

## Sprachtechnologie und maschinelles Lektorat im Unternehmen

# Die Kontrolle der Kontrolle

Von Petra Drewer, Sybille Horend und Michael Enders

Ohne Sprachkontrolle ist Technische Redaktion heute kaum mehr vorstellbar. Um Technische Dokumentation effizient und kostengünstig zu erstellen, zu verwalten, zu übersetzen und zu publizieren, sind sprachliche Standardisierung und Konsistenz unerlässlich.



**Prof. Dr. Petra Drewer** ist Professorin für Sprache und Sprachwissenschaft im Studiengang Technische Redaktion an der Hochschule Karlsruhe. Zu ihren Forschungsschwerpunkten, in denen sie auch Beratertätigkeiten wahrnimmt, gehören mehrsprachige Dokumentationserstellung und Übersetzen, Terminologiemanagement und Verständlichkeitsforschung.

Ohne Lektorat ist der ehrgeizigste Redaktionsleitfaden sinnlos. Diese Einsicht ist so alt wie der Beruf des Lektors, doch hat in den letzten Jahren der menschliche Lektor maschinelle Konkurrenz bekommen. Sprachkontroll-Tools erobern den Markt. Dies liegt nicht nur daran, dass das menschliche Lektorat meist kostenintensiv und leider nicht immer fehlerfrei arbeitet, sondern auch an der Konstellation, dass ein Lektorat der Erstellung stets nachgeschaltet werden muss.

Im Folgenden soll erläutert werden, was ein solches Sprachkontroll-Tool (Controlled Language Checker – CLC) eigentlich ist, wozu es dient, wie es grundsätzlich funktioniert und wo seine Stärken und Schwächen liegen.

### CLC – Definition und Funktionsweise

Bei den folgenden Ausführungen beziehen wir uns auf zwei Produkte, die

speziell für Texte der Technischen Redaktion entwickelt wurden:

- CLAT vom IAI (Institut der Gesellschaft zur Förderung der Angewandten Informationsforschung an der Universität des Saarlandes e. V.) in Saarbrücken
- acrocheck von der acrolinx GmbH in Berlin (Spin-off des DFKI Deutschen Forschungszentrums für künstliche Intelligenz, Saarbrücken/Berlin)

Die Werkzeuge kommen bereits während der Erstellung der Texte durch den Autor zum Einsatz. Hat der Autor ein Modul oder mehrere Module erstellt, so kann er sie sofort auf die richtige Rechtschreibung, Grammatik und Terminologie sowie auf die für seine Texte festgelegten Stilregeln prüfen.

Der Controlled Language Checker wird über eine Schnittstelle mit dem jeweiligen Erstellprogramm des Autors verbunden und prüft sowohl Standardregeln, die beim Kauf mitgeliefert werden, als auch kundenspezifische

**D**ie sprachlichen Regeln aus einem Redaktionsleitfaden oder einer kontrollierten Sprache sind im Normalfall gut gemeint und wohlgedacht. Doch wird auch ihre Einhaltung überwacht? Wer prüft, ob sich alle Beteiligten an die Vorgaben halten? Inwiefern können so genannte Controlled Language Checker die sprachliche Kontrolle unterstützen?

### Flare ohne Hilfe ist wie Hilfe ohne Flare



**Flare**  
das XML basierte  
Autorensystem für  
Online-Hilfen

+



**Mimic**  
Software  
Simulation  
ganz einfach

+



**Capture**  
Bildschirmabzüge  
leicht gemacht

+



**Echo**  
Dynamische  
Ton-Integration

=



**MadPak™**

Die „Ultimative Autoren Suite“  
für echtes Single-Sourcing

Online-Hilfen  
leicht gemacht -  
neueste Produkte,  
modernste Technologie,  
einzeln oder im Paket!

Jetzt gratis Testversionen anfordern!  
[www.cognitas.de](http://www.cognitas.de)

**cognitas.**  
Wissen auf den Punkt gebracht



**Dr. Sybille Horend** arbeitet nach mehrjähriger Tätigkeit als Lektorin, Produktmanagerin und Technische Redakteurin aktuell als Terminologiebeauftragte bei DaimlerChrysler. Sie ist zudem Lehrbeauftragte im Studiengang Technische Redaktion an der Hochschule Karlsruhe sowie der Hochschule der Medien in Stuttgart.



**Michael Enders**, Automobil-Journalist mit Wirtschaftsdiplom und Ingenieur-ausbildung, führte fachverantwortlich Sprachtechnologie in Redaktionssysteme bei DaimlerChrysler ein. Ab 1992 war er als Projektleiter für die Erstellung und Produktion von mehr als 100 mehrsprachigen Publikationen zuständig, schrieb Redaktionshandbücher und -leitfäden. Seit Anfang April ist er für Technische Dokumentation bei der Kverneland Group in Gottmadingen verantwortlich.

Sprachregeln, die individuell erstellt werden müssen. Jeder Regelverstoß bei der Textproduktion wird gemeldet und erscheint meist mit einer Fehlerbeschreibung sowie einem konkreten Verbesserungsvorschlag, den der Autor annehmen oder verwerfen kann.

Eine automatische Korrektur kann nicht stattfinden, da die menschliche Sprache meist zu komplex ist für eine automatische Umformulierung. So liegt das letzte Wort wie auch bisher beim Autor. Über die Unterstützung bei der Erstellung hinaus können zudem Fehlerprotokolle erstellt und damit die Qualität der Texte überwacht werden.

### CLC – Einführung und Betrieb

Um Sprachtechnologie optimal in einem Unternehmen einzuführen, sollten idealerweise schon im Vorfeld die Unternehmensstrukturen auf die neuen Anforderungen ausgerichtet werden und entsprechende Vorarbeiten bereits geleistet sein, zum Beispiel die Festlegung von Terminologie und Stilregeln. Beginnen die Vorarbeiten zu spät, so entsteht das Risiko, dass der erwartete Effekt, einen signifikant verbesserten Text zu erhalten, ausbleibt. Die Akzeptanz des CLC sinkt.

Eine hohe Effizienz des CLC wird zudem erreicht, wenn die Verantwortlichkeiten für die Sprachkontrolle innerhalb eines Unternehmens an einer

zentralen Stelle zusammenlaufen. Hierzu sollte eine feste Rolle, zum Beispiel die der „Linguistischen Administration“, eingeführt werden. Diese Position bildet dann die Schnittstelle zwischen Redaktion (Redakteur/Lektor), IT-Betrieb und dem Entwickler der Sprachtechnologie.

Lange vor der Einführung der Sprachtechnologie sollten aus dieser Rolle heraus die im Unternehmen zu prüfenden Dokumentationsarten im Rahmen einer Sprachanalyse geprüft und ausgewertet werden. Als Quintessenz resultiert aus dieser Arbeit ein Basispaket von firmenspezifischen Stilregeln und Prüfmechanismen, die in der Folge implementiert werden können.

Ein weiterer wichtiger Schritt der Sprachanalyse ist die Ermittlung von Termkandidaten als Basisvokabular. Diese Kandidaten sollten sprachlich aufbereitet und bereinigt als Terminologie zum Einführungsstermin des CLC zur Verfügung stehen. Es empfiehlt sich dabei, schon von Anfang an bei der Festlegung von Terminologie auf die Belange der Autoren hinsichtlich der tatsächlichen Verwendung der Inhalte einzugehen. Sind Terminologie und Regeln erarbeitet, so kann sich die „Linguistische Administration“ der jeweiligen Schnittstelle widmen und die ersten Anbindungstests durchführen.

## 3D-CAD Daten

weiterverwenden,  
aufbereiten & publishen



Animationen leicht gemacht  
Automatisch explodieren  
Automatisch demontieren  
Assoziative Positionsnummern  
2D-Ausleitung

www.kisters.de/viewer  
E-mail: info@kisters.de  
Telefon: 0241 9671-0



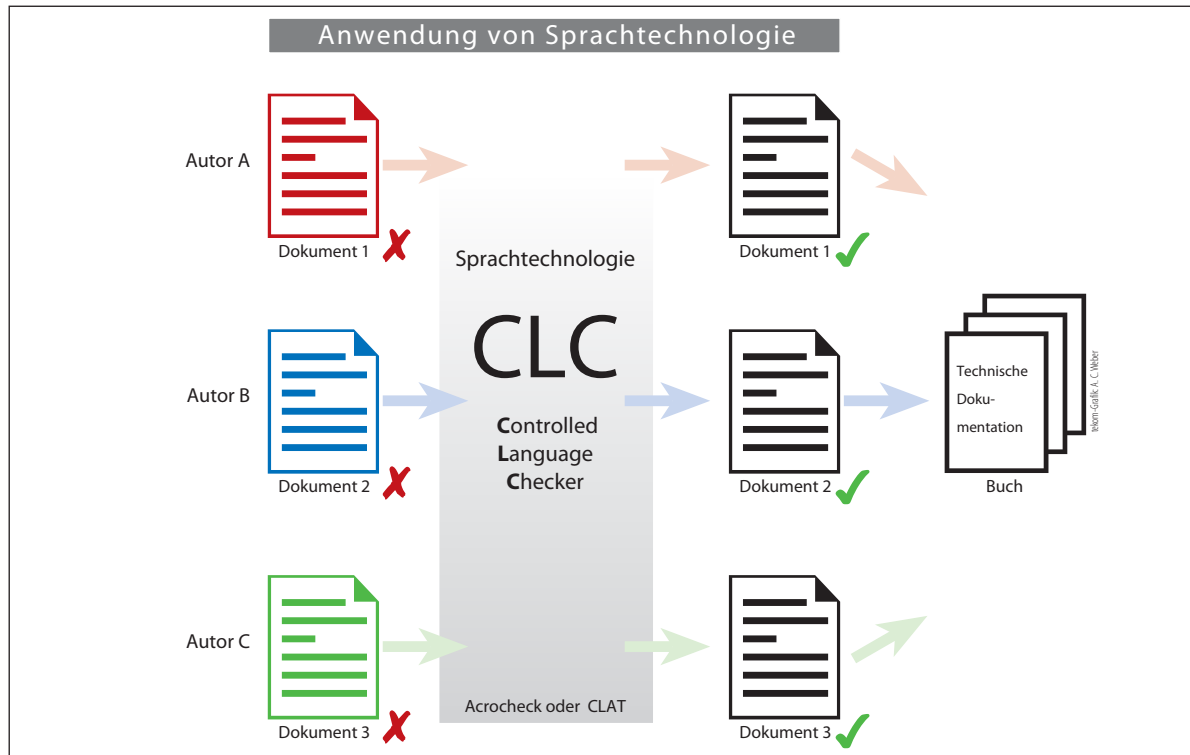


Abb. 1: Anwendung von Sprachtechnologie – vor Anwendung der Sprachtechnologie erweisen sich die Dokumente (1–3) der verschiedenen Autoren (A, B, C) als sprachlich inkonsistent: Unterschiedliche Schreibstile sowie fehlende beziehungsweise inkonsistente Terminologie erschweren den Lesefluss und den Informationstransfer ebenso wie Grammatik- und Rechtschreibfehler. Nach erfolgreicher Korrektur sind die Dokumente terminologisch einwandfrei, konsistent und entsprechen in Sprache und Stil der unternehmensspezifischen Fachsprache. Darüber hinaus sind sie konform zur Corporate Language des Unternehmens.

Von der „Linguistischen Administration“ ist zudem große Durchsetzungsstärke gefordert. Denn bei der Einführung von Sprachtechnologie müssen gewachsene und starre Erstellungsstrukturen und -prozesse innerhalb eines Unternehmens aufgebrochen und neu ausgerichtet werden. Dies ist besonders problematisch, da die Redaktionen neben dem vorhandenen Termindruck beim Erstellungsprozess den Mehraufwand durch die Einführung der Sprachtechnologie auffangen müssen.

Bei fortschreitender Integration wird dieser Aufwand immer geringer, tendiert jedoch nie gegen null, denn Sprache lebt und ist dynamisch, weshalb immer wieder Anpassungen nötig sind. Folglich ist auch ein Änderungsmanagement zu berücksichtigen ebenso wie ein möglicherweise notwendiges Migrationskonzept für den Altbestand von Dokumenten, die je nach Verwendung nicht immer verändert werden dürfen.

Um die Sprachqualität im Gesamtunternehmen zu steigern, ist es sinnvoll, die Sprachkontrolltechnologie schon zu Beginn der Prozess- und Wertschöpfungskette einzusetzen – beispielsweise in der Entwicklungsabteilung. So wird ein hoher Grad an Konsistenz erreicht,

da alle nachgeschalteten Prozesse schon mit geprüften Inhalten und Terminologie versorgt werden.

### Akzeptanz durch Kommunikation

Die Einführung einer neuen Form der Kontrolle wird nicht auf allen Seiten sofort als nutzbringend akzeptiert. Schließlich wird die Kreativität der Autoren eingeschränkt. Außerdem ergeben sich neue Möglichkeiten des Fehlernachweises. Durch eine intensive Vorbereitung der Beteiligten auf die bevorstehende Einführung eines CLC kann man diese Probleme jedoch gut abfedern. Es muss dabei das Verständnis reifen, dass der neue Arbeitsablauf, der nicht mehr der gewohnten Ergonomie entspricht, als Erstellhilfe dienen soll, die die Arbeit erleichtert.

Die Einführung neuer Prozesse und Werkzeuge ist ohnehin ein Szenario, mit dem sich nicht jeder Anwender gern auseinandersetzt. Somit muss die Einführung von Sprachtechnologie zusammen mit den Anwendern und Autoren erfolgen. Notwendig ist hier vor allem eine begleitende Unterstützung der Autoren bei linguistischen und techni-

schen Themen per Redaktionskonferenz, per E-Mail oder per Telefon.

### CLAT und acrocheck

Aus theoretischer Perspektive ist bereits einiges über „Sinn und Nutzen von CLC“ diskutiert worden. Aber was können diese anscheinend so hilfreichen Produkte nun tatsächlich und welche Unterschiede gibt es? Um diese Frage zu beantworten, betrachten wir die beiden Controlled Language Checker acrocheck und CLAT.

Ein direkter Vergleich der beiden Produkte im laufenden Betrieb kann nur in einer konkreten, für beide Tools identischen Systemumgebung durchgeführt werden, die wir für den angestrebten Kurzvergleich nicht einrichten konnten. Aufgrund der mangelnden Vergleichbarkeit der Installationen sowie der sehr großen Funktionsumfänge haben wir uns darum bemüht, die jeweiligen relevanten Produktumfänge darzustellen und auf Besonderheiten hinzuweisen, wobei die Darstellung keinen Anspruch auf Vollständigkeit erhebt. Die Preise der Produkte können jeweils beim Hersteller angefragt werden und richten sich nach dem notwen-

digen Funktionsumfang, der Anzahl der Nutzer und der Art der Integration in die bestehende IT-Landschaft des Unternehmens.

### 1. Prüfungen

Grundsätzlich bieten beide Produkte folgende Prüfungen an:

- Rechtschreibung: bei CLAT neue deutsche Rechtschreibung nach Duden oder progressive/konservative Varianten, bei acrocheck spezielles Modul „Rechtschreibung für die Technische Dokumentation“
- Grammatik: bei CLAT entsprechend zum Duden KORREKTOR, bei acrocheck basierend auf Duden und linguistischer Analyse großer Datenmengen aus der Technischen Dokumentation
- Stil: einige Stilregeln greifen auf Informationen der DTD zurück und können strukturelle Vorgaben mit linguistischen Vorgaben verknüpfen; die Optimierung der Stilregeln erfolgt bei acrocheck in Zusammenarbeit mit TÜV SÜD DocLab
- Terminologie: Konzeptbasierter Ansatz, der drei Benennungsarten unterscheidet:
  - a) Vorzugsbenennung
  - b) Negativbenennungen (CLAT) gesperrte Terme (acrocheck),
  - c) zugelassene Benennungen (CLAT)/erlaubte Terme (acrocheck)

Es gibt zudem bei beiden Tools die Möglichkeit, Termkandidaten automatisch zu erkennen und in den Bestand aufzunehmen sowie Termvarianten (Abkühlvorgang/Abkühlungsvorgang/Vorgang des Abkühlens ...) zu identifizieren. Außerdem kann die konsistente Nutzung von besonderen Schreibweisen zum Beispiel für Marketing-Benennungen sichergestellt werden. Das Programm acrocheck enthält zudem eine eigene Termdatenbank. Der Terminologieimport kann bei beiden Produkten manuell und automatisch zu festgelegten Zeiten und in regelmäßigen Abständen erfolgen.

Die Anwendung acrocheck verfügt über die Zusatzfunktion, Standardphrasen und -sätze zu hinterlegen und vorzuschlagen. Sie erfüllt damit auch die grundlegende Funktion eines Authoring-Memory-Systems.

Die Prüfung selbst kann bei beiden Produkten jeweils online oder offline erfolgen. In diesem Zusammenhang beinhaltet acrocheck die Möglichkeit, Prüfberichte zu erstellen. Diese Auswertung von Dokumenten bietet das IAI mit der Applikation „ZertiFakt“.

### 2. Randbedingungen

Im Hinblick auf Ergänzungen, zum Beispiel neue Regeln, ist es bei beiden Produkten möglich, die linguistischen Ressourcen zu aktualisieren, ohne dabei die Software anzupassen und umgekehrt. Bei beiden werden die lingu-

istischen Ressourcen zentral auf einem Server verwaltet.

Bei acrocheck und bei CLAT können vom Kunden „Projekte“ mit unterschiedlichen linguistischen Ressourcen zum Beispiel für verschiedene Textsorten eingerichtet und verwaltet werden.

Der linguistische und der technische Support erfolgen ebenfalls ähnlich. Es gibt eine Telefon-Hotline, eine Fehlerdatenbank im Internet sowie Support über E-Mail. Auf Wunsch kann dieser Kontakt durch Vor-Ort-Support erweitert werden.

### 3. Rollen und Rechte

Bei beiden Produkten gibt es sehr ähnliche Rollen- und Rechtekonzepte.

CLAT unterscheidet zwischen Redakteur, Lektor, Terminologe und Linguistischem Manager; acrocheck unterscheidet zwischen Anwender, Linguistischem Manager und IT-Administrator.

Welche Rechte sich hinter diesen Rollen verbergen können, verdeutlicht das folgende Beispiel:

Rolle	Erlaubte Aktion
Redakteur	Prüfen und korrigieren Neue Terme an Linguistischen Manager schicken
Administrator	Neue Nutzer anlegen Termdatenbanken anlegen Terminologie synchronisieren

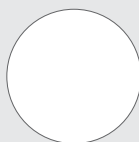
## cognitas - Prüfprogramme für Adobe FrameMaker



### TermControl

Ihre sprachunabhängige Terminologie-Prüfsoftware

### DUDEN Korrektor für Adobe FrameMaker



### Language Checker

Ihre Unterstützung für Simplified English

## Eigentlich sollten wir diese Tools nicht verkaufen -

Doch weil uns die Qualität jeglicher Technischer Dokumentation am Herzen liegt, verkaufen wir sie auch - sogar an Wettbewerber!

Jetzt gratis Testversionen anfordern!  
www.cognitas.de



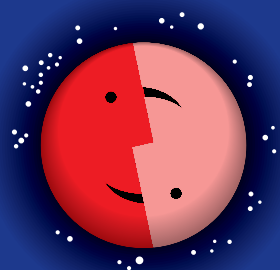


COMING  
SOON

2007

PASSOLO

BROUGHT TO  
YOU BY  
THE EXPERTS  
IN .NET  
LOCALIZATION



www.passolo.com

#### 4. Editoren

Bei der notwendigen Integration in ein Content-Management-System ist das Zusammenspiel der CLC mit dem jeweiligen Editor von großer Bedeutung.

Das IAI bietet eine Integration von CLAT in PTC Arbortext, MS Word und Adobe FrameMaker an. Es bestehen Partnervereinbarungen mit folgenden Firmen: AGI, Dosco GmbH, ESG GmbH und Sun Microsystems Inc. Zudem besteht seit 2000 eine enge Kooperation mit Duden und Brockhaus.

Acrocheck verfügt über Integrationen in PTC Arbortext, Adobe FrameMaker, MS Word, Just Systems XMetaL und AuthorIT. Partnerschaften und enge Kontakte bestehen auch zu Documentum, Schema, Fischer, Noxum und Docufy sowie zu TÜV DocCert, Idiom, Star, across und SDL/Trados.

#### 5. Betriebssysteme

- Betriebssystem Server CLAT: Windows XP, 2000, 2003, Linux Suse und RedHat, Mac OS X, Solaris 8 und höher
- Betriebssystem Server acrocheck: Windows, Unix (Linux, Solaris)
- Betriebssystem Client CLAT: Windows XP, 2000, 2003, Linux Suse und RedHat, Mac OS X, Solaris 8 und höher
- Betriebssystem Client acrocheck: Windows (Batchversion unter Windows, Unix)

#### 6. Die Teams

- Das IAI hat derzeit 19 fest angestellte und zwei freie Mitarbeiter. Davon sind sechs promovierte Linguisten beziehungsweise Computerlinguisten.
- Das Team von acrolinx besteht aus 25 festen und zwei freien Mitarbeitern. Davon sind fünf promovierte Computerlinguisten.

#### Fazit

Das maschinelle Lektorat ist nicht der heilige Gral der Technischen Dokumentation – aber es erleichtert die Arbeit im Tagesgeschäft spürbar. Unternehmen, die Sprachtechnologie zum Einstellen und Prüfen von Sprachregeln im eigenen Hause realisieren und umsetzen möchten, haben mit acrocheck und CLAT die Wahl zwischen zwei sehr mächtigen Systemen, die Kompetenz und umfangreiche Gestaltungsmöglichkeiten im CLC-Umfeld bieten. Generell haben beide Systeme ihre besonderen Stärken – und Schwächen, die es individuell abzuwägen gilt. Wir schlagen interessierten Redaktionen vor, sich anhand entsprechender Präsentationen der Anbieter selbst einen Eindruck darüber zu verschaffen, welche Lösung für ihre jeweilige Systemumgebung und ihre jeweiligen linguistischen Anforderungen die richtige ist. Diese kleine Investition für einen großen Schritt in die Zukunft der Technischen Redaktion lohnt sich auf jeden Fall.

#### Autorenanschrift

##### Prof. Dr. Petra Drewer

Hochschule Karlsruhe – Technik und Wirtschaft  
petra.drewer@hs-karlsruhe.de  
www.technischeredaktion.com

##### Dr. Sybille Horend

horend@gmx.de

##### Michael Enders

Kverneland Group  
michael.enders@kvernelandgroup.com