



TU Clausthal

Diplomarbeit

Unterstützung kollaborativer Schreibprozesse durch ausgewählte
Verfahren der Computerlinguistik

vorgelegt von

Sebastian Groß
Wirtschaftsinformatik
Matrikelnummer 324650

Clausthal-Zellerfeld, den 22. Juni 2011

Erstgutachter

Prof. Dr. Niels Pinkwart
Institut für Informatik

Zweitgutachter

Prof. Dr. Jörg P. Müller
Institut für Informatik

Zusammenfassung

Die Arbeit untersuchte die Frage, ob Verfahren der Computerlinguistik einen Beitrag zur Verbesserung kollaborativer Schreibprozesse leisten können. Die Antwort sollte durch den praktischen Einsatz ausgewählter Verfahren in einem Mehrbenutzereditor erfolgen. Aus diesem Grund lag der Schwerpunkt auf der praktischen Umsetzung der theoretischen Konzepte. Diese wurden unter morphologischen, syntaktischen und semantischen Aspekten betrachtet und hinsichtlich einer konkreten Problemstellung eingeordnet.

Die Durchführung mehrerer Laborstudien zeigte Schwächen aber auch Potenziale für zukünftige Entwicklungen auf. Die Arbeit konnte aufgrund ihres explorativen Charakters und der Vielschichtigkeit der Problemstellung keine abschließende Bewertung vornehmen. Im Bereich des kollaborativen Schreiben kann sie allerdings zumindest eine Motivation sein, die Verbesserung von Schreibprozessen unter linguistischen Gesichtspunkten noch tiefgreifender zu untersuchen.

Schlüsselwörter

Kollaboratives Schreiben, CSCW, Automatische Indexierung, Computerlinguistik

Inhaltsverzeichnis

Abkürzungsverzeichnis	II
Abbildungsverzeichnis	IV
Tabellenverzeichnis	IV
1 Einleitung	1
1.1 Problemstellung	2
1.2 Zielsetzung	3
1.3 Erwartete Ergebnisse	4
2 Grundlagen der Indexierung	5
2.1 Indexierung und Information Retrieval	5
2.2 Verfahren der automatischen Indexierung	5
2.3 Kriterien der Indexierungsgüte	9
3 Computerlinguistische Verfahren	10
3.1 Fehlertoleranz	10
3.2 Morphologische Analyse	12
3.3 Syntaktische Analyse	17
3.4 Semantische Analyse	20
4 Konzeption und Implementierung des Mehrbenutzereditors	22
4.1 Ausgangssituation	22
4.2 System- und Datenstrukturbeschreibung	23
4.3 Systemfunktionen	26
4.4 Kritische Systembetrachtung	39
5 Evaluierung des Mehrbenutzereditors	40
5.1 Ausgangssituation	40

5.2	Organisation	41
5.3	1. Laborstudie	45
5.4	2. Laborstudie	55
5.5	3. Laborstudie	65
5.6	Gesamtfazit und Ausblick	71
Literaturverzeichnis		73
Anhang		75
A	Hörspiel I (2. Studie)	76
B	Hörspiel II (2. Studie)	79
C	Hörspiel III (3. Studie)	83
D	DVD-ROM	87

API Application Programming Interface, Programmierschnittstelle
CSCW Computer Supported Cooperative Work, computerunterstützte Gruppenarbeit
ERM Entity-Relationship-Model, Gegenstand-Beziehung-Modell
FIFO First In – First Out
GUI Graphical User Interface, grafische Benutzeroberfläche
HTTP Hypertext Transfer Protocol, Hypertext-Übertragungsprotokoll
IR Information Retrieval, Informationsabruf
JMS Java Messaging Service
NLP Natural Language Processing, Verarbeitung natürlicher Sprachen
POS Part-Of-Speech
RMI Remote Method Invocation, Aufruf entfernter Methoden
RST Rhetorical Structure Theory, Theorie rhetorischer Strukturen
STTS Stuttgart-Tübingen-Tagset
Stud.IP Studienbegleitender Internetsupport von Präsenzlehre
XML Extensible Markup Language, erweiterbare Auszeichnungssprache

Abbildungsverzeichnis

2.1	Übersicht Indexierungsverfahren	6
3.1	Lexikalische Ordnung GermaNet	21
4.1	ERM: Objekte in der Datenhaltungsschicht	24
4.2	Klassendiagramm: Term, Word, Morph	24
4.3	Übersicht RMI	25
4.4	Screenshot: Serverparameter	26
4.5	Aktivitätsdiagramm: Lemmatisierung	31
4.6	Beispielsatz mit POS-Tags	33
4.7	Ausschnitt indexierter Sätze	37
4.8	Screenshot: Client	38
5.1	Indexierte Wörter und Informationsgehalt (1. Studie)	49
5.2	SUS-Auswertung (1. Studie)	49
5.3	Anzahl angezeigter Ergebnisse und Ergebnisaufrufe (1. Studie)	50
5.4	Zeitprofil Teilnehmer 01 und 10 (1. Studie)	51
5.5	Fehlerquotienten (1. Studie)	52
5.6	Anzahl geschriebener Wörter (2. Studie)	58
5.7	SUS-Auswertung (2. Studie)	58
5.8	Anzahl angezeigter Ergebnisse und Ergebnisaufrufe (2. Studie)	59
5.9	Anzahl Auslassungen und Inkonsistenzen (2. Studie)	60
5.10	Fehlerquotienten (2. Studie)	61
5.11	SUS-Auswertung (3. Studie)	67
5.12	Fehlerquotienten und Anzahl Auslassungen und Inkonsistenzen (3. Studie)	68
5.13	Zeitprofil Teilnehmer 19 (3. Studie)	68

Tabellenverzeichnis

3.1	Fehlerklassen im Information Retrieval	10
3.2	Beispiel Soundex und Phonix	11
3.3	Ersetzungstabelle Kölner Phonetik	12
3.4	Kompositionstypen	16
4.1	Einstellbare Zugriffsrechte	27
4.2	Ausschnitt Ereignisprotokoll	28
4.3	Ergebnisse Lemmatisierung	32
5.1	Themenauswahl und Perspektiven (1. Studie)	46
5.2	Variation (1. Studie)	46
5.3	Zuordnung Benutzerkennung / Gruppe (1. Studie)	47
5.4	Hörspiele (2. Studie)	56
5.5	Variation (2. Studie)	56
5.6	Zuordnung Benutzerkennung / Gruppe (2. Studie)	57
5.7	Vergleichsergebnisse ausgewählter Sätze (2. Studie)	62

1 Einleitung

Computer Supported Cooperative Work (CSCW) ist ein inter- bzw. multidisziplinäres Forschungsgebiet, das von der (Angewandten) Informatik, der Wirtschaftsinformatik und auch der (Arbeits-)Psychologie geprägt und beeinflusst wird. Zentraler Gedanke der wissenschaftlichen Untersuchung ist dabei der Einsatz von computergestützten Technologien, um die Zusammenarbeit mehrerer Teilnehmer zu unterstützen. Dazu wird die soziale Interaktion analysiert und Konzepte entwickelt und evaluiert, um eine Steigerung der Effektivität, der Effizienz und des Komforts zu erzielen.¹

Kollaboration kann als koordinierte Kooperation verstanden werden, um ein gemeinsames Ziel durch Arbeitsteilung zu erreichen. Im kollaborativen Schreiben, einer klassischen Disziplin von CSCW, schließen sich zwei oder mehr Autoren zusammen, um in verteilter Arbeit ein Dokument zu verfassen. Das Ergebnis dieses Vorgangs ist im optimalen Fall ein syntaktisch, stilistisch, semantisch und inhaltlich konsistentes Dokument, das sich nicht von dem eines einzelnen Autors unterscheidet. Dem Vorgang selbst liegt der Ablauf aufeinander aufbauender Prozesse der Planung, des Schreibens und der Kontrolle² zugrunde.

In den Prozess des Schreibens fließen die Kreativität bzw. das Wissen des Autors ein. Dieses wird ergänzt durch gedruckte, geschriebene oder digitale Informationen, die in traditionellen oder digitalen Bibliotheken organisiert sein können. Die Verteilung der Informationen und das intellektuelle Potenzial, diese zu erfassen, zu verarbeiten und in eigene Arbeit einfließen zu lassen, sind individuell verschieden. Ziel und bedeutender Vorteil der Kooperation und insbesondere der Kollaboration ist, die Fähigkeiten der Individuen zu kombinieren und dadurch die Leistung, die ein Einzelner erbringen würde, zu steigern. Dieser Vorteil muss nicht auf die Dauer der Zusammenarbeit beschränkt sein, vielmehr kann der Einzelne nachhaltig von der Kollaboration profitieren und dazulernen. Zusam-

¹Vgl. Gross und Koch, 2007, S. IX-5.

²Olivier und Pinkwart, 2009.

menarbeit und Verteilung von Aufgaben birgt aber auch die Gefahr, dass das Ergebnis durch mangelnde Planung, fehlende Kontrolle oder soziale Faktoren negativ beeinflusst wird.

Die automatische Analyse von Dokumenten kann eine wichtige Unterstützung des Benutzers darstellen und bedeuten, dass Mängel gar nicht erst entstehen oder erkannt und korrigiert werden können. Eine Funktion zur automatischen Rechtschreib- und Grammatikprüfung in klassischen Textverarbeitungsprogrammen wie Microsoft Word³ oder OpenOffice⁴ ist heute für viele Benutzer nicht mehr wegzudenken.

Im kollaborativen Schreiben lag der Fokus der Forschung, Dokumente mehrerer Autoren automatisch zu überprüfen, bislang weniger auf der inhaltlichen Konsistenz. Praktische Ansätze, die Konsistenz zu gewährleisten, bezogen sich vielmehr auf

- die algorithmische Auflösung von syntaktischen Konflikten wie z. B. in XML Dokumenten,⁵
- die Überprüfung auf stilistische Konsistenz anhand statistischer Verfahren⁶ oder
- die Analyse der semantischen Konsistenz mittels der Rhetorical Structure Theory (RST).⁷

Die Überprüfung der inhaltlichen Konsistenz fällt in die Kontrollphase und setzt einen manuellen Vergleich des Dokuments durch den Benutzer voraus. Dieser muss dazu das gesamte Dokument intellektuell erfassen, Inkonsistenzen zwischen Textpassagen identifizieren und korrigieren und die Korrekturen wiederum in der weiteren Überprüfung berücksichtigen.

1.1 Problemstellung

Naive Ansätze, Dokumente automatisch abzugleichen, scheitern oftmals an den Eigenschaften natürlicher Sprachen. Bereits die automatische Überprüfung auf Wortebene, also der Vergleich jedes einzelnen Wortes mit jedem anderen Wort im Dokument, wird

³<http://office.microsoft.com/de-de/word/>.

⁴<http://de.openoffice.org/>.

⁵Skaf-Molli, Molli und Rahhal, 2009.

⁶Glover und Hirst, 1995.

⁷Rahhal u. a., 2007.

durch Konjugation⁸, Deklination⁹ und Komparation¹⁰ erheblich gestört. Gleiche Wörter, die sich nur in Form bzw. Schreibweise unterscheiden, werden nicht als solche erkannt. Wortkompositionen und Textphrasen bestehend aus mehreren Wörtern werden nicht oder nicht angemessen berücksichtigt. Der Ansatz, Wörter nach Übereinstimmungen in Zeichenketten zu überprüfen, vernachlässigt die Tatsache, dass Wörter zwar unterschiedlich in ihrer Schreibweise sein, aber die gleiche oder eine ähnliche Bedeutung haben können. Umgekehrt können sich gleiche Wörter in unterschiedlichen grammatischen Kontexten, die deren Bedeutung beeinflussen, wiederfinden. Daher kann für eine umfassende Analyse auch die grammatische Struktur von Bedeutung sein.

Dokumente müssen zunächst inhaltlich erschlossen werden. Dazu gehört, dass grammatische Strukturen identifiziert und die strukturellen und semantischen Eigenschaften von Wörtern beachtet werden. Erst dann können Automatismen greifen, die den Autor bei der manuellen Überprüfung des Dokuments unterstützen.

1.2 Zielsetzung

Im theoretischen Teil der Arbeit werden zunächst die Grundlagen der Indexierung besprochen. An diese schließen die computerlinguistischen Verfahren unter besonderer Berücksichtigung der deutschen Sprache an. Es werden die Möglichkeiten und Schwächen der Verfahren analysiert und dahin überprüft, ob sie zur Unterstützung eines Autors sinnvoll und effektiv eingesetzt werden können.

Im praktischen Teil der Arbeit wird der Prototyp eines Mehrbenutzereditors entwickelt, in dem ausgewählte Konzepte des theoretischen Teils Anwendung finden. Das Ziel ist die Unterstützung des Autors während des Schreibens eines Dokuments. Die Dokumente werden dazu fortlaufend analysiert und auf Satzebene miteinander verglichen. Die Verfahren werden dazu unstrukturierte bzw. schwach strukturierte Dokumente¹¹ in deutscher Sprache analysieren können. Die Vergleichsergebnisse werden entsprechend ihrer Relevanz gewichtet und dem Benutzer angezeigt. Der Benutzer wird auf Grundlage der Vergleichsergebnisse aktiv Änderungen an seinem Dokument vornehmen können, die wiederum in der Analyse berücksichtigt werden.

⁸Beugung des Verbs (Bsp.: er geht, er ging, er ist gegangen).

⁹Beugung des Substantivs (Bsp.: der Baum, den Baum, dem Baum, des Baumes).

¹⁰Steigerung des Adjektivs (Bsp.: schnell, schneller, am schnellsten).

¹¹Lewandowski, 2005, S. 16.

Der Anwendung wird eine Client-Server-Architektur zugrunde liegen, die den Einsatz in einer verteilten Umgebung ermöglicht. Der Zugriff auf die Dokumente, die zentral gespeichert werden, wird asynchron erfolgen; ein gleichzeitiges Schreiben an einem Dokument durch mehrere Benutzer wird nicht möglich sein. Es wird allerdings die Möglichkeit bestehen, auf Dokumente lesend zuzugreifen und den Schreibprozess anderer Benutzer mitzuverfolgen. In abschließenden Studien werden die im Prototyp eingesetzten Verfahren im praktischen Einsatz getestet.

1.3 Erwartete Ergebnisse

Als Ergebnis der Arbeit wird eine Übersicht über computerlinguistische Verfahren im Kontext der automatischen Indexierung und dem automatischen Vergleich von Dokumenten vorliegen. Es wird versucht, die Frage zu beantworten, inwiefern diese Konzepte genutzt werden können, um den Benutzer während des Schreibens zu unterstützen. Diese Erkenntnisse fließen direkt in die Entwicklung des Prototyps ein. Die im theoretischen Teil analysierten Potenziale der Verfahren werden in der praktischen Anwendung des Prototyps überprüft und ggf. Schwachstellen bzw. Verbesserungsmöglichkeiten aufgezeigt.

In der Bewertung der abschließenden Studien wird von zentraler Bedeutung sein, ob und in welcher Ausprägung die Leistung des Autors verbessert werden kann. Dabei beschränkt sich die Untersuchung nicht auf inhaltliche Konsistenz, auch Rechtschreibung, Grammatik und sprachlicher Ausdruck werden berücksichtigt. Zudem soll geklärt werden, wie sich die Ergebnisse in die bisherigen wissenschaftlichen Erkenntnisse im Bereich des kollaborativen Schreibens einfügen.

2 Grundlagen der Indexierung

2.1 Indexierung und Information Retrieval

Die Indexierung kann als Teilprozess des Information Retrieval (IR) verstanden werden. Das Ziel von IR-Systemen ist, Informationen aufzubereiten und bei Bedarf zur Verfügung zu stellen. Die Güte eines IR-Systems misst sich dabei an der Präzision und der Vollständigkeit der Informationen, die durch Suchstrategien und -techniken abgerufen werden.¹²

Das Ziel der Indexierung ist, Deskriptoren zu extrahieren, die das Dokument inhaltlich repräsentieren. Es wird unterschieden zwischen der manuellen bzw. intellektuellen und der automatischen Indexierung. Während bei der intellektuellen Indexierung der Text manuell von einem (menschlichen) Indexierer analysiert und indexiert wird, läuft die automatische Indexierung computer-gestützt ab. Die halbautomatische Indexierung nutzt im ersten Schritt die computer-gestützte Indexierung, im zweiten Schritt werden die extrahierten Deskriptoren durch einen Indexierer geprüft und ggf. korrigiert.

Die Indexierung kann zwei verschiedene Ansätze verfolgen. Im dokumentenorientierten Ansatz werden die Deskriptoren aus dem Volltext, einer Zusammenfassung oder dem Titel des Dokuments gewonnen. Der begriffsorientierte Ansatz nutzt ein Indexierungsvokabular, auf das die bedeutungstragenden Wörter eines Textes abgebildet werden.¹³

2.2 Verfahren der automatischen Indexierung

Abb. 2.1 zeigt eine mögliche Einteilung der Indexierungsverfahren, die im Folgenden kurz vorgestellt werden. In [Nohr, 2001, S. 23] wird zusätzlich die *Einfache Stichwortex-*

¹²Vgl. Nohr, 2001, S. 13.

¹³Vgl. Nohr, 2001, S. 21-24.

traktion / Volltextinvertierung als Indexierungsverfahren genannt, das zwar vollautomatisch arbeitet, jedoch keine Auswahl trifft und keine Terme bearbeitet.

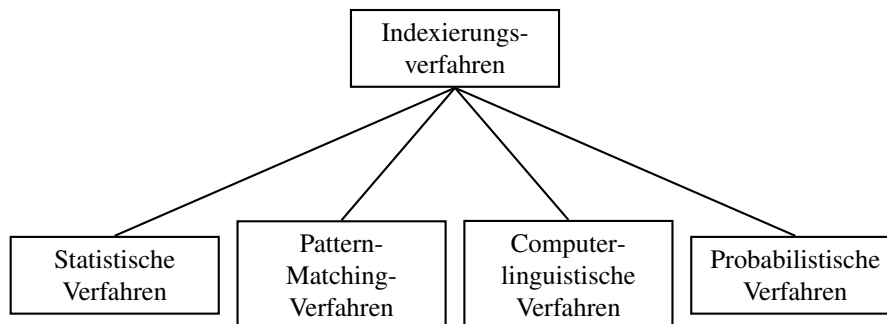


Abb. 2.1: Übersicht der Indexierungsverfahren nach Nohr, 2001, S. 23-25 und Stock, 2008, S. 366

2.2.1 Statistische Verfahren

Statistische Verfahren stellen einen Zusammenhang zwischen der Bedeutung von Wörtern für ein Dokument und der Häufigkeit ihres Auftretens her. Es wird von zwei Annahmen¹⁴ ausgegangen:

- Nicht alle Wörter eines Textes sind geeignet, um das Dokument zu repräsentieren, folglich muss eine Auswahl getroffen werden.
- Nicht alle Wörter haben hinsichtlich ihrer inhaltlichen Bedeutung die gleiche Wertigkeit, folglich müssen sie gewichtet werden.

Die Häufigkeit des Auftretens (Termfrequenz) eines Wortes in einem Text kann als Indiz für die Bedeutung angesehen werden. Da Texte sich in der Länge unterscheiden, wird die Worthäufigkeit in Relation zu der Gesamtanzahl der Wörter in einem Text gesetzt.

$$TF_{td} = \frac{FREQ_{td}}{GESAMT_{td}}$$

$FREQ_{td}$ = Frequenz eines Termes im Dokument

$GESAMT_{td}$ = Gesamtzahl der Terme im Dokument

Um die Spannweite der Termfrequenzen zu begrenzen, kann eine logarithmische Berechnungsbasis verwendet werden.¹⁵

¹⁴Vgl. Nohr, 2001, S. 26.

¹⁵Hamran, 1986.

$$TF_{td} = \frac{\log_2(FREQ_{td}+1)}{\log_2 GESAMT_{td}}$$

$FREQ_{td}$ = Frequenz eines Terms im Dokument

$GESAMT_{td}$ = Gesamtzahl der Terme im Dokument

Die Annahme, dass Wörter, die am Anfang stehen, eine andere Bedeutung als solche, die am Ende stehen, haben, kann in der Termgewichtung berücksichtigt werden. Neben der Position des Terms kann auch die Struktur des Dokuments berücksichtigt werden. So wird ein Term im Titel anders bewertet als ein Term in der Zusammenfassung.¹⁶

Werden mehrere Dokumente untersucht und indexiert, kann die inverse Dokumentenhäufigkeit (IDF) berechnet werden. Diese gibt die Relation der Texte, die ein bestimmtes Wort enthalten, zur Gesamtzahl der Texte an. In der Bewertung können die Wörter, die nur in wenigen Dokumenten vorhanden sind, entsprechend gewichtet werden. Auch hier kann zur Begrenzung eine logarithmische Berechnungsbasis verwendet werden.¹⁷

$$IDF(t) = \frac{FREQ_{td}}{DOKFREQ_t}$$

$FREQ_{td}$ = Frequenz eines Terms im Dokument

$DOKFREQ_t$ = Gesamtzahl der Terme im Dokument

Ein anderes Modell geht von der Annahme aus, dass niedrig- bzw. hochfrequente Wörter nicht als Deskriptoren geeignet sind. Dazu werden für jeden Begriff die Häufigkeiten innerhalb aller Dokumente gezählt. Anschließend werden alle Begriffe eliminiert, die einen bestimmten Schwellenwert übersteigen bzw. unterschreiten. Die Wörter mit mittlerer Häufigkeit werden dann als Deskriptoren gewählt.¹⁸

2.2.2 Pattern-Matching-Verfahren

Der Ansatz bei Pattern-Matching-Verfahren beruht darauf, Muster in Dokumenten zu erkennen. Dieses Verfahren beschränkt sich dabei nicht ausschließlich auf Textdokumente, vielmehr können auch Bild- und Filmdokumente analysiert werden. Ein großer Vorteil dieses Ansatzes ist, dass lernfähige Systeme konzipiert werden können. Dabei wird sich

¹⁶Vgl. Stock, 2000, S. 162.

¹⁷Vgl. Stock, 2000, S. 161-162.

¹⁸Vgl. Salton und McGill, 1987, S. 67.

das Verhalten eines menschlichen Indexierers zu Nutze gemacht und aus diesem gelernt, Dokumente zuvor definierten Klassen zuzuordnen.¹⁹

2.2.3 Probabilistische Verfahren

Probabilistische Verfahren ermitteln die Relevanzwahrscheinlichkeiten von Termen. Dazu liegen Wahrscheinlichkeitsinformationen vor, die aus einer intellektuell indexierten Menge von Dokumenten (Trainingskorpus) ermittelt wurden. Der Assoziationsfaktor z , der „die Näherung für die bedingte Wahrscheinlichkeit, dass ein bestimmter Deskriptor s zuzuteilen ist, wenn ein Term t auftritt“,²⁰ ausdrückt, wird folgendermaßen berechnet:

$$z(t, s) = \frac{h(t, s)}{f(t)}$$

$f(t)$ = Anzahl der Dokumente, in denen t vorkommt

$h(t, s)$ = Anzahl derjenigen unter diesen Dokumente, denen der Deskriptor s intellektuell zugeteilt wurde

Damit lässt sich die Wahrscheinlichkeit bestimmen, ob der Deskriptor s für ein Dokument relevant ist.²¹

2.2.4 Computerlinguistische Verfahren

Computerlinguistische Verfahren analysieren Texte unter linguistischen Aspekten. Sie werden unterschieden in regelbasierte, die auf algorithmische Lösungen aufbauen, und wörterbuchbasierte Verfahren. Regelbasierte Verfahren streben eine generalisierte Verarbeitung und Indexierung an, die sprachliche Regelmäßigkeiten in Regeln fassen und mittels Algorithmen lösen. Wörterbuchbasierte Verfahren hingegen vollziehen die Indexierung unter Zuhilfenahme von Wörterbüchern.²²

Da die computerlinguistischen Verfahren im Mittelpunkt dieser Arbeit stehen, werden einzelne Lösungsansätze sowohl algorithmischen als auch wörterbuchbasierten Ursprungs in Kap. 3 im Detail betrachtet.

¹⁹Vgl. Nohr, 2001, S. 65-66.

²⁰Vgl. Nohr, 2001, S. 72.

²¹Vgl. Stock, 2007, S. 367.

²²Vgl. Nohr, 2001, S. 24.

2.3 Kriterien der Indexierungsgüte

Im Information Retrieval sind vor allem die Präzision und die Vollständigkeit der Ergebnisse, die ein System ermittelt, von Bedeutung. Ergebnisse, die nicht zur Suchanfrage passen, stören. Relevante Ergebnisse, die zur Suchanfrage passen, dürfen hingegen nicht fehlen.²³ Die Qualität einer Ergebnismenge wird durch die beiden Kenngrößen *Recall* und *Precision* ausgedrückt. *Recall* misst die Anzahl der relevanten Dokumente, die auch nachgewiesen wurden. *Precision* misst den Anteil der nachgewiesenen Dokumente, die relevant sind.²⁴

$RETREL$ = gefundene relevante Treffer

$RETNREL$ = gefundene nicht relevante Treffer

$NRETREL$ = nicht gefundene relevante Einheiten

$$\text{Recall} = \frac{RETREL}{RETREL + NRETREL}$$

$$\text{Precision} = \frac{RETREL}{RETREL + RETNREL}$$

Beispiel:

Für eine gegebene Suchanfrage sind 100 Dokumente relevant sind. Die Ergebnismenge umfasst 80 Dokumente, von denen aber nur 40 Dokumente tatsächlich relevant sind.

$$\text{Recall} = \frac{40}{100} = 0.4 = 40\%$$

$$\text{Precision} = \frac{40}{80} = 0.5 = 50\%$$

²³Vgl. Stock, 2007, S. 63.

²⁴Vgl. Salton und McGill, 1987, S. 187.

3 Computerlinguistische Verfahren

Die Verarbeitung natürlicher Sprachen, im Englischen als Natural Language Processing (NLP) bezeichnet, wird besonders von der Komplexität der Sprache, die verarbeitet werden soll, beeinflusst. Diese Kapitel beschäftigt sich mit der morphologischen, syntaktischen und semantischen Analyse²⁵ der deutschen Sprache, es zeigt aber auch im Vergleich mit der englischen Sprache Unterschiede der verschiedenen Lösungsansätze auf.

3.1 Fehlertoleranz

Textdokument können Schreibfehler enthalten, die die Analyse stören können. Daher ist es sinnvoll, Texte mit einer gewissen Toleranz zu verarbeiten. In Sprachen, in denen Groß- und Kleinschreibung unterschieden wird, kann eine Fehlerquelle ausgeschlossen werden, indem nur die Kleinschreibung eines Wortes betrachtet wird. Untersuchungen ergaben, dass im Information Retrieval bis zu 80% der Suchanfragen, die Rechtschreibfehler aufweisen, vier Fehlerklassen zugeordnet werden können.²⁶

Auslassung	Chmikalie
Einfügung	Chemiekalie
Substitution	Chemekalie
Vertauschung	Chmeikalie

Tab. 3.1: Fehlerklassen am Beispiel *Chemikalie*
Quelle: Nohr, 2001, S. 41

Bei einem direkten Vergleich der Zeichenketten aus Tab. 3.1 würde keine vollständige Übereinstimmungen festgestellt werden. Ein erster Ansatz wäre, die notwendigen Ände-

²⁵Kaiser, 1993.

²⁶Vgl. Nohr, 2001, S. 41.

rungen zu berechnen, um eine Zeichenkette in eine andere zu überführen. Der errechnete Wert kann als Maß für die Ähnlichkeit beider Zeichenketten verstanden werden.

Ursprünglich für ein nach phonetischen Merkmalen sortiertes Namenregister entwickelte Robert C. Russel Anfang des 20. Jahrhunderts ein Verfahren, welches im Information Retrieval als Soundex Bedeutung erlangte. Die Buchstaben eines Wortes werden in diesem Verfahren nach bestimmten Regeln durch Ziffern ersetzt, wobei der erste Buchstabe als Index erhalten bleibt. Buchstaben, die für den Klang des Wortes unbedeutend sind, werden dabei eliminiert. Das Ergebnis ist, dass gleichklingende Wörter die gleiche Zeichenfolge zugeordnet wird. Eine Weiterentwicklung des Soundex-Verfahrens ist Phonix²⁷. Das Verfahren behebt die Schwächen von Soundex, indem es „gewisse Buchstabenfolgen durch phonetisch analoge Gegenstücke“²⁸ ersetzt. Dabei wird der Kontext, in dem sich der Buchstabe befindet, berücksichtigt.²⁹

Wort	Soundex	Phonix
KNIGHT	K714	N53
NIGHT	N14	N53
NITE	N14	N53

Tab. 3.2: Beispiel für Soundex und Phonix
Quelle: Stock, 2007, S. 310

Phonix arbeitet in diesem Fall hinsichtlich des Klanges genauer als Soundex, da es keine Unterscheidung zwischen KNI und NI vornimmt.

Ein an Phonix angelehntes und an die deutsche Sprache angepasstes Verfahren ist die *Kölner Phonetik*. Es stammt von Hans Joachim Postel. Das Verfahren weist einem Wort eine Ziffernfolge zu. Von links nach rechts wird zunächst Buchstabe für Buchstabe nach den Ersetzungsregeln in Tab. 3.3 durch Ziffern ersetzt. Zu beachten ist ggf. der Kontext, d. h. der Buchstabe der vor bzw. nach dem zu ersetzenden Buchstaben steht.

²⁷phonetic indexing, auch bekannt als Phonex

²⁸Vgl. Stock, 2007, S. 309.

²⁹Vgl. Stock, 2007, S. 307-311.

Buchstabe	Kontext	Code
A, E, I, J, O, U, Y		0
H		-
B		1
P	nicht vor H	
D, T	nicht vor C, S, Z	2
F, V, W		3
P	vor H	
G, K, Q		4
C	im Anlaut vor A, H, K, L, O, Q, R, U, X vor A, H, K, O, Q, U, X außer nach S, Z	
X	nicht nach C, K, Q	48
L		5
M, N		6
R		7
S, Z		8
C	nach S, Z	
	im Anlaut außer vor A, H, K, L, O, Q, R, U, X	
	nicht vor A, H, K, O, Q, U, X	
D, T	vor C, S, Z	
X	nach C, K, Q	

Tab. 3.3: Ersetzungstabelle Kölner Phonetik
Quelle: Postel, 1969

Im zweiten Schritt werden alle doppelten Codes entfernt. Zuletzt werden alle Nullen außer am Anfang entfernt.³⁰

Für das Wort *Chemikalie* ergibt sich im ersten Schritt die Ziffernfolge 4060040500. Nach Entfernen aller doppelten Code bleibt 40604050. Schlussendlich bleibt die Ziffernfolge 4645, nachdem alle Nullen entfernt wurden. Derselbe Code ergibt sich für alle Schreibweisen des Wortes *Chemikalie* aus Tab. 3.1. Das Verfahren ist somit in diesem konkreten Fall geeignet, Fehler nach den vier erwähnten Fehlerklassen zu kompensieren.

3.2 Morphologische Analyse

Die morphologische Analyse eines Dokuments findet auf Wortebene statt. Dazu werden Wortformen verarbeitet und sinntragende Wortbestandteile untersucht.³¹ Wörter natür-

³⁰Postel, 1969.

³¹Vgl. Salton und McGill, 1987, S.275.

licher Sprachen treten in verschiedenen Wortformen auf. Beeinflusst werden sie durch Flexion, Derivation und Komposition.³²

3.2.1 Wortfilterung

Nicht jedes Wort eines Dokuments trägt eine semantische Information. Im Rahmen der Indexierung sind nur solche Wörter von Interesse, die den Inhalt des Dokuments repräsentieren. Daher ist es sinnvoll, nicht alle Wörter zu verarbeiten.

Die einfachste Lösung ist der Abgleich mit Stoppwortlisten. Stoppwortlisten sind Negativlisten und enthalten alle Wörter, die von der Verarbeitung ausgeschlossen werden sollen. Diese können von Hand, halb- oder vollautomatisch erstellt werden und sind sprachspezifisch. Automatisch erstellte Listen machen sich dabei statistische Merkmale zu Nutze. Hochfrequente Wörter, die in vielen Dokumenten auftauchen, zeichnen sich durch Beliebtheit aus und sollten eliminiert werden.³³

Eine weitere Möglichkeit bietet die Filterung nach syntaktischen Merkmalen, also der Bedeutung des Wortes in einem konkreten grammatischen Kontext. Part-of-speech Tagger, die in Kap. ?? noch näher betrachtet werden, analysieren die Wörter in einer grammatischen Einheit (z. B. eines Satzes) und identifizieren die entsprechende Wortart. Bestimmte Wortarten wie Artikel (der, die, das, ...), Präpositionen (in, auf, ab, ...) etc. sind isoliert betrachtet inhaltslos und können daher eliminiert werden.

Problematisch ist die Wortfilterung bei Wortphrasen, deren Bestandteile isoliert betrachtet, eliminiert werden würden. Die Phrase selbst kann allerdings für die weitere Verarbeitung von Bedeutung sein. Ein klassisches Beispiel ist „to be or not to be“,³⁴ deren Bestandteile u. U. vollständig eliminiert werden würden.

3.2.2 Wortstammbildung

Die Wortstammbildung, auch Stemming genannt, verfolgt das Ziel, eine Wortform durch die Anwendung von Regeln auf ihren Wortstamm zu reduzieren. Aufgrund des regelbasierten Ansatzes eignet sich dieses Konzept vor allem für regelmäßige Sprache wie dem Englischen. Das Ergebnis des Stemming ist oftmals nicht die eigentliche Grundform des

³²Hausser, 2000, S. 265-266.

³³Stock, 2007, S. 225.

³⁴Vgl. Stock, 2000, S. 151.

Wortes, sondern eine künstliche Form. In der Praxis existieren zwei Ansätze, Endungen in Wortformen zu löschen und den Wortstamm zu erhalten. Das Prinzip des Longest-Match-Stemmers ist, in einem einzigen Arbeitsschritt die längstmögliche Endung abzuschneiden. Dafür werden neben kontextspezifischen Regeln eine Liste mit Endungen verwendet. Im iterativen Ansatz werden Schritt für Schritt Endungen abgeschnitten, bis die Stammform erreicht wurde.³⁵

Der Porter-Stemmer ist ein häufig eingesetzter iterativer Stemmer von Martin F. Porter. Für die englische Sprache entwickelt, existiert heute auch eine Variante für die deutsche Sprache. Der Algorithmus verwendet Ersetzungsregeln, die die Wortform in mehreren Schritten so lange verkürzt, bis der Wortstamm gefunden wurde. Die Ersetzungsregeln in einem Arbeitsschritt sind abhängig von der m -Bedingung. Die m -Bedingung zählt Abfolgen von VC -Gruppen in einem Wort. V steht für Vokale sowie Y , wenn diese einem Konsonanten folgt, und C steht für Konsonanten. V und C können jeweils mehrere aufeinanderfolgende Vokale bzw. Konsonanten enthalten. Ein Wort kann dann dargestellt werden als $[C](VC)^m[V]$. Die Zahl m gibt die Anzahl der VC -Gruppen an.

Die Ersetzungsregeln eines Arbeitsschritt enthalten immer eine Bedingung, die erfüllt sein muss, und eine Anweisung, wie die Endung eines Wortes ersetzt wird. Dabei kann die Endung komplett eliminiert oder durch eine andere Endung ersetzt werden. In fünf Iterationsrunden werden verschiedene Ersetzungsregeln unter Einhaltung der m -Bedingung angewandt, bis das Wort auf den Wortstamm reduziert wurde. Die m -Bedingung sorgt dabei dafür, dass kein Overstemming stattfindet.

Stemming beeinflusst Recall und Precision (siehe Kap. 2.3) im IR gleichermaßen. Es erhöht zwar die Vollständigkeit der Ergebnisse, schadet aber andererseits auch der Präzision, da womöglich viele unterschiedliche Wortformen auf den gleichen Wortstamm reduziert werden.

3.2.3 Grundformbildung

Die Grundformbildung, auch Lemmatisierung genannt, verfolgt das Ziel, für eine Wortform durch Anwendung spezifischer Regeln oder durch Nachschlagen in einem Wörterbuch die Grundform zu ermitteln. Die Grundform ist bei Substantiven der Nominativ Singular, bei Verben der Infinitiv Präsens und bei Adjektiven die unmarkierte Form im Positiv. Regelgeleitete Verfahren ähneln dem Stemming, da gleichermaßen Ersetzungsre-

³⁵Vgl. Stock, 2007, S. 233-234.

geln verwendet werden. Sie greifen auf Grundformenlexika zurück. Ein Beispiel für die englische Sprache ist der S-Lemmatierer, der nach vier Regeln Ersetzungen vornimmt

- Wortlänge ≤ 3 , dann beende Bearbeitung,
- Wortform endet mit IES, aber nicht mit EIES oder AIES, dann ersetze IES durch Y,
- Wortform endet mit ES, aber nicht mit AES, EES oder OES, dann ersetze ES durch E,
- Wortform endet mit S, aber nicht mit US oder SS, dann lösche S,

und anschließend die reduzierte Form in einem Grundformenlexikon nachschlägt. Für Verben bietet sich zudem die Erweiterung auf -ING- und -ED-Endungen an.³⁶

Das Wörterbuchbasierte Verfahren nutzt ein Vollformenlexikon, das allen möglichen Wortformen ihre Grundform zuordnet. Die Erstellung und Wartung ist aufwändig, zudem kann eine Vollständigkeit nie gewährleistet werden. Gerade die deutsche Sprache ist reich an Wortzusammensetzungen, die nicht vollständig erfasst werden können. Dennoch stellt ein Vollformenlexikon besonders in Sprachen mit starken Flexionen³⁷ wie dem Deutschen eine wichtige Möglichkeit zur Lemmatisierung dar.

Derivation

Derivation bezeichnet die Bildung neuer Wörter, in dem sogenannten Morphemen Präfixe bzw. Suffixe hinzugefügt werden.³⁸

Es kann sinnvoll sein, Wortformen verschiedener Wortklassen auf dasselbe Lemma zurückzuführen. Dies leistet die Derivation, die wortklassenübergreifend eine Lemmatisierung z. B. von einem Adjektiv zu einem Substantiv durchführt:

- 1. Schritt (Lemmatisierung innerhalb der Wortklasse): mikroskopischen -> mikroskopisch
- 2. Schritt (Lemmatisierung durch Derivation): mikroskopisch -> Mikroskop

Aufgrund der Regelmäßigkeit der Derivation in der deutschen Sprache ist es möglich, über Listen die Präfixe bzw. Suffixe zu identifizieren und zu entfernen. Lemmatisierung

³⁶Vgl. Stock, 2007, S. 229.

³⁷Starke Flexion bei Verben: er liest, er las, er hat gelesen.

³⁸Vgl. Meibauer u. a., 2002, S. 55.

durch Derivation erhöht zwar den Recall in IR-Systemen, dafür verschlechtert sich die Precision.

Komposition

Die Komposition stellt neben der Derivation die wichtigste Art der Wortbildung in der deutschen Sprache dar. Als Komposita werden Einheiten bestehend aus mehreren Wörtern verstanden. Dabei können u. a. Verben, Adjektive und Substantive miteinander kombiniert werden.³⁹

Komposition	Beispiel
$N \rightarrow N + N$	Holz+haus
$N \rightarrow V + N$	Web+stuhl
$A \rightarrow N + A$	haut+freundlich
$V \rightarrow N + V$	kopf+stehen

Tab. 3.4: Kompositionstypen
Quelle: Meibauer u. a., 2002, S. 48-49

Die Zerlegung von Komposita kann unter Zuhilfenahme eines Wörterbuchs geschehen. Durch eine zeichenweise Abtrennung wird die längste Übereinstimmung (longest match) beider Wörter im Wörterbuch gesucht. Kann nur eines der Wörter gefunden werden, wird die Zerlegung verworfen.⁴⁰

Von rechts nach links ergeben sich für das Wort *Staubecken* die Bestandteile *Staub* und *Ecken*, von links nach rechts ergeben sich *Stau* und *Becken*. Die Zerlegung kann den Recall erhöhen, senkt u. U. aber auch die Präzision. Der Gebrauch des Wortes *Staubecken* im Sinne eines Wasserbeckens in einer Suchanfrage kann zu Ergebnissen führen, die „stau-bige Ecken“ enthalten.

3.2.4 Phrasenerkennung

Ein Begriff, im Sinne einer Bedeutungseinheit, muss sich nicht auf ein Wort beschränken, vielmehr kann er aus mehreren einzelnen Wörtern zusammengesetzt sein. Es können vier Ansätze⁴¹ unterschieden werden, um Phrasen zu erkennen.

³⁹Vgl. Meibauer u. a., 2002, S. 48-55.

⁴⁰Vgl. Stock, 2007, S. 263.

⁴¹Vgl. Stock, 2007, S. 250-259.

Statistische Methode

Die statistische Methode ermittelt Phrasen anhand von statistischen Merkmalen wie die Anzahl gemeinsamen und isolierten Auftretens, die Phrasenlänge, dem Abstand und der Umgebung.

Syntaktische Methode

Die syntaktische Methode identifiziert Phrasen anhand syntaktischer Eigenschaften. Dazu ist eine syntaktische Analyse (siehe Kap. 3.3) erforderlich.

Verklumpung von Textteilen

Durch Eliminierung von Wörtern, die kein Teil einer Phrasen sein können, werden Texte auf Textklumpen reduziert. Dazu können Stoppwortlisten verwendet werden. In der englischen Sprache kann man sich zudem zu Nutze machen, dass Eigennamen großgeschrieben werden.

Nutzung von Wörterbüchern

Indikatorlisten für z. B. Eigennamen können Hinweise darauf geben, ob es sich um Phrasen handelt.

3.3 Syntaktische Analyse

Die syntaktische Analyse beschäftigt sich mit der Struktur von Sätzen und der Bedeutung einzelner Wörter für diese. Während die morphologische Analyse sich zunächst auf Wortebene beschränkt, können durch die syntaktische Analyse größere Zusammenhänge erkannt werden. Es existieren zwei Anwendungsfelder, die im Folgenden betrachtet werden.

3.3.1 Part-Of-Speech Tagging

Ziel des Part-Of-Speech Tagging ist, die Wortart einer gegebenen Wortform zu bestimmen. Dies muss unter Beachtung der syntaktischen Struktur, in der das Wort verwendet wird, erfolgen. Dazu wird der Wortform ein Tag zugeordnet, das aus einem zuvor festgelegten Tag-Set stammt. Es existieren verschiedene Tag-Sets, die sprachabhängig sind. Das bekannteste deutsche Tagset ist das Stuttgart-Tübingen-Tagset (STTS). Im Englischen wird oftmals das Brown-Tagset bzw. die vereinfachte Form davon, das Penn-Treebank-Tagset (PTTS), verwendet. POS Tagger können dabei u. a. folgende Paradigmen⁴² verfolgen:

Transformationsbasiertes Tagging

Aus einem Regelwerk wird ein endlicher Automat mit Zuständen, Zustandsübergängen und Aktionen erzeugt. Mithilfe eines Wörterbuchs werden einer Wortform alle möglichen Tags zugeordnet. Anschließend werden Mehrdeutigkeiten durch Regeln unter Berücksichtigung der syntaktischen Umgebung aufgelöst. Die Regeln müssen dabei von Hand erstellt werden und erfordern entsprechendes Wissen.

(Hidden) Markov Modelle

Hidden Markov Modelle werden in Statistik-basierten Taggern eingesetzt. Sie bilden endliche Automaten mit Übergangswahrscheinlichkeiten zwischen den Zuständen. Für eine Wortsequenz wird die höchste Wahrscheinlichkeiten für eine mögliche Tagzuordnung berechnet und ausgewählt. Dazu muss anhand eines getaggten Trainingskorpus Wissen über mögliche Tags und Übergangswahrscheinlichkeiten vorliegen.

Maximum Entropy Modelle

Maximum Entropy Modelle fanden im Bereich NLP das erste Mal in der maschinellen Übersetzung⁴³ und der Satzgrenzenerkennung⁴⁴ praktische Anwendung. Sie eignen sich aber auch für POS Tagging. Auf Grundlage eines Wahrscheinlichkeitsmodells, das mit-

⁴²Vgl. Schröder, 2002.

⁴³Berger, Pietra und Pietra, 1996.

⁴⁴Reynar und Ratnaparkhi, 1997.

hilfe eines handgetaggten Trainingskorpus erstellt wird, werden Wörtern unter Berücksichtigung von kontextuellen „*Histories*“ Tags zugeordnet.

Heutige Tagger erzielen eine Genauigkeit abhängig vom Textkorpus von über 96%.⁴⁵

3.3.2 Syntax Parsing

Ein Parser ermittelt durch Zerlegung (Parsing) sogenannte Konstituenten (syntaktische Einheiten) und Abhängigkeiten zwischen diesen. Parser arbeiten nach einem Regelsystem. Dieses wird als Grammatik bezeichnet. Zu syntaktischen Strukturen wie z. B. einem Satz wird im Vorgang des Parsens eine Strukturbeschreibung produziert. Dabei sind zwei Arbeitsweisen zu unterscheiden. Der Top-Down-Parser arbeitet die Gesamtheit der syntaktischen Struktur von oben nach unten ab und zerlegt diese schrittweise in die Konstituenten. Der Bottom-Up-Parser geht von den einzelnen Konstituenten von unten nach oben und fasst diese zu syntaktischen Einheiten zusammen, bis die Spitze erreicht wurde.⁴⁶

Beispiel für Top-Down-Parsing:

Gegeben ist der Satz „Der Hund jagt die Katze“ mit den Wortklassensymbolen

ART, SUB, VER, ART, SUB

und der Grammatik

$S \rightarrow NP + VP$

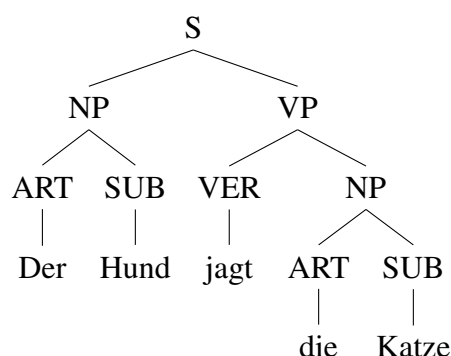
$NP \rightarrow ART + SUB$

$VP \rightarrow VER + NP.$

Beginnend beim Startsymbol *S* werden die Regeln durchlaufen und die Variablen ersetzt. Das Ergebnis ist folgende Baumstruktur:

⁴⁵Vgl. Schröder, 2002, S. 17.

⁴⁶Vgl. Lenders und Willée, 1986, S. 107-109.



Die Grammatik im Beispiel besteht nur aus wenigen Regeln. Diese reichen für einfache Sätze aus. Für komplexe Satzstrukturen, wie sie z. B. in der deutschen Sprache durch verschachtelte Relativsätze o. ä. auftreten, würde die Anzahl der Regeln entsprechend größer ausfallen.

Die Grammatik ist ein Beispiel für eine Phrasenstrukturgrammatik, die anhand von kontextfreien Ersetzungsregeln grammatisch korrekte Sätze erzeugen kann. Ein anderes Beispiel ist die Transformationsgrammatik, die kontextsensitive Ersetzungsregel zulässt.⁴⁷

3.4 Semantische Analyse

In der semantische Analyse wird das gesamten Dokument analysiert und versucht, Bedeutungen von Wörtern mithilfe von kontextuellem Wissen zu untersuchen.⁴⁸ Dazu gehört z. B. die Auflösung von Doppeldeutigkeiten (Word Sense Disambiguation), wie bei dem Wort *Bank* als Kreditinstitut oder Sitzmöglichkeit. Dafür reicht das Wissen über morphologische oder syntaktische Eigenschaften nicht aus. Es wird Wissen benötigt, dass sich aus dem Kontext, in dem das Wort verwendet wird, ergibt. Kontextuelles Wissen kann z. B. in der Form eines Wortnetzes, in dem Wörter und ihre Beziehungen zueinander abgebildet werden, vorliegen. Das WordNet⁴⁹ ist für die englische Sprache erstelltes lexikalisch-semantisches Netzwerk. Für die deutsche Sprache existiert GermaNet, das von der Universität Tübingen⁵⁰ erstellt und gepflegt wird. GermaNet bildet ein lexikalisch-semantisches Netz, das Substantive, Verben und Adjektive enthält und miteinander verbindet. Es existieren Kategorien wie z. B. Gesellschaft, Geist und Bewegung, deren Wör-

⁴⁷Salton und McGill, 1987, S. 291.

⁴⁸Vgl. Nohr, 2001, S. 44.

⁴⁹<http://wordnet.princeton.edu/>.

⁵⁰<http://www.sfs.uni-tuebingen.de/GermaNet/>

ter über Relationen miteinander verbunden sind. Es werden u. a. folgende Relationen abgebildet:

- **Hyperonymie:** Oberbegriff eines Begriffs (*Blume* ist Oberbegriff von *Rose*).
- **Holonymie und Meronymie:** Teil-von-Beziehung (*Rinde* (Meronym) ist Bestandteil von *Baum* (Holonym)).
- **Synonymie:** Begriffe, die die gleiche oder eine ähnliche Bedeutung haben (*Bauer* ist synonym zu *Landwirt*).
- **Antonymie:** Begriffe, die miteinander inkompatibel sind (*heiß* ist inkompatibel zu *kalt*).

Eine vollständige Auflistung der Relationen findet sich auf der Webseite⁵¹ des Herausgebers.

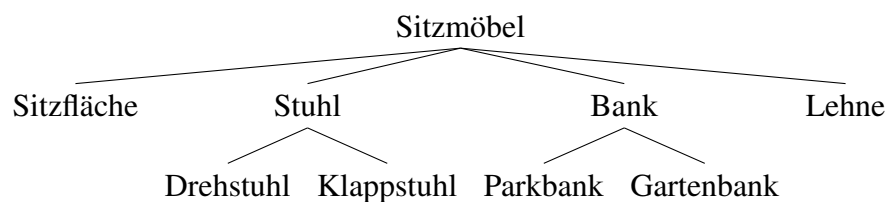


Abb. 3.1: Lexikalische Ordnung GermaNet
Quelle: Lemnitzer und Kunze, 2007, S.142

Abb 3.1 zeigt einen Auszug aus GermaNet. Doppeldeutigkeiten können nun mithilfe zusätzlicher Informationen einer syntaktischen Einheit aufgelöst werden. Enthält der Satz, in dem das Wort *Bank* verwendet wird, zusätzlich das Verb *sitzen*, kann über die Relationen auf den Oberbegriff *Sitzmöbel* geschlossen werden.

⁵¹<http://www.sfs.uni-tuebingen.de/lsd/rerelations.shtml>

4 Konzeption und Implementierung des Mehrbenutzereditors

4.1 Ausgangssituation

Die im theoretischen Teil der Arbeit erlangten Erkenntnisse sollten sich in einer praktischen Umsetzung eines Mehrbenutzereditors wiederfinden. Dafür wurde ein Prototyp entwickelt, der dem Benutzer die grundlegenden Funktionen eines Schreibprogramms bieten sollte. Zudem sollte eine automatische Indexierung unter Zuhilfenahme von computerlinguistischen Verfahren und der Vergleich von Dokumenten erfolgen.

Die in diesem Kapitel beschriebenen Funktionen und Funktionsweisen geben dabei den letzten Stand des Systems wieder. Während der Evaluierung des Prototyps (siehe Kap. 5) wurden immer wieder Fehler identifiziert und korrigiert. Änderungen am System, die aufgrund der Erkenntnisse in den Studien vorgenommen wurden, sind in den Studienbeschreibungen vermerkt.

Es wurde verzichtet, existierende Softwarelösungen zu nutzen bzw. auf diese aufzubauen. Zwar existieren Softwarelösungen für das kollaborative Schreiben in verteilten Umgebungen, die Anforderungen, z. B. hinsichtlich einstellbaren Zugriffsrechten, wurden allerdings nicht erfüllt oder die notwendigen Anpassungen hätten den Aufwand nicht gerechtfertigt.

Die Auswahl der Technologie, um den Prototyp zu implementieren, orientierte sich vor allem an vier zuvor identifizierten Anforderungen:

- Die Technologie muss die Implementierung einer verteilten Anwendung durch die eigene API oder durch frei verfügbare Bibliotheken ermöglichen.

- Die Technologie muss die Möglichkeit der nebenläufigen Programmierung unterstützen, damit zeit- und rechenintensive Vorgänge ausgelagert werden können.
- Die Technologie muss den Zugriff auf eine relationale Datenbank ermöglichen, da die zu erwartenden Datenmengen gespeichert und für spätere Zugriffe verfügbar gemacht werden müssen.
- Die Technologie muss die Gestaltung einer grafischen Benutzeroberfläche unterstützen. Von Vorteil sind bekannte und intuitiv bedienbare GUI-Elemente.

Die Wahl fiel auf die Programmiersprache Java⁵² von Oracle. Java erfüllte die vier Anforderungen. Außerdem zeigte sich während der Recherche, dass zu einer Vielzahl von computerlinguistischen Verfahren Implementierungen in Java existieren.

4.2 System- und Datenstrukturbeschreibung

Beim Prototyp handelt es sich um eine klassische Client-Server-Anwendung, der eine Drei-Schichten-Architektur bestehend aus Präsentations-, Anwendungs- und Datenhaltungsschicht zugrunde liegt. Der Verteilungsschnitt zwischen Client und Server liegt zwischen Präsentations- und Anwendungsschicht.

Die Datenhaltungsschicht bilden das Framework Hibernate⁵³ und eine relationale MySQL⁵⁴-Datenbank. Hibernate ermöglicht eine Abbildung von Attributen relationaler Datenbanktabellen auf Attribute in Java-Objekten.

⁵²<http://www.oracle.com/technetwork/java/index.html>

⁵³<http://www.hibernate.org/>

⁵⁴<http://www.mysql.de/>

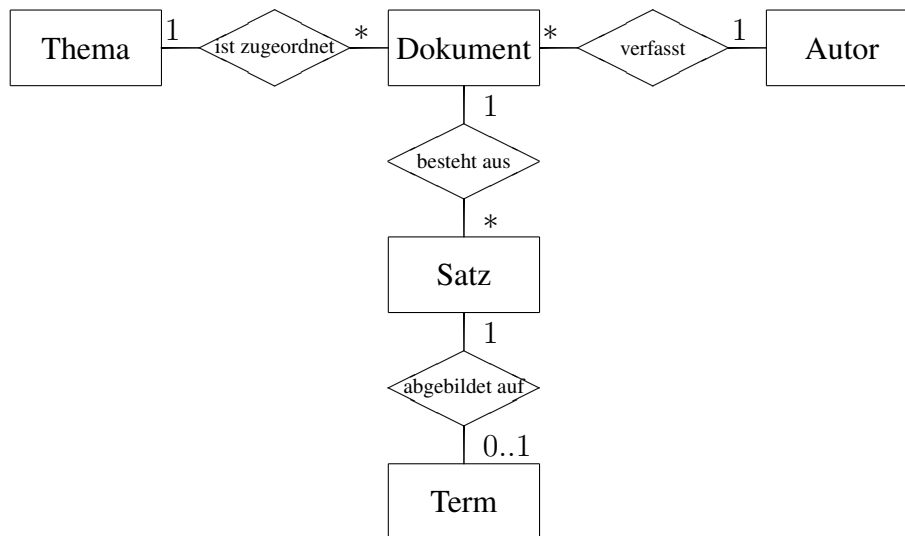


Abb. 4.1: ERM: Objekte in der Datenhaltungsschicht

Abb. 4.1 stellt einen Ausschnitt der Objekte in der Datenhaltungsschicht dar. Der Einfachheit halber wurde auf die Darstellung von Attributen verzichtet. Ein Thema besteht aus Dokumenten, die von einem Autor erstellt wurden. Die Dokumente enthalten Sätze, denen ein Term zugeordnet ist. Der Term bildet den Index ab. Dem Index liegt eine Baumstruktur zugrunde. Diese wurde in Java mit den Klassen *Term* und *Word*, wie in Abb. 4.2 dargestellt, abgebildet.

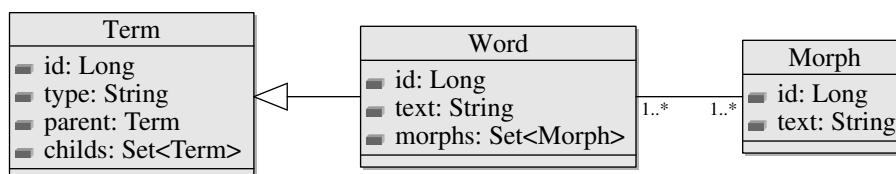


Abb. 4.2: Klassendiagramm: Term, Word und Morph
Quelle: Eigene Darstellung

Term steht allgemein für eine grammatische Struktur. *Word* ist eine Spezialisierung von *Term* und bildet ein indexiertes Wort in der extrahierten Form ab, dem eine oder mehrere morphologische Grundformen, durch *Morph* repräsentiert, zugeordnet werden.

Zur Kommunikation zwischen Client und Server wird RMI eingesetzt. Dafür wurde zunächst ein Observer-Muster genutzt. Dieses wurde nach einem Artikel von Prof. Dr. Thomas Steiner⁵⁵ implementiert und ermöglichte sogenannte Callbacks vom Server zu den

⁵⁵Steiner, 2005.

Clients, um diese über globale Änderungen zu informieren. In einer ersten Pilotstudie zeigte sich allerdings eine Schwäche dieses Konzepts. Die Technik versagte aufgrund gesperrter Netzwerkports. RMI weicht in diesem Fall auf das zustandslose Hypertext-Transfer-Protokoll (HTTP) aus. Da das Protokoll keine Beziehung zwischen Server und Client herstellt, werden Callbacks vom Server zum Client verhindert. Aus diesem Grund wurde die Kommunikation kurzfristig auf Polling umgestellt. Jeder Client fragt nun regelmäßig den Server nach Änderungen ab. Dies ist bei vielen verbundenen Clients nicht die optimale Lösung, in der Praxis hat sich diese Art der Kommunikation mit nur wenigen Benutzern (≤ 6) aber als sehr stabil und somit praktikabel erwiesen.

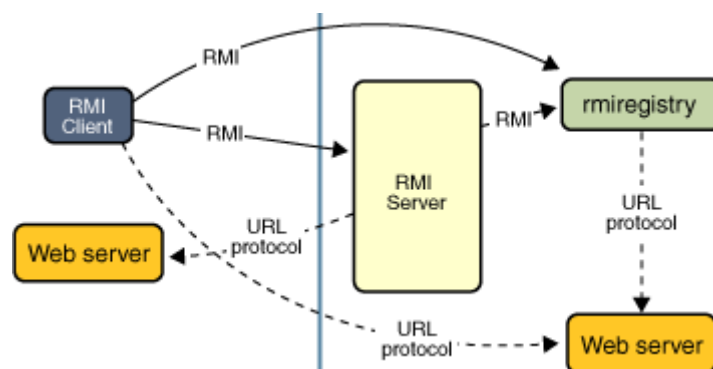


Abb. 4.3: Übersicht RMI

Quelle: <http://download.oracle.com/javase/tutorial/rmi/overview.html>

Abb. 4.3 stellt die Abläufe der Registrierung und den Aufruf von entfernten Methoden dar. Das Prinzip von RMI ist, dass der Server Objekte bei der RMI-Registry anmeldet, auf die der Client über vordefinierte Schnittstellen zugreifen kann. Die Objekte stellen Methoden zur Verfügung, denen Parameter übergeben werden und die Werte zurückgeben können.

RMI bietet eine stärkere Kopplung zwischen Client und Server als z. B. eine nachrichtenbasierte Lösung wie JMS⁵⁶. Der Client erhält, sofern der Server verfügbar ist und antworten kann, eine sofortige Rückmeldung, ansonsten wird ein Kommunikationsfehler, ausgelöst z. B. durch einen Timeout, erkannt. Dies ist in vielen Fällen, wie z. B. der Anmeldung eines Benutzers am Server, die optimale Lösung. In anderen Fällen, wie z. B. dem Laden großer Datenmengen bzw. komplexer Datenstrukturen, zeigen sich dabei aber Schwächen. Bis der Server auf die Anfrage mit einer Antwort reagiert, blockiert der Client. Dies ist bei der Verwendung von Hibernate und komplexen Datenstrukturen nicht ungewöhnlich. Aus diesem Grund wurde angelehnt an eine nachrichtenbasierte Lösung eine Warteschlange für clientseitige Anfragen implementiert. Die Anfragen werden der Warte-

⁵⁶http://download.oracle.com/docs/cd/B14099_19/web.1012/b14012/jms.htm

schlange hinzugefügt und nach dem FIFO-Prinzip abgearbeitet. Die Antwort, sofern eine erforderlich ist, wird einer clientspezifischen Warteschlange hinzugefügt, die der Client in regelmäßigen Abständen abfragt (Polling) und verarbeitet. Der Benutzer bekommt zwar Änderungen, wie z. B. das Öffnen eines Dokuments, mit einer entsprechenden Verzögerung angezeigt, allerdings blockiert die Anwendung währenddessen nicht.

4.3 Systemfunktionen

Die Anwendungsschicht liegt auf dem Server, der sowohl für die Indexierung als auch den Vergleich der Dokumente verantwortlich ist. Es ist möglich, beim Starten des Servers verschiedene Optionen einzustellen, die das Systemverhalten und die Funktionen des Clients beeinflussen.

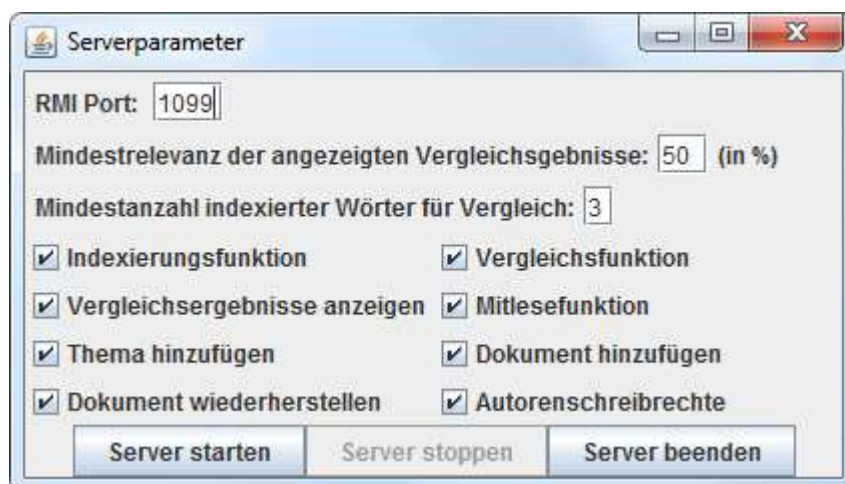


Abb. 4.4: Einstellbare Serverparameter
Quelle: Eigene Darstellung

Abb. 4.4 stellt die einstellbaren Optionen dar. Je nach Bedarf kann die Indexierungs- und Vergleichsfunktion eingestellt werden. Zudem kann festgelegt werden, ob Vergleichsergebnisse im Client angezeigt werden sollen.

Der Client des Schreibprogramm ist in seinem Funktionsumfang stark limitiert. Es können Themen erstellt werden. Zu einem Thema können Dokumente hinzugefügt werden. Die Dokumente eines Themas werden miteinander verglichen. Themenübergreifende Vergleiche von Dokumenten finden nicht statt. Ein Dokument kann immer nur von einem Benutzer schreibend geöffnet werden, alle anderen Zugriffe erfolgen ausschließlich lesend, ohne dass Änderungen am Dokument vorgenommen werden können. Änderungen

des Autors werden zeitnah (innerhalb weniger Sekunden) bei den anderen Benutzern, die das Dokument ebenfalls geöffnet haben, angezeigt. Diese Funktion wird im Folgenden als *Mitlesefunktion* bezeichnet. Beim Starten des Servers können verschiedene Rechte eingestellt werden, die den Zugriff auf die Dokumente regelt.

Mitlesefunktion	Autorenschreibrechte	Zugriff
aktiviert	aktiviert	Autor schreibend, andere Benutzer lesend
deaktiviert	aktiviert	Autor schreibend, andere Benutzer kein Zugriff
aktiviert	deaktiviert	1. Benutzer, der das Dokument öffnet, schreibend, andere Benutzer lesend
deaktiviert	deaktiviert	1. Benutzer, der das Dokument öffnet, schreibend, andere Benutzer kein Zugriff

Tab. 4.1: Einstellbare Zugriffsrechte

Wie Tab. 4.1 zeigt, ist der Zugriff auf Dokumente je nach Einstellung abhängig davon, ob der Benutzer auch gleichzeitig der Autor ist oder ob bereits ein anderer Benutzer schreibend auf das Dokument zugreift. Ist die Mitlesefunktion deaktiviert, kann das Dokument grundsätzlich nur von einem Benutzer geöffnet werden.

Das konfigurierbare Rechtesystem war für die Durchführung der Studien notwendig, da abhängig von der gestellten Aufgabe der Zugriff auf Dokumente gestattet bzw. verhindert werden musste.

4.3.1 Protokollfunktion

Schon zu Beginn der Bearbeitung wurde vom Betreuer angeregt, Benutzereingaben und -aktionen zu protokollieren, um in einer späteren Auswertung diese mit den Angaben der Teilnehmer vergleichen zu können und ggf. subjektive Fehlangaben durch objektiv ermittelte Daten zu kompensieren. Aus diesem Grund wurde eine Protokollfunktion in den Editor integriert. Von dieser werden drei verschiedene Ereignistypen aufgezeichnet:

- **[INSERT]** Einfügen einer Zeichenkette in das Dokument,
- **[REMOVE]** Entfernen einer Zeichenkette in dem Dokument,

- **[HIGHLIGHT]** Hervorheben einer Zeichenkette im Dokument durch Aufruf eines Vergleichsergebnisses.

Um die Anzahl der protokollierten Ereignisse zu minimieren, werden zusammenhängende Änderungen gleichen Typs zusammengefasst. Neben der eigentlichen Änderungen und dem Ereignistyp werden zudem die Anfangs- und Endposition im Dokument und ein Zeitstempel gespeichert. Dadurch ist es möglich, das Dokument von der Entstehung bis zur Fertigstellung in allen Schritten zu rekonstruieren.

Ereignistyp	eingefügte/markierte Zeichenkette	gelöschte Zeichenkette	Anfangsposition	Endposition
INSERT	Mein kleines blaues Auto fährt ganz schnell.		0	44
REMOVE		kleines	5	13
HIGHLIGHT	Mein blaues Auto fährt ganz schnell.		0	37

Tab. 4.2: Ausschnitt Ereignisprotokoll

Tab. 4.2 stellt den Auszug eines Ereignisprotokolls dar.

4.3.2 Indexierungsfunktion

Der Vergleich der Dokumente soll auf Satzebene stattfinden. Dazu müssen zunächst die Sätze eines Dokuments identifiziert werden. Erst dann kann Satz für Satz inhaltlich erschlossen werden.

Im Prototyp werden drei verschiedene Bibliotheken eingesetzt, die natürliche Sprachen verarbeiten können. Die Bibliothek des OpenNLP-Projekts basiert auf dem Maxent-Framework⁵⁷ und arbeitet mit Maximum Entropy Modellen (siehe Kap. 3.3.1). Die zweite Bibliothek ist BananaSplit⁵⁸ von Niels Ott. BananaSplit ist ein Werkzeug, um zusammengesetzte Wörter (siehe Kap. ??) in ihre Bestandteile zu zerlegen. Es basiert auf einem Ispell-Wörterbuch⁵⁹. Für das Stemmen unbekannter Wortformen wird der Snowball Stemmer⁶⁰, der auf dem

⁵⁷<http://maxent.sourceforge.net/about.html>

⁵⁸<http://niels.drni.de/s9y/pages/bananasplit.html>

⁵⁹<http://www.gnu.org/software/ispell/ispell.html>

⁶⁰<http://snowball.tartarus.org/>

Porter-Stemmer (siehe Kap 3.2.2) basiert und an die deutsche Sprache angepasst ist, eingesetzt.

Identifizierung von Sätzen, Tokens und POS-Tags

Die OpenNLP-Bibliothek stellt Methoden zur Verfügung, um Sätze und Tokens⁶¹ zu identifizieren und die Wortart von Wörtern zu ermitteln. Die Basis ist jeweils ein sprachabhängiges Modell⁶², das mit einem Textkorpus trainiert wurde. Es ist möglich, Modelle mit eigenen Trainingsdaten zu erstellen. Da dies allerdings aufwändig ist und mit den verfügbaren Modellen ausreichend gute Ergebnisse erzielt wurden, wurde darauf verzichtet.

Der *Sentence Detector* erkennt Satzgrenzen anhand von Satzzeichen, wobei er unterscheidet, ob ein Punkt z. B. eine Abkürzung markiert. Der *Tokenizer* identifiziert die einzelnen Tokens des Satzes. Der *POS Tagger* ordnet anschließend den Tokens die entsprechende Wortart nach dem Stuttgart-Tübingen-Tagset⁶³ (STTS) zu. Das Tagset besteht aus 54 Tags für Wortarten wie Nomen, Verben, Adjektive etc. und Satzzeichen.

Lemmatisierung

In einem mehrstufigen Verfahren wird versucht, die Grundform eines Wortes zu ermitteln. Da, wie in Kap. 3.2.1 festgestellt, nicht jedes Wort in einem Satz von Bedeutung ist, werden nicht alle Wörter morphologisch untersucht. Auf Grundlage der zuvor durch den *POS Tagger* zugeordneten Wortarten werden folgende Wörter untersucht und versucht, sie zu lemmatisieren:

- Adjektive (schnell, schön, ...),
- Substantive (Haus, Bahnhof, ...),
- Kardinalzahlen (drei, 2011, ...),
- Vollverben,
 - finites Verb (geht, kommst, ...),
 - Imperativ (geh, komm, ...),
 - Infinitiv (gehen, kommen, ...),

⁶¹ Tokens sind im Sinne der Linguistik einzelne Wörter.

⁶² <http://opennlp.sourceforge.net/models-1.5/>.

⁶³ <http://www.ims.uni-stuttgart.de/projekte/corplex/TagSets/stts-table.htm>

- Partizip (gegangen, gekommen, ...),
- Eigennamen (Hans, Schmidt, ...).

Bei Kardinalzahlen wird vor der Lemmatisierung versucht, sie in das entsprechende Wort (1053 => eintausenddreißig) umzuwandeln. Da vor allem bei niedrigen Zahlen (≤ 20) mal Zeichen- mal Ziffernfolgen auftreten, bietet sich eine vorherige Umwandlung an, um diese besser vergleichen zu können.

Fremdsprachliches Material (z. B. „A big fish“) wird ohne Untersuchung in den Index aufgenommen.

Satzzeichen und Wörter wie Hilfsverben (sein, wollen, ...), Pronomen, Artikel, Adverbien, etc. werden nicht untersucht, sondern zu einem späteren Zeitpunkt eliminiert. Gleiche Wörter bzw. Wortformen werden nicht gefiltert, es werden somit auch Wörter mehrfach in den Index aufgenommen, sofern sie sich im Satz wiederholen.

Zur Lemmatisierung kommt ein Vollformenlexikon zum Einsatz. Dazu wurde zunächst ein Export⁶⁴ der internen Datenbank von *Morphy*⁶⁵, „ein frei verfügbares Softwarepaket für die morphologische Analyse und Synthese und die kontextsensitive Wortartenbestimmung des Deutschen“⁶⁶ von Wolfgang Lezius, verwendet. Nach der 1. Studie wurde das Lexikon erweitert. Hierfür wurde zusätzlich das von *BananaSplit* verwendete Ispell-Wörterbuch mithilfe von Morphy morphologisch analysiert. Das Ergebnis dieser Analyse wurde in eine Datenbank eingelesen. Das Vollformenlexikon umfasst in der aktuellen Fassung ca. 384.000 Substantive (davon ca. 2.500 Eigennamen), ca. 22.000 Adjektive und ca. 8.300 Verben. Das Wörterbuch macht keine Unterscheidung zwischen Singular und Plural, die Wortformen *Hauses* und *Häuser* werden gleichermaßen auf die Grundform *Haus* lemmatisiert.

Das Ziel der Lemmatisierung ist, einer gegebenen Wortform alle möglichen Grundformen zuzuordnen. Dass es sinnvoll sein kann, bei der Lemmatisierung die Wortklasse (Substantiv, Verb, Adjektiv) zu berücksichtigen, um Mehrdeutigkeiten von Wörtern aufzulösen, macht folgendes Beispiel deutlich:

Das Wort *laden* kann auf das Verb *laden* und die Substantive *Lade* im Sinne von Schublade und *Laden* im Sinne von Geschäft lemmatisiert werden. Eine eindeutige Zuordnung

⁶⁴Naber, 2011.

⁶⁵<http://www.wolfganglezius.de/doku.php?id=cl:morphy>

⁶⁶Vgl. Lezius, 2010.

muss nicht in jedem Fall möglich sein, wie die Wortform *abgewogen* in Tab. 4.3 verdeutlicht. In diesem Fall müsste zusätzlich eine semantische Analyse (siehe Kap. 3.4) des Satzes durchgeführt werden, um den semantischen Kontext, in dem das Wort verwendet wird, zu ermitteln.

Bei der fehlertoleranten Abfrage wird für ein gegebenes Wort zunächst der phonetische Code (siehe Kap. 3.1) mithilfe einer Implementierung der Kölner Phonetik erzeugt. Anhand dieses Codes wird eine Vorauswahl an gleichklingenden Wörtern im Lexikon gesucht. Anschließend wird mit einer Implementierung des Levenshtein-Algorithmus⁶⁷ das Wort oder die Wörter mit der geringsten Distanz zum Ursprungswort ermittelt. Um nicht beliebig schlechte Distanzen zuzulassen, ist die Anzahl der Änderungen, die erforderlich sind, begrenzt. Wird mehr als ein Wort ermittelt, werden dem Ursprungswort sämtliche Grundformen zugeordnet.

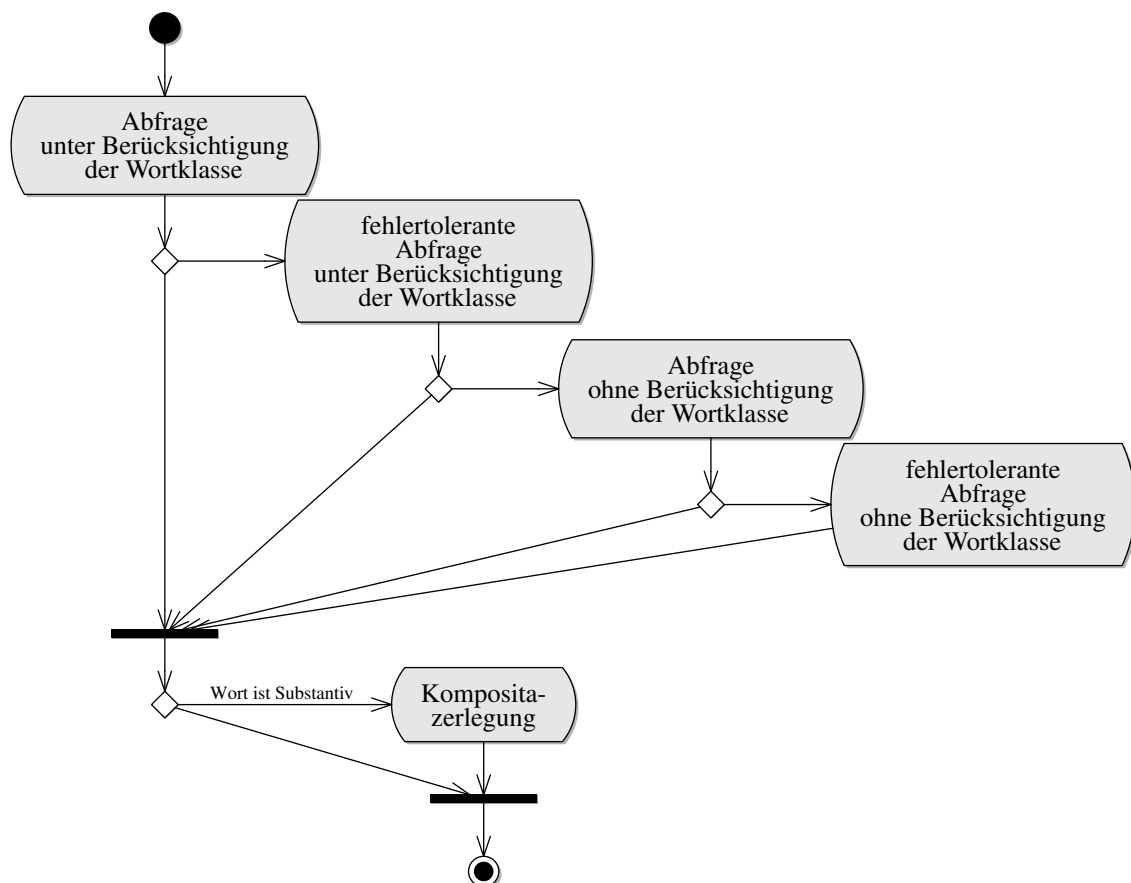


Abb. 4.5: Aktivitätsdiagramm: Lemmatisierung
Quelle: Eigene Darstellung

⁶⁷Der Levenshtein-Algorithmus berechnet die minimale Anzahl an Änderungen, die notwendig sind, um eine Zeichenkette in eine andere zu überführen.

Abb. 4.5 stellt die Suche im Vollformenlexikon dar. Bei Substantiven wird zusätzlich versucht, sie mithilfe von BananaSplit in ihre Bestandteile zu zerlegen. Die einzelnen Bestandteile werden dann wiederum lemmatisiert.

Konnte keine Grundform mithilfe des Lexikons ermittelt werden, wird im letzten Schritt das Wort mithilfe des Snowball Stemmers gestemmt (siehe Kap. 3.2.2).

#	Wortform	ermittelte Wortklasse	Grundform(en)
1	abgewogen	Verb	abwiegen, abwägen
2	schneller	Adjektiv	schnell
3	schneller	Substantiv	schnell, schnellere
4	feuerwermann	Substantiv	feuerwehrmann
5	berueete	Verb	beraten, bereiten, beruhen, bereuen

Tab. 4.3: Ergebnisse der Lemmatisierung

Tab. 4.3 zeigt eine Auswahl an lemmatisierten Wortformen. Wenn die Wortklasse, wie in Zeile 3, nicht korrekt durch den *POS Tagger* ermittelt werden konnte, kann es passieren, dass zusätzliche Grundformen zugeordnet werden. In Zeile 4 konnte der Fehler im Wort durch die fehlertolerante Abfrage kompensiert werden. Zeile 5 ist ein Extrembeispiel für ein Wort mit Schreibfehlern, dem sehr unterschiedliche Grundformen zugeordnet werden.

Identifizierung von grammatischen Strukturen

Im letzten Schritt der Indexierung wird versucht, grammatische Strukturen zu identifizieren und zusammenfassen. Dies erfolgt nach vier Regeln:

- Wortgruppen der Wortklassen *Adjektive* und *Verben*, die aufeinander folgen, nur durch ein Komma oder eine Und-Konjunktion voneinander getrennt sind,
- Wortgruppen der Wortklassen *Substantive* und *Eigennamen*, die aufeinander folgen,
- Nominalphrasen [NP] bestehend aus einem oder mehreren Adjektiven und einem Substantiv,
- Verbalphrasen [VP] bestehend aus einem Verb und einem oder mehreren Adjektiven

werden unter einer neuen grammatischen Struktur zusammengefasst.

Dies wird an folgendem Beispielsatz demonstriert.

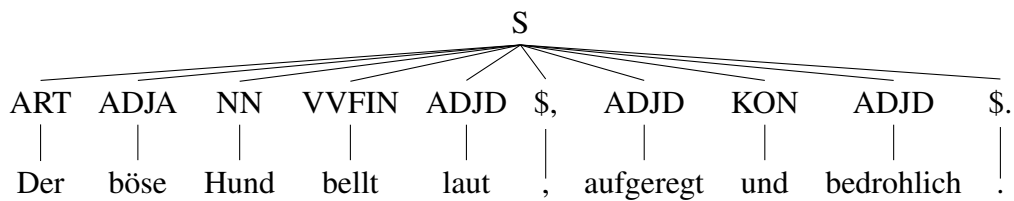
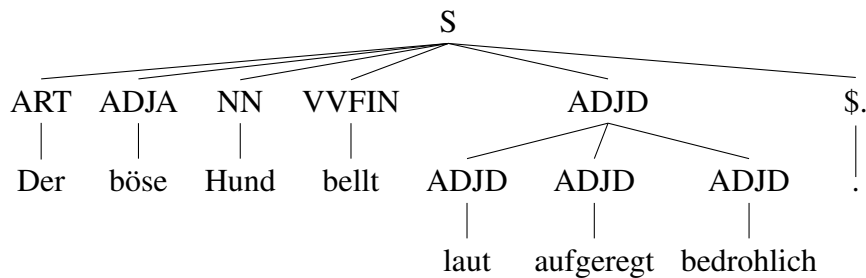
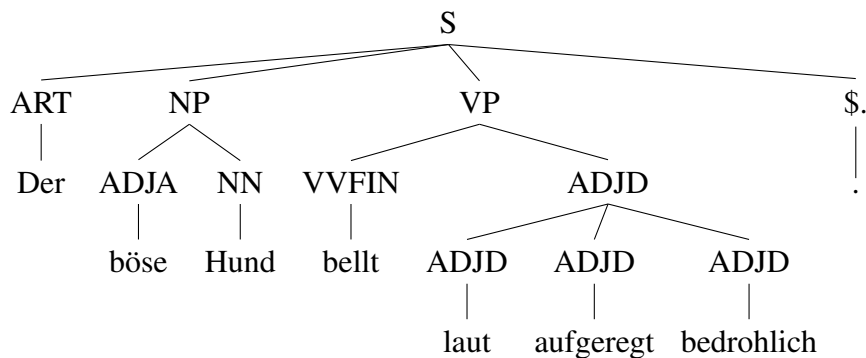


Abb. 4.6: Beispielsatz mit POS-Tags
Quelle: Eigene Darstellung

1. Zunächst werden die drei Adjektive *laut*, *aufgeregt* und *bedrohlich* zusammengefasst (Regel 1). Das Komma und die Konjunktion entfallen dabei.

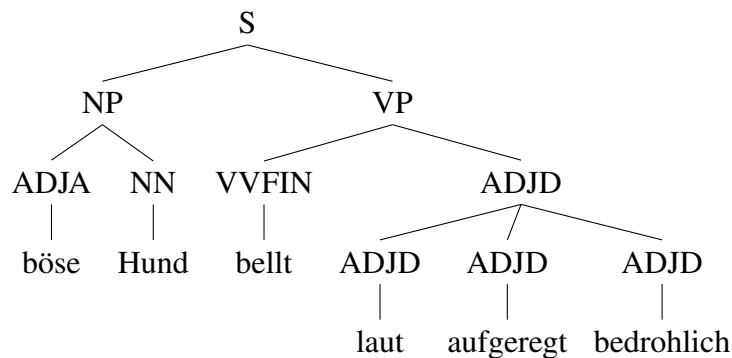


2. Anschließend werden Phrasen identifiziert und zusammengefasst (Regel 3 und 4).



Wortfilterung

Im letzten Schritt werden sämtliche Wörter und Zeichen entfernt, die keine Bedeutung haben.



Der so entstandene Indexbaum wird auf die in Abb. 4.2 dargestellte Klassenstruktur abgebildet. Jedes indexierte Wort wird durch die Klasse *Word* und jede grammatische Struktur durch die Klasse *Term* repräsentiert. Die zugeordneten Grundformen entsprechen der Klasse *Morph*.

4.3.3 Vergleichsfunktion

Nach der Indexierung werden sämtliche Dokumente eines Themas miteinander verglichen. Die Anzahl der Vergleiche ist anhängig von der Anzahl der Dokumente und der Anzahl der Sätze, die ein Dokument enthält. Insgesamt werden während des Schreibens mindestens

$$\sum_{i=1}^n m_i * \left(\sum_{j=1, j \neq i}^n m_j \right)$$

$n =$ Anzahl Dokumente, $m_i =$ Anzahl Sätze in Dokument i

Vergleiche durchgeführt. Da Sätze während des Schreibens geändert oder gelöscht werden können, handelt es sich nur um die Mindestanzahl an Vergleichen.

Beispiel:

In einem Thema mit 4 Dokumenten, wovon jedes Dokument aus 10 Sätzen besteht, werden mindestens

$$4 * (10 * (10 + 10 + 10)) = 1200$$

Vergleiche durchgeführt. Die Anzahl der Vergleichsergebnisse, die dem Benutzer angezeigt werden, kann durch Angabe einer Mindestrelevanz, wie in Abb. 4.4 zu sehen, be-

grenzt werden. Es werden grundsätzlich alle Vergleichsergebnisse gespeichert, um Vergleiche nicht wiederholen zu müssen. Die Gültigkeit eines Vergleichsergebnisses ist abhängig von den Sätzen, für die es erzeugt wurde. Wird der Ursprungs- oder Vergleichssatz gelöscht oder verändert, wird eine neue Indexierung und somit einer neuer Vergleich des Dokuments erforderlich. Das bisherige Vergleichsergebnis entfällt und wird dem Benutzer nicht mehr angezeigt.

Semantischer Vergleich

Der Vergleich der Wörter soll nicht ausschließlich auf Ebene der morphologischen Grundformen vollzogen werden, es sollte zudem auch die Möglichkeit bestehen, semantische Zusammenhänge zu identifizieren. Dazu wurde das in Kap. 3.4 beschriebene Lexikon GermaNet eingesetzt. Der Herausgeber von GermaNet stellt verschiedene Werkzeuge⁶⁸ zur Editierung und den Zugriff auf die Daten, die als XML-Dateien vorliegen, zur Verfügung. Es existiert eine API in Java. Diese stellt Methoden zur Verfügung, um zu einem gegebenen Wort sämtliche bzw. ausgewählte Relationen zu ermitteln. Ausgehend vom ersten Wort wird das semantische Netz durchlaufen, bis das zweite Wort gefunden wurde. Die Wörter *Haus* und *Garage* sind über 6 Kanten miteinander verbunden, *Haus* und *Hochhaus* hingegen über 2 Kanten. Eine Verbindung über nur eine Kante bedeutet eine sehr starke semantische Beziehung wie z. B. Synonymie bei den Verben *fischen* und *angeln*. Die Anzahl der Kanten, die durchlaufen werden können, wird begrenzt, um nicht beliebig schlechte Ergebnisse zu erhalten. Es ist grundsätzlich möglich, den Vergleich auf bestimmte Relationen wie z. B. Synonymie oder Antonymie (um Widersprüche aufzudecken) zu beschränken. Die konkrete Bewertung einer semantischen Relation soll allerdings vom Benutzer vorgenommen werden, der diese ggf. in einer Änderung seines Dokuments umsetzt. Daher werden alle möglichen Relationen erfasst.

Relevanzbewertung und Gewichtungsfaktoren

Anders als z. B. im Booleschen Retrieval, in dem eine Suchanfrage bzw. mehrere Suchanfragen verknüpft durch Operatoren eindeutig beantwortet werden können,⁶⁹ ist beim Vergleich von Sätzen oftmals keine Eindeutigkeit gegeben. Vielmehr ist von Interesse, wie „ähnlich“ sich zwei Sätze sind.

⁶⁸<http://www.sfs.uni-tuebingen.de/GermaNet/tools.shtml>

⁶⁹Vgl. Stock, 2007, S. 141-142.

Die Relevanz spiegelt die strukturelle Ähnlichkeit von Sätzen bzw. von Satzteilen und die semantische Ähnlichkeit von Wörtern, nicht jedoch die semantische Ähnlichkeit zweier Sätze wider. Zwei Sätze können z. B. durch das Wort *nicht*, das bei der Indexierung und somit beim Vergleich nicht berücksichtigt wird, unterschiedlicher Bedeutung. Daher ist die Bezeichnung *Ähnlichkeit* in diesem Kontext auch nicht ganz korrekt.

Die Bewertung der Relevanz konzentriert sich auf Übereinstimmungen von Wörtern und grammatischen Strukturen, die gewichtet werden. Folgende Faktoren werden dabei berücksichtigt:

- Art der grammatischen Struktur: Ein Wort erhält eine andere Gewichtung als z. B. eine Nominalphrase.
- Stärke der semantischen Beziehung zweier Wörter (ermittelt unter Zuhilfenahme von GermaNet): Die Anzahl der Kanten im semantischen Netz, die durchlaufen werden müssen, wird berücksichtigt.
- Tiefe der Wortebene bei zusammengesetzten Wörtern: Stimmt nur ein Teil eines zusammengesetzten Wortes mit dem Vergleichswort überein, wird dies in der Bewertung berücksichtigt.
- Verhältnis der Anzahl indexierter Wörter: Ein Satz, der weniger indexierte Wörter als der Vergleichssatz hat, wird anders bewertet als ein Satz mit mehr indexierten Wörtern.

Es werden grundsätzlich alle ermittelten Grundformen im Vergleich berücksichtigt. Es reicht aus, wenn eine der Grundformen des Ursprungswortes mit einer der Grundformen des Vergleichswortes übereinstimmt. Kann keine Übereinstimmung festgestellt werden, wird die semantische Beziehung mit GermaNet untersucht bzw. die nächste Wortebene bei zusammengesetzten Wörtern verglichen.

Vergleichsebenen

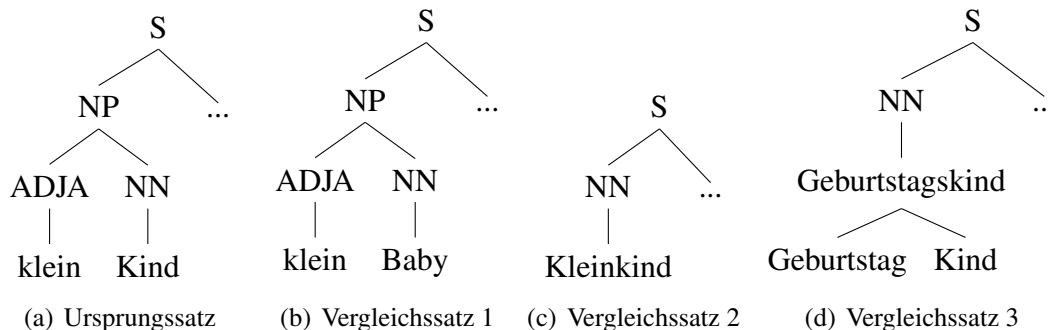


Abb. 4.7: Ausschnitte von vier indexierten Sätzen

Abb. 4.7 stellt Ausschnitte von vier Sätzen dar. Der Vergleich erfolgt im Tiefensuchverfahren. Ausgehend vom Ursprungssatz werden im Vergleichssatz entsprechende grammatische Strukturen und Wörter gesucht. Dabei wird nur die Übereinstimmung mit der höchsten Bewertung erfasst.

Der Vergleich des Ursprungssatz mit Vergleichssatz 1 ergibt, dass sich sowohl das Wort *klein* als auch das Wort *Kind* wiederfinden. *Kind* ist der Oberbegriff von *Baby*. Die grammatische Struktur [NP] wird in diesem Fall in der Bewertung berücksichtigt, da sie sich vollständig im Vergleichssatz wiederfindet.

Der Vergleich des Ursprungssatz mit Vergleichssatz 2 ergibt, dass sich nur das Wort *Kind* wiederfindet. *Kind* ist der Oberbegriff von *Kleinkind*. Da das Wort *klein* nicht gefunden wurde, wird die grammatische Struktur nicht in der Bewertung berücksichtigt.

Der Vergleich des Ursprungssatz mit Vergleichssatz 3 ergibt, dass sich das Wort *Kind* auf der zweiten Wortebene des Wortes *Geburtstagskind* wiederfindet. Dies wird in der Bewertung entsprechend berücksichtigt.

4.3.4 Darstellung der Vergleichsergebnisse

Die Darstellung der Vergleichsergebnisse im Client wurde bewusst zurückhaltend gestaltet. Da das Schreiben ein dynamischer Prozess ist, werden die Vergleichsergebnisse laufend aktualisiert. Hinzugefügte Sätze bedeuten mögliche neue Vergleichsergebnisse. Änderungen an bestehenden Sätzen können hingegen bedeuten, dass Vergleichsergebnisse

nicht mehr angezeigt werden. Daher wurde darauf verzichtet, den Benutzer durch Markierungen, Hervorhebungen oder selbstöffnenden Dialogen auf neue Vergleichsergebnisse hinzuweisen. Stattdessen muss der Benutzer die Ergebnisse aktiv aufrufen.

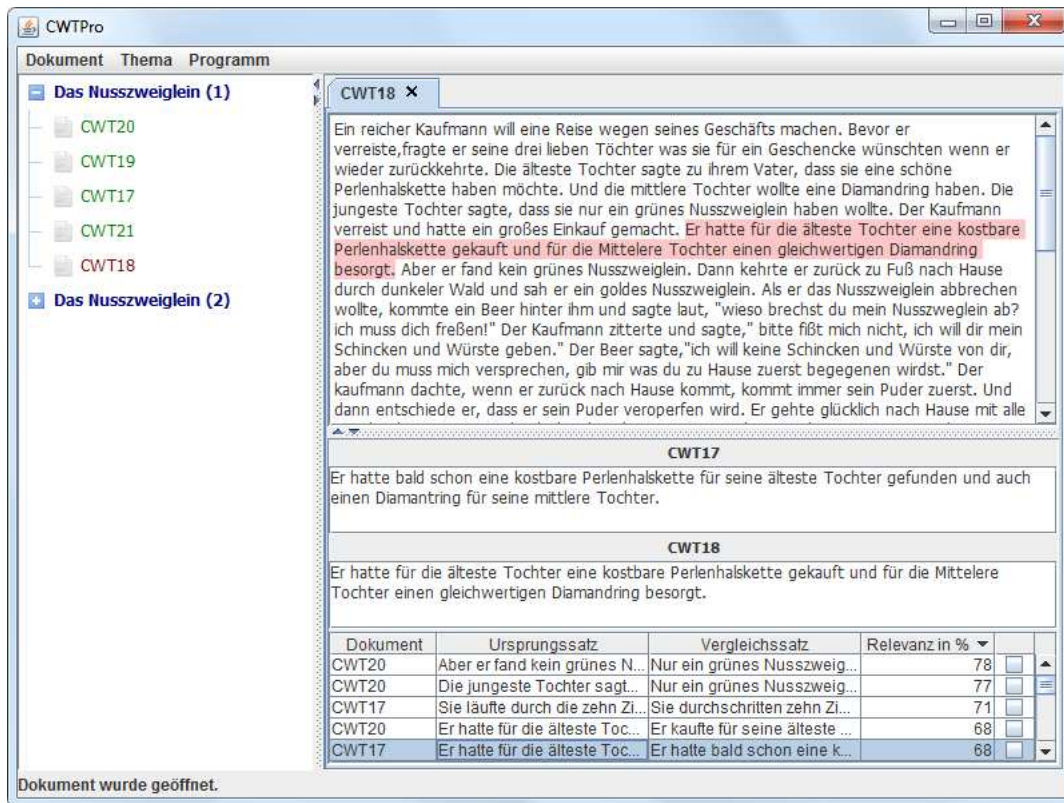


Abb. 4.8: Screenshot des Clients
Quelle: Eigene Darstellung

Abb. 4.8 zeigt die Aufteilung des Hauptfensters. Auf der linken Seite wird eine Übersicht mit verfügbaren Themen und Dokumenten angezeigt. Unterhalb des Texteingabefeldes werden die Vergleichsergebnisse in Form einer Tabelle angezeigt. Diese enthält folgende Spalten:

- Name des Vergleichsdokuments,
- Ursprungssatz,
- Vergleichssatz,
- Relevanz.

Die Ergebnisse können mithilfe der Spalten aufsteigend bzw. absteigend sortiert werden. Die Sortierung nach Ursprungs- bzw. Vergleichssatz entspricht der Reihenfolge des Auf-

tretens im Dokument. Wird ein Ergebnis aufgerufen, wird der Ursprungs- und der Vergleichssatz oberhalb der Tabelle angezeigt. Zudem wird der Ursprungssatz im Textfenster farbig markiert, damit der Benutzer diesen nicht suchen muss.

4.4 Kritische Systembetrachtung

Das Identifizieren und Zusammenfassen von grammatischen Strukturen (siehe Kap. 4.3.2) ist nicht mit letzter Konsequenz verfolgt worden. Das war vor allem der Tatsache geschuldet, dass kein Syntaxparser (siehe Kap. 3.3.2) für die deutsche Sprache gefunden werden konnte, der frei verfügbar ist und zufriedenstellende Ergebnisse liefert. Vielmehr ging es um die Überlegung, dass es durchaus sinnvoll sein kann, grammatische Strukturen im Indexbaum abzubilden und entsprechend in der Bewertung zu berücksichtigen.

Die Zerlegung von zusammengesetzten Wörtern mit BananaSplit beschränkt sich ausschließlich auf Substantive. Bei diesen wird versucht, sie in ihre Bestandteile zu zerlegen, wenn das Wort selbst nicht im verwendeten Wörterbuch vorhanden ist. Die Problematik wird vor allem in Abb. 4.7 (c) deutlich. Das Wort *Kleinkind* ist im Wörterbuch vorhanden und wird deswegen nicht in *klein* und *Kind* zerlegt. Die Bewertung des Vergleichs zwischen Satz 1 und Satz 3 beschränkt sich daher auf *Kind* und *Kleinkind*. Eine weitere Schwäche von BanaSplit ist, dass Fugenbindungen bei der Zerlegung nicht entfernt werden. *Gerichtsurteil* wird in *Gerichts* und *Urteil* zerlegt. Aufgrund dessen ist eine weitere Lemmatisierung der Bestandteile notwendig.

Zur Bestimmung der Grundform wird ein Vollformenlexikon verwendet, das Wörter nach der alten Rechtschreibung enthält. Das Wörterbuch von BananaSplit zur Zerlegung von Komposita wie auch GermaNet beruhen hingegen auf der neuen deutschen Rechtschreibung. Werden Schreibweise nach alter und neuer Rechtschreibung bei der Lemmatisierung teilweise durch die fehlertolerante Abfrage kompensiert, kann der Vergleich unter Zuhilfenahme von GermaNet allerdings in Einzelfällen gestört werden, da die Grundform immer der alten Rechtschreibung entspricht.

Die Gewichtungsfaktoren bei der Bewertung der Relevanz sind willkürlich gewählt und nicht evaluiert worden. Es existieren keine vergleichbaren Ansätze, die durch praktischen Einsatz oder wissenschaftliche Untersuchungen eine Orientierung hätten geben können. Es wurde lediglich darauf geachtet, dass die Ergebnisse plausibel sind.

5 Evaluierung des Mehrbenutzereditors

5.1 Ausgangssituation

Die Entwicklung des Prototyps war geprägt von umfangreichen Testläufen, die die Stabilität und Funktionalität des Systems immer wieder bestätigten. Als Textkorpus dienten verschiedene Zeitungsartikel, auf die über das Presseportal news.google.de⁷⁰ zugegriffen wurde. Die gewählten Zeitungsartikel stammten von unterschiedlichen Autoren bzw. Nachrichtenagenturen, behandelten aber das gleiche Thema. Diese waren aufgrund ihrer meist fehlerfreien und an der neuen Rechtschreibung orientierten Schreibweise besonders geeignet, um Übereinstimmungen in Sätzen zu identifizieren und hinsichtlich ihrer Relevanz zu bewerten. Einzelne Funktionen der morphologischen Analyse wie z. B. die fehlertolerante Ermittlung der Grundform oder die Kompositazerlegung wurden von Hand mit ausgewählten Wörtern getestet.

Die Tests beschränkten sich auf die technischen Komponenten. Inwieweit diese einen Benutzer während des Schreiben unterstützen können, konnte auf diese Weise nicht ermittelt werden. Aus diesem Grund wurden verschiedene Studienkonzepte erarbeitet, die das Ziel verfolgten, das System in verschiedenen Aufgabenstellungen und unter verschiedenen Arbeitsbedingungen zu beobachten und dadurch Aussagen über die unterstützenden Eigenschaften des Systems treffen zu können. Diese sollten anhand subjektiver Angaben des Benutzers und zum anderen objektiv durch protokollierte und ausgewertete Fakten getroffen werden.

⁷⁰<http://news.google.de>.

5.2 Organisation

5.2.1 Ausschreibung

Zunächst wurde die Durchführung von zwei Studien geplant. Ein Konzept für eine dritte Studie wurde nach den Erkenntnissen der ersten beiden Studien erstellt. Die Ausschreibung erfolgte über die Lernplattform Stud.IP der TU Clausthal, die Uni-internen Newsgroups *tuc.misc*⁷¹ und *tuc.announce*⁷² und die Mail-Verteiler für Studierende der Informatik/Wirtschaftsinformatik des Instituts für Informatik. Insgesamt meldeten sich 114 Studierende an, von denen 24 Teilnehmer für die drei Studien ausgewählt wurden. Bei der Auswahl der Teilnehmer wurde darauf geachtet, dass nicht an mehreren Studien teilgenommen werden konnte. Eine Befragung ergab, dass die Mehrzahl der Teilnehmer über Stud.IP auf die jeweilige Studie aufmerksam wurde. Dies lässt vermuten, dass die Fachrichtungen breit verteilt waren, da das System von der gesamten Universität genutzt wird.

Die ausgewählten Teilnehmer wurden per E-Mail über ihre Teilnahme informiert und darum gebeten, den Termin bzw. die Termine innerhalb einer Frist von einigen Tagen zu bestätigen. Bestätigte ein Teilnehmer den Termin nicht fristgerecht, wurde ein Ersatzkandidat angeschrieben, bis die benötigte Teilnehmerzahl erreicht war. Es wurden keine Ersatzkandidaten zur Studie eingeladen, die ggf. einen nicht erschienenen Teilnehmer hätten ersetzen können.

Die TU Clausthal zeichnet sich durch einen vergleichsweise hohen Anteil ausländischer Studierender von über 33%⁷³ aus. Der Bundesdurchschnitt deutscher Universitäten liegt bei 11,5%.⁷⁴ Da der zu testende Prototyp und die verwendeten Verfahren für die deutsche Sprache konzipiert und entwickelt wurden und die Aufgabenstellungen erforderten, Texte in deutscher Sprache verfassen zu können, waren die Voraussetzung zur Teilnahme an der Studie *„hervorragende Deutschkenntnisse“*. Dies wurde nach der ersten Studie durch die Formulierung

„Von den Teilnehmern wird die Beherrschung der deutschen Sprache in Wort und Schrift auf Muttersprachniveau erwartet.“

ersetzt, da sich in der 1. Studie bei den ausländischen Studierenden große sprachliche

⁷¹Newsgroup für Verschiedenes.

⁷²Newsgroup für Ankündigungen.

⁷³TU Clausthal, 2010, Stand: 15. November 2010.

⁷⁴statistica, 2011.

Defizite gezeigt hatten. Die Voraussetzung wurde nicht (z. B. durch eine Schreibprobe oder ein Telefonat) überprüft, stattdessen wurde erwartet, dass sich die Interessenten bzgl. ihrer Fähigkeiten selbst einschätzten.

5.2.2 Allgemeine Vorgaben

Die Teilnehmer wurden zu Beginn jeder Studie darauf hingewiesen, dass die Bewertung der sprachlichen Qualität auf Grundlage der neuen deutschen Rechtschreibung erfolgt. Zudem wurde ausdrücklich darauf hingewiesen, dass vor Satzzeichen kein Leerzeichen und nach Satzzeichen zur Kennzeichnung grammatischer Strukturen genau ein Leerzeichen stehen darf bzw. muss.

- Kathrin spielt mit Tom.Gerda spielt alleine.⁷⁵
- Kathrin spielt Fußball . Tom und Gerda spielen Tennis.⁷⁶

Der Hinweis war dem Umstand geschuldet, dass die eingesetzte Bibliothek (siehe Kap. 4.3.2) in beiden Fällen die Grenzen der Sätze nicht korrekt erkennt und als einen Satz identifiziert.

5.2.3 Laborbedingungen

Die Studien waren so konzipiert, dass die Teilnehmer während des Schreibens nicht direkt miteinander kommunizieren sollten. Deswegen wurden die Teilnehmer der 1. Studie während des 1. Termins auf verschiedene Räume verteilt. Da dies mit einem erheblichen Aufwand hinsichtlich Planung und Organisation verbunden war, wurde dies zum 2. Termin versuchsweise geändert und alle Teilnehmer verblieben im gleichen Raum. Da die Teilnehmer die Frage im Fragebogen

„Hatte die Art der räumlichen Aufteilung einen Einfluss auf Ihre Arbeitsweise?“

ausnahmslos negativ beantworteten, wurde in den folgenden Studien keine räumliche Trennung der Teilnehmer mehr vorgenommen. Stattdessen wurde darauf geachtet, dass während der Studien nicht miteinander kommuniziert wurde.

⁷⁵Im Netzjargon als Klempen bezeichnet.

⁷⁶Im Netzjargon als Plenken bezeichnet.

Das Computersystem, auf dem der Server des Schreibprogramms lief, verfügte über einen Pentium IV Single-Core mit 3.0 GHz und 2048 MB Arbeitsspeicher. Die Dimensionierung war ausreichend. An die Computersysteme für die Clientsoftware waren keine besonderen Ansprüche gestellt.

5.2.4 Auswertungsgrundlage

Zur Auswertung der Studien standen die gespeicherten Protokolle der Protokollfunktion (siehe Kap. 4.3.1), die handschriftlichen Notizen der Teilnehmer, sofern sie erlaubt waren, die ausgefüllten Fragebögen und die Texte samt Metadaten zur Verfügung. Die Metadaten umfassten folgende Werte:

- Anzahl geschriebener Sätze,
- Anzahl geschriebener Wörter⁷⁷,
- Anzahl indexierter Wörter.

Die Texte wurden durch den Diplomanden auf Grundlage des Duden von 2006⁷⁸ (24. Auflage) und Hinweisen im Internet zu Regeländerungen⁷⁹ seit Verabschiedung der neuen Rechtschreibung geprüft. Dabei wurden sämtliche Fehler, also auch Wiederholungs- und Folgefehler, gezählt und der Fehlerquotient (FQ)

$$FQ = \frac{\text{Anzahl Fehler} \cdot 100}{\text{Anzahl Wörter}}$$

berechnet. Dieser gibt die Anzahl der Rechtschreib- und Grammatikfehler auf 100 Wörter an.

Die Fragebögen enthielten allgemeine Fragen zur Person sowie studienspezifische Fragen und Fragen zu den Funktionen des Prototyps. Neben offenen Fragen waren in den Fragebögen überwiegend Skalafragen mit Werten zwischen 1 und 5 enthalten. Eine 1 entsprach dem schlechtesten Wert, eine 5 entsprach dem besten Wert. Außerdem enthielt jeder Fragebogen zehn Fragen zur Benutzerfreundlichkeit. Diese waren der *System Usability Scale* von John Brooke⁸⁰ entnommen und ins Deutsche übersetzt worden. Der Test misst anhand

⁷⁷Ein Vergleich mit der Zählfunktion in Microsoft Word 2007 ergab eine Abweichung von $\leq 1\%$.

⁷⁸Dudenredaktion, 2006.

⁷⁹Schitteck, 2010.

⁸⁰Brooke, 1986.

einer Skala von 0 (schlechtester Wert) bis 100 (besten Wert) die Benutzerfreundlichkeit eines Systems. Die Fragebögen sind als PDF-Datei auf der beiliegenden DVD-ROM zu finden.

5.3 1. Laborstudie

5.3.1 Ausgangspunkt

Nach der erfolgreichen Pilotierung des Prototyps sollte das Schreibprogramm unter realen Bedingungen getestet werden. Dabei sollte die Studie auf die Erkenntnisse von bereits zuvor am Institut für Informatik durchgeführten Studien⁸¹ aufbauen.

Zunächst sollte beobachtet werden, ob die Indexierungs- (siehe Kap. 4.3.2) und Vergleichsfunktion (siehe Kap. 4.3.3) des Servers einwandfrei funktionieren. Dabei standen weniger die unterstützenden, sondern mehr die technischen Eigenschaften im Vordergrund. Außerdem sollten mögliche Schwächen in der Benutzerfreundlichkeit der Client-Software identifiziert werden.

5.3.2 Studienbeschreibungen

Die Studie fand vom 13. bis 19. April 2011 an insgesamt vier Terminen statt. Eingeladen waren 8 Studierende, die auf zwei Gruppen aufgeteilt waren. Geplant war, dass jede Gruppe aus je zwei weiblichen und zwei männlichen Teilnehmern besteht. Aufgrund von zwei Absagen, die nur durch Einladung von zwei männlichen Studierenden kompensiert werden konnten, bestand Gruppe I aus vier männlichen Teilnehmern. In beiden Gruppen waren jeweils zwei deutsche und zwei ausländische Studierende.

Zu Beginn wurde die Aufgabe und der Ablauf der Studie erklärt. Die Aufgabe bestand darin, zu einem Thema eine Geschichte aus vier unterschiedlichen Perspektiven zu schreiben. Zunächst hatten die Teilnehmer 20 Minuten Zeit, sich in der Gruppe abzusprechen und Notizen zu machen. Anschließend wurde ihnen ein Computer mit dem Schreibprogramm zugeteilt, an dem sie ihren Teil der Geschichte schreiben konnten. Für das Schreiben hatten die Teilnehmer 80 Minuten Zeit. Währenddessen war die Kommunikation zwischen den Teilnehmern unterbunden, dafür war es möglich, auf die Texte der anderen Teilnehmer zuzugreifen und mitzulesen (siehe Kap. 4.3), ohne selbst Änderungen an diesen vornehmen zu können. Während des zweiten Termins hatten die Teilnehmer zusätzlich die Möglichkeit, Ergebnisse der Vergleichsfunktion aufzurufen. Das Schreibprogramm wurde den Teilnehmern kurz anhand von zwei Screenshots vorgestellt. Auf die einzelnen Funktionen und die Bedienung wurde nicht im Detail eingegangen.

⁸¹Olivier und Pinkwart, 2009.

Themen

Die Themen sollten so beschaffen sein, dass sich die Teilnehmer einerseits eine klare Vorstellung machen konnten, andererseits aber auch Freiräume für Kreativität gegeben waren. Die Perspektiven wurden so gewählt, dass sie sich untereinander verbinden ließen.

Thema	A	B
Titel	Popkonzert	Beachvolleyballturnier
Perspektiven	Sänger (m/w) Fan (m/w) Manager (m/w) Ordner (m/w)	Spieler (m/w) Gegenspieler (m/w) Zuschauer (m/w) Stadionsprecher (m/w)

Tab. 5.1: Themenauswahl und Perspektiven
Quelle: Eigene Darstellung

Das Geschlecht (männlich/weiblich) der Perspektiven wurde nicht vorgegeben.

Variation

Um vergleichen zu können, ob die Vergleichsfunktion Auswirkungen auf das Ergebnis hat und ausschließen zu können, dass die Themen oder die Zusammensetzung der Gruppen einen Einfluss auf diese haben, variierten die Bedingungen. Dadurch konnte ein Kreuzvergleich zwischen den Gruppen durchgeführt werden.

	Gruppe I		Gruppe II	
	<i>Thema</i>	<i>Vergleichsergebnisse</i>	<i>Thema</i>	<i>Vergleichsergebnisse</i>
1. Termin	A	nein	B	nein
2. Termin	B	ja	A	ja

Tab. 5.2: Variation der Bedingungen
Quelle: Eigene Darstellung

Die Mitlesefunktion war während der gesamten Studie aktiviert, so dass die Teilnehmer auf sämtliche Dokumente der Gruppe lesend zugreifen konnten.

Gruppeneinteilung

Gruppe	Benutzerkennungen
I	01, 02, 03, 04
II	09, 10, 11, 12

Tab. 5.3: Zuordnung Benutzerkennung zu Gruppe

Besondere Vorkommnisse

Die Protokollfunktion (siehe Kap. 4.3.1) hat während des ersten Termins von Gruppe I aufgrund eines bis dahin unerkannten Fehlers nicht funktioniert. Die Benutzeraktionen wurden nur teilweise protokolliert. Der Fehler konnte direkt nach der Studie identifiziert und behoben werden, so dass die Protokollfunktion beim ersten Termin von Gruppe II uneingeschränkt sämtliche Benutzeraktionen aufzeichnete.

Während des zweiten Termins von Gruppe I kam es bei Teilnehmer 02 zu einem Verbindungsabbruch zum Server. Vermutet wird ein lokaler Netzwerkfehler, da im Serverprotokoll kein Fehler auftauchte und auch die anderen Teilnehmer nicht betroffen waren. Der Verbindungsabbruch hatte zur Folge, dass Teile des im Client zwischengespeicherten Ereignisprotokolls verloren gingen. Das Dokument selbst war nicht von dem Datenverlust betroffen, da es zusätzlich auf dem Server zwischengespeichert wurde. Der Teilnehmer konnte sich sofort wieder mit dem Client am Server anmelden und ohne sichtbaren Datenverlust an seinem Dokument weiterarbeiten.

5.3.3 Auswertung und Interpretation der Ergebnisse

Studiendesign

Die Frage im Fragebogen, ob die Zeit zur Planung und zum Schreiben angemessen war, wurde fast ausschließlich positiv beantwortet. Nur zwei Teilnehmer hätten sich geringfügig mehr Zeit gewünscht. Es konnte allerdings beobachtet werden, dass einige Teilnehmer während der Durchführungsphase des ersten Termins zum Ende hin zeitlich etwas in Bedrängnis gerieten. Während des zweiten Termins hatten die Teilnehmer ihre Zeitplanung allerdings entsprechend angepasst, so dass die Zeit ausreichend war. Bei der Frage, welches Thema besser gefallen hätte, gab es eine leichte Tendenz zu Thema B.

Planungsphase: Inhalt, Interaktion und soziale Aspekte

Die Planungsphase war in beiden Gruppen ähnlich strukturiert. Zunächst wurden die Perspektiven verteilt, anschließend wurde der Rahmen der Geschichte anhand von W-Fragen (Wann?, Wo?, Wer?, ...) grob abgesteckt. Erst danach wurden Details besprochen und inhaltliche Verbindungen zwischen den Perspektiven hergestellt.

Große Unterschiede zeigten sich in der Art und Weise der Interaktion. Gruppe I wurde von den beiden deutschen Teilnehmern dominiert, die die inhaltliche Gestaltung durch Fragen und Diskussionen vorantrieben. Insgesamt waren die Beteiligung und die Motivation aller Teilnehmer in dieser Gruppe sehr hoch, ungeachtet sprachlicher Defizite der ausländischen Teilnehmer. Die zweite Gruppe war sowohl während des ersten als auch des zweiten Termins zunächst etwas verhalten. Es fehlte eine bestimmende Person, die die anderen Teilnehmer animierte und die Beteiligung an der Planung und die Motivation förderte. Dies spiegelt sich auch in der Anzahl der Notizen wider, die bei Gruppe I sichtbar höher ausfällt als bei Gruppe II. Die Planung war überwiegend fakten- und weniger handlungsorientiert. Es standen weniger Handlungen, die mehrere Perspektiven betreffen, im Vordergrund, sondern mehr Fakten wie Uhrzeit, Ort, verkaufte Tickets etc.

Während des zweiten Termins wurde deutlich, dass die Teilnehmer auf die Erfahrungen während der ersten Planungs- und Durchführungsphase aufbauten. So wurde von beiden Gruppen während der zweiten Planungsphase geäußert, dass durch die Mitlesefunktion Handlungsstränge, die die eigene und eine andere Perspektive betreffen, erkannt und in den eigenen Teil der Geschichte übernommen werden können. Konkret wurde gesagt, dass Handlungen während des Schreibens ausgedacht werden sollen, die dann mithilfe der Mitlesefunktion übernommen werden.

Indexierung

Die Indexierung der Texte hat während der gesamten Studie ausnahmslos funktioniert. Im Schnitt ergibt sich eine Indexierungsquote von fast 46% (siehe Abb. 5.1 (b)) der Wörter eines Satzes. Auffällig sind die Unterschiede zwischen den Gruppen in der Durchschnittszahl indexierter Wörter pro Satz (siehe Abb. 5.1 (a)). Gruppe II hat beim zweiten Termin kürzere Sätze verfasst als Gruppe I. Ein Zusammenhang zwischen Satzlänge und dem Informationsgehalt kann nicht eindeutig festgestellt werden, da die Unterschiede beim Informationsgehalt innerhalb und zwischen den Gruppen eher gering sind.

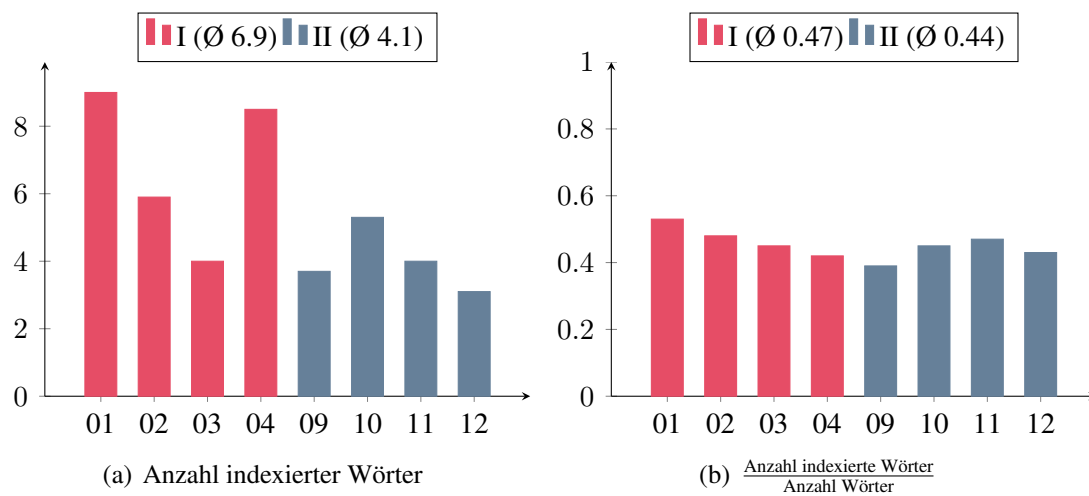


Abb. 5.1: Ø Anzahl indexierter Wörter und Informationsgehalt pro Satz
Quelle: Eigene Darstellung

Bewertung des Prototyps

Der Prototyp hat abgesehen von kleineren, von den Teilnehmern unbemerkten Fehlern funktioniert. Die Bedienung des Schreibprogramms wurde zu Beginn nicht ausführlich erklärt, dennoch schienen die Teilnehmer während der Studie keinen größeren Schwierigkeiten ausgesetzt, ihren Text zu schreiben. Dies bestätigt die Auswertung des SUS-Tests, dargestellt in Abb. 5.2, der im Durchschnitt einen Wert von über 82 Punkten ermittelt hat.

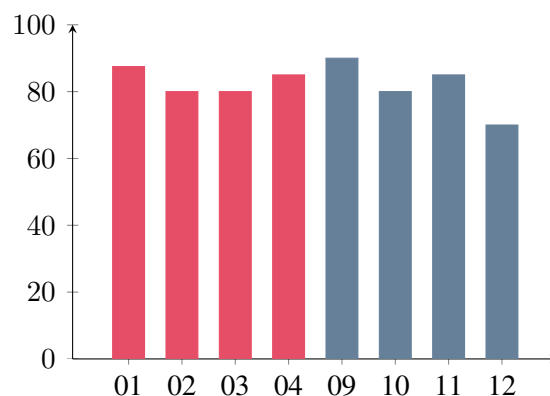


Abb. 5.2: Auswertung des SUS-Tests
Quelle: Eigene Darstellung

Die Nützlichkeit der Mitlesefunktion wurde in beiden Gruppen hoch bewertet. Auf einer Skala von 1 bis 5 ergab sich für Gruppe I ein Wert von 4,75, für Gruppe II ein Wert von 4,25.

Die Nützlichkeit der Vergleichsfunktion hingegen wurde sehr unterschiedlich bewertet. Insgesamt nutzten drei Teilnehmer aus Gruppe I und zwei Teilnehmer aus Gruppe II die Funktion und konnten somit eine Bewertung abgeben. Für Gruppe I ergab sich auf einer Skala von 1 bis 5 im Durchschnitt ein Wert von 3,67, für Gruppe II hingegen nur 2,5.

Festzuhalten ist, dass sich die Nutzung der Mitlese- und der Vergleichsfunktion nicht ausschließen. Beide Funktionen wurden von einigen Teilnehmern parallel genutzt.

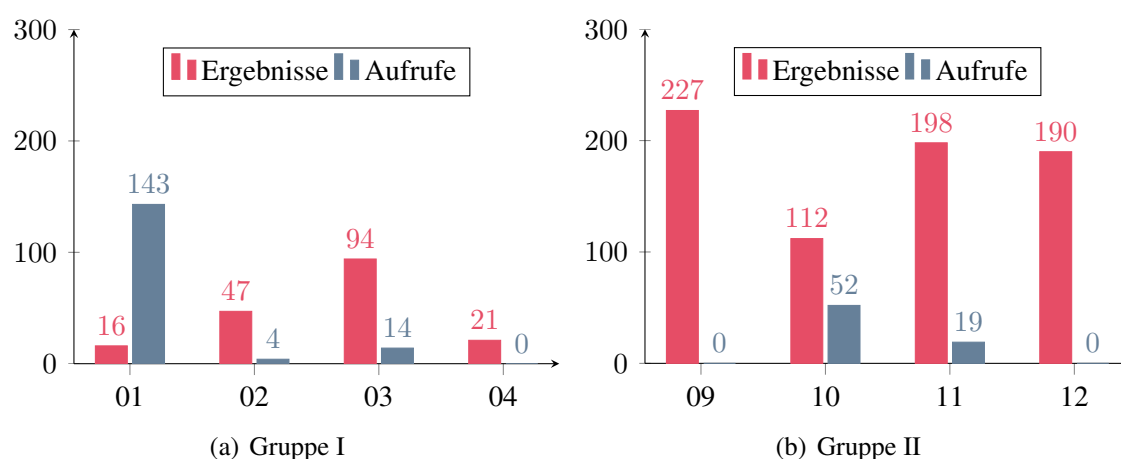


Abb. 5.3: Anzahl angezeigter Ergebnisse und Ergebnisaufrufe
Quelle: Eigene Darstellung

[**Anm.:** Die Anzahl der Ergebnisaufrufe von Teilnehmer 02 ist aufgrund des beschriebenen Verbindungsabbruchs (siehe Kap. 5.3.2) unvollständig.]

Eingestellt war für die Anzeige der Vergleichsergebnisse eine Mindestrelevanz von 20%. Abb. 5.8 (a) gibt die Anzahl der Vergleichsergebnisse, die zum Ende der Durchführungsphase angezeigt wurden, wieder. Aufgrund der Dynamik der Vergleichsfunktion kann die Anzahl während der Durchführungsphase höher gewesen sein, wahrscheinlich ist allerdings, dass sie währenddessen niedriger war und immer mehr zunahm.

Es kann ein Zusammenhang zwischen der durchschnittlichen Satzlänge und der Anzahl der angezeigten Vergleichsergebnisse hergestellt werden. Kurze Sätze bedeuteten viele Vergleichsergebnisse, da sich die wenigen indexierten Wörter aufgrund der Themengleichheit in vielen anderen Dokumenten wiederfanden. Lange Sätze mit vielen indexierten Wörtern führten hingegen zu wenigen Vergleichsergebnissen. Verstärkt wurde dies in Gruppe II, da alle Teilnehmer im Durchschnitt kurze Sätze schrieben.

Auffällig ist die Tatsache, dass Teilnehmer 01 die meisten Vergleichsergebnisse aufrief,

obwohl dieser, bedingt durch lange Sätze, die wenigsten Vergleichsergebnisse angezeigt bekam. Die Teilnehmer 10 und 11 aus Gruppe II hingegen riefen nicht so viele Vergleichsergebnisse auf, wie ihnen angezeigt wurde.

Arbeitsweise

Während der Studie konnte keine klare Trennung zwischen Schreib- und Kontrollphase beobachtet werden. Vielmehr wurde während der gesamten Durchführungsphase zwischen dem eigenen Text und den Texten der anderen Teilnehmer gewechselt. Es konnte kein ununterbrochener Schreibprozess beobachtet werden, vielmehr wurden während des Schreibens von den Teilnehmern immer wieder Korrekturen am eigenen Text vorgenommen. Dies belegt auch die Auswertung der Ereignisprotokolle, die Sprünge im Text protokolliert hat. Inwieweit die Mitlese- bzw. Vergleichsfunktion Anlass für diese Änderungen gegeben hat, ist gar nicht bzw. schwer nachvollziehbar. Alle drei Teilnehmer aus Gruppe I, die die Vergleichsfunktion nutzten, gaben an, Änderungen am Dokument aufgrund dieser Funktion vorgenommen zu haben. Die angegebene Anzahl konnte allerdings nur in einem Fall anhand des Ereignisprotokolls nachvollzogen werden. Teilnehmer 01 gab an, mehr als 9 Änderungen vorgenommen zu haben. Die Auswertung des Ereignisprotokolls ergab allerdings, dass in keinem Fall ein Zusammenhang zwischen dem Ursprungssatz, dessen Vergleichsergebnis aufgerufen wurde, und der Änderung am Dokument hergestellt werden konnte.

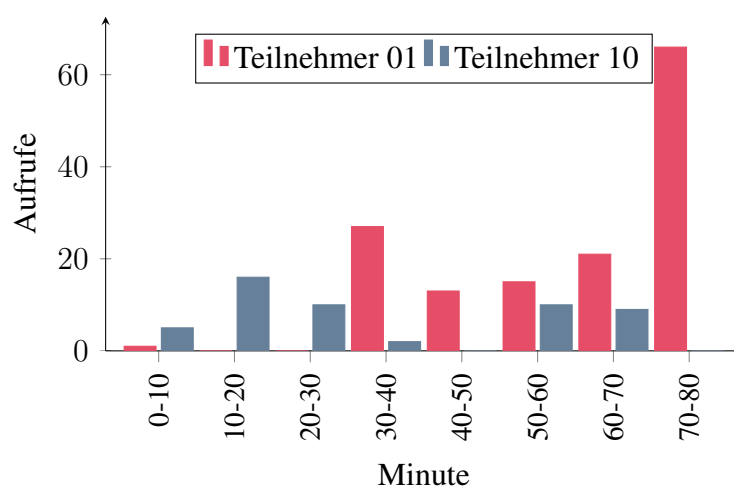


Abb. 5.4: Zeitprofil der Ergebnisaufrufe von Teilnehmer 01 und 10
Quelle: Eigene Darstellung

Die zeitliche Auswertung der Ergebnisaufrufe von Teilnehmer 01 und Teilnehmer 10,

dargestellt in Abb. 5.4, bestätigt, dass es zwischen Schreiben und Kontrolle keine klare Trennung gab. Teilnehmer 01 nutzte die Vergleichsfunktion verstärkt in den letzten 10 Minuten. Die Auswertung des Ereignisprotokolls ergab, dass zu diesem Zeitpunkt nur noch wenige Änderungen am Dokument vorgenommen wurden.

Bewertung der Texte

Die Geschichten von Gruppe I wiesen jeweils eine Inkonsistenz auf. Die Geschichten von Gruppe II sind inhaltlich konsistent.

Die Überprüfung der Rechtschreibung und Grammatik ergab, dass es beim 2. Termin eine Tendenz zu mehr Fehler gab.

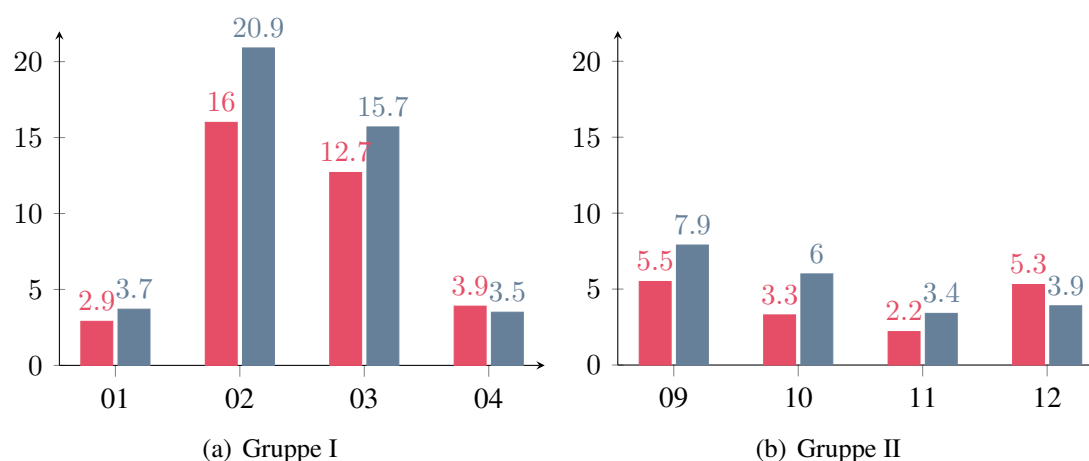


Abb. 5.5: Fehlerquotienten
Quelle: Eigene Darstellung

Auswertung der offenen Fragen

Auf die Frage, ab welcher Relevanz die Vergleichsergebnisse hilfreich waren, bewegten sich die Angaben bei vier Antworten zwischen 30% und 60%. Im Mittel lag sie bei ca. 50%.

Zu möglichen Verbesserungsmöglichkeiten im Schreibprogramm befragt, bezogen sich drei Antworten auf die Mitlesefunktion. Dort wurde die Möglichkeit gewünscht, die Texte nebeneinander anordnen zu können. Ein Teilnehmer wünschte sich die Möglichkeit, andere Teilnehmer auf inhaltliche Widersprüche hinweisen zu können. Außerdem hat ein

Teilnehmer die Möglichkeit, die Texte der anderen mit eigenen Suchbegriffen durchsuchen zu können, vermisst.

5.3.4 Fazit

Die gestellte Aufgabe und die damit verbundene Problematik, eine in sich schlüssige Geschichte verteilt auf mehrere Personen zu schreiben, wurde von den Teilnehmern ohne Erklärung sofort verstanden. Die Umsetzung weist zwar quantitative und qualitative Unterschiede auf, insgesamt kann aus der Beteiligung und Motivation der Teilnehmer beider Gruppen ein positives Fazit gezogen werden.

Die Vergleichsfunktion zeigt eine deutliche Schwäche. Bei kurzen Sätzen, unabhängig vom Informationsgehalt, werden viele mögliche Übereinstimmungen in anderen Sätzen gefunden. Dadurch nimmt die Anzahl der angezeigten Vergleichsergebnisse erheblich zu. Es ist nicht auszuschließen, dass die Anzahl der angezeigten Ergebnisse einen Einfluss auf die Nutzung und die Bewertung der Nützlichkeit der Vergleichsfunktion hat. Teilnehmer 10 und 11 hatten mehr Ergebnisse als Ergebnisaufrufe. Dass die zur Verfügung stehende Zeit eine stärkere Nutzung der Funktion verhinderte, kann ausgeschlossen werden. Wahrscheinlicher ist, dass die Sätze, zu denen ein Vergleichsergebnis aufgerufen wurde, nur in wenigen Wörtern bzw. Bedeutungen übereinstimmten und somit für den Teilnehmer nicht relevant waren.

Die Anhebung der Mindestrelevanz auf die im Fragebogen angegebenen 50% hätte möglicherweise Abhilfe schaffen können. Die Berücksichtigung in der Gewichtung unterschiedlicher Satzlängen (siehe Kap. 4.3.3) konnte diesen Umstand nicht kompensieren, da alle Sätze in Gruppe II relativ kurz waren.

Auffällig sind die hohen Fehlerquotienten der ausländischen Teilnehmer 02 und 03, dargestellt in Abb. 5.10 (a). Die hohe Fehlerquote spiegelt sich zwar nicht im Informationsgehalt (siehe Abb. 5.1 (b)) wider, allerdings ermittelt die fehlertolerante Lemmatisierung u. U. viele verschiedene Grundformen (siehe Tab. 4.3). Dies kann die Anzahl und die Qualität der Vergleichsergebnisse (siehe Abb. 5.8 (a)) ähnlich wie sehr kurze Sätze negativ beeinflussen.

Teilnehmer 01, der der Vergleichsfunktion eine Nützlichkeit von 4 bescheinigte, bekam zwar die wenigsten Vergleichsergebnisse angezeigt, nutzte die Funktion aber am intensivsten. Er nahm nach eigene Angaben mehr als 9 Änderungen vor, die im Ereignisprotokoll

nicht nachvollzogen werden können. Dies lässt die Vermutung zu, dass Änderungen aufgrund eines Vergleichsergebnisses nicht zwangsläufig den Ursprungssatz betreffen müssen.

Deutlich wurde, dass die Mitlesefunktion unter den gegebenen Bedingungen (keine direkte Kommunikation) eine wichtige Funktion darstellte. Die Nutzung beschränkte sich dabei nicht auf die Kontrolle der inhaltlichen Konsistenz. Sie wurde auch genutzt, um nachträglich Planungen zu konkretisieren.

Drei Teilnehmer nutzten die Vergleichsfunktion nicht und zogen der automatischen die intellektuelle Indexierung (siehe Kap. 2.1) bzw. das Mitlesen vor. Dies kann unter anderen Voraussetzungen (größere Dokumentensammlung, weniger Zeit) zumindest problematisch werden.

5.4 2. Laborstudie

5.4.1 Ausgangspunkt

Die erste Studie deckte zwar Schwächen der Vergleichsfunktion auf, zeigte aber auch positive Ergebnisse. Daher sollte in einer zweiten Studie der Fokus auf der Unterstützung der Teilnehmer liegen. Hierfür wurde eine Aufgabenstellung gewählt, die die qualitativen bzw. unterstützenden Eigenschaften der Vergleichsfunktion (siehe Kap. 4.3.3) aufzeigen sollte.

5.4.2 Änderungen am Prototyp

Aufgrund der Ergebnisse der ersten Studie wurden technische Änderungen am Prototyp vorgenommen. Diese bezog sich auf die schlechten Ergebnisse, die bei sehr kurzen Sätzen mit wenigen indexierten Wörtern auftraten. Der Prototyp wurde dahingehend erweitert, dass eine Mindestanzahl indexierter Wörter, die für einen Vergleich benötigt werden, festgelegt werden konnte.

Beim ersten Termin wurde auf die Festlegung einer Mindestanzahl indexierter Wörter verzichtet, da die Teilnehmer keine Vergleichsergebnisse angezeigt bekamen. Nachdem die Daten des ersten Termins ausgewertet waren, konnte festgestellt werden, dass das Problem kurzer Sätze in dieser Studie nicht zum Tragen kam, daher wurde auch beim zweiten Termin darauf verzichtet, eine Mindestanzahl festzulegen.

5.4.3 Studienbeschreibungen

Die Studie fand vom 04. bis 11. Mai 2011 an insgesamt vier Terminen statt. Eingeladen waren 8 Studierende, die auf zwei Gruppen aufgeteilt waren. Jede Gruppe bestand aus je zwei weiblichen und zwei männlichen Teilnehmern. In beiden Gruppen waren ausschließlich deutsche Studierende.

Zu Beginn wurde den Teilnehmern die Aufgabe und der Ablauf der Studie erklärt. Die Aufgabe bestand darin, den Inhalt eines Hörspiels in eigenen Worten wiederzugeben. Hierfür wurde den Teilnehmern ein ca. 9 Minuten langes Hörspiel mit kurzer Unterbrechung zweimal vorgespielt. Es konnten sich währenddessen Notizen machen. Anschlie-

ßend hatten die Teilnehmer 50 Minuten Zeit, den Inhalt wiederzugeben. Ihnen wurde ein Computer mit dem Schreibprogramm zugeteilt, an dem sie schreiben konnten. Während der Durchführungsphase war die Kommunikation zwischen den Teilnehmern unterbunden. Beim zweiten Termin hatten die Teilnehmer zusätzlich die Möglichkeit Vergleichsergebnisse aufzurufen. Das Schreibprogramm wurde den Teilnehmern ausführlich anhand von zwei Screenshots erklärt. Auf Fragen der Teilnehmer wurde dabei näher eingegangen.

Hörspiele

Die Hörspiele waren so beschaffen, dass sie akustisch verständlich und inhaltlich nachvollziehbar waren. Es wurden bewusst klassische Werke gewählt, da in diesen teils nicht mehr zeitgemäße Begriffe und Formulierungen verwandt wurden. Dies sollte dazu führen, dass sich die Teilnehmer inhaltlich mit dem Hörspiel auseinandersetzen, den Text nicht Wort für Wort übernehmen und sich somit eine sprachliche Vielfalt ergibt.

	Hörspiel A	Hörspiel B
Titel	Wie Eulenspiegel dem Pfarrer sein Pferd abschwatzte	Die weiße Schlange
Autor(en)	Hermann Bote	Brüder Grimm
Länge	8:57 min.	9:35 min.

Tab. 5.4: Ausgewählte Hörspiele

Variation

Um vergleichen zu können, ob die Unterstützung des Teilnehmers Auswirkungen auf das Ergebnis hat und ausschließen zu können, dass die Geschichten einen Einfluss auf diese haben, variierten die Bedingungen. Dadurch konnte ein Kreuzvergleich zwischen den Gruppen durchgeführt werden.

	Gruppe I		Gruppe II	
	<i>Geschichte</i>	<i>Vergleichsergebnisse</i>	<i>Geschichte</i>	<i>Vergleichsergebnisse</i>
1. Termin	A	nein	B	nein
2. Termin	B	ja	A	ja

Tab. 5.5: Variation (2. Studie)
Variation der Bedingungen

Die Mitlesefunktion (siehe Kap. 4.3) war in dieser Studie deaktiviert, da die Teilnehmer

nicht auf die Dokumente der anderen Teilnehmer zugreifen sollten.

Gruppeneinteilung

Gruppe	Benutzerkennungen
I	05, 06, 07, 08
II	13, 14, 15, 16

Tab. 5.6: Zuordnung Benutzerkennung zu Gruppe

Besondere Vorkommnisse

Teilnehmer 15 aus Gruppe II blieb dem zweiten Termin ohne Angabe von Gründen fern.

5.4.4 Auswertung und Interpretation der Ergebnisse

Wiedergabephase

Die Teilnehmer machten sich während des ersten Abspielens umfangreiche Notizen, die während des zweiten Abspielens nur noch ergänzt wurden. Die Durchsicht der Notizzettel ergab, dass nur vereinzelt Korrekturen an den Notizen vorgenommen wurden.

5.4.5 Durchführungsphase

Während der Durchführung konnte bei beiden Gruppen zu Beginn eine hohe Schreibgeschwindigkeit festgestellt werden, die nach ca. 30 Minuten mehr und mehr abnahm. Weder beim ersten Termin noch beim zweiten Termin wurden die 50 Minuten von irgendeinem Teilnehmer voll ausgeschöpft. Die ersten Teilnehmer stellten bereits nach ca. 35 Minuten das Schreiben ein. Die Notizen wurden von allen Teilnehmern intensiv genutzt.

5.4.6 Hörspiele

Bei der Frage, welches Hörspiel besser gefallen hätte, waren die Teilnehmer unentschieden. Abb. 5.6 anhand der geschriebenen Wörter ist kennbar, dass Geschichte A länger

bzw. inhaltlich umfangreicher war.

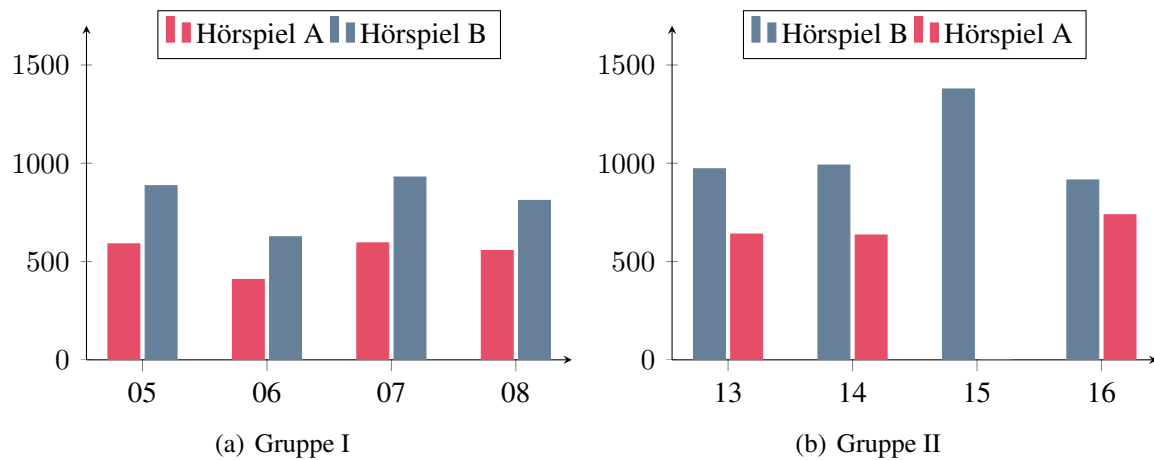


Abb. 5.6: Anzahl geschriebener Wörter
Quelle: Eigene Darstellung

Bewertung des Prototyps

Die Bedienung des Schreibprogramms wurde zu Beginn ausführlich erklärt. Im Vergleich zur ersten Studie nahm der Durchschnittswert des SUS-Test ab, er lag bei 76 Punkten. Bemängelt wurde in zwei Fällen die schlechte Integration der Vergleichsfunktion ins Systems.

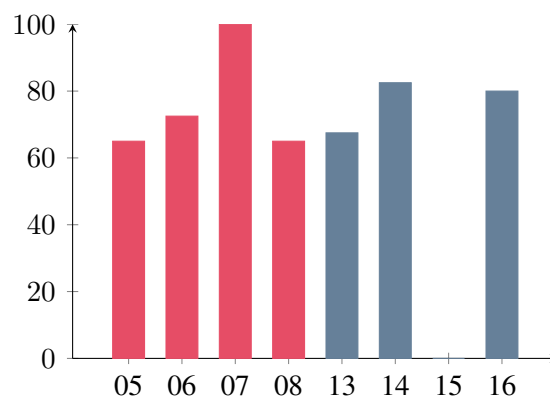


Abb. 5.7: Auswertung des SUS-Tests
Quelle: Eigene Darstellung

Die Vergleichsfunktion wurde von allen Teilnehmer genutzt. Die Bewertung in Gruppe I lag bei 2,5, wobei einmal eine sehr hohe und einmal überhaupt keine Nützlichkeit bescheinigt wurde. Gruppe II gab im Durchschnitt eine neutrale Bewertung mit 3,0.

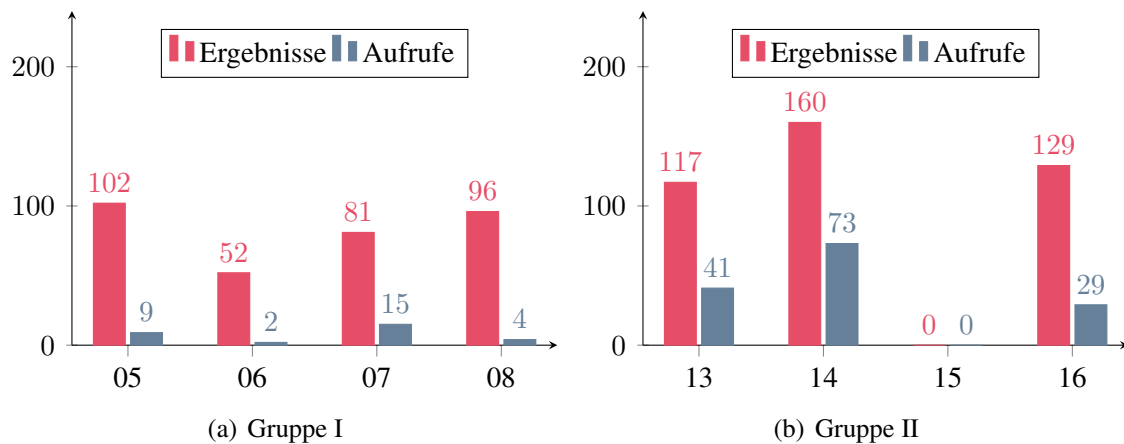


Abb. 5.8: Anzahl angezeigter Ergebnisse und Ergebnisaufrufe
Quelle: Eigene Darstellung

Auffällig ist, dass in Gruppe II mehr Ergebnisse angezeigt wurden als in Gruppe I, obwohl Gruppe I aufgrund der Beschaffenheit der Geschichte, mehr Text schrieb und Gruppe II aus einem Teilnehmer weniger bestand, der hätte Vergleichssätze beisteuern können. Die Auswertung ergab, dass dies in Gruppe II nicht mit kurzen Sätzen erklärt werden konnte. Der Schnitt lag bei ca. 15 Wörtern pro Satz, wovon ca. 50% indiziert wurden. Eine Vermutung ist, dass die Texte der Teilnehmer aus Gruppe II sich inhaltlich ähnlicher waren und somit mehr Übereinstimmungen gefunden wurden.

Arbeitsweise

Alle Teilnehmer nutzten in den ersten 5 Minuten das erste Mal die Vergleichsfunktion. Anschließend war die Nutzung sehr unterschiedlich. Teilnehmer 06 rief das zweite und letzte Vergleichsergebnis nach ca. 30 Minuten auf. Danach nutzte er die Funktion nicht mehr. Teilnehmer 14 rief die meisten Vergleichsergebnisse auf. Die Mehrzahl (≥ 60) davon rief er erst in den letzten 20 Minuten auf, Änderungen nahm er zu diesem Zeitpunkt nur noch vereinzelt vor. In diesem Fall kann von einer Trennung zwischen Schreiben und Kontrolle gesprochen werden. Bei anderen Teilnehmern war die Trennung nicht eindeutig erkennbar.

Bei allen Teilnehmern konnte beobachtet werden, dass der Notizzettel intensiv genutzt wurde. Schreibpausen dienten dazu, die Notizen zu lesen oder Vergleichsergebnisse aufzurufen.

Bewertung der Texte

Die Texte aller Teilnehmer wiesen Inkonsistenzen bzw. Auslassungen auf. Sachverhalte wurde falsch dargestellt oder hinzuerfunden bzw. für das Verständnis wichtige Sachverhalte wurden nicht erwähnt.

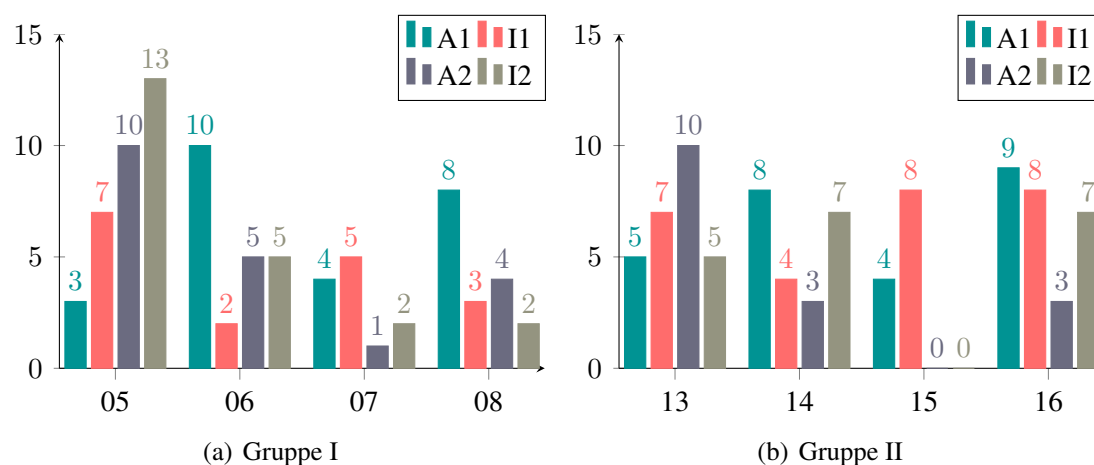


Abb. 5.9: Anzahl Auslassungen und Inkonsistenzen
Quelle: Eigene Darstellung

Abb. 5.9 stellt die Auslassungen (A1) und Inkonsistenzen (I1) der Texte des 1. Termins und die Auslassungen (A2) und Inkonsistenzen (I2) der Texte des 2. Termins dar. Es ist keine klare Tendenz erkennbar, dass die Vergleichsfunktion, die während des 2. Termins genutzt werden konnte, die Anzahl der Auslassungen und Inkonsistenzen beeinflusst hätte. Einzig bei Teilnehmer 07 ist eine deutliche Verbesserung zu erkennen. Im Schnitt ergaben sich beim zweiten Termin 10,5 Inkonsistenzen bzw. Auslassungen für Gruppe I und 11,7 Inkonsistenzen bzw. Auslassungen für Gruppe II.

Es besteht kein Zusammenhang zwischen Textlänge (siehe Abb. 5.6) und der Qualität der Texte. Viele Wörter bedeuteten nicht zwangsläufig, dass der Text wenig Auslassungen oder umgekehrt viele Inkonsistenzen enthielt.

Die Fehlerquotienten bewegten sich, abgesehen von zwei Ausreißern, auf dem Niveau der deutschen Teilnehmer der ersten Studie. Die Texte wiesen überwiegend einfache Rechtschreib- und Grammatikfehler auf.

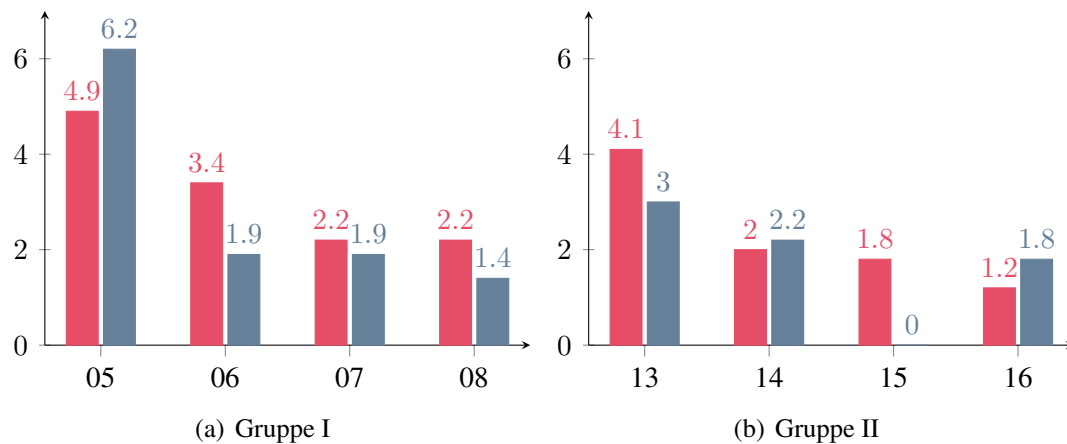


Abb. 5.10: Fehlerquotienten
Quelle: Eigene Darstellung

Auswertung der offenen Fragen

Auf die Frage, ab welcher Relevanz die Vergleichsergebnisse hilfreich waren, bewegten sich die Angaben bei vier Antworten zwischen 45% und 60%. Im Mittel lag sie bei ca. 52%.

Nach Verbesserungsmöglichkeiten befragt bemängelte ein Teilnehmer, dass die Anzeige des Vergleichssatzes unterhalb des Textfenster irritierend wirke. Drei Teilnehmer hätten sich eine andere Möglichkeit, die Vergleichsergebnisse zu sortieren, gewünscht. Zwei davon sagten ausdrücklich, dass sich die Sätze nach Auftreten im Dokument sortieren lassen sollten. [**Anm.:** Zu diesem Zeitpunkt war es lediglich möglich, Ursprungs- bzw. Vergleichssätze alphabetisch zu sortieren.] Ein weiterer Teilnehmer hätte sich die Anzeige eines größeren Kontexts, also die Anzeige des Vorgänger- bzw. Nachfolgesatzes eines Vergleichssatzes, gewünscht.

Vergleichsergebnisse

Die Relevanz der Vergleichsergebnisse beruht auf inhaltlichen und semantischen Übereinstimmungen. Inwieweit ein Ergebnis allerdings für einen Teilnehmer relevant war, beruhte auf subjektiven Faktoren. Die Angabe, ab wann ein Vergleichsergebnis relevant war, sagte nicht aus, dass alle Ergebnisse ab dieser Relevanz für den Teilnehmer relevant waren. Anhand eines Beispiels soll demonstriert werden, wie die angezeigten Vergleichsergebnisse beschaffen waren. Dazu wurde von jedem Teilnehmer aus Gruppe I ein Satz so

ausgewählt, dass sie in etwa den selben Sachverhalt wiedergeben.

Textpassage im Hörspiel:

„Als er schon durch drei Königreiche gewandert war und abends in einen Wald kam, setzte er sich unter einen Baum und wollte schlafen; da hörte er in den Ästen ein Geräusch, und ein goldener Apfel fiel in seine Hand. Zugleich flogen drei Raben zu ihm herab, setzten sich auf seine Knie und sagten: „Wir sind die drei jungen Raben, die du vom Hungertod errettet hast; als wir groß geworden waren und hörten, dass du den goldenen Apfel suchtest, so sind wir über das Meer geflogen bis ans Ende der Welt, wo der Baum des Lebens steht, und haben dir den Apfel geholt.“

Teilnehmer 05:

„Plötzlich kamen 3 junge Raben mit dem goldenen Apfel zu ihm, sie waren über das Meer und bis an das Ende der Welt geflogen, um ihm zu helfen.“

Teilnehmer 06:

„Doch als er sich zum Schlafen unter einen Baum legt, kommen die drei geretteten Raben vorbei und bringen ihm den Apfel vom Ende der Welt, denn sie hatten von seinem Unglück erfahren.“

Teilnehmer 07:

„Sie erzählten ihm, dass sie, als sie groß geworden waren und von seiner misslichen Lage gehört hatten, über alle Meere geflogen seien, bis an das Ende der Welt, um dort vom Baum des Lebens einen Apfel für ihn zu pflücken.“

Teilnehmer 08:

„Sie sprachen zum Jüngling, dass sie bis ans Ende der Welt geflogen seien, um ihm den Apfel vom Baum des Lebens zu beschaffen.“

	05	06	07	08
05	-	33%	31%	21%
06	42%	-	26%	34%
07	36%	24%	-	40%
08	34%	42%	53%	-

Tab. 5.7: Vergleichsergebnisse ausgewählter Sätze (2. Studie)
Vergleichsergebnisse

In Tab. 5.7 wurde jeweils der Satz in der Zeile (Ursprungssatz) mit dem Satz in der Spalte (Vergleichssatz) verglichen. Rot markierte Vergleichsergebnisse wurden aufgrund der

eingestellten Mindestrelevanz von 30% nicht angezeigt.

Die Vergleiche zwischen Teilnehmer 08 und 07 mit 56% und Teilnehmer 06 und 05 mit 42% zeigen, dass die Teilnehmer 08 und 06 durchaus von dem Vergleichsergebnis hätten profitieren können. Die Ergebnisse machen aber auch deutlich, dass durchaus auch größere Zusammenhänge, die im Hörspiel in mehreren Sätzen dargestellt sind, in einem Satz Erwähnung finden können. Der Wunsch eines Teilnehmers, einen größeren Kontext angezeigt zu bekommen, lässt sich in diesem Fall nachvollziehen.

5.4.7 Fazit

Die Auswertung der 2. Studie ist in vielerlei Hinsicht unbefriedigend. Die Aufgabenstellung, den Inhalt eines Hörspiels in eigenen Worten wiederzugeben, wurde sehr individuell ausgelegt. Die Anzahl der Ungenauigkeiten in der inhaltlichen Darstellung, wie auch teilweise vermeidbare Rechtschreib- und Grammatikfehler können zumindest nicht mit zeitlicher Überforderung erklärt werden, da kein Teilnehmer die zur Verfügung stehende Zeit ausschöpfte.

Auffällig ist die im Vergleich zur ersten Studie die schlechte Bewertung der Vergleichsfunktion. In Gruppe I wurden im Verhältnis zu den angezeigten nur wenige Ergebnisaufrufe (siehe Abb. 5.8 (a)) aufgerufen. Es besteht allerdings kein direkter Zusammenhang zwischen der Anzahl der aufgerufenen Ergebnisse und der Bewertung der Funktion.

Gruppe II hat die Vergleichsfunktion öfter benutzt als Gruppe I, hat aber im Schnitt mehr Auslassungen und Inkonsistenzen in ihren Texten. Die Vermutung, dass die Texte sich ähnlicher waren, könnte wiederum bedeuten, dass fehlerhafte Darstellungen übernommen wurden.

Zwar zeigt die Auswertung der Ereignisprotokolle wie auch der Fragebögen, dass von den Teilnehmern inhaltliche und sprachliche Änderungen vorgenommen wurden, jedoch ist in keinem Fall eine konsequente Nutzung erkennbar. Änderungen wurden nur vereinzelt vorgenommen.

Die Aufrufe der Vergleichsergebnisse ist individuell verschieden, eine klare Trennung ist meistens nicht ersichtlich. Inwieweit der Aufruf eines Vergleichsergebnis das Weiter-schreiben beeinflusst hat, kann nicht festgestellt werden. Es kann jedoch vermutet werden, dass der Aufruf eines Vergleichsergebnisses nicht zwangsläufig Änderungen am Ursprungssatz zur Folge haben muss. Wie in der ersten Studie bereits vermutet, kann es auch

das Weiterschreiben beeinflusst haben.

Die Leistungsverbesserung von Teilnehmer 07 kann nicht direkt mit der Nutzung der Vergleichsfunktion erklärt werden. Es kann eine Änderung (Korrektur eines Rechtschreibfehlers) im Ereignisprotokoll nachvollzogen werden, die der Teilnehmer auch im Fragebogen angab. Die anschließende Nutzung der Vergleichsfunktion hatte allerdings keine weiteren Auswirkungen auf das Ergebnis.

5.5 3. Laborstudie

5.5.1 Ausgangspunkt

Aufgrund der bis dahin unbefriedigenden Ergebnisse die Planungen für eine weitere Studie kurzfristig geändert. Aufbauend auf die 2. Studie wurde die 3. Studie mit geänderten Rahmenbedingungen konzipiert.

5.5.2 Änderungen am Studiendesign

Die Ergebnisse der 2. Studie ließen die Vermutung zu, dass die Ergebnisse möglicherweise durch fehlende Motivation beeinflusst wurden. Zwar hatten alle Teilnehmer die Aufgabe bearbeitet, jedoch zeigten die Texte im Vergleich zum Hörspiel viele inhaltliche Schwächen. Zudem hätten viele Dokumente hinsichtlich Rechtschreibung und Grammatik verbessert werden können, wenn Korrektur gelesen worden wäre. Dabei hatte die zur Verfügung stehende Zeit keinen Einfluss, da während der Studie beobachtet werden konnte, dass die Zeit völlig ausreichend war.

Um einen zusätzlichen Anreiz zu schaffen, die gestellte Aufgabe gewissenhaft und bestmöglich zu bearbeiten, wurde neben einer erfolgsunabhängigen Entlohnung eine erfolgsabhängige Entlohnung in Aussicht gestellt. Diese orientierte sich einerseits an der individuellen Leistung des Teilnehmers und andererseits an der Leistung der Gruppe. Es wurde festgelegt, dass die Teilnehmer einen bestimmten Fehlerquotienten, der sich an den Ergebnissen der 2. Studie orientierte, und eine bestimmte Anzahl an Inkonsistenzen und Auslassungen im Vergleich zum Hörspiel nicht überschreiten durften. Zur Bewertung der Gruppenleistung wurde der Durchschnitt dieser Faktoren ermittelt.

5.5.3 Änderungen am Prototyp

Am Prototyp wurde der Vorschlag der Teilnehmer aus der 2. Studie umgesetzt. Die Sortierung der Ergebnisse nach Ursprungs- bzw. Vergleichssatz erfolgte nicht mehr in alphabetischer Reihenfolge, sondern in der Reihenfolge des Auftretens im Dokument. Außerdem wurde die Mindestanzahl indexierter Wörter eines Satzes, die für einen Vergleich vorhanden sein müssen, auf 3 festgelegt.

5.5.4 Studienbeschreibungen

Die Studie fand vom 01. bis 06. Juni 2011 an insgesamt zwei Terminen statt. Eingeladen waren 8 Studierende, die auf zwei Gruppen aufgeteilt waren. Jede Gruppe bestand aus je zwei weiblichen und zwei männlichen Teilnehmern. In beiden Gruppen war jeweils ein ausländischer Studierender.

Die Aufgabe bestand wie in der 2. Studie darin, den Inhalt eines Hörspiels in eigenen Worten wiederzugeben. Im Gegensatz zur 2. Studie durften sich die Teilnehmer allerdings keine Notizen machen. Stattdessen hatten diese nach der Abspielphase 5 Minuten Zeit, innerhalb der Gruppe Fragen zu klären und den Inhalt zu wiederholen.

Als Hörspiel wurde *Das Nusszweiglein* von Ludwig Bechstein ausgewählt. Dieses hat eine Länge von 9:04 Minuten.

Besondere Vorkommnisse

Drei der vier eingeladenen Teilnehmer aus Gruppe II erschienen ohne Angabe von Gründen nicht zum Termin. Eine sinnvolle Durchführung wäre somit nicht möglich gewesen, da mangels Vergleichsdokumenten keine Vergleiche stattgefunden hätten. Aus diesem Grund wurde das Dokument von Teilnehmer 21 dem Thema von Gruppe I zugeordnet.

5.5.5 Auswertung und Interpretation der Ergebnisse

Abspiel- und Diskussionsphase

Während der Abspielphase durften keine Notizen gemacht werden. In der anschließenden Diskussionsphase wurden zunächst Verständnisfragen geklärt. Dem ausländischen Teilnehmer wurden die Begriffe *Pudel* und *Hirte* erklärt. Danach wurde der Inhalt der Geschichte noch einmal grob skizziert und vereinzelt Details besprochen.

Durchführungsphase

Während der Durchführungsphase war das Schreibtempo deutlich geringer als es in der 2. Studie beobachtet werden konnte. Bei fast allen Teilnehmern konnten Pausen festgestellt

werden.

Bewertung des Prototyps

Der Prototyp hat ohne Auffälligkeiten funktioniert. Die Bedienung des Schreibprogramms wurde zu Beginn sehr ausführlich erklärt. Die Auswertung des *System User Scalability*, dargestellt in Abb. 5.11, zeigt allerdings, dass nicht jeder Teilnehmer mit der Benutzerfreundlichkeit zufrieden war.

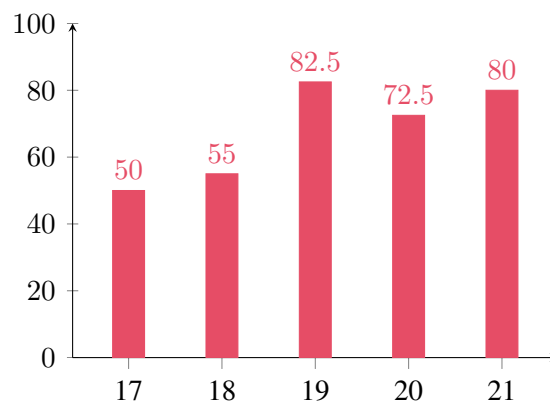


Abb. 5.11: Auswertung des SUS-Tests
Quelle: Eigene Darstellung

Die vergleichbaren Ergebnisse der SUS-Auswertung in der 1. Studie (siehe Abb. 5.2) und in der 2. Studie (siehe Abb. 5.7) hatten keine Ausreißer nach unten.

Die Vergleichsfunktion wurde mit 3,4 bewertet. Ein Teilnehmer bescheinigte der Funktion nur eine geringe Nützlichkeit, die anderen Bewertungen waren neutral bis gut.

Bewertung der Texte

Abgesehen von einer Ausnahme wiesen die Texte ähnlich viele Inkonsistenzen und Auslassungen, dargestellt in Abb. 5.12, auf, wie die Texte in der zweiten Studie.

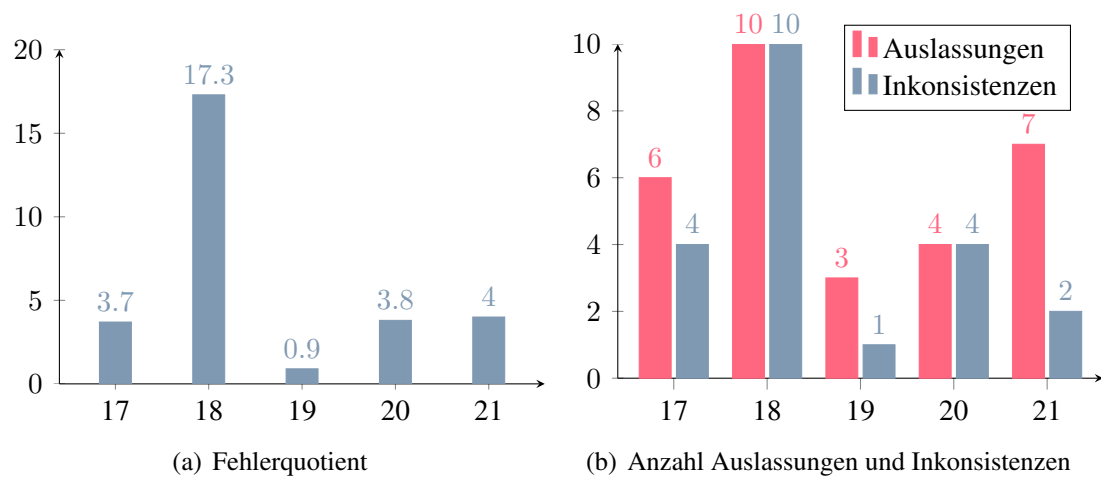


Abb. 5.12: Fehlerquotienten und Anzahl Auslassungen und Inkonsistenzen
Quelle: Eigene Darstellung

[**Anm.:** Der Text von Teilnehmer 18 konnte aufgrund der vielen Rechtschreibfehler und erheblicher sprachlicher Defizite hinsichtlich Auslassungen und Inkonsistenzen nicht bewertet werden. Daher wurde pauschal ein Maximalwert von 10 festgelegt.]

Arbeitsweise

Die Auswertung der Ereignisprotokolle ergab bei den meisten Teilnehmer keine klare Trennung zwischen Schreiben und Kontrolle. Einzig bei Teilnehmer 19, der mit Abstand das beste Ergebnis hinsichtlich Fehlerquotient, Anzahl Auslassungen und Inkonsistenzen erzielte, konnte eine strikte Trennung festgestellt werden.

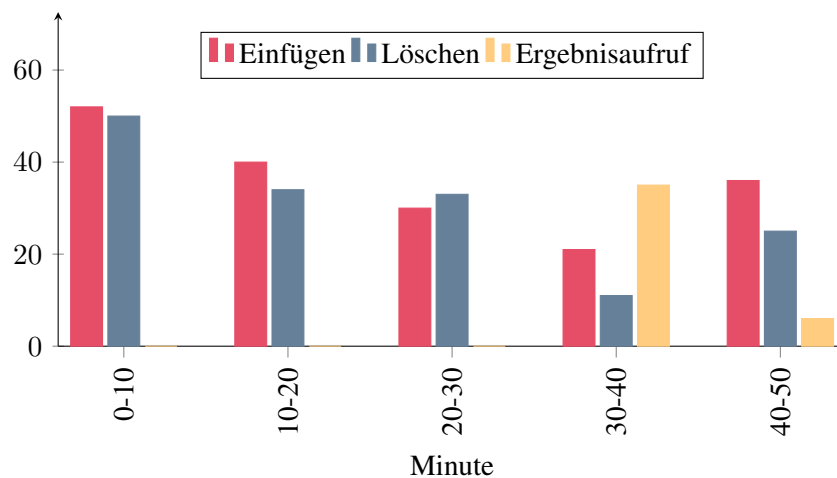


Abb. 5.13: Zeitprofil der Ergebnisaufufe von Teilnehmer 19
Quelle: Eigene Darstellung

Im Zeitprofil von Teilnehmer 19, dargestellt in Abb. 5.13, wird eine strikte Trennung zwischen Schreiben und Kontrolle deutlich. Wie die Auswertung des Ereignisprotokolls zeigte, war das Dokument nach 30 Minuten in den Grundzügen fertiggestellt. Die letzten 20 Minuten waren bestimmt von Ergebnisaufrufen und kleinen Ergänzungen bzw. Korrekturen.

Insgesamt konnten 14 zusammenhängende inhaltliche Änderungen, eine Korrektur der Rechtschreibung und drei Änderungen im sprachlichen Ausdruck im Protokoll festgestellt. Davon standen sechs Änderungen in einem direkten Zusammenhang mit einem vorherigen Ergebnisaufruf des Ursprungssatzes, an dem anschließend die Änderung vorgenommen wurde.

Vergleichsergebnisse

Die Auswertung der Vergleichsergebnisse von Teilnehmer 19 zeigte, dass zwei der vier Fehler hätten vermieden werden können.

Ursprungssatz:

„Nur das von seiner jüngsten Tochter gewünschte Nusszweiglein konnte er nicht auftreiben, was ihn traurig stimmte.“

Vergleichssatz (Teilnehmer 18):

„Nur ein grünes Nusszweiglein für sein jüngste Tochter konnte er nicht finden, was ihn sehr betrückte, [da ihm diese Tochter die Liebste war.](#)“

Relevanz: 38%

Ursprungssatz:

„Im letzten Raum waren Drachen, Schlangen und Lindwürmer.“

Vergleichssatz (Teilnehmer 17):

„Sie durchschritten zehn Zimmer und kamen in ein elftes, in dem Schlangen, [Basiliske](#) und Lindwürmer waren.“

Relevanz: 40%

Zum dritten und vierten Fehler existierte kein entsprechendes Vergleichsergebnis, da kein Teilnehmer dieses Detail im Text erwähnte.

Auswertung der offenen Fragen

Auf die Frage, ab welcher Relevanz die Vergleichsergebnisse hilfreich waren, bewegten sich die Angaben zwischen 50% und 65%. Im Mittel lag sie bei ca. 59%.

Nach Verbesserungsmöglichkeiten befragt bemängelte ein Teilnehmer, dass die Anzeige der Vergleichsergebnis zu unübersichtlich sei. Außerdem wünschte sich der Teilnehmer einen größeren Kontexte, also die Anzeige des Vorgänger bzw. Nachfolgesatzes. Teilnehmer 19, der die Vergleichsfunktion intensiv und zielgerichtet nutzte, hätte sich gewünscht, dass man Vergleichsergebnisse von Dokumenten bestimmter Nutzer ausblenden und manuell eine eigene Mindestrelevanz für die Vergleichsergebnisse festlegen kann.

Zwei Teilnehmer gaben an, dass sie mehr Zeit benötigt hätten. Ein Teilnehmer hätte weitere 10 Minuten, ein anderer hätte weitere 60 Minuten benötigt.

Fazit

Isoliert betrachtet hat Teilnehmer 19 die Vergleichsfunktion in der erwarteten Weise genutzt. Durch gezielte Aufrufe von Vergleichsergebnissen und die manuelle Überprüfung zwei gegenübergestellter Sätze ist es gelungen, das mit deutlichem Abstand beste Ergebnis aller Teilnehmer dieser Studie zu erreichen. Inwieweit dieses Ergebnis durch das intellektuelle Potenzial beeinflusst ist, kann mangels Vergleichswerten nicht ermittelt werden. Die Auswertung des Ereignisprotokolls lässt aber den Schluss zu, dass die Vergleichsfunktion zu einem gewissen Teil dazu beigetragen hat. Der Teilnehmer bescheinigte der Vergleichsfunktion entsprechend eine Nützlichkeit von 4.

Bei drei Teilnehmern wurde eine zeitliche bzw. sprachliche Überforderungen deutlich. Es kann daher nicht ausgeschlossen werden, dass die Qualität der Texte und die Nutzung und Bewertung der Vergleichsfunktion dadurch beeinflusst wurde.

Teilnehmer 21, der getrennt von der Gruppe schrieb, konnte von dem guten Ergebnis von Teilnehmer 19 nicht profitieren. Hier konnte ein ähnliches Verhalten wie bei den Teilnehmern der 2. Studie beobachtet werden. Nach exakt 41 Minuten stellte dieser Teilnehmer das Schreiben ein und war auf Nachfrage mit seinem Ergebnis zufrieden.

5.6 Gesamtfazit und Ausblick

Die Organisation der Studien war im Nachhinein betrachtet nicht optimal. Es hätte eine vorherige Überprüfung der sprachlichen Fähigkeiten der Teilnehmer stattfinden müssen. Die teilweise sehr hohen Fehlerquotienten von bis zu 20 haben eine sinnvolle Indexierung und somit einen Vergleich der Dokumente erheblich beeinträchtigt.

Ob die teils erfolgsabhängige Entlohnung in der 3. Studie tatsächlich die erhoffte Motivationssteigerung erzielt hat, kann an dieser Stelle nicht geklärt werden. Allerdings war die Anzahl der Anmeldungen (2. Studie: 54, 3. Studie: 34) trotzdem ausreichend hoch, so dass auch für zukünftige Studien diese Art der Entlohnung gewählt werden würde.

Aus technischer Sicht hat der Prototyp die Anforderungen erfüllt. Die automatische Indexierung und der Vergleich der Dokumente hat während der Studien ausnahmslos funktioniert. Die Benutzerfreundlichkeit des Clients wurde überwiegend positiv bewertet. Kritisiert wurde die Darstellung der Vergleichsergebnisse und die Art der Sortierung.

Die Relevanz, ab der Vergleichsergebnisse hilfreich waren, bewegte sich zwischen 50% und 60%. Dabei handelt es sich durchaus um einen plausiblen Wert, der sowohl eine mögliche Ähnlichkeit als eben auch Unterschiede widerspiegelt.

Die Erwartung, dass die automatische Indexierung und der anschließende Vergleich von Dokumenten einen positiven Einfluss auf das Ergebnis haben würde, wurde nicht erfüllt. Es konnte keine signifikante Verbesserung festgestellt werden. Zwar gab es in jeder Studie positive Ausnahmen, in der Summe jedoch konnten nur geringe unterstützende Eigenschaften beobachtet werden.

Der Transfer von Wissen, für den die 2. und 3. Studie konzipiert war, blieb größtenteils aus. Die Erwartung, dass ein Teilnehmer die intellektuellen und / oder sprachlichen Fähigkeiten anderer Teilnehmer zu seinem Vorteil nutzt, erfüllte sich nicht. Die Ursachen können dabei nicht klar benannt werden. Die Vermutung, dass fehlende Motivation eine Rolle spielte, konnte in der dritten Studie zumindest nicht widerlegt werden. Ein anderer Aspekt ist die fehlende soziale Interaktion während der Studien. Die 2. und 3. Studie waren so konzipiert, dass jeder Teilnehmer für sich arbeitete. Das Fehlen eines gemeinsamen Ziels kann durchaus die Leistung des Einzelnen beeinflusst haben. In der 3. Studie wurde versucht, dies durch eine an der Gruppenleistung orientierten Entlohnung zu simulieren. Leider zeigte sich bei drei der fünf Teilnehmer eine zeitliche und teils sprachliche Überforderung, die sich möglicherweise auf die Motivation niedergeschlagen hat.

Die wichtigsten Aspekte in zukünftigen Entwicklungen sind sowohl die Darstellung der Vergleichsergebnisse als auch die Berechnung der Relevanz. Die zurückhaltende Gestaltung (siehe Kap. 4.3.4) kann durchaus Einfluss auf die Ergebnisse gehabt haben. Deutlichere Hinweise z. B. durch Hervorheben von Textstellen, die Rechtschreibfehler, Widersprüche oder Gemeinsamkeiten aufweisen, könnten die Aufmerksamkeit des Benutzers erhöhen und so dazu beitragen, dass Mängel eher erkannt werden. Die Bewertung der Relevanz unterliegt zwar mathematischen Berechnungen, es fehlt allerdings ein mathematisches Modell, das in der Lage wäre, die Probleme, die sich während der Studien zeigten, zu kompensieren. Kurze Sätze bedeuten wenige indexierte Wörter, die für einen Vergleich herangezogen werden können. Die Wahrscheinlichkeit, dass ein Vergleichssatz durch eine hohe Bewertung als relevant eingestuft wird, steigt im Vergleich zu Sätzen, die viele indexierte Wörter enthalten. Dies sollte berücksichtigt werden. Andererseits kann auch das isolierte Auffinden eines Wortes von Bedeutung sein. Die Berücksichtigung statistischer Faktoren wie in statistischen Verfahren der Indexierung (siehe Kap. 2.2.1) können die Relevanzbewertung verbessern und somit dem Benutzer gezielter relevante Vergleichsergebnisse zur Verfügung stellen.

Die Idee, Dokumente auf Satzebene zu indexieren und zu vergleichen, würdigt den Satz als semantische Einheit, vernachlässigt aber, dass für einen Benutzer relevante Informationen auf mehrere Sätze verteilt sein können. Von zwei Teilnehmer wurde gewünscht, den Vorgänger- bzw. Nachfolgesatz des Vergleichssatzes anzuzeigen. Die Anzeige eines größeren Kontexts kann ein richtiger Ansatz sein. Andererseits muss auch der Fall berücksichtigt werden, dass eben nicht immer ein größerer Kontext die gewünschten Informationen bringt. Daher sollte nur dann ein größerer Kontext gewählt werden, wenn auf Grundlage der Vergleichsfunktion eine entsprechende Relevanz zwischen Ursprungssatz und den benachbarten Sätzen des Vergleichssatzes festgestellt werden kann.

Abschließend sollte ein Aspekt nicht unerwähnt bleiben, der zwar unterschwellig im Studiendesign der 2. und 3. Studie vorhanden war, aber nie konkret in dieser Arbeit erwähnt wurde. Im Zuge der Plagiatsaffären verschiedener politischer Persönlichkeiten kam es zumindest in den Sinn, dass ein solches System auch zur Identifizierung von Plagiaten eingesetzt werden könnte. Für wortwörtliches Plagieren bieten sich andere Verfahren, die Texte auf bloße Übereinstimmungen hin überprüfen, eher an. Der geübte Plagiator, der Formulierungen oder den strukturellen Aufbau eines Satzes ändert, um ein Plagiat zu vertuschen, kann jedoch durch die in dieser Arbeit beschriebenen Konzepte möglicherweise überführt werden.

Literatur

- Berger, A. L., S. A. Della Pietra und V. J. Della Pietra: "A Maximum Entropy Approach to Natural Language Processing". In: *Computational Linguistics*. Bd. 22. Abruf: 19.05.2011. 1996. URL: <http://www.aclweb.org/anthology-new/J/J96/J96-1002.pdf>.
- Brooke, John: *SUS. A quick and dirty usability scale*. Techn. Ber. Abruf: 09.03.2011. 1986. URL: www.hostserver150.com/usabilit/trump/documents/Suschapt.doc.
- Dudenredaktion: *Duden. Die deutsche Rechtschreibung*. Wissenschaftlicher Rat der Dudenredaktion, 2006.
- Glover, A. und G. Hirst: *Detecting stylistic inconsistencies in collaborative writing*. Techn. Ber. Abruf: 04.10.2010. 1995. URL: <http://www.cs.toronto.edu/pub/gh/Glover+Hirst-CollabWriting-96.ps>.
- Gross, T. und M. Koch: *Computer-Supported Cooperative Work*. Rosenheimer Straße 145, D-81671 München: Oldenbourg Wissenschaftsverlag GmbH, 2007.
- Hamran, D.: *An Experimental Study of Factors Important in Document Ranking*. Techn. Ber. Abruf: 22.10.2010. 1986. URL: http://www2.denizyuret.com/ref/harman/Document_Ranking___Harman_1986.pdf.
- Hausser, R.: *Grundlagen der Computerlinguistik*. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg New York, 2000.
- Kaiser, A.: *Computer-gestütztes Indexieren in Intelligenten Information Retrieval Systemen. Ein Relevanz-Feedback orientierter Ansatz zur Informationserschließung in unformatierten Datenbanken*. Wien, 1993.
- Lemnitzer, L. und C. Kunze: *Computerlexikographie*. Tübingen, 2007.
- Lenders, W. und G. Willée: *Linguistische Datenverarbeitung. Ein Lehrbuch*. Opladen: Westdeutscher Verlag, 1986.

- Lewandowski, D.: *Web Information Retrieval. Technologien zur Informationssuche im Internet*. Abruf: 17.10.2010. 2005. URL: <http://www.durchdenken.de/lewandowski/web-ir/>.
- Lezius, W.: *www.wolfganglezius.de [cl:morphy]*. Abruf: 18.01.2011. 2010. URL: <http://www.wolfganglezius.de/doku.php?id=cl:morphy>.
- Meibauer, J. u. a.: *Einführung in die germanistische Linguistik*. Stuttgart: Verlag J.B. Metzler, 2002.
- Naber, D.: *Deutsches Morphologie-Lexikon*. Abruf: 18.01.2011. 2011. URL: <http://www.danielnaber.de/morphologie/>.
- Nohr, H.: *Automatische Indexierung. Einführung in betriebliche Verfahren, Systeme und Anwendungen*. Potsdam: Verlag für Berlin Brandenburg GmbH, 2001.
- Olivier, H. und N. Pinkwart: *Towards supporting Phases in Collaborative Writing Processes*. Techn. Ber. Abruf: 04.10.2010. 2009. URL: http://hcis.in.tu-clausthal.de/pdf/HOlivierDE_final.pdf.
- Postel, H. J.: "Die Kölner Phonetik. Ein Verfahren zur Identifizierung von Personennamen auf der Grundlage der Gestaltanalyse". In: *IBM Nachrichten* 19 (1969), S. 925–931.
- Rahhal, C. u. a.: *SemCW: Semantic Collaborative Writing using RST*. Techn. Ber. Abruf: 18.10.2010. 2007. URL: <http://hal.inria.fr/docs/00/16/76/41/PDF/RR-6271.pdf>.
- Reynar, J. C. und A. Ratnaparkhi: "A Maximum Entropy Approach to Identifying Sentence Boundaries". In: *Proceedings of the Fifth Conference on Applied Natural Language Processing*. Abruf: 19.05.2011. 1997. URL: <http://www.aclweb.org/anthology-new/A/A97/A97-1004.pdf>.
- Salton, G. und M. J. McGill: *Information Retrieval. Grundlegendes für Informationswissenschaftler*. Hamburg, 1987.
- Schitteck, W.: *Deutsche Rechtschreibung – was hat sich geändert? Eine knappe und dementsprechend grobe Übersicht über alle heute gültigen Änderungen seit 1996*. Abruf: 07.04.2011. 2010. URL: <http://www.staff.uni-marburg.de/~schitteck/rechtsch.htm>.
- Schröder, I.: *A Case Study in Part-of-Speech Tagging Using the ICOPOST Toolkit*. Techn. Ber. Vogt-Kölln-Str. 30, D-22527 Hamburg, 2002.
- Skaif-Molli, H., P. Molli und C. Rahhal: *Collaborative Writing of XML Documents*. Techn. Ber. Abruf: 09.10.2010. 2009. URL: <http://hal.inria.fr/docs/00/43/25/93/PDF/ICTTA08.pdf>.
- statistica: *Anteil ausländischer Studenten an deutschen Hochschulen vom Wintersemester 1998/99 bis 2009/10*. Abruf: 06.06.2011. 2011. URL: <http://de.statista>.

com/statistik/daten/studie/222/umfrage/anteil-auslaendischer-studenten-an-hochschulen/.

Steiner, T.: *Observer-Muster über RMI*. Techn. Ber. Abruf: 07.12.2010. 2005. URL: www.sigs.de/publications/js/2005/05/steiner_JS_05_05.pdf.

Stock, W. G.: *Information Retrieval*. Rosenheimer Straße 145, D-81671 München: Oldenbourg Wissenschaftsverlag GmbH, 2007.

— *Informationswirtschaft*. Rosenheimer Straße 145, D-81671 München: Oldenbourg Wissenschaftsverlag GmbH, 2000.

— *Wissenrepräsentation*. Rosenheimer Straße 145, D-81671 München: Oldenbourg Wissenschaftsverlag GmbH, 2008.

TU Clausthal, Pressereferent Dezernat 2 (EDV und Statistik) der: *Zahlenspiegel WS 2010/11*. Abruf: 01.06.2011. 2010. URL: http://www.tu-clausthal.de/hv/d5/vhb/system9/9_30_10_10_11.pdf.

A Hörspiel I (2. Studie)

Titel: Wie Eulenspiegel dem Pfarrer sein Pferd abschwatzte

Autor: Hermann Bote (1467 - um 1520)

Quelle:

- http://www.hekaya.de/txt.hx/wie-eulenspiegel-dem-pfarrer-zu-kissenbrueck-sein-pferd-mit-einer-falschen-beichte-abschwatzte--sage--eulenspiegel_38
- <http://www.vorleser.net/html/bote.html>

Eine böse Schalkheit ließ sich Eulenspiegel nicht entgehen in dem Dorfe Kissenbrück im Asseburger Gerichtsbezirk. Da wohnte ein Pfarrer, der eine gar schöne Haushälterin hatte und dazu ein kleines, hübsches, munteres Pferd. Die hatte der Pfarrer alle beide sehr gern, das Pferd und auch die Magd. Nun war der Herzog von Braunschweig zu dieser Zeit in Kissenbrück gewesen und hatte den Pfarrer durch andere Leute mehrfach gebeten, ihm das Pferd zu überlassen, er wolle ihm dafür mehr geben, als es wert sei. Der Pfarrer schlug es aber dem Fürsten allezeit ab. Er wollte das Pferd nicht verlieren, weil er es so gern hatte. Der Fürst wagte auch nicht, ihm das Pferd wegnehmen zu lassen, denn das Gericht unterstand dem Rat von Braunschweig.

Eulenspiegel hatte diese Dinge gehört und wohl verstanden und sprach zu dem Fürsten: »Gnädiger Herr, was wollt Ihr mir schenken, wenn ich Euch das Pferd des Pfaffen zu Kissenbrück herbeischaffe?« »Wenn du das tust«, sprach der Herzog, „will ich dir den Rock geben, den ich jetzt an habe.“ Und das war ein roter, mit Perlen bestickter Schamlot.

Eulenspiegel nahm das an und ritt von Wolfenbüttel in das Dorf zur Herberge beim Pfarrer. Er war in des Pfarrers Haus wohlbekannt, denn er war oft vorher bei ihm gewesen und ihm willkommen. Als er nun etwa drei Tage dort gewesen war, da gebärdete er sich, als ob er ganz krank sei, ächzte laut und legte sich nieder. Dem Pfaffen und seiner Haushälterin tat es leid, und sie wussten keinen Rat, was sie tun sollten. Zuletzt wurde Eulenspiegel so krank, dass ihn der Pfaffe anredete und ihn bat, er möge beichten und das Abendmahl nehmen. Eulenspiegel war durchaus dazu geneigt. Der Pfarrer wollte ihm selbst die Beichte abnehmen und ihn aufs schärfste befragen. Er sprach, Eulenspiegel möge an seine Seele denken, denn er habe sein Leben lang viel Abenteuer getrieben. Er Sorge sich, ob ihm Gott der Allmächtige seine Sünden vergeben werde. Eulenspiegel sprach ganz kränklich zu dem Pfarrer: er wisse nichts, das er getan habe, außer einer Sünde; die aber dürfe er ihm nicht beichten. Er möge ihm einen anderen Pfaffen holen, dem wolle er sie beichten.

Denn wenn er sie ihm offenbare, so besorge er, dass er ihm darum zürnen würde.

Als der Pfarrer das hörte, meinte er, dahinter sei etwas verborgen, und das wollte er wissen. Er sprach: „Lieber Eulenspiegel, der Weg ist weit, ich kann den anderen Pfaffen nicht so schnell erreichen. Wenn du aber inzwischen stirbst, so hätten du und ich vor Gott dem Herrn die Schuld, wenn es deshalb mit dir versäumt würde. Sage es mir! Die Sünde wird so schwer nicht sein, ich will dich davon lossprechen. Was hülfte es auch, wenn ich böse würde? Ich darf doch die Beichte nicht offenbaren.“ Da sagte Eulenspiegel: „So will ich das wohl beichten.“ Die Sünde sei auch nicht so schwer. Sondern ihm sei es nur leid, dass der Pfarrer zornig werden würde, denn es beträfe ihn. Da verlangte es den Pfarrer noch mehr, es zu wissen. Und er sprach: wenn er ihm etwas gestohlen, sonst etwas angetan, ihn geschädigt habe oder was es auch sei, Eulenspiegel möge es ihm beichten. Er wolle es ihm vergeben und ihn nimmer darum hassen.

Eulenspiegel sprach: „Ach, lieber Herr, ich weiß, Ihr werdet mir darum zürnen. Doch ich fühle und fürchte, dass ich bald von hinnen scheiden muss. Ich will es Euch sagen. Gott weiß, ob Ihr zornig oder böse werdet. Lieber Herr, das ist es: ich habe bei Eurer Magd geschlafen.“ Der Pfaffe fragte, wie oft das geschehen sei. Eulenspiegel antwortete: „Nur fünfmal.“ Der Pfaffe dachte: dafür soll sie fünf Hiebe bekommen.

Er absolviertes Eulenspiegel sogleich, ging in die Kammer und ließ seine Magd zu sich kommen. Er fragte sie, ob sie bei Eulenspiegel geschlafen habe. Die Köchin sprach nein, das sei gelogen. Der Pfaffe sagte, Eulenspiegel habe es ihm doch gebeichtet, und er glaube es ihm auch. Die Haushälterin sprach: „Nein“, der Pfaffe sprach: „Ja“ und erwischte einen Stecken und schlug sie braun und blau. Eulenspiegel lag im Bett, lachte und dachte bei sich selbst: Nun will das Spiel gut werden und ein rechtes Ende nehmen. Und er lag den ganzen Tag so.

In der Nacht aber wurde er gesund, stand des Morgens auf und sprach, es gehe ihm besser, er müsse in ein anderes Land. Der Pfarrer möge berechnen, was er während der Krankheit verzehrt habe. Der Pfaffe rechnete mit ihm ab, war aber so irr in seinem Sinn, dass er nicht wusste, was er tat. Er berechnete Geld und nahm doch kein Geld und war mit allem zufrieden, wenn Eulenspiegel nur von dannen ritt. Ebenso ging es der Köchin, die um seinetwillen geschlagen worden war.

Als Eulenspiegel bereit war und gehen wollte, sprach er zu dem Pfaffen: „Herr, seid daran erinnert, dass Ihr die Beichte offenbart habt! Ich will nach Halberstadt zum Bischof gehen und ihm das von Euch berichten.“ Der Pfaffe vergass seinen Zorn, als er hörte, dass Eulenspiegel ihn in Schwierigkeiten bringen wollte. Er fiel ihm zu Füßen und bat ihn mit großem Ernst zu schweigen. Es sei im Jähzorn geschehen. Er wolle ihm zwanzig Gulden geben, damit er ihn nicht anzeige. Eulenspiegel sprach: „Nein, ich wollte nicht einmal hundert Gulden nehmen, um das zu verschweigen. Ich will gehen und es vorbringen, wie

es sich gebührt.“Der Pfaffe bat die Magd mit tränenden Augen, sie solle Eulenspiegel fragen, was er von ihm haben möchte; das wolle er ihm geben. Schließlich sagte Eulenspiegel, wenn der Pfaffe ihm sein Pferd geben wolle, so wolle er schweigen, und es solle ungemeldet bleiben. Er wolle aber nichts anderes nehmen als das Pferd. Der Pfaffe hatte das Pferd sehr gern und hätte Eulenspiegel lieber seine ganze Barschaft gegeben, als von dem Pferde zu lassen. Und doch trennte er sich von ihm, wenn auch gegen seinen Willen, denn die Not brachte ihn dazu.

Er gab Eulenspiegel das Pferd und ließ ihn damit fortreiten. Also ritt Eulenspiegel mit des Pfaffen Pferd nach Wolfenbüttel. Als er auf den Stadtwall kam, stand der Herzog auf der Zugbrücke und sah Eulenspiegel mit dem Pferd dahertraben. Sogleich zog der Fürst den Rock aus, den er Eulenspiegel versprochen hatte, ging zu ihm und sprach: „Schau her, mein lieber Eulenspiegel, hier ist der Rock, den ich dir versprochen habe!“Da sprang Eulenspiegel vom Pferd und sagte: „Gnädiger Herr, hier ist Euer Pferd.“Er hatte sich den großen Dank des Herzogs verdient und musste ihm erzählen, wie er das Pferd von dem Pfaffen an sich gebracht hatte. Darüber lachte der Fürst und war fröhlich und gab Eulenspiegel ein anderes Pferd zu dem Rock.

Der Pfarrer aber trauerte um das Pferd und schlug die Köchin noch oft und heftig darum, so dass sie ihm entlief. Da war er ihrer beide ledig, des Pferdes und der Magd.

B Hörspiel II (2. Studie)

Titel: Die weiße Schlange

Autoren: Brüder Grimm

Quelle:

- http://www.hekaya.de/txt.hx/die-weisse-schlange--maerchen--grimm_khm_17
- <http://www.vorleser.net/html/grimm.html>

Es ist nun schon lange her, da lebte ein König, dessen Weisheit im ganzen Lande berühmt war. Nichts blieb ihm unbekannt, und es war, als ob ihm Nachricht von den verborgensten Dingen durch die Luft zugetragen würde. Er hatte aber eine seltsame Sitte. Jeden Mittag, wenn von der Tafel alles abgetragen und niemand mehr zugegen war, musste ein vertrauter Diener noch eine Schüssel bringen. Sie war aber zugedeckt, und der Diener wusste selbst nicht, was darin lag, und kein Mensch wusste es, denn der König deckte sie nicht eher auf und aß nicht davon, bis er ganz allein war.

Das hatte schon lange Zeit gedauert, da überkam eines Tages den Diener, der die Schüssel wieder wegtrug, die Neugierde, dass er nicht widerstehen konnte, sondern die Schüssel in seine Kammer brachte. Als er die Tür sorgfältig verschlossen hatte, hob er den Deckel auf, und da sah er, dass eine weiße Schlange darin lag. Bei ihrem Anblick konnte er die Lust nicht zurückhalten, sie zu kosten; er schnitt ein Stückchen davon ab und steckte es in den Mund. Kaum aber hatte es seine Zunge berührt, so hörte er vor seinem Fenster ein seltsames Gewisper von feinen Stimmen. Er ging hin und horchte, da merkte er, dass es die Sperlinge waren, die miteinander sprachen und sich allerlei erzählten, was sie im Felde und Walde gesehen hatten. Der Genuss der Schlange hatte ihm die Fähigkeit verliehen, die Sprache der Tiere zu verstehen. Nun trug es sich zu, dass gerade an diesem Tage der Königin ihr schönster Ring fortkam und auf den vertrauten Diener, der überall Zugang hatte, der Verdacht fiel, er habe ihn gestohlen.

Der König ließ ihn vor sich kommen und drohte ihm unter heftigen Scheltworten, wenn er bis morgen den Täter nicht zu nennen wüsste, so sollte er dafür angesehen und gerichtet werden. Es half nichts, dass er seine Unschuld beteuerte, er ward mit keinem besseren Bescheid entlassen. In seiner Unruhe und Angst ging er hinab auf den Hof und bedachte, wie er sich aus seiner Not helfen könne. Da saßen die Enten an einem fließenden Wasser friedlich nebeneinander und ruhten, sie putzten sich mit ihren Schnäbeln glatt und hielten ein vertrauliches Gespräch. Der Diener blieb stehen und hörte ihnen zu. Sie erzählten

sich, wo sie heute morgen all herumgewackelt wären, und was für ein gutes Futter sie gefunden hätten, da sagte eine verdrießlich: „Mir liegt etwas schwer im Magen, ich habe einen Ring, der unter der Königin Fenster lag, in der Hast mit hinuntergeschluckt.“ Da packte sie der Diener gleich beim Kragen, trug sie in die Küche und sprach zum Koch: „Schlachte doch diese ab, sie ist wohl genährt.“

„Ja“, sagte der Koch und wog sie in der Hand, „die hat keine Mühe gescheut, sich zu mästen, und schon lange darauf gewartet, gebraten zu werden.“ Er schnitt ihr den Hals ab, und als sie ausgenommen ward, fand sich der Ring der Königin in ihrem Magen. Der Diener konnte nun leicht vor dem Könige seine Unschuld beweisen, und da dieser sein Unrecht wieder gutmachen wollte, erlaubte er ihm, sich eine Gnade auszubitten, und versprach ihm die größte Ehrenstelle, die er sich an seinem Hofe wünschte. Der Diener schlug alles aus und bat nur um ein Pferd und Reisegeld, denn er hatte Lust die Welt zu sehen und eine Weile darin herumzuziehen.

Als seine Bitte erfüllt war, machte er sich auf den Weg und kam eines Tags an einem Teich vorbei, wo er drei Fische bemerkte, die sich im Rohr gefangen hatten und nach Wasser schnappten. Obgleich man sagt, die Fische wären stumm, so vernahm er doch ihre Klage, dass sie so elend umkommen müssten. Weil er ein mitleidiges Herz hatte, so stieg er vom Pferde ab und setzte die drei Gefangenen wieder ins Wasser. Sie zappelten vor Freude, streckten die Köpfe heraus und riefen ihm zu: „Wir wollen es dir gedenken und es dir vergelten, dass du uns errettet hast.“

Er ritt weiter, und nach einem Weilchen kam es ihm vor, als hörte er zu seinen Füßen in dem Sand eine Stimme. Er horchte und vernahm, wie ein Ameisenkönig klagte: „Wenn uns nur die Menschen mit den ungeschickten Tieren vom Leib blieben! Da tritt mir das dumme Pferd mit seinen schweren Hufen meine Leute ohne Barmherzigkeit nieder.“ Er lenkte auf einen Seitenweg ein, und der Ameisenkönig rief ihm zu: „Wir wollen es dir gedenken und es dir vergelten.“

Der Weg führte ihn in einen Wald, und da sah er einen Rabenvater und eine Rabenmutter, die standen bei ihrem Nest und warfen ihre Jungen heraus. „Fort mit euch, ihr Galgenschwengel“, riefen sie, „wir können euch nicht mehr satt machen, ihr seid groß genug, und könnt euch selbst ernähren.“ Die armen Jungen lagen auf der Erde, flatterten und schlugen mit ihren Fittichen und schrie: „Wir hilflosen Kinder, wir sollen uns selbst ernähren und können noch nicht fliegen was bleibt uns übrig, als hier Hungers zu sterben.“ Da stieg der gute Jüngling ab, tötete das Pferd mit seinem Degen und überließ es den jungen Raben zum Futter. Die kamen herbeigehüpft, sättigten sich und riefen: „Wir wollen es dir gedenken und es dir vergelten.“

Er musste jetzt seine eigenen Beine gebrauchen, und als er lange Wege gegangen war, kam er in eine große Stadt. Da war großer Lärm und Gedränge in den Straßen, und kam einer

zu Pferde und machte bekannt, die Königstochter suche einen Gemahl, wer sich aber um sie bewerben wolle, der müsse eine schwere Aufgabe vollbringen, und könne er es nicht glücklich ausfahren, so habe er sein Leben verwirkt. Viele hatten es schon versucht, aber vergeblich ihr Leben daran gesetzt. Der Jüngling, als er die Königstochter sah, ward er von ihrer großen Schönheit so verblendet, dass er alle Gefahr vergaß, vor den König trat und sich als Freier meldete.

Als bald ward er hinaus ans Meer geführt und vor seinen Augen ein goldener Ring hingeworfen. Dann hieß ihn der König diesen Ring aus dem Meeresgrund wieder hervorzuholen, und fügte hinzu: „Wenn du ohne ihn wieder in die Höhe kommst, so wirst du immer aufs neue hinabgestürzt, bis du in den Wellen umkommst.“ Alle bedauerten den schönen Jüngling und ließen ihn dann einsam am Meere zurück. Er stand am Ufer und überlegte, was er wohl tun sollte, da sah er auf einmal drei Fische daherschwimmen, und es waren keine andere als jene, welchen er das Leben gerettet hatte. Der mittelste hielt eine Muschel im Munde, die er an den Strand zu den Füßen des Jünglings hinlegte, und als dieser sie aufhob und öffnete, so lag der Goldring darin. Voll Freude brachte er ihn dem Könige und erwartete, dass er ihm den verheißenen Lohn gewähren würde. Die stolze Königstochter aber, als sie vernahm, dass er ihr nicht ebenbürtig war, verschmähte ihn und verlangte, er sollte zuvor eine zweite Aufgabe lösen.

Sie ging hinab in den Garten und streute selbst zehn Säcke voll Hirsen ins Gras. „Die muss er morgen, ehe die Sonne hervorkommt, aufgelesen haben“, sprach sie, „und darf kein Körnchen fehlen.“ Der Jüngling setzte sich in den Garten und dachte nach, wie es möglich wäre, die Aufgabe zu lösen, aber er konnte nichts ersinnen, saß da ganz traurig und erwartete, bei Anbruch des Morgens zum Tode geführt zu werden. Als aber die ersten Sonnenstrahlen in den Garten fielen, so sah er die zehn Säcke alle wohl gefüllt nebeneinander stehen, und kein Körnchen fehlte darin. Der Ameisenkönig war mit seinen tausend und tausend Ameisen in der Nacht angekommen, und die dankbaren Tiere hatten den Hirsen mit großer Emsigkeit gelesen und in die Säcke gesammelt. Die Königstochter kam selbst in den Garten herab und sah mit Verwunderung, dass der Jüngling vollbracht hatte, was ihm aufgegeben war. Aber sie konnte ihr stolzes Herz noch nicht bezwingen und sprach: „Hat er auch die beiden Aufgaben gelöst, so soll er doch nicht eher mein Gemahl werden, bis er mir einen Apfel vom Baume des Lebens gebracht hat.“

Der Jüngling wusste nicht, wo der Baum des Lebens stand, er machte sich auf und wollte immer zugehen, solange ihn seine Beine trugen, aber, er hatte keine Hoffnung, ihn zu finden. Als er schon durch drei Königreiche gewandert war und abends in einen Wald kam, setzte er sich unter einen Baum und wollte schlafen; da hörte er in den Ästen ein Geräusch, und ein goldener Apfel fiel in seine Hand. Zugleich flogen drei Raben zu ihm herab, setzten sich auf seine Knie und sagten: „Wir sind die drei jungen Raben, die du

vom Hungertod errettet hast; als wir groß geworden waren und hörten, dass du den goldenen Apfel suchtest, so sind wir über das Meer geflogen bis ans Ende der Welt, wo der Baum des Lebens steht, und haben dir den Apfel geholt.“

Voll Freude machte sich der Jüngling auf den Heimweg und brachte der schönen Königstochter den goldenen Apfel, der nun keine Ausrede mehr übrig blieb. Sie teilten den Apfel des Lebens und aßen ihn zusammen: da ward ihr Herz mit Liebe zu ihm erfüllt, und sie erreichten in ungestörtem Glück ein hohes Alter.

C Hörspiel III (3. Studie)

Titel: Das Nusszweiglein

Autor: Ludwig Bechstein (1801 - 1860)

Quelle:

- http://www.hekaya.de/txt.hx/das-nusszweiglein--maerchen--bechstein_8
- <http://www.vorleser.net/html/bechstein.html>

Es war einmal ein reicher Kaufmann, der musste in seinen Geschäften in fremde Länder reisen. Da er nun Abschied nahm, sprach er zu seinen drei Töchtern: „Liebe Töchter, ich möchte euch gerne bei meiner Rückkehr eine Freude bereiten, sagt mir daher, was ich euch mitbringen soll?“ Die Älteste sprach: „Lieber Vater, mir eine schöne Perlenhalskette!“ Die andere sprach: „Ich wünschte mir einen Fingerring mit einem Diamantstein.“ Die Jüngste schmiegte sich an des Vaters Herz und flüsterte: „Mir ein schönes, grünes Nusszweiglein, Väterchen.“ „Gut, meine lieben Töchter!“ sprach der Kaufmann, „ich will mir’s aufmerken und dann lebet wohl.“ Weit fort reiste der Kaufmann und machte große Einkäufe, gedachte aber auch treulich der Wünsche seiner Töchter. Eine kostbare Perlenhalskette hatte er bereits in seinen Reisekoffer gepackt, um seine Älteste damit zu erfreuen, und einen gleich wertvollen Diamantring hatte er für die mittlere Tochter eingekauft. Einen grünen Nusszweig aber konnte er nirgends gewahren, wie er sich auch darum bemühte. Auf der Heimreise ging er deshalb große Strecken zu Fuß und hoffte, da sein Weg ihn vielfach durch Wälder führte, endlich einen Nussbaum anzutreffen; doch dies war lange vergeblich, und der gute Vater fing an betrübt zu werden, dass er die harmlose Bitte seines jüngsten und liebsten Kindes nicht zu erfüllen vermochte. Endlich, als er so betrübt seines Weges dahinzog, der ihn just durch einen dunklen Wald und an dichtem Gebüsch vorüberführte, stieß er mit seinem Hut an einen Zweig, und es raschelte, als fielen Schlossen darauf; wie er aufsah, war’s ein schöner, grüner Nusszweig, daran eine Traube goldner Nüsse hing. Da war der Mann sehr erfreut, langte mit der Hand empor und brach den herrlichen Zweig ab. Aber in demselben Augenblicke schoss ein wilder Bär aus dem Dickicht und stellt sich grimmig brummend auf die Hintertatzen, als wollte er den Kaufmann gleich zerreißen. Und mit furchtbarer Stimme brüllte er: „Warum hast du meinen Nusszweig abgebrochen, du? Warum? Ich werde dich auffressen.“ Beugend vor Schreck und zitternd sprach der Kaufmann: „O lieber Bär, friss mich nicht, und lass mich mit dem Nusszweiglein meines Weges ziehen, ich will dir auch einen großen Schinken und viele

Würste dafür geben!“ Aber der Bär brüllte wieder: „Behalte deinen Schinken und deine Würste! Nur wenn du mir versprichst, mir dasjenige zu geben, was dir zu Hause am ersten begegnet, so will ich dich nicht fressen.“ Dies ging der Kaufmann gerne ein, denn er gedachte, wie sein Pudel gewöhnlich ihm entgegenlaufe, und diesen wollte er, um sich das Leben zu retten, gerne opfern. Nach derbem Handschlag tappte der Bär ruhig ins Dickicht zurück; und der Kaufmann schritt, aufatmend, rasch und fröhlich von dannen. Der goldene Nusszweig prangte herrlich am Hut des Kaufmanns, als er seiner Heimat zueilte. Freudig hüpfte das jüngste Mägdlein ihrem lieben Vater entgegen; mit tollen Sprüngen kam der Pudel hinterdrein, und die ältesten Töchter und die Mutter schritten etwas weniger schnell aus der Haustüre, um den Ankommenden zu begrüßen. Wie erschrak nun der Kaufmann, als seine jüngste Tochter die erste war, die ihm entgegenflog! Bekümmert und betrübt entzog er sich der Umarmung des glücklichen Kindes und teilte nach den ersten Grüßen den Seinigen mit, was ihm mit dem Nusszweig widerfahren. Da weinten nun alle und wurden betrübt, doch zeigte die jüngste Tochter den meisten Mut und nahm sich vor, des Vaters Versprechen zu erfüllen. Auch ersann die Mutter bald einen guten Rat und sprach: „Ängstigen wir uns nicht, meine Lieben, sollte je der Bär kommen und dich, mein lieber Mann, an dein Versprechen erinnern, so geben wir ihm, anstatt unsrer Jüngsten, die Hirtentochter, mit dieser wird er auch zufrieden sein.“ Dieser Vorschlag galt, und die Töchter waren wieder fröhlich und freuten sich recht über diese schönen Geschenke. Die Jüngste trug ihren Nusszweig immer bei sich; sie gedachte bald gar nicht mehr an den Bären und an das Versprechen ihres Vaters. Aber eines Tages rasselte ein dunkler Wagen durch die Straße vor das Haus des Kaufmanns, und der hässliche Bär stieg heraus und trat brummend in das Haus und vor den erschrockenen Mann, der Erfüllung seines Versprechens begehrend. Schnell und heimlich wurde die Hirtentochter, die sehr hässlich war, herbeigeholt, schön geputzt und in den Wagen des Bären gesetzt. Und die Reise ging fort. Draußen legte der Bär sein wildes zotteliches Haupt auf den Schoß der Hirtin und brummte:

„Graue mich, grabble mich,
Hinter den Ohren zart und fein,
Oder ich fress dich mit Haut und Bein!“

Und das Mädchen fing an zu grabbeln; aber sie machte es dem Bären nicht recht, und er merkte, dass er betrogen wurde; da wollte er die geputzte Hirtin fressen, doch diese sprang rasch in ihrer Todesangst aus dem Wagen. Darauf fuhr der Bär abermals vor das Haus des Kaufmanns und forderte furchtbar drohend die rechte Braut. So musste denn das liebliche Mägdlein herbei, um nach schwerem bittren Abschied mit dem hässlichen Bräutigam

fortzufahren. Draußen brummte er wieder, seinen rauen Kopf auf des Mädchens Schoß legend:

„Graue mich, grabble mich,
Hinter den Ohren zart und fein,
Oder ich fress dich mit Haut und Bein!“

Und das Mädchen grabbelte, und so sanft, dass es ihm behagte und dass sein furchtbarer Bärenblick freundlich wurde, so dass allmählich die arme Bärenbraut einiges Vertrauen zu ihm gewann. Die Reise dauerte nicht gar lange, denn der Wagen fuhr ungeheuer schnell, als brause ein Sturmwind durch die Luft. Bald kamen sie in einen sehr dunklen Wald, und dort hielt plötzlich der Wagen vor einer finstergähnenden Höhle. Diese war die Wohnung des Bären. Oh, wie zitterte das Mädchen! Und zumal da der Bär sie mit seinen furchtbaren Klauenarmen umschlang und zu ihr freundlich brummend sprach: „Hier sollst du wohnen, Bräutchen, und glücklich sein, so du drinnen dich brav benimmst, dass mein wildes Getier dich nicht zerreißt.“ Und er schloss, als beide in der dunklen Höhle einige Schritte getan, eine eiserne Türe auf und trat mit der Braut in ein Zimmer, das voll von giftigem Gewürm angefüllt war, welches ihnen gierig entgegnungelte. Und der Bär brummte seinem Bräutchen ins Ohr:

„Sieh dich nicht um!
Nicht rechts, nicht links;
Gerade zu, so hast du Ruh!“

Da ging auch das Mädchen, ohne sich umzublicken, durch das Zimmer, und es regte und bewegte sich so lange kein Wurm. Und so ging es noch durch zehn Zimmer, und das letzte war von den scheußlichsten Kreaturen angefüllt, Drachen und Schlangen, giftgeschwollenen Kröten, Basilisken und Lindwürmern. Und der Bär brummte in jedem Zimmer:

„Sieh dich nicht um!
Nicht rechts, nicht links;
Gerade zu, so hast du Ruh!“

Das Mädchen zitterte und bebte vor Angst und Bangigkeit wie in Espenlaub, doch blieb sie standhaft, sah sich nicht um, nicht rechts, nicht links. Als sich aber das zwölfte Zimmer öffnete, strahlte beiden ein glänzender Lichtschimmer entgegen, es erschallte drinnen eine liebliche Musik, und es jauchzte überall wie Freudengeschrei, wie Jubel. Ehe sich die Braut nur ein wenig besinnen konnte, noch zitternd vom Schauen des Entsetzlichen und nun wieder dieser überraschenden Lieblichkeit - tat es einen furchtbaren Donnerschlag,

also dass sie dachte, es breche Erde und Himmel zusammen. Aber bald ward es wieder ruhig. Der Wald, die Höhle, die Gifttiere, der Bär - waren verschwunden; ein prächtiges Schloss mit goldgeschmückten Zimmern und schön gekleideter Dienerschaft stand dafür da, und der Bär war ein schöner junger Mann geworden, war der Fürst des herrlichen Schlosses, der nun sein liebes Bräutchen an das Herz drückte und ihr tausendmal dankte, dass sie ihn und seine Diener, das Getier, so liebeich aus seiner Verzauberung erlöset. Die nun so hohe, reiche Fürstin trug aber noch immer ihren schönen Nussweig am Busen, der die Eigenschaft hatte, nie zu verwelken, und trug ihn jetzt nur noch um so lieber, da er der Schlüssel ihres holden Glückes geworden. Bald wurden ihre Eltern und ihre Geschwister von diesem freundlichen Geschick benachrichtigt und wurden für immer, zu einem herrlichen Wohlleben, von dem Bärenfürsten auf das Schloss genommen.

D DVD-ROM

Inhalt der DVD-ROM:

- Quellcodes des Prototyps,
- Fragebögen im PDF-Format,
- Präsentationsfolien im PDF-Format,
- Hörspiele im MP3-Format,
- XML-Dateien von GermaNet,
- Sicherungsdateien der MySQL-Datenbanken.

Eidesstattliche Versicherung

Ich versichere hiermit an Eides Statt, dass ich die vorliegende Arbeit selbständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus Veröffentlichungen oder anderen Quellen entnommen sind, sind als solche eindeutig kenntlich gemacht. Die Arbeit ist in gleicher oder ähnlicher Form noch nicht veröffentlicht und noch keiner Prüfungsbehörde vorgelegt worden. Mit der Überprüfung der vorliegenden Arbeit hinsichtlich dieser Angaben durch ein informationstechnisches Prüfverfahren erkläre ich mich ausdrücklich einverstanden.

Des Weiteren erkläre ich mich damit einverstanden, dass meine Diplomarbeit in der Instituts- und der Universitätsbibliothek ausgelegt und zur Einsichtnahme aufbewahrt werden darf.

Ort, Datum

Unterschrift