



## Echt gut der Text!

### Zur Messbarkeit von Textqualität und Textverständlichkeit

tekom - Jahrestagung  
09. November 2005



## Qualität - Aspekte

- Qualität ist ein relativer Begriff:
  - ...bezogen auf die Eignung zur Erfüllung festgelegter oder vorausgesetzte Erfordernisse
  - Ergebnis-/Produktqualität, Prozessqualität, Potenzialqualität,...
- Qualität technischer Dokumente:
  - gut strukturiert; sorgfältig gesetzt; edle Aufmachung; anschauliche Abbildungen; zielgruppengerecht; verkauft sich gut; genau die Information, die ich brauche; keine fachlichen Fehler; orthografisch und grammatikalisch korrekt; abwechslungsreich; normgerecht; ansprechend; spannend; gut formuliert; leicht verständlich; preisgünstig;....





## Am Anfang stand das Zählen: Lesbarkeitsformeln

- Empirische Lesbarkeitsforschung:  
Beginn in der 1. Hälfte des 20. Jahrhunderts (USA)
- Basis: objektiv feststellbare und  
quantitativ auswertbare  
Textmerkmale
- Parameter:
  - Wortlänge
  - Satzlänge
  - Wort-Gebräuchlichkeit
  - Type/Token-Verhältnis (Verhältnis  
„Anzahl unterschiedlicher Wörter“ zu  
„Anzahl der Wörter“)

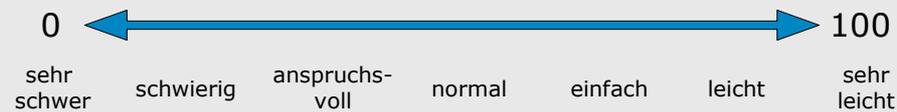


## Am weitesten verbreitet: „Reading Ease“-Formel von Flesch (1948)

- $RE = 206,835 - 0,846wl - 1,015sl$

wl = Anzahl der Silben pro 100 Wörtern

sl = durchschnittliche Anzahl der Wörter pro Satz





## Übertragung auf das Deutsche (Mihm 1974)

Reading Ease für deutsche Texte (entsprechender RE-Score für englische Texte)	Charakteristik	Typischer Text	Mittlere Wortlänge	Mittlere Satzlänge
-20 bis +10 (0-30)	sehr schwer	wissenschaftl. Abhandlung	über 2,20	über 30
+10 bis 30 30 bis 40	(30-50) schwierig (50-60) anspruchsvoll	Fachliteratur, Sachbuch, Roman (z. B. „Buddenbrooks“)	1,90 1,78	25 21
40 bis 50	(60-70) normal	Roman (z. B. „Stiller“)	1,70	17
50 bis 60	(70-80) einfach	Unterhaltungsliteratur (z. B. „Karl May“)	1,62	14
60 bis 70 70 bis 80	(80-90) leicht (90-100) sehr leicht	Heftchenroman, Comics	1,54 unter 1,45	11 unter 9



## Am kompliziertesten: Formel von Dick und Steiwer (1977)

$$\begin{aligned}
 \text{Verständlichkeitsindex} = & 235,96 - \left[ \ln \left( \frac{\text{Anzahl Buchstaben}}{\text{Anzahl Wörter}} + 1 \right) \right] \cdot 73,021 \\
 & - \left[ \ln \left( \frac{\text{Anzahl Wörter}}{\text{Anzahl Sätze}} + 1 \right) \right] \cdot 12,564 \\
 & - \left[ \ln \left( \frac{\text{Anzahl unterschiedlicher Wörter}}{\text{Anzahl Wörter}} + 1 \right) \right] \cdot 50,003
 \end{aligned}$$





## Online-Berechnung (z.B. [www.readability.info](http://www.readability.info))

**Readability report for Test2.doc**

```

readability grades:
Kincaid: 11.3
ARI: 11.7
Coleman-Liau: 14.4
Flesch Index: 46.3
Fog Index: 12.1
Lix: 47.0 = school year 8
SMOG-Grading: 13.3

sentence info:
1103 characters
275 words, average length 3.11 characters = 1.00 syllables
11 sentences, average length 28.1 words
30% (4) short sentences (at most 13 words)
7% (1) long sentences (at least 38 words)
7 paragraphs, average length 3.9 sentences
0% (0) questions
30% (4) passive sentences
longest sent 30 wds at post 2; shortest sent 4 wds at post 11

word usage:
verb types:
to be (2) auxiliary (0)
types 44 % of total
noun/adjective (11) pronoun 2 (4) preposition 13 (25)
conjunctions 3 (6)

sentence beginnings:
prepositional (1) locative/prepositional (0) article (4)
subordinating conjunction (0) coordination (0) preposition (4)

```

Web page readability score derived: 12079

Upload Microsoft Word document or enter a URL.

Average readability

<http://www.readability.info>

Flesch Index: 46.3

Echt gut der Text!

Martin Jung

09. November 2005 | Seite: 7

cognitas Gesellschaft für Technik Dokumentation GmbH  
Telefon: (089) 61001-154 | E-Mail: martin.jung@cognitas.de  
www.cognitas.de



## Berechnung in Microsoft Word

Lesbarkeitsstatistik			
Zahlen			
Zeichen	232	Kurze Sätze	7
Silben	75	Lange Sätze	0
Wörter	46	Einfache Sätze	1
Sätze	8	Lange Wörter	8
Absätze	0		
- Durchschnitt			
Silben pro Wort			1,63
Wörter pro Satz			5,75
Sätze pro Absatz			8,0
- Lesbarkeit			
Flesch Lesbarkeitsgrad (100 = sehr leicht)			63
Passiv (% der finiten Verbphrasen)			20
Satzkomplexität (100 = sehr komplex)			5
Wortschatzkomplexität (100 = sehr komplex)			37
		Hilfe	OK

„Die Lesbarkeitsstatistiken enthalten verschiedenartige Angaben, anhand derer der Lesbarkeitsgrad Ihres Dokuments beurteilt wird. Diese Statistik deutet lediglich auf Anzeichen für mögliche Probleme hin und ist nicht als direkte Handlungsanweisung zu verstehen.“

(Word-Hilfetext)

Echt gut der Text!

Martin Jung

09. November 2005 | Seite: 8

cognitas Gesellschaft für Technik Dokumentation GmbH  
Telefon: (089) 61001-154 | E-Mail: martin.jung@cognitas.de  
www.cognitas.de



## Lesbarkeitsformeln: Vorteile / Nachteile

- Vorteile:
  - objektiv messbar
  - einfach anwendbar
- Nachteile:
  - teilweise der Intuition widersprechende Ergebnisse
  - keinerlei Bezug auf Inhalte

Heute zwei Haus mit Badewanne und dann nicht. Dabei genauso dumm wie Stroh und überhaupt. Und überhaupt! Wenn sie nicht verwendet werden, dann aber hallo. Bitteschön. Danke. Die Textmerkmale sind auch nicht mehr das, was ich vergessen habe. Daher: Wer andern eine Grube gräbt, währt am längsten.

(Beispiel von Stefan Hejnk)

→ RE=64



## Hamburger Verständlichkeitsmodell

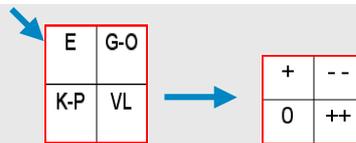
- Entwickelt von:
  - Inghard Langer
  - Friedemann Schulz von Thun
  - Reinhard Tausch
- Professoren am Psychologischen Institut der Uni Hamburg
- **der** Klassiker : „Sich verständlich ausdrücken“
- empirisch induktiver Ansatz  
200 Texte, mehr als 30 Themen mit insgesamt ca. 4500 Testpersonen
- „Schätzen statt Zählen“





## Dimensionen der Verständlichkeit (nach Schulz von Thun 2001)

Einfachheit	++ + 0 - --	Kompliziertheit
Gliederung- Ordnung	++ + 0 - --	Unübersichtlichkeit Zusammenhangs- losigkeit
Kürze- Prägnanz	++ + 0 - --	Weitschweifigkeit
Ver- lebendigung	++ + 0 - --	keine Verlebendigung



## Hamburger Verständlichkeitsmodell: Wichtige Ergebnisse

- Die Zahl der Kriterien lässt sich zu vier übergeordneten Dimensionen zusammenfassen.
- Einschätzungen unterschiedlicher Leser decken sich signifikant.
- Optimierte Texte werden besser verstanden.
- Alter, Intelligenz- und Bildungsgrad spielen keine Rolle: Alle profitieren gleichermaßen.
- Optimierungsprinzip lässt sich auf unterschiedlichste Themen übertragen.
- Textoptimierung ist lernbar.





## Hamburger Verständlichkeitsmodell: Kritik

- gilt als theoretisch nicht ausreichend fundiert
  - Beschreibungsvokabular aus der Alltagssprache, keine explizit definierten Prädikatoren
  - Subjektivität der Experteneinschätzungen
  - Kriterienausprägung nicht messbar
- Textimmanenz
  - Nicht berücksichtigt: Lesererwartungen, Lesesituationen, Textsorten die gleichen Prinzipien der Textoptimierung
  - **Aber:** Gerade die Textimmanenz des Ansatzes machte deutlich: Es lohnt sich, Texte zu verbessern.
- Aspekt „Visualisierungen“ nicht berücksichtigt



## Hamburger Verständlichkeitsmodell – gibt es sonst noch was?

- Verständlichkeitsdimensionen nach Groeben
  - Zentral: „kognitive Strukturierung“ (Subsumtion)
- Zahlreiche neue Ansätze aus der Linguistik und Kognitionspsychologie:
  - Zentral: Untersuchung des Verstehensprozess
  - Verstehen vergleichbar mit dem Zusammenbauen eines Regalmodul (Nickel)
- Diese neueren Ansätze sind „bei der Basis“ weitgehend unbekannt.





## Überprüfung der Textqualität

- Selbstprüfung
- Korrekturlesen / Lektorat
  - Redaktionsleitfäden/Checklisten/Nachschlagewerke/Ratgeber
- Experimentelle Überprüfung
  - Usability-Tests
  - Abfrage: Leser-Zufriedenheit
- Toolgestützte Qualitätsprüfung?



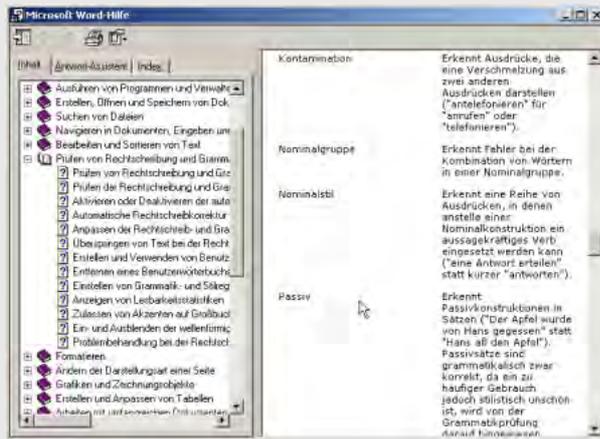
## Beispiele messbarer Kriterien (toolgestützt möglich)

- |  |   |
|--|---|
| <ul style="list-style-type: none"><li>▪ Wortwahl<ul style="list-style-type: none"><li>▪ Anzahl der Rechtschreibfehler</li><li>▪ Wortlänge</li><li>▪ Terminologische Konsistenz</li><li>▪ Anzahl von Fremdwörtern</li><li>▪ Anteil seltener Wörter</li><li>▪ Type/Token-Verhältnis</li><li>▪ Anzahl der Abkürzungen</li><li>▪ Anzahl der aktiven Verben</li></ul></li></ul> | <ul style="list-style-type: none"><li>▪ Satzbau<ul style="list-style-type: none"><li>▪ Anzahl der Interpunktionsfehler</li><li>▪ Anzahl der Grammatikfehler</li><li>▪ Anzahl der Pronominalisierungen</li><li>▪ Satzlänge</li><li>▪ Satzkomplexität (Anzahl der Nebensätze)</li><li>▪ Anzahl der Passivkonstruktionen</li><li>▪ Anzahl der Verneinungen</li><li>▪ Anzahl der Ellipsen</li></ul></li></ul> |
|--|---|





## Tools: Grammatik- und Stilprüfung in Microsoft Word



Viele Optionen,  
die meisten  
wenig  
überzeugend!



Echt gut der Text!

Martin Jung

09. November 2005 | Seite: 17

cognitas Gesellschaft für Technik Dokumentation GmbH  
Telefon: (089) 61001-154 | E-Mail: martin.jung@cognitas.de  
www.cognitas.de



## Tools: Grammatik- und Stilprüfung in Microsoft Word

### Diskussionsgruppe: Textverarbeitung Word 2003

Hallo, irgendwie will das bei mir mit der Grammatikprüfung in WORD nicht hinhauen. Ich habe unter den Optionen alles eingeschaltet, doch werden mir kaum Sachen grün unterstrichen. Ich kann ja sogar einen Satz wie folgenden eingeben, der kein Sinn macht: "Der die das gemacht hab ich." und Word nimmt das einfach so hin.



Echt gut der Text!

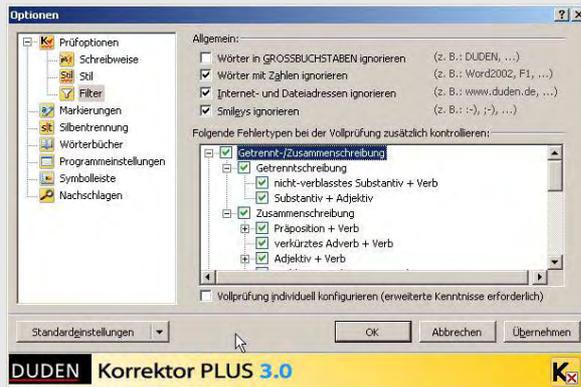
Martin Jung

09. November 2005 | Seite: 18

cognitas Gesellschaft für Technik Dokumentation GmbH  
Telefon: (089) 61001-154 | E-Mail: martin.jung@cognitas.de  
www.cognitas.de



## DUDEN Korrektor



Wesentlich besser  
als die Word-  
Prüfung!



## Eine andere Liga: Multilint, CLAT, ...

- hochentwickelte Linguistic Engines
- ausgereifte Logik der Regelanwendung (Disambiguierung/Fehlererkennung)
- Ausrichtung auf die Anforderungen des Kunden
- Integration in die Prozesse





## Ableitung von Qualitäts-Kennzahlen möglich?

- Gewichtung innerhalb einer Merkmalsklasse  
Verstoß ist nicht gleich Verstoß  
*ein Öl verschmierter Lappen*  
*Ein ölverschmierter Lappen*  
*Wegen der Störung mussten alle Wagen diesen Typs zurück ins Depot.*  
*Wegen der Störung mussten alle Wagen diesem Typs zurück ins Depot.*
- Gewichtung der Merkmalsklassen untereinander:
  - Was ist „schlimmer“:  
2 Passivkonstruktionen oder ein 3 Pronominalisierungen?



## Grenzen der Messbarkeit: Graduelle Phänomene

- Viele mehrdeutige Konstruktionen treten ausgesprochen häufig auf, führen aber bei weitem nicht in allen Fällen zu Missverständnissen.
- Nominalkomposita:  
*Eisenbohrer, Diamantbohrer, Holzbohrer*
- Attribute  
*Fehler vom Fehlergrad FATAL, die zum Systemabsturz führen, ...*
- Gradpartikel  
*Ich trinke nur italienischen Rotwein.*





## Fazit

- Die Leser „fühlen“ den Unterschied zwischen guten und schlechten Texten – und diese Gefühle wiederum halten einer empirischen Überprüfung stand.
- Aus der Verständlichkeitsforschung kommen keine nennenswerten Impulse für die Praxis.
- Einige Textmerkmale können wir messen, wissen aber nicht, wie wir die Ergebnisse zu interpretieren sind.
- Tools könnten bei der Qualitätssicherung unterstützen, die Qualitätsstandards aber nicht selbst formulieren.
- **Ungesteuerte** Einführung solcher Checker-Tools kostet Zeit, Nerven und Geld und führt zu keiner Verbesserung.



## Literatur / Links

- [www.tekom.de](http://www.tekom.de) Stichwörter:  
„Verständlichkeit“; „Textanalyse“; „Textoptimierung“.
- Ballstaed, Steffen: Texte optimieren.  
technische kommunikation 2/2000.
- Fottner-Top, Claudia: Machineller Lektor.  
technische kommunikation 1/2002.
- Gottschling, Stefan: Die Texterfibel für das Direktmarketing.  
(kostenlos beziehbar über [www.textakademie.de](http://www.textakademie.de))
- Hennig, Jörg; Tjarks-Sobhani, Marita (Hrsg.):  
Verständlichkeit und Nutzungsfreundlichkeit in der  
technischen Dokumentation. Lübeck 1999.  
(tekom - Schriften zur technischen Dokumentation Bd.1)





## Literatur / Links (2)

- Hochhaus, Stephan: Der verständliche Text. Perspektiven für die Textoptimierung. Masterarbeit Ruhruniversität Bochum 2004 ([www.yauh.de/files/masterarbeit.pdf](http://www.yauh.de/files/masterarbeit.pdf))
- Krings, Hans: Vertrauen ist gut, Testen ist besser. technische kommunikation 5 / 2003
- Langer, Inghard; Schulz von Thun, Friedemann; Tausch, Reinhard: Sich verständlich ausdrücken. München 2002 (6. Auflage)
- Nickl, Markus: Verständlich Schreiben – eine Einführung. Unter [www.direktportal.de](http://www.direktportal.de) (2005).



## Literatur / Links (3)

- Reuther, Ursula: Linguistisch modelliertes Wissen für die technische Dokumentation. [www.doculine.com/news/2002/0602/sprachtechnologie.shtml](http://www.doculine.com/news/2002/0602/sprachtechnologie.shtml) 2002.
- Roller, Marianne: Verfahren der Textanalyse (I). tekomp nachrichten 1 / 1980.
- Roller, Marianne: Verfahren der Textanalyse (IV). tekomp nachrichten 2 / 1981.
- Schmidt-Wigger, Antje: Grammar and Style Checking for German. [www.iai.uni-sb.de/iaide/de/pub.htm](http://www.iai.uni-sb.de/iaide/de/pub.htm) 1998.
- Schulz, Matthias: Kampf der Blähung. Wie man Anleitungen verständlich schreibt. Aalen 2004.

