

Angelika Storrer (Mannheim)

Text und Interaktion im Internet

Abstract: Der Beitrag behandelt das Zusammenspiel von Text und Interaktion im Internet. Abschnitt 2 erläutert am Beispiel der Wikipedia, wie sich die textorientierte Arbeit an den Artikeln und das interaktionsorientierte Diskutieren funktional ergänzen. Abschnitt 3 untersucht Links als digitale Kohärenzbildungshilfen und zeigt an einem Fallbeispiel, wie diese in den schriftlichen Diskussionen dafür genutzt werden, relevante Informationen im „virtuellen“ Aufmerksamkeitsbereich präsent und für phorische und deiktische Bezugnahmen zugänglich zu machen. Abschnitt 4 diskutiert Ergebnisse aus zwei Vergleichsstudien zum Gebrauch der Konnektoren *weil* sowie *sprich* und *d.h.* in Wikipedia-Artikeln und Diskussionen, die auf der Basis von Wikipedia-Korpora in der DEREKO-Sammlung des IDS durchgeführt wurden.

1 Einführung

Nicht-linear organisierte Texte und Kommunikationsformen wie E-Mail, Chat und Diskussionsforen waren bereits in den 90er-Jahren des vergangenen Jahrhunderts verbreitet. Durch soziale Medien und deren Nutzung über mobile Endgeräte sind Hypertexte und internetbasierte Kommunikation inzwischen für viele Menschen alltäglich geworden. Obwohl die Besonderheiten internetbasierter Kommunikation und sozialer Medien mittlerweile ausführlich beschrieben sind,¹ werden sie in deutschen Grammatiken, mit Ausnahme von Hoffmann (2014), allenfalls am Rande behandelt. Selbst neuere Ansätze zur Textanalyse, z.B. Ágel (2017), konzentrieren sich auf gestaltstabile, linear organisierte Schrifttexte. Viele Kategorien zur Textanalyse – z.B. Oppositionspaare wie Anapher vs. Katapher, Thema vs. Rhema – setzen implizit voraus, dass Texte eine stabile Gestalt haben und linear organisiert sind, d.h. in einer eindeutig identifizierbaren Abfolge vom Textanfang zum Textende führen. In linearen Texten kann ein Autor antizipieren, welche Konzepte und Diskursreferenten im Vortext eingeführt wurden und welche Wissensbestände zum aktuellen Zeitpunkt im Aufmerksamkeitsbereich des Rezipienten „erreichbar“ sein sollten. Dass diese an linearen, gestaltstabilen Texten entwickelten Kategorien nicht ohne Adaption auf nicht-linear organisierte digitale Texte

¹ Vgl. u.a. Beißwenger (2007); Androutsopoulos et al. (2013); Barton/Lee (2013); Marx/Weidacher (2014); Storrer (2018).

übertragen werden können, wurde in der Literatur zu Hypertexten schon früh thematisiert (vgl. u.a. Kuhlen 1991; Storrer 2004).

Dieser Beitrag versteht sich als Plädoyer dafür, den Gegenstandsbereich von Textanalyseansätzen auf digitale Hypertexte und interaktionsorientierte Schreibformen zu erweitern. Im Jahr 2018 sind digitale Kohärenzhilfen, wie Links, Hashtags oder Adressierungen, keine Randerscheinungen mehr. Auch gibt es inzwischen öffentlich zugängliche Korpora, an denen man den Sprachgebrauch in der digitalen Kommunikation untersuchen kann (vgl. Beißwenger 2018). Am Beispiel der Wikipedia, einem der ältesten und erfolgreichsten Projekte des Social Web, rekapituliert dieser Beitrag in Abschnitt 2.1 zunächst zentrale Eigenschaften, durch die sich digitale Hypertexte von gestaltstabilen Texten unterscheiden. Abschnitt 2.2 erläutert das Zusammenspiel des textorientierten Verfassens von Wikipedia-Artikeln und der interaktionsorientierten Aushandlung von Fragen zur Artikelarbeit auf den Diskussionsseiten des Projekts. Die Opposition von Text- und Interaktionsorientierung bildet dann auch den Hintergrund der Vergleichsstudien zum Umgang mit Hyperlinks in Abschnitt 3 und zur Verwendung der Konnektoren *weil*, *d.h.* und *sprich* in Abschnitt 4.

Die Wikipedia steht unter einer freien Lizenz zur Verfügung und bietet damit der Forschung die Möglichkeit, die Daten in linguistisch aufbereitete Korpora zu überführen. Das Institut für Deutsche Sprache bietet über seine Korpusinfrastruktur Zugriff auf linguistisch aufbereitete Korpora zu Wikipedia-Artikeln und Wikipedia-Diskussionen zu verschiedenen Jahrgängen in mehreren Sprachen; der Bestand wird laufend erweitert und ausgebaut (Lüngen/Kupietz 2017). Zwei im Jahr 2015 erzeugte Korpora mit Artikeltexten (Wiki-A-2015) und mit Artikeldiskussionen (Wiki-D-2015) aus dieser Korpusfamilie dienen als die Datengrundlage für die Fallstudien dieses Beitrags.

2 Hypertext und interaktionsorientiertes Schreiben im Social Web: das Beispiel Wikipedia

In diesem Kapitel geht es um zwei Konzepte, die für das Zusammenspiel von Text und Interaktion im Social Web² zentral sind: (1) das Hypertext-Konzept und (2) das

² Die Bezeichnung *Social Web* verwenden wir in der von Ebersbach/Glaser/Heigl (2016) präzisierten Bedeutung, die auf die sozialen und kommunikativen Aspekte, den Aufbau und die Pflege

Konzept des interaktionsorientierten Schreibens. Beide sollen am Beispiel der Wikipedia erläutert werden, die sich sehr gut für diesen Zweck eignet:

Ad (1): Wikipedia wird von einem Hypertextsystem verwaltet, das die nicht-lineare Aufbereitung von Inhalten auf miteinander durch Hyperlinks verknüpften Seiten unterstützt. Wenn man Wikipedia-Artikel oder die darauf bezogenen Diskussionen mit den am Printtext ausgerichteten Ansätzen zur Text- und Diskursanalyse bearbeiten möchte, stellen sich neue Herausforderungen, die in Abschnitt 2.1 erläutert werden.

Ad (2): An der Wikipedia können sich prinzipiell alle Interessierten beteiligen, sofern sie sich an die in der Community ausgehandelten Leitlinien und Prinzipien halten. Da nicht nur die dabei entstehenden Produkte – die enzyklopädischen Artikel –, sondern auch die Diskussions- und Revisionsprozesse schriftlich verfügbar sind, bietet die Wikipedia interessante Möglichkeiten, nicht nur die Produkte, sondern auch die gemeinschaftlichen Schreib- und Aushandlungsprozesse zu erforschen.³ Am Vergleich der Artikel- und der Diskussionsbereiche kann man aber auch sehr gut eine neue Haltung zum Umgang mit Schriftsprache illustrieren, die in Abschnitt 2.2 als interaktionsorientiertes Schreiben vom textorientierten Verfassen der Artikeltexte abgegrenzt wird.

2.1 Text und Hypertext am Beispiel Wikipedia

Hypertexte sind digitale Texte, die von Softwaresystemen verwaltet werden. Auch die Wikipedia wird von einem Wiki-System namens MediaWiki verwaltet, das die gemeinschaftliche Bearbeitung von Hypertexten mit speziellen Funktionen unterstützt (vgl. Beißwenger/Storrer 2010; van Dijk 2010). Die von Wiki-Systemen verwalteten Hypertexte haben verschiedene Bereiche, sogenannte Namensräume, von denen wir zwei herausgreifen: (1) den Artikelnamensraum mit den enzyklopädischen Artikeln, die in der Wikipedia den Bezugspunkt der gemeinsamen Arbeit im Projekt bilden; (2) den Namensraum der Artikeldiskussionen, in denen über die Arbeit an den Artikeln und deren Optimierung diskutiert wird. Am Beispiel der Wikipedia werden im Folgenden vier charakteristische Merkmale von Hypertexten erläutert: Adaptivität, Dynamik, Vernetzung (Nicht-Linearität) und Interaktivität.

Das Merkmal „Adaptivität“ bedeutet, dass die Art und Weise, wie die Daten präsentiert werden, nicht fest vorgegeben ist. Aus derselben Datenstruktur kann

sozialer Beziehungen, den Informationsaustausch und die gemeinsame Wissensproduktion mithilfe von webbasierten Anwendungen fokussiert (ebd., S. 32).

3 Vgl. z.B. Liebert/Kohl (2004); Kallass (2013); Beißwenger (2016); Gredel (2017).

man unterschiedliche Sichten für verschiedene Interaktionsgeräte oder Nutzungssituationen generieren: In einer „App“-Sicht auf die Wikipedia sind die Navigations- und UI-Elemente anders angeordnet als in einer für die Desktop-Nutzung konzipierten Browser-Sicht. Druckerfreundliche Darstellungen kann man mit der Funktion „Als PDF herunterladen“ erzeugen.

Das Merkmal „Dynamik“ lässt sich gerade an Wikipedia-Artikeln sehr gut illustrieren. Diese sind nicht für eine Publikation als abgeschlossener, gestaltungstabiler Text konzipiert, sondern für eine fortlaufende Aktualisierung und Optimierung. Alle Artikelversionen werden in der öffentlich einsehbaren Versionsgeschichte gespeichert,⁴ dort kann man Versionen vergleichen und neue Versionen durch frühere ersetzen. Jede gespeicherte Version kann über einen permanenten Link angesteuert werden. Im Beispiel in Abschnitt 3 werden wir sehen, wie die Nutzer Links zu verschiedenen Artikelversionen in den Diskussionsseiten als Kontextualisierungshilfen einsetzen.

Variabilität im Layout und das Vorhandensein verschiedener Textversionen sind natürlich kein genuin neues Merkmal digitaler Hypertexte. Auch in der Printwelt gab es Texte in verschiedenen Ausgaben, sodass die Bezugnahme auf eine stabile Textform implizit schon immer eine Vorarbeit für die eigentliche Textanalyse bildete. Die Merkmale Adaptivität und Dynamik basieren allerdings auf Computertechnik und Algorithmen, die Textformen überhaupt erst aus einer digitalen Datenbasis generieren. Aus dieser Datenbasis muss man sich eine stabile Datengrundlage als eindeutig rekonstruierbaren Bezugspunkt für die Analyse generieren. Dabei kann man entweder bei den vom Wiki-System ausgegebenen Datenstrukturen ansetzen, wie dies z.B. beim Aufbau und der linguistischen Aufbereitung von Wikipedia-Korpora geschieht (Margaretha/Lüngen 2014). Man kann aber auch eine bestimmte Sicht abspeichern, z.B. als PDF-Dokument oder durch Bildschirmschnappschüsse, wie in den Abbildungen 1–2 dieses Artikels. Der Vorteil der datenstrukturbezogenen Analyse besteht darin, dass man die Daten computergestützt und damit auch quantitativ auswerten kann. Der Vorteil der sichtbezogenen Analyse, bei der eine bestimmte Präsentationsform der Daten durch Speicherung „stabilisiert“ wird, besteht darin, dass die Artikel und Diskussionen in einem „authentischen“ Layout erscheinen, wie es für die Nutzung der Wikipedia entwickelt wurde. Natürlich lassen sich auch Erkenntnisse aus datenstrukturbezogenen und aus sichtbezogenen Analysen kombinieren.⁵

⁴ In dem in Abbildung 1 gezeigten Artikelausschnitt kann man die Versionsgeschichte über den Link rechts oben aufrufen.

⁵ Vgl. z.B. die Untersuchung in Wessler et al. (2017).

Adaptivität und Dynamik sind keine Eigenschaften, die nur Hypertexte haben. Auch mit einem Textverarbeitungssystem können Texte in verschiedenen Versionen gespeichert und in unterschiedlichen Ansichten angezeigt werden. Die zentrale Idee von Hypertext, die mit dem Merkmal „Nicht-Linearität“ erfasst werden soll, besteht in der Verknüpfung von Hypertexteinheiten durch Hyperlinks, im Weiteren kurz „Links“ genannt. Links verknüpfen Texte und Textausschnitte, z.B. Wikipedia-Artikel mit weiterführenden Informationen und mit Quellen. Mit Links wird der Text um eine operative Schicht angereichert, die beim Ausdrucken oder in den Schnappschüssen wie in den Abbildungen 1 und 2 verloren geht. Typen von Links und ihr Stellenwert für die Kohärenzbildung wurden bereits früh ausführlich diskutiert (Kuhlen 1991; Hammwöhner 1997). Als neuer Aspekt soll in Abschnitt 3 an einem Beispiel herausgearbeitet werden, wie sich der Gebrauch von Links in den Artikeltexten von dem in den Diskussionen unterscheidet und welche Konsequenzen sich daraus für die Analyse ergeben.

Bereits vor der Verbreitung des Social Web bezeichnete man in der Informatik Software als interaktiv, wenn die Nutzer deren Verhalten über Interaktionsgeräte selbst steuern können, indem sie z.B. Hyperlinks aktivieren oder Suchfunktionen nutzen (vgl. Haack 1995). In der Medien- und Kulturwissenschaft wird der Ausdruck „interaktiv“ hingegen meist für die sprachliche Interaktion zwischen Nutzern über internetbasierte Kommunikationssoftware verwendet. Will man zwischen beiden Formen der Interaktivität begrifflich differenzieren, kann man die Mensch-Computer-Interaktion (MCI) von der Interaktion zwischen Menschen über internetbasierte Kommunikationsdienste (IBK) unterscheiden. Allerdings werden im Social Web zunehmend beide Aspekte relevant, wie man auch am Beispiel der Wikipedia schnell sehen kann: Die Bearbeitung eines Wikipedia-Artikels folgt zwar einem Handlungsmuster der computergestützten Textbearbeitung (Text aufrufen, bearbeiten, Bearbeitung speichern), das man der MCI zuordnen würde. Die Nutzer sind aber dazu angehalten, die Änderungen in einem dafür vorgesehenen Kommentarfeld kurz zu erläutern. Diese Kommentare werden in den Versionengeschichten gespeichert und richten sich primär an die anderen an der Artikelarbeit beteiligten Nutzer. Sie gehören also zum Bereich IBK und müssen, wie wir in Abschnitt 2.3 an einem Beispiel zeigen werden, bei einer vollständigen Analyse und Rekonstruktion von Interaktionsverläufen mit berücksichtigt werden.

2.2 Text- und interaktionsorientiertes Schreiben in der Wikipedia

In der Gegenüberstellung des Artikelnamensraums mit dem Namensraum der Artikeldiskussionen kann man sehr gut die Unterschiede zwischen zwei Haltun-

gen zum Gebrauch der Schriftsprache illustrieren: die Textorientierung und die Interaktionsorientierung.⁶

Im Kapitel „Grammatik von Text und Diskurs“ der Grammatik der deutschen Sprache (Zifonun/Hoffmann/Strecker 1997) werden Texte definiert als „Produkte sprachlichen Handelns, die in ihrer medialen Repräsentation und Gestaltkonstanz darauf angelegt sind, abgelöst von der Entstehungssituation an anderen Orten und zu anderen Zeiten (immer neu) rezipierbar zu sein“ (ebd., S. 249). An diesen funktionalen Textbegriff schließt auch die Opposition von textorientiert und interaktionsorientiert an. Zentral für diesen Textbegriff ist die Ablösbarkeit der Texte von einer aktuellen Sprachproduktionssituation. Texte sollten so formuliert sein, dass sie ohne direkte Interaktion zwischen Schreiber und Leser verständlich sind. Charakteristisch für das textorientierte Schreiben sind Prozesse des Planens und (oft mehrfachen) Redigierens. Auch die Arbeit an den Wikipedia-Artikeln ist in diesem Sinne textorientiert, auch wenn Wikipedia-Artikel als dynamische Hypertexte nicht auf einen Abschluss hin konzipiert sind, sondern fortlaufend optimiert und aktualisiert werden können. Die textorientierte Grundhaltung spiegelt sich auch in den Leitlinien der Community für das Verfassen guter Wikipedia-Artikel: Artikeltexte sollten so formuliert sein, dass die Inhalte aus sich selbst heraus, ohne Interaktion mit den Autoren, verständlich sind. Der Funktion der Textsorte „enzyklopädischer Artikel“ entsprechend sollen die Formulierungen sachorientiert und möglichst ohne Spezialwissen verstehbar sein. Die Nutzer sind gehalten, den Normen und Konventionen der schriftlichen Standardsprache zu folgen.

Das interaktionsorientierte Schreiben ist hingegen auf einen Interaktionsverlauf bezogen, an dem mehrere Akteure beteiligt sind und bei dem die Möglichkeit besteht, Verstehensprobleme direkt zu bearbeiten. Die Versprachlichungsstrategien zielen auf das Gelingen der laufenden Interaktion. Hierfür kann eine schnelle Reaktion wichtiger sein als sprachliche Elaboriertheit. Wie bei mündlichen Alltagsgesprächen spielen Aspekte der Beziehungsgestaltung und der gemeinsamen Bearbeitung einer Kommunikationsaufgabe eine zentrale Rolle. Diese Orientierung am Gelingen der sozialen Interaktion spiegelt sich in den Leitlinien und Empfehlungen für das Schreiben auf Diskussionsseiten. In der Wikiquette hat die

⁶ Die Opposition *textorientiert* vs. *interaktionsorientiert* wird ausführlich in Storrer (2018) erläutert. Ich präferiere die Bezeichnung *interaktionsorientiertes Schreiben* bzw. *interaktionsorientiertes Schreibprodukt* gegenüber dem kürzeren *interaktionale Sprache*, wie sie u.a. in Imo (2017) gebraucht wird, weil es sich – im Rahmen des in Hoffmann (2018) skizzierten ressourcenbasierten Ansatzes – nicht um eine neue Art von Sprache handelt, sondern um Unterschiede im Gebrauch sprachlicher und digitaler Ressourcen, die der Orientierung an einer laufenden sozialen Interaktion entgegenkommen.

Community Leitlinien für den sozialen Umgang zusammengestellt, die eine konstruktive und störungsfreie Zusammenarbeit fördern wollen. Ein „Kritik-Knigge“ stellt zusätzlich Empfehlungen für die Kritik an Artikeln zusammen, die verhindern sollen, dass sich gute Autoren durch harsch formulierte Kritik aus dem Projekt zurückziehen.

3 Hyperlinks als digitale Kohärenzbildungshilfen in Wikipedia-Artikeln und Wikipedia-Diskussionen

Produkte des textorientierten Schreibens sind im Internet sehr häufig Bezugspunkt für digitale Interaktionen: Online-Artikel, Beiträge in Blogs, Facebook, Instagram werden von anderen Nutzern kommentiert und bewertet. Aus den Kommentaren kann sich eine Diskussion mit dem Autor oder zwischen den kommentierenden Nutzern entwickeln. In der Wikipedia ist das Verhältnis von Text und Interaktion noch komplexer, weil es, wie bereits erläutert, verschiedene Namensräume gibt, denen im Rahmen des Projekts unterschiedliche Funktionen zukommen. Der Artikel fokussiert auf das Zusammenspiel zwischen den Artikeln der Wikipedia und den darauf bezogenen Artikeldiskussionen. Dieses Zusammenspiel wird am Beispiel des in Abbildung 1 dargestellten Ausschnitts aus dem Wikipedia-Artikel zum Lemma „Panama Papers“ und dem Ausschnitt aus einem darauf bezogenen Diskussionsverlauf in Abbildung 2 erläutert. Im Mittelpunkt stehen die Funktionen, die Hyperlinks als digitale Kohärenzbildungshilfen in beiden Bereichen übernehmen.

Als Produkte des textorientierten Schreibens sind Wikipedia-Artikel darauf ausgerichtet, Wissen zum Artikelgegenstand so zu präsentieren, dass ein breiter Kreis von Nutzern schnell und ohne direkte Rückfragen an die Autoren Informationen entnehmen kann. Die Links, die im Artikeltext enthalten sind, führen entweder zu Fußnoten, in denen die im Text gemachten Aussagen durch Quellen belegt werden.⁷ Oder es handelt sich um farblich als Links gekennzeichnete Textwörter, die mit anderen Wikipedia-Artikeln verknüpft sind. Letztere bezeichne ich als textintegrierte Links, weil sie neben ihrer Funktion als Link als Textsegmente regulär zum Aufbau der Satz- und Textbedeutung beitragen.

⁷ Beispiele hierfür sind die mit 5 und 6 gekennzeichneten Links am Ende des Absatzes des in Abbildung 1 gezeigten Ausschnitts.

Textintegrierte Links haben wie andere Textwörter auch eine Form- und Inhaltsseite. Als Links besitzen sie zudem auch prozedurale Eigenschaften, die festlegen, wie das Linkziel auf einem Interaktionsgerät angezeigt wird. Das Aktivieren des textintegrierten Links *Geldwäschedelikte* führt in dem in Abbildung 1 gezeigten Schnappschuss dazu, dass ein Fenster mit einer Kurzdefinition zu *Geldwäsche* eingeblendet wird. Diese Kurzdefinition wird aus dem verlinkten Wikipedia-Artikel zu *Geldwäsche* generiert, den der Nutzer durch entsprechende Aktivierung auch vollständig einsehen kann. Damit geben die textintegrierten Links in den Artikeltexten den Nutzern die Möglichkeit, weiterführendes Sprach- und Weltwissen zum Artikelgegenstand abzurufen. Es sind Vertiefungsangebote, die bei Bedarf genutzt werden können. Die Artikeltexte sind aber so konzipiert, dass sich auch ohne Rezeption der verlinkten Seiten eine kohärente Textbedeutung aufbauen lässt.

Artikel Diskussion Lesen Bearbeiten Quelltext bearbeiten Versionsgeschichte Wikipedia durchsuchen

Panama Papers

Als **Panama Papers** [pana.ma:ˈpepaˈʔ] (deutsch *Panama-Papiere*) werden vertrauliche Unterlagen des panamaischen Offshore-Dienstleisters Mossack Fonseca bezeichnet, die infolge eines 2,6 Terabyte großen Datenlecks am 3. April 2016 an die Öffentlichkeit gelangten. Nach Einschätzung der beteiligten Medien belegen die Unterlagen legale Strategien der Steuervermeidung, aber auch Steuer- und **Geldwäschedelikte**, den Bruch von UN-Sai

Enthüllung: **Geldwäsche** bezeichnet das Verfahren zur Einschleusung illegal erwirtschafteten Geldes bzw. von illegal erworbenen Vermögenswerten in den legalen Finanz- und Wirtschaftskreislauf. Da das zu „waschende“ Geld aus illegalen Tätigkeiten wie Korruption, Bestechung, Raub, Erpressung,

Mossack Fonseca.^[1] Die gen Politiker und andere flöcher, Briefkastenfirmen,

Mails, Briefe, gen und Bankauszüge als Ein anonymer sten Bastian Obermayer koordinierte das

International Consortium of Investigative Journalists (ICIJ) die einjährige Datenauswertung und weitere Recherchen. Am 3. April 2016 präsentierten 109 Zeitungen, Fernsehstationen und Online-Medien in 76 Ländern gleichzeitig die ersten Ergebnisse.^{[5][6]}

Panama Papers

Herkunftsländer politisch exponierter Personen, die bisher in den Panama Papers genannt wurden (Stand: 15. April 2016)

–	Beschreibung
Zeitraumen	1977 bis Frühjahr 2016
Veröffentlichung	3. April 2016
	2,6 Terabyte, 11,5 Millionen Dokumente
Umfang der Originaldaten	214.488 Briefkastenfirmen

Abb. 1: Ausschnitt aus dem Artikel „Panama Papers“⁸

Auch die Diskussionsseiten enthalten Links zu Quellen, zu Wikipedia-Artikeln und zu Seiten des Projekts, in denen die Leitlinien und Konventionen der gemein-

⁸ Permanente URL: https://de.wikipedia.org/w/index.php?title=Panama_Papers&oldid=179829168 (Stand: 4.10.2018).

schaftlichen Arbeit erläutert sind.⁹ Häufig dienen Links aber auch dazu, die für die aktuelle Diskussion relevanten Internetseiten mit dem eigenen Diskussionsbeitrag zu verknüpfen. Die verlinkten Seiten werden damit Teil des schriftlich konstituierten gemeinsamen Aufmerksamkeitsbereichs. Diese für das interaktionsorientierte Schreiben charakteristische Funktion von Links soll im Folgenden an dem in Abbildung 2 gezeigten Ausschnitt illustriert werden.

Wikipedia-Diskussionen beginnen mit einem Initialbeitrag, der das Thema der Diskussion in einer Überschrift umreißt und das Anliegen im Folgetext noch weiter erläutert. Auf diesen Initialbeitrag können andere Nutzer Bezug nehmen. Normalerweise werden Beiträge signiert, d.h. mit dem Nutzernamen und dem Zeitpunkt des Postings markiert. Der in Abbildung 2 gezeigte Ausschnitt umfasst vier Beiträge von unterschiedlichen Nutzern, die am 4. April 2016 im Zeitraum von 15:57 Uhr bis 16:32 Uhr gepostet wurden.



Abb. 2: Ausschnitt aus einer Diskussion zum Artikel „Panama Papers“¹⁰

⁹ Ein Beispiel dieser Art ist der Link zur Seite „Keine Theoriefindung“ in Beleg 6 in Abschnitt 4.1.

¹⁰ Permanente URL: [https://de.wikipedia.org/wiki/Diskussion:Panama_Papers/Archiv/1#Infobox_\(erl.\)](https://de.wikipedia.org/wiki/Diskussion:Panama_Papers/Archiv/1#Infobox_(erl.)) (Stand: 5.10.2018)

Der in der Überschrift des Initialbeitrags benannte Gegenstand der Diskussion ist eine Infobox im Wörterbuchartikel, aus der sich in tabellarischer Form wichtige Informationen entnehmen lassen. Die aktuelle Version dieser Infobox ist im Artikelausschnitt in Abbildung 1 an der rechten Seite zu sehen. Zum Zeitpunkt der Diskussion standen zwei Versionen des Artikels zur Debatte, eine mit Infobox und eine ohne Infobox. Der Nutzer des Initialbeitrags gestaltet die Präpositionen *mit* und *ohne* als Links, die zu den jeweiligen Seitenvarianten führen, sodass alle, die zur strittigen Infobox Stellung nehmen möchten, schnell beide Varianten einsehen können.

Diskussionen sind zeitlich situiert. Sie beziehen sich auf die Versionen, die im Zeitraum des Diskussionsverlaufs aktuell waren. Die durch den Link *mit* verknüpfte Artikelversion ist auf den 4. April 2016 um 10:47 Uhr datiert; die durch den Link *ohne* verknüpfte Version auf den 4. April 2016 um 15:54 Uhr. Die Debatte um die beiden Versionen ist auf die am selben Tag laufende Arbeit am Artikel bezogen; der in Abbildung 2 gezeigte Ausschnitt ist Teil einer Interaktion, die in verschiedenen Namensräumen der Wikipedia stattfand.

Der Bezug zum Artikeltext wird bereits im ersten Satz des Initialpostings deutlich: *Was spricht gegen diese harmlose Infobox im Artikel, außer ‚das hatten wir noch nie‘*. Die Nominalphrase *diese harmlose Infobox im Artikel* referiert auf das strittige Objekt, das bemerkenswerterweise in der Artikelversion, die zum Zeitpunkt des Postings aktuell war, nicht enthalten war. Dieser Einstieg macht deutlich, dass es bereits eine Vorkommunikation zum Thema gegeben hat. Das zeigt sich nicht nur am Adjektiv *harmlos*, mit dem sich der Nutzer in der Streitfrage als Befürworter der Infobox positioniert. Es wird auch ersichtlich an dem als Zitat eingefügten *das hatten wir noch nie* als Argument gegen die Infobox. Um die Bezüge nachvollziehbar zu machen, legt der Nutzer Links zu drei Seiten aus der Versionengeschichte des Artikels, die für die Vorgeschichte des Initialbeitrags zentral waren. Wenn man die mit den Zahlen (2), (3) und (4) gekennzeichneten Links aktiviert, kann man folgenden Ablauf rekonstruieren:

- Aus der mit (2) verlinkten Seite kann man entnehmen, dass Nutzer Aschmidt am 4. April 2016 um 14:45 die fragliche Infobox entfernt. Im Kommentarfeld zu dieser Änderung hinterlässt er den Eintrag: *-infobox; kein konsens in solchen artikeln; nur auf enwiki geläufig*. Als Insider kann man daraus entnehmen, dass er die Infobox entfernt hat, weil es über die Verwendung von Infoboxen in der deutschen Wikipedia noch keinen Konsens gibt und diese bislang nur in der englischen Wikipedia¹¹ „geläufig“ sind.

¹¹ Das Kürzel *enwiki* steht in der Werkstattsprache für die englische Sprachversion der Wikipedia.

- Aus der mit (3) verlinkten Seite geht hervor, dass der Nutzer Benqo, der auch den Initialbeitrag des Diskussionsverlaufs in Abbildung 2 verfasst hat, diesen Eingriff drei Minuten später wieder rückgängig macht. Im Kommentarfeld äußert er sich zum Entfernen der Infobox und der dafür angeführten Begründung folgendermaßen: *nicht geläufig reicht mir nicht als grund. ich bitte doch auf der disk darzulegen, wo das problem dieser infobox liegt, die kurz und bündig infos zusammenfasst*. Mit seiner Bitte, das Problem auf der Diskussionsseite des Artikels (*disk* abgekürzt) auszuhandeln, soll ein Bearbeitungskonflikt (ein sogenannter *edit war*) vermieden werden, in dem zwei Nutzer ihre Änderungen wechselseitig wieder zurücksetzen.
- Die mit (4) verlinkte Seite dokumentiert, dass der Nutzer Aschmidt dieser Bitte nicht nachkommt, sondern drei Minuten später wieder die von ihm präferierte Version ohne Infobox aktuell setzt. Aus seinem beigefügten Kommentar: *wie bereits gesagt: das hatten wir noch nie; wenn du das auf dewiki neu einführen möchtest, begründe du es bitte zuvor; danke* stammt dann auch das Zitat des Initialbeitrags.

Unmittelbar nach dem erneuten Entfernen der Infobox postet Nutzer Benqo den Initialbeitrag des in Abbildung 2 gezeigten Diskussionsverlaufs. Wenn man den gesamten Hypertextraum der Wikipedia betrachtet, ist dieser Beitrag aber nur Teil einer Interaktionssequenz, die bereits mit dem erstmaligen Entfernen der Infobox und dem beigefügten Kommentar beginnt. Durch das Einfügen der Links im Initialbeitrag macht der Nutzer Benqo die gesamte Sequenz zugänglich und damit auch für Außenstehende rekonstruierbar.

Denn der Beitrag auf der Diskussionsseite ist nicht nur an seinen Kontrahenten adressiert – diesem ist die Vorkommunikation ja präsent. Vielmehr richtet sich der Beitrag auch an andere Nutzer, die den Artikel bearbeiten oder an der Debatte aus anderen Gründen interessiert sind. Die Verlinkung vereinfacht es den Außenstehenden, sich vom Gegenstand und vom Verlauf ein Bild zu machen und selbst dazu Stellung zu beziehen. Wie man sieht, hat er damit Erfolg: Im weiteren Verlauf plädieren mehrere Nutzer für die Beibehaltung der Box, sie wird – wie man an der Versionengeschichte rekonstruieren kann – wieder eingefügt und bleibt fortan im Artikel verankert. Die Initiative des Nutzers, den Streitpunkt auf die Diskussionsseite zu verlagern, hat einen längeren Bearbeitungskonflikt verhindert.

Meine Analyse des Ausschnitts in Abbildung 2 war fokussiert auf den Initialbeitrag und dessen Verlinkung mit den Artikeln und der Versionengeschichte. Interessante Aspekte, wie die Verwendung der alten Rechtschreibung durch den Nutzer Aschmidt, der auch zum Thema Infoboxen eine konservative Haltung vertritt, konnten nicht vertieft werden. Das Beispiel sollte vor allem deutlich machen, dass die durch Links verknüpften Informationen für das Verständnis von Diskus-

sionsbeiträgen unabdingbar sind. Diskussionen sind zeitlich situiert und können nur mit Bezug auf die Artikelversionen interpretiert werden, die zum Zeitpunkt der Diskussion aktuell waren. Weiterhin sollte deutlich werden, dass auch beim Umgang mit Links systematische Unterschiede bestehen zwischen der textorientierten Arbeit an den Artikeln einerseits und dem interaktionsorientierten Diskutieren andererseits. Der Gebrauch von Links auf Artikelseiten ist in den Leitlinien zum Verfassen von Artikeln eng reglementiert. Auf den Diskussionsseiten haben die Nutzer mehr Freiräume, dort bilden sich Praktiken heraus, die der auf effiziente, gemeinsame Problemlösung ausgerichteten Schreibhaltung förderlich sind. Links sind ein praktisches Mittel, um interaktionsrelevante Informationen in den gemeinsamen Aufmerksamkeitsbereich zu rücken und – wie man am Beispiel sieht – auch für phorische und deiktische Bezugnahmen zugänglich zu machen. Aus Produktionssicht können Links Schreibarbeit sparen, weil die verlinkten Informationen nicht neu ausformuliert werden müssen. Aus Rezeptionsicht erleichtern sie den Einstieg in die Diskussion für bislang Unbeteiligte, die sich aus den verlinkten Seiten schnell ein Bild von der Vorgeschichte und den Gegenständen der Diskussion machen können.

4 Konnektoren in Wikipedia-Artikeln und Wikipedia-Diskussionen: Zwei korpusbasierte Fallstudien

In Abschnitt 3 ging es um Links als neue digitale Kohärenzbildungshilfen und um Unterschiede, die im Gebrauch von Links beim textorientierten Verfassen von Artikeln einerseits und beim interaktionsorientierten Diskutieren andererseits deutlich werden. Auch in diesem Abschnitt vergleichen wir Verwendungsweisen in Artikeln und in Diskussionen, es geht aber um „traditionelle“ Kohäsionsmittel, nämlich die Konnektoren *weil* sowie *d.h.* und *sprich*, die sich aus in den folgenden Teilkapiteln erläuterten Gründen für einen solchen Vergleich anbieten. Als Datenbasis dienen zwei Korpora aus der Wikipedia-Korpusfamilie des IDS, die aus einer Wikipedia-Kopie vom 17.11.2015 aufgebaut wurden: das Korpus der Artikelseiten mit ca. 790 Millionen Textwörtern (als Wiki-A-2015 abgekürzt) und das Korpus der Artikeldiskussionen mit ca. 270 Millionen Textwörtern (als Wiki-D-2015 abgekürzt). Es handelt sich um erste Fallstudien, die demonstrieren sollen, dass es lohnenswert ist, den Gebrauch von Konnektoren beim text- und beim interaktionsorientierten Schreiben auf der Basis dieser beiden großen Korpora vergleichend zu untersuchen.

4.1 Der Konnektor *weil* in Wikipedia-Artikeln und Wikipedia-Diskussionen

Der Konnektor *weil* gehört nicht nur zu den sehr gut untersuchten, sondern auch zu den sehr häufig verwendeten Konnektoren des Deutschen. Zur Suchform *weil* findet man im Artikelkorpus Wiki-A-2015 mehr als 100.000 Treffer, im Diskussionskorpus Wiki-D-2015 annähernd 300.000 (vgl. Tab. 1).

Tab. 1: Ergebnisse der vergleichenden Untersuchung zum Konnektor *weil* in Korpora zu Wikipedia-Artikeln (Wiki-A-2015) und Artikeldiskussionen (Wiki-D-2015)

Suchform	Korpus	Frequenz absolut	Sample	Vletzt		V2		Rest		Pseudo	
				abs.	%	abs.	%	abs.	%	abs.	%
<i>weil</i>	Wiki-A	102.145	1.000	944	94,4	0	0	5	0,5	51	5,1
	Wiki-D	297.300	1.000	918	91,8	15	1,5	45	4,5	22	2,2
<i>weil,</i>	Wiki-A	91	91	60	65,9	0	0	12	13,2	19	20,9
	Wiki-D	570	570	368	64,6	77	13,5	79	13,8	46	8,1
<i>weil:</i>	Wiki-A	14	14	10	71,4	0	0	3	21,4	1	7,2
	Wiki-D	736	736	224	30,4	315	42,8	175	23,7	22	3

Dass *weil* in den expositorischen Artikeltexten seltener vorkommt als in den Diskussionen, in denen auch Argumente ausgetauscht und damit Begründungszusammenhänge hergestellt werden müssen, war erwartbar. Was uns am Vergleich der beiden Teilkorpora vor allem interessierte, war die Frage nach dem Anteil der Verwendungen mit Verbzweitstellung (im Folgenden als V2-*weil* abgekürzt) wie in den folgenden Beispielen:

- (1) Ich würde einen LA unterstützen, **weil**, so ist das wirklich nix. (<https://de.wikipedia.org/wiki/Diskussion:Kissamos/Archiv>; Beleg aus Wiki-D-2015)
- (2) Die Art der Inhalte ist ebenfalls nicht egal, **weil**: Wikipedia ist kein Nachrichtenportal oder Veranstaltungskalender und dient nicht der aktuellen Berichterstattung. (https://de.wikipedia.org/wiki/Diskussion:Deutschland_schafft_sich_ab/Archiv/5; Beleg aus Wiki-D-2015)

Die semantischen und diskursstrukturellen Unterschiede zwischen V2-*weil* und dem Subjunktor *weil*, der einen Nebensatz mit dem finiten Verb am Ende einleitet (im Folgenden als Vletzt-*weil* abgekürzt), sind sehr gut erforscht und beschrieben (vgl. u.a. Dittmar/Bressem 2005; Breindl/Volodina/Waßner 2014, S. 840 ff; Duden-

Grammatik 2016, S. 1222f.). Dabei zeigt sich, dass V2-*weil* in der mündlichen Interaktion spezielle Funktionen übernimmt und gar nicht in allen Kontexten durch Vletzt-*weil* ausgetauscht werden kann. Dessen ungeachtet bewertet die nicht-wissenschaftliche Sprachkritik die Verwendung von V2-*weil* auch in mündlichen Gesprächen als eine falsche Variante des Subjunktors.¹² Eichinger (2017) bezeichnet V2-*weil* zu Recht als ein Paradebeispiel dafür, wie in der nicht-wissenschaftlichen Sprachkritik „Veränderungen kritisiert oder zumindest kritisch kommentiert werden, die beim genauen Hinsehen darauf zurückzuführen sind, dass schriftsprachliche Explizitheit und ihre Regularitäten zum Maßstab der Beurteilung gemacht werden, obwohl es um Erscheinungen geht, die für natürliche Mündlichkeit kennzeichnend sind“ (ebd., S. 284). Er zeigt an Beispielen, wie die Verwendung von V2-*weil* zu den beim Sprechen generell bevorzugten Formulierungsstrategien passt und warum V2-*weil* in der mündlichen Alltagssprache nicht einfach als falsche Variante des Subjunktors betrachtet werden sollte (ebd., S. 286).

Unsere Untersuchung beschäftigt sich mit der Frage, ob sich V2-*weil* auch in der Schriftsprache ausbreitet und ob sich Unterschiede zwischen den redigierten Artikeltexten und den oft spontanschriftlich verfassten Diskussionen nachweisen lassen. Hierfür haben wir aus beiden Teilkorpora – dem Teilkorpus mit redigierten Artikeln Wiki-A-2015 und dem Teilkorpus mit Diskussionen zu Artikeln Wiki-D-2015 – zwei Zufallsstichproben mit jeweils 1.000 Treffern gezogen. Alle Treffer wurden intellektuell analysiert und einer der folgenden vier Klassen zugeordnet: Vletzt (Subjunktors mit Verbletzstellung), V2 (Konnektor mit Verbzweitstellung), Pseudo (falsch positive Treffer) und Rest (falsch positive bzw. nicht zuordenbare Treffer).

Zu den falsch positiven „Pseudotreffern“ gehören Imperativformen des Verbs *weilen* bzw. davon abgeleitete Eigennamen wie *Weil am Rhein* sowie Belege wie (3), in denen *weil* als Ausdrucksform erwähnt, aber nicht als Konnektor verwendet wird.

- (3) [...] und kann jemand einige der suboptimalen Kurzsätze sinnstiftend und mit einer der schönen Konjunktionen des Deutschen (**weil**, nachdem, indem usw.) besser in den Gesamtkontext einbinden? (https://de.wikipedia.org/wiki/Diskussion: Bahnhof_Berlin_Gesundbrunnen; Beleg aus Wiki-A-2015)

Zur Restklasse gehören Beispiele wie (4) und (5), in denen *weil* zwar als Konnektor verwendet wird, das zweite Konnekt aber kein finites Verb enthält.

¹² Vgl. das Zitat von Bastian Sick in Eichinger (2017, S. 285).

- (4) Beides wird jedenfalls durch Gordon dargestellt, ganz im Gegensatz zum Link zur FSF, der lediglich die Lizenzinkompatibilität beleuchtet, welche hier gar keine Rolle spielt (**weil**: Sammelwerk!). (<https://de.wikipedia.org/wiki/Diskussion:Cdrtools/Archiv/2010>; Beleg aus Wiki-D-2015)
- (5) Zum Diskussionsthema Das Große Freie ist zu sagen, dass der immer noch vorhandene Hinweis auf fränkische Siedler historisch unseriös ist, **weil**, nicht belegbar und erklärbar. (https://de.wikipedia.org/wiki/Diskussion:Gro%C3%9Fes_Freies; Beleg aus Wiki-D-2015)

Die mit *weil* angeschlossenen verblosen Konnekte lassen sich zwar problemlos durch Sätze mit Kopulaverb paraphrasieren. Es ist aber nicht entscheidbar, ob solche Paraphrasen als V2-Variante (*weil: es ist ein Sammelwerk!*) oder als Vletzt-Variante (*weil es ein Sammelwerk ist*) zu formulieren sind. Zur Gruppe der verblosen Konnekte in der Restklasse gehören auch Beispiele wie (6), in denen das Konnekt als Link fungiert, der mit einer Seite verknüpft ist, auf der das zur Begründung angeführte Argument im Detail nachlesbar ist.

- (6) Nein, wir müssen bis Montag warten und das nehmen, was in einem Zeitungsartikel steht, **weil**: [Wikipedia:Keine Theoriefindung#Theoriedarstellung](https://de.wikipedia.org/wiki/Diskussion:Alternative_f%C3%BCr_Deutschland/Archiv/004) (https://de.wikipedia.org/wiki/Diskussion:Alternative_f%C3%BCr_Deutschland/Archiv/004; Beleg aus Wiki-D-2015)

Eine andere Fallgruppe der Restklasse bilden Beispiele wie (7), in denen das zweite Konnekt als Zitat aus einer anderen Sprache stammt und deshalb nicht den auf das Deutsche bezogenen topologischen Klassen zugeordnet werden kann.

- (7) Wenn dann mal in einem Review „no serious complications“ beobachtet werden, muss das für die Praxis nicht viel heißen, **weil**: „It is difficult to estimate the incidence of SMT complications, as they are probably under-reported in the literature“. (<https://de.wikipedia.org/wiki/Diskussion:Chiropraktik/Archiv/1>; Beleg aus Wiki-D-2015)

Die Ergebnisse unserer Auswertung in Tabelle 1 zeigen, dass V2-*weil* in den 1.000 Belegen der Artikelstichprobe kein einziges Mal vorkommt. Dieses Ergebnis mag die Sprachkritiker beruhigen, ist aber weniger überraschend als die niedrige Zahl von V2-*weil*-Belegen in den Diskussionen. Nur 15 der 1.000 analysierten Belege gehören zur Klasse V2-*weil*. Damit liegt der Anteil mit 1,5% deutlich niedriger als der Anteil von 30% an V2-*weil*-Sätzen, der in einer Untersuchung von Dittmar/Bressemer (2005) an einem Korpus mit mündlichen Interviews aus den 90er-Jahren nachgewiesen wurde (vgl. ebd., S. 102 ff., 118). Es sei dahingestellt, ob man aus

unserem Befund ableiten kann, dass die V2-weil-Variante tatsächlich vor allem den Versprachlichungsstrategien beim mündlichen Formulieren entgegenkommt.¹³ Vielleicht überwiegt auch einfach in der schriftlichen Kommunikation, selbst in der interaktionsorientierten, die Scheu vor V2-weil-Sätzen, die in der populären Sprachkritik selbst im mündlichen Sprachgebrauch verpönt sind.

Bei der Durchsicht der 15 V2-weil-Belege fiel uns auf, dass der Konnektor *weil* häufig durch ein Komma, wie in Beispiel (1), oder durch einen Doppelpunkt, wie in Beispiel (2), von den Verbzweit-Sätzen abgegrenzt wurde. Diese Interpunktionszeichen stehen an der Stelle, an der in der gesprochenen Sprache auch ein separierender Grenzton oder eine Pause gemacht werden könnte. In einer Nachuntersuchung haben wir in Wiki-A-2015 und Wiki-D-2015 nach der Suchform *weil* gefolgt von einem Komma bzw. von einem Doppelpunkt gesucht. Wie die Ergebnisse in Tabelle 1 zeigen, findet man auch bei diesen beiden Anfragen sehr viel mehr Treffer in den Diskussionen als in den Artikeltexten. Da die Trefferlisten sehr viel kleiner waren, konnten wir sie vollständig intellektuell analysieren und alle Treffer den oben genannten vier Klassen – Vletzt, V2, Pseudo und Rest – zuordnen. Belege wie (8), bei denen das Komma hinter *weil* einen darauf folgenden Einschub von einem mit *weil* eingeleiteten Nebensatz abtrennt, wurden der Vletzt-Klasse zugeordnet.¹⁴

- (8) Das erste Jahr der Gruppe gilt als bedeutend, **weil**, Zitat nach Ursula Prinz, „sie die wichtigsten jungen Künstler der Zeit, die später zu unterschiedlichem Ruhm gelangen sollten, zusammenfaßte und bekannt machte.“ (https://de.wikipedia.org/wiki/Großgörschen_35; Beleg aus Wiki-A-2015)

Tabelle 1 zeigt, dass sich auch bei dieser Untersuchung kein einziger Beleg für V2-weil in den Artikeltexten fand. In den Trefferlisten der Diskussionsseiten ist der Anteil von V2-weil-Belegen mit 13,5% für die Separierung durch Komma und 42,8% für die Separierung durch Doppelpunkt jedoch deutlich höher als in dem 1.000-Belege-Sample der ersten Untersuchung. Dass der Doppelpunkt als Trennzeichen häufiger genutzt wird als das Komma, könnte daran liegen, dass das Amtliche Regelwerk die Verwendung des Kommas an dieser Position nicht lizenziert, während das Abtrennen durch Doppelpunkt mit einer weiten Auslegung

¹³ Diskussionsbeiträge sind Produktdaten, aus denen man nicht ersehen kann, wie schnell formuliert wurde und wie viele und welche Überarbeitungen vor dem Veröffentlichenden vorgenommen wurden.

¹⁴ Ich erwähne dies wegen einer Nachfrage am Rande der IDS-Jahrestagung hier explizit, obwohl diese Zuordnung eigentlich die einzig naheliegende ist.

des Paragraphen § 81 (2) im Amtlichen Regelwerks (Regelwerk 2018, S. 90) kompatibel ist, nach dem ein Doppelpunkt zur Ankündigung einer danach folgenden Erklärung genutzt werden kann. Ob die Schreiber wirklich so gut mit dem Regelwerk vertraut sind, sei dahingestellt. Vielleicht wird der Doppelpunkt auch einfach wegen seiner auffälligeren Erscheinungsform als besser erkennbares Grenzzeichen präferiert.

Zusammenfassend lässt sich festhalten, dass der Konnektor *weil* mit Verbzweitstellung, also V2-*weil*, in den von uns untersuchten Belegen aus dem Wikipedia-Artikelkorpus Wiki-A-2015 nicht ein einziges Mal vorkommt. Auch in den Wikipedia-Diskussionen liegt der Anteil deutlich niedriger als in Untersuchungen, die an Gesprächskorpora durchgeführt wurden. Wenn V2-*weil* in den Diskussionen verwendet wird, ist der Konnektor *weil* häufig durch ein Interpunktionszeichen vom V2-Satz getrennt; als Separator wird der mit dem Regelwerk kompatible Doppelpunkt häufiger verwendet als das an dieser Position nicht regelkonforme Komma.

4.2 Die Konnektoren *sprich* und *d.h.* in Wikipedia-Artikeln und Wikipedia-Diskussionen

Die zweite Fallstudie vergleicht die Vorkommensfrequenz der Konnektoren *das heißt* in der Schreibform *d.h.* und *sprich* in Beispielen wie (9), (10) und (11) in Wikipedia-Artikeln und Wikipedia-Diskussionen.

- (9) Die solchen Werken zuvor erteilte Beteiligungsziffer, die sogenannte Absatzquote, konnte auf andere Werke übertragen (**sprich**: verkauft) werden. (http://de.wikipedia.org/wiki/Kaliwerk_Gewerkschaft_Wils; Beleg aus Wiki-A-2015)
- (10) Gibt es überhaupt andere Artikel, die dem ähneln? Also **sprich** über Organhandel in anderen Ländern? Oder ist es nur der Kosovo? (http://de.wikipedia.org/wiki/Diskussion:Organhandel_im_Kosovo; Beleg aus Wiki-D-2015)
- (11) Es gibt da einen Unterschied zwischen dem A340 und dem iPhone: Erster wurde vor dem Zeitalter der Kommunikation entwickelt, **sprich**, deswegen findet man auch wenig Infos in Zeitungen etc. Und Zweitens ist das ganze zeitnaher geschehen – **sprich**, dort finden sich viele Informationen, die zwar momentan gerade relevant sind, aber wer weiß, ob sich in zehn Jahren noch wer dafür interessiert? (http://de.wikipedia.org/wiki/Diskussion:Airbus_A340/Archiv/1; Beleg aus Wiki-D-2015)

Im Handbuch der Konnektoren des IDS gehören *d.h.* und *sprich* zu den metakommunikativen Konnektoren, und zwar zur formulierungsbezogenen, identifizierenden Subklasse (Breindl/Volodina/Waßner 2014, S. 1140 ff). Metakommunikative Konnektoren verknüpfen ihre Konnekte „nicht auf der Ebene der Sachverhalte, sondern beziehen sich auf die Form des Ausdrucks der Konnekte oder deren informationsstrukturelle Rolle im Diskurs (ebd., S. 1138). Mit formulierungsbezogenen metakommunikativen Konnektoren wird eine Formulierungsalternative zu einem bereits versprachlichten Sachverhalt angeschlossen. Identifizierend bedeutet im engeren Sinne, dass ein Sprecher die beiden Konnekte als bedeutungsäquivalent ausweist, wie in Beispiel (9). Die Autoren weisen aber darauf hin, dass die Verwendungsmöglichkeiten – wie an Beispielen wie (10) und (11) deutlich wird – weit über diese enge Bestimmung hinausgehen (ebd., S. 1144).

Im Mittelpunkt unserer Untersuchung steht nicht die semantische Analyse von Einzelbeispielen, sondern der Vergleich der Vorkommensfrequenzen in den Artikeltexten der Wikipedia einerseits und in den Wikipedia-Diskussionen andererseits. Dieser Vergleich ist durch folgende Überlegung motiviert: Beim interaktionsorientierten Diskutieren ist es zwar prinzipiell möglich, sprachliche Formulierungen zu überarbeiten, bevor man den Beitrag veröffentlicht. Allerdings steht – wie in Abschnitt 2 erläutert wurde – die gemeinsame Bearbeitung eines aktuellen Problems im Vordergrund. Wenn man davon ausgeht, dass die an der Problemlösung Beteiligten gleichzeitig online und auf die gemeinsame Interaktion fokussiert sind, dann ist eine schnelle Reaktion oft zweckdienlicher als eine geschliffene und präzise Formulierung. Metakommunikative Konnektoren können beim schnellen und spontanen Schreiben ein probates Mittel sein, um bereits formulierte Textsegmente im weiteren Verlauf zu spezifizieren, zu generalisieren, zu erläutern oder zusammenzufassen. Die metakommunikativen Konnektoren übernehmen also Funktionen, die Kaiser/Schmidt (2016) für mündliche Gespräche an einem Sample von 45 *sprich*-Beispielen aus dem Gesprächskorpus FOLK herausgearbeitet haben.

Tab. 2: Ergebnisse der vergleichenden Untersuchung zu den Konnektoren *d.h.* und *sprich* in Korpora zu Wikipedia-Artikeln (Wiki-A-2015) und Artikeldiskussionen (Wiki-D-2015)

Suchform	Korpus	Frequenz absolut	Frequenz pmw	Sample	true positives im Sample
d.h.	Wiki-A	19.634	24,4	100	99
	Wiki-D	29.250	94,4	100	97
sprich	Wiki-A	1.698	2,1	100	61
	Wiki-D	8.732	28,2	100	91

Wie die Ergebnisse der Frequenzerhebungen in Tabelle 2 zeigen, kommen beide Konnektoren tatsächlich im Wiki-D-Korpus sehr viel häufiger vor als im Artikelkorpus Wiki-A. Da die beiden Korpora unterschiedlich groß sind, sind für den Vergleich vor allem die relativen Frequenzwerte aussagekräftig, die als pmw-Werte (Instanzen pro Million Wortformen; vgl. Keibel 2009) angegeben sind.

Die Vorkommenszahlen beziehen sich auf unbereinigte Trefferlisten. Um einen Eindruck von deren Präzision zu erhalten, haben wir in jedem Teilkorpus an einer Zufallsstichprobe von 100 Treffern manuell überprüft, wie häufig die Suchformen *d.h.* bzw. *sprich* tatsächlich als Konnektor verwendet werden. Wenig überraschend war, dass der Anteil der Konnektor-Verwendungen (true positives) bei *d.h.* deutlich höher liegt als bei der Suchform *sprich*, die auch als Imperativ verwendet wird. Interessant ist eher, dass der Anteil dieser – als falsch positiv gewerteten – Imperativbelege im Artikelkorpus Wiki-A-2015 höher ist als im Diskussionskorpus Wiki-D-2015, wo man eine direkte Ansprache und damit auch den Gebrauch des Imperativs eher erwarten würde. Der niedrigere Anteil in Wiki-D-2015 mag daran liegen, dass der Imperativ *sprich* in einer Konstellation, in der nun genau nicht gesprochen, sondern geschrieben wird, als ungewöhnlich empfunden wird. Die Durchsicht der Trefferliste aus Wiki-A-2015 zeigt außerdem, dass der Imperativ *sprich* dort sehr häufig als Bestandteil von Buch- und Filmtiteln bzw. von Zitaten vorkommt, also erwähnt, aber nicht verwendet wird. Jedenfalls zeigt die Auswertung der Stichproben, dass *sprich* als Konnektor vor allem im Artikelkorpus noch seltener vorkommt, als es der für die unbereinigte Trefferliste berechnete pmw-Wert nahelegt.

Aus Tabelle 2 kann man auch interessante Frequenzunterschiede zwischen den Konnektoren *sprich* und *d.h.* entnehmen. Einerseits wird der Konnektor *d.h.* in beiden Korpora deutlich häufiger verwendet. Andererseits ist die Diskrepanz zwischen den pmw-Werten in Wiki-A-2015 und Wiki-D-2015 bei *d.h.* sehr viel geringer als bei *sprich*. Wenn also *sprich* als Konnektor verwendet wird, dann sehr viel häufiger beim Diskutieren als beim Verfassen von Artikeltexten.

Intuitiv würde man diesen Unterschied zwischen *d.h.* und *sprich* darauf zurückführen, dass der Konnektor *d.h.* stilistisch unmarkiert ist, während der Konnektor *sprich* eher der informellen Alltagssprache zugeordnet wird. Leider gibt es kein professionell erarbeitetes Referenzwerk, gegen das man diese Intuition abgleichen könnte. Die Konnektorverwendung von *sprich* ist weder im gedruckten Duden-1 noch in Internetwörterbüchern wie dem DWDS-Wörterbuch oder dem Duden online erfasst. Diese Beschreibungslücke verwundert, weil die Verwendung von *sprich* als Konnektor weder neu noch selten ist. Die folgende Skizze des Entwicklungspfad der Imperativform von *sprechen* zum Konnektor *sprich* anhand von Belegen aus den DWDS-Korpora kann eine sorgfältige lexikografische Rekonstruktion nicht ersetzen. Sie soll lediglich deutlich machen, dass es sich bei die-

sem Konnektor – und vermutlich auch bei anderen metakommunikativen Konnektoren – lohnen kann, die diachrone Entwicklung und die Textsortenspezifika mit einzubeziehen.

Am Anfang des Grammatikalisierungsprozesses stehen Verwendungen wie in (12) und (13), bei denen der Autor eine Schreibform mit einer anderen Form verknüpft, aus der sich die Aussprache des Wortes rekonstruieren lässt.

- (12) Laubadere und sein General-Adjutant Doxon (**sprich** Dosson) befragten mich sofort über die Beschaffenheit der Belagerung der Stadt Landau. (Laukhard, Friedrich Christian: F. C. Laukhards Leben und Schicksale. Bd. 4,1. Leipzig, 1797; Beleg aus DWDS-DTA)
- (13) Die je nach Fassung etwa 2400 Strophen des Nibelungenlieds sind in 39 âventiuren (**sprich**: Aventüren) untergliedert. (<https://de.wikipedia.org/w/index.php?title=Nibelungenlied&oldid=153490598>; Beleg aus Wiki-D-2015)

Beispiel (12) aus dem Jahr 1797 ist ein früher Beleg für diesen Verwendungstyp im Deutschen Textarchiv (DWDS-DTA); in diesem Korpus findet man in wissenschaftlichen und enzyklopädischen Texten des 19. Jahrhunderts weitere Belege u.a. in Texten von Friedrich Schlegel, Alexander von Humboldt oder auch in Brehms Tierleben. Das Verwendungsmuster ist auch in verschiedenen Dekaden des DWDS-Kernkorpus (DWDS-KK) für das 20. Jahrhundert belegt. Wie das Beispiel (13) zeigt, findet man es auch in den enzyklopädischen Artikeln der Wikipedia.

Der Bezug zum Imperativ von *sprechen* ist in dieser Verwendung noch deutlich: Man könnte *sprich* auch als elliptische Form eines Satzes – etwa *Sprich es folgendermaßen aus!* – interpretieren. Ein Austausch durch *d.h.* ist bei diesem Verwendungstyp nicht möglich. Die Verknüpfung von Schreibformen mit Lautformen verschiebt die Aufmerksamkeit vom Lesen zum „inneren Hören“. Hieraus könnten sich die ab den 70er-Jahren des 20. Jahrhunderts im DWDS-KK belegten Verwendungsweisen wie (14) und (15) entwickelt haben, bei denen eine Verschiebung auf einer abstrakteren Ebene stattfindet, ein Wechsel zwischen verschiedenen Perspektiven, der durch den Konnektor *sprich* eingeleitet wird.

- (14) Dies ist der Grund dafür, daß die Vereinigten Staaten (die ersten Produzenten der neuen Technologien und Sitz der wichtigsten Firmen) zum Zwecke der Globalisierung der Wirtschaft ihr ganzes Gewicht in die Waagschale der Deregulierung warfen, um zu erreichen, daß möglichst viele Länder ihre Grenzen dem ›freien Informationsfluß‹ öffnen, **sprich** den Giganten der US-amerikanischen Medien- und Unterhaltungsindustrie. (Beck, Ulrich: Was ist Globalisierung?, Frankfurt a.M.: Suhrkamp 1997, S. 84; Beleg aus DWDS-KK)

- (15) Genau wie Willy Brandt bittet auch Schmidt die Abgeordneten, daß sie gerade in dieser sich mehr und mehr zuspitzenden Lage Disziplin zeigen und der FDP nicht die von Genscher so dringlich erwünschten Stichworte, **sprich** Vorwände, liefern. (Der Spiegel, 11.10.1982; Beleg aus DWDS-KK)

In Beispiel (14) referiert der Autor mit dem Ausdruck *freier Informationsfluss* auf ein Konzept, von dessen Gebrauch er sich aber – formal durch die Verwendung von Anführungsstrichen – distanziert. Indem er den Nebensatz durch *sprich* mit dem Ausdruck *Giganten der US-amerikanischen Medien- und Unterhaltungsindustrie* verknüpft, leitet er einen Wechsel von der Außenperspektive zur Innenperspektive ein. Das aus Innensicht formulierte Konnekt präzisiert die Zielrichtung des *freien Informationsflusses* und benennt die Profiteure der gewünschten Öffnung. Implizit macht der Autor in seiner Gegenüberstellung der Formulierungen aus Außen- und Innenperspektive auch deutlich, warum er sich von der Außenperspektive distanziert. Der Leser wird dazu eingeladen, die referierte Außenposition aus der Innensicht des Autors kritisch zu reflektieren. In beiden Beispielen wird dieser Übergang zwischen den Perspektiven durch *sprich* markiert, das in diesen Beispielen nur noch als Konnektor und nicht mehr als eine Kurzform des Imperativs interpretiert werden kann.

Im Beispiel (15) verknüpft *sprich* den neutralen Ausdruck *Stichworte* durch den negativ wertenden Ausdruck *Vorwände*, wobei beide Formulierungen auf dasselbe Denotat Bezug nehmen. Mit der Reformulierung vollzieht der Autor implizit einen Wechsel von der Perspektive Genschers zur Perspektive von Brandt und Schmidt und begründet damit implizit auch, warum die von Genscher *erwünschten Stichworte* aus der Perspektive von Brandt und Schmidt unerwünscht sind.

Beispiele wie (14) und (15) finden sich bis heute vor allem in journalistischen Texten. In den Wikipedia-Korpora sind sie – nach unserer Durchsicht der Belege – eher selten. Vielmehr wird *sprich* – wie die Beispiele (9), (10) und (11) zeigen – als ein sehr flexibel einsetzbarer Konnektor verwendet, der syntaktisch sehr unterschiedliche Konnekte verknüpft und als formulierungsbezogener metakommunikativer Konnektor ein breites semantisches und diskursfunktionales Spektrum abdeckt. Die Ergebnisse unserer quantitativen Studie legen nahe, dass *sprich* im Vergleich zu *d.h.* eher beim informellen Schreiben genutzt wird; dieser Eindruck müsste aber an Korpora aus verschiedenen Textsortenbereichen genauer überprüft werden.

Literatur

Wissenschaftliche Literatur

- Ágel, Vilmos (2017): Grammatische Textanalyse. Textglieder, Satzglieder, Wortgruppenglieder. Berlin/Boston.
- Androutsopoulos, Jannis (2011): Language change and digital media. A review of conceptions and evidence. In: Kristiansen, Tore/Coupland, Nikolas (Hg.): Standard languages and language standards in a changing Europe. (= Standard Language Ideology in Contemporary Europe 1). Oslo, S. 145–161.
- Barton, David/Lee, Carmen (2013): Language online. Investigating digital texts and practices. London/New York.
- Beißwenger, Michael (2007): Sprachhandlungskoordination in der Chat-Kommunikation. (= Linguistik – Impulse & Tendenzen 26). Berlin/New York.
- Beißwenger, Michael (2016): Praktiken in der internetbasierten Kommunikation. In: Deppermann, Arnulf/Feilke, Helmuth/Linke, Angelika (Hg.): Sprachliche und kommunikative Praktiken. (= Jahrbuch 2015 des Instituts für Deutsche Sprache). Berlin/New York, S. 279–310.
- Beißwenger, Michael (2018): Internetbasierte Kommunikation und Korpuslinguistik. Repräsentation basaler Interaktionsformate in TEI. In: Lobin, Henning/Schneider, Roman/Witt, Andreas (Hg.): Digitale Infrastrukturen für die germanistische Forschung. (= Germanistische Sprachwissenschaft um 2020 6). Berlin/New York, S. 307–349.
- Beißwenger, Michael/Storrer, Angelika (2010): Kollaborative Hypertextproduktion mit Wiki-Technologie. Beispiele und Erfahrungen im Bereich Schule und Hochschule. In: Jakobs, Eva-Maria/Lehnen, Katrin/Schindler, Kirsten (Hg.): Schreiben und Medien. Schule, Hochschule, Beruf. (= Textproduktion und Medium 10). Frankfurt a.M., S. 13–36.
- Breindl, Eva/Volodina, Anna/Waßner, Hans-Ulrich (2014): Handbuch der deutschen Konnektoren 2. Semantik der deutschen Satzverknüpfen. Teilbd. 2. (= Schriften des Instituts für Deutsche Sprache 13). Berlin/München/Boston.
- Deppermann, Arnulf/Reineke, Silke (Hg.) (2018): Sprache im kommunikativen, interaktiven und kulturellen Kontext. (= Germanistische Sprachwissenschaft um 2020 3). Berlin/Boston.
- van Dijk, Ziko (2010): Wikipedia. Wie Sie zur freien Enzyklopädie beitragen. München.
- Dittmar, Norbert/Bressemer, Jana (2005): Syntax, Semantik und Pragmatik des kausalen Konnektors *weil* im Berliner ‚Wendekorpus‘ der neunziger Jahre. In: Schwitalla, Johannes/Wegstein, Werner (Hg.): Korpuslinguistik deutsch: synchron, diachron, kontrastiv. Würzburger Kolloquium 2003. Tübingen, S. 99–125.
- Duden-Grammatik (2016): Der Duden in zwölf Bänden. Bd. 4: Die Grammatik. Hrsg. v. Angelika Wöllstein und der Dudenredaktion. 9., völlig überarb. u. aktual. Aufl. Berlin.
- Ebersbach, Anja/Glaser, Markus/Heigl, Richard (2016): Social web. 3., überarb. Aufl. Konstanz/München.
- Eichinger, Ludwig M. (2017): Gesprochene Alltagssprache. In: Deutsche Akademie für Sprache und Dichtung/Union der deutschen Akademien der Wissenschaften (Hg.): Vielfalt und Einheit der deutschen Sprache. Zweiter Bericht zur Lage der deutschen Sprache. Tübingen, S. 283–331.
- Gredel, Eva (2017): Digital discourse analysis and Wikipedia: Bridging the gap between Foucauldian discourse analysis and digital conversation analysis. In: Journal of Pragmatics 115, S. 99–114.

- Haack, Johannes (1995): Interaktivität als Kennzeichen von Multimedia und Hypermedia. In: Issing, Ludwig J./Klimsa, Paul (Hg.): Information und Lernen mit Multimedia. Weinheim, S. 151–166.
- Hammwöhner, Rainer (1997): Offene Hypertextsysteme. Das Konstanzer Hypertextsystem (KHS) im wissenschaftlichen und technischen Kontext. (= Schriften zur Informationswissenschaft 32). Konstanz.
- Hoffmann, Ludger (2014): Deutsche Grammatik. Grundlagen für Lehrerbildung, Schule, Deutsch als Zweitsprache und Deutsch als Fremdsprache. 2., neu bearb. u. erw. Aufl. Berlin.
- Hoffmann, Ludger (2018): Grammatik und gesprochene Sprache im Diskurs. In: Deppermann/Reineke (Hg.), S. 5–28.
- Imo, Wolfgang (2017): Interaktionale Linguistik und die qualitative Erforschung von CvK. In: Beißwenger, Michael (Hg.): Empirische Erforschung internetbasierter Kommunikation. (= Empirical Linguistics 9). Berlin/New York, S. 81–108.
- Kaiser, Julia/Schmidt, Thomas (2016): *sprich* als Reformulierungsindikator. Handreichung. Einführung in die Benutzung der Ressourcen DGD und FOLK für gesprächsanalytische Zwecke. Internet: <https://ids-pub.bsz-bw.de/frontdoor/index/index/docId/5539> (Stand: 15.7.2018).
- Kallass, Kerstin (2013): Schreibprozesse in der Wikipedia. Eine linguistische Analyse. Internet: <https://kola.opus.hbz-nrw.de/frontdoor/index/index/docId/726> (Stand: 12.3.2018). [Dissertation Universität Koblenz-Landau].
- Keibel, Holger (2009): Mathematische Häufigkeitsmaße in der Korpuslinguistik. Eigenschaften und Verwendung. Mannheim. Internet: www.ids-mannheim.de/kl/dokumente/freqMeasures.html (Stand: 12.3.2018).
- Kuhlen, Rainer (1991): Hypertext. Ein nicht lineares Medium zwischen Buch und Wissensbank. Berlin.
- Liebert, Wolf-Andreas/Kohl, Christian (2004): Selbstorganisation der Wissenschaftsvermittlung: Quellentransparenz, Kontroversität und Qualitätssicherung in der Internet-Enzyklopädie Wikipedia. In: Fachsprache 3–4, S. 134–148.
- Lüngen, Harald/Kupietz, Marc (2017): CMC corpora in DeReKo. In: Bański, Piotr et al. (Hg.): Proceedings of the Workshop on Challenges in the Management of Large Corpora and Big Data and Natural Language Processing (CMLC-5+BigNLP) 2017. Mannheim, S. 20–24. Internet: https://ids-pub.bsz-bw.de/frontdoor/deliver/index/docId/6259/file/Luengen_Kupietz_CMC+Corpora_2017.pdf (Stand: 15.7.2018).
- Margaretha, Eliza/Lüngen, Harald (2014): Building linguistic corpora from Wikipedia articles and discussions. In: Journal of Language Technology and Computational Linguistics (JLCL) 29, 2, S. 59–83.
- Marx, Konstanze/Weidacher, Georg (2014): Internetlinguistik. Ein Lehr- und Arbeitsbuch. Tübingen.
- Regelwerk 2018 = Deutsche Rechtschreibung. Regeln und Wörterverzeichnis. Aktualisierte Fassung des amtlichen Regelwerks entsprechend den Empfehlungen des Rats für deutsche Rechtschreibung 2016. Redigierte Fassung, veröffentlicht im Januar 2018. Internet: www.rechtschreibrat.com/DOX/rfdr_Regeln_2016_redigiert_2018.pdf (Stand: 15.7.2018).
- Storrer, Angelika (2004): Kohärenz in Hypertexten. In: Zeitschrift für germanistische Linguistik 31, 2, S. 274–292.
- Storrer, Angelika (2018): Interaktionsorientiertes Schreiben im Internet. In: Deppermann/Reineke (Hg.), S. 219–244.

- Wessler, Hartmut et al. (2017): Wikiganda. Detecting bias in multimodal Wikipedia entries. In: Seizov, Ognyan/Wildfeuer, Janina (Hg.): *New studies in multimodality*. London u.a., S. 201–224.
- Zifonun, Gisela/Hoffmann, Ludger/Strecker, Bruno (1997): *Grammatik der deutschen Sprache*. 3 Bde. (= Schriften des Instituts für Deutsche Sprache 7). Berlin/New York.

Korpusquellen

- DWDS-DTA: Deutsches Textarchiv (1600–1900): Verfügbar im Wortauskunftssystem DWDS der Berlin-Brandenburgischen Akademie der Wissenschaften. Internet: www.dwds.de/ (Stand: 2.7.2018).
- DWDS-KK: DWDS-Kernkorpus (1900–1999): Verfügbar im Wortauskunftssystem DWDS der Berlin-Brandenburgischen Akademie der Wissenschaften. Internet: www.dwds.de/ (Stand: 2.7.2018).
- Wiki-A-2015: Vollständige Kopie der Artikelseiten und der Artikeldiskussionen der deutschsprachigen Wikipedia (Version 17.11.2015). Institut für Deutsche Sprache Mannheim.
- Wiki-D-2015: Vollständige Kopie der Artikeldiskussionen der deutschsprachigen Wikipedia (Version 17.11.2015). Institut für Deutsche Sprache Mannheim. [Wiki-A-2015 und Wiki-D-2015 Download unter: <http://corpora.ids-mannheim.de/pub/wikipedia-deutsch/2015/> (Stand: 2.8.2016), recherchierbar mit dem Recherchesystem COSMAS II: www.ids-mannheim.de/cosmas2/ (Stand: 2.8.2016)].