

Narrative Muster und Diskursanalyse: Ein datengeleiteter Ansatz

Noah Bubenhofer, Nicole Müller, Joachim Scharloth, TU Dresden

Summary. Narratives are pre-existent socially accepted forms of interpretation, which delineate and confine the interpretation strategies in individual texts. Even though narratives shape the 'sayable' in a Foucaultian sense, they have so far been widely neglected by linguistic discourse analysis. In this paper we discuss data-driven methods in the analysis of narratives. Using 3376 stories of the first sexual encounter as an example, we propose using hierarchical ngram-collocation-graphs as a means of visualizing the complex co-occurrence-structure that forms narratives on the textual surface.

Zusammenfassung. Narrative sind sozial akzeptierte Interpretationsmuster, die unsere Wahrnehmung und Darstellung von Zusammenhängen ermöglichen und gleichzeitig begrenzen. Obwohl sie eine wichtige Funktion bei der Konstruktion der Grenzen des Sagbaren im Foucaultschen Sinn haben, wurden sie von der linguistischen Diskursanalyse bislang weitgehend vernachlässigt. Am Beispiel von 3376 Geschichten vom „Ersten Mal“ illustrieren wir in diesem Aufsatz methodische Ansätze zur Rekonstruktion von Narrativen. Das von uns vorgeschlagene Verfahren zur Analyse und Visualisierung von Narrativen ist die datengeleitete Berechnung von hierarchischen n-Gramm-Kollokationsgraphen.

1. Einleitung

Ich hatte mein erstes Mal -> mein erstes Mal mit # -> nahm mich in den Arm -> fragte er mich ob ich -> wir bei ihm zu Hause -> seine Eltern nicht da waren -> kam er auf mich zu -> mich zu küssen und ich -> legten uns auf sein Bett -> fragte mich was los sei -> noch nie einen Freund gehabt -> zogen wir uns gegenseitig aus -> Wir küsstet uns leidenschaftlich und -> Dann zog ich ihm seine -> Er schaute mich an und -> schaute mich an und fragte -> an und fragte ob ich -> mit ihm schlafen wolle und -> Er holte ein Kondom aus -> Dann drang er vorsichtig in -> er vorsichtig in mich ein -> Er fragte mich ob ich -> Als er merkte dass ich -> nahm mich in den Arm -> seit # Jahren zusammen und

Diese Phrasen bleiben von einer Geschichte vom „Ersten Mal“, wenn man von ihr das Vereinzelnde, Individualisierende wegnimmt und nur jene Teile der sprachlichen Gestaltung übrig lässt, die auch in anderen Geschichten zum gleichen Thema häufig vorkommen.

Geschichten vom „Ersten Mal“ sind aus paarsoziologischer Perspektive Darstellungen der „Aufnahme und Aufrechterhaltung einer als legitim anerkannten Sexualinteraktion“ als Schwellen-Wendepunkt in einer Zweierbeziehung (Lenz 2009: 105). War die Aufnahme einer sexuellen Beziehung früher an die Eheschließung geknüpft, hat das „Erste Mal“ in der Gegenwart den Status der ersten Stufe der Vertiefung und Verfestigung einer Beziehung (Lenz 2009: 109). Junge Menschen schreiben ihre Erlebnisse auf und publizieren sie in Webforen und schaffen so ein durch „Narration erzeugtes Bild des Paarseins“, konstruieren eine „narrative Paaridentität“ (Maier 2007: 49). Die einzelnen Geschichten weisen eine Serialität auf und bestehen in ihrer Summe aus typischen Versatzstücken, die immer

wieder auftreten. Solche typische Geschichten folgen also narrativen Mustern, die wiederum aus bestimmten sprachlichen Mustern bestehen.

Unser Ziel ist es, narrative Muster mit einer datengeleiteten Methode aus Textkorpora herauszuarbeiten. Damit verbunden ist das Plädoyer, das Konzept der Narrativität in diskursanalytische Arbeiten einzubeziehen und dabei statistisch-korpuslinguistische Methoden anzuwenden. In der Diskursanalyse spielen narrative Muster als Analysekategorie bislang kaum eine Rolle. Bei korpuslinguistischen Zugängen zu Diskursen wird zwar auf sprachliche Ausdrucksmittel auf verschiedenen Ebenen eingegangen (Lexeme, syntaktische Strukturen, Mehrworteinheiten etc.), jedoch kaum auf ihre Verkettung innerhalb eines Textes. Ein korpuspragmatischer, datengeleiteter Zugang zu Narrativität fokussiert aber genau auf solche Phänomene: Die musterhafte Abfolge von sprachlichen Ausdrucksmitteln wird dabei als Indiz für narrative Muster gedeutet.

In der vorliegenden Arbeit klären wir zuerst, was wir unter einem korpuspragmatischen, datengeleiteten Ansatz verstehen (Abschnitt 2). Anschließend gehen wir auf das Narrativitätskonzept ein und stellen Überlegungen an, wie sich dieses an diskursanalytische Ansätze anschließen lässt (Abschnitt 3). Im vierten Abschnitt entwickeln wir einen Ansatz, wie mit datengeleiteten korpuslinguistischen Methoden narrative Muster in Geschichten vom „Ersten Mal“ identifiziert und analysiert werden können.

2. Korpuspragmatische Analysemethoden

Wir verorten unsere Arbeit im Kontext eines noch jungen Forschungsparadigmas, der Korpuspragmatik (Scharloth & Bubenhofer 2011). 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 geht davon aus, dass sich pragmatische Informationen „im pragmatischen Mehrwert oder Gebrauchswert von Einheiten aller sprachlicher Strukturbereiche“ (Feilke 2000: 78) zeichenhaft manifestieren. Damit werden pragmatische Spuren an der sprachlichen Oberfläche, in die sich ein Gebrauchswert eingeschrieben hat, als Muster sichtbar (vgl. zu dieser Rehabilitierung der Textoberfläche auch den Sammelband von Feilke und Linke 2009). Diese „Sprachgebrauchsmuster“ (Bubenhofer 2009) werden damit als Ergebnis von sprachlich-sozialem Handeln gelesen und gedeutet.

Attraktiv an dieser Orientierung an sprachlichen Mustern auf der Textoberfläche sind die methodischen Möglichkeiten, die sich dadurch ergeben. Dank der Verfügbarkeit von großen Korpora und der maschinellen Verarbeitung von Textdaten wird es möglich, quantitativ zu arbeiten und Algorithmen zu entwickeln, um die Musterhaftigkeit in den Daten induktiv zu entdecken. Digitale Korpora sind somit nicht nur „Belegsammlungen oder Zettelkästen in elektronischer Form“, sondern ermöglichen eine eigene „korpuslinguistische Perspektive“ (Perkuhn und Belica 2006: 2).

Was ist mit dieser „korpuslinguistischen Perspektive“ gemeint? Die oben gemachten Ausführungen deuten es bereits an. Offensichtlich ist das quantitative Vorgehen: Nicht der Einzelbeleg interessiert, sondern die Musterhaftigkeit von Belegen, die als Ergebnis rekurrenter Sprachhandlungen gelesen werden können. Doch es kommt ein weiterer Aspekt hinzu: In Ergänzung zu den klassischen datenbasierten korpuslinguistischen Analysen, die der Überprüfung von Forschungshypothesen dienen (corpus-based-Paradigma), ist es fruchtbar, ein induktives Vorgehen der datengeleiteten Analyse (corpus-driven-Paradigma) hinzuzuziehen. Dieses Paradigma wird von Tognini-Bonelli (2001: 84 ff) vor dem Hintergrund der Arbeiten von Sinclair (1991) expliziert und im deutschen Sprachraum von meh-

ren Forscherinnen und Forschern verbreitet (Belica und Steyer 2006; Bubenhofer 2009; Perkuhn u.a. 2005; Steyer 2004; Scharloth u.a. 2013; Bubenhofer und Scharloth 2013a). Statt eine Hypothese mit vorher festgelegten Analysekatoren zu überprüfen, werden in einem Korpus sämtliche Zeichenkonfigurationen berechnet, die sich bei der Anwendung vorher festgelegter Algorithmen ergeben.

Eine sehr einfache Anwendung einer datengeleiteten Analyse ist das Berechnen des statistisch auffälligen Vokabulars in einem Untersuchungskorpus im Vergleich zu einem Referenzkorpus. Dabei werden die jeweiligen Frequenzen jedes Lemmas in den beiden Korpora miteinander verglichen. Je stärker die Frequenzen voneinander differieren, desto typischer ist das Lemma für das eine oder andere Korpus. Dieses Grundprinzip kann auf größere sprachliche Einheiten, zum Beispiel Wortgruppen, ausgedehnt werden.

Im Falle von narrativen Mustern müssen ebenfalls typische Mehrworteinheiten in den Untersuchungskorpora aufgespürt werden. Hinzu kommt dann aber ein weiterer Aspekt, nämlich deren typische Abfolge. Wir werden weiter unten genau spezifizieren, wie dies operationalisiert werden kann.

3. Narrativität und Diskurse

Bei der Konstitution, Formation und Vermittlung kultureller Wissensordnungen spielt das Erzählen als allgegenwärtige sprachliche Praxis eine nicht zu unterschätzende Rolle. Im Erzählen entfaltet sich stets eine Anordnung ausgewählter Versatzstücke von Wirklichkeitszugängen und findet eine Umwandlung von Kontingenz in anschlussfähige Sinnhaftigkeit statt (vgl. Meuter 2004). Aufgrund jener Prozesse der Sinn- und Selbsterzeugung müssen Erzählungen und ihre orchestrierten Verdichtungen in Form kultureller Narrative als Werkzeuge der Bedeutungsstiftung, als Mittel des Verstehens und Erklärens und somit als Modi der Welt- und Identitätskonstruktion verstanden werden.

Dieser wirklichkeitskonstruierende Effekt von Narrativität ergibt sich zunächst aus der im Prinzip des Erzählens angelegten Notwendigkeit zur Selektion. Um Sinnhaftigkeit herzustellen, muss zwangsläufig Komplexität reduziert und Information ausgewählt werden: „Sosehr das Erzählen sich in der Unendlichkeit sprachlicher Möglichkeiten entfaltet, an seinem Anfang – gleichsam als Einlasstor – steht der Vorgang der Reduktion. [...] Was über einen Gegenstand ausgesagt werden kann, ist gegenüber seinem Wahrnehmungsbild in hohem Maße ungenau, unvollständig und ergänzungsbedürftig.“ (Koschorke 2012: 28). Die Auswahlprozesse der Versprachlichung können nur diejenigen Ausschnitte einschließen, durch welche der Fortgang der Geschichte angetrieben und das ‚Rauschen‘ der Umwelt des jeweils Erzählten gefiltert wird.

Welche Elemente und Perspektiven bei diesem Prozess integriert werden, orientiert sich wiederum an sozial anschlussfähigen, eingespielten Selektionsverfahren und einer an kollektiven Wertigkeiten und Normen ausgerichteten Auswahlstruktur (vgl. u.a. Bruner 1991). In reflexiv-zirkulären dynamischen Prozessen werden diese konventionalisierten Repertoires erzählerischer Elemente durch die weitere Erzählproduktion in einem situativ, medial und sozial jeweils angemessenen und durchsetzungsfähigen Verhältnis sowohl aktualisiert als auch variiert¹, solange ein notwendiger Grad der Anschlussfähigkeit an be-

¹ Anschaulich beobachtet dies auch Doris Tophinke bei ihrer Untersuchung von Wirklichkeitserzählungen im Internet: „Geht die semantische Schematisierung zu weit und wird die Besonderheit des Einzelfalls zu wenig herausgearbeitet, wird dies moniert (‚Diese Geschichte gibts tausendmal hier‘, ‚die geschichte is sowas von klischeehaft, ich glaubs nich!‘). [...] Diese Formen des Erzählens lassen sich – vor dem Hintergrund dieser typischen Eigenschaften – in einem Feld verorten, das in der Horizontalen von der Unterscheidung ‚alltäg-

stehende Formen erhalten bleibt. Narrative sind daher immer auch Schauplätze der Ver- und Aushandlung von Sinn und Bedeutung; Bekanntes und Außerordentliches werden integrativ verknüpft durch den Effekt von „narrative’s apparatus for dealing simultaneously with canonicity and exceptionality“ (Bruner 1990: 47). Diese wechselseitige Dynamik zwischen Orchestrierung und Individualisierung von Erzählkompositionen (Koschorke 2012: 51-61). lässt sich exemplarisch an den Variationen der hier untersuchten Erzählungen des „Ersten Mals“ beobachten: In den dargestellten Graphen lassen sich mehrere charakteristische Erzählmuster bzw. ‚Ereignisketten-Muster‘ beobachten, aus denen verschiedene Durchgänge typisierbarer Schemata ableitbar sind.

Neben der Auswahl der Erzählelemente ist es des Weiteren ihre Sequenzialität, also ihre Anordnung, Reihenfolge und Kombination, die hochgradig bedeutungstiftende Effekte hervorbringt: Durch die Verortung auf einer zeitlich-chronologischen Achse werden ausgewählte Ereignisse bzw. Geschehensmomente mit Konzepten wie Ursache und Wirkung, Grund und Folge, Anfang und Ende, Vorher und Nachher verknüpft und Kategorien wie Aufstieg und Abstieg, Gewinn und Verlust, Schuld und Unschuld usw. erzeugt (Klausnitzer 2008: 216). Allgemein konstitutiv für das Erzählen ist die sprachliche Darstellung einer Veränderung, eines Wandels in der Zeit in linear geordneten Sätzen, die aufeinander bezogene episodenhafte Gestalten bilden. Die durch solche sinnstiftenden Ordnungsprozesse entstehenden Kompositionen dienen – um zwei zentrale Funktionen des Erzählens in Gesellschaften zu nennen – der Kohärenzstiftung sowie der Geschehensintegration des Erlebens aus Sicht der Erfahrenden. Handlungen, Ereignisse und Geschehnisse werden miteinander verknüpft und erhalten somit Kausalität und Bedeutung, auf die jeder menschliche Akteur zum Erhalt seiner subjektiv sinnhaften Handlungsfähigkeit angewiesen ist, um seine amorphe gelebte Erfahrung zu strukturieren (vgl. Lucius-Hoene 2010). Im Erzählprozess werden in der Retrospektive durch interpretative Kompositionen eigenen sowie fremden Handlungen und Entscheidungen Motiviertheit und Formen von Rationalität unterstellt, wodurch auch zukünftiges Handeln und Entscheiden sinnhaft gerichtet und somit zuallererst möglich werden. Durch diese sozial überformte Aktivität des ‚emplotment‘ (White 1990) – also der narrativen Organisation amorpher Geschehnisse bzw. der Überführung einer Ereignisabfolge in die Handlung einer Geschichte – stellen Narrative immer auch kulturelle Orientierungsfolien für individuelles Handeln und Erleben zur Verfügung, und werden so als Scharnier an der Schnittstelle zwischen Individuum und Kollektiv, zwischen dem Persönlichen und dem Allgemeinen wirksam.

Vor diesem Hintergrund wird besonders im Zusammenhang mit den hier untersuchten Geschichten erster sexueller Erfahrungen der Aspekt der Moralität, der Tabuisierung und der ‚Grenzen des Sagbaren‘ relevant. Da Narrative stets auf kulturell verfügbare Plotmuster, bevorzugt verwendete Erzählformen und dadurch entstehende präfigurierte Sinnmuster zurückgreifen, mittels derer Geschehensfragmente ausgewählt und in kausale Verhältnisse zueinander angeordnet und kontextualisiert werden, entsteht ein Erfahrungswissen über Sagbares und Nicht-Sagbares, wodurch eine stark wertende Dimension wirksam wird. Kollektive Vorstellungen und Werte werden hier aktiv konstituiert und aktualisiert, die Weltbilder und Wertehierarchien mitformen, nicht zuletzt auch durch die performative Kraft, die dem Erzählen innewohnt (vgl. Nünning 2013: 44 ff.). In Narrativen verstrickte Wertehierarchien und normativ gefärbte Geschehensabfolgen sind jedoch freilich keine fixierten Handlungsanweisungen, sondern Plattformen für die Ver- und Aushandlung von Bedeu-

lich‘ und ‚exzeptionell‘ und in der Vertikalen von der Unterscheidung ‚schematisch‘ und ‚spezifisch‘ aufgespannt wird.“ (Tophinke 2009: 252 f.)

tung und dadurch auch für die Bewältigung von Konflikten (vgl. etwa Koschorke 2011), in denen sich Gesellschaften ihre Versionen von gutem und schlechtem Handeln, von erstrebenswerten Zielen und den Mitteln zu ihrem Erreichen sowie ihren Kosten, Risiken und Folgen erzählen. Erzählen beinhaltet immer auch eine Stellungnahme und Positionierung des Erzählenden zu herrschenden sozialen Ordnungen.

Die bereits angeklungene Funktion des Erzählens, Vertrautes und Exceptionelles integrativ miteinander zu verknüpfen und somit auch potenziell Konflikthafes in eine veränderliche und dynamische Ordnung produktiv einzuflechten, ist im hier relevanten Kontext des „Ersten Mals“ zentral und verleiht der Untersuchung der Narrative ‚junger Liebe‘ eine besondere Bedeutsamkeit. Denn in diesem Fall geht es maßgeblich auch um die Integration eines außergewöhnlichen Geschehnisses in die für Handeln und Selbstverstehen des Akteurs ausschlaggebende Ordnung der eigenen Identitätskonstruktion. Die erste sexuelle Erfahrung stellt ein diskursiv stark aufgeladenes, schwellenhaftes Ereignis dar, das in der Regel als Bruch, als Neues, Unbekanntes und Irritierendes diskursiviert und erlebt wird. Als solches kann das Erzählen als sinn- und kohärenzstiftender Prozess der Konstruktion von Identität bei bruchhaften, ungewöhnlichen und irritierenden Erlebnisse gewissermaßen ‚domestizierende‘ Wirkung entfalten:

„The role of narrative in meaning construction becomes especially clear following anomalous or otherwise disturbing events. [...] Following such disturbing events, people generally become talkative. They tell and retell the story until the events are gradually domesticated into one or more coherent or shared narratives that circulate among the community of sufferers. The meanings are emergent in the narrative process. Through narrative, the strange and the familiar achieve a working relationship“ (Shore 1996: 58).

Im Zentrum der Geschichte steht somit oftmals ein das Vorher und Nachher bestimmendes Ereignis, das einen Wandel herbeiführt und eine ‚Diskontinuität‘ oder ein ‚destabilisierendes Element‘ (Lucius-Hoene/ Deppermann 2002: 22) markiert. Speziell biographisches Erzählen in seinen vielfältigen Formen – in der Alltagskommunikation (vgl. bspw. Fuchs-Heinritz 2009), in Therapiegesprächen (vgl. etwa Rabelhofer 2013 oder Illouz 2006), in schriftlichen autobiographischen Dokumenten usw. – dient in entscheidendem Maße der Identitätsarbeit, die Praktiken der interaktiv und situativ eingesetzten Selbstverständigung und Verhandlung von Erfahrungen beschreibt. Die Herstellung narrativer Identität als kommunikative Praxis dient nicht nur der Bewältigung von Erlebtem, sondern der prozessualen Vergewisserung, dem Erhalt und der sozialen Positionierung des Selbst. In Anlehnung an Paul Ricœurs Konzept „narrativer Identität“ (Ricœur 2005) wird Identität hier als eng an das Narrative gekoppelt verstanden: Durch das Medium erzählter Erfahrung konstituiert sich Identität performativ und wird somit im Erzählen zuallererst hergestellt.

„Narrative Identität ist somit nicht Teil der erzählten Geschichte, die ihrerseits auf eine wie auch immer geartete Wirklichkeit verweist, oder Ergebnis eines inneren Monologs oder Erinnerungsprozesses, sondern sie entsteht in der Situation des Erzählens selbst, ist interaktiv und situationsspezifisch hergestellt mittels sprachlicher Praktiken. Identität ist ein ‚practical accomplishment‘ [...] sie ist die Art und Weise, wie Menschen in Erzählungen sich selbst als Protagonisten oder Akteure erscheinen lassen und für sich und den Hörer Sinn zu stiften suchen“ (Lucius-Hoene 2010: 154f).

Erzählen Akteure autobiographisch von Selbsterlebtem, so „praktizieren wir beim Erzählen eine Erkenntnisform und vollbringen eine soziale Leistung, in der sich unsere Identität wie unser Verhältnis zur Welt vollziehen.“ (Lucius-Hoene/ Deppermann 2002: 21).

Doch auch in ihren Selbsterzählungen bedienen sich Akteure zur Herstellung narrativer Identität erzählerischer und diskursiver Muster und kultureller Orientierungsfolien: Speziell in kommunikativen Austauschsituationen – sowohl in persönlichen Gesprächen als auch in Internetforen – ist die interaktive Konstruktion von Identität und sozialer Positionierung als kreative Inszenierung, Aushandlungs- und Selbstverhandlungsprozess angewiesen auf sprachliche Ressourcen, welche die Sprecher im Laufe ihres Lebens erworben haben. Diese Vorlagen enthalten spezifische, bewegliche narrative Strategien und dynamische Darstellungsregeln, die sich als Repertoire situativ angemessener und durchsetzungsfähiger Muster sedimentieren. Das hier interaktiv hergestellte, stets variierte und aktualisierte Repertoire unterliegt zweifellos auch gender-, milieu- und altersspezifischen Schwankungen. Dennoch findet hier eine „Rückbindung der lokalen sprachlichen an die allgemeinen sozialen und kulturellen Praktiken“ (Lucius-Hoene/ Deppermann 2002: 163) statt.

Angesichts dieser Scharnierfunktion von Narrativen ist es bemerkenswert, dass die Diskursforschung es bislang weitgehend versäumt hat, Narrativität als Form- und Strukturprinzip in seiner Ordnung und Sinn generierenden Funktion sowie Erzählen als diskursivierte und diskursivierende soziale Praxis angemessen zu integrieren. Außer Zweifel steht, dass einerseits Erzählungen, ihre deutende und bedeutende Organisation an diskursiver Sinnproduktion maßgeblich beteiligt sind sowie andererseits diskursive Wissensgenerierung kulturelle Narrative grundlegend mitprägt, denn „Erzählungen gedeihen im Medium sozial geteilten Wissens.“ (Koschorke 2012: 37) Aus diesem Grund möchte der vorliegende Aufsatz einen Beitrag dazu leisten, die Lücken zwischen Biographie-, Narrativitäts- und Diskursforschung zu schließen. Denn erst der „konfigurative Akt des Erzählens, die *Narrativierung* von Ereignissen als diskursive Handlung, erlaubt es weiterhin, die episodenhafte Zerstreung der Erzählelemente zu integrieren.“ (Viehöfer 2012: 207) Also solche fungieren Narrative auch als „soziale[] Skripte“ (ebd: 193) und Sinnfiguren, die diskursiv erzeugtes Wissen über narrative Identitäten in das Selbstverstehen sozialer Akteure hinein vermitteln. Aus diesem Grund soll im Folgenden das Auftauchen typischer Versatzstücke und Bausteine von Erzählungen und ihre jeweilige Anordnung und Verknüpfung in der Diskursgestalt untersucht werden. Narrative Strukturen zeichnen sich durch abgrenzbare Episoden aus, die durch ihre Anordnung zu Kausalitäten verwoben werden und dadurch den roten Faden bzw. die *story line* konstituieren:

„In synchroner Hinsicht verknüpfen sie die unterschiedlichen Deutungselemente eines Diskurses zu einem zusammenhängenden, erzählbaren Gebilde. [...] Durch den Rückgriff auf eine *story line* können Akteure diskursive Kategorien sehr heterogener Herkunft in einem mehr oder weniger kohärenten Zusammenhang aktualisieren“ (Keller 2011: 252).

Auf diese Weise werden die diskursiven Bausteine in einer Dramaturgie herstellenden Weise verwendet, deren nähere Untersuchung für die Diskursforschung durchaus lohnenswert wäre.

4. Datengeleitete Analyse narrativer Muster: Eine Illustration von Methoden am Beispiel von Erzählungen vom Ersten Mal

Im Folgenden wollen wir unseren methodischen Ansatz zur Identifizierung narrativer Muster in Erzählungen vorstellen und am Beispiel eines umfangreichen Korpus mit Geschichten vom Ersten Mal illustrieren.

4.1 Das Korpus „Junge Liebe“

Das für die folgende exemplarische Analyse zusammengestellte Junge-Liebe-Korpus besteht aus 3376 Geschichten von meist jungen Personen, die die Geschehnisse rund um „Das erste Mal“ autobiographisch darstellen. Mit wenigen Ausnahmen wurden alle Erzählungen von heterosexuell orientierten Personen verfasst. Weil die Anzahl der Erzählungen von gleichgeschlechtlichen Erlebnissen keine aussagekräftigen statistischen Analysen ermöglicht hätte, besteht das Korpus ausschließlich aus Geschichten mit heterosexueller Konstellation. Sämtliche Geschichten wurden von den – inzwischen teilweise nicht mehr erreichbaren – Internet-Plattformen rockundliebe.de (2094 Erzählungen), Erstes-Mal.com (385 Erzählungen) und planet-liebe.de (897 Erzählungen) entnommen. Die Webseiten wurden automatisiert heruntergeladen, die Texte extrahiert, mit Metainformationen (Alter beim Ersten Mal und Geschlecht) versehen, mit Hilfe des TreeTagger lemmatisiert und mit Part-of-speech-Informationen annotiert. Zusätzlich wurden alle Zahlen durch ein Raute-Symbol ersetzt.

Insgesamt umfasst das Korpus 1.886.588 laufende Wortformen. Im Hinblick auf die Dimension Geschlecht ist das Korpus ungleich verteilt: rund 73% der Geschichten stammen von Frauen, nur rund 27% von Männern (vgl. Tabelle 1).

	Anzahl Geschichten	laufende Wortformen	durchschnittliche Länge	Durchschnittsalter
Frauen	2458	1.395.896	567.9	15,8
Männer	918	490.692	534.5	16,8

Auffällig ist zudem, dass Geschichten von Frauen im Durchschnitt rund 33 Wörter länger sind als die von Männern.

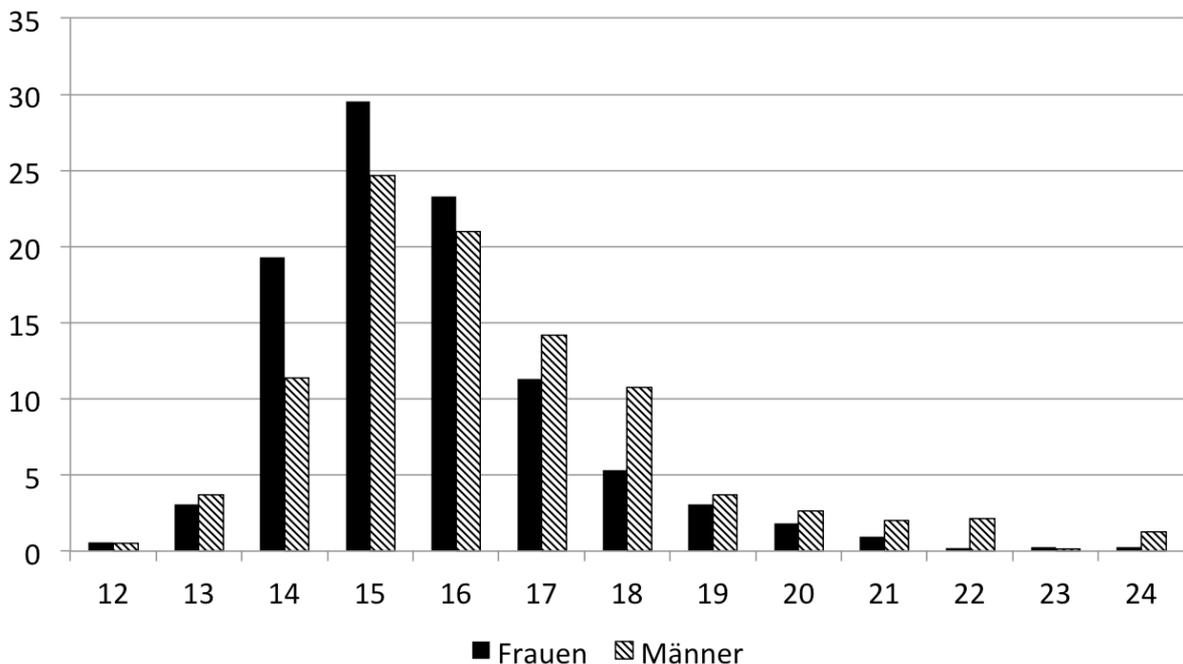


Abbildung 1: Alter beim Ersten Mal in Abhängigkeit von Geschlecht im Junge-Liebe-Korpus.

Auch hinsichtlich des Alters, das die Autorinnen und Autoren nach eigenen Angaben bei ihrem Ersten Mal hatten, unterscheiden sich die Korpora: Die Männer sind in ihren eigenen Texten im Durchschnitt 16,8 Jahre alt, die Frauen in ihren eigenen Texten im Durchschnitt ein Jahr jünger. Das Alter des Partners war leider nicht in hinreichend vielen Fällen zu ermitteln.

Im Folgenden soll also exemplarisch untersucht werden, welche geschlechtsspezifischen Unterschiede Erzählungen vom „Ersten Mal“ aufweisen. Die Illustration datengeleiteter methodischer Zugänge zu Erzählungen steht dabei im Vordergrund.

4.2 Frequente und typische Mehrworteinheiten

Als geeignete sprachliche Muster, aus denen sich narrative Muster zusammensetzen, scheinen uns Wortformen-n-Gramme als besonders geeignet. N-Gramme sind allgemein Folgen von n Einheiten, in unserem Falle von Wörtern. Für die Berechnung der Mehrworteinheiten müssen eine Reihe von Parametern gesetzt werden, wie Länge der Mehrworteinheit oder maximale Spannweite zwischen erstem und letztem Wort der Mehrworteinheit. Die Berechnungen in diesem Aufsatz sind größtenteils mit den folgenden Parametern durchgeführt worden:

- Länge des n-Gramms: 5 Wörter (Pentagramme)
- Spannweite (max. Distanz zwischen erstem und letztem Wort): 6 Wörter
- Basis: Wortformen
- Mindestfrequenz: 5 pro Mio. Wörter

Die Berechnung der n-Gramme erfolgte teils mit dem *Ngram Statistics Package* (NSP, Banerjee und Pedersen 2003), teils mit eigener Software.

Traditionelle korpuspragmatische Zugriffe auf Textkorpora suchen nach frequenten Mustern bzw. nach Mustern, die für ein Korpus im Vergleich zu einem oder mehreren anderen Korpora typisch sind. Untersucht man etwa, welche Pentagramme für das Männer- und das Frauenkorpus jeweils typisch sind, so erhält man folgende Listen:

Männer-Korpus				Frauen-Korpus			
llr	n-gram	f(1)	f(2)	llr	n-gram	f(1)	f(2)
145,33	fragte sie mich ob ich	0	54	80,84	drang er in mich ein	134	0
88,81	fragte ich sie ob sie	0	33	77,82	ob ich mit ihm schlafen	129	0
75,36	drang ich in sie ein	0	28	68,97	fragte er mich ob ich	167	5
67,28	Ich fragte sie ob sie	0	25	60,93	in mich ein Es tat	101	0
64,59	drang langsam in sie ein	0	24	60,93	legte er sich auf mich	101	0
64,59	setzte sie sich auf mich	0	24	47,66	legte sich auf mich und	79	0
64,59	und zog es mir ueber	0	24	47,66	und drang in mich ein	79	0
61,9	setzte sich auf mich und	0	23	45,85	nahm mich in den Arm	76	0
59,21	sie sich auf mich und	0	22	44,64	und legte sich auf mich	74	0

56,52	ob ich mit ihr schlafen	0	21	44,04	fing er an mich zu	73	0
53,83	Sie fragte mich ob ich	0	20	43,43	er sich auf mich und	72	0
53,83	in sie ein Es war	0	20	42,83	in mich ein Es war	71	0
53,83	mir ein Kondom ueber und	0	20	41,81	Er fragte mich ob ich	123	6
53,83	und ich fragte sie ob	0	20	41,02	und zog es sich ueber	68	0
51,13	fluesterte sie mir ins Ohr	0	19	40,42	ihn in mir zu spueren	67	0
51,13	ich fragte sie ob sie	0	19	40,42	Er legte sich auf mich	67	0
48,44	an mir einen zu blasen	0	18	38	er fragte mich ob ich	63	0
48,44	ich drang in sie ein	0	18	38	mich ob ich mit ihm	63	0
48,44	legte sich auf den Ruecken	0	18	35,59	fragte mich ob ich es	59	0
48,44	mir das Kondom ueber und	0	18	34,38	Ich war mit meinem Freund	57	0

Aus diesen Listen wird unter anderem erkennbar, dass die verbale Handlung des Fragens, oder präziser: des Einholens von Einverständnis, offenbar häufig Bestandteil von Erstes-Mal-Erzählungen sind. Ebenso zeigen sich einige wenige geschlechtsspezifische Unterschiede: etwa die Referenz auf die Dauer der Beziehung („Ich war mit meinem Freund“).

Für unsere Fragestellung ist die Methode dennoch wenig geeignet, weil sich in den obersten Rängen vor allem Mehrworteinheiten finden, die im Genus der Personalpronomina auf das Geschlecht des Autors / der Autorin bzw. des Partners / der Partnerin verweisen. So ist die gleiche Handlung, etwa das Fragen nach Einverständnis oder das Eindringen in beiden Textkorpora ähnlich präsent, erscheinen hier aber aufgrund der Pronomina als hochsignifikante Differenzen. Aber nicht nur wegen dieser Spezifik autobiographischer Erzählungen ist das Ergebnis nur wenig zielführend. Diese Vorgehensweise ist auch insofern unbefriedigend, als sie uns zwar etwas über die Typizität des Auftretens von Mehrworteinheiten verrät, nichts jedoch über ihre Abfolge oder gar Verkettung.

4.3 Typische Musterpositionen und Positionsdifferenzen

Als eine erste Annäherung an die narrative Struktur haben wir die typischen Positionen von n-Grammen in den Texten bestimmt. Hierfür haben wir alle Texte in mehrere jeweils gleich große Teile geteilt und dann untersucht, in welchem Teil der Erzählungen sich die n-Gramme im Durchschnitt wiederfinden.

In den beiden Abbildungen (2 und 3), die die geschlechtsspezifische Verteilung von n-Grammen auf Textabschnitte zeigen, springt zunächst die ungleiche Füllung der einzelnen Abschnitte mit n-Grammen ins Auge. Während in den Abschnitten 3, 4 und 5 (mit Maximum in Abschnitt 4) sehr viele n-Gramme zu finden sind, sind es in den anderen deutlich weniger. Dies ist ein Effekt der gewählten Grenzwerte, der gleichwohl aussagekräftig ist: in den Abschnitten 1, 2 und 6 gibt es weniger hochfrequente n-Gramme, was bedeutet, dass die Erzählungen hier eine größere Variation aufweisen. Dies ist insofern bemerkenswert, als üblicherweise Geschichtenanfänge und -enden rituell geformt und damit in höherem Maße mit Phrasen durchdrungen sind. In unserem Korpus ist es allerdings so, dass die

Schilderung des Kerngeschehens, der sexuellen Handlung, offenbar sprachlich sehr stereotyp erfolgt. Diese Beobachtung trifft für die männlichen wie weiblichen Erzählungen zu. Geschlechtsspezifische Unterschiede sind auf dieser Ebene nicht feststellbar.

Diese können sichtbar werden, wenn man die Positionsdifferenzen einzelner n-Gramme bei in den Erzählungen von Männern und Frauen vergleicht. Um die Verteilung der n-Gramme vergleichbar zu machen, wurden die Frequenzvektoren 0/1-normalisiert. Der Wert 1 steht also für die höchste Frequenz, der Wert 0 für die niedrigste. Ein Wert nah bei 1 verweist auf eine Frequenz nahe dem Höchstwert.

<p>Mein erstes Mal hatte ich (64) Ich war mit meinem Freund (57) Mein Freund und ich waren (50) # Monate mit meinem Freund (44) Ich hatte mein erstes Mal (43) Monate mit meinem Freund zusammen (39) hatte mein erstes Mal mit (32)</p>	<p># Jahre älter als ich (56) mit meinem Freund zusammen und (46) über Gott und die Welt (31)</p>	<p>fragte er mich ob ich (167) Lisa Lisa Lisa Lisa Lisa (153) fing er an mich zu (72) dass ich mit ihm schlafen (67) fing wir die uns zu (46) uns tief die Augen (43) Ich mit ihm schlafte wollte (41) Ich sagte ihm dass ich (41) an uns zu küssen und (41) fragte er mich ob wir (40) wir an uns zu küssen (37) wusste dass ich noch Jungfrau (35) zogen wir uns gegenseitig aus (32) an und fragte ob ich (30) an mich zu küssen und (30) ich drehte mich um und (30)</p>	<p>ob ich mit ihm schlafen (129) und fragte mich ob ich (125) Er fragte mich ob ich (123) legte er sich auf mich (111) tief in die Augen und (88) legte sich in die Augen (80) mich tief in mich und (79) und legte sich auf mich (74) dass ich mich auf mich und (72) er zog es sich über (68) Er legte sich auf mich (67) mich ob ich mit ihm (63) er fragte mich ob ich (63) fragte mich ob ich es (60) sich ein Kondom über und (59) mich auf den Rücken und (51) und er fragte mich ob (49) mich in den Arm und (47) Ich zog uns gegenseitig (43) Ich mit ihm schlafen will (41) und flüsterte mir ins Ohr (41) sah mir in die Augen (41) dass fragte er mich ob (40) Das Kondom über und (40) ich mit ihm schlafen wollte (39) fragte mich ob ich (36) sagte mir dass ich (36) dass ich mein erstes Mal (34) zog sich ein Kondom über (34) fragte mich ob ich mir (34) fing an mich zu lecken (33) und streifte es sich über (33) Dann fragte er mich ob (33) beugte sich über mich und (33) ich mein erstes Mal mit (33) ob ich mir stehet sei (32) zu küssen und zu stechen (32) wir an uns zu küssen (32) fingen an uns zu küssen (32) schauten wir tief in die (31) Er nahm ein Kondom aus (31) mich ob ich es wirklich (30) sein Penis in den Mund (30) Da hing er in mich (30) ob ich es wirklich will (30)</p>	<p>drang er in mich ein (132) in mich ein Es tat (101) und drang in mich ein (79) in mich ein Es war (71) ihm in mir zu spüren (66) drang langsam in mich ein (66) drang vorsichtig in mich ein (67) sich auf mich und drag (62) ganz vorsichtig in mich ein (56) in mich ein und ich (48) in mich ein und ich (48) als er in mich eindrang (45) Als er mich eindrang (43) Es tat überhaupt nicht wie (40) langsam und vorsichtig in mich (39) und vorsichtig in mich ein (38) und drang vorsichtig in mich (34) zog es sich über und (34) und drang langsam in mich (33) langsam in mich ein Es (33) er in mich ein Es (32) er legte sich auf mich (32) dem drang er in mich (32) mich und drang in mich (31) in mich eindringt wie es (31) Er nahm mich in den (31) drang in mich ein Es (30) mich ein Es war ein (30) mein erstes Mal mit ihm (30)</p>	<p>sind immer noch zusammen und (34) Wir sind immer noch zusammen (33) Ich liebe dich über alles (30)</p>
<p>Abschnitt 1</p>	<p>Abschnitt 2</p>	<p>Abschnitt 3</p>	<p>Abschnitt 4</p>	<p>Abschnitt 5</p>	<p>Abschnitt 6</p>

Abbildung 2: Typische Positionen von n-Grammen in den Erzählungen vom Ersten Mal (Frauen). Pentagramme, Window: 6, Grenzwert: 30.

<p>Mein erstes Mal hatte ich (27) erstes Mal hatte ich mit (16) Mein Mal hatte ich mit (14) hatte mein erstes Mal mit (13) Jahr mit meiner Freundin zusammen (13) Mal hatte ich mit # (13) Meine Freundin hatte ich mit (13) Meine Freundin und ich waren (12) Mein erstes Mal ich mit (12) # mit meiner Freundin zusammen (12) euch von meinem ersten Mal (12) erstes Mal hatte ich # (12)</p>	<p>wir an uns zu küssen (20) fragte mich ob ich nicht (19) gingen wir an uns zu (18) mich ob ich Lust hätte (15) ob ich Lust hätte mit (13) ich mich in sie verliebt (13) wir uns das erste Mal (13) sie mich ob ich ihr (13)</p>	<p>fragte sie mich ob ich (59) fragte ich sie ob sie (40) ich fragte sie ob sie (32) gingen an uns zu küssen (31) und fragte mich ob ich (27) sie mich ob ich mit (26) mich ob ich mit ihr (25) fragte mich ob ich mit (25) und fragte sie ob sie (24) tief in die Augen und (22) Sie fragte mich ob ich (22) sie ob sie mit mir (19) und fing an mich zu (19) fing an mich zu küssen (19) sagte sie mir dass sie (19) wir uns in die Augen (17) ob sie mit mir schläfen (17) sie mich ob ich nicht (17) fragte sie ob sie mit (17) zog mir mein Shirt aus (16) uns tief in die Augen (16) mir tief in die Augen (16) fragte sie ob ich mit (16) mich ob ich ihr schläfen (16) mich ich mit ihr schläfen (15) an und fragte ob ich (15) zogen wir uns gegenseitig aus (15) mich ob ich mit ihr schläfen (15) mich ob ich mit ihr (15) ich sie ob sie mit (15) sie ob ich mit ihr (15) und lingen an uns zu (15)</p>	<p>Wir zogen uns aus und (30) legte sich auf den Rücken (29) nahm ihn in den Mund (29) fing an mir einen zu (25) Wir zogen uns gegenseitig aus (24) sich auf den Rücken und (24) ob ich mit ihr schläfen (24) zogen uns gegenseitig aus und (23) und legte mich auf sie (22) und ich fragte sie ob (21) Sie fing an zu stöhnen (21) flüsterte sie mir ins Ohr (21) ich frachte sie ob sie (21) ich legte mich auf sie (21) sie fragte mich ob ich (21) setzte sie sich auf und (21) und zogen uns gegenseitig aus (20) zog mir mein aus und (20) mir das Kondom über und (19) und fing an sie zu (19) ich mich auf den Rücken (19) setzte sie sich mich und (18) nahm sie ihn in den (18) an mir einen zu blasen (18) ich legte mich sie und (17) und sie sich auf mich (17) und ich fragte ob sie (17) Dann setzte sie sich auf (16) nahm sie in den Mund (16) fing an sie zu lecken (16) zog mir die Hose aus (16) ihn in den Mund und (16) kam sie zu mir und (16) sie ihn in den Mund (16) Sie setzte sich auf mich (16) meinen Penis in den Mund (15) fing an sie zu küssen (15) in die Hand und ihn (15) Wir uns gegenseitig aus und (15) fing an mir zu blasen (15) fing an einen zu blasen (15) legte sich auf den und (15) Wir zogen gegenseitig aus und (15) Sie zog mir mein aus (15) und ich sie ob sie (15)</p>	<p>und drang in sie ein (57) drang langsam in sie ein (40) setzte sich auf mich und (39) drang ich in sie ein (38) ich drang in sie ein (33) sie sich auf mich und (32) und zog es mir über (31) und setzte sich auf mich (26) und ich drang in sie (26) setzte sie sich auf mich (26) ein Kondom aus Tasche und (25) legte mich auf sie und (25) zog mir ein Kondom über (23) sie mir ein Kondom über (22) drang vorsichtig in sie ein (22) mir ein Kondom über und (22) und drang langsam in sie (21) in sie ein Es war (21) und streifte es mir über (21) drang in sie ein und (21) ich langsam in sie ein (21) mich auf den Rücken und (20) drang in sie ein Es (19) ich drang in sie ein (19) in sie ein und sie (18) Ich auf den Rücken und (17) streifte mir ein Kondom über (17) und ich drang in ein (16) und ich drang in sie ein (16) und ich drang sie ein (16) streifte es mir über und (16) Ich legte mich auf sie (16) auf den Rücken und sie (15) sie auf den Rücken und (15) und fing an zu stöhnen (15) ein Kondom aus der und (14) setzte sie auf mich und (14) und drang vorsichtig in sie (14) drang in sie ein Sie (14) legte ich auf den Rücken (14)</p>	<p>Wir sind immer noch zusammen (19) sind immer noch zusammen und (17) ein Am nächsten Morgen wir (13)</p>
<p>Abschnitt 1</p>	<p>Abschnitt 2</p>	<p>Abschnitt 3</p>	<p>Abschnitt 4</p>	<p>Abschnitt 5</p>	<p>Abschnitt 6</p>

Abbildung 3: Typische Positionen von n-Grammen in den Erzählungen vom Ersten Mal (Männer). Pentagramme, Window: 6, Grenzwert: 12 (teilweise gekürzt).

Die folgenden Abbildungen zeigen die Distribution einiger n-Gramme, deren Positionierung im Text geschlechtsspezifische Unterschiede aufweisen. Dies sind beispielsweise n-Gramme, die sexuelle Erfahrung und Beziehungsstatus betreffen (Abbildung 4).

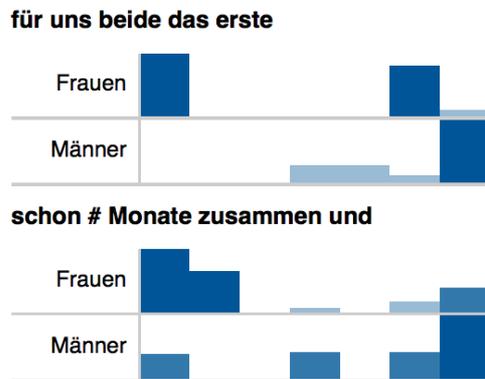


Abbildung 4: Vergleich der Distribution von n-Grammen, die auf sexuelle Erfahrung und Beziehungsstatus Bezug nehmen, in Texten von Männern und Frauen.

Während das n-Gramm „für uns beide das erste“ von Frauen im ersten und vorletzten Abschnitt am häufigsten gebraucht wird, erwähnen Männer die Tatsache, dass es für beide das Erste Mal war, erst am Ende ihrer Erzählungen. Auch das n-Gramm „schon # Monate zusammen und“ wird von Frauen dominant in den ersten Teilen ihrer Geschichten verwendet, Männer hingegen benutzen es am Ende. Eine Kontextanalyse zeigt allerdings, dass bei Verwendung des n-Gramms am Ende einer Erzählung der Geschlechtsakt der Auftakt der Beziehung war, die ihre Fortsetzung bis in die Gegenwart zum Zeitpunkt des Schreibens hat; die Verwendung des n-Gramms zu Beginn einer Erzählung stellt die Dauer der bereits bestehenden Beziehungen dar. Die Tatsache, dass Frauen häufiger am Anfang einer Erzählung die Beziehungsdauer erwähnen, kann als Indiz dafür angesehen werden, dass sich Frauen auf die Einhaltung einer subsistenten Norm (im Sinne von Gloy 1975) verpflichtet fühlen, einer unausgesprochenen Norm, die allzu frühe sexuelle Kontakte innerhalb einer Beziehung sanktioniert.

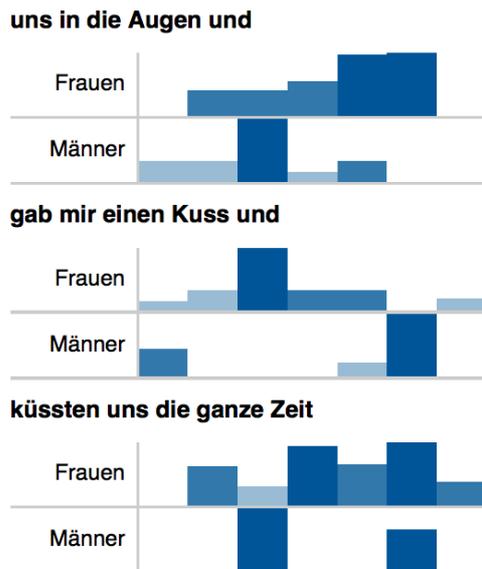


Abbildung 5: Vergleich der Distribution von n-Grammen, die auf Schlüsselhandlungen Bezug nehmen in Texten von Männern und Frauen.

Größere Differenzen in der Distribution zeigen sich auch bei n-Grammen, die auf Schlüsselhandlungen im Kerngeschehen verweisen (Abbildung 5). So sind die n-Gramme „uns in die Augen und“ und „gab mir einen Kuss und“ je gegensätzlich verteilt. Während in den Erzählungen der Frauen der Kuss am Anfang jener Abschnitte zu finden ist, die sich mit sexuellen Handlungen befassen, berichten Männer hier vorwiegend von Blicken in die Augen; Männer berichten, am Ende der sexuellen Aktivitätsphase geküsst zu werden, Frauen erzählen hier dagegen vom Austausch von Blicken. Dies könnte man so deuten, dass für Frauen mit dem Vollzug des Geschlechtsaktes eine Intensivierung der Beziehung einhergeht, die für den Mann durch die Gabe des Einverständnisses zum Geschlechtsakt durch den tiefen Blick bereits erreicht ist und sich dann im Akt manifestiert. Ein weiterer Aspekt könnte sein, dass Männer narratologisch versichern wollen, dass Einverständnis vorgelegen hat, Frauen dagegen, dass emotionale Nähe vorgelegen hat. Dies könnte ein Hinweis darauf sein, dass sich aufgrund kultureller Stereotype geschlechtsspezifische Ängste mit dem „Ersten Mal“ verbinden. In diesen Kontext passen auch die Positionsdifferenzen des n-Gramms „küssten uns die ganze Zeit“. Während das fortwährende Küssen in den Erzählungen der Männer Teil von „Vor-“ bzw. „Nachspiel“ zu sein scheint, schildern Frauen ihr Erstes Mal so, dass das Küssen Bestandteil aller Phasen des Kerngeschehens sein kann.

4.4 Musterverknüpfungen

Die bisherigen Analysen haben die Typizität und Positionierung von Mustern in Erzählungen in den Blick genommen. Sie sind aber insofern defizitär im Hinblick auf die Fragestellung, als sie die Muster jeweils atomistisch betrachteten. Zwar war die Aussage möglich, dass „Wir zogen uns gegenseitig aus“ und „und fing an sie zu“ im gleichen Abschnitt von Erzählungen von Männern auftreten; dass sie aber tatsächlich gemeinsam in ein und derselben Geschichte auftreten, wurde bislang nicht untersucht. Es wäre nämlich durchaus statistisch möglich, dass sie zwar durchschnittlich im gleichen Abschnitt, aber in völlig unterschiedlichen Gruppen von Erzählungen vorkommen.

Zur Ermittlung des gemeinsamen Auftretens von sprachlichen Entitäten in Texten hat die Linguistik die Kollokationsanalyse entwickelt. In ihr wird ermittelt, wie hoch die Typizität

des gemeinsamen Auftretens sprachlicher Einheiten in einem Korpus ist (Evert 2009). Die Grundlage für die Berechnung von Kollokationen ist die Bestimmung von Kollokatoren. Kollokatoren sind sprachliche Einheiten, die gemeinsam mit einer bestimmten als Basis bezeichneten sprachlichen Einheit auftreten. Je nach linguistischem Erkenntnisinteresse legt man ein Fenster von n Wörtern links und m Wörtern rechts der Basis fest und wertet dieses Fenster bei jedem Auftreten der Basis im Korpus aus. So erhält man eine Liste von sprachlichen Einheiten, die gemeinsam mit der gesuchten Basis auftreten, sowie ihrer Frequenzen. Die Liste der Kollokatoren alleine ist jedoch noch nicht aussagekräftig. Um zu überprüfen, ob eine Basis und ein Kollokator überzufällig häufig miteinander auftreten, muss untersucht werden, ob der Kollokator in anderen Kontexten nicht ähnlich häufig auftritt wie im Kontext der Basis. Hierfür teilt man das Korpus in zwei Teilkorpora. Das erste Teilkorpus besteht aus allen Basen und den dazugehörigen Textfenstern, also von den n Wörtern links und m Wörtern rechts der Basis. Das zweite Teilkorpus besteht aus allen restlichen Wörtern. Nun zählt man die Frequenzen der Kollokator-Einheit im zweiten Korpus und berechnet, wie häufig sie in beiden Korpora relativ zur Wortzahl des jeweiligen Korpus vorkommt. Zur Beantwortung der Frage, ob das gemeinsame Auftreten von Kollokator und Basis als zufällig oder überzufällig gelten kann, setzt man die relativen Frequenzen miteinander in Beziehung: als statistische Verfahren werden unterschiedliche Signifikanztests verwendet. Je nach Wahl der Parameter (Korpus, Fenstergröße, Grenzwerte, Signifikanztest und Signifikanzniveau) erhält man unterschiedliche Ergebnisse.

In unserer hier vorgestellten Analyse sind Basen und Kollokatoren Pentagramme; als Fenster wurden 30% des Textes rechts bzw. links einer Basis bestimmt, Pentagramme mussten mindestens 6x mit einer Basis gemeinsam auftreten, als Signifikanztest diente der Likelihood-Ratio-Test.

Während sich Informationen über die Kollokationen zu einer einzigen Basis leicht als Liste oder Tabelle überblicken lassen, ist die Analyse von Kollokationsnetzen, also Informationen über die Kollokationen mehrerer Basen, die selbst Kollokationen anderer Basen sein können, bei einer nennenswerten Menge an Verbindungen ohne Visualisierung nicht mehr möglich. Basis-Kollokationen-Verbindungen und Kollokationsnetze lassen sich aber beispielsweise als gerichtete Graphen visualisieren. Graphen sind – einfach gesprochen – eine Menge von Punkten („Knoten“/„vertices“/„nodes“), zwischen denen Linien („Kanten“/„edges“) verlaufen. In unserer Untersuchung repräsentieren die Punkte Pentagramme, die Kanten signifikantes gemeinsames Auftreten zweier n -Gramme in Texten. Die folgende Abbildung zeigt den Pentagramm-Kollokationsgraphen für die von Frauen verfassten Texte. Die Kollokationsgraphen wurden mit der Software GraphViz erzeugt (<http://www.graphviz.org>, Ellson et. Al. 2003).

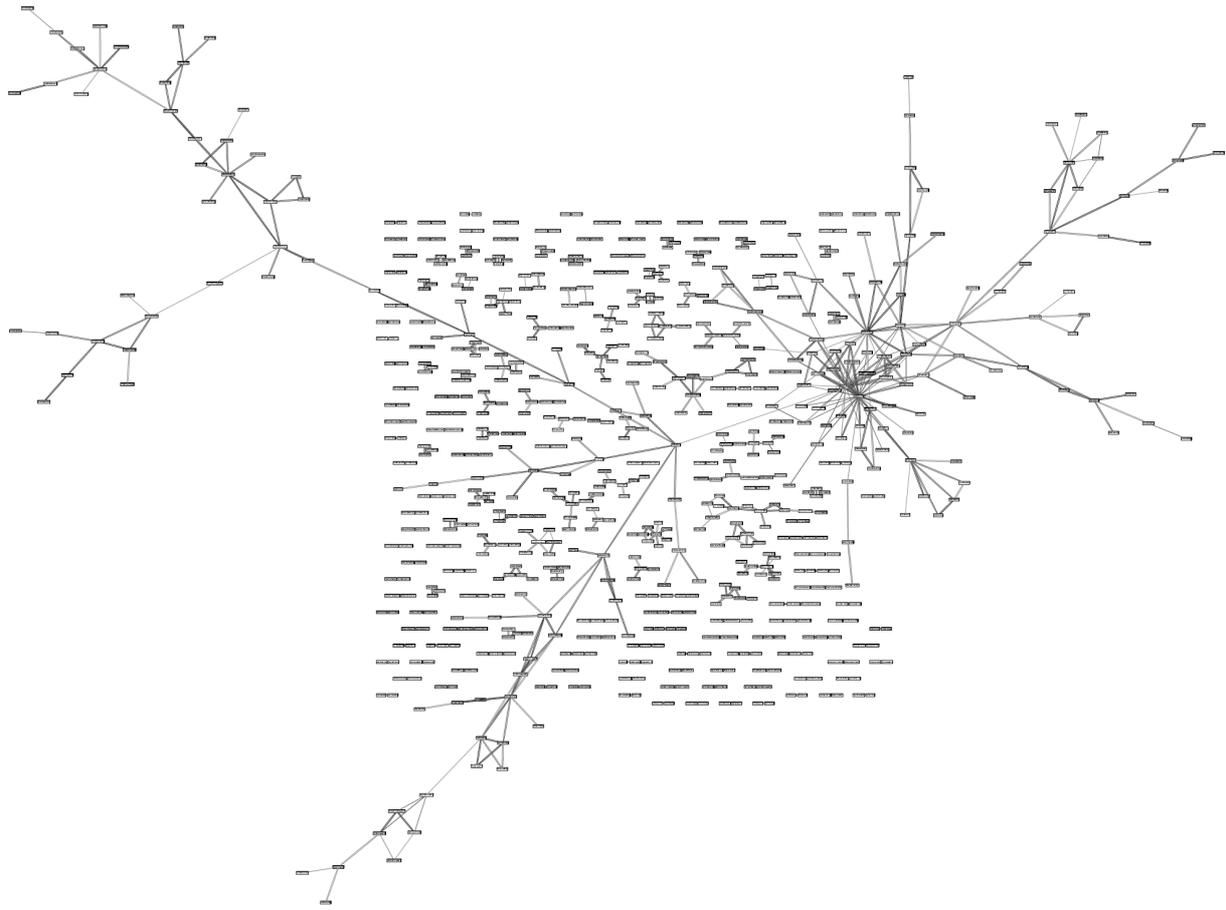


Abbildung 6: Pentagramm-Kollokationsgraph des Frauen-Korpus.

Der Graph zeigt Verdichtungsbereiche und Verknüpfungachsen. Verdichtungsbereiche entstehen dadurch, dass mehrere Knoten ähnliche Verknüpfungsmuster haben. Die meisten Verknüpfungen weisen die Pentagramme „ich mit ihm schlafen wollte“ und „und ich sagte ja Er“ auf (vgl. Abbildung 7). Sie markieren offenbar den immer wieder auftretenden Punkt in weiblichen Erzählungen, an dem der Mann nach dem Einverständnis zum Beischlaf fragt und die Erzählerin dieses Einverständnis gibt.

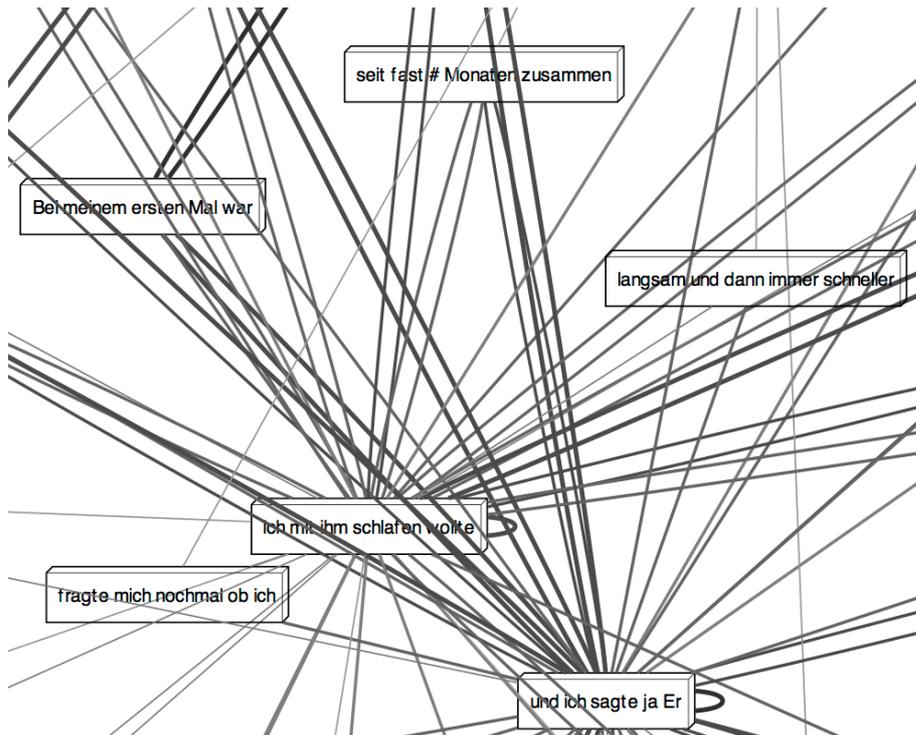


Abbildung 7: Pentagramm-Kollokationsgraph des Frauen-Korpus: Ausschnitt Verdichtungs-bereich. Je dicker die Kanten, desto höher die Signifikanz der Verbindung.

Auch wenn Analyse und Visualisierung die n-Gramm-Verknüpfung relativ gut sichtbar machen, so geht bei der Darstellung der Verknüpfungen mit Hilfe eines (nicht-hierarchischen) Force-directed-Algorithmus doch die Dimension der Linearität, die für die Texte prägend ist, verloren.

Ein Verfahren zur Rekonstruktion narrativer Muster sollte jedoch die Ergebnisse der Kollokationsanalyse mit den Erkenntnissen aus der Analyse typischer Musterpositionen kombinieren. Hierfür bietet sich eine Darstellung als hierarchischer Graph an (vgl. Abbildung 8). Der Graph, der Tetragrammverkettungen in den Geschichten von Frauen illustriert, bildet die Abfolge von Mustern in der vertikalen Dimension (von oben nach unten) ab. Mehrere voneinander unabhängige narrative Muster im gleichen Abschnitt, das heißt an ähnlichen Erzählpositionen, werden nebeneinander dargestellt. Auch in diesem Graphen sind Bereiche von geringer phraseologischer Durchdringung und Verdichtungs-bereiche sichtbar.

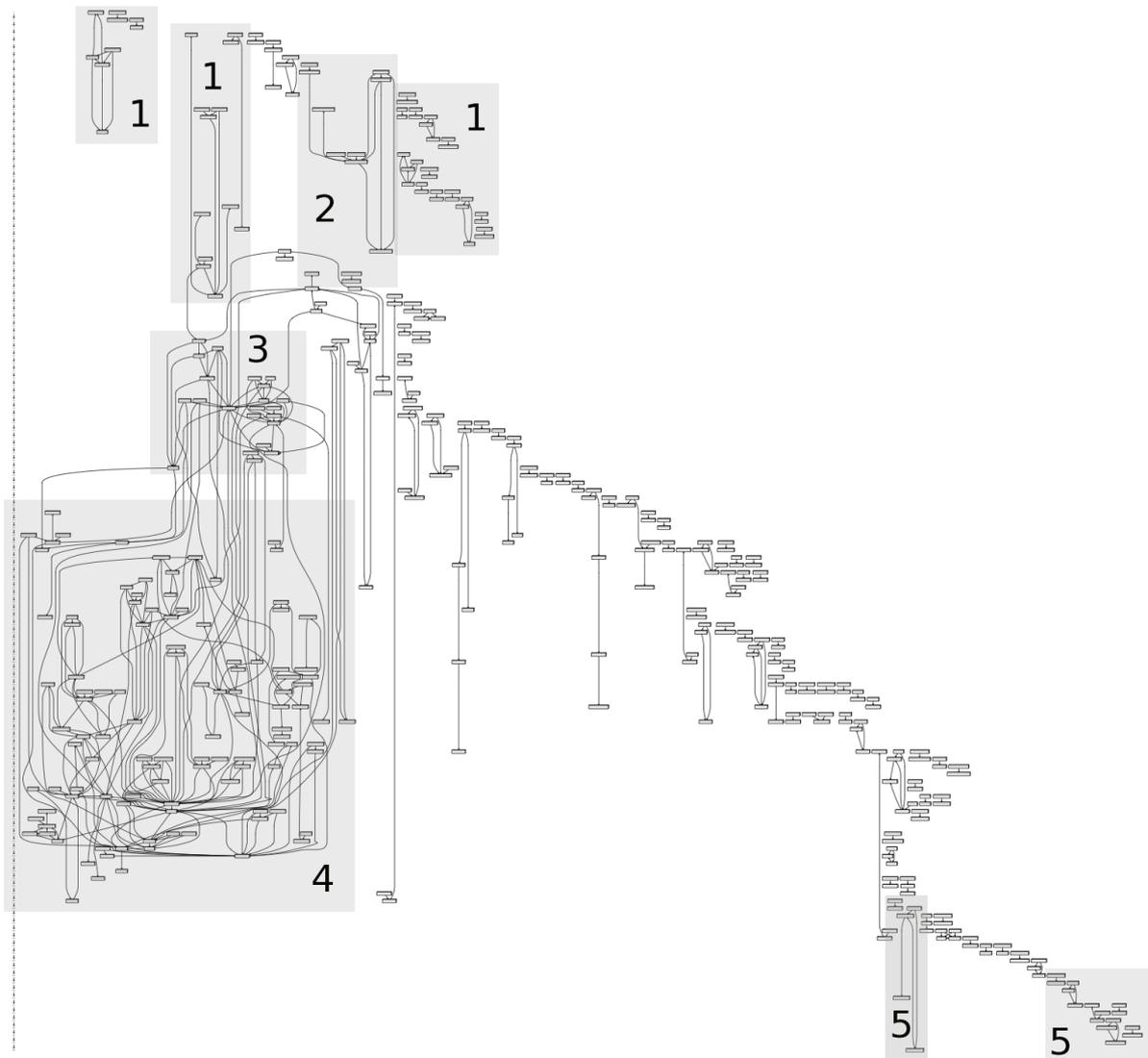


Abbildung 8: Narrationsgraph der Erzählungen von Frauen (Knoten: Tetragramme).

Einige für die Fragestellung relevante Bereiche des Graphen sind in der Abbildung grau hinterlegt und nummeriert. Sie sollen im Folgenden vorgestellt und die enthaltenen sprachlichen Muster (in Auswahl) zitiert werden:

Muster in 1 referieren auf das Alter der Hauptpersonen der Erzählung:

Mein erstes Mal hatte – ich mit meinem Freund – hatte ich mit # – erstes Mal mit # – Bei meinem ersten Mal – ersten Mal war ich – Freund und ich waren – Ich war damals # – Ich war # und – # und er war – älter als ich und – ist # Jahre älter

Muster in 2 referieren auf die Dauer der Beziehung:

Monate mit meinem – Monate mit meinem Freund – # Wochen mit meinem – mit meinem Freund zusammen – # Monate mit ihm – Monate mit ihm zusammen

Muster in 3 referieren auf die Frage des Mannes nach dem Einverständnis (Abbildung 9):

schaute mir tief in die – schaute mir lange in die – in die Augen und – fragte mich ob ich – Er fragte mich ob - mit ihm schlafen – ich es wirklich will – ich es wirklich wollte

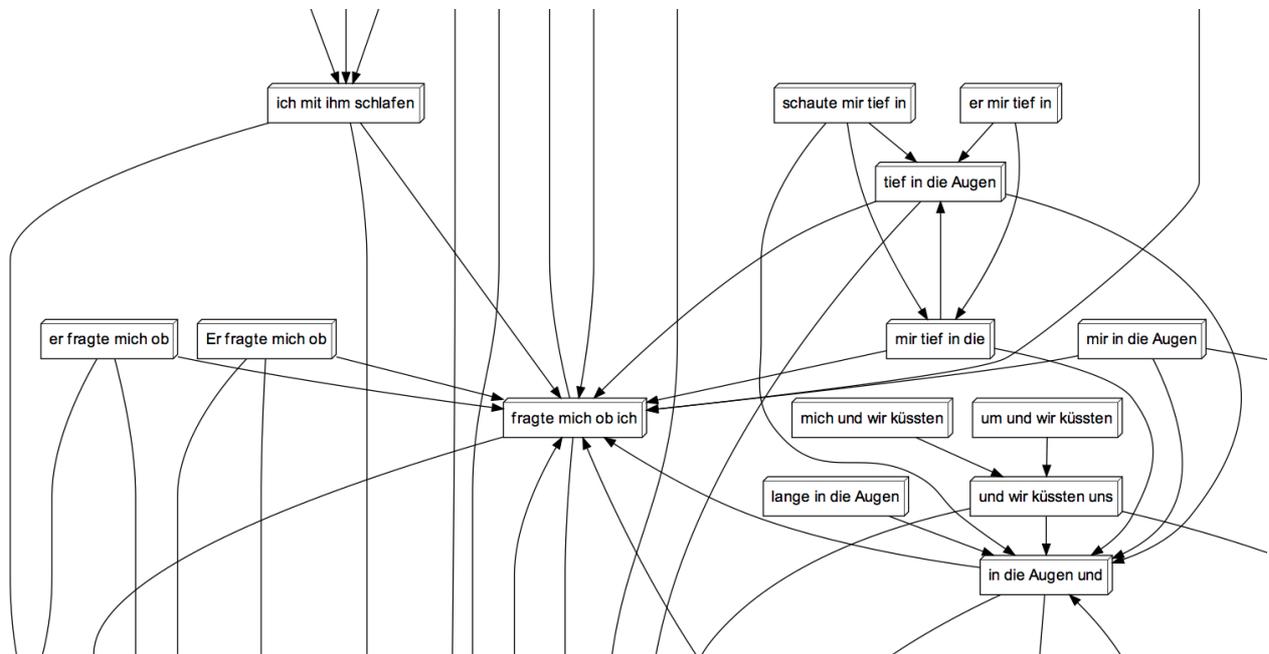


Abbildung 9: Narrationsgraph der Erzählungen von Frauen (Ausschnitt 3)

Muster in 4 referieren auf das sexuelle Geschehen, in dem vor allem der Mann aktiv ist (Abbildung 10):

Er holte ein Kondom – Kondom aus seiner Hosentasche – aus seiner Tasche – Kondom aus seinem Nachttisch – holte ein Kondom raus – und streifte es sich – zog es sich über – sich über und drang – ganz vorsichtig in mich – langsam und vorsichtig in – langsam in mich ein – drang in mich ein – in mich ein Es – Es tat überhaupt nicht – tat überhaupt nicht weh

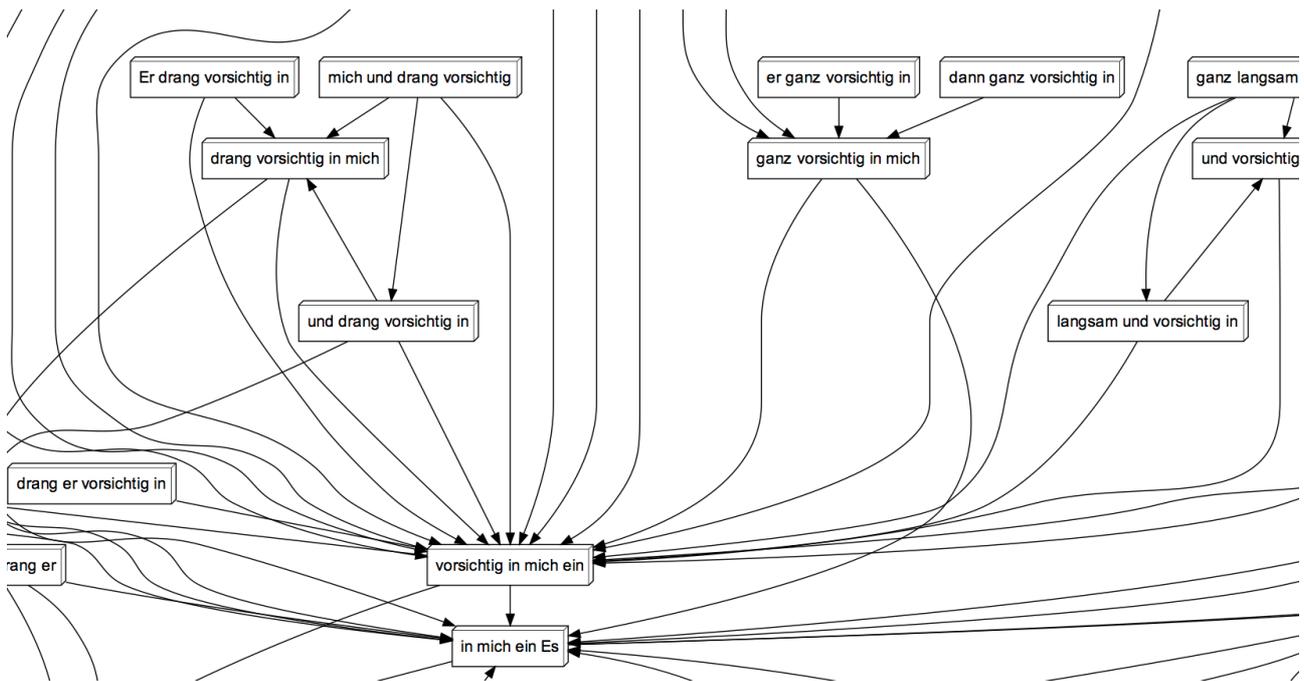


Abbildung 10: Narrationsgraph der Erzählungen von Frauen (Teilbereich aus Ausschnitt 3)

Muster in 5 referieren auf den gegenwärtigen Beziehungsstatus:

Und wir sind immer – immer noch zusammen und – immer noch mit ihm – noch mit ihm zusammen – Schatz ich liebe dich – liebe dich über alles

Die Umriss der typischen Erzählung vom Ersten Mal aus der Sicht von Frauen werden anhand dieses Verfahrens gut sichtbar. Alternative Erzählstränge, die sich teilweise parallel zu den grau hinterlegten Teilen befinden, beziehen sich auf die Aspekte Schmerz („erst tat es ein“, „ein bisschen weh aber“, „dann war es einfach“, „es einfach nur noch“), praktische Unerfahrenheit („versuchte in mich einzudringen“) und die Evaluation („Es war ein wunderschönes“, „Es war ein unbeschreibliches“, „war ein unbeschreibliches Gefühl“, „Ich hätte nie gedacht“).

Im Narrationsgraph der aus männlicher Perspektive (Abbildung 11) verfassten Geschichten soll nur auf zwei Bereiche eingegangen werden, die im Vergleich zum Narrationsgraph der Frauen von besonderem Interesse sind.

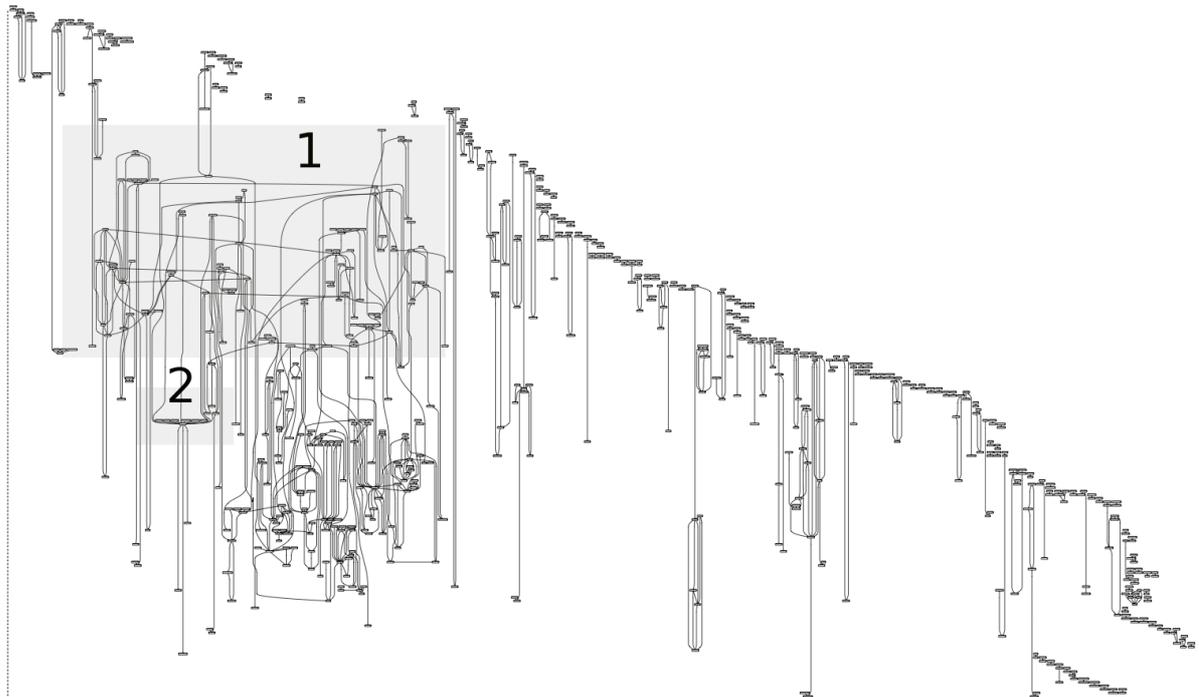


Abbildung 11: Narrationsgraph der Erzählungen von Männern (Knoten: Tetragramme)

Zum einen sind dies jene sprachlichen Muster im mit 1 bezeichneten Bereich, die auf die Einholung des Einverständnisses zum Geschlechtsakt verweisen. Hier ist es so, dass die Frage von männlicher wie weiblicher Seite kommen kann („fragte sie mich ob“, „ich fragte sie ob“). Zum anderen findet sich im mit 2 bezeichneten (Abbildung 12) Bereich eine auffällige Verbindung mehrerer n-Gramme mit der Mehrworteinheit „Sie meinte ich solle“.

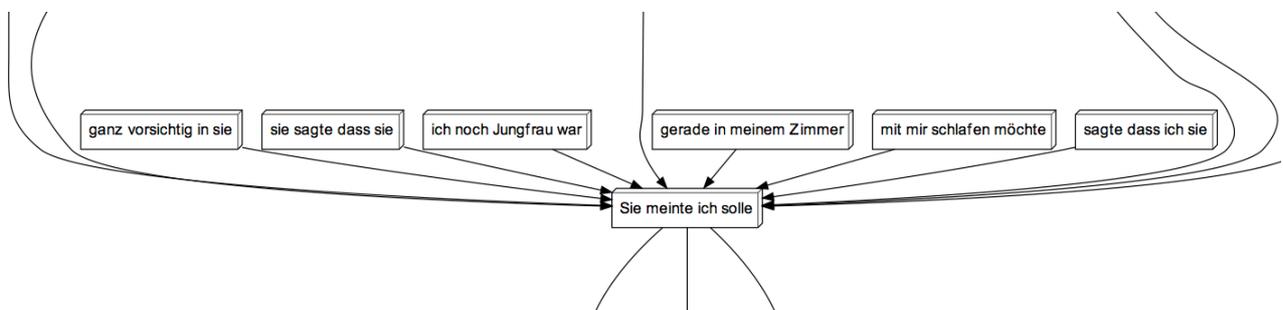


Abbildung 12: Narrationsgraph der Erzählungen von Männern (Ausschnitt 2)

Dies ist ein Hinweis darauf, dass aus der Perspektive von Männern Handlungen durch die Frauen als legitim ratifiziert werden müssen, ehe sie vollzogen werden können. Darüber hinaus sind in den Erzählungen von Männern Frauen aktiver als in ihren eigenen weiblichen Erzählungen.

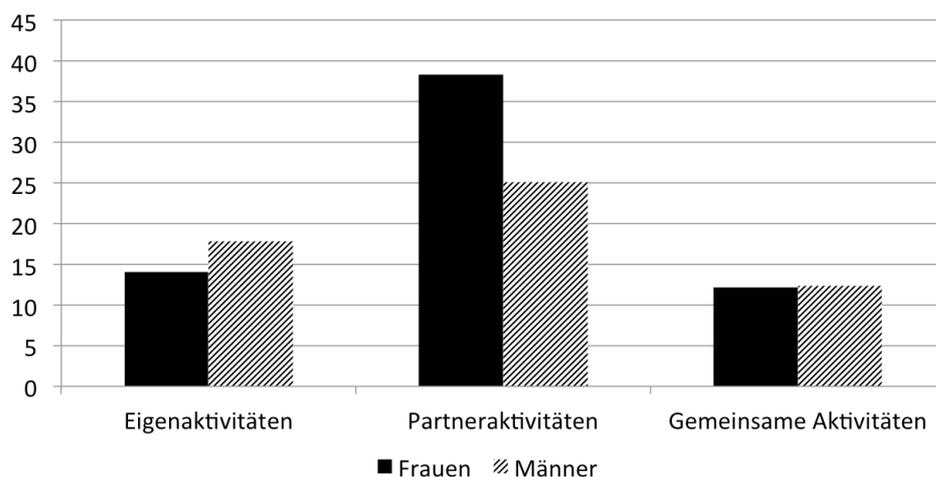


Abbildung 13: Aktivitätsanteile in Narrationsgraphen der Geschichten von Männern und Frauen

Dies bestätigt auch die Analyse der Aktivitätsanteile in beiden Graphen (Abbildung 13). Bei dieser Analyse wurden die n-Gramme daraufhin ausgewertet, ob sie auf eine Aktivität verweisen und ob diese Aktivität von der Ich-Erzählinstanz vorgenommen, vom Partner bzw. der Partnerin durchgeführt oder gemeinsam vollzogen wurde. Auffällig ist, dass in beiden Graphen die Partneraktivitäten dominieren, in denen der Frauen freilich sehr viel deutlicher als bei den Männern.

5. Fazit

Die Analysen zeigen, dass Geschichten vom Ersten Mal von Männern und Frauen recht ähnlich erzählt werden und zwar nicht nur im Hinblick auf das sexuelle Geschehen, sondern auch im Hinblick auf die verbalen Handlungen, die ihm vorausgehen und es begleiten. Zentraler Bestandteil typischer Erzählungen beider Geschlechter ist die verbale Verständigung über die Bereitschaft zum Geschlechtsakt und die explizite Gabe des Einverständnisses durch die Frau. Das von der Paarsoziologie als Schwellen-Wendepunkt bezeichnete Erste Mal wird also als eine durch Einverständnis der Frau legitimierte Handlungsfolge erzählt, in der der Mann mehr Handlungsmacht hat als die Frau.

In methodischer Hinsicht hat sich gezeigt, dass die beschriebenen Verfahren geeignet sind, narrative Muster an der Textoberfläche zu identifizieren. Als für die Analyse relevante Kategorien haben sich herauskristallisiert:

- Inhalte von n-Grammen
- n-Gramm-Dichte (wie groß ist die Variation, wo finden sich gruppen-, institutionen- oder kommunikationsbereichsspezifische Verdichtungen)
- n-Gramm-Positionen und Positionsdifferenzen
- spezifische Verkettungsmuster der n-Gramme

Auch wenn uns die Ergebnisse der exemplarischen Analyse vielversprechend scheinen, so sind doch einige Schwächen sichtbar geworden:

- Um zu einer übersichtlicheren Darstellung zu gelangen, wäre es wohl sinnvoll, semantisch oder an der sprachlichen Oberfläche sehr ähnliche n-Gramme auf Typen zurückzuführen.
- Bei der Berechnung der n-Gramme könnte es sinnvoll sein, statt ausschließlich auf Textwörter zu bauen, auch semantisch und / oder funktional bestimmte Lexemklassen als Bestandteile von n-Grammen zuzulassen. Auf diese Weise etwa könnten n-Gramme wie „es war total schön“ und „es war unglaublich schön“ auf das Muster „es war INTENSIFIER schön“ oder „es war INTENSIFIER EVALUATIVES_ADJEKTIV“ zurückgeführt werden.
- Statt wie im vorliegenden Fall von zwei Subkorpora auszugehen, könnte es je nach Fragestellung sinnvoll sein, durch strukturentdeckende Verfahren wie die Clusteranalyse ähnliche Geschichten und damit typische Narrative datengeleitet zu gruppieren. Einige Proberechnungen haben jedoch gezeigt, dass für ein erfolgreiches Clustern der Texte die Länge der Geschichten möglichst ähnlich sein sollte.

Trotz der zweifellos bestehenden Desiderate sind wir dennoch überzeugt, dass die datengeleitete Analyse narrativer Muster das Methodenrepertoire der Diskursanalyse um eine wertvolle Dimension erweitert.

Literatur

- Banerjee, Satanjeev und Ted Pedersen (2003), „The Design, Implementation, and Use of the Ngram Statistic Package“. In: *Proceedings of the Fourth International Conference on Intelligent Text Processing and Computational Linguistics*. Mexico City.
- Belica, Cyril und Kathrin Steyer (2006), „Korpusanalytische Zugänge zu sprachlichem Usus“. *AUC (Acta Universitatis Carolinae), Germanistica Pragensia*, XX.
- Bruner, Jerome (1990), *Acts of Meaning*. Cambridge: Harvard University Press.
- Bruner, Jerome (1991), „The Narrative Construction of Reality“. *Critical Inquiry* 18: 1-21.
- Bubenhofer, Noah (2009), *Sprachgebrauchsmuster. Korpuslinguistik als Methode der Diskurs- und Kulturanalyse*. Berlin, New York: de Gruyter.
- Bubenhofer, Noah und Joachim Scharloth (2013), „Korpuslinguistische Diskursanalyse: Der Nutzen empirisch-quantitativer Verfahren“. In: Ingo H. Warnke, Ulrike Meinhof und Martin Reisigl (Hrsg.), *Diskurslinguistik im Spannungsfeld von Deskription und Kritik*. Diskursmuster – Discourse Patterns. Berlin: Akademie-Verlag: 147-168.
- Ellson, John et al. (2003): „Graphviz and Dynagraph – Static and Dynamic Graph Drawing Tools.“ In: Michael Jünger, Petra Mutzel (Hrsg.): *Graph Drawing Software*. Heidelberg: Springer-Verlag: S. 127-148.
- Evert, Stefan (2009), „Corpora and Collocations“. In: Anke Lüdeling und Merja Kytö (Hrsg.), *Corpus Linguistics: An International Handbook*. Vol. 2: 1212-1248. (=Handbooks of linguistics and communication-science; 29.2)
- Feilke, Helmuth (2000), „Die pragmatische Wende in der Textlinguistik“. In: Klaus Brinker (Hrsg.), *Text- und Gesprächslinguistik/Linguistics of Text and Conversation*. Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science. Berlin/New York: de Gruyter: 64-82.
- Feilke, Helmuth und Angelika Linke (Hrsg.) (2009), *Oberfläche und Performanz. Untersuchungen zur Sprache als dynamische Gestalt*. Berlin, New York: de Gruyter.
- Fuchs-Heinritz, Werner (2009), *Biographische Forschung: Eine Einführung in Praxis und Methoden*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Gloy, Klaus (1975), *Sprachnormen I. Linguistische und soziologische Analysen*. Stuttgart: Frommann.
- Illouz, Eva (2006), „Leiden, emotionale Felder und emotionales Kapital“. In: Dies.: *Gefühle in Zeiten des Kapitalismus*. Frankfurt am Main: Suhrkamp: 65-96.
- Keller, Rainer (2011), *Wissenssoziologische Diskursanalyse: Grundlegung eines Forschungsprogramms*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Klausnitzer, Ralf (2008), *Literatur und Wissen: Zugänge, Modelle, Analysen*. Berlin: de Gruyter.

- Koschorke, Albrecht (2011), „Wie Bürgerkriege erzählt werden. Feldtheoretische Überlegungen zur Konfliktsemantik“. In: Sabrina Ferhadbegović und Brigitte Weiffen (Hrsg.), *Bürgerkriege erzählen: zum Verlauf unziviler Konflikte*. Konstanz: Konstanz University Press: 35-54.
- Koschorke, Albrecht (2012), *Wahrheit und Erfindung: Grundzüge einer allgemeinen Erzähltheorie*. Frankfurt am Mai: Fischer.
- Lenz, Karl (2009), *Soziologie der Zweierbeziehung. Eine Einführung*. 4. Auflage. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Lucius-Hoene, Gabriele (2010), „Narrative Identitätsarbeit im Interview“. In: Birgit Griese (Hrsg.), *Subjekt - Identität – Person? Reflexionen zur Biographieforschung*. Wiesbaden: VS Verlag für Sozialwissenschaften: 149-171.
- Lucius-Hoene, Gabriele und Arnulf Deppermann (2002), *Rekonstruktion narrativer Identität: ein Arbeitsbuch zur Analyse narrativer Interviews*. Opladen: Leske & Budrich.
- Maier, Maja S. (2007), *Paaridentitäten. Biografische Rekonstruktionen homosexueller und heterosexueller Paarbeziehungen*. Weinheim: Juventa.
- Meuter, Norbert (2004), „Geschichten erzählen, Geschichten analysieren. Das narrativistische Paradigma in den Kulturwissenschaften“. In: Friedrich Jaeger (Hrsg.), *Handbuch der Kulturwissenschaften*. Stuttgart: Metzler: 140-155.
- Nünning, Ansgar (2013), „Wie Erzählungen Kulturen erzeugen: Prämissen, Konzepte und Perspektiven für eine kulturwissenschaftliche Narratologie“. In: Alexandra Strohmaier (Hrsg.), *Kultur-Wissen-Narration: Perspektiven transdisziplinärer Erzählforschung für die Kulturwissenschaften*. Bielefeld: transcript: 15-54.
- Perkuhn, Rainer u.a. (2005), „Korpustechnologie am Institut für Deutsche Sprache“. In: Johannes Schwitalla und Werner Wegstein (Hrsg.), *Korpuslinguistik deutsch: synchron – diachron – kontrastiv. Würzburger Kolloquium 2003*. Tübingen: Niemeyer: 57-70.
- Perkuhn, Rainer & Cyril Belica (2006), „Korpuslinguistik – Das unbekannte Wesen. Oder Mythen über Korpora und Korpuslinguistik.“ *Sprachreport 22, 1*: 2-8.
- Rabelhofer, Bettina (2013), „Erzählen in der Psychoanalyse“. In: Alexandra Strohmaier (Hrsg.), *Kultur - Wissen - Narration: Perspektiven transdisziplinärer Erzählforschung für die Kulturwissenschaften*. Bielefeld: transcript: 343-358.
- Ricœur, Paul (2005), „Narrative Identität“. In: Ders.: *Vom Text zur Person: hermeneutische Aufsätze (1970-1999)*. Hamburg: Meiner: 209-227.
- Scharloth, Joachim und Noah Bubenhofer (2011), „Datengeleitete Korpuspragmatik: Korpusvergleich als Methode der Stilanalyse“. In: Ekkehard Felder, Marcus Müller und Friedemann Vogel (Hrsg.), *Korpuspragmatik. Thematische Korpora als Basis diskurslinguistischer Analysen von Texten und Gesprächen*. Berlin, New York: de Gruyter: 195-230.
- Scharloth, Joachim, David Eugster und Noah Bubenhofer (2013), „Das Wuchern der Rhizome. Linguistische Diskursanalyse und Data-driven Turn“. In: Dietrich Busse und Wolfgang Teubert (Hrsg.), *Linguistische Diskursanalyse. Neue Perspektiven*. Wiesbaden: Springer VS: 345-380.
- Sinclair, John (1991), *Corpus, Concordance, Collocation*, Oxford: Oxford University Press.
- Shore, Bradd (1996), *Culture in mind: cognition, culture, and the problem of meaning*. New York: Oxford University Press.
- Steyer, Kathrin (2004), *Kookkurrenz. Korpusmethodik, linguistisches Modell, lexikografische Perspektiven*. In: Kathrin Steyer (Hrsg.), *Wortverbindungen – mehr oder weniger fest*. Institut für Deutsche Sprache. Jahrbuch 2003. Berlin/ New York: de Gruyter: 87-116.
- Tognini-Bonelli, Elena (2001), *Corpus Linguistics at Work*, Amsterdam: Benjamins.
- Tophinke, Doris (2009), „Wirklichkeitserzählungen im Internet“. In: Christian Klein und Matias Martinez (Hrsg.), *Wirklichkeitserzählungen: Felder, Formen und Funktionen nicht-literarischen Erzählens*. Stuttgart: Metzler: 245-274.
- Viehöfer, Willy (2012), „Narrative Diskurse, personale Identitäten und die ästhetisch-plastische Chirurgie“. In: Rainer Keller u.a. (Hrsg.), *Diskurs - Macht - Subjekt*. Wiesbaden: VS Verlag für Sozialwissenschaften: 191-228.
- White, Hayden (1990), *Die Bedeutung der Form: Erzählstrukturen in der Geschichtsschreibung*. Frankfurt a.M.: Fischer.