
Abschlussbericht Projekt Gendercheck “Gleichstellung in der Sprache”

Der Gleichstellungseditor: Ein Tool zur sprachlichen Gleichbehandlung der Geschlechter in der deutschen Sprache

Anne Altmayer¹, — Michael Carl² — Sandrine Garnier² — Johann Haller^{1,2} — Bärbel Miemietz¹

¹Universität des Saarlandes
66123 Saarbrücken, Germany

²Institut für Angewandte Informationsforschung,
Martin-Luther-Straße 14,
66111 Saarbrücken, Germany

ABSTRACT. In this working paper we describe the “Gendercheck Project”, a project which uses the controlled-language authoring tool (CLAT) as a platform to check German texts for gender discriminatory use of language. As the use of non-discriminatory language is nowadays compulsory in administration and official documents, a number of guidelines and recommendations exist, which help to avoid gender imbalance and stereotypes in language use.

Within the gender-check project, we have modified the style proofing function and implemented a new set of style rules which detect woman-discriminatory formulations. We investigate three main problems:

- 1) generic use of masculine nouns which are meant to implicate both sexes
- 2) use of masculine pronouns following masculine nouns or gender-neutral indefinite pronouns
- 3) missing agreement between two nominal phrases

The results of the “Gendercheck Editor” show impressive results for precision and recall of 90% on German legal texts.

KEYWORDS: gender mainstreaming, controlled language

Contents

1	Einleitung	3
2	Sprachliche Ungleichbehandlung der Geschlechter	4
2.1	Historischer Hintergrund	4
2.2	Administrative Richtlinien	5
3	Umsetzung des Projekts	7
3.1	Controlled-Language Authoring Technology	7
3.2	Der Gleichstellungseditor	8
3.3	Fehlertypen und Umformulierungsmöglichkeiten	9
3.3.1	Generische Personenbezeichnungen im Maskulinum	10
3.3.2	Pronomen im Maskulinum	10
3.3.3	Fehlende Kongruenz von Subjekt und Prädikatsnomen im Genus	10
3.3.4	Korrektur und Lösungsvorschläge	10
4	Markierungsstrategie für den Gleichstellungseditor	11
4.1	Fehlerklasse 1: Markierung der Peronenbezeichnungen im Maskulinum	11
4.1.1	Ausfiltern “gegendeter” Formulierungen	12
4.1.2	Nicht markierte Formulierungen	14
4.2	Fehlerklasse 2: Markierung von Pronomen	15
4.3	Fehlerklasse 3: Prädikatsnomen	16
5	Bisheriger Stand der Korrekturvorschläge	17
5.1	Nomen im Maskulinum	17
5.2	Substantivierte Adjektive	17
5.3	Indefinitpronomen	17
5.4	Das Indefinitpronomen “mancher”	18
5.5	Die Konstruktion “derjenige, der”	18
5.6	Pronomen	18
5.7	Relativpronomen	19

5.8	Prädikatsnomen	19
6	Auswertung des Gleichstellungseditors	20
7	Grenzen und Möglichen des Gleichstellungseditors	22
7.1	Word-Integration	22
7.2	Spezifizierung der Korrekturvorschläge	23
7.3	Verbreitung und Bekanntmachung	23
8	Bibliographie	23

1. Einleitung

Mit der Verabschiedung des Bundesgleichstellungsgesetzes am 30.11.2001 und dem darin vorgeschriebenen Gleichheitsgrundsatz auch in der Sprache hat das Thema “Geschlechtergerechte Formulierung” an Aktualität gewonnen.

Das Überprüfen von Texten auf geschlechtergerechte Formulierungen ist jedoch zeitaufwendig und die notwendigen Umformulierungen erfordern ein hohes Maß an Übung und Erfahrung, das besonders durch ungeschulte Autoren und Autorinnen nicht immer gewährleistet ist. Darüber hinaus sind die entsprechenden Richtlinien nicht allgemein bekannt. Ein Werkzeug, das Texte auf geschlechtergerechte Formulierungen hin überprüft und zweifelhafte Stellen anmerkt ist deshalb ein wichtiger Schritt, um Geschlechtergleichstellung in geschriebenen Verwaltungstexten durchzusetzen und die Schreibenden für dieses Thema zu sensibilisieren. Das Projekt “Gendercheck” hat sich die Entwicklung eines solchen Werkzeugs zum Ziel gesetzt, den so genannten “Gleichstellungseditor”. Dieser Bericht fasst die Ergebnisse des Projekts zusammen.

Zunächst geben wir einen historischen Hintergrund über die sprachliche Ungleichbehandlung der Geschlechter und die Entwicklung einer Gegenbewegung bis heute. Obwohl eine grosse Anzahl von Richtlinien und Leitfäden für geschlechtergerechte Formulierungen heute vorliegen, fehlt es an konkreten Hilfen für Autoren.

Der Gleichstellungseditor benutzt ein Softwarepaket zur linguistischen Kontrolle technischer Texte. Im Abschnitt 3 stellen wir die Technologie vor, auf der die Realisierung des Gleichstellungseditors basiert. Dieser Abschnitt untersucht die Fehlertypen und Umformulierungsmöglichkeiten für die Erstellung geschlechtergerechter Texte. Zwei Hauptkategorien lassen sich isolieren: geschlechterneutrale Formulierungen und — wo dies nicht möglich ist — Nennung von Paarform. Ein Modul zur Detektion syntaktischer Fehlformulierungen wurde entsprechend auf geschlechtsspezifische den Vorgaben entwickelt und angepasst.

Zunächst werden potentielle Fehler markiert. Dann werden die Markierungen in korrekten Formulierungen ausgefiltert. Die verbleibenden Markierungen werden mit Fehlercodes versehen und können im Editor sichtbar gemacht werden. Die Abschnitte 4 und 5 beschreiben die Strategie und verwendeten Techniken im Detail.

Im Abschnitt 6 werten wir den Editor aus. In den von uns untersuchten Texten konnte der Editor eine Präzision von über 98% erreichen und ein Recall von 85%.

Im Abschnitt 7 nennen wir die Grenzen des Möglichen mit der bestehenden Technik.

2. Sprachliche Ungleichbehandlung der Geschlechter

2.1. *Historischer Hintergrund*

Die sprachliche Ungleichbehandlung der Geschlechter im Deutschen ist eine historisch gewachsene Tatsache, die lange Zeit ein in gewisser Weise zutreffendes Abbild der Gesellschaft lieferte und kaum in Frage gestellt wurde. Mit der Veränderung geschlechtstypischer Rollenbilder vor allem ab der Mitte des 20. Jahrhunderts wurde der Widerspruch zwischen der Lebenswirklichkeit von Frauen und der Art und Weise, wie über sie gesprochen wurde, immer deutlicher. Frauen wollten sich nun nicht mehr durch Maskulina mitgemeint fühlen müssen und keine männlichen Berufsbezeichnungen mehr tragen, sie wollten explizit als Frauen angesprochen werden, und zwar symmetrisch zur Ansprache der Männer.

Engagierte Linguistinnen setzten sich - durch erste Arbeiten in den USA inspiriert - seit den 70er Jahren auch in Deutschland intensiv mit der sprachlichen Ungleichbehandlung von Frauen und Männern auseinander. Die Arbeiten von Luise F. Pusch[PUS 84, PUS 90] und Senta Trömel-Plötz[TRö 82, TRö 88] wurden wegweisend für die (west-) deutsche "feministische Linguistik". Die von Pusch und Trömel-Plötz gemeinsam mit Ingrid Guentherodt und Marlies Hellingeri [GüN 80] verfassten "Richtlinien zur Vermeidung sexistischen Sprachgebrauchs" die in diesem Aufsatz gegebene Definition von "sexistischer Sprache" inhaltlich bis heute ihre Bedeutung behalten und dürften in der Literatur zum Thema einen ungeschlagenen impact factor haben.

Für die feministische Linguistik (die, nebenbei bemerkt, auch die fachwissenschaftliche Diskussion bereicherte, z.B. bei Prototypensemantik, Referenztheorie und Theorie der Markiertheit) gab es große Unterstützung auch aus anderen Wissenschaften (vor allem Soziologie, Psychologie und Rechtswissenschaft). Durch Tests konnte gezeigt werden, dass prototypische Vorstellungen z.B. "den Rechtsanwalt" "den Arzt" besetzten und dass Frauen sich keineswegs automatisch angesprochen fühlten, wenn in Stellenanzeigen ein "Ingenieur" ein "Betriebswirt" wurde; es wurde der Nachweis erbracht, dass Frauen tatsächlich mit sprachlichen Argumenten von bestimmten Rechten ausgeschlossen worden waren (so eine studierte Schweizer Juristin, der das Bundesgericht 1887 keine Zulassung als Rechtsanwältin erteilte, weil sie eben kein

“Schweizer”, dem verfassungsgemäß dieses Recht hätte erteilt werden müssen); und es wurden Textbelege dafür zusammengetragen, dass angeblich geschlechtsneutrale Maskulina Frauen nicht wirklich mitmeinen, wie bis dahin geglaubt und später noch lange behauptet wurde.

Die Ergebnisse (sprach-)wissenschaftlicher Untersuchungen wurden in die Gleichstellungspolitik aufgenommen, was wiederum auf die Sprache zurückwirkte. Aufgrund einer EG-Richtlinie von 1975 zur “Verwirklichung des Grundsatzes der Gleichbehandlung von Frauen und Männern beim Zugang zu Arbeitsplätzen, Berufsausbildung, Beförderung und in bezug auf die Arbeitsbedingungen” 1980 [] das Bürgerliche Gesetzbuch der Bundesrepublik Deutschland (nach einem Vertragsverletzungsverfahren wegen Untätigkeit) um das Verbot, Arbeitsplätze nur für Frauen oder nur für Männer auszuschreiben, ergänzt werden. Deutsche Stellenanzeigen haben sich in der Folge sichtbar gewandelt, und das jährlich neu aufgelegte Verzeichnis der Ausbildungsberufe führt seitdem konsequent männliche und weibliche Berufsbezeichnungen an.

Insbesondere von Mitte der 80er bis Mitte der 90er Jahre folgten parlamentarische Debatten und ministerielle Arbeitsgruppen, die in Gleichstellungsgesetze und in amtliche Richtlinien für einen geschlechtsneutralen Sprachgebrauch mündeten. Heute ist in weiten Bereichen zumindest für den offiziellen Sprachgebrauch eine geschlechtergerechte Sprache verbindlich vorgeschrieben: Es gibt eine Empfehlung des Ministerkomitees des Europarates zur Vermeidung von Sexismus in der Sprache, es gibt einen Bundestagsbeschluss mit entsprechendem Inhalt, und im Saarländischen Landesgleichstellungsgesetz gibt es einen Paragraphen, der die sprachliche Gleichbehandlung von Frauen und Männern vorschreibt. Diese Regelungen haben auch in einem inzwischen veränderten Kontext Bestand: In der Politik hat ein Paradigmenwechsel stattgefunden weg von einem defizitorientierten Ansatz der Frauenförderung hin zum Gender mainstreaming, d.h. hin zu einer Gleichstellungsarbeit, die in allen Ressorts und bei allen Entscheidungen Führungsstrategie und Querschnittsaufgabe sein soll und die beide Geschlechter im Blick hat. Parallel dazu haben sich die Wissenschaften, hier also die Linguistik, von einem einseitig frauenbezogenen Ansatz zu einem umfassenden, weibliche und männliche Perspektive einschließenden Ansatz erweitert: Heute befasst sich nicht mehr die feministische Linguistik mit der sprachlichen Diskriminierung von Frauen, sondern die Gender Linguistik mit einer Sprache, die beiden Geschlechtern gerecht wird (und unter dem Blickwinkel der “Verschiedenheit” in den diversity studies auch die Dichotomie männlich-weiblich noch aufgelöst).

2.2. Administrative Richtlinien

Eine wirklich geschlechtergerechte Sprache zu gebrauchen, ist noch immer in mancher Hinsicht ungewohnt und kann in Einzelfällen sogar schwierig sein. Das ist nach Jahrhunderten eingeschliffener Sprechweisen nicht erstaunlich, und zurecht können die Sprecherinnen und Sprecher des Deutschen, von denen nur wenige über eine sprachwissenschaftliche Ausbildung verfügen, erwarten, bei der Änderung ihrer Sprechgewohnheiten Hilfestellung zu erhalten.

Teilweise gibt es diese Hilfestellung auch bereits: Das Deutsche gehört nicht nur - neben dem Englischen - zu den unter dem Geschlechteraspekt derzeit am besten erforschten Sprachen, insbesondere in den letzten zwei Jahrzehnten ist auch eine Vielzahl von Leitfäden, Richtlinien o.ä. entstanden, die all denen, die eine geschlechtergerechte Ausdrucksweise verwenden wollen, Regeln und Beispiele für eine "politisch korrekte" zugleich stilistisch gute und abwechslungsreiche Sprache an die Hand geben. Es gibt Leitfäden von Linguistinnen, die sich gewissermaßen an alle wenden, und es gibt auch Leitfäden, die in einem bestimmten Kontext stehen und eine ganz bestimmte Zielrichtung haben, z.B. von der Deutschen UNESCO-Kommission [Deu93] oder von der Schweizerischen Bundeskanzlei [Sch96], von kommunalen Frauenbeauftragten oder von einzelnen Hochschulen [GÜN 80, HEL 93, HÄB 92, Deu93, Bun02, Arb97, Sch96]

Dass der Bedarf groß ist, zeigt die Erfahrung aus drei Jahren Gleichstellungsarbeit an der Universität des Saarlandes. Vom Vorlesungsverzeichnis und der Internetpräsentation der Universität über Studien- und Prüfungsordnungen bis hin zu Formularen und offiziellen Schreiben gibt es kaum einen Bereich, in dem nicht schon Nachbesserung angemahnt werden musste. Die Kooperationsbereitschaft derjenigen, die die Texte schreiben, ist groß, und nur selten muss auf den § 28 des Saarländischen Landesgleichstellungsgesetzes hingewiesen werden. Wie jedoch die Überarbeitung von statten geht, ist bisher weitgehend dem Zufall überlassen. Bestimmte Abteilungen der Universität haben hierfür inzwischen, teilweise in Zusammenarbeit mit der Frauenbeauftragten, Routinen entwickelt, eine in allen Bereichen anwendbare, effiziente Methode, die Texte zu überarbeiten, gibt es jedoch nicht. Die Benutzung der existierenden Leitfäden erfordert eine Einarbeitung und einen Zeitaufwand, der aus den verschiedensten Gründen nur von den wenigsten erbracht werden kann. Im übrigen sind diese Leitfäden außerhalb von linguistisch oder gleichstellungspolitisch interessierten Kreisen bisher so gut wie unbekannt.

An diesem Punkt liegt es nahe, eine Software zu erstellen, die ähnlich einem Rechtschreibprogramm arbeitet und die den Verfasserinnen und Verfassern von Texten eine nicht geschlechtergerechte Sprache signalisiert und ihnen evtl. auch bereits Korrekturvorschläge anbietet. Diese Software zu erstellen, ist auf der Grundlage der allgemeinen Linguistik und mit den bisherigen Ergebnissen der sprachwissenschaftlichen Geschlechterforschung möglich. Die nichtkorrekten Ausdrucksweisen, lassen sich, wie dies bereits in den schriftlich vorliegenden Leitfäden geschehen ist, zu bestimmten Typen von Entgleisungen zusammenfassen, die unterschiedliche Korrekturvorschläge nach sich ziehen. Ein Computerprogramm kann diese Korrekturvorschläge nicht nur erheblich schneller, sondern auch passgenauer zur Verfügung stellen als ein schriftlicher (und damit endlicher, veraltender und möglicherweise das eigene Sachgebiet nicht berücksichtigender) Leitfaden.

3. Umsetzung des Projekts

Im Rahmen des Gendercheck Projekts haben wir einen Editor zur Kontrolle technischer Dokumente (controlled-language authoring technology (CLAT)) an die Besonderheiten zur Kontrolle von Verwaltungstexten auf geschlechtergerechte Formulierungen angepasst. In diesem Abschnitt beschreiben wir zunächst den Hintergrund und die Funktionsweise von CLAT. Wir zeigen dann, wie die Technologie für den Gendercheckeditor nutzbar gemacht wurde. Im Abschnitt 4 und 5 gehen wir dann detailliert auf die Markierungsstrategie ein und geben Beispiele, wie dies in Lösungsvorschläge umgesetzt wird.

3.1. *Controlled-Language Authoring Technology*

CLAT wurde ursprünglich entwickelt aus der Notwendigkeit einiger Firmen, technische Texte automatisch auf allgemeine und firmenspezifische Sprachkonventionen hinzu überprüfen. Dabei wird der Text auf mehrere Gesichtspunkte hin untersucht:

- 1) Orthographische Korrektheit
- 2) Abgleich mit firmeneigener Terminologie und Abkürzungen
- 3) Grammatische Kontrolle nach allgemeinsprachlichen und firmenspezifischen Konventionen
- 4) Stilistische Kontrolle nach allgemeinsprachlichen und firmenspezifischen Konventionen

In der Orthographiekontrolle wird der Text auf mögliche Rechtschreibfehler untersucht und gegebenenfalls ein (oder mehrere) Alternativen vorgeschlagen. Die Terminologiekomponente gleicht den Text mit einer Terminologie- und Abkürzungsdatenbank ab, wobei auch terminologische Varianten gefunden werden. Die grammatische Kontrolle überprüft den Text auf grammatische Richtigkeit und desambiguiert Bedeutungen soweit es die allgemeinsprachlichen und firmenspezifischen Konventionen zulassen, während die stilistische Kontrolle Texte auf firmenspezifische stilistische Vorschriften untersucht. Diese Komponenten bauen zwar aufeinander auf, sie sind aber so gestaltet, dass einzelne Module ausgetauscht werden können und so einfach an andere Anforderungen und Texttypen angepasst werden können.

Neben den inhaltlichen Kontrollmechanismen bietet CLAT auch eine graphische Oberfläche, die es erlaubt, markierte Textsegmente farblich zu unterlegen. Einzelne Fehlercodes können aus- oder zugeschaltet, und Textsegmente bearbeitet oder übersprungen werden, je nach Augenmerk des Autors. CLAT erlaubt jedoch auch, die graphische Oberfläche auszuschalten und statt dessen eine XML Annotation zu erzeugen, die dann z.B. in einem Batch Mode weiterverarbeitet werden kann.

CLAT wurde bewusst modular konzipiert, um den Anforderungen verschiedener Firmen gerecht zu werden: auf einer funktionell-inhaltlichen Ebene können einzelne

Kontrollmodule zu- oder abgeschaltet werden, in seinem graphischen Interface können Fehlermeldungen editiert und gesteuert werden.

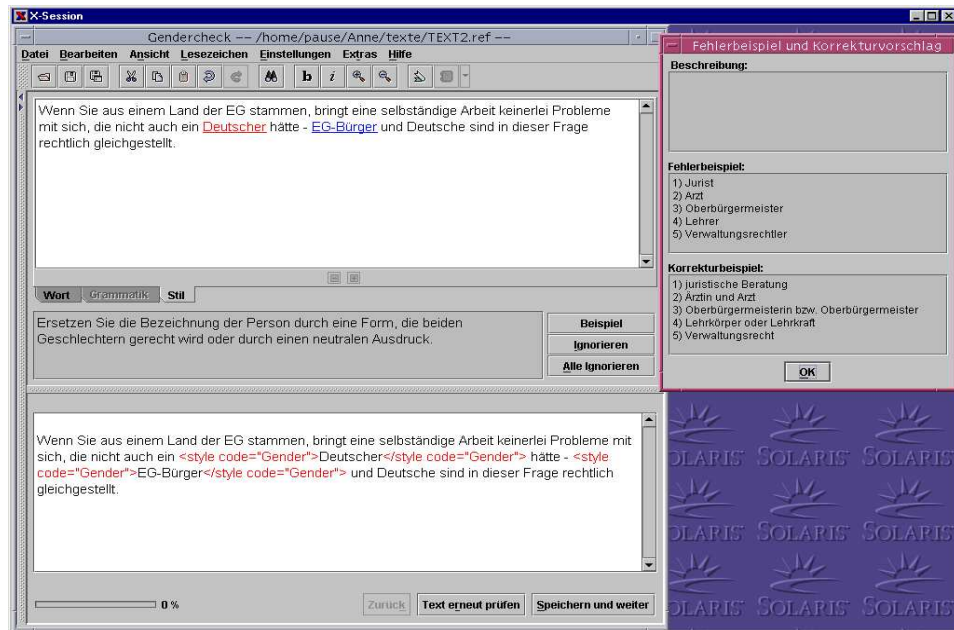


Figure 1. *Der Gleichstellungseditor*

3.2. *Der Gleichstellungseditor*

Das Projekt “Gendercheck” nutzt diese offene Struktur von CLAT, wobei ein Großteil der vorhandenen Ressourcen auch hier nutzbar gemacht werden konnte. Während die Orthographie- und allgemesprachliche Grammatikkontrolle ohne Änderung auch für den Gleichstellungseditor genutzt werden kann, wurden die Terminologie- und Abkürzungskontrolle ausgeschaltet. Im Rahmen dieses Projekts werden geschlechtergerechte Umformulierungsvorschläge als stilistische Kontrolle angesehen; entsprechend wurde die Style-Kontrolle an die Erfordernisse des Projekts angepasst.

Das Bild 1 zeigt den graphischen Editor von CLAT. Im unteren Teil des Editors wird der Eingabetextes dargestellt während der obere Teil des Editors das gleiche Textsegment mit den automatisch detektierten Fehler zeigt. Der Eingabesatz ist in Beispiel 1 wiederholt.

Beispiel 1:

Wenn Sie aus einem Land der EG stammen, bringt eine selbständige Arbeit keinerlei Probleme mit sich, die nicht auch ein Deutscher hätte - EG Bürger und Deutsche sind in dieser Frage rechtlich gleichgestellt.

Wie wir in Abschnitt 6 besprechen, werden geschlechtsunneutrale Formulierungen manuell in einer SGML annotation markiert. Die SGML-markierten Stellen “Deutscher” und “EG Bürger” sind hier identisch in beiden Fenstern.

Der Autor kann auf eines der im oberen Fenster markierten Stellen klicken, um eine entsprechende Fehlermeldung im mittleren Fenster zu erhalten. Weiter Informationen können in einem zusätzlichen Fenster (hier rechts) angezeigt werden.

Das Hauptaugenmerk der Neuentwicklung lag auf der Implementierung bzw. Modifikation einer stilistischen Regelkomponente und den zugehörigen Meldungen des graphischen Interface. Die Arbeit in der ersten Hälfte des Projektes kennzeichnete sich hauptsächlich durch einen Iterationszyklus, in dem wiederholt die folgenden beiden Fragen verfeinert wurden:

- 1) Welche Umformulierungen sollen vorgeschlagen werden?
- 2) Welche Fehlformulierungen können automatisch erkannt werden?

Während die erste Fragestellung einen benutzerorientierten Standpunkt einnimmt und versucht, möglichst detaillierte Lösungsvorschläge für geschlechterneutrale Formulierungen zu erarbeiten, geht die zweite Fragestellung von den technischen Möglichkeiten aus, Fehlformulierungen präzise und automatisch zu erkennen und zu benennen. Beide Ziele sind nur bedingt vereinbar, da die Fehlerquote zunimmt, je präziser und feinkörniger die automatische Fehleranalyse ausfällt. Die automatische Fehlererkennung dessen Markierung und eine entsprechende Fehlermeldung sind in CLAT jedoch unabhängig und werden in den folgenden Abschnitten gesondert besprochen.

3.3. Fehlertypen und Umformulierungsmöglichkeiten

Zur Verwirklichung des Gleichheitsgrundsatzes auf sprachlicher Ebene gilt es in erster Linie, Alternativen für die als Oberbegriff verstandene Personenbezeichnung im Maskulinum (“generisches Maskulinum”) zu finden. Da es sich bei den untersuchten Texten ausschließlich um Formulierungen aus den Bereichen Recht und Verwaltung handelt, können wir grundsätzlich von einer unspezifischen Referenz ausgehen, d.h. eine Personenbezeichnung im Maskulinum bezeichnet nicht eine konkrete Person, sondern bezieht sich auf alle Personen gleichermaßen, gleich welchen Geschlechts. Wir haben uns also zunächst darauf konzentriert, diese Personenbezeichnungen im Maskulinum sowie die ihnen nachfolgenden maskulinen Pronomen herauszufiltern und entsprechende Lösungsvorschläge anzuzeigen.

Die zu berücksichtigenden Fehler werden in drei Klassen unterteilt: Generisches Maskulinum, Pronomen im Maskulinum und fehlende Kongruenz. Im Folgenden

führen wir diese Fehlerklassen näher aus. Die daraus resultierenden Umformulierungsmöglichkeiten können in zwei Klassen abgehandelt werden.

3.3.1. *Generische Personenbezeichnungen im Maskulinum*

Beispiel 2:

“Der Beamte muss den Anforderungen Genüge leisten.”

Dieser Klasse können auch die Indefinitpronomen “jeder” “jedermann” “mancher” gezählt werden, und zwar sowohl in substantivierter Form als auch in Artikelfunktion.

Beispiel 3:

“Jeder Kunde kann selbst entscheiden.”

“Jeder soll mithelfen.”.

3.3.2. *Pronomen im Maskulinum*

Pronomen im Maskulinum (Relativpronomen, Possessivpronomen, Personalpronomen), die einem Nomen oder einem substantivierten Indefinitpronomen im Maskulinum folgen.

Beispiele 4:

“Der Beamte muss seine Wohnung in der Nähe des Arbeitsplatzes suchen.”

“Jeder muss seinen Beitrag dazu leisten.” “Wer Rechte hat, der hat auch Pflichten.”

Nicht als fehlerhaft gelten die Indefinitpronomen “jemand” und “wer” wenn diesen keine weiteren Pronomen folgen.

3.3.3. *Fehlende Kongruenz von Subjekt und Prädikatsnomen im Genus*

Beispiel 5:

“Ihr Ansprechpartner ist Frau Müller”

3.3.4. *Korrektur und Lösungsvorschläge*

Die vorgesehenen Korrekturen sehen ausschließlich eine Verbesserung des Sprachgebrauchs vor und orientieren sich an der regelkonformen Schreibweise. Dementsprechend gilt die Schreibung mit einem großen I im Wortinnern nicht als Korrekturmöglichkeit. Das Gleiche gilt für die Ersetzung des Indefinitpronomens “man” durch “frau”

Die Lösungsvorschläge lassen sich demnach in zwei Hauptkategorien aufteilen:

1) Nennung von Paarformen, d.h. die feminine und die maskuline Form der Personenbezeichnung wird angezeigt, beide Formen werden mit “und” “oder” “bzw.”

verknüpft. Eine Verknüpfung mit Schrägstrich '/' wird ebenfalls als Lösung vorgeschlagen, soll aber nur in Formularen und Verordnungen Verwendung finden. Bei der Nennung der Paarformen wird die feminine vor der maskulinen Form genannt. Die Nennung einer der beiden Formen in Klammern wird ausgeschlossen.

2) Wann immer möglich, werden geschlechtsneutrale bzw. geschlechtsabstrakte Ausdrücke verwendet. Diesen soll gegenüber Paarformen der Vorzug gegeben werden, da letztere nicht selten die Lesbarkeit des Textes erschweren. Zur neutralen Formulierung sind insbesondere Kollektiva (z.B. "Lehrkörper" Arbeitnehmerschaft" sowie substantivierte Partizipien ("Studierende" und substantivierte Adjektive ("Berechtigte" geeignet. Beschreibung von CLAT

4. Markierungsstrategie für den Gleichstellungseditor

In diesem Abschnitt beschreiben wir die Markierungsstrategie im Detail, mit der, entsprechend der Fehlertypisierung in Abschnitt 3.3, automatisch ungedenderte Formulierungen ermittelt werden. Wir unterscheiden zwischen Markierungen von Personenbezeichnungen im Maskulinum, Markierungen von Pronomen, die sich auf ungedenderte Personenbezeichnungen beziehen und nicht-kongruenten Prädikatsnomen.

Die Fehlererkennungsstrategie besteht aus zwei Schritten. Zunächst werden alle Personenbezeichnungen markiert, sowohl die im Femininum als auch im Maskulinum. In einem zweiten Schritt werden Markierungen gelöscht und ausgefiltert, für die korrekte Formulierungen identifiziert wurden. Den verbleibenden markierten Stellen werden dann entsprechende Fehlercodes zugewiesen, die im CLAT Editor angezeigt werden können.

Die Markierungsstrategie benutzt flache Pattern, die im KURD Formalismus beschrieben sind [CAR 98].

4.1. Fehlerklasse 1: Markierung der Personenbezeichnungen im Maskulinum

Die Markierung der Personenbezeichnungen beruht im Wesentlichen auf zwei Mechanismen:

a) Die morphologische Analyse der Texte berechnet nicht nur Ableitungs- und Flexionsmerkmale der Wörter, sondern vergibt auch semantische Merkmale. So werden menschliche Agenten, wie "Soldat", "Bürgermeister", "Beamte", aber auch "Krankenschwester" und "Hausfrau" mit einem semantischen Merkmal *s=agent* versehen. Die Syle Komponente vom Gleichstellungseditor beruht im Wesentlichen auf diesem semantischen Merkmal und der Genusinformation.

b) Probleme bereiten Nomen, deren Grundwort als Ableitung eines nominalisierten Verbs gebildet wird. Beispielsweise werden "Gewichtheber" und "Busfahrer" nicht automatisch als menschliche Agenten markiert, da diese Wörter sowohl Dinge als auch Menschen bezeichnen können. Geschlechtergerechte Umformulierungen sind

aber nur für menschliche Agenten relevant. Um die Erfassung von Substantiven dieser Klasse dennoch gewährleisten zu können, wurden einige Wörter in eine Lexemliste aufgenommen, die mit großer Wahrscheinlichkeit menschliche Agenten in den untersuchten Texten bezeichnen. Der Vorteil einer mit Lexemen bestückten Einschlussliste besteht darin, dass ein Eintrag alle Wortformen des Lexems erfasst. So deckt beispielsweise das Lexem “absolvieren” die Formen “Absolvent”, “Absolventin”, “Absolventen” und “Absolventinnen” ab. Diese Liste wurde in der Form einer KURD Regel erstellt und enthält bislang die Lexeme folgender Wörter:

```
EinschlussListe =
    Ae{c=noun,
        ls:tour$;
        absolvieren$;
        dezernieren$;
        richten$;
        fahren$;
        administrieren$;
        vorstand$}
: Ag{style=agent}.
```

Weiterhin kann jedes dieser Lexeme alleine oder als Grundwort eines Kompositums erscheinen. Die o.g. KURD Regel prüft ob ein Lexem im zu prüfenden Text mit einem Eintrag in der Lexemliste endet. Falls dies der Fall ist wird ein entsprechendes Merkmal `style=agent` vergeben. So wird beispielsweise auch der “Netzadministrator” und der “Omnibusfahrer” als fehlerhaft markiert.

4.1.1. *Ausfö ltern ‘gegendeter’ Formulierungen*

Die markierten Stellen durchlaufen dann mehrere Filtermechanismen, die den Kontext überprüfen und gegebenenfalls Markierungen wieder löschen. Im Folgenden listen wir einige dieser Mechanismen auf.

a) Von der Fehlererkennung bewusst ausgenommen werden Nomina im Maskulinum, denen ein Familienname folgt, da dies auf ein spezifischen Referenten schliessen lässt. So wird in Beispiel 6 das Nomen “Beamte” nicht markiert, da ihm der Familienname “Meier” folgt:

Beispiel 6:

“Der Beamte Meier hat gegen die Vorschrift verstoßen.”

In KURD lässt sich das mit der Regel `AgentMitFname` folgendermassen darstellen: wenn ein Eigenname `s=fname` einer Sequenz von markierten Agenten `style=agent` folgt, werden die Agentmarkierungen durch die Operation `r{style=nil}` wieder gelöscht.

```

1 AgentMitFname =
2     +Ae{style=agent},
3     Ae{c=noun,s=fname}
4     : Ar{style=nil}.

```

b) Ebenfalls ausgenommen wurden substantivierte Adjektive und Partizipien im Plural, da diese sich zur geschlechtsneutralen Personenbezeichnung bestens eignen und ausdrücklich als Korrekturmöglichkeit vorgeschlagen werden.

Beispiel 7:

- a) “Ist bereits ein unabhängiger Sachverständiger bestellt?”
- b) “Sind bereits Sachverständige bestellt?”

In Beispiel 7a wird “Sachverständiger” die Markierung belassen, während bei Beispilesatz 7b die Markierung wieder gelöscht wird, da “Sachverständige” ambig in Bezug auf Genus ist und bereits feminine und maskuline Agenten bezeichnen kann.

c) Manche Nomen werden fälschlicherweise als maskuline Agenten markiert, für die entweder keine alternativen Formulierungen existieren, oder die ambige Lesarten haben. In die erste Kategorie fallen Wörter wie “Mensch”, “Gast”, “Flüchtling”, die alle maskulin sind, für die aber keine analoge feminine Form existiert. Diese Wörter wurden in eine Ausschlussliste aufgenommen, um eine Fehlermarkierung zu unterdrücken. Ähnlich wie bei der Lexemliste wird auch die Ausschlussliste in Form einer KURD regel realisiert. Auch hier ist nur das letzte Lexem eines Kompositas relevant.

```

1 Ausschlussliste =
2     Aa{style=agent,
3     lu:mensch$;
4     flüchtling$;
5     säugling$;
6     gast$;
7     rat$}
8     : Ar{style=nil}.

```

Eine Ambige Lesart hat beispielsweise das Wort “Rat”. Es kann eine Institution, ein Gebäude oder eine Person bezeichnen. Diese Lesarten können in aller Regel nicht hinreichend desambiguiert werden, sodass das Wort als fehlerhaft markiert würde. Dies gilt auch für Komposita, die semantisch eindeutig nicht als Personenbezeichnung identifizierbar sind, wie zum Beispiel in “Ältestenrat”:

Beispiel 8:

- “Der Präsident bestimmt im Benehmen mit dem Ältestenrat, welche Tage als Sitzungstage gelten.”

d) Auch Markierungen in bereits “gegenderten” Formulierungen werden gelöscht. In Beispiel 9 erscheint sowohl die Paarform “Die Beamtin oder der Beamte” als auch die sich darauf beziehenden Pronomen “sie oder er” und “ihrer oder seiner” als geschlechtergerechte Formulierung.

Beispiel 9:

“Die Beamtin oder der Beamte auf Lebenszeit oder auf Zeit ist in den Ruhestand zu versetzen, wenn sie oder er infolge eines körperlichen Gebrechens oder wegen Schwäche ihrer oder seiner körperlichen oder geistigen Kräfte zur Erfüllung ihrer oder seiner Dienstpflichten dauernd unfähig (dienstunfähig) ist.”

Eine entsprechende KURD Regel ist unten abgebildet. Der Beschreibungsteil der Regel in den Zeilen 2 bis 5 passt auf eine Konjunktion bei der die beiden markierten Agenten (`style=agent`) das gleiche Lexem `ls=_L` aber mit unterschiedlichem Genus haben müssen. Letzteres ist ausgedrückt durch die beiden Variablen `ehead={g=_G}` und `ehead={g~=_G}`, die nur dann unifizieren, wenn sie nicht identisch sind. Die Markierung wird in diesem Fall gelöscht `Ar{style=nil}`

```

1  gegendert =
2      Ae{style=agent,ls=_L,ehead={g=_G}},
3      e{lu=oder;und;bzw.;/},
4      *a{style~=agent}e{c=w},
5      Ae{style=agent,ls=_L,ehead={g~=_G}}
6      :  Ar{style=nil}.

```

Die Regel erlaubt nicht nur die Konjunktionen “und” sondern auch mit anderen üblichen Konjunkoren wie “oder”, “bzw.” und “/”.

4.1.2. Nicht markierte Formulierungen

Nicht als fehlerhaft erkannt werden Determinativkomposita, deren Bestimmungswort eine Person, deren Grundwort aber einen Gegenstand bezeichnet (z.B. “Rednerpult”, “Teilnehmerliste”). Wünschenswert wäre es auch hier Alternativvorschläge bereitzustellen, etwa “Redepult” oder “Teilnehmendeliste”. Ein entsprechender Mechanismus könnte aber einfach analog der Ausschluss- und Einschlusslisten implementiert werden, bei dem das Lexem der Liste nicht letzter sondern erster Bestandteil des Kompositums ist.

Nicht vorgesehen ist die Markierung von Artikeln und Adjektiven, die einem Nomen im Maskulinum vorausgehen. Diese Regelung wurde eingeführt, da bei mehreren Einfügungen zwischen Artikel und zugehörigem Nomen eine konsequente Markierung ohnehin nicht möglich bzw. unübersichtlich wäre. Im Beispielsatz 10 würde eine Markierung der gesamten Nominalphrase “der jeweils für die Aufgaben zuständige Dezernent” vermutlich eher von der eigentlichen Fehlerquelle ablenken.

Beispiel 10:

“Den Vorsitz führt der jeweils für die Aufgaben zuständige Dezernent.”

Nicht möglich ist eine Untersuchung über die Satzgrenze hinaus; dies hat zur Folge, dass ein nicht näher spezifiziertes Nomen im Maskulinum als fehlerhaft markiert, wenn der entsprechende Hinweis auf spezifische Referenz (z.B. ein Name) im darauffolgenden oder vorhergehenden Satz genannt wird: z.B.

“Herr Müller hat die Dienstvorschrift verletzt. Der Beamte ist somit zu entlassen.”

Hier würde “Beamte” als Fehler markiert werden, da die Information darüber, dass es sich um spezifische Referenz handelt, im vorhergehenden Satz enthalten ist.

4.2. Fehlerklasse 2: Markierung von Pronomen

Markiert werden auch Personalpronomen, Possesivpronomen, Relativpronomen und Indefinitpronomen. Die Strategie ist analog zu der Markierungsstrategie von Personenbezeichnungen. Zunächst werden alle in Frage kommenden Pronomen markiert, und in einem zweiten Schritt werden korrekte Formulierungen ausgefiltert. Ausser bei den Indefinitpronomen werden Ausdrücke dieser Klasse nur dann markiert, wenn sie sich auf ein markierten Agenten der Fehlerklasse 1 beziehen. Für Relativpronomen wird zunächst das erste Komma gesucht, das einem markierten Agenten folgt. Das unmittelbar darauf folgende Relativpronomen wird markiert, falls es den gleichen Genus wie der markierte Agent hat. Die entsprechende KURD Regel ist unten abgebildet.

```

1 RelativPronomen =
2     e{style=agent, ehead={g=_G}},
3     *a{lu~=&cm},
4     e{lu=&cm},
5     Ae{lu=d_rel, ehead={g=_G}}
6     : Ag{estyle=Gen2, bstyle=Gen2}.

```

Auch Personal und Possesivpronomen werden nur dann markiert, falls sie sich möglicherweise auf eine vorhergehende markierte Personenbezeichnung beziehen und kein weiter markierter Agent zwischen ihnen steht.

```

1 PersonalPronomen =
2     e{style=agent},
3     1Ae{lu=er; er_es, c=w, sc=pers} | e{s~=agent, sc~=punct}
4     : Ar{bstyle=Gen2, estyle=Gen2}.

```

Ausdrücklich ausgenommen von der Fehlererkennung wird “jeder” in der Funktion eines Artikels - entsprechend der oben erwähnten Regel, wonach Artikel und Adjektive, die einem Nomen im Maskulinum vorausgehen, nicht als fehlerhaft markiert werden.

```

1 PossesivPronomen =
2     e{style=agent},
3     1Ae{lu=sein, c=w, sc=poss} | e{s~=agent, sc~=punct}
4     :   Ar{style=agent, bstyle=Gen2, estyle=Gen2}.

```

In einem zweiten Schritt werden die Markierungen in “gegenderten” Formulierungen analog zu dem Vorgehen in Regel gegendert aus Abschnitt 4.1.1 wieder gelöscht.

Als problematisch hat sich die Erkennung von Pronomen in aus mehreren Satzteilen bestehenden Sätzen erwiesen. Hier werden Pronomen, die auf mehrere verschiedene Substantive folgen, nicht dem entsprechenden Substantiv zugeordnet und als fehlerhaft markiert.

Beispiel 11:

“Das gleiche gilt, wenn ein Mitglied des Bundestages in ein solches Dienstverhältnis berufen wird, von dem Tage an, mit dem er Ernennung wirksam wird”.

Wie schon in der Klasse der Nomen stellt auch hier Homonymie ein Problem dar: so wird “sein” auch dann als Possessivpronomen im Maskulinum erkannt wird, wenn es sich um ein Hilfsverb handelt.

4.3. Fehlerklasse 3: Prädikatsnomen

Die Kongruenz von Prädikatsnomen wird mit der folgenden KURD Regel erfasst

```

1 kongruenz =
2     +Ae{mark=np, style=agent} e{ehead={g=_G}},
3     *Ae{mark=np},
4     e{ls=sein, c=verb},
5     *Ae{style~=agent},
6     Ae{mark=np, style=agent} e{ehead={g~=_G}},
7     *Ae{mark=np}
8     :   Ar{bstyle=Gen3, estyle=Gen3}.

```

Die Zeilen 1 und 2 beschreiben das Subjekt, während die Zeilen 5 bis 7 das Prädikatsnomen erfassen. Beide Satzteile sind durch eine Kopula “sein” verbunden. Ähnlich der Regel wird auch hier nur eine Markierung erzeugt, wenn der Genus von beiden Satzteilen nicht identisch ist.

5. Bisheriger Stand der Korrekturvorschläge

Die Korrekturvorschläge bieten bislang eine Reihe verschiedenartiger Typen von Korrekturen an, die in diesem Abschnitt besprochen werden.

5.1. *Nomen im Maskulinum*

Wird ein Nomen im Maskulinum markiert, so erfolgt ein allgemeiner Korrekturvorschlag:

“Ersetzen Sie die Bezeichnung der Person durch eine Form, die beiden Geschlechtern gerecht wird oder durch einen neutralen Ausdruck.”

Ein solcher allgemeiner Korrekturvorschlag wird auch in Figure 1 dargestellt.

Zur Veranschaulichung werden verschiedene Fehler- und Korrekturbeispiele angeboten, wie auch im rechten Fenster in Figure 1 dargestellt. Die Fehlerbeispiele zeigen fünf verschiedene Substantive im Maskulinum, denen jeweils eine Korrekturmöglichkeit zugeordnet wird. Im Fall 1) wird die Personenbezeichnung “Jurist” durch die Bezeichnung der Tätigkeit “juristische Beratung” ersetzt; im zweiten, dritten und vierten Fall werden die Paarformen mit verschiedenen Formen der Verknüpfung vorgeschlagen; in Fall 5) soll die Personenbezeichnung “Verwaltungsrechtler” durch das Tätigkeitsgebiet “Verwaltungsrecht” ersetzt werden.

Welche der vorgeschlagenen Möglichkeiten Anwendung findet, liegt (noch) im Ermessen der Schreibenden.

5.2. *Substantivierte Adjektive*

Einen gesonderten Korrekturvorschlag erhalten substantivierte Adjektive und Partizipien im Singular. Hier wird die Pluralbildung vorgeschlagen, da Personenbezeichnungen in dieser Form stets geschlechtsneutral sind; z.B. die “Studierenden”, “die Beschäftigten”, “die Angehörigen”.

“Ersetzen Sie die Singularform des Substantivs durch die Pluralform. Beachten Sie die Kongruenz.”

5.3. *Indefinitpronomen*

Werden die Indefinitpronomen “jeder” und “jedermann” markiert, so erfolgt die Meldung:

“Überprüfen Sie, ob Sie diesen Ausdruck durch ‘Alle’ ersetzen können. Beachten Sie, dass auch Pronomen und Verben in Numerus übereinstimmen.”

Beispielsweise kann der Satz 12a durch 12b ersetzt werden.

Beispiel 12:

- a) "Jeder Kunde wird freundlich bedient"
- b) "Alle Kunden werden freundlich bedient"

5.4. Das Indefinitpronomen "mancher"

Eine eigene Meldung erhält das Indefinitpronomen "mancher". Hier soll geprüft werden, ob sich die Singularform des Pronomens durch die geschlechtsneutrale Pluralform ersetzt werden kann: Überprüfen Sie, ob Sie diesen Ausdruck durch die Pluralform "Manche" ersetzen können. Beachten Sie, dass auch Pronomen und Verben in Numerus übereinstimmen müssen.

Beispielsweise der Satz 13a durch 13b ersetzt werden.

Beispiel 13:

- a) "Manch einer macht sich seine Gedanken"
- b) "Manche machen sich ihre Gedanken"

5.5. Die Konstruktion "derjenige, der"

Ebenfalls durch Pluralbildung läßt sich im Falle von "derjenige, der" eine neutrale Formulierung finden. Der Korrekturvorschlag lautet:

"Überprüfen Sie, ob Sie diesen Ausdruck durch 'Diejenigen' ersetzen können. Beachten Sie, dass auch Pronomen und Verben in Numerus übereinstimmen müssen."

Als Beispiel erscheint in dem Hilfsfenster der Vorschlag 14a im Sinne von 14b zu verändern:

Beispiel 14:

- a) "Wir weisen darauf hin, daß derjenige, der den von seinem Hund hinterlassenen Kot nicht beseitigt, ordnungswidrig handelt."
- b) "Wir weisen darauf hin, daß diejenigen, die den von ihrem Hund hinterlassenen Kot nicht beseitigen, ordnungswidrig handeln."

5.6. Pronomen

Folgt auf ein markiertes Nomen oder Indefinitpronomen im Maskulinum ein Pronomen, so lautet der Korrekturvorschlag:

“Überprüfen Sie das Subjekt, auf das sich das Pronomen bezieht. Verändern Sie das Subjekt ggf. so, dass es beiden Geschlechtern gerecht wird und passen Sie das Pronomen an.”

Auch hier kann 15a durch 15b im Sinne geschlechtergerechter Formulierung verbessert werden:

Beispiel 15:

- a) “Es kann eine Geldbuße verhängt werden, wenn jemand seiner Verpflichtung zur Auskunft nicht nachkommt”
- b) “Es kann eine Geldbuße verhängt werden, wenn eine Person ihrer Verpflichtung zur Auskunft nicht nachkommt.”

5.7. Relativpronomen

Eine eigene Meldung erhalten die einem markierten Nomen nachfolgenden Relativpronomen, da diese in manchen Fällen auch ganz entfallen können. Auf diese Weise kann gegebenenfalls durch Streichung des Relativpronomens ein geschlechtsneutraler Satz formuliert werden. Da diese Möglichkeit für jeden Einzelfall geprüft werden muß, lautet der Korrekturvorschlag:

“Überprüfen Sie, ob das Relativpronomen entfallen kann.”

In Beispiel 16a kann das Relativpronomen problemlos entfallen.

Beispiel 16:

- a) “Wer zahlt, der hat auch mitzureden.”
- b) “Wer zahlt, hat auch mitzureden.”

5.8. Prädikatsnomen

Stimmen Nomen und Prädikatsnomen im Genus nicht überein, so erfolgt die Meldung:

“Achten Sie auf Übereinstimmung des Genus der Satzteile.”

Das Hilfsfenster enthält folgenden Beispielsatz 17a, der dementsprechend korrekt 17b lauten müsste:

Beispiel 17:

- a) “Ihr Ansprechpartner ist Frau Meier”
- b) “Ihre Ansprechpartnerin ist Frau Meier.”

6. Auswertung des Gleichstellungseditors

Die Auswertung des Gleichstellungseditors beruht auf zwei Texten.

ET₁ Ein Korpus bestehend aus unzusammenhängenden negativ Beispielsätzen aus der Literatur [Ges99, Sch96].

TT₂ Auszug aus dem Abgeordnetengesetz des Deutschen Bundestages

Text	Fehlertypen				Textgröße		
	Type 1	Type 2	Type 3	\sum	#Sätze	#Wörter	Fehler/Satz
ET ₁	97	62	15	174	95	1062	1,95
TT ₂	95	21	—	111	251	6473	0,44

Die geschlechtsspezifischen Stellen der Texte wurden zunächst manuell mit einer SGML Annotation versehen, wobei jeder SGML-code einem anderen Umformulierungsvorschlag entsprach. Während Text ET₁ eine gleichmäßige Verteilung verschiedener Fehlformulierungen aufweist, enthält Text TT₂ vor allem “ungegenderte” Pronomina.

Die Auswertung der Texte in Tabelle 6 zeigt, dass die Fehlertypen eine sehr unterschiedliche Verteilung und Dichte aufweisen. Am dichtesten gesät sind die Fehler im Text ET₁ in dem durchschnittlich jedes sechste Wort und jeder Satz fast zwei als “ungegendert” markierte Stellen enthält. Da es sich um negativ - Beispielsätze handelt sind die Fehlertypen hier relativ gleichmäßig verteilt. Die Verteilung und Dichte im Text TT₂ weicht davon sehr ab. Es fällt auf, dass der Fehlertyp 3, falsches Prädikatsnomen, in TT₂ nicht vorkommt. Auch sind problematische Stellen in diesem Text mit 0.44 Fehlern pro Satz sehr viel dünner gesät. Dabei ist zu beachten, dass ein “Satz” mit Durchschnittlich 25 Wörtern hier sehr viel länger ist. Der Grund hierfür ist, dass lange Aufzählungen als ein “Satz” segmentiert wurden. Daraus ergibt sich, dass nur annähernd jedes 60ste Wort ungegendert ist.

Zunächst wurde der text ET₁ benutzt, um das Markierungssystem und die Fehlermeldungen zu entwickeln und zu trimmen. Basierend auf fehlenden Meldungen (misses) und Falschalarm (noise) wurden dann Änderungen sowohl im CLAT Modul als auch an der Konzeptionierung der Fehlermeldungen und ihrer Granularität vorgenommen. So enthielt der Text ET₁ in seinem ersten Annotationsschema mehr als 30 verschiedene Fehlercodes für die drei Fehlerklassen, die den Umformulierungsvorschlägen der Gesellschaft für Informatik [Ges99] bzw. der Universität von Zürich [Uni00] entsprachen. Diese Granularität stellte sich aber als zu feinkörnig heraus, um als Modell einer automatischen Annotation zu dienen, und wurde vom Gleichstellungseditor in nur sehr unzureichender Weise reproduziert. In einem zweiten Durchlauf wurde nur ein Fehlercode annotiert, entsprechend etwa einer sehr unpräzisen Meldung “Überprüfen Sie diese Textstelle auf geschlechtergerechte Formulierung”. Nach einer Anpassung der Regeln konnten mehr als 95% Präzision und mehr als 89% Recall erreicht werden. Diese Ergebnisse dienten dann als Grundlage für eine Verfeinerung der Fehlermeldungen und der entsprechenden Umformulierungsvorschläge. Dabei kristallisierten sich die drei oben genannten Klassen

von Fehlermeldungen heraus, die einerseits wünschenswert erscheinen und die andererseits auch automatisch mit hoher Sicherheit unterschieden werden können.

Text ET ₁					
Fehler	hit	misses	noise	precision	recall
Type 1	85	12	1	0.988	0.876
Type 2	55	7	5	0.917	0.887
Type 3	15	0	1	0.937	1.000
Σ	155	19	7	0.957	0.891
Text TT ₂					
Type 1	86	9	5	0.945	0.905
Type 3	14	7	4	0.778	0.667
Σ	100	16	9	0.917	0.862

Die beiden Texte wurden dann zur Endauswertung im Batch Modus durch den Gleichstellungseditor geschickt und die automatisch annotierten Segmente wurden mit den manuell annotierten Segmenten verglichen. Das Resultat ist in Tabelle 6 abgebildet. Die Ergebnisse sind für den Gesetzestext TT₂ ein wenig schlechter als für den "Trainingstext" ET₁.

Rauschen für Meldungen vom Fehlertyp 1 (Personenbezeichnungen im generischen Maskulinum) werden zum einen vor allem durch Formen erzeugt, für die es keine femininen Äquivalente gibt, wie z.B. Ableitungen auf "-ling" wie "Abkömmling". Diese Wörter könnten dem Ausschlusslexikon beigefügt werden das in Zukunft auch benutzergesteuert erweiterbar gestaltet werden müsste. Andere Gründe für Falschalarm sind nominalisiert Partizipien, wie "Angestellte", "Hinterblieben", die zwar beide Geschlechter einschliessen, aber von dem Grammatikmodul desambiguiert wurden.

Fehlende Markierungen (misses) sind zum Grossteil auf einige Wörter zurückzuführen, die wie "Schriftführer" und "Ehegatte" nicht als Personenbezeichnungen analysiert werden. Diese Wörter könnten analog in das Einschlusslexikon aufgenommen werden. Auch das Einschlusslexikon müsste auf die Dauer benutzergesteuert erweiterbar sein.

Fehlende Markierungen (misses) bei Fehlern vom Typ 2 können auf Pronomen mit Bezugswort im vorhergehenden Satz zurückgeführt werden. Da das System satzbasiert arbeitet besteht hier keine einfache Möglichkeit, dies zu verändern. Im Satzbeispiel 18 bezieht sich das Possesivpronomen seiner im zweiten Satz auf den Agenten Bewerber im ersten Satz. Dieser Bezug kann nicht hergestellt werden, wenn die Analyse satzweise geschieht.

Beispiel 18

Einem Bewerber um einen Sitz im Bundestag ist zur Vorbereitung seiner Wahl innerhalb der letzten zwei Monate vor dem Wahltag auf Antrag Urlaub von bis zu zwei Monaten zu gewähren. Ein Anspruch auf Fortzahlung seiner Bezüge besteht für die Dauer der Beurlaubung nicht.

Ein Beispiel für Falschalarm vom Fehlertyp 2 ist in Beispiel 19 gezeigt. Das Relativpronomen der wurde vom Gleichstellungseditor erkannt und fälschlicherweise Beamte als Bezugswort identifiziert. Tatsächlich bezieht sich das Relativpronomen aber auf “Antrag” und ist somit geschlechterunbedenklich.

Beispiel 19

Der Beamte ist auf seinen Antrag, der binnen drei Monaten seit der Beendigung der Mitgliedschaft zu stellen ist, ...

Sehr viel mächtigere Mechanismen sind hier notwendig, um die Anzahl der misses und noise für Fehlertyp 2 zu verringern.

7. Grenzen und Möglichen des Gleichstellungseditors

Das ursprüngliche Vorhaben einer Unterscheidung nach Referenztypen, konnte nicht verwirklicht werden. Die hierfür erforderliche Einbeziehung des Kontextes, d.h. eine automatische Textkontrolle über die Grenzen des einzelnen Satzes hinaus ist bislang nicht möglich. Dieses Problem konnten wir durch Einschränkung der Textsorte auf den Bereich Rechts und Verwaltung teilweise umgehen, da es sich hierbei in der Regel um Texte mit unspezifischer Referenz handelt.

Das Problem der Referenz bleibt dennoch bestehen. Die Fehler, die dadurch entstehen, dass die einem Nomen zugrundeliegende Referenz nicht erkannt oder falsch bestimmt wird, können nicht vermieden werden.

7.1. Word-Integration

Eine breitere Verwendung kann der Editor nur finden, wenn er anwendungsfreundlich konzipiert ist. Aus diesem Grunde ist in einem nächsten Schritt die Integration in das Textverarbeitungsprogramm Word geplant. Testphase Nach erfolgreicher Word-Einbindung ist vorgesehen, das Programm einer ersten Testphase zu unterziehen. Eine erste Erprobung wird in der Verwaltung der Universität des Saarlandes erfolgen. Es ist vorgesehen, täglich anfallende Schreibarbeiten sowie bereits bestehende Verordnungen, Studienordnungen, Vorlesungsverzeichnisse etc. einer Prüfung auf geschlechtergerechte Formulierung hin zu unterziehen. Auswertung und Anpassung Die Ergebnisse dieser etwa einmonatigen Testphase werden anschließend in Bezug auf “noises” (fälschlicherweise angemahnte Formulierungen) und “misses” (nicht markierte Fehler) ausgewertet. Nach eingehender Analyse der Ergebnisse wird das Programm weitestgehend angepasst und die Fehlererkennung optimiert.

7.2. Spezifizierung der Korrekturvorschläge

Ist der Quotient der Fehlererkennung zufriedenstellend, so können in einem weiteren Schritt die Korrekturmeldungen weiter differenziert werden: anstelle der allgemeinen Korrekturvorschläge, die sich auf eine Fehlerklasse beziehen, soll sich dann die Lösungsmöglichkeit auf den jeweiligen konkreten Fehler beziehen. Nach Möglichkeit sollte bei den Korrekturvorschlägen die konkreten syntaktischen und semantischen Gegebenheiten berücksichtigt werden.

7.3. Verbreitung und Bekanntmachung

Zur Bekanntmachung des Projektes ist bereits mit dem Bundesverwaltungsamt eine Vorstellung in der online-Fachzeitschrift "WIN" vereinbart. Da das BVA eine eigene Broschüre zur sprachlichen Gleichstellung von Mann und Frau herausgegeben hat und zudem durch das Bundesgleichstellungsgesetz und die Gemeinsame Geschäftsordnung der Bundesministerien zur Kontrolle seiner Veröffentlichungen verpflichtet ist, stößt das Projekt auf großes Interesse. Geplant ist weiterhin eine Beantragung von Fördermitteln im Rahmen des 6. Europäischen Forschungsrahmenprogramms. Die zu diesem Zweck notwendige Bildung eines internationalen Konsortiums soll sich auf verschiedene Staaten der Europäischen Union (Frankreich, Spanien, Portugal, Großbritannien, Polen, Litauen) erstrecken und hat das Ziel, ein Tool zur geschlechtergerechten Formulierung für mehrere Sprachen anzubieten.

8. Bibliographie

- [Arb97] Arbeitsgemeinschaft der kommunalen Gleichstellungs- und Frauenbeauftragten Mittelfrankens, Nürnberg, 'Frauen & Sprache. Tipps zur Gleichbehandlung von Frauen und Männern in der Sprache', 1997.
- [Bun02] Bundesstelle für Büroorganisation und Bürotechnik (Hg.), Köln, 'Sprachliche Gleichbehandlung von Frauen und Männern. Hinweise, Anwendungsmöglichkeiten und Beispiele.', 2002.
- [CAR 98] CARL M., SCHMIDT-WIGGER A., 'Shallow Postmorphological Processing with KURD', *Proceedings of NeMLaP3/CoNLL98*, Sydney, 1998, p. 257-265.
- [Deu93] Deutsche UNESCO-Kommission (UNESCO) (Hg.), Bonn, 'Eine Sprache für beide Geschlechter. Richtlinien für einen nicht-sexistischen Sprachgebrauch', 1993.
- [Ges99] Gesellschaft für Informatik (Hg.), Bonn, 'Gleichbehandlung im Sprachgebrauch: Reden und Schreiben für Frauen und Männer', 1999.
- [GüN 80] GÜNTHERODT I., HELLINGER M., PUSCH L. F., TRÖMEL-PLÖTZ S., *Richtlinien zur Vermeidung sexistischen Sprachgebrauchs*, vol. 69, 1980.
- [HEL 93] HELLINGER M., BIERBACH C., 'Eine Sprache für beide Geschlechter. Richtlinien für einen nicht-sexistischen Sprachgebrauch', Bericht, 1993, Bonn.
- [HäB 92] HÄBERLIN S., SCHMID R., WYSS E. L., *Übung macht die Meisterin. Ratschläge für einen nicht-sexistischen Sprachgebrauch*, München, 1992.

- [PUS 84] PUSCH L. F., *Das Deutsche als Männersprache*, Frankfurt/M., 1984.
- [PUS 90] PUSCH L. F., *Alle Menschen werden Schwestern. Feministische Sprachkritik*, Frankfurt/M., 1990.
- [Sch96] Schweizerische Bundeskanzlei (Hg.), Bern, ‘Leitfaden zur sprachlichen Gleichbehandlung im Deutschen’, 1996.
- [TRö 82] TRÖMEL-PLÖTZ S., *Frauensprache: Sprache der Veränderung*, Frankfurt/M., 1982.
- [TRö 88] TRÖMEL-PLÖTZ S., ‘Sexismus in der Sprache —Feminin - Maskulin. Konventionen Kontroversen Korrespondenzen’, *Friedrich Jahresheft*, vol. VII, 1988, Erhard Friedrich.
- [Uni00] Universität Zürich (Hg.), Zürich, ‘Leitfaden zur sprachlichen Gleichbehandlung von Frau und Mann’, 2000.