* bekanntlich beim Gehen, Fahren, Reiten, Fliegen etc. ein plötzlicher und auch starker Richtungswechsel (zumeist um 180 Grad in die komplette Gegenrichtung) gemeint ist. Erneut ist fraglich, ob sich die mit diesem Wort transportierte bildliche Bedeutung eines sofortigen Umschwungs auf den Zielbereich eines realen (längerfristigen) Wandels in der Energieerzeugung so leicht übertragen lässt. Ein metasprachliches Konkordanzbeispiel aus dem untersuchten Diskurs selbst zeigt, dass man dieses Wort nicht unbedingt kritiklos hinnehmen muss:

- (36) „Die ‚*Energiewende*‘ ist ein schönes Wort: Die Bezeichnung ‚Wende‘ suggeriert ein leichtes Umschalten wie von Zauberhand und einen schnellen Wandel. Doch zeigt die Technikgeschichte, dass das Energiesystem zu den langsamsten Strukturen der Entwicklung gehört. Eine ‚Wende der Energie‘ ist ein Widerspruch in sich – eine ganz und gar ideologische Konstruktion, die keine Kilowattstunde produziert.“

Welt kompakt, Dienstag 17. Januar 2012, RUBRIK: FORUM; Leitartikel Ausg. 12

Bemerkenswert ist die Begründung „eine ganz und gar ideologische Konstruktion“, die praktisch schon den Kern des Diskursbegriffes mit seiner Idee von der sozial bedingten Konstitution von Wissen durch Sprache und durch bestimmte Wörter (wie das betreffende Exempel) trifft. Insgesamt jedoch weisen die positiven Konnotationen zur *Energiewende* dieses Kenn-Wort resp. Schlagwort des Fukushima-Diskurses eindeutig als ein ‚Fahnenwort‘ bzw. ‚Hochwertwort‘ in der Terminologie Hausmanns (1994) aus.

### **Das SGM im Atomkraftwerk Fukushima**

Die MWE *im Atomkraftwerk Fukushima* rangiert mit einer Frequenz von 232 sowie einer Log-Likelihood von 328,89 auf 14. Position in der Tabelle der Dreiwort-Einheiten. Bereits wenige Einsichten in die Kontexte und unmittelbaren Konkordanzen links von diesem SGM zeigen, dass damit ein Ort gemeint ist, der das Zentrum eines Desasters beschreibt, wie folgende Beispiele belegen: *Unfall im Atomkraftwerk Fukushima, Störfall im Atomkraftwerk Fukushima, Super-GAU im Atomkraftwerk Fukushima, Explosion im Atomkraftwerk Fukushima, Pannenserie im Atomkraftwerk Fukushima, Kernschmelze im Atomkraftwerk Fukushima, Katastrophe im Atomkraftwerk Fukushima*. Das Muster wird als ein Präzedenzfall für die Schattenseiten der zivilen Atomkraft benutzt:

- (37) „Eine Störung *im Atomkraftwerk Fukushima* löste eine der größten Nuklearkatastrophen der Geschichte aus.“

Berliner Morgenpost Online, Dienstag 10. Mai 2011, RUBRIK: WIRTSCHAFT

Bis es zu dieser Aussage überhaupt kam, bzw. auf dem Weg dorthin interessiert sich die Berichterstattung für alle technischen, menschlichen und physikalischen Vorgänge in dem AKW, immer mit der Angst im Hintergrund vor dem „Größten Anzunehmenden Unfall“ (GAU), in welchen eine Störung in einem Kernkraftwerk münden kann:

- (38) „*Im Atomkraftwerk Fukushima* droht nach Angaben des Regierungssprecher Yukio Edano eine Kernschmelze in drei Reaktoren.“

Berliner Morgenpost Online, Montag 14. März 2011 10:44 PM MET (GMT+1), Rubrik: VERMISCHTES

(39) „Im Kampf gegen die Zeit versuchen Techniker und Ingenieure, den Super-GAU *im Atomkraftwerk Fukushima* Daiichi noch zu verhindern. Sie riskieren durch die Strahlenbelastung ihr Leben.“

Stuttgarter Zeitung, Freitag 18. März 2011, RUBRIK: INNENPOLITIK; S. 4

Es entsteht beim Lesen der Konkordanzzeilen zu diesem Muster der starke Eindruck, dass das Innere des AKWs ein Kern des Chaos darstellt, welches unkontrolliert ist, höchstgefährlich und von außen nur schwer beherrsch- und einsehbar. Mehrmals ist die Rede von der ungewissen Lage und Situation im Kraftwerk:

(40) „Die grösste Sorge ist die noch immer unklare Situation *im Atomkraftwerk Fukushima*.“

Stocks, 18. März 2011, RUBRIK: MÄRKTE; S. 10

(41) „Physiker Helmut Hirsch hält die Lage *im Atomkraftwerk Fukushima* Eins nicht für aussichtslos, sie bleibe aber unkalkulierbar.“

Aachener Zeitung, Freitag 18. März 2011, RUBRIK: DIE SEITE DREI; S. 3

Die oben genannten Beispiele zeigen, was noch typisch für dieses SGM ist, nämlich neben den technisch-physikalischen Vorgängen der „menschliche Anteil“ – soll heißen: die Versuche der Arbeiter, das „Chaos“ einzudämmen und das daraus resultierende Leid (z.B. infolge radioaktiver Belastung). Da die Bemühungen, diese so gering wie möglich zu halten, allein nicht ausreichen, ändert die japanische Regierung kurzerhand die Definition dessen, was überhaupt als radioaktive Belastung respektive Limit gilt, wie folgende interessante Textstelle zeigt:

(42) „Die normale Dosis für Arbeiter in Atomkraftwerken sei 20 Millisievert, nie jedoch über 50 Millisievert pro Jahr, erklärt der australische Nuklearberater Tony Irwin. ‚Sie sollten die Leute also rotieren lassen, um unterhalb des Grenzwerts zu bleiben. Es gibt nicht viele Länder, die ein Notfalllimit von 100 Millisievert pro Jahr zulassen‘, sagt Irwin. Das japanische Gesundheitsministerium erhöhte den Grenzwert für die Arbeiter *im Atomkraftwerk Fukushima* Daiichi am Mittwoch jedoch von 100 auf 250 Millisievert. Wegen der besonderen Umstände sei dieser Schritt unvermeidlich gewesen, teilte das Ministerium mit.“

Stuttgarter Zeitung, Freitag 18. März 2011

Dies ist ein extremes Beispiel dafür, wie der Diskurs die Wirklichkeit definiert und festlegt, was (in Hinblick auf Radioaktivität) als Gefahr anzusehen ist und was nicht. Zusammenfassend gesagt, offenbaren die vielen Konkordanzbeispiele, dass die Lokalisierungsformel *im Atomkraftwerk Fukushima* gewissermaßen den Status eines weiteren Epizentrums der Dreifach-Katastrophe von Japan hat; zumindest in der Wahrnehmung der deutschen Presse. Nicht nur im Erdbeben, sondern auch hier, im besagten AKW, scheint der „Kern des Übels“ zu liegen, welcher seiner Umgebung und Umwelt Leid und Schaden zufügt. Im Gegensatz zu einem Erdbeben, welches plötzlich und menschlich unverschuldet auftritt, was bereits die Bezeichnung *Naturkatastrophe* ausdrückt, ist der menschliche Anteil an einer *Atomkatastrophe* sehr viel größer. Das oszillierende, letzten Endes aus den Fugen geratende Gleichgewicht zwischen den Störfällen *im Atomkraftwerk Fukushima*, die sich zum nuklearen Desaster ausweiten könnten und den Bemühungen und Widrigkeiten der Arbeiter und Techniker (bzw. deren Ohnmacht), die dagegen anzukämpfen versuchen; genau dieser Widerstreit zwischen unkontrollierter Technik und um Kontrolle bemühte Menschen macht das SGM aus.

### **Das SGM Menschen in Japan**

Untersucht man die Konkordanzen zum SGM *Menschen in Japan* (Freq.=142 im FK und 6 Mal im RK), so wird relativ klar, dass mit dieser Sammelbezeichnung insbesondere jene Einwohner Japans gemeint sind, welche von dem Erdbeben, dem Tsunami oder der

Fukushima-Katastrophe betroffen sind. Es wird also speziell auf die Opfer referiert bzw.: als solche werden die Japaner dargestellt; als wehrlose Menschen, die von dem Desaster unerwartet getroffen wurden:

(43) „Das Ausmaß der Katastrophe lässt sich noch nicht wirklich erfassen. Auch das Leid, das über die *Menschen in Japan* durch das verheerende Erdbeben hereingebrochen ist.

Rheinische Post Duesseldorf, 14. März 2011 Montag

Das wirklich kennzeichnende Element dieses Sprachgebrauchsmusters ist, dass damit eine Solidarisierung erreicht werden soll. Fast alle Konkordanzbeispiele zeigen, dass sich hinter der Formel *Menschen in Japan* Mitgefühl und Hilfsbereitschaft in der einen oder anderen Form verbirgt:

(44) „Die atomare Katastrophe nach dem verheerenden Erdbeben und Tsunami in Japan beschäftigt auch die Menschen im Kreis Pinneberg intensiv. Der Kreistag hat seine Marathonsitzung am Mittwochabend mit einer Gedenkminute eingeleitet, um der Tausenden von toten, vermissten und obdachlosen *Menschen in Japan* zu gedenken und seine Anteilnahme auszudrücken.“

Hamburger Abendblatt, Freitag 18. März 2011, RUBRIK: REGIONAL; S. 1 Ausg. 65

Überlegt man sich für dieses Sprachgebrauchsmuster paradigmatische Konstellationen wie *Einwohner Japans, Leute in Japan* etc., so wird klar, dass mit der Formel *Menschen in Japan* die unschuldige Opferrolle stärker akzentuiert wird. Die Bezeichnung signalisiert, dass es sich bei den Opfern der Katastrophe gewissermaßen um *Menschen* „wie du und ich“ handelt, wobei die nachgelagerte Wortgruppe *in Japan* lediglich eine geografische Information darstellt, bei der die Nationalität des leidenden Menschen stärker in den Hintergrund gerückt werden soll. Wesentlich ist, dass diese Wortverbindung fast immer in einem „Hilfekontext“ (Gedanken bei bzw. Hilfe für die *Menschen in Japan*) benutzt wird.

### **Die SGM *Tokyo Electric Power* bzw. *AKW-Betreiber Tepco***

Die EWE *Tepco* ist in der Liste mit den typischen EWE auf Position 9. Das Akronym findet sich ausgeschrieben auch unter den typischen Dreiwort-Einheiten wieder: *Tokyo Electric Power* (119 Mal im FK). Und was hiermit wiederum gemeint ist, zeigt ein Beispiel aus dem Diskurs selbst:

(45) „Auch Asiens größter Stromversorger *Tokyo Electric Power* (Tepco), der drei große Nuklearkomplexe mit 17 Reaktorblöcken betreibt, stand immer wieder wegen der Informations-politik am Pranger.“

Frankfurter Rundschau, Montag 14. März 2011, RUBRIK: POLITIK; S. 8

Nicht immer vertreten in den Konkordanzzeilen zu diesem SGM ist unmittelbar danach das Wort „company“. Möglicherweise wird die Kenntnis von der Bedeutung des Kürzels *Co* allgemein vorausgesetzt. Die unmittelbare pragmatische Funktion dieses SGM ist beinahe zu banal, um erklärt zu werden: es dient dazu dem Leser der deutschen Presse den Namen eines ihm höchstwahrscheinlich unbekanntem japanischen Konzerns klar zu machen. Der Kontext von fast allen Konkordanzen zu diesem SGM ist inhaltlich praktisch derselbe wie in den Konkordanzen zur zweithäufigsten Dreiwort-Einheit im Zusammenhang mit dem Type *Tepco*, nämlich *AKW-Betreiber Tepco* (74 Mal im FK). Die sprachliche Oberfläche dieses Musters gibt hier bereits den inhaltlichen Kontext der Konkordanzen vor: *Tepco* ist der Energieversorger, welcher, das AKW Fukushima Daichii betreibt und somit für dessen Finanzen und Sicherheit zuständig ist. Damit wird ein eindeutiger Fokus auf einen ganz gezielten Tätigkeitsbereich des Konzerns gelegt, er wird diskursiv geprägt, denn das Unternehmen unterhält zugleich viele dutzend Wasser- und Wärmekraftwerke. Wie bereits das erste Beispiel oben andeutet, enthält die Berichterstattung über den Konzern überwiegend Negatives; das betrifft insbesondere deren Informationspolitik, die (gescheiterten) Versuche

zur Eindämmung der Kernschmelze in Fukushima, der daraus resultierende Image- wie Wirtschaftsschaden für den Konzern usw.:

- (46) „AKW-Betreiber *Tepco* unter Beschuss. [...] Einer der größten Energieversorger der Welt hat schon in den vergangenen Jahren gegen viele Affären und Pannen angekämpft. Pannen, Lügen und Vertuschungsaffären. All das pflasterte den Weg eines der größten Energiekonzerne der Welt: der *Tokyo Electric Power*, kurz *Tepco*. Der negative Höhepunkt ist nun der Zusammenbruch des Atomkraftwerks Fukushima.“

Die Presse, Montag 14. März 2011.

Auf eine schlichte Formel gebracht, besteht die Funktion des SGM *AKW-Betreiber Tepco* darin, mit lediglich drei Wörtern einen der wichtigsten sozialen Akteure des hier untersuchten Diskurses mit einem der wichtigsten Objekte bzw. Gegenstände desselben, nämlich dem von enormen Störfällen betroffenen Atomkraftwerk in Fukushima kausal zu verknüpfen.

### Weitere SGM sozialer Akteure

Einige weitere SGM stellen Institutionen oder Personen dar, welche im Diskurs offenbar eine tragende Rolle spielen und sich in der generierten Liste wiederfinden: darunter u.a. *die japanische Regierung* (Freq. = 289 im FK), *Ministerpräsident Naoto Kan* (170), *Regierungssprecher Yukio Edano* (120), *der japanischen Regierung* (115), *Regierung in Tokio* (111), *die Nachrichtenagentur Kyodo* (132) sowie auch *Bundeskanzlerin Angela Merkel* (Freq. = 270<sup>68</sup>). Im folgenden Zitat finden sich zufällig mehrere dieser Muster in ein und demselben Artikel wieder:

- (47) „Einen Tag nach der angekündigten Kehrwende in der Energiepolitik macht *die japanische Regierung* bereits wieder einen Schritt zurück: Der von *Ministerpräsident Naoto Kan* in Aussicht gestellte Atomausstieg nach der Fukushima-Katastrophe sei nicht offizielle Regierungslinie, sagte sein *Sprecher Yukio Edano* am Donnerstag. Eine Gesellschaft ohne Atomkraft sei lediglich eine ‚*Hoffnung für die ferne Zukunft*‘, zitierte *die Nachrichtenagentur Kyodo* Edano.“

Spiegel Online, 14. Juli 2011 Donnerstag 1:32 PM GMT+1, RUBRIK: KONFLIKT UM KEHRTWENDE

Dieses Beispiel steht paradigmatisch für den gemeinsamen Inhalt vieler Konkordanzanzen zu den genannten SGM: auch sie werden in den meisten Fällen in einem direkten oder indirekten Zusammenhang mit der Katastrophenserie, insbesondere dem atomaren Vorfall in Japan gestellt. Zum einen erfüllen die SGM die pragmatische Funktion als eine von Journalisten genutzte Quellenbezeichnung, um (nicht selten) die Gewähr für die Richtigkeit dargebotener Fakten und Informationen entsprechend zu relativieren. Dies trifft gemäß den Konkordanzanzen nicht nur für *die Nachrichtenagentur Kyodo* zu, sondern auch für alle anderen offiziellen japanischen Akteure:

- (48) „Die Informationen aus Japan widersprechen sich. Fest steht aber: *Nach Angaben der*<sup>69</sup> *Betreiberfirma* ist die radioaktive Strahlung in der Unglücksanlage gestiegen. Grund könnte der Beginn einer Kernschmelze sein. *Nach Angaben der Regierung in Tokio* droht ein solcher Prozess in drei der insgesamt sechs Blöcke von Fukushima I.“

Spiegel Online, 14. März 2011 Montag 6:31 PM GMT+1, RUBRIK: AKW FUKUSHIMA

---

<sup>68</sup> Trotz dieser hohen Frequenz befindet sich das SGM *Bundeskanzlerin Angela Merkel* „lediglich“ an 89. Position der typischen Dreiwort-Einheiten im FK. Dies hängt damit zusammen, dass diese MWE im RK ebenfalls immerhin 68 Mal vorkommt und der Log-Likelihood-Wert mit 135,55 entsprechend niedriger ist (jedoch noch immer hochsignifikant). Als einer der wichtigsten politischen Akteure ist die deutsche Kanzlerin auch in vielen anderen Diskursen erwähnt.

<sup>69</sup> Auf Position 85 in der Typikliste der MWE steht das SGM *nach Angaben der* (Freq. = 303). Eine Formel, welche durch eine gewisse Festigkeit charakterisiert ist, da sie im medialen Kontext sehr häufig auftaucht. Dass sie für den hiesigen Fukushima-Diskurs typisch ist, könnte als ein weiteres Indiz für die Unsicherheit gegenüber den berichteten Fakten zu den Ereignissen in Japan gedeutet werden.

Bezüglich der preisgegebenen Informationen, jedoch insbesondere in Hinsicht auf den Umgang bzw. das allgemeine Krisenmanagement in Japan werden die genannten Akteure im Diskurs hierzulande publikationsübergreifend und in der Mehrzahl eindeutig in einen negativen Kontext gesetzt:

(49) „Die *Regierung in Tokio* versagt. Nach Beben, Flut und dem Atomkraftwerk-GAU in Fukushima wird der Wiederaufbau vor allem von Unternehmern getragen.“

manager magazin, 20. Mai 2011, RUBRIK: TRENDS; Japan; S. 94 Ausg. 6

Stets wenn sich ein katastrophales Geschehen ereignet und in der Konsequenz großer Schaden (für Menschen, Umwelt, Wirtschaft etc.) entsteht, sind wichtige Entscheidungen gefragt: z.B. in der Organisation von Hilfsmaßnahmen, der Finanzierung des Wiederaufbaus, der Entschädigung von Betroffenen, der korrekten Informationspolitik oder im künftigen Umgang mit ähnlichen Bedrohungen zwecks eventueller Prävention. Entscheidungen dergestalt werden von den Verantwortlichen, d.h. der (politischen) Führung erwartet. Aus diesem Grund unterliegt deren Handeln einer ständigen (kritischen) Beobachtung. Dieser Umstand kommt am Beispiel des sich in Japan abspielenden Desasters in dem dafür zuständigen obersten Akteur, dem SGM *Regierung in Tokio* semantisch-pragmatisch zum Ausdruck.

### **SGM zu den ökonomischen Auswirkungen**

Die corpus-driven-Berechnung hat sprachliche Muster hervorgebracht, welche dem wirtschaftlichen Kontext angehören. Hierzu gehören etwa SGM wie *Die japanische Wirtschaft, im ersten Quartal, im zweiten Quartal* sowie zahlreiche typische Einworteinheiten wie *Wirtschaft, Mrd* (in Konkordanzen meist im Kontext von Währungsangaben), *Quartal* aber auch konkrete Unternehmen wie *Eon, EnBW, Siemens* etc. Es zeigt sich in zahlreichen Zusammenhängen derartiger SGM, dass ökonomische Fakten und Entwicklungen oftmals in direktem Zusammenhang mit dem Desaster in Japan gestellt werden:

(50) „Die *japanische Wirtschaft* hat den größten Schock nach dem Tsunami-Desaster verdaut. Erstmals seit vier Quartalen vermeldete die Regierung wieder Wachstum.“

Spiegel Online, 14. November 2011 Montag 8:25 AM GMT+1

(51) „Der Schweizer Versicherungskonzern Zurich Financial Services (ZFS) konnte seinen Gewinn im *zweiten Quartal* insbesondere durch einen Beteiligungsverkauf deutlich steigern. Trotz hoher Belastungen durch Naturkatastrophen legte das Ergebnis um 88 % auf 1,3 Mrd. Dollar zu. [...] In der größten Unternehmenssparte, der Schaden- und Unfallversicherung, sackte das Betriebsergebnis um ein Fünftel auf 1,1 Mrd. Dollar ab. Dies lag in erster Linie an den hohen Schäden durch die Erdbeben in Japan und Neuseeland (mehr als 500 Mill. Dollar) und durch die Wetterereignisse in den USA (Gesamtschaden 200 Mill. Dollar).“

Börsen-Zeitung, Freitag 12. August 2011, RUBRIK: S. 5 Ausg. 145

So wie hier der Konzern Zurich Financial Services werden in den Artikel häufig große Versicherungsunternehmen thematisiert (Munich Re u.a.). Sie spüren demnach die wirtschaftlichen Auswirkungen in erster Linie aufgrund von Verwüstungen durch *Naturkatastrophen*, wie es explizit in dem entsprechenden Artikel dazu heißt. Die atomaren Unglücksvorfälle in Japan haben daran nur wenig Anteil. Sofern speziell der GAU in Fukushima finanzielle Konsequenzen hat, sind es vorwiegend deutsche Unternehmen, die wegen der politisch beschlossenen Abkehr von atomar erzeugter Energie in der BRD diverse wirtschaftliche Konsequenzen verzeichnen (etwa in den Konkordanzen zu *Siemens* oder den deutschen Stromkonzernen). Somit indizieren diese SGM, dass die Reihe der Desaster in Japan sowohl für die Privat- als auch für die Volkswirtschaft als Ganzes (von verschiedenen Ländern) einen Impaktfaktor abgeben, der als Erklärungsmuster für bestimmte Schwingungen, Kursvariationen und andere ökonomische Entwicklungen herangezogen wird.

## 6 Zusammenfassung der Ergebnisse – Hypothesenbildung

Die gefundenen Sprachgebrauchsmuster und Phänomene werden hier in Form von Hypothesen zum Fukushima-Diskurs kurz zusammengefasst. Dabei verstehen sich die folgenden Aussagen nicht als endgültig und absolut, sondern als ein verallgemeinernder Vorschlag bezüglich der inhaltlichen Kernpunkte. Die Ergebnisse, sowohl der berechneten typischen Einwort- wie Mehrwort-Einheiten und der daraus abgeleiteten SGM führen unmittelbar zu einer ersten zentralen Beobachtung über die Eigenart des Fukushima-Diskurses:

1. Die Unglücksserie in Japan vom März 2011 wird in der Deutschen Presse hauptsächlich mit Blick auf Fukushima und die (dadurch angefachte) Atomenergiedebatte geführt.

Diese Tatsache zeigt sich, wenn man allein die gewonnenen Kenn-Wörter betrachtet, welche zum großen Teil alle im Umfeld entweder der Ereignisse von Fukushima oder aber der dadurch in Deutschland entfachten (Atom-)Energiedebatte stehen. Dazu gehören u.a. *Fukushima, Atomkraft, Radioaktivität, Strom, Tschernobyl, RWE, Atomausstieg, Reaktoren, Reaktor, Energiewende*. Auch in den spezifisch sprachoberflächlichen Ausformungen der Mehrwort-Einheiten zum einen aber auch in deren inhaltlichen Kontexten zum anderen spiegelt sich diese Gewichtung auf den ‚Atomdiskurs‘ im weitesten Sinne wieder, was u.a. die analysierten SGM (*Ausstieg aus der Atomkraft* und *im Atomkraftwerk Fukushima*) bereits zeigen. Hinzu kommen eine Fülle weiterer in der Liste typischer MWE, z.B.: *aus der Kernenergie, Atomkatastrophe in Fukushima, Atomkraftwerk Fukushima Daichii, der Präfektur Fukushima, auf erneuerbare Energien* u.v.m.

2. Die Katastrophe in Japan allgemein, jedoch die Vorfälle in Fukushima im Besonderen werden als eine starke Zäsur begriffen.

Diese Beobachtung wurde anhand eines der quantitativ wichtigsten SGM in diesem Diskurs ausgemacht, nämlich *nach (der Katastrophe von) Fukushima*, auf welches man im Korpus in zahlreichen weiteren Abwandlungen stößt.

3. Es wird auf sprachlicher Ebene oft ein Gegensatz aufgemacht zwischen der negativ konnotierten *Atomkraft* und der positiv konnotierten *Energiewende*.

Dies drückt sich bereits an der sprachlichen Oberfläche in SGM aus wie (*Ausstieg aus der Atomkraft, für die Energiewende, für den Atomausstieg, aus der Atomkraft, aus der Kernenergie, auf erneuerbare Energien* usw.). Die Konkordanz weisen die *erneuerbaren Energien* und *Energiewende* eindeutig als Fahnenwörter, *Atomkraft* und *Atomenergie* hingegen als Unwertwörter im Diskurs aus.

4. Die quantitativ wichtigsten sozialen Akteure des Fukushima-Diskurses *Tepco* und die *Regierung in Tokio* werden oftmals negativ kontextualisiert.

Dies betrifft, wie zahlreiche Textstellen belegen, die Informationspolitik, Bewertungen und Handlungen der genannten Akteure im Umgang mit der Unglücksserie in Japan.

5. Die desaströsen Vorfälle in Japan werden sehr häufig als ein Impakt bzw. Einflussfaktor begriffen, der wirtschaftliche Auswirkungen auf bestimmte Unternehmen oder Volkswirtschaften hat.

Letzteres erkennt man an der sprachlichen Oberfläche der SGM *Mrd Yen, Mrd Euro, Eon und RWE, AKW-Betreiber Tepco, im ersten Quartal, im zweiten Quartal* etc. Musterübergreifend tauchen die ökonomischen Konsequenzen (für die Wirtschaft Deutschlands und der Welt) der desaströsen Vorfälle in Japan inhaltlich in den Konkordanzen immer wieder auf.

6. Die Geschehnisse in Japan werden hinsichtlich ihrer direkten Zerstörungen der japanischen Infrastruktur und der Schilderung von Opferschicksalen hierzulande verhältnismäßig gering rezipiert.

Bezüglich dieser Hypothese lässt sich sagen, dass trotz der wirklich immensen Opferzahlen, tausender Vermisster, Verletzter und persönlichen Leidensgeschichten, welche mit den Naturkatastrophen in Japan verbunden sind, die quantitative Widerspiegelung dessen im hier untersuchten Diskurs stark verschoben sind „zugunsten“ der Darstellungen der Ereignisse im japanischen Kernkraftwerk Fukushima Daichii. Ohne Zweifel wird der langfristige Schaden des zerstörten Reaktors für Mensch und Umwelt exorbitant sein. So wird noch immer eine Sperrzone mit einem Radius von 20 km um das Akw herum aufrecht erhalten. Doch gibt es bis zum gegenwärtigen Zeitpunkt noch keinen offiziell an der Strahlung Fukushimas gestorbenen Menschen. Dieser Fokus der Deutschen Presse auf den Unglücksreaktor hat zum einen etwas mit der in der BRD traditionsreichen Atomenergie-debatte zu tun. Zum anderen aber spielt generell unter psychologischem Gesichtspunkt gewiss auch die Gefahrenantizipation der hiesigen Diskursgemeinschaft eine wichtige Rolle: Erdbeben mit einer so hohen Magnitude und derart zerstörerische Tsunamis wie in Japan sind hierzulande fast unmöglich – eine atomare Havarie hingegen ist im Vergleich dazu in Deutschland sehr viel wahrscheinlicher.

## 7 Fazit

Zunächst einmal ein paar Aussagen zu den allgemeinen Gütekriterien bezüglich einer korpuslinguistischen Diskursanalyse (vgl. Spitzmüller/Warnke 2011, S. 199) anhand der durchgeführten Untersuchung. Was die technische Ermittlung der Mehrwort-Einheiten anbelangt, so wurde klar versucht, darzulegen, nach welchen Parametern dies geschah. Wenn letztere gleichbleiben und das Verfahren unabhängig wiederholt wird, so werden unter den besprochenen Prämissen die gleichen Muster von einer Software ermittelt, wie in dieser Arbeit. Die Resultate des corpus-driven-Verfahrens in dieser Arbeit können somit als ‚verlässlich‘ im Sinne einer Reproduzierbarkeit gelten. Die ständige theoretische Reflexion und Erläuterung der gewählten Methode sollte für ein Höchstmaß an Transparenz sowie intersubjektiver Nachvollziehbarkeit sorgen. Dass sich die hier erfolgte Diskursanalyse einer intra- wie intertextuellen Perspektive verschreibt, wurde mehrfach dargelegt (siehe Kap. 1 und 2.1). Auch wurden einige soziale Akteure dieses Diskurses benannt: neben den zuvor geklärten, beteiligten Medien konnten in den Sprachgebrauchsmustern selbst handelnde Personen, Unternehmen, Organisationen usw. ausgemacht werden. Damit wurden die drei wichtigsten Ebenen von DIMEAN angeschnitten und somit eine Triangulation erreicht. Die Interpretation der erhaltenen Cluster stellt wie erwähnt den qualitativen corpus-based-Part dar, und dieser ist selbstverständlich in einem gewissen Maß an die Subjektivität des Forschers gebunden. So gibt es garantiert linguistische Phänomene, die bei der Ergebnisinterpretation von mir nicht berücksichtigt, gar übersehen oder vom Leser eventuell vermisst wurden. Es sei daran erinnert, dass die Arbeit auch keinen Anspruch auf Vollständigkeit erhebt (lediglich den einer Diskursskizzierung), denn der Umfang der Daten muss auf die eine oder andere Art und Weise in eine Beschränkung und Auswahl münden. Andere Untersuchungen kommen möglicherweise zu anderen Ergebnissen respektive Interpretationen, die alle auf den Kern

einer konstruktivistischen Sicht der Diskurslinguistik zielen. Mit den Worten Wolfgang Teuberts:

“Linguistics is not a science like the natural sciences whose remit is the search for ‘truth’.[...] Interpretation, and not verification, is the proper response to the quest for meaning. There is no true meaning. The corpus linguist is not privileged as an ‘expert’ to pass judgment on what is permissible and what not. He or she is part of a discourse community, not outside of it. (Teubert 2005, S. 7f.)

Dies betrifft insbesondere wie gesagt die Ausdeutung der errechneten Befunde. Denn auf der anderen Seite: betrachtet man die typischen Ein- wie Mehrwort-Einheiten des corpus-driven-Teils der Untersuchung, so finden sich viele, welche sich möglicherweise auch oft mit Vorannahmen hinsichtlich mutmaßlich wichtiger oder dominanter Schlüsselwörter eines Diskurses decken können. Dies spricht aber gleichermaßen nicht gegen die Methodik. Denn es ist ein genuines Interesse der Wissenschaft im Allgemeinen auch gewisse Spekulationen empirisch zu bestätigen oder anders ausgedrückt: implizite Vermutungen explizit zu machen. Oder wie es Michael Stubbs ausdrückt:

“Yet even if quantification only confirms what we already know, this is no bad thing. Indeed, in developing a new method, it is perhaps better not to find anything too new, but to confirm findings from many years of traditional study, since this gives confidence that the method can be relied on.” (Stubbs 2005; S. 6)

Tatsächlich führen diese Befunde darüberhinaus etwa zu der Frage, wie kognitiv semantische Netze darüber gebildet werden, welche Begriffe, Schlag- bzw. Schlüsselwörter für ein bestimmtes Thema, Sachverhalt, Person oder sonstiges besonders kennzeichnend sind.<sup>70</sup> Des Weiteren lohnt sich auf einer allgemeineren Ebene die Erforschung von frequenzbasierten Mehrworteinheiten, denn als ‚prefabricated chunks‘ (vgl. Biber et al. 2004, S. 372) bzw. ‚Routineformeln‘ bilden sie psycholinguistisch gesehen wichtige Bausteine des Formulierens in der Sprache (vgl. Burger 2010, S. 57). Letztere ist also nicht eine gänzliche Neukombination von Einzelwörtern, sondern wesentlich stärker aus solchen sprachlichen Schablonen zusammengesetzt und für diese gilt: „Der Grad ihrer Beherrschung spiegelt die Fähigkeit wieder, kulturell angemessen zu kommunizieren“. (Steyer 2004, S. 91). Dabei kann es sich um Sprachgebrauchsmuster in der ganz alltäglichen Umgangssprache handeln, welche für Fremdsprachenlerner eine Herausforderung darstellen; es können spezielle Formulierungen gemeint sein, die in akademischen Texten häufig gebraucht werden; Muster, welche charakteristische Phrasen von massenmedialer Sprache verkörpern; oder eben auch typische Mehrwort-Einheiten ganz bestimmter, zuvor definierter Themenkreise. Ganz gleich, welcher Sprachausschnitt betrachtet wird: diskurslinguistisch interessant ist stets die Frage, welches Muster sich warum als ein vorgefertigter Baustein des Formulierens (sei es mündlich oder schriftlich) herausgebildet hat und wieso nicht etwa eine andere mögliche, paradigmatische Konstellation. Gebrauchshäufigkeiten bilden in diesem Zusammenhang ein wichtiges Indiz zum Aufspüren derartiger Formeln.

Hohe Frequenzen erklären selbstverständlich nicht alle Phänomene und die hier praktizierte Methode ersetzt nicht „klassische“ diskurslinguistische Ansätze. Dennoch sollten

---

<sup>70</sup> Auch wenn so ein „Versuch“ keine linguistische Aussagekraft hat, so habe ich probierhalber mündlich einmal mehrere Personen darum gebeten, sich geistig die Vorfälle um das Desaster in Japan vom März 2011 zu vergegenwärtigen – mit der Bitte, mir ihrer Meinung nach typische Schlagwörter zu nennen (ohne selbstverständlich vorher etwas von den gefundenen Resultaten zu erzählen). Das könnten Themen, Personen, Institutionen, Sachverhalte etc. sein; Hauptsache, es handelt sich dabei jeweils um einzelne Wörter. Beinahe alle der mir genannten Schlagwörter (so bspw. etwa *Radioaktivität*, *Katastrophe*, *Röttgen*, *Fukushima*; *Grünen*) waren zugleich Teil der hier berechneten Kenn-Wort-Liste. Einmal abgesehen von der „Validität“ dieser Ergebnisse, stellt sich die Frage, ob ganz bestimmte Begriffe zu einem Sachverhalt u.a. auch deshalb stärker memoriert werden, weil es sich um Wörter handelt, die zum einen sehr häufig genannt werden und zum anderen zugleich häufiger als bei anderen kognitiv repräsentierten Entitäten (andere Themen, Menschen, Gegenstände usw.).

quantitative Verfahren in diesem Zweig der Sprachwissenschaft nicht unterschätzt werden (vgl. Bubenhofer/Scharloth 2013, S. 160). Im Gegenteil: die fortschreitende Digitalisierung von Texten aller Art bildet große Datenmengen, deren induktive Auswertung zahlreiche neue Erkenntnisse über Sprache bereithalten. Sofern immer mehr ältere Texte digitalisiert werden, lohnt es sich, die diachrone Perspektive mit der hier praktizierten korpuslinguistischen Methode zu verfolgen: hat man ein genügend großes Korpus z.B. vom Kolonial-, Menschenrechts-, Migrations-, Globalisierungsdiskurs usw. auch aus vergangenen Jahrzehnten in digitaler Form, so ist es ein Leichtes, diese Daten mit beliebigen Referenzkorpora unterschiedlicher Zeiträume zu kontrastieren.<sup>71</sup> Der Sprachgebrauch ganz spezifischer Sprachausschnitte aus dem politischen, medialen, literarischen, laiengebundenen Diskurs (oder anderen Ebenen, je nach Forschungsinteresse und -perspektive) kann wie gesagt – dem jeweiligen, beliebigen Referenzkorpus gegenübergestellt – sehr schnell dessen Schlüsselwörter (oder Kenn-Wortverbindungen) zutage fördern. An dieser Stelle sollte auch bei künftiger Forschung weiter angesetzt werden: ein Spezialkorpus könnte man verschieden zusammengesetzten Referenzkorpora gegenüberstellen, um die in der Folge gewonnenen, wahrscheinlich voneinander abweichenden Sprachgebrauchsmuster zu kategorisieren und so zu einen neuen, relativeren Begriff von ‚Bedeutung‘ in einem jeweiligen Diskurs zu finden.

Der Aufsatz konnte hoffentlich zeigen, dass eine datengeleitete Analyse nicht bloß frequenzfixiert und inhaltsleer wirkt. Stattdessen können mit zuvor möglichst offenen Hypothesen vor dem Hintergrund empirisch gestützter Daten fundierte Aussagen getroffen werden; so führt man im Idealfall eine ‚Quantitativ informierte qualitative Diskursanalyse‘ (vgl. Bubenhofer 2013) durch. Mit ihrer Hilfe können bisherige diskurslinguistische, qualitative Erkenntnisse im einfachsten Fall bestätigt, im radikalsten Fall falsifiziert und ersetzt, zumindest aber ergänzt werden. Doch ganz gleich, was die Resultate erbringen: in jedem Fall wäre der Gewinn für diskursanalytische Fragestellungen immens – und ein solches Potential sollte auf keinen Fall verschenkt werden.

## Quellenverzeichnis

- Atkins, Sue/Clear, Jeremy/Ostler, Nicholas (1992): Corpus Design Criteria. In: *Journal of Literary and Linguistic Computing* 7 (1), S. 1–16. URL: <http://www.natcorp.ox.ac.uk/archive/vault/tgaw02.pdf> [Stand: 22.10.2012].
- Baker, Paul (2004): Querying Keywords: Questions of Difference, Frequency, and Sense in Keywords Analysis. In: *Journal of English Linguistics* 32 (4), S. 346–359.
- Belica, Cyril (2001ff.): Kookkurrenzdatenbank CCDB. Eine korpuslinguistische Denk- und Experimentierplattform für die Erforschung und theoretische Begründung von systemisch-strukturellen Eigenschaften von Kohäsionsrelationen zwischen den Konstituenten des Sprachgebrauchs. Institut für Deutsche Sprache, Mannheim. URL: <http://corpora.ids-mannheim.de/ccdb/> [Stand: 20.10.2012].
- Belica, Cyril/Steyer, Kathrin (2005): Korpusanalytische Zugänge zu sprachlichem Usus. In: *AUC (Acta Universitatis Carolinae), GERMANISTICA PRAGENSIA XX*. URL: [http://www1.ids-mannheim.de/fileadmin/ids/Downloads/II/Usuelle\\_Wortverbindungen/Belica\\_Steyer\\_2005.pdf](http://www1.ids-mannheim.de/fileadmin/ids/Downloads/II/Usuelle_Wortverbindungen/Belica_Steyer_2005.pdf) [Stand: 06.10.2012].
- Biber, Douglas/Conrad, Susan/Cortes, Viviana (2004): If you look at...: Lexical Bundles in University Teaching and Textbooks. In: *Applied Linguistics* 25 (3), S. 371–405.

---

<sup>71</sup> Vgl. z.B. Bubenhofer/Schröter (2012) und Bubenhofer/Scharloth (2013, ab S. 152) über den diachronen Wandel von Schlagwörtern im „Text+Berg“-Korpus.

- Bubenhofer, Noah (2008): Diskurse berechnen? Wege zu einer korpuslinguistischen Diskursanalyse. In: Warnke, Ingo H./Spitzmüller, Jürgen (Hgg.): Methoden der Diskurslinguistik. Sprachwissenschaftliche Zugänge zur transtextuellen Ebene. Berlin/New York: W. de Gruyter (=Linguistik – Impulse & Tendenzen. 31), S. 407–437.
- (2009): Sprachgebrauchsmuster. Korpuslinguistik als Methode der Diskurs- und Kulturanalyse. Berlin/New York: W. de Gruyter (=Sprache und Wissen. 4).
  - (2006-2011): Einführung in die Korpuslinguistik: Praktische Grundlagen und Werkzeuge. Elektronische Ressource. URL: <http://www.bubenhofer.com/korpuslinguistik/> [Stand: 03.04.2012].
  - (2013): Quantitativ informierte qualitative Diskursanalyse. Korpuslinguistische Zugänge zu Einzeltexten und Serien. In: Roth, Kersten Sven/Spiegel, Carmen (Hgg.): Angewandte Diskurslinguistik. Felder, Probleme, Perspektiven. Berlin: Akademie-Verlag (=Diskursmuster – Discourse Patterns. 2), S. 109–134.
- Bubenhofer, Noah/Scharloth, Joachim (2013): Korpuslinguistische Diskursanalyse: Der Nutzen empirisch-quantitativer Verfahren. In: Warnke, Ingo H./Meinhof, Ulrike/Reisigl, Martin (Hgg.): Diskurslinguistik im Spannungsfeld von Deskription und Kritik. Berlin: Akademie-Verlag (=Diskursmuster – Discourse Patterns. 1), S. 147–168.
- Burger, Harald (<sup>4</sup>2010): Phraseologie. Eine Einführung am Beispiel des Deutschen. 4., neu bearb. Aufl. Berlin: E. Schmidt (=Grundlagen der Germanistik. 36).
- Busse, Dietrich; Teubert, Wolfgang (1994): Ist Diskurs ein sprachwissenschaftliches Objekt? Zur Methodenfrage der historischen Semantik. In: Busse, Dietrich/Hermanns, Fritz/Teubert, Wolfgang (Hgg.): Begriffsgeschichte und Diskursgeschichte. Methodenfragen und Forschungsergebnisse der historischen Semantik. Opladen: Westdeutscher Verl., S. 10–28.
- Cheng, Winnie/Greaves, Chris/Warren, Martin (2006): From n-gram to skipgram to concgram. In: International Journal of Corpus Linguistics 11 (4), S. 411–433.
- Church, Kenneth W./Gale, William A. (1995): Poisson mixtures. In: Natural Language Engineering 1 (2), S. 163–90.
- Cortes, Viviana (2004): Lexical bundles in published student disciplinary writing: Examples from history and biology. In: English for Specific Purposes 23, S. 397–423.
- Coulmas, Florian/Stalpers, Judith (2011): Fukushima. Vom Erdbeben zur atomaren Katastrophe. München: C.H.Beck.
- Dunning, Ted (1993): Accurate Methods for the Statistics of Surprise and Coincidence. In: Computational Linguistics 19 (1), S. 61–74.
- Firth, John Rupert (1957a [1951]): Modes of Meaning. In: Papers in Linguistics 1934 – 1951. London: Oxford University Press, S. 190–215.
- (1957b): A synopsis of linguistic theory, 1930–1955. In: Studies in linguistic analysis. Oxford: The Philological Society, S. 1–32.
- Fix, Ulla (2002): Knappe Historie und kurze Laudatio – auf alle. In: Fix, Ulla/Adamzik, Kirsten/Antos, Gerd/Klemm, Michael (Hgg.): Brauchen wir einen neuen Textbegriff? Antworten auf eine Preisfrage. Frankfurt a. M. [u.a.]: Peter Lang (=Forum Angewandte Linguistik. 40), S. 7–16.
- Foucault, Michel (1973): Archäologie des Wissens. Titel der Originalausgabe: L'archéologie du savoir. Paris 1969. Aus dem Französischen von Ulrich Köppen. Frankfurt a. M.: Suhrkamp.
- (1982): Die Ordnung des Diskurses. Inauguralvorlesung am Collège de France – 2. Dezember 1970. Französischer Originaltitel: L'ordre du discours. Übersetzt von Walter Seitter. Ungek. Ausg. Frankfurt a. M./Berlin/Wien: Ullstein (=Anthropologie).
  - (1996 [1980]): Der Mensch ist ein Erfahrungstier. Gespräch mit Ducio Trombadori. Mit einem Vorwort von Wilhelm Schmid. Mit einer Bibliographie von Andrea Hemminger.

- Frankfurt a. M.: Suhrkamp [ital.: *Conversazione con Michel Foucault*. In: *Il Contributo 1* (1980), S. 23–84].
- Fujii, Yushiro/Satake, Kenji (2011): *Off Tohoku-Pacific Tsunami on March 11, 2011*. IISEE (International Institute of Seismology and Earthquake Engineering). URL: <http://iisee.kenken.go.jp/staff/fujii/OffTohokuPacific2011/tsunami.html> [Stand: 06.10.2012].
- Greaves, Chris (2009): *Concgram 1.0. A phraseological search engine*. [CD-Rom und Anleitung]. Amsterdam [u.a.]: John Benjamins (=Studies in Corpus Linguistics. Software 1). URL: <http://benjamins.com/jbp/catalogs/concgrammanual.pdf> [Stand: 25.04.2012].
- Halliday, Michael Alexander Kirkwood (2002 [1961]): *Categories of the theory of grammar*. In: Webster, Jonathan (Ed.): *On Grammar*. M. A. K. Halliday. London/New York: Continuum (=Collected Works of M. A. K. Halliday. 1), S. 37–94.
- Hano, Johannes (2011): *Das japanische Desaster. Fukushima und die Folgen*. Freiburg i. Br.: Herder.
- Hausmann, Franz J. (2004): *Was sind eigentlich Kollokationen?* In: Steyer, Kathrin (Hg.): *Wortverbindungen – mehr oder weniger fest*. Berlin/New York: W. de Gruyter (=Institut für deutsche Sprache, Jahrbuch 2003), S. 309–334.
- Hermanns, Fritz (1994): *Schlüssel-, Schlag- und Fahnenwörter. Zu Begrifflichkeit und Theorie der lexikalischen „politischen Semantik“*. Bericht Nr. 81. Arbeiten aus dem Sonderforschungsbereich 245 „Sprache und Situation“. Universität Heidelberg/Mannheim.
- IAEA (International Atomic Energy Agency) (2009): *INES. The International Nuclear And Radiological Event Scale User’s Manual*. 2008 Edition. Co-Sponsored by the International Atomic Energy Agency and OECD/Nuclear Energy Agency. Wien. URL: [http://www-pub.iaea.org/MTCD/publications/PDF/INES-2009\\_web.pdf](http://www-pub.iaea.org/MTCD/publications/PDF/INES-2009_web.pdf) [Stand: 07.10.2012].
- (2011): *Mission Report. The Great East Japan Earthquake Expert Mission*. IAEA International Fact Finding Expert Mission of the Fukushima Dai-ichi NPP Accident Following the Great East Japan Earthquake and Tsunami. Tokyo, Fukushima Dai-ichi NPP, Fukushima Dai-ni NPP and Tokai Dai-ni NPP, Japan. 24 May – 2 June 2011. URL: [http://www-pub.iaea.org/mtcd/meetings/pdfplus/2011/cn200/documentation/cn200\\_final-fukushima-mission\\_report.pdf](http://www-pub.iaea.org/mtcd/meetings/pdfplus/2011/cn200/documentation/cn200_final-fukushima-mission_report.pdf) [Stand: 07.10.2012].
- Japanese Government (2011): *Report of Japanese Government to the IAEA Ministerial Conference on Nuclear Safety. The Accident at TEPCO’s Fukushima Nuclear Power Stations*. URL: [http://www.kantei.go.jp/foreign/kan/topics/201106/iaea\\_houkokusho\\_e.html](http://www.kantei.go.jp/foreign/kan/topics/201106/iaea_houkokusho_e.html) [Stand: 07.10.2012].
- JMA (Japan Meteorological Agency) (2011): *The 2011 off the Pacific coast of Tohoku Earthquake ~first report~*. URL: [http://www.jma.go.jp/jma/en/News/2011\\_Earthquake\\_01.html](http://www.jma.go.jp/jma/en/News/2011_Earthquake_01.html) [Stand: 05.10.2012].
- Jung, Matthias (1994): *Zählen oder deuten? Das Methodenproblem der Diskursgeschichte am Beispiel der Atomenergiedebatte*. In: Busse, Dietrich/Hermanns, Fritz/Teubert, Wolfgang (Hgg.): *Begriffsgeschichte und Diskursgeschichte. Methodenfragen und Forschungsergebnisse der historischen Semantik*. Opladen: Westdeutscher Verl., S. 60–83.
- Jung, Matthias/Wengeler, Martin (1999): *Wörter – Argumente – Diskurse. Was die Öffentlichkeit bewegt und was die Linguistik dazu zu sagen hat*. In: Stickel, Gerhard (Hg.): *Sprache – Sprachwissenschaft – Öffentlichkeit*. Berlin/New York: W. de Gruyter (=Institut für deutsche Sprache. Jahrbuch 1998).
- Keller, Reiner/Hirsland, Andreas/Schneider, Werner/Viehöver, Willy (2011): *Zur Aktualität sozialwissenschaftlicher Diskursanalyse – Eine Einführung*. In: dies.

- (Hgg.): Handbuch Sozialwissenschaftliche Diskursanalyse. Band 1: Theorien und Methoden. 3., erw. Aufl. Wiesbaden: Verlag für Sozialwissenschaften (=Interdisziplinäre Diskursforschung), S. 7–34.
- Kilgarriff, Adam (2001): Comparing Corpora. In: *International Journal of Corpus Linguistics* 6 (1), S. 97–133. URL: <http://www.kilgarriff.co.uk/Publications/2001-K-CompCorpIJCL.pdf> [Stand: 10.10.2012].
- Köhler, Reinhard (2005): Korpuslinguistik – zu wissenschaftlichen Grundlagen und methodologischen Perspektiven. In: *LDV Forum* 20 (2), S. 1–16.
- Lehr, Andrea (1996): Kollokationen und maschinenlesbare Korpora. Ein operationales Analysemodell zum Aufbau lexikalischer Netze. Tübingen: Niemeyer (=Reihe Germanistische Linguistik. 168).
- Lemnitzer, Lothar/Zinsmeister, Heike (2006): *Korpuslinguistik. Eine Einführung*. Tübingen: Narr.
- LexisNexis® Wirtschaft (o.J.). Elektronische Literaturliteraturdatenbank. URL: <http://www.lexisnexis.com/de/business/search/loadForm.do?formID=DE01NBNewsSrch&random0.33312378479518034> [Stand: 02.10.2012].
- Liebert, Wolf-Andreas (1994): Das analytische Konzept „Schlüsselwort“ in der linguistischen Tradition. Bericht Nr. 83. Arbeiten aus dem Sonderforschungsbereich 245 „Sprache und Situation“. Universität Heidelberg/Manheim.
- Meier, Stefan (2011): Multimodalität im Diskurs: Konzept und Methode einer multimodalen Diskursanalyse. In: Keller, Reiner/Hirsland, Andreas/Schneider, Werner/Viehöver, Willy (Hgg.): *Handbuch Sozialwissenschaftliche Diskursanalyse. Band 1: Theorien und Methoden*. 3., erw. Aufl. Wiesbaden: Verlag für Sozialwissenschaften (=Interdisziplinäre Diskursforschung), S. 499–532.
- National Police Agency of Japan (2013): Damage Situation and Police Countermeasures associated with 2011Tohoku district – off the Pacific Ocean Earthquake. December 10, 2013. URL: [http://www.npa.go.jp/archive/keibi/biki/higaijokyo\\_e.pdf](http://www.npa.go.jp/archive/keibi/biki/higaijokyo_e.pdf) [Stand: 02.01.2014].
- Oakes, Michael P. (1998): *Statistics for Corpus Linguistics*. Edinburgh: Edinburgh University Press (=Edinburgh Textbooks in Empirical Linguistics).
- Partington, Alan (2006): Metaphors, motifs and similes across discourse types: Corpus-Assisted Discourse Studies (CADS) at work. In: Stefanowitsch, Anatol/Gries, Stefan Thomas (Hgg.): *Corpus-based Approaches to Metaphor and Metonymy*. Berlin/New York: M. de Gruyter (=Trends in Linguistics – Studies and Monographs. 171), S. 267–304.
- Partington, Alan/Morley, John (2004): From frequency to ideology: investigating word and cluster/bundle frequency in political debate. In: Barbara Lewandowska-Tomaszczyk (Hg.): *Practical applications in language and computers. PALC 2003*. Frankfurt a. M. [u.a.]: Lang (=Łódź studies in language. 9), S. 179–192.
- Perkuhn, Rainer (2007): „Corpus driven“: Systematische Auswertung automatisch ermittelter sprachlicher Muster. In: Kämpfer, Heidrun/Eichinger, Ludwig M. (Hgg.): *Sprach-Perspektiven. Germanistische Linguistik und das Institut für Deutsche Sprache*. Tübingen: Narr (=Studien zur Deutschen Sprache. Forschungen des Instituts für Deutsche Sprache. 40), S. 465–492.
- Perkuhn, Rainer/Belica Cyril (2006): Korpuslinguistik – Das unbekannte Wesen. Oder Mythen über Korpora und Korpuslinguistik. In: *Sprachreport* 22 (1), S. 2–8. URL: <http://pub.ids-mannheim.de/laufend/sprachreport/pdf/sr06-1a.pdf> [Stand: 26.06.2012].
- Rayson, Paul/Garside, Roger (2000): Comparing Corpora using Frequency Profiling. In: *Proceedings of the workshop on Comparing Corpora. The 38th Annual Meeting of the Association for Computational Linguistics. 1-8 October 2000: Hong Kong*, S. 1–6. URL: [http://www.comp.lancs.ac.uk/~paul/publications/rg\\_acl2000.pdf](http://www.comp.lancs.ac.uk/~paul/publications/rg_acl2000.pdf) [Stand: 25.07.2012].

- Scharloth, Joachim/Bubenhof, Noah (2012): Datengeleitete Korpuspragmatik: Korpusvergleich als Methode der Stilanalyse. In: Felder, Ekkehard/Müller, Marcus/Vogel, Friedemann (Hgg.): Korpuspragmatik. Thematische Korpora als Basis diskurslinguistischer Analysen von Texten und Gesprächen. Berlin/New York: W. de Gruyter (=Linguistik – Impulse & Tendenzen. 44), S. 195–230.
- Scharloth, Joachim/Eugster, David/Bubenhof, Noah (2013): Das Wuchern der Rhizome. Linguistische Diskursanalyse und Data-driven Turn. In: Busse, Dietrich/Teubert, Wolfgang (Hgg.): Linguistische Diskursanalyse. Neue Perspektiven. Wiesbaden: Springer VS (=Interdisziplinäre Diskursforschung), S. 345–380. URL: [http://www.scharloth.com/files/Rhizom\\_Zeit.pdf](http://www.scharloth.com/files/Rhizom_Zeit.pdf) [Stand: 07.01.2014].
- Scott, Mike (2006): In Search of a Bad Reference Corpus. Dokument zur Vortragsreihe: Word Frequency and Keyword Extraction. AHRC ICT Methods Network Expert Seminar on Linguistics. Lancaster University, UK, 08.09.2006. URL: [http://www.methodsnetwork.ac.uk/redist/pdf/es1\\_05scott.pdf](http://www.methodsnetwork.ac.uk/redist/pdf/es1_05scott.pdf) [Stand: 21.06.2012].
- (2008): WordSmith Tools version 6. Liverpool: Lexical Analysis Software.
- Scott, Mike/Tribble, Chris (2006): Textual Patterns. Key words and corpus analysis in language education. Amsterdam/Philadelphia: John Benjamins (=Studies in Corpus Linguistics. 22).
- Sinclair, John (1991): Corpus, Concordance and Collocation. Oxford/New York: Oxford University Press (=Describing English Language).
- (2004): Trust the Text. Language, corpus and discourse. Edited with Ronald Carter. London/New York: Routledge.
- Spitzmüller, Jürgen/Warnke, Ingo (2011): Diskurslinguistik. Eine Einführung in Theorien und Methoden der transtextuellen Sprachanalyse. Berlin/Boston: W. de Gruyter.
- Steyer, Kathrin (2004): Kookkurrenz. Korpusmethodik, linguistisches Modell, lexikografische Perspektiven. In: dies. (Hg.): Wortverbindungen – mehr oder weniger fest. Berlin/New York: W. de Gruyter (=Institut für Deutsche Sprache. Jahrbuch 2003), S. 87–116.
- Steyer, Kathrin/Lauer, Meike (2007): „Corpus-Driven“: Linguistische Interpretation von Kookkurrenzbeziehungen. In: Kämper, Heidrun/Eichinger, Ludwig M. (Hgg.): Sprach-Perspektiven. Germanistische Linguistik und das Institut für Deutsche Sprache. Tübingen: Narr (=Studien zur Deutschen Sprache. 40), S. 493–509.
- Storjohann, Petra (2007): Wie viel Diskurs braucht ein Wörterbuch? Das Stichwort Globalisierung im Elexiko-Wörterbuch. In: German Life and Letters 60/4, S. 569–592.
- Stubbs, Michael (1996): Text and Corpus Analysis. Computer-assisted Studies of Language and Culture. Oxford/Cambridge (Massachusetts): Blackwell (=Language in Society. 23).
- (2001): Words and Phrases. Corpus Studies of Lexical Semantics. Oxford/Malden (Massachusetts): Blackwell.
- (2005): Conrad in the computer: examples of quantitative stylistic methods. In: Language and Literature 14 (1), S. 5–24.
- Teubert, Wolfgang/Čermáková, Anna (2004): Direction in corpus linguistics. In: Halliday, M. A. K. et al. (Hgg.): Lexicology and Corpus Linguistics. An introduction. London/New York: Continuum (=Open Linguistic Series), S. 113–166.
- Teubert, Wolfgang (2005): My Version of Corpus Linguistics. In: International Journal of Corpus Linguistics 10 (1), S. 1–13.
- (2012): Von der Korpuslinguistik zur Analyse thematischer Diskurse. In: Felder, Ekkehard/Müller, Marcus/Vogel, Friedemann (Hgg.): Korpuspragmatik. Thematische Korpora als Basis diskurslinguistischer Analysen von Texten und Gesprächen. Berlin/New York: W. de Gruyter (=Linguistik – Impulse & Tendenzen. 44), S. 231–280.
- Tognini-Bonelli, Elena (2001): Corpus linguistics at work. Amsterdam/Philadelphia: John Benjamins (=Studies in Corpus Linguistics. 6).

- Töpfer, Klaus/Yogeshwar, Ranga (2011): *Unsere Zukunft. Ein Gespräch über die Welt nach Fukushima*. München: C.H.Beck.
- Tribble, Chris (2000): *Genres, Keywords, Teaching: towards a pedagogic account of the language of Project Proposals*. In: Bernard, Lou/McEnery, Tony (Hgg.): *Rethinking language pedagogy from a corpus perspective. Papers from the Third International Conference on Teaching and Language Corpora*. Frankfurt a. M./New York: P. Lang (=Łódź studies in language. 2), S. 75–90. URL: <http://www.ctribble.co.uk/text/Genre.htm> [Stand: 12.10.2012].
- UCREL (University Centre for Computer Corpus Research on Language) (o.J.): *Log-likelihood calculator. How to calculate log likelihood*. Lancaster University, UK. URL: <http://ucrel.lancs.ac.uk/llwizard.html> [Stand: 09.10.2012].
- USGS (United States Geological Survey) (2011): *Magnitude 9.0 - NEAR THE EAST COAST OF HONSHU, JAPAN*. URL: <http://earthquake.usgs.gov/earthquakes/recenteqsww/Quakes/usc0001xgp.php> [Stand: 05.09.2012].
- Warnke, Ingo H. (2002): *Adieu Text – bienvenue Diskurs? Über Sinn und Zweck einer poststrukturalistischen Entgrenzung des Textbegriffs*. In: Fix, Ulla/Adamzik, Kirsten/Antos, Gerd/Klemm, Michael (Hgg.): *Brauchen wir einen neuen Textbegriff? Antworten auf eine Preisfrage*. Frankfurt a. M. [u.a.]: Peter Lang (=Forum Angewandte Linguistik. 40), S. 125–142.
- (2007): *Diskurslinguistik nach Foucault – Dimensionen einer Sprachwissenschaft jenseits textueller Grenzen*. In: ders.: *Diskurslinguistik nach Foucault. Theorien und Gegenstände*. Berlin/New York: W. de Gruyter (=Linguistik – Impulse & Tendenzen. 25), S. 3–24.
- Warnke, Ingo H./Spitzmüller, Jürgen (2008): *Methoden und Methodologie der Diskurslinguistik – Grundlagen und Verfahren einer Sprachwissenschaft jenseits textueller Grenzen*. In: dies. (Hgg.): *Methoden der Diskurslinguistik. Sprachwissenschaftliche Zugänge zur transtextuellen Ebene*. Berlin/New York: W. de Gruyter (=Linguistik – Impulse & Tendenzen. 31), S. 3–54.
- Wengeler, Martin (2003): *Topos und Diskurs. Begründung einer argumentationsanalytischen Methode und ihre Anwendung auf den Migrationsdiskurs (1960–1985)*. Tübingen: Niemeyer (=Reihe Germanistische Linguistik. 244).