



Maschinelle Autorenunterstützung für Simplified Technical English

Simplified Technical English (STE) kann, richtig angewendet, zu Texten führen, die schnell und leicht verständlich sind, und zwar für alle Leser. Dies wird durch eine Einschränkung von Wortschatz, Syntax und Satzmuster erreicht. Der Fokus dieses Dokuments liegt darauf, zu zeigen, wie Software zur Autorenunterstützung beim Schreiben nach STE unterstützen kann. Dabei werden auch ganz konkrete Anwendungsbeispiele illustriert.

| Inhalt

1	Einleitung	3
2	Vorteile einer maschinellen Sprachprüfung	3
3	Einordnung präskriptiver vs. proskriptiver Ansatz	4
4	Zur maschinellen Prüfbarkeit des Standards	6
5	Wie kann eine maschinelle Unterstützung beim Schreiben nach STE aussehen	7
5.1	Maschinelle Prüfbarkeit in der Praxis	7
	Wortregeln	7
	Nomenregeln	9
	Verbregeln	10
	Satzregeln	10
	Informationsart-spezifische Regeln	11
	Regeln zu Zeichensetzung und Wortanzahl	13
	Schreibvorgaben	14
5.2	Analyse und Bewertung der maschinellen Umsetzbarkeit	16
6	Fazit	18

1 Einleitung

Das ASD-STE100 Simplified Technical English, kurz STE, ist eine der bekanntesten Kontrollierten Sprachen in der Technischen Dokumentation. STE hat seinen Ursprung in der Luft- und Raumfahrt, wird aber auch in anderen Bereichen der Technischen Dokumentation genutzt. Das Regelwerk limitiert und standardisiert die englische Sprache. Es besteht aus einer Sammlung von Schreibregeln und einem Wörterbuch mit eingeschränktem Vokabular.

STE kann, richtig angewendet, zu Texten führen, die schnell und leicht verständlich sind, und zwar für alle Leser: für ein internationales Zielpublikum, für Muttersprachler und Nichtmuttersprachler, für Experten und Laien. Dies wird durch eine Einschränkung von Wortschatz, Syntax und Satzmuster erreicht.

Theoretisch kann man als Technischer Redakteur mit STE starten, nachdem man sich die aktuellste Ausgabe des Standards besorgt und aufmerksam gelesen hat. Das ist jedoch kein leichtes Unterfangen und nach Einschätzung der Herausgeber des Standards wenig sinnvoll, denn niemand kann sich 53 Schreibregeln und über 3000 Wortverwendungsregeln einfach so merken. Aus diesem Grund sind eine einleitende Schulung sowie maschinelle Unterstützung bei der Umsetzung des Standards unumgänglich.

Dieses Dokument beginnt mit einem Kapitel zu den Vorteilen einer maschinellen Autorenunterstützung. Daraufhin folgen eine Verortung des STE in die verschiedenen Ansätze Kontrollierter Sprachen und eine Bewertung der maschinellen Prüfbarkeit des Standards. Der Fokus des Whitepapers liegt schließlich darauf, am Beispiel des Congree Authoring Servers zu zeigen, wie Software zur Autorenunterstützung beim Schreiben nach STE unterstützen kann. Dabei werden auch ganz konkrete Anwendungsbeispiele illustriert.

2 Vorteile einer maschinellen Sprachprüfung

Der Einsatz einer maschinellen Sprachprüfung birgt zahlreiche Vorteile – besonders, aber nicht nur, wenn nach STE geschrieben wird.

Der zentrale Punkt ist, dass sich durch eine entsprechende Software sowohl Style Guide oder Redaktionsleitfaden, als auch Standards wie STE abbilden lassen. Autoren werden darin unterstützt, die Regeln zu beachten und umzusetzen. Dabei ist es ein großer Motivationsfaktor, dass der manuelle Aufwand zum Nachschlagen und Recherchieren im Style Guide oder Standard wegfällt.

Eine maschinelle Sprachprüfung ermöglicht eine objektive und konsistente Prüfung der Inhalte. Sie

ersetzt nicht vollständig das Humanlektorat, da manche Aspekte mit dem Weltwissen eines Menschen beurteilt werden müssen. Sie verringert jedoch den Aufwand im Lektorat spürbar.

Eine maschinelle Sprachprüfung geschieht objektiv und konsistent. Über die Prüfung hinaus gibt es die Möglichkeit, alle Prüfergebnisse detailliert über Berichte nachzuvollziehen und auszuwerten.

Indem die maschinelle Sprachprüfung dem Anwender seine Fehler aufzeigt und die entsprechenden Regeln sowie Korrekturvorschläge anbietet, kommt es zu einem automatischen Lerneffekt. Bestenfalls wird sogar der Schulungsbedarf für die Anwendung des Style Guides oder des anzuwendenden Standards reduziert.

3 Einordnung präskriptiver vs. proskriptiver Ansatz

Hinsichtlich Kontrollierter Sprachen kann zwischen dem so genannten präskriptiven und proskriptiven Ansatz differenziert werden. Kontrollierte Sprachen nach dem präskriptiven Ansatz schreiben „erlaubte“ Begriffe und Strukturen vor, während bei proskriptiven Ansätzen „nicht erlaubte“ sprachliche Entitäten festgelegt werden.

Für den präskriptiven Ansatz spricht, dass durch die Vorgabe von erlaubten Strukturen im Vergleich zum proskriptiven Ansatz weniger Formulierungsfreiheit gegeben ist, was sich durchaus positiv auf die Textkonsistenz auswirken kann, unter anderem auch, wenn mehrere Autoren beteiligt sind. Die eingeschränkte Formulierungsfreiheit könnte jedoch auch als Nachteil aus der Sicht der Autoren gewertet werden. Außerdem kann man aus der größeren Anzahl zu lernender Regeln einen vergleichsweise höheren Schulungsbedarf für den präskriptiven Ansatz vermuten.

Vorteile des proskriptiven Ansatzes sind, dass er den technisch Schreibenden größere Formulierungsfreiheit lässt und die Fehlermeldungen durch Prüfprogramme vergleichsweise spezifischer gestaltet sind – schließlich benennen diese konkret die Verstöße der Autoren. Nachteil des proskriptiven Ansatzes ist, dass bei vielen „nicht-falschen“ Möglichkeiten, einen Text zu gestalten, mehr Raum für inkonsistente Gestaltung bleibt.

Vereinfacht gesagt sind die Schwächen eines Standards die Stärken des anderen.

STE gehört überwiegend zu den so genannten präskriptiven Ansätzen, beinhaltet aber auch proskriptive Elemente. Durch die Regeln wird zumeist vorgegeben, welche sprachlichen Konstruktionen erlaubt sind. Alles, was nicht explizit erlaubt ist, gilt als verboten. Einige wenige Regeln sind proskriptiv formuliert, z. B. Regel **3.4 Do not use helping verbs to make complex verb struc-**

tures. Das STE-Wörterbuch unterliegt im Prinzip dem präskriptiven Ansatz: „The dictionary gives all the general words that are approved in STE [...]“¹. Es werden jedoch auch immer nicht-erlaubte Wörter angegeben, um die Verwendung ihrer erlaubten Synonyme zu empfehlen.

Grundlagenwissen: Chancen und Grenzen des Einsatzes von STE

✓ **Verständlichkeit:** Durch die Anwendung des Standards können Sie z. B. sicherstellen, dass Ihre englischen Texte leicht verständlich und schnell erfassbar sind. Durch die Reduktion sprachlicher Komplexität und Vielfalt werden die Texte für alle Leser verständlicher: für ein internationales Zielpublikum, für Muttersprachler und Nichtmuttersprachler, für Experten und Laien. All das kann durch die eingeschränkte Syntax und Satzmuster von STE erreicht werden.

✓ **Bessere interne Kommunikation:** Die Anwendung von STE vereinfacht die englische Sprache so weit, dass diese für Nichtmuttersprachler einfacher verständlich ist. Eine abteilungs- und länderübergreifende Kommunikation auf Englisch wird somit durch den Einsatz von STE erleichtert.

✓ **Wirtschaftlichkeit:** Die Anwendung des STE bietet wirtschaftliche Vorteile. Einsparpotenzial wird besonders im Bereich Übersetzung deutlich. Durch die vereinfachte und standardisierte Syntax und Wortwahl wird die Übersetzung aus dem Englischen heraus erleichtert und damit kostengünstiger gemacht: „As the text volume is reduced by at least 20% and the remaining text becomes more repetitive, the use of Simplified Technical English typically results in 30 to 40% less translation cost“². Darüber hinaus ist es durch den STE-Standard leichter, als Nichtmuttersprachler englische Texte direkt selbst zu verfassen, anstatt erst Quelltexte in der eigenen Sprache zu verfassen. Übersetzungen werden durch STE also in manchen Anwendungsszenarien bestenfalls obsolet.

● **Arbeitserleichterung:** Direkt zu Beginn der Anwendung bringt STE keine automatische Arbeitserleichterung mit sich. Der Standard muss erst gelernt und in die tägliche Arbeit integriert werden. Hierbei kann Software zur maschinellen Autorenunterstützung eine Hilfe sein. Wenn der Standard erst einmal verinnerlicht und aus den Arbeitsabläufen nicht mehr wegzudenken ist, wird das Schreiben durch die festen Regeln vereinfacht – der Aspekt der „kreativen Arbeit“ („Wie drücke ich das nun aus? Mit welchen Wörtern spreche ich den Leser am besten an?“) weicht dem Prinzip der vorgegebenen Sprache und Struktur.

✗ **Ergänzungen benötigt:** Im Regelwerk des Standards wird die Aussage getroffen, dass man ihn nicht alleine verwenden sollte, sondern in Kombination mit weiteren Standards und Methoden: „It is intended to be used with other applicable specifications for technical publications, style guides, and official directives.“ (ASD 2017: ASD-STE100 – Issue 7, page 14). Ein Beispiel für die Notwendigkeit der Kombination mit anderen Ressourcen stellt die STE-Wortliste dar. Diese gibt allgemeinsprachliches und technisches Vokabular vor und klassifiziert die Wörter als erlaubt oder verboten, manchmal auch in Bezug auf deren Wortart. Neben diesem Vokabular hat jedes Unternehmen auch noch seine eigene Terminologie mit Vorzugsbenennungen und Negativbenennungen. Diese Unternehmensterminologie wird vom STE nicht berührt und muss entsprechend zusätzlich eingesetzt werden.

✗ **STE nur im Gesamtpaket:** „Ein bisschen STE“ widerspricht der Konzeption des Standards. Theoretisch könnte man als Technischer Redakteur einige Schreibregeln isoliert anwenden. Der Standard selbst versteht sich jedoch als Gesamtpaket und sieht die partielle Anwendung bzw. das individuelle Heraussuchen von passenden Regeln nicht vor.

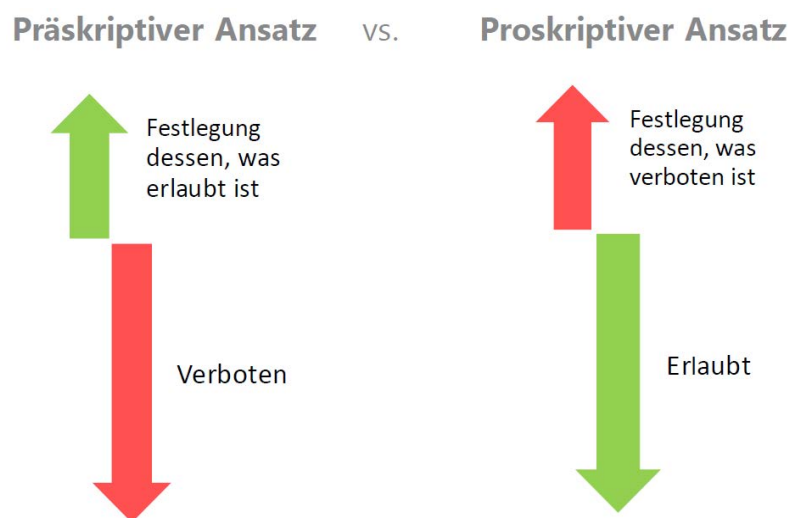
¹ ASD (2017): STE Issue 7, S. 2-0-3

² S. <http://www.shufrans-techdocs.com/benefits-of-simplified-technical-english/> (letzter Abruf: 26.11.2018, 9.00 Uhr)

4 Zur maschinellen Prüfbarkeit des Standards

Die maschinelle Prüfbarkeit eines Standards hängt immer auch vom zu Grunde liegenden Ansatz ab. Das STE gehört, wie bereits erläutert, zu den überwiegend präskriptiven Ansätzen. Im Vergleich zu proskriptiv formulierten Regeln (s. Abbildung, Bereich „Redaktionsvorgaben“), definieren präskriptive Ansätze wie das STE genau, welche sprachlichen Konstruktionen und Wörter erlaubt sind.

Dies impliziert, dass alles Abweichende verboten ist:



Um einen Text zu verbessern, ist es sinnvoll, potenzielle Fehler auszumerzen. Entsprechend muss ein Prüfprogramm dem Autor diese zurückmelden. Wenn in einem Standard jedoch mehr Dinge verboten als erlaubt sind (s. Grafik), kann dies zu einem relativ hohen Meldungsaufkommen führen.

Ein wichtiger Bestandteil des STE sind die so genannten Technical Names. Darunter versteht man technische Benennungen, die von Unternehmen zu Unternehmen verschieden sein können. Man kann sagen, dass die Technical Names im Wesentlichen den Benennungen in einer unternehmensspezifischen Terminologie entsprechen. Daraus lässt sich ableiten, dass ihre Überprüfung nur indirekt über die Schreibregeln des STE abgedeckt werden kann. Regel 1.5 bis 1.11 befassen sich mit Technical Names und legen fest, wie diese gebildet und verwendet werden sollen. Die konkreten Benennungen gibt jedoch das Unternehmen vor, das den Standard jeweils anwendet, nicht der Standard selbst. Eine Hinterlegung der Technical Names z. B. in einer Terminologiedatenbank ist somit nötig, um die Benennungen maschinell prüfbar zu machen.

5 Wie kann eine maschinelle Unterstützung beim Schreiben nach STE aussehen

5.1 Maschinelle Prüfbarkeit in der Praxis

Wortregeln

Unter Wortregeln verstehen sich die STE-Regeln 1.1 bis 1.14, deren Geltungsbereich sich nur auf die Form und die Verwendung einzelner Wörter bezieht. Neun von 14 Regeln sind maschinell prüfbar, bei drei weiteren ist eine maschinelle Überprüfung theoretisch umsetzbar, zwei weitere nicht umsetzbar.

Eine der unterstützten STE-Wortregeln ist die folgende:

1.1 You can use words that are:

- » Approved in the dictionary
- » Technical names
- » Technical verbs.

Verkürzt: Use words that are approved. Die Regel schlägt z. B. bei folgendem Beispielsatz an: „This revolutionary device will significantly improve the quality of your technical documentation without **implementing** an additional work step.“

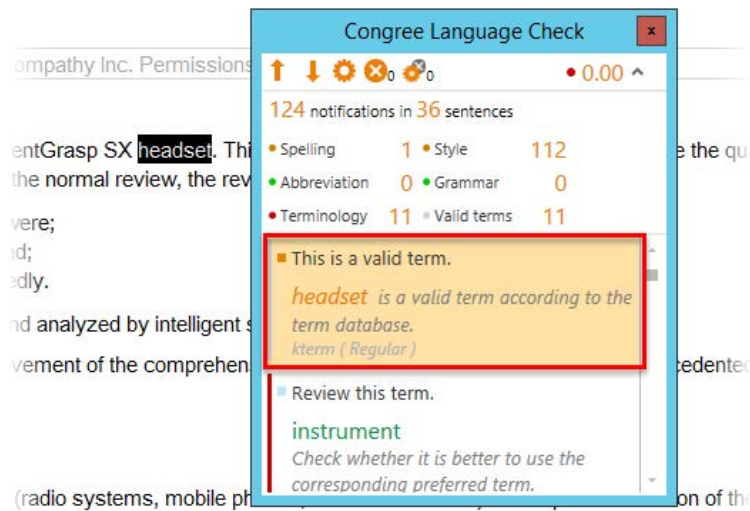
Im Rahmen einer maschinellen Prüfung, hier am Beispiel von Congree, wird auf ein eingebautes STE-Wörterbuch zurückgegriffen. Die Software identifiziert das Verb “(to) implement“ als nicht erlaubt und meldet dies mit folgender Anweisung:

The screenshot shows the Congree Language Check interface overlaid on a document. The document text includes a preface and a section on radio interference. The word "implementing" is highlighted in the preface. The Congree Language Check window displays the following information:

- 124 notifications in 36 sentences
- Spelling: 1, Style: 112, Abbreviation: 0, Grammar: 0, Terminology: 11, Valid terms: 11
- A notification box with the following text:

Use words that are approved.
 Consider rewriting the sentence using the following alternative(s): **do(v)**
 According to the STE dictionary the following word(s) is/are not approved:
implement(v)
 Style sSTE11en(Regular)
- A checkbox labeled "Use approved words as the part of" is visible at the bottom of the notification box.

Ein weiteres Beispiel ist die Regel **1.8 Use technical names that agree with approved nomenclature**. Hier geht es im Wesentlichen darum, einen Abgleich mit der vorhandenen Terminologie durchzuführen und zurückzumelden, wenn der Autor eine Vorzugsbenennung korrekt im Text verwendet hat:



Ein letztes Beispiel für die Kategorie der Wortregeln soll **1.14 Use American English spelling** sein. Bei dieser Wortregel handelt es sich um eine Orthografierregel, die nach einer bestimmten Schreibweise sucht und in der Regel einen direkt anwendbaren Korrekturvorschlag liefert:



Zur Beantwortung der eingangs gestellten Frage, welche Regeln nicht maschinell prüfbar sind – und warum –, werden die betreffenden Regeln betrachtet. Es handelt sich um:

- » 1.9 When you must select a technical name, use one which is short and easy to understand.
- » 1.10 Do not use slang or jargon words as technical names.

Bei beiden Regeln geht es um die Festlegung von Technical Names. Dabei handelt es sich jedoch nicht um eine Routine in der Texterstellung, die maschinell überprüft werden kann, sondern um Vorgaben zur Terminologearbeit. Hier kann ein Feature zum Tragen kommen, das den Autor dabei unterstützt, Benennungen für die Terminologie vorzuschlagen. Der Congree Authoring Server z. B. stellt dieses Feature in Form der Termkandidaten-Erfassung zur Verfügung.

Nomenregeln

Zum sprachlichen Komplex der Nomen-Cluster (Verbindungen mehrerer Nomen) gibt es drei STE-Regeln, von denen zwei maschinell geprüft werden können. Dabei handelt es sich um **2.1 Write noun clusters of no more than three words** und **2.3 When applicable, use an article (the, a, an) or a demonstrative adjective (this, these) before a noun**. Beide werden regulär durch eine Stilprüfung gefunden. Der Autor erhält eine Anweisung – für Regel 2.1 sieht dies am Beispiel von Congree folgendermaßen aus:

sages were;
were read;
repeatedly.

adset and analyzed by an intelligent software information collector system using a sophis
c improveme

y fields (radio

e of the Conte

points during

ises the follow

e
remote control

Congree Language Check

0.00 ^

126 notifications in 36 sentences

• Spelling	2	• Style	113
• Abbreviation	0	• Grammar	0
• Terminology	11	• Valid terms	11

Write noun clusters of no more than three words.
Use prepositions to clarify the semantic relations between nouns.
Long noun clusters are not easy to understand because the words in the noun cluster can connect to each other in different ways.
Example
Style sSTE21en(Regular)

mentation at an unprecedented

) can impair the centre of the d

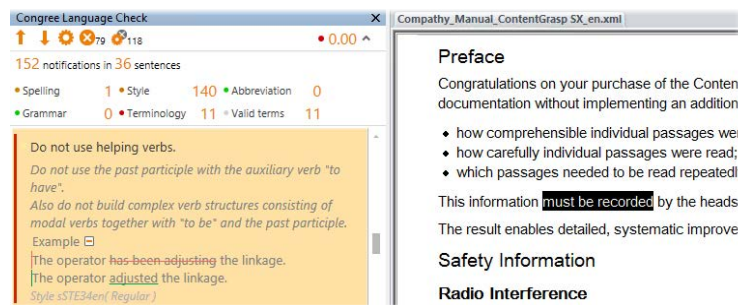
nd increased irritability. Avoid s

to protect your skin.

Verbregeln

Alle sieben Verbregeln des STE können maschinell geprüft werden. Die Verbregeln definieren nicht, welche Verben erlaubt sind, dies wird über das STE-Wörterbuch geregelt und entsprechend über die Wortregeln prüfbar gemacht.

Um ein Beispiel anhand des Congree Authoring Servers zu demonstrieren, wurde Regel **3.4 Do not use helping verbs.** ausgewählt

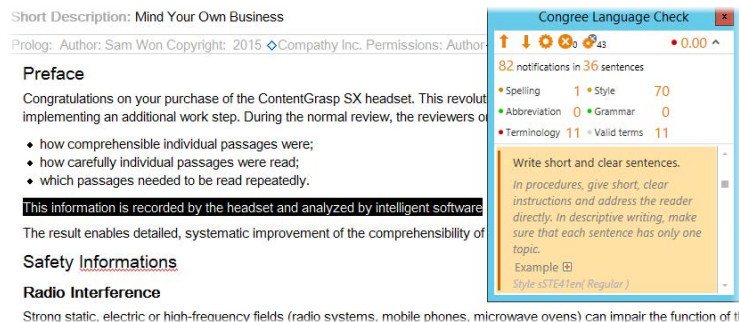


In diesem Anwendungsbeispiel wird die Modalverb-Konstruktion *must be recorded* erkannt und moniert.

Satzregeln

Der STE-Issue 7 bietet vier Satzregeln, von denen drei maschinell prüfbar sind.

Als Beispiel sei Regel **4.1 Write short and clear sentences.** genannt. Diese vereint gleich zwei Aspekte und gilt abhängig davon, ob z. B. eine Anweisung oder eine Beschreibung vorliegt. Zum einen bezieht die Regel sich auf die direkte Ansprache des Lesers mit kurzen, klaren Sätzen in Anweisungen. Zum anderen legt sie fest, dass in Beschreibungen jeder Satz nur einen Inhalt transportieren soll. In der beispielhaften Umsetzung mithilfe der Software Congree wird entsprechend moniert, dass der Autor gegen letzteren Aspekt verstoßen und zwei Vorgänge in einem Satz beschrieben hat:



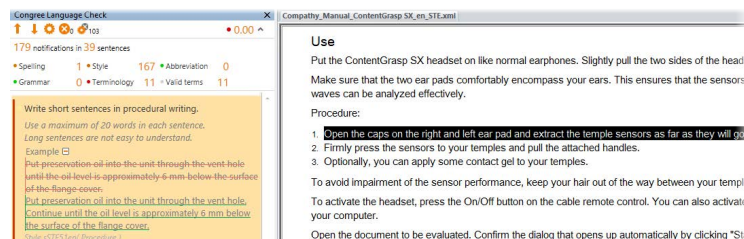
Bei der nicht maschinell prüfbaren Satzregel handelt es sich um **4.4 Use connecting words and phrases to connect sentences that contain related topics**. Der Grund für die Nicht-Umsetzbarkeit liegt daran, dass hier die Ebene des Inhalts einbezogen wird. Während sprachliche Phänomene grundsätzlich auch satzübergreifend prüfbar sind, kann das Kriterium „[...] sentences] that contain related topics“ nicht auf die maschinelle Prüfung übertragen werden. Hier ist das Weltwissen des Autors gefragt, der inhaltliche Verwandtschaft erkennen und daraufhin die passende Konjunktion auswählen kann.

Informationsart-spezifische Regeln

Die drei STE-Kapitel **5 Procedural writing**, **6 Descriptive writing** und **7 Safety instructions** beziehen sich auf bestimmte Informationsarten und sollen daher in der Folge als Informationsart-abhängige Regeln gemeinsam betrachtet werden.

Das Kapitel 5 (Handlungen bzw. Vorgänge) umfasst fünf Regeln, die alle maschinell prüfbar sind. Die Regeln beziehen sich auf unterschiedliche sprachliche Merkmale von Handlungs- und Vorgangsbeschreibungen, z. B. auf die zu verwendenden Verbformen oder auf die Anzahl von Anweisungen pro Satz. Eine Regel, mit der im Folgenden die maschinelle Prüfung von **Procedural writing**-Regeln beispielhaft gezeigt wird, ist **5.1 Write short sentences. Use a maximum of 20 words in each sentence**.

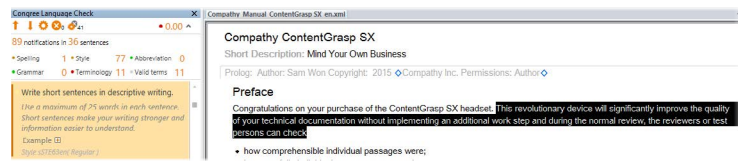
Am Beispiel der Software Congree gestaltet sich die Prüfung folgendermaßen:



Dabei muss die Textstruktur der Software bekanntgemacht werden, damit Handlungen bzw. Vorgänge anhand von Auszeichnungselementen erkannt werden. Auf dieser Basis kann die Satzlänge ermittelt und mit der vorgegebenen Länge von maximal 20 Wörtern abgeglichen werden.

Das Kapitel 6 (Beschreibungen) beinhaltet sechs Regeln, von denen zwei maschinell prüfbar sind. Regel **6.1 Give information gradually** benötigt aufgrund ihrer eher allgemeinen Formulierung eine Umsetzung mithilfe mehrerer Einzelregeln.

Eine exemplarische Prüfung durch Congree soll anhand von **6.3 Write short sentences. Use a maximum of 25 words in each sentence.** gezeigt werden:



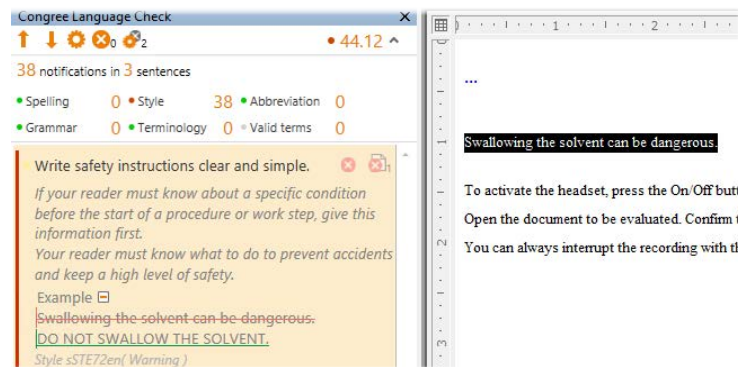
Äquivalent zu Regel 5.1 wird hier die Wortanzahl pro Satz in der entsprechenden Informationseinheit ermittelt und mit dem erlaubten Maximum abgeglichen.

Nun soll untersucht werden, warum vier der sechs Regeln nicht maschinell geprüft werden können. Es handelt sich um:

- » 6.2 Use key words and phrases to organize your text logically.
- » 6.4 Use paragraphs to show related information.
- » 6.5 Make sure that each paragraph has only one topic.
- » 6.6 Make sure that no paragraph has more than six sentences.

Bei Regel 6.2 handelt es sich nicht um eine formalisierbare Regel, sondern um eine allgemein formulierte Anweisung. Software kann nicht entscheiden, wo, wie und welche Schlüsselwörter und -phrasen in einen Text eingebracht werden sollen. Hier ist das Weltwissen und Fachwissen des Autors gefragt, um Schlüsselwörter und -phrasen zu identifizieren und korrekt zu nutzen. Regel 6.4 bezieht sich darauf, verwandte Inhalte in Absätze zu bündeln. Ähnlich wie bei Regel 4.4 wird auch hier die Inhaltsebene einbezogen. Die inhaltliche Verwandtschaft von Sätzen muss der Autor mit Welt- und Fachwissen ermitteln. Hier kann Software kein adäquates Ergebnis bringen, womit die Regel nicht maschinell umsetzbar ist. Dasselbe gilt für Regel 6.5, die sich auf den Grundsatz „Ein Absatz – ein Thema“ bezieht. Regel 6.6 befasst sich damit, dass in Beschreibungen kein Absatz mehr als sechs Sätze beinhalten soll. Diese Regel ist nicht maschinell prüfbar, da das Vorliegen eines Absatzes der Software keine Information darüber übermittelt, wie viele Prüfeinheiten in Form von Sätzen der Absatz enthält.

Kapitel 7 (Sicherheitshinweise) beinhaltet drei Regeln, von denen eine maschinell prüfbar ist. Dabei handelt es sich um **7.2 Start a safety instruction with a clear or simple command or condition.**



Die Software ermittelt, dass ein Sicherheitshinweis weder mit einer Anweisung noch mit klaren Bedingung beginnt, und moniert diesen Umstand.

Nicht umsetzbar sind:

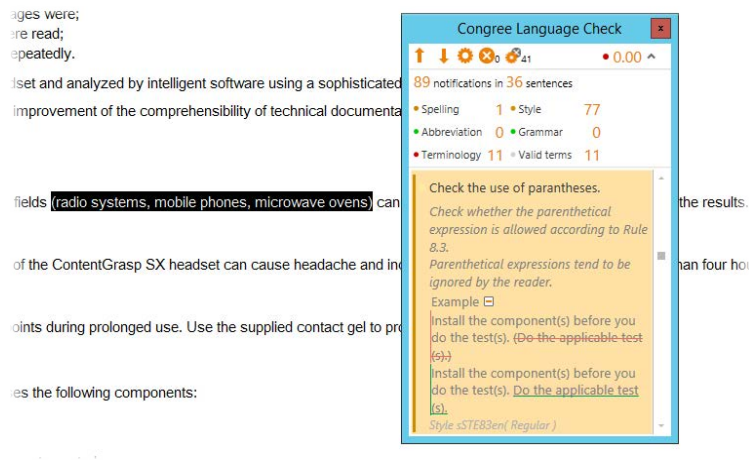
- » 7.1 Use an applicable word (for example, “warning” or “caution”) to identify the level of risk.
- » 7.3 Give an explanation to show the specific risk or possible result.

Bei 7.1 ist die Nicht-Umsetzbarkeit durch den Bezug auf die Inhaltsebene begründbar. Es geht darum, das verwendete Signalwort dem Risiko anzupassen, das dem Sicherheitshinweis zu Grunde liegt. Diese Risikobewertung muss die inhaltlich verantwortliche Person treffen, daher kann eine Software zum Prüfen der STE-Regeln hier nicht zum Einsatz kommen. Auch 7.3 ist eine rein inhaltsbezogene Anweisung, bei deren Anwendung Software nicht unterstützen kann.

Regeln zu Zeichensetzung und Wortanzahl

Alle sieben STE-Regeln zu Zeichensetzung und Wortanzahl sind maschinell prüfbar. Regel 8.6 und 8.7 befassen sich damit, welche Einheiten als ein Wort gezählt werden. Diese beiden STE-Regeln lassen sich technisch in die Wortanzahl-Regeln, 5.1 und 6.3 (s. oben), einbeziehen, und benötigen keine separate „Maschinenregel“. Die verbleibenden sechs Regeln lassen sich problemlos 1:1 in maschinelle Prüfregeln umwandeln.

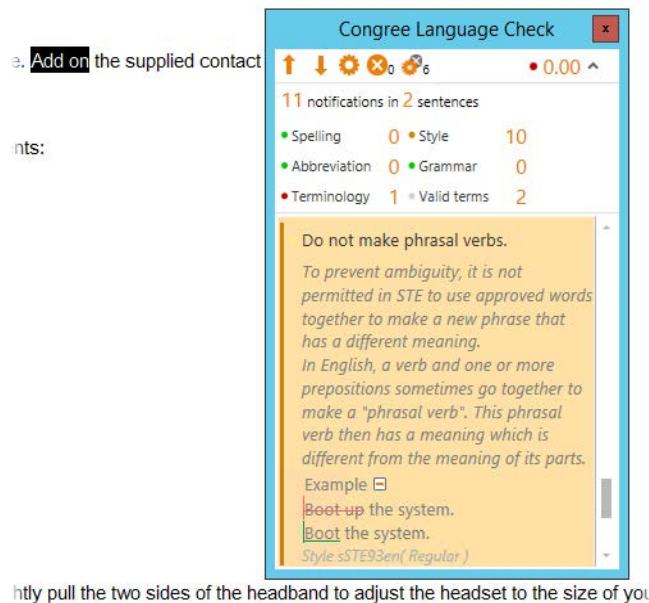
Als Beispiel für die maschinelle Prüfung, gezeigt anhand des Congree Authoring Servers, soll Regel **8.3 Check the use of parentheses.**³ dienen:



Die Klammersetzung wird maschinell ermittelt. Der Abgleich mit der Liste an zugelassenen Nutzungsarten muss jedoch der Autor selbst vollziehen.

Schreibvorgaben

Die STE-Schreibvorgaben umfassen vier Regeln, von denen nur eine maschinell prüfbar ist. Es handelt sich um Regel **9.3 When you use two words together, do not make phrasal verbs.**, deren Umsetzung z. B. so aussehen kann:



³ Langform s. STE-Issue 7, S. 1-7-5

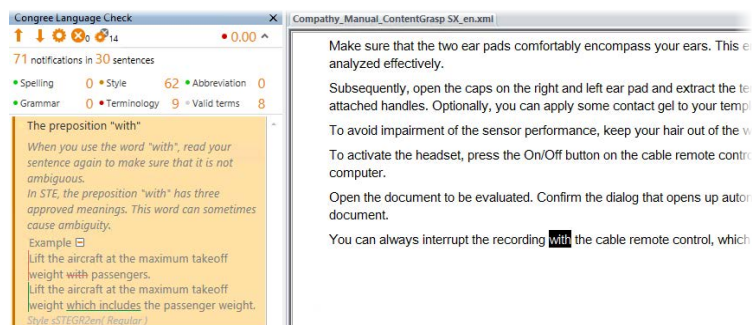
Die Umsetzung basiert darauf, dass die Software Verbphrasen erkennt. Dabei handelt es sich um feststehende Verbindungen zwischen einem Verb und mindestens einer Präposition. Diese Phrasen sind in der maschinellen Prüfung hinterlegt und können so gefunden und moniert werden.

Bei den nicht maschinell prüfbaren Regeln handelt es sich um:

- » 9.1 Use a different construction to write a sentence when a word-for-word-replacement is not sufficient.
- » 9.2 Use each approved word correctly.
- » 9.4 When you select terminology or wording, always use a consistent style.

Regel 9.1 bezieht sich auf das Ersetzen nicht erlaubter durch erlaubte Wörter aus dem STE-Wörterbuch. Wenn das Ersetzen von Wörtern die Satzbedeutung ändern würde, gibt die Regel vor, dass in solchen Fällen der ganze Satz umformuliert werden müsse. Die maschinelle Nicht-Prüfbarkeit liegt auch hier darin begründet, dass die Regel sich auf den Textinhalt bezieht. Software kann nicht erkennen, welche Bedeutung einem Satz innewohnt und wie sich diese durch ein verändertes Wort ändert. Hier sind Weltwissen und Einschätzungsgabe des Autors gefragt. Dasselbe gilt für Regel 9.2, die darauf abzielt, dass erlaubte Wörter korrekt verwendet werden. Bei Regel 9.4 handelt es sich um eine Regel, die sich, äquivalent zu 1.9 und 1.10, nicht auf den Texterstellungsprozess bezieht. Stattdessen beinhaltet sie Vorgaben zur Terminologiearbeit, bei der die Termkandidaten-Erfassung des Congree Authoring Servers Unterstützung bieten kann (s. oben).

Neben den vier geschilderten Regeln beinhaltet das Schreibpraktiken-Kapitel auch vier allgemeine Empfehlungen. Diese beziehen sich auf die Verwendung von bestimmten Konjunktionen, Präpositionen und Pronomen. Alle vier allgemeinen Empfehlungen sind maschinell prüfbar. Als Umsetzungsbeispiel anhand der Software Congree wird hier die Regel **GR-2 The preposition "with"** gezeigt:



Wie auch in den drei weiteren allgemeinen Empfehlungen geht es hier darum, Mehrdeutigkeiten zu vermeiden. Die maschinelle Umsetzung besteht darin, die potenziell mehrdeutige Präposition zu melden. Die Überprüfung, ob es sich wirklich um eine Mehrdeutigkeit handelt und, falls ja, die Behebung derselben obliegt dem Autor.

5.2 Analyse und Bewertung der maschinellen Umsetzbarkeit

Bei einer Bewertung der maschinellen Prüfbarkeit spielt die Wahl des verwendeten Prüfwerkzeugs eine essenzielle Rolle. Ein einfacher STE-Checker kann einfache Textparameter ermitteln, z. B. die Satzlänge, das Vorliegen von Nomen oder von Passivkonstruktionen. Wenn ein linguistisch basierter Checker verwendet wird, können weitaus komplexere sprachliche Phänomene erkannt werden:

- » Wortart
- » Ganze Satzstrukturen
- » Mögliche Ambiguitäten
- » Textstrukturabhängige Phänomene etc.

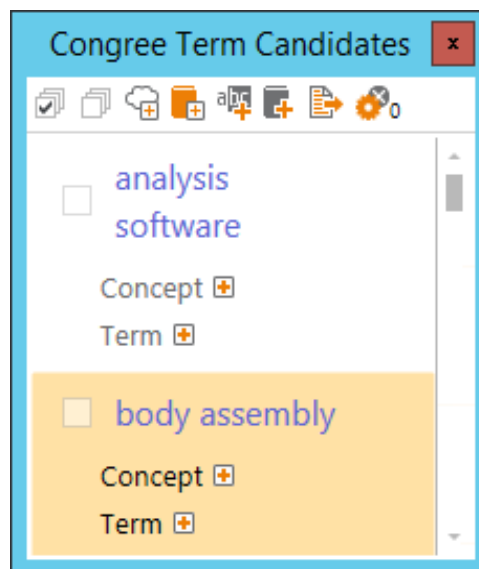
Unter Verwendung eines linguistischen Checkers ist die Mehrheit der STE-Regeln maschinell prüfbar. Mit dem Congree Authoring Server ergibt sich die folgende Abdeckung:

Kategorien	Anzahl umsetzbar	Anzahl nicht umsetzbar
Wörter	9 bzw. 12	2
Nomen-Cluster	2	1
Verben	7	-
Sätze	3	1
Informationsart – Vorgänge	5	-
Informationsart – Beschreibungen	2	4
Informationsart – Sicherheitshinweise	1	2
Zeichensetzung und Wortanzahl	7	-
Schreibvorgaben	1	3
Allgemeine Empfehlungen	4	-

Interessant sind neben der insgesamt sehr hohen Abdeckung jene Fälle, in den die Software doch an ihre Grenzen stößt – hier in Orange gekennzeichnet. Die Gründe sind jeweils ähnlich:

1. Die Regel bezieht sich nicht auf die Schreibpraxis, sondern z. B. auf Terminologiearbeit.
2. Oder: Die Regel bezieht sich auf die Inhaltsebene, die auch eine Software mit KI-Features Stand heute nicht ausreichend analysieren und bewerten kann.
3. Oder: Die Regel benötigt Welt- und Fachwissen seitens des Autors.

Punkt 1 zeigt auf, dass der Standard sich vermutlich nicht an ein homogenes Publikum richtet, sondern an verschiedene Lesergruppen. Aufgrund ihres Einsatzgebiets können Regeln, die sich auf Terminologiefestlegung beziehen, nicht komplett mit linguistisch basierten Schreibregeln geprüft werden. Dennoch bietet eine maschinelle Autorenunterstützung Features, die bei der Festlegung von Terminologie unterstützen. Der Congree Authoring Server z. B. verfügt über die Termkandidaten-Erfassung, mit der unternehmensspezifische Benennungen und Technical Names zur Aufnahme in die Terminologie vorgeschlagen werden können:



Die Punkte 2 und 3 machen deutlich, dass an vielen Stellen offensichtlich wird, dass eine STE-Prüfung den Autor nicht vollständig ablösen wird. Weltwissen und Expertenwissen müssen vom Menschen selbst kommen, die Maschine unterstützt nur, wo Regelfülle zu einer kognitiven Überlastung führen würde. Auch wenn das Feature zur Termkandidaten-Erfassung potenzielle Terme bzw. Technical Names gesammelt hat, wird die Expertise eines Menschen benötigt, um zu entscheiden, ob die Benennungen den STE-Regeln entsprechen. Ähnliches gilt für Regel **8.3 Check the use of parentheses.**, die zwar maschinell prüfbar ist, aber dennoch eine Konsultierung des STE-Standards durch den Autor erfordert.

6 Fazit

Der STE-Standard ist zum größten Teil maschinell prüfbar. Eine linguistisch basierte Autorenunterstützung wie der Congree Authoring Server deckt alle texterstellungrelevanten Regeln ab und kennt alle erlaubten und verbotenen Wörter des STE-Wörterbuchs.

Durch eine maschinelle Überprüfung ergibt sich eine Vielzahl von Mehrwerten. Software unterstützt den Autor, der sich durch die maschinelle Unterstützung nicht mehr alle Regeln und Wörter merken muss. Sie hilft durch das Aufzeigen der Regeln, dass diese dem Autor im Sinne von „Training on the Job“ bekanntgemacht werden. Und nicht zuletzt bietet eine linguistisch basierte Autorenunterstützung auch Hilfe beim Terminologieaufbau.

Allerdings ersetzt Software nicht das Welt- und Expertenwissen des Autors. Vielmehr stellt sie sich an dessen Seite und assistiert ihm, damit er seine kognitiven Ressourcen auf den Inhalt seiner Dokumente ausrichten kann.

Über uns:

Konsistenz, Verständlichkeit und übersetzungsgerechtes Schreiben sind die Schwerpunkte unserer Autorenunterstützung. Doch dies allein genügt uns nicht: Mit dem gebündelten Wissen aus Forschung und Praxis entwickeln wir Produkte, die auf den ersten Blick begeistern sollen. Diesen Anspruch verfolgen wir mit einem kompetenten Team, das weiß, worauf es bei anwenderfreundlicher Software ankommt.

**Congree Language
Technologies GmbH**
Im Stoeckmaedle 13
76307 Karlsbad
www.congree.com

