

MULTILINT

Gefördert vom BMWi - Gesch.-Z.: II C 7-00 3048/5

Abschlußbericht

Ursula Reuther (IAI)
Antje Schmidt-Wigger (IAI)

Mit einem Beitrag von
Claudia Fottner-Top (itl)

Juli 1998

Inhaltsverzeichnis

1	Einleitung	1
2	Benutzerumgebung und -anforderungen	1
2.1	Dokumenterstellung bei BMW	1
2.2	Vorhandenes Datenmaterial	2
2.3	Terminologiebestand	2
2.4	MULTILINT aus BMW Benutzersicht	3
3	Oberfläche und Integration	4
3.1	Beschreibung der Benutzeroberfläche	4
3.2	Integration von MULTILINT bei BMW	4
4	Kontrollwerkzeuge	5
4.1	Basiskonzeption	5
4.1.1	Konzeptuelle Entwicklung	5
4.1.2	Technische Umsetzung	6
4.2	Grammatische Richtigkeit	7
4.2.1	Rechtschreibkontrolle	7
4.2.2	Grammatikkontrolle	7
4.3	Terminologische Konsistenz und Korrektheit	8
4.3.1	Abkürzungskontrolle	8
4.3.2	Konsistenzkontrolle	9
4.3.3	Terminologiekontrolle	9
4.4	Stilistische Angemessenheit	9
5	Weitere MULTILINT-Werkzeuge	10
5.1	Zusätzliche Aufgaben der Autoren	10
5.2	Übersetzungsunterstützende Werkzeuge	11
5.3	Suchfunktionen	12
5.4	Werkzeuge zur Textcharakterisierung	12
6	Evaluierung	12
6.1	Zyklisch-interaktive Evaluierung	12
6.2	Beispielbasierte, vergleichende Evaluierung	13
6.3	Einschätzung des Prototypen durch die Technischen Autoren	13
7	Publikationen und Präsentationen	14

1 Einleitung

Vorliegender Bericht faßt die Ergebnisse des Projekts MULTILINT zusammen und beschreibt den Prototypen, der im Rahmen des Projekts gemeinsam von den Projektpartnern BMW AG und IAI entwickelt, implementiert und in einem spezifischen Anwendungskontext bei BMW validiert wurde.

Zielsetzung des Projekts war es, linguistisch intelligente Werkzeuge zu entwickeln, die die Technischen Autoren dabei unterstützen, qualitativ hochwertige technische Dokumentation zu erstellen.

Diese Kontrollwerkzeuge wurden gemäß der Bedürfnisse des Autors entwickelt und stehen ihm konfigurierbar in einer Art Toolbox zur Verfügung. Durch ihren Einsatz kann der Autor bereits bei der Erstellung von Dokumenten eine Qualitätskontrolle vornehmen.

Für den gegenwärtigen Prototypen wurde der Schwerpunkt auf die Entwicklung der Quelltext-Kontrolltools gelegt, da die Dokumenterstellung in Deutsch bei BMW der Einsatz- und Validierungsbereich von MULTILINT war. Hinzu kommt, daß bei dem Einsatz der MULTILINT-Tools die Strategie verfolgt wird, die Werkzeuge so früh wie möglich in den Gesamtarbeitsablauf der Dokumenterstellung einzubinden, was zu einer Optimierung aller nachfolgenden Prozesse, inklusive des Übersetzungsprozesses, führt.

Neben diesen Quelltext-Kontrolltools, die Kontrollen zur grammatischen Korrektheit, zu terminologischer Konsistenz und Korrektheit sowie zu stilistischen Regeln der technischen Dokumentation anbieten, stehen den Autoren ebenfalls übersetzungsunterstützende Tools, wie z.B. Glossarerstellung oder Grobübersetzung, und andere Funktionalitäten, wie. z.B. Statistikfunktionen und Abstracting, zur Verfügung.

Der Prototyp, der nunmehr am Ende des Projekts verfügbar ist, ist bei BMW installiert und kann innerhalb des BMW eigenen Redaktionssystems aufgerufen werden. Der Autor kann somit die MULTILINT-Tools direkt auf das Dokument anwenden, das er gerade bearbeitet.

2 Benutzerumgebung und -anforderungen

Die Situation bezüglich technischer Dokumentation beim MULTILINT-Anwender BMW ist durch folgende Gegebenheiten charakterisiert, die von Projektbeginn an mit berücksichtigt wurden und dabei einen entscheidenden Einfluß auf die Prototypenentwicklung hatten.

2.1 Dokumenterstellung bei BMW

Das BMW Redaktionssystem (RS), das als Autoren- und Dokumentmanagement-System dient, ist das zentrale Element im Bereich der Dokumenterstellung und -verwaltung bei BMW. Es ermöglicht die Koordination der wesentlichen Geschäftsprozesse im Technischen Kundendienst und hat das Ziel, die volumenträchtigen Produkt- und Serviceinformationen gebündelt im funktionalen Zusammenhang über eine CD-ROM den BMW Handelsorganisationen zur Verfügung zu stellen. Bei den BMW Vertragshändlern sind derzeit ca. 3500 Systeme weltweit installiert. Von den Daten dieses Informationssystems werden monatliche updates erstellt, die dann auch in alle relevanten Sprachen übersetzt werden müssen. Die Übersetzung dieser Dokumente wird nicht von BMW selbst durchgeführt, sondern an eine Firma vergeben, die die Übersetzungen unter Einsatz eines eigenen Translation Memory Systems liefert.

Das Redaktionssystem hat unter anderem das Ziel, daß den Autoren nur noch standardisierte Prozesse zugänglich sind. Die Autoren arbeiten täglich damit und entwickeln ca. 600 Seiten

Dokumente pro Jahr und Autor.

In diese Situation fügt sich MULTILINT als eine erste Kontrollinstanz zur Qualitätssicherung im Dokumenterstellungsprozeß sehr gut ein.

2.2 Vorhandenes Datenmaterial

Um die linguistischen Tools optimal auf die Benutzeranforderungen abzustimmen, war es nötig, eine benutzerspezifische und repräsentative Datenmenge aufzubauen, die als Ausgangsbasis für die Entwicklung der Tools benutzt werden konnte.

Zu diesem Zweck wurde dem IAI von BMW als Entwicklungs- und Testkorpus in mehreren Etappen umfangreiches Textmaterial aus dem täglichen Arbeitsumfang der Technischen Autoren zur Verfügung gestellt. Es handelt sich fast ausnahmslos um im Redaktionssystem erstellte Texte, weshalb sie durchgehend in SGML-Format vorliegen. Der Aufbau der Texte ist entsprechend strukturiert und vereinheitlicht. Sie haben eine durchschnittliche Länge von einer halben bis zwei Seiten und umfassen vor allem die Dokumenttypen *Reparaturanleitung* (Dokumentation von Reparaturabläufen), *Service Information* (Ankündigungen von Neuerungen oder Erklärung eines Problems mit der Lösung) und *Hotline Information* (Dokumentation aktueller Probleme bei den Händlern).

Auf eigens für diesen Zweck entworfenen Formblättern wurde für jede Datenlieferung festgehalten, um was für ein Dokument es sich handelt, welchem Arbeits- bzw. Testbereich es zugeführt wurde und welche linguistischen Ressourcen damit weiter entwickelt wurden.

Auf der Basis dieses Korpus wurden die einzelnen linguistischen Komponenten und die zugrundeliegenden linguistischen Daten weiterentwickelt und den Erfordernissen der Anwendungsdomäne angepaßt.

Das überlassene Datenmaterial weist einen heterogenen Sprachstil auf, der neben Strukturen des Standard-Deutschen vielfach elliptische Elemente und Einschübe in allen Variationen enthält. Die Länge der zu verarbeitenden Einheiten variiert zwischen 1 und 75 Wörtern (Satzzeichen incl.), wobei Sätze mit mehr als 40 Wörtern nur sehr selten auftreten. Nur etwa ein Drittel aller Bearbeitungseinheiten enthält ein Verb, so daß die Mehrzahl aller Bearbeitungseinheiten als Strukturen unterhalb der Satzebene beschrieben werden müssen.

2.3 Terminologiebestand

Im Bereich Terminologie konnte das ursprünglich gesetzte Ziel, die quell- und zielsprachliche Terminologie aufzuarbeiten und sie in die ein- und zweisprachigen Lexika zu integrieren, nur bis zu einem gewissen Grad erreicht werden, da die Situation bei BMW bezüglich gültiger Terminologie oftmals nicht geklärt werden konnte.

Eine verbindliche einsprachige Terminologie war nicht verfügbar, deshalb wurden zunächst nur zwei- und mehrsprachige Terminologiesammlungen in die Lexika integriert:

Zu Beginn des Projekts wurden seitens BMW zwei zweisprachige Wortlisten geliefert, die zum einen (mit 4000 Wortpaaren) aus der Übersetzerabteilung bei BMW, zum anderen (mit 7500 Wortpaaren) aus einem älteren Normierungsversuch bei BMW stammen. Später kamen mehrere Versionen der zweisprachigen TERMI-Datenbank und Auszüge aus einer achtsprachigen Teilleiste hinzu. Alle Termini mit ihren Übersetzungsäquivalenten sind in die zweisprachigen Wörterbücher und somit in die übersetzungsunterstützenden Tools (z.B. Erstellung eines zweisprachigen Glossars) eingegangen.

Diese zwei- oder mehrsprachigen Terminologien mußten auch als Basis für die einsprachige

Terminologiekontrolle benutzt werden, die laut Aufgabenbeschreibung sowohl die konsistente Verwendung von Termini als auch ihre BMW Adäquatheit überprüfen soll.

Während der Entwicklung der Terminologie-Kontrolltools wurden verschiedene der Datensammlungen als potentielle Kandidaten für eine MULTILINT-Terminologie identifiziert (Befundnummernkatalog, Teileliste, Symptombaum, TERMI-Datenbank), wobei nun, teils aus pragmatischen teils aus inhaltlichen Gründen, die aktuelle MULTILINT-Terminologie aus allen deutschen Einworttermini der TERMI-Datenbank besteht, die bereits einer gewissen Vereinheitlichung unterzogen wurden.

Zudem wurde dem IAI eine offizielle Liste BMW spezifischer Abkürzungen mit etwas über 1000 Akronymen übermittelt, zu der alle Abkürzungen aus dem ganzen Textkorpus hinzugenommen wurden, die laut BMW zulässige Abkürzungen sind. Auf dieser Gesamtliste basiert die MULTILINT-Abkürzungskontrolle.

2.4 MULTILINT aus BMW Benutzersicht

In folgenden Einsatzbereichen kann der Einsatz von MULTILINT entscheidende Qualitätsverbesserungen und Kostenvorteile bringen:

- Unterstützung der Autoren (“Autorenlösung”):
Die Autoren überprüfen ihren Text mit MULTILINT, bevor sie ihr Dokument an das Redaktionssystem übergeben. Wenn die Autoren die Korrekturanweisungen von MULTILINT umsetzen, ist sichergestellt, daß nur Dokumente mit einem guten sprachlichen Qualitätsstandard in Umlauf kommen. Gleichzeitig haben die Autoren den Vorteil, daß sie bei der Erstellung ihrer Dokumente ihr Hauptaugenmerk auf den Inhalt, d.h. auf die technische Information, legen können, weil sie in sprachlichen Fragen unterstützt werden.
- Basissystem für das Lektorat (“Lektoratslösung”):
Ein Lektor identifiziert mit Hilfe von MULTILINT sprachliche Qualitätsprobleme und beseitigt sie, bevor das Dokument übersetzt wird oder auf andere Weise in Umlauf kommt.
Mit MULTILINT können maschinell innerhalb kürzester Zeit diejenigen Stellen herausgefiltert werden, die redaktionell überarbeitet werden müssen. Damit greift MULTILINT bereits auf die Wurzel von möglichen Verständlichkeitsproblemen zu.

Sowohl die Autorenlösung als auch die Lektoratslösung haben entschiedene Vorteile. Deshalb sind die beiden Lösungen als jeweilige Ergänzung zu sehen und nicht als Alternativen.

Zum gegenwärtigen Zeitpunkt gibt es im Ansatz bereits die Autorenlösung: MULTILINT ist in das Redaktionssystem integriert und kann über einen Knopf im FrameBuilder-Menü ‘Spezial’ aufgerufen werden.

Aufgrund technischer Gegebenheiten ist diese Integration jedoch erst als eine “Light-Integration” anzusehen. Denn um Änderungen vorzunehmen, müssen die Benutzer zum einen im Frame-Dokument die zu ändernden Textstellen selbst identifizieren, z.B. über die Frame-Suche, und zum anderen zwischen den MULTILINT-Ergebnisfenstern und dem Frame-Dokument hin- und hernavigieren.

Innerhalb einer echten Integration könnten die Benutzer interaktiv verbessern, d.h. bereits im MULTILINT-Ergebnisfenster die zu ändernden Textstellen editieren. Solch eine echte Integration, die die Akzeptanz von MULTILINT noch weiter erhöhen würde, ist derzeit aus technischen, BMW internen Gründen nicht möglich.

Die Lektoratslösung gibt es zum gegenwärtigen Zeitpunkt ebenfalls noch nicht.

Folgende **begleitende Maßnahmen** auf der Seite des Industriepartners sind für den Einsatz von MULTILINT notwendig:

- Schulungen und Betreuung der Autoren:
Verständliches Schreiben, Einführung in das Arbeiten mit MULTILINT sowie Betreuung bei Training on the job (nach und nach für alle Autoren)
- Linguistische Betreuung von MULTILINT:
Problemanalysen, Spezifikation der linguistischen und redaktionellen Anforderungen, Systematisierung von Fehlern und Benutzeranforderungen
- “Lokalisierung” von MULTILINT: Didaktische Reduzierung der linguistischen Informationen und Anpassung an die Zielgruppe (technisch orientierte Autoren)
- Terminologische Betreuung von MULTILINT:
Pflege des firmeneigenen Wörterbuchs und der Liste der Abkürzungen, Pflege von Rechtschreibregeln und Kriterien für Verständlichkeit

3 Oberfläche und Integration

3.1 Beschreibung der Benutzeroberfläche

Um die verschiedenen Komponenten von MULTILINT zu steuern und dem Benutzer den Systemzugriff zu erlauben, wurde eine graphische Benutzeroberfläche entwickelt. Ein Menü dient der Auswahl (Aktivierung) der linguistischen Funktionen und der benutzerspezifischen Parametrisierung des Systems (Benutzerprofil). Die Ergebnisse sind benutzerfreundlich dargestellt, und die Navigation von einem Fehler zum anderen wird unterstützt. Da keine vollständige Integration in den BMW Editor möglich war, wurden separate Anzeigefenster implementiert, um die Fehler dem Benutzer im Kontext anzeigen zu können.

Die Systemdokumentation und die Benutzerhilfe sind im HTML-Format verfügbar.

3.2 Integration von MULTILINT bei BMW

Nach etwa einjähriger Vorlaufzeit konnte die erste Stand-alone Version des Prototypen von MULTILINT bei BMW installiert werden. Die Installierung erfolgte auf einer SUN-Workstation unter dem Betriebssystem Solaris 2.5. Im durchschnittlichen Rhythmus von drei Monaten wurden insgesamt fünf Folgeversionen nachgeliefert, die stetigen Tests seitens der Benutzer unterzogen wurden.

Aufgrund der Anforderungen der Pilotanwender wurden im folgenden die ersten Diskussionen mit BMW und Hewlett Packard (HP, Hersteller des RS) zur Integration von MULTILINT in die BMW RS-Umgebung geführt. Es wurde beschlossen, daß aus firmenpolitischen Gründen im Rahmen dieses Förderprojekts keine komplette Integration in das RS durchgeführt wird, sondern eine sogenannte “leichte” Integration, auch “Autorenlösung” genannt. Hierbei steht dem Technischen Autor ein Menüpunkt in seinem RS-Texteditor zur Verfügung. Bei Knopfdruck wird die aktuelle Datei an MULTILINT geschickt. In weiteren o.g. Anzeigefenstern werden dem Autor die Fehler, die das System gefunden hat, angezeigt.

Dieser Ansatz ermöglicht weder eine automatische Korrekturübernahme, wie für manche der Funktionalitäten ursprünglich vorgesehen, noch eine Markierung der Fehler im Originaldokument. Dies kann nur durch eine “volle” Integration von MULTILINT in die RS-Umgebung erzielt werden

Um den Anforderungen zu genügen, die generell für bei BMW eingesetzte Software gelten, mußte MULTILINT den BMW Softwarequalitätsstandards angepaßt werden. Dazu wurde von BMW bereits eine Quellcode-Kontrolle mit dem statischen Analysewerkzeug Logiscope C der Firma Verilog durchgeführt, die für einen Prototypen ganz zufriedenstellende Ergebnisse gezeigt hat.

4 Kontrollwerkzeuge

Von seiten BMW wurde sehr früh innerhalb der Projektlaufzeit der Schwerpunkt des Interesses auf das Potential von MULTILINT gelegt, die sprachlichen Aspekte in der technischen Dokumentation zu kontrollieren und zu verbessern. Die Schwachpunkte, die durch die BMW Kundendienstabteilung identifiziert wurden, betreffen folgende Bereiche:

- Wortwahl
- Satzstruktur
- Grammatik
- Rechtschreibung
- Interpunktion
- Terminologie
- Herstellerspezifische Begriffe
- Abkürzungen

Somit war der Schwerpunkt innerhalb des Projekts auf die Entwicklung von Werkzeugen gesetzt, die o.g. Phänomene analysieren und kontrollieren können.

4.1 Basiskonzeption

Alle linguistischen Werkzeuge basieren auf den gleichen linguistischen Ressourcen. Um größtmögliche Flexibilität und Kohärenz der einzelnen Kontrollwerkzeuge zu gewährleisten, mußte ihre Entwicklung einerseits so modular wie möglich gehalten werden, andererseits sollten jedoch Überschneidungen und Verdoppelungen von Verarbeitungsschritten weitestgehend vermieden werden.

4.1.1 Konzeptuelle Entwicklung

Zunächst wurde zwischen allgemeinsprachlicher Problematik und fachgebiets- bzw. textsortenabhängigen Fragestellungen unterschieden. Hier eine deutliche Trennung zu gewährleisten, erlaubt in der Folge, Teile der entwickelten Werkzeuge ohne Modifikation in anderen Bereichen einzusetzen. So wurden Fragen der Grammatik, Rechtschreibung und Interpunktion deutlich getrennt von terminologischen oder stilistischen Fragen behandelt. Terminologie, herstellerspezifische Begriffe und Abkürzungen sind zumindest fachspezifisch und können in der Praxis meist nur unter Einbezug der jeweiligen Herstellervorgaben behandelt werden. Stilistische Fragen, u.a. bezüglich der Satzstrukturen, sind weniger hersteller- oder fachgebiets- als textsortenabhängig. Dennoch wurde bei der Durchsicht der gängigen Literatur deutlich, daß allgemeingültige

Schreibregeln nicht definiert sind und herstellersistem-spezifische Präferenzen daher beim jetzigen Stand der wissenschaftlichen Diskussion nicht außer acht gelassen werden können.

Eine andere Modularisierungsebene der Kontrollwerkzeuge ist in der Komplexität der Verarbeitung und der daraus resultierenden Abhängigkeit eines Werkzeugs von den Ergebnissen eines anderen begründet. So werden Probleme in der Rechtschreibung und bezüglich terminologischer Einheiten in der Regel am einzelnen Wort festgemacht, wohingegen grammatikalische Fehler und stilistisch ungewünschte Satzstrukturen die Analyse mehrerer Wörter und deren Beziehung zueinander nötig machen.

Diese Einteilung ist sicherlich als Vereinfachung zu verstehen, aber um eine optimale Bearbeitungszeit zu erreichen, sollte für jedes Problem nur die minimal nötige Komplexität herangezogen werden.

Als Konsequenz dieser Modularisierung sollte für die komplexen Module angenommen werden, daß Probleme der wortbezogenen Module nicht mehr behandelt werden müssen. So sollte die Grammatikkontrolle so implementiert sein, daß sie keine Grammatikfehler an falsch geschriebenen Worten erkennen braucht. Ebenso sollte die Stilkontrolle davon ausgehen können, daß ihre Eingabe sprachlich richtig ist. Der Anwender muß die einzelnen Werkzeuge also entweder nacheinander starten oder mehrere komplette Kontrolldurchgänge aktivieren. Dies ist auch insoweit nötig, als daß keine automatische Kontrolle der manuellen Verbesserungen erfolgt.

4.1.2 Technische Umsetzung

Die beschriebene wünschenswerte Modularisierung der Kontrollwerkzeuge wurde in MULTILINT umgesetzt. So stehen dem Benutzer in der Endversion des Prototypen sechs verschiedene Quelltext-Kontrollfunktionen zur Verfügung, die er je nach vorhandenen Daten, Textsorte oder persönlichem Arbeitsstil automatisch alle zusammen ablaufen lassen oder sukzessive, frei kombiniert anwenden kann. Ein zweiter Kontrolldurchlauf verlangt dabei jedoch einen Neustart von MULTILINT auf dem verbesserten Quelltext. Eine Staffelung der Anwendung ist aufgrund der Teilintegration nicht möglich. Damit hieraus kein Nachteil entsteht, wurden die einzelnen Werkzeuge in ihrem Anwendungsbereich teilweise robuster gemacht; so ist die Stilkontrolle z.B. unempfindlich gegenüber einem Großteil der auftretenden Rechtschreib- und Grammatikfehler.

Für die einzelnen Kontrollen wurden jeweils entsprechende Programmteile entwickelt, die auf den Bestand der linguistischen Ressourcen zugreifen. Dabei sind Eingabe- und Ausgabeinformationen vollständig definiert. So beruhen alle Werkzeuge auf der morphologischen Analyse der Einzelwörter, die jeweils eindeutig einem segmentierten Satzobjekt zugeordnet sind. Die Rechtschreibkontrolle und die terminologischen Kontrollen bearbeiten wiederum nur diese Wortobjekte. Grammatik- und Stilkontrolle behandeln wortübergreifende Phänomene mittels eines formalismus-basierten Programmes, wobei im Vorfeld auch eine wortübergreifende Disambiguierung eingesetzt wird.

Die Weitergabe der identifizierten sprachlichen Probleme erfolgt mittels Attribut-Wert-Paaren, die MULTILINT durchgängig als interne Darstellungsform verwendet. Dabei kann pro Wort momentan nur ein Fehlertyp pro Kontrollwerkzeug übermittelt werden, was jedoch aufgrund der starken Modularisierung keinen spürbaren Nachteil für den Benutzer darstellt.

Jeder Fehlertyp ist mit einem spezifischen Kommentar gekoppelt, der dem Benutzer auf Anfrage angezeigt wird. Als weitere Hilfe kann der Benutzer zwischen verschiedenen Meldungsarten auswählen oder sich ein erklärendes Beispiel anzeigen lassen. Bei der Rechtschreib- und Terminologiekontrolle wird zusätzlich ein Verbesserungsvorschlag gegeben.

4.2 Grammatische Richtigkeit

Zu der Gruppe der Kontrollwerkzeuge, die die grammatische Korrektheit des Textes überprüfen soll, gehören zum einen eine morphologisch basierte Rechtschreibkontrolle und zum anderen eine Grammatikkontrolle, die im folgenden näher beschrieben sind.

4.2.1 Rechtschreibkontrolle

Die Rechtschreibkontrolle ist das direkte Ergebnis der Analysekomponente von MULTILINT. Sie basiert auf einer morphologischen Analyse, die sowohl Kompositions-, Derivations- und Inflektionsinformation berücksichtigt und somit gängigen Rechtschreibprüfprogrammen in dieser Hinsicht überlegen ist.

Zusätzlich zur Markierung des fehlerhaften Worts werden dem Autor Verbesserungsvorschläge angeboten. Die Verbesserungsvorschläge basieren auf einer Liste von ca. 160.000 deutschen Wortformen, die noch um ca. 22.000 BMW spezifische Wörter erweitert wurde. Zur Identifikation möglicher ähnlicher Wörter wird das FindLink Tool der Firma CONNEX eingesetzt.

Tests zeigten gute Ergebnisse: So wurden 77 Rechtschreibfehler getestet, die willkürlich aus den in den BMW Korpora automatisch entdeckten Fehlern ausgewählt waren. FindLink schlug bei ca. 75% der falsch geschriebenen Wörter die richtige Variante als Verbesserung vor, wobei der Vorschlag zu 80% an erster, sonst an zweiter Stelle erschien. Die Vorschlagsliste wurde dementsprechend auf zwei Vorschläge beschränkt.

Über die übliche Kontrollfunktion eines Spell Checkers hinaus werden Fehler der Groß- und Kleinschreibung teilweise bereits innerhalb der Rechtschreibkontrolle abgefangen. So wird die Wortstellung im Satz miteinbezogen und Tippfehler, bei denen innerhalb eines Wortes großgeschrieben wird. Außerdem wird innerhalb dieser Komponente auf Besonderheiten der Groß- und Kleinschreibung und der Getrenntschreibung geachtet, die der Corporate Identity von BMW entsprechen.

4.2.2 Grammatikkontrolle

Die Arbeiten an der Grammatikkontrolle wurden relativ unabhängig von BMW Vorgaben am IAI durchgeführt, da die Ziele einer Grammatikkontrolle durch die gängigen Rechtschreib- und Grammatikregeln des Deutschen definiert sind. Einzig die Auswahl von typischen Fehlern, wie sie in der technischen Dokumentation von BMW vorkommen, mußte im Vorfeld geschehen. Hierfür wurde das gesamte BMW Korpus manuell abgearbeitet und die darin enthaltenen grammatischen Fehler festgehalten. Interessant zu erwähnen ist in diesem Zusammenhang, daß Interpunktionsfehler ungefähr die Hälfte aller Fehler ausmachen; die restlichen Fehler verteilen sich mehr oder weniger gleichmäßig auf die restlichen Bereiche.

Auf der Basis dieser Fehlerklassifikation wurden dann in einem nächsten Schritt die Fehler und die dazugehörigen Fehlermeldungen formal linguistisch in den entsprechenden Analysemodulen beschrieben und integriert.

Zuerst wurden zwei unterschiedliche Versionen der Grammatikkontrolle, die auf verschiedenen Formalismen beruhen und demnach unterschiedlich konzipiert sind, integriert, um sie auf ihre Effizienz und Brauchbarkeit innerhalb der Anwendungsdomäne zu testen.

Im Laufe der Test- und Erprobungsphase sowohl am IAI als auch bei BMW wurde eine Entscheidung zugunsten des flachen Parseansatzes getroffen, der innerhalb eines eigens dafür entwickelten Formalismus (KURD) implementiert wurde.

Spezifisch für die Textsorten des Korpus und die vorkommenden Fehler sind im KURD-Formalismus 234 Regeln für 55 Fehlerklassen implementiert worden. Diese Regeln decken 181 Fehler der 378 initial gesammelten Fehler ab. Zusätzlich wurden im Testkorpus 87 Fehler mit dem Regelsatz automatisch identifiziert, die der manuellen Korrektur entgangen waren. Legt man diese erweiterte Fehlersammlung zugrunde, kommt man demnach für das Testkorpus auf einen Recall von 0,57, was das Verhältnis zwischen den tatsächlich gefundenen Fehlern und den im Text vorhandenen Fehlern ausdrückt. Neben den richtig erkannten Fehlern hat die automatische Bearbeitung des Testkorpus natürlich auch zu Fällen von 'falschem Alarm' geführt. Momentan betrifft dies 62 Sätze des Testkorpus. Hiermit kann von einer Präzision von 0,81 gesprochen werden, was das Verhältnis der falschen Alarme zu allen gefundenen Fehlern ausdrückt.

4.3 Terminologische Konsistenz und Korrektheit

Diese Gruppe der Kontrollwerkzeuge, die sich im weitesten Sinne mit terminologischen Defiziten auseinandersetzt, umfaßt neben einer Abkürzungskontrolle eine Kontrollfunktion zur konsistenten Verwendung von Begriffen und eine weitere Überprüfung der Termini in bezug auf ihre Zulässigkeit bei BMW.

4.3.1 Abkürzungskontrolle

Mittels der Abkürzungskontrolle kann überprüft werden, ob die im Text vorkommenden Abkürzungen und Akronyme den BMW üblichen Abkürzungsstandards entsprechen und, wenn diese Voraussetzung erfüllt ist, welche Bedeutung die entsprechende Abkürzung hat. Neben allgemeinen Abkürzungen, wie z.B. *km*, *z.B.*, *BMW*, werden auch benutzerspezifische Abkürzungen berücksichtigt.

Ein Problem bei der Integration dieses Kontrolltools war, daß bei unbekanntem Abkürzungen die Satzsegmentierung fehlschlug und der Satzzusammenhang nicht mehr rekonstruiert werden konnte, nur bei bekannten Abkürzungen die Satzsegmentierung nach dem Punkt unterdrückt wird.

Um diesem Problem vorzubeugen, wurde unter Rücksprache mit BMW eine Abkürzungsliste erstellt. Sie enthält die offizielle Abkürzungsliste sowie alle Abkürzungen der überlassenen Texte, zu denen BMW die Zustimmung gegeben hat. An dieser Liste wurden in der Folge umfangreiche Erweiterungs- und Abgleicharbeiten durchgeführt.

Der Benutzer hat nun die Möglichkeit, bei Abkürzungen aus dieser Liste um deren Erläuterung zu bitten. Dies ist hilfreich, falls der Benutzer in der Wahl der Abkürzung unsicher ist.

Nicht in dieser Liste enthaltene Abkürzungen werden dem Benutzer im Text kenntlich gemacht. Es wird dabei unterschieden zwischen Akronymen, die der Benutzer bei neu eingeführten Konzepten in begrenztem Rahmen benutzen darf, und Ad-Hoc-Abkürzungen, die gänzlich verboten sind.

Zusätzlich eingeführt wurde eine Sonderbehandlung für nicht mehr gängige Maßeinheitsbezeichnungen (nach DIN-Norm 1301, Teil 3).

4.3.2 Konsistenzkontrolle

Die Konsistenzkontrolle wurde als Zusatz zur listenbasierten Terminologiekontrolle neu entwickelt und arbeitet ohne terminologische Information auf dem vorliegenden Text als Einheit. Sie basiert auf der Idee, daß ähnliche Wörter möglicherweise auf ein und dasselbe Konzept referieren und der Technische Autor angehalten werden sollte, sich in diesem Fall für eine Variante zu entscheiden. Die Ähnlichkeitsfunktion basiert hier auf der linguistischen Analyse, faßt also morphologisch verwandte Wörter zusammen. Zusätzlich werden die semantischen Eigenschaften der Wörter hinzugezogen, um Synonymität bei einigen Kombinationen auszuschließen.

Die Ähnlichkeitsfunktion wurde hinsichtlich zusätzlicher Abstraktionsmöglichkeiten weiter untersucht. So wurde einerseits als Randphänomen der Derivation die Vorsilbe unterdrückt, andererseits wurden Versuche unternommen, Synonymbeziehungen aufgrund von Kompositareduktion zu untermauern. Die Ergebnisse waren jedoch nicht sehr vielversprechend und auf eine Erweiterung wurde verzichtet.

Die Benutzer zeigen sich mit dem implementierten Ansatz zur dokumentbasierten Konsistenzkontrolle sehr zufrieden. Er entspricht den redaktionellen Bedürfnissen und ist wenig fehleranfällig.

4.3.3 Terminologiekontrolle

Die Problematik der Terminologiekontrolle wurde während der gesamten Projektlaufzeit mehrfach diskutiert und mehrere Lösungsversuche wurden im Ansatz implementiert.

In der ersten Version des Prototypen war als rudimentäre Version ein reiner Abgleich mit der BMW Terminologie implementiert, die die Allgemeinsprache gänzlich außer acht gelassen hat. Dieser Ansatz hat verständlicherweise wenig Anklang bei den Benutzern gefunden.

U.a. auf Anregung der Benutzer wurden andere Wissensquellen auf ihre Eignung zur Terminologiekontrolle hin untersucht: die mehrsprachigen BMW Wörterbücher, ein allgemeines Lexikon zur Automobilterminologie und die Hauptgruppennummern, die sowohl in den mehrsprachigen Wörterbüchern als auch an den Texten selbst spezifiziert sind. Hier konnten die Ergebnisse jedoch nicht als ausreichend angesehen werden, um sie dem Benutzer in einer Kontrollfunktion anbieten zu können.

In der endgültigen Version basiert die Kontrolle auf dem Ansatz zur Termkandidaten-Definition. Der Ansatz ist unserer Erachtens vielversprechend, da er anwenderunabhängig linguistisches Wissen verarbeiten und umsetzen kann. Er liegt auch bereits anderen Realisierungen im Bereich der Terminologiekontrolle und der automatischen Termextraktion zugrunde (wie z.B. bei Mercedes Benz in Zusammenarbeit mit der Universität Stuttgart). Eine Verbesserung wurde erreicht durch die integrierte Kombination mit der Konsistenz-Kontrollstrategie. So werden Komposita, die ähnlich einem Term sind, jedoch nicht gleich, dem Benutzer kenntlich gemacht und der BMW Term als Verbesserungsvorschlag mitgeliefert.

4.4 Stilistische Angemessenheit

In einem von BMW durchgeführten eintägigen Seminar wurden Ansätze zur Qualitätskontrolle technischer Texte vorgestellt. Zentrale Kriterien für Textqualität sind demnach Vollständigkeit, Konsistenz und Lesbarkeit der Texte. Im Rahmen dieser Aufgabenstellung wurde zunächst versucht, das Kriterium der **Lesbarkeit** genauer zu spezifizieren. Die Ergebnisse fassen zusammen, welche Eigenschaften von Sätzen deren Lesbarkeit einschränken:

- Zu lange Informationseinheiten lassen sich nicht aufnehmen.
- Zentrale Elemente dürfen nicht zu spät erwähnt werden.
- Ambige Strukturen können das Verstehen beeinträchtigen oder zu Mißverständnissen führen.
- Schwerfälliger Stil kann die Aufnahmefähigkeit einschränken.
- Anweisungen sollten so deutlich, stringent und zwingend wie möglich sein.

Zu jeder dieser Eigenschaften wurden gemeinsam mit BMW und durch Literaturrecherche konkrete Stilregeln definiert. Dabei wurden die verschiedenen Textsorten berücksichtigt, die das MULTILINT-Korpus ausmachen: Reparaturanleitung als instruktive Texte und Service Informationen als informative Texte. Während eines Kontrollaufs werden textsortenspezifische Regeln nur aktiviert, wenn für den aktuellen Text die jeweilige Textsorte erkannt wird.

Da an jeder Textstelle nur jeweils ein Fehler angezeigt werden soll, wurden die definierten Fehler von seiten BMW auf ihre Wichtigkeit hin überprüft und in drei Prioritätsklassen aufgeteilt. Kriterien waren dabei neben der Stärke des Verstoßes auch die Größe des markierten Bereichs und die Spezifität des Verbesserungsvorschlags.

Die Stilregeln wurden formalisiert und in einem eigens für MULTILINT entwickelten Formalismus implementiert. Die Regeln wurden an einem ausgewählten Teilkorpus abgesichert, bis die fehlerhafte Erkennung bei unter 10% lag. Die Sätze des Teilkorpus waren nicht nur aufgrund ihrer Repräsentativität, sondern auch aufgrund enthaltener komplexer Probleme ausgewählt worden, so daß bei freiem Input sogar von einem besseren Ergebnis ausgegangen werden kann.

Dem Benutzer der Stilkontrolle werden die stilistisch schwachen Textstellen farblich unterlegt im Text präsentiert. Auf Knopfdruck erscheint die entsprechende Stilregel im Wortlaut und kann auf Wunsch noch mit einem Umformungsbeispiel erläutert werden. Es stehen drei verschiedene Formulierungsvarianten für die Stilregeln zur Verfügung, so daß via Parameter zwischen *'für Technische Redakteure'*, *'knapp'* und *'linguistisch'* gewählt werden kann.

5 Weitere MULTILINT-Werkzeuge

5.1 Zusätzliche Aufgaben der Autoren

Bevor ein Technischer Autor bei BMW seinen Text verfaßt, muß er alle verfügbaren Wissensquellen ausschöpfen, um die zu beschreibenden Fahrzeugteile, Problemsituationen oder Reparaturhandgriffe vollständig zu erfassen. Neben der wichtigsten Informationsquelle, der hauseigenen Werkstatt, steht ihm natürlich das gesamte, bereits erstellte Textmaterial des Redaktionssystems zur Verfügung. Dasselbe Textmaterial wird er auch einsetzen, um Textbausteine bei der Erstellung einer überarbeiteten Version oder einer verwandten Reparaturbeschreibung wiederzuverwenden, wodurch Zeitersparnis sowohl bei der Erstellung als auch bei der Übersetzung der Texte (insbesondere beim Einsatz eines Translation Memory) erreicht wird. Die Hauptaufgabe dabei ist also die Suche nach passenden Problembeschreibungen. Momentan ist der Bearbeiter bei der Suche u.a. auf den Symptombaum und die angebotenen Parameter angewiesen, eine einfache Volltext-Suche ist nicht möglich. Eine linguistisch intelligente Suche könnte ihn dabei ebenso unterstützen wie die Möglichkeit, den Inhalt ausgewählter Texte in Kurzform wiederzugeben.

In einigen Arbeitsumgebungen stellt sich den Autoren das Suchproblem zusätzlich zweisprachig: um eine schnelle Hilfe von der Zentrale zu erhalten, formulieren einige nicht deutschsprachige

Vertriebsgesellschaften ihre Anfragen auf Englisch. Bei der Bearbeitung dieser Anfragen könnte eine mehrsprachige Suche oder eine Online-Übersetzung hilfreich sein.

Sicherlich hilfreich wären Übersetzungsunterstützende Tools auch für diejenigen Autoren, die von Zeit zu Zeit Übersetzungen selbständig anfertigen müssen. Das Bemühen der Übersetzungsabteilung oder der Sekretariate könnte so in einigen Fällen reduziert werden.

5.2 Übersetzungsunterstützende Werkzeuge

Zur Übersetzungsunterstützung stehen dem Autor bei BMW momentan zwei verschiedene Werkzeuge zur Verfügung: die Erstellung eines textbezogenen fremdsprachlichen Glossars und die sogenannte "Grobübersetzung", die den vollständigen Text übersetzt. In der Stand-alone Version von MULTILINT, die zu Beginn der Projektlaufzeit bei BMW installiert wurde (siehe Kapitel 3.2), existieren zusätzlich zwei Varianten dieser Werkzeuge, bei denen der Benutzer direkten Zugriff auf den MULTILINT-Editor haben muß: die Wörterbuchsuche und die interaktive Übersetzung.

Alle Werkzeuge greifen auf bestehende bilinguale Lexika zurück, in die auch die BMW spezifischen Terme der TERMI-Datenbank übernommen wurden. Die BMW spezifischen Übersetzungen werden bevorzugt, da das Glossar-Erstellungstool, das auch die Basis für die Übersetzung liefert, fachgebiets-sensitiv arbeitet.

Vom Umfang her zeigen sich die bidirektionalen bilingualen Lexika recht unterschiedlich: das Sprachpaar Deutsch-Englisch ist mit Abstand das am meisten entwickelte mit fast 450.000 Einträgen. Danach folgen Deutsch-Spanisch und Deutsch-Französisch mit jeweils etwa 25.000 Einträgen. Bei diesen drei Sprachpaaren haben BMW spezifische Einträge einen Anteil von mehr als 5%. Englisch-Französisch, Deutsch-Bulgarisch, Deutsch-Russisch und Englisch-Spanisch bieten auch einen großen Wortschatz (ca. 15.000, 9.500, 9.000 bzw. 6.000 Einträge), wobei hier vom Anwender kein Vokabular eingespielt wurde. Deutsch-Italienisch, Deutsch-Niederländisch und Deutsch-Schwedisch wurden mit 1.000 - 2.000 Einträgen vollständig aus BMW Ressourcen übernommen. Deutsch-Griechisch und Deutsch-Portugiesisch hingegen, mit jeweils über 2.500 Einträgen, sind Eigenentwicklungen. Für alle anderen Sprachpaare existieren keine speziell entwickelte zweisprachige Wörterbücher; hierfür muß Deutsch (oder Englisch) als Zwischensprache fungieren.

Zweisprachiges Glossar

Das zweisprachige Glossar soll dem Benutzer in erster Linie terminologische Hilfe anbieten. Hierzu werden für jedes Inhaltswort des Textes ein oder mehrere Übersetzungen ausgegeben. Enthalten die zweisprachigen Wörterbücher keinen Eintrag für ein Wort, wird im Falle eines Kompositums ein Übersetzungsversuch unternommen, indem die Einzelteile des Kompositums übersetzt werden. Die Glossarzeile ist dann mit (+) als 'mit Vorsicht zu betrachten' markiert. Für unübersetzbare Simplizita wird keine Glossarzeile erstellt.

Dieselbe Funktionalität kann der Benutzer der Stand-alone Version auch als Wörterbuchsuche auf einen im Text markierten Bereich anwenden.

Übersetzung

Auf der Basis der bilingualen Lexika wurde auch ein Programm zur Volltext-Übersetzung entwickelt. Mittels einiger sprachpaarbezogener Transformationsregeln wird aus einem Satz in der Ausgangssprache eine Übersetzung in die Zielsprache erstellt, die nicht den Anspruch erhebt, eine komplette, fehlerfreie Übersetzung zu sein, sondern die zu informativen Zwecken gelesen werden kann. Diese Funktion existiert jedoch nur für die Sprachen Deutsch, Französisch und Englisch.

Mittels der interaktiven Übersetzungsfunktion kann der Benutzer der Stand-alone Version diese Grobübersetzung auch direkt für Einzelfragmente aktivieren.

5.3 Suchfunktionen

In der Stand-alone Version von MULTILINT ist exemplarisch eine Suchfunktion integriert, die die Möglichkeiten eines linguistisch intelligenten Retrieval-Werkzeugs vorstellt. Die reine Volltext-Suche kann hier durch linguistische Mittel verfeinert werden und das gesuchte Konzept in den verschiedenen sprachlichen Realisierungen (flektiert, in Kompositaform, als Nomen- oder Verbderivat) aufgefunden werden.

Für einen Benutzer mit passiven Fremdsprachenkenntnissen ist die mehrsprachige Suche von Vorteil, in der er die Suchanfrage in seiner Muttersprache formulieren kann. Diese Art von Suche erlaubt es, Sätze zu finden, in denen die Übersetzungen der Suchwörter enthalten sind, wobei ebenfalls alle sprachlichen Realisierungen berücksichtigt werden.

5.4 Werkzeuge zur Textcharakterisierung

Dokumente können durch diese zusätzlichen Werkzeuge sowohl qualitativ als auch inhaltlich eingestuft werden.

Zur qualitativen Einstufung bietet MULTILINT die Berechnung einer Fehlerstatistik an. Es wurden bisher nicht alle Fehlertypen in die Statistik miteinbezogen, da die Akzeptanz der Autoren ein solches Bewertungswerkzeug betreffend sehr gering war. Die Projektpartner gehen jedoch davon aus, daß die wachsende Vertrautheit der Autoren mit MULTILINT eine Ausweitung dieser Funktion möglich machen wird.

Zur inhaltlichen Einstufung bietet MULTILINT zum einen die Textstatistik-Funktion, die die häufigsten Inhaltswörter des Textes auflistet, zum anderen die "Abstract"-Funktion. Hier wird versucht, auf der Basis der semantischen Kodierung der Inhaltswörter eine Schlüsselwortliste und eine davon abgeleitete Zusammenfassung des Textes zu berechnen.

6 Evaluierung

6.1 Zyklisch-interaktive Evaluierung

In der Entwicklungsphase wurden die Kontrollfunktionen von den Entwicklern anhand aller von BMW überlassenen Dokumente regelmäßig auf die Richtigkeit ihrer Anwendung hin überprüft; je nach Dokumentart ein oder mehrere Male.

Danach wurde das System nach jeder Versionslieferung von zunächst zwei Pilotanwendern bei BMW getestet. Im Laufe der Systementwicklung wurde der Kreis der Anwender stetig erweitert. Nach jeder Lieferung übermittelten die Anwender detaillierte und systematische Test- und Validierungsergebnisse, deren Auswertung wiederum in die nächste Version einfließen. Durch diese zyklische, interaktive Evaluierungsmethode konnte die Benutzeroberfläche aufgrund der Rückmeldungen der Pilotanwender erheblich verbessert und so die Benutzerfreundlichkeit erhöht werden. Diese Rückmeldungen betrafen aber auch Programmierfehler sowie die Ergebnisqualität und die Performanz der linguistischen Komponenten.

6.2 Beispielbasierte, vergleichende Evaluierung

Ab der dritten Projektphase (November 96) wurde die Evaluierung hauptsächlich von der Benutzerseite aus gesteuert. So wurde z.B. das Korrekturverhalten eines menschlichen Korrektors mit den Vorschlägen des MULTILINT-Systems verglichen, um die realistischen Anwendungsmöglichkeiten der Tools einschätzen zu können und konkrete Hinweise für ihre Verbesserung zu erhalten.

Aufgrund der Ergebnisse läßt sich folgendes formulieren:

- Die Entwicklung eines Tools zur Stilkontrolle für die Postedition von technischen Texten ist durchaus gerechtfertigt, da Stilfehler den weitaus größten Teil der manuellen Korrekturen ausmachten.
- Der Abdeckungsgrad der Rechtschreib- und Stilkontrolle von MULTILINT konnte auf 65% gesteigert werden, was einem befriedigenden Ergebnis entspricht.
- Das ‘Geräusch’ der Stilkontrolle konnte auf ca. 20% vermindert werden, was ebenfalls einem akzeptablen Ergebnis entspricht.

Eine vollständige Überlappung der automatischen und der manuellen Korrekturen der vorliegenden Evaluierung kann und soll nicht erreicht werden. Jeder Korrektor bringt anderes Wissen und andere Vorlieben in den Korrekturprozeß mit ein und unterscheidet sich somit von seinen Kollegen und dem automatischen Tool. Eine Optimierung der Korrekturergebnisse kann nur im Arbeitsprozeß selbst erreicht werden.

6.3 Einschätzung des Prototypen durch die Technischen Autoren

MULTILINT kann zwei Zielgruppen bedienen: Technische Autoren und Lektoren. Zum gegenwärtigen Zeitpunkt sind die Benutzer von MULTILINT Technische Autoren.

Folgende Anforderungen sind für beide Zielgruppen zentral: MULTILINT muß auf Knopfdruck im eigenen Arbeitsumfeld zu Verfügung stehen, muß schnell sein und gute Ergebnisse liefern. Hinter diesen Anforderungen stehen folgende Entwicklungsbereiche: Integration, Design und Funktionalität.

Für alle drei Entwicklungsbereiche war während des Projektverlaufs ein stetiger Fortschritt zu erkennen. Mit Ende des Projekts sind diese drei zentralen Anforderungen im Ansatz realisiert, zeigen jedoch deutlich einen neuen Startpunkt für die Weiterentwicklung auf: von der derzeitigen Prototypen-Version in Richtung Produktiv-Version.

Die Vorteile, die die Autoren in MULTILINT sehen, sind: Die Autoren können erstens MULTILINT in ihrem Workflow genau dann einsetzen, wenn sie Hilfe benötigen, und akzeptieren zweitens von einer Maschine besser sprachliche Korrekturen als von einem menschlichen Korrektor. Grundbedürfnis der Autoren ist es, MULTILINT freiwillig einzusetzen und die einzelnen Tools frei zu wählen.

Die Autoren sehen die Notwendigkeit ein, sachlich richtige UND sprachlich verständliche Dokumente zu erstellen. Sie haben folgende Faustregel bereits im eigenen Arbeitsumfeld erfahren: “Je früher ein Fehler erkannt und behoben werden kann, um so kostengünstiger ist die Korrektur”.

MULTILINT ist genau das Werkzeug, das Probleme im sprachlichen Bereich erkennen und dort die Autoren unterstützen kann. Da in der technischen Dokumentation das Hauptaugenmerk auf dem Inhalt, also auf der technischen Information liegt, ist ein Werkzeug zur sprachlichen

Qualitätssicherung nicht nur eine sinnvolle, sondern sogar notwendige Ergänzung, um Qualität zu erreichen und Kosten zu sparen. Durch die Beseitigung sprachlicher Probleme wird verhindert, daß sich die Verständlichkeitsprobleme in der Übersetzung potenzieren oder allgemein Mißverständnisse auf Benutzerseite zu negativen Folgen führen.

Seit MULTILINT über einen Knopf im SGML-Editor FrameBuilder aufgerufen werden kann, haben die Autoren die Möglichkeit, MULTILINT an Ihrem eigenen Arbeitsplatz zu benutzen. Eine entscheidende Voraussetzung für die Akzeptanz der Zielgruppe war - zusätzlich zur Integration in den Editor - die Gestaltung der Benutzeroberfläche.

Die Autoren haben den Prototypen MULTILINT akzeptiert, da sie die Möglichkeit sehen, daß MULTILINT *das* Standard-Werkzeug des Technischen Autors werden kann. Im Prinzip ist MULTILINT das Werkzeug, das sich jeder Technische Autor wünscht und wünschen muß. Dennoch sehen die Autoren am Prototypen nicht nur das Potential, das in MULTILINT steckt, sondern auch zahlreiche Möglichkeiten zur Verbesserung und Erweiterung, nicht nur im Bereich Integration und Akzeptanz, sondern auch in den Bereichen Design und Funktionalität.

7 Publikationen und Präsentationen

Der Projektfortgang, die wissenschaftlichen Ergebnisse und das Entstehen des Prototypen wurden in verschiedenen Formen dokumentiert und den Projektbeteiligten und der Öffentlichkeit zugänglich gemacht.

Für die interne Projektarbeit wurde der vollständige Schriftverkehr und alle erstellten Dokumente und Veröffentlichungen auf einer CD systematisiert und abgespeichert, die den Projektbeteiligten als Arbeitsmittel zur Verfügung steht. Diese Sammlung ist als vertraulich zu betrachten und nur für die teilnehmenden Partner und die Projektträger und -förderer verfügbar.

Als besonders wichtig für die Öffentlichkeitsarbeit haben sich die Präsentationen von MULTILINT auf der CeBit in Hannover (1997 und 1998) erwiesen, wo ein breites Anwenderpublikum auf die Möglichkeiten des Prototypen aufmerksam gemacht werden konnte. Mit geringerer Breitenwirkung gilt dies auch für den Technologie-Transfertag "Informations- und Kommunikationstechnologien" am 4. Dezember 1997 in Kaiserslautern.

Zusätzlich ist MULTILINT auf zahlreichen wichtigen wissenschaftlichen und/oder zielgruppenspezifischen Tagungen und Kongressen vertreten gewesen, die innerhalb der Projektlaufzeit im In- und Ausland stattgefunden haben. Die entsprechenden Publikationen sind im folgenden chronologisch aufgeführt.

- Jörg Schütz, "Terminology and Information Engineering", Proceedings of TKE 96, Wien, 26.-28. August 1996
- Jörg Schütz, "Combining Language Technology and Web Technology to Streamline an Automotive Hotline Support Service", Proceedings of AMTA 96, Montreal, 2.-4. Oktober 1996
- Ursula Reuther, "Linguistic Knowledge in an Industrial Application" Vortrag an der Universidad Autónoma de México (UNAM), Departamento Lingüística Aplicada, 23. Oktober 1996
- Johann Haller, *MULTILINT - Multilingual Documentation with Linguistic Intelligence*, Aslib, 'Translating and the Computer', London, 13./14. November 1996

- Johann Haller, Jörg Schütz, *Integration linguistischer Intelligenz in die multilinguale technische Dokumentation*, Proceedings “Sprache ohne Grenzen”, DLR, Projektträger Informationstechnik, Köln, 1997
- Schweiberer, O., *Dokumente prüfen mit MULTILINT*, itl, 1997.
- Johann Haller, *MULTILINT*, MT Anwenderkreis, Luxemburg, 6./7. Februar 1997
- Jörg Schütz/Johann Haller, *MULTILINT Goes European*, Kommission der EU, Luxemburg, 25./26. Februar 1997
- Jörg Schütz, *Multilingual Product Documentation*, UNU, Tokio, 27. Februar 1997
- Johann Haller, *Technische Dokumentation mit MULTILINT*, Deutsch-Chinesischer Kooperations-Workshop, Konstanz, 24./25. Juni 1997
- Ursula Reuther, *Multilinguale Informationsverarbeitung*, Open House am IÜD, Universität des Saarlandes, 27. Juni 1997
- Johann Haller, Jörg Schütz, *Integration linguistischer Intelligenz in die multilinguale technische Dokumentation*, Forum “Sprache ohne Grenzen”, München, 4./5. November 1997
- Ursula Reuther, *MULTILINT - Multilinguale Dokumentation mit linguistischer Intelligenz*, Tagungsband der tekOM Herbsttagung 1997 in Mannheim, 13./14. November 1997
- Johann Haller, Projektvorstellung MULTILINT bei der GMD, Darmstadt, 2. Dezember 1997
- Michael Carl, Antje Schmidt-Wigger, Munpyo Hong, *KURD - A Formalism for Shallow Post Morphological Processing*, Proceedings of the Natural Language Processing Pacific Rim Symposium 1997, Phuket, Thailand, Dezember 1997
- Michael Carl, Antje Schmidt-Wigger, *Shallow Post Morphological Processing with KURD*, Proceedings of the International Conference on New Methods in Language Processing, Sydney, Australia, Januar 1998
- Antje Schmidt-Wigger, *Grammar and Style Checking for German*, Proceedings of the 2nd International Workshop on Controlled Language Applications, Pittsburgh, 21./22. Mai 1998
- Ursula Reuther, *Controlling Language in an Industrial Application*, Proceedings of the 2nd International Workshop on Controlled Language Applications, Pittsburgh, 21./22. Mai 1998
- Antje Schmidt-Wigger, *Building Consistent Terminologies*, erscheint in: Proceedings of the First Workshop on Computational Terminology, Montréal, 15. August 1998