

Überlegungen zur empirischen Validierung von Sprichwörter-Dummies

Peter Grzybek & Christoph Chlosta

Im Rahmen empirisch experimenteller Erhebungen zur Bekanntheit von Sprichwörtern richtet sich eine zentrale Frage auf geeignete Erhebungsmethoden. Dabei muss der Berücksichtigung sprachlicher Varianten und Variationen besondere Aufmerksamkeit geschenkt werden, weshalb die Methode der Teiltex-Präsentation von Sprichwörtern zunehmend favorisiert wird. Dennoch stellen nach wie vor auch Ganztext-Präsentationen ein in der Sprichwortforschung wichtiges Instrument dar. Die im Hintergrund stehende Untersuchung stellt einen Beitrag zur Prüfung der Validität des Ganztext-Verfahrens dar: In einem Experimental-Korpus von ca. 2000 deutschen Sprichwörtern wurden zusätzlich zu den authentischen Sprichwörtern auch Sprichwort-Dummies präsentiert und auf ihre Bekanntheit hin getestet. Die vorliegende Studie untersucht die sprachliche Struktur dieser Sprichwort-Dummies im Vergleich zu den authentischen Sprichwörtern unter besonderer Berücksichtigung des Grades ihrer Lesbarkeit/Verständlichkeit.

Im Rahmen des weiten Bereichs von Sprichwortforschungen hat sich in den vergangenen Jahren die Empirische Parömiologie als ein wichtiger Teilbereich der modernen Parömiologie etabliert. Damit haben neben dominant philologischen Vorgangsweisen auch experimentelle und testbasierte Verfahren Eingang in die Parömiologie gefunden. Im Prinzip lassen sich in der Empirischen Parömiologie zwei Fokussierungen unterscheiden, die letztlich beide stark miteinander zusammenhängen. Vereinfacht gesagt, richtet sich auf der einen Seite das empirische Interesse auf das sprichwörtliche Material, auf der anderen Seite auf diejenigen, die mit diesem Material umgehen, d.h. die SprichwortbenutzerInnen. Diese grundsätzliche Aufteilung in zwei Ausrichtungen der empirischen Sprichwortforschung entspricht dem, was wiederholt als *informantenbezogene* vs. *textbezogene* Empirie in der Parömiologie bezeichnet worden ist.

Ausgangspunkt der Empirischen Parömiologie war ursprünglich die Frage nach sogenannten Sprichwörter-Minima; sie zielt in letzter Konsequenz auf die Erhebung derjenigen Menge von Sprichwörtern ab, die allen (oder fast allen) Angehörigen einer bestimmten Kultur oder eines bestimmten Soziums bekannt sind. Es liegt auf der Hand, dass diese Frage sich nur teilweise mit der Frage nach der Dokumentation von Sprichwörtern in (der Geschichte) einer Kultur deckt; allein der positivistische Sammeleifer des 19. Jahrhunderts, der die jeweiligen nationalen Sprichwortgüter in möglichst umfassenden Maße „dokumentieren“ wollte (und sich dabei nicht selten durch ausgeprägte Erfinder-

und Übersetzertätigkeit auszeichnete), hat einen noch heute kaum zu hebenden Nebelschleier über die Teilmenge der in den jeweiligen Kulturen tatsächlich zum aktiven Sprichwortbestand gehörenden Sprichwörter geworfen. Damit ist ein Fragenkomplex ins Zentrum der Diskussion gerückt, der sich in maximal komprimierter Form wie folgt zusammenfassen lässt:

Wer innerhalb einer gegebenen Kultur kennt welche Sprichwörter in welcher Form und wovon hängt diese Kenntnis ab?

Eine Antwort auf diese Frage(n) ist nicht nur von Interesse für Phraseologie und Parömiologie, sondern z.B. auch für spezifisch linguistische und/oder poetologische Untersuchungen von Sprichwörtern, für Fragen des Fremdsprachenunterrichts, für experimentelle Untersuchungen zur Sprichwortkenntnis bestimmter Zielgruppen (z.B. Kinder, Heranwachsende, bestimmte soziale Gruppierungen, Patienten mit psychisch oder neuropsychologisch bedingten Störungen, etc.), u.a.m.

Traditionelle Verfahren mit dem Ziel, auf textbasierter Basis Aufschluss über die Bekanntheit von Sprichwörtern zu erhalten, stoßen zwangsläufig an die Grenze ihrer Leistungsfähigkeit, nicht zuletzt deshalb, weil sie bei genauerer Betrachtung das Konzept der Bekanntheit auf die eine oder andere Art und Weise mit der Vorkommenshäufigkeit gleichsetzen (nicht selten, um die aufwendige informantenbezogene Untersuchung zu umgehen). Zwar gibt es, wie neueste Forschungen zeigen, systematische Zusammenhänge zwischen der Bekanntheit und der Vorkommenshäufigkeit von Sprichwörtern, doch ist dieses Verhältnis weitaus vielschichtiger als bisher vermutet (Grzybek 2008; Grzybek & Chlosta 2009), so dass jegliche Art von Frequenzanalyse Untersuchungen zur Bekanntheit nicht ersetzen kann.

Will man nun wissen, welche Sprichwörter allgemein bekannter Bestandteil einer Kultur sind, so bieten sich experimentelle Untersuchungen an. In der Vergangenheit sind in dieser Hinsicht verschiedene Verfahren angewendet worden; die mit ihnen verbundenen Vor- und Nachteile können hier nicht im Detail erörtert werden (vgl. Grzybek, Chlosta & Grotjahn 2009):

1. Das **spontane Nennen bzw. Aufschreiben** von Sprichwörtern zeigt in der Regel eine Menge von nicht mehr als ca. 30-50 Sprichwörtern je Proband; bei diesen muss es sich nicht, um die tatsächlich bekanntesten Sprichwörter handeln, vielmehr dürfte es sich zu einem großen Teil um (auch in sprachlicher Hinsicht) prototypische Einheiten handeln.
2. Das **Ankreuzen auf einer normierten Skala** – z.B. von 0 („unbekannt“) bis 7 („hochgradig bekannt“) ist ein vor allem im Rahmen der anglo-amerikanischen Psycholinguistik beliebtes Verfahren; ein Problem stellt hierbei die intersubjektiv nur schwer zu vereinheitlichende Skalierung bzw. Anwendung der vorgegebenen Skala dar.
3. **Ganztext-Präsentationen** mit der dichotomischen Frage „Bekannt“ vs. „Unbekannt“. Als ein wesentliches Manko dieser Methode (und z.T. auch der

vorangegangenen) ist u.a. die mögliche Gefahr von Fehleinschätzungen herausgestellt worden.

4. Das Verfahren von **Teilttext-Präsentationen** besteht darin, den befragten Personen nur jeweils den Anfang eines Sprichworts zu präsentieren, den es zu vervollständigen gilt. In Einklang mit der Annahme, dass ein Sprichwort ein sprachliches Klischee im Sinne Permjakovs (1970) ist, das man entweder als Ganzes kennt oder gar nicht, hat das Verfahren eine Reihe von Vorteilen; zu nennen sind hier besonders, dass man Varianten und Variationen eines Sprichworts erheben kann.

Letztendlich wird in der Empirischen Parömiologie das Verfahren der Teilttext-Präsentation favorisiert. Was freilich auf den ersten Blick von der methodischen Einfachheit einleuchtend sein mag, bedeutet aber nicht nur für die Befragten einen vergleichsweise hohen Arbeitsaufwand, sondern auch für die Auswertungen. Nicht zuletzt aus diesem Grunde gibt es bislang zwar eine ganze Reihe von Untersuchungen zu verschiedenen Sprachen mit diesem Verfahren, doch wurde für keine Sprache wirklich das gesamte mit diesem Vorgehen verbundene Programm vollständig umgesetzt. Weit häufiger finden sich die anderen vorgestellten Verfahren, was letztlich aber vor allem an der Ökonomie liegt.

Ohne den Status der verschiedenen Untersuchungen hier darzustellen, wollen wir uns auf das Deutsche beschränken, für das in langjährigen Untersuchungen ein „Experimental-Korpus“ von ca. 2000 Sprichwörtern erarbeitet wurde.¹ Aufgenommen wurden nach entsprechenden Studien nicht nur Sprichwörter aus traditionellen Sammlungen, sondern auch weitere, die von Befragten zusätzlich genannt wurden, die als Ergebnis umfangreicher Medienanalysen hervorgingen u.a.m. Diese wurden in Zwischenabfragen vorgetestet, um ihren potentiellen Bekanntheitsgrad sicherzustellen. Da die Sprichwörter aus den verschiedenen Quellen somit nicht einheitlich empirisch validiert wurden, steht derzeit eine vollständige und einheitlich basierte Validierung des Experimental-Korpus mittels Ganztextpräsentation an.

Aus Gründen der Absicherung und Konsolidierung, aber auch um die Zuverlässigkeit der Ganztextpräsentation empirisch zu überprüfen, wurden in das Korpus der 2000 Sprichwörter auch 15 Dummy-Sprichwörter eingebaut und nach dem Zufallszahlenprinzip platziert. Bei diesen Dummy-Sprichwörtern handelt es sich um erfundene, d.h. ausgedachte Texte, die auf intuitive Art und Weise sprachlich so konstruiert wurden, dass sie möglichst echt klingen. Denn es gibt ja ungeachtet der Tatsache, dass man als MuttersprachlerIn offenbar ein „Gefühl“ dafür hat, wie ein Sprichwort klingt, im Hinblick auf die sprachliche Struktur des Sprichworts praktisch keine obligatorischen Merkmale, wohl aber eine Reihe von fakultativen Merkmalen; dabei handelt es sich in der Regel um

¹ Das Experimental-Korpus soll letztlich alle potenziell allgemein bekannten Sprichwörter der deutschen Sprache beinhalten, damit diese dann mittels Teilttextpräsentation getestet werden können.

intuitive, subjektiv basierte Annahmen zur sprachlichen Struktur wie etwa Tendenz zur Kürze, spezifische Rhythmik, Metaphorizität u.a.m (vgl. Chlosta & Ostermann 2003) – Faktoren, deren Präsenz und Relevanz nur in seltenen Fällen einer systematischen empirischen Überprüfung unterzogen und in der Regel eher nur in Form symptomatischer Beispiele belegt werden. Tabelle 1 enthält die 15 Dummy-Sprichwörter.

- 1 Bier ist nicht zum Waschen da.
- 2 Lieber ein Brot als ein Brötchen.
- 3 Nach der Arbeit schmeckt Wasser wie Wein.
- 4 Flinke Zunge weiß nicht, was sie spricht.
- 5 Einen rostigen Nagel soll man nicht hauen.
- 6 Wein und Essig stammen von derselben Rebe.
- 7 Die Schnecke gewinnt das Rennen durch Beharrlichkeit.
- 8 Aus einem Fisch kann man keinen Vogel machen.
- 9 Je dicker das Seil, desto stärker der Knoten.
- 10 Der Mensch hat zwei Arme, aber nur einen Kopf.
- 11 Wie ihr die Blinden behandelt, so werdet ihr behandelt.
- 12 Im Wald ist die Ameise ein König, im Haus Ungeziefer.
- 13 Für den Bäcker ist es gestern, für den Studenten morgen.
- 14 Auch die schönste Rose ist Unkraut, wenn sie beim Nachbarn wächst.
- 15 Ein Freund schärft Dir das Beil, der Bruder fällt Dir den Baum.

Tabelle 1: Liste der 15 Dummy-Sprichwörter

In Anbetracht dieser Tatsache liegt es nahe, die sprachliche Struktur dieser Dummy-Sprichwörter genauer unter die linguistische Lupe zu nehmen und sie dabei unter Bezugnahme auf Verfahren der Quantitativen Linguistik zu reflektieren. Die methodischen Ansätze der Quantitativen Linguistik haben die Sprichwortforschung bislang nur peripher berührt; erst seit wenigen Jahren gibt es Bemühungen, allgemeine Regularitäten z.B. der Wortfrequenz, der Wort- oder Satzlänge von Sprichwörtern sowie deren Vorkommenshäufigkeiten in Sprichwortkorpora im Sinne allgemeiner Gesetzmäßigkeiten zu formulieren (zur Satzlänge deutscher Sprichwörter etwa vgl. Grzybek & Schlatter 2002).

Zum Vergleich ziehen wir die Sammlung *Deutsche Sprichwörter* von Karl Simrock (1846) mit ihren 12981 sprichwörtlichen Belegen heran. Die durchschnittliche (in der Anzahl der Silben pro Wort gemessene) Wortlänge der Sprichwörter des Dummy-Korpus beträgt $\bar{x} = 1.46$ bei einer Standardabweichung von $s = 0.21$; der Vergleich mit den 12981 Sprichwörtern aus Simrock stellt sich bei einer identischen mittleren Wortlänge von $\bar{x} = 1.46$ ($s = 0.66$) der authentischen Sprichwörter als nicht signifikant heraus ($p > 0.99$).

Die praktisch identische Struktur der Wortlängen in den authentischen und den Dummy-Sprichwörtern drückt sich auch darin aus, dass die Dummy-Items

ebenso wie die authentischen Sprichwörter ein und demselben Verteilungsmodell folgen; hierbei handelt es sich freilich nicht um die negative Binomialverteilung, die Grotjahn (1982) zunächst als allgemeines Modell der Häufigkeitsverteilung von Wortlängen ins Spiel brachte, sondern um die Poisson-Verteilung, gegen welche die erstgenannte für $k \rightarrow \infty, q \rightarrow 0, kq \rightarrow a$ konvergiert²; für beide Sprichworttypen ergibt sich sogar ein und derselbe Parameterwert von $a = 0.47$. Tabelle 2a/b führt die beobachteten und theoretischen Häufigkeiten auf; die Anpassungsgüte ist in Form des als X^2/N berechneten Determinationskoeffizienten C angegeben, der für $C < 0.02$ eine gute, für $C < 0.01$ eine sehr gute Anpassung indiziert.

Tab. 2a			Tab 2b		
x	f(x)	NP(x)	x	f(x)	NP(x)
1	62719	63793,27	1	77	77,68
2	32288	29792,01	2	39	36,33
3	5473	6956,56	3	6	8,50
4	1166	1082,93	4	2	1,50
5	106	126,43			
6	12	12,79			
$\alpha = 0,47; C = 0,0054$			$\alpha = 0,47; C = 0,0089$		

Tabelle 2a/b

2 Mittlerweile geht man, nicht zuletzt unter dem Einfluss der Arbeiten von Grotjahn (vgl. Grotjahn & Altmann 1993; Wimmer, Köhler, Grotjahn & Altmann 1994), von einem System von Verteilungen aus, innerhalb dessen die negative Binomialverteilung einen der Spezialfälle darstellt (zu einem historischen Überblick siehe Grzybek 2006).

Die Ergebnisse sind graphisch in Abbildung 1a und 1b dargestellt:

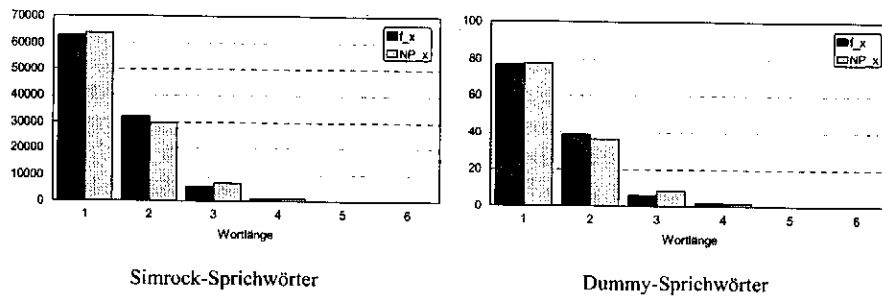


Abbildung 1: Häufigkeit von Wortlängen in authentischen und Dummy-Sprichwörtern (Anpassung der Poisson-Verteilung)

Auch die durchschnittlichen Satzlängen³ weisen keine signifikanten Unterschiede auf: Bei den authentischen Sprichwörtern aus Simrock beträgt die mittlere, in der Anzahl der Worte pro Satz gemessene Satzlänge $\bar{x} = 7.84$ ($s = 3.40$); im Vergleich dazu beträgt die mittlere Satzlänge der Dummy-Sprichwörter $\bar{x} = 8.27$ ($s = 1.83$). Berechnet man die Satzlänge in der Anzahl der Silben pro Satz, so beläuft sich der Durchschnitt der Simrock-Sprichwörter auf $\bar{x} = 11.46$ ($s = 4.85$), derjenige der Dummies auf $\bar{x} = 12.00$ ($s = 2.73$). Beide Unterschiede erweisen sich als nicht signifikant ($z = 0.91$, $p = 0.37$ bzw. $z = 0.76$, $p = 0.44$).

Aufgrund der Analyse von Wort- und Satzlänge können wir also festhalten, dass sich die Sprichwörter unseres Dummy-Korpus „normgerecht“ verhalten und sprachlich nicht von den authentischen Sprichwörtern abweichen.

Aufgrund von Textuntersuchungen im Kontext der Quantitativen Linguistik ist bekannt, dass Wort- und Satzlänge in einem spezifischen Verhältnis zueinander stehen, und zwar sowohl intra-textuell vor dem Hintergrund des sog. Menzerath-Gesetzes als auch inter-textuell in Form des sog. Arens-Gesetzes (d.h. in Form eines Mittelwert-Vektors für verschiedene Texte) – zur Relevanz dieser Unterscheidungen für das Verhältnis von Satz- und Wortlänge siehe: Grzybek & Stadlober (2007), Grzybek, Stadlober & Kelih (2007), sowie Grzybek, Kelih & Stadlober (2008). Ebenso ist bekannt, dass diese Abhängigkeit nicht nur textsorten-spezifisch ist, sondern sich gerade in Sprichwörtern vollkommen anders als in „üblichen“ Fließtexten darstellt; dies hängt vermutlich damit zusammen, dass bei Sprichwörtern die eigentliche *Text-Ebene* insofern

³ Für die hiesigen Analysen wurde ein Sprichwort jeweils als ein Satz gewertet; bei den Texten aus der Simrock-Sammlung ist das in ca. 4% der Fälle nicht unproblematisch, wenn ein Sprichwort aus mehr als einem „Satz“ besteht, wobei freilich einerseits Fragen der Satzdefinition, andererseits Fragen der Interpunktion eine wesentliche Rolle spielen.

fehlt, als jeder einzelne sprichwörtliche Satz einen in sich geschlossenen Text repräsentiert. Vor diesem Hintergrund liegt es nahe, neben der Ermittlung der Satz- und Wortlänge als quasi-isolierter Einheiten (Best 2006) einen weiteren Vergleich zwischen den authentischen Sprichwörtern und den Dummy-Items anzustellen, der sich auf die Textverständlichkeit beider Einheiten bezieht.

Textverständlichkeit ist ein Konzept, das ungeachtet prinzipiell verschiedener Ausgangspunkte in der Forschung der vergangenen Jahre in wesentlichen bestimmten Teilbereichen auch unter dem Stichwort ‚Lesbarkeit‘ abgehandelt wurde: Lesbarkeit, Verstehbarkeit bzw. Verständlichkeit und Schwierigkeit eines Textes sind somit Konzepte, die in der Forschung zunehmend zueinander in Bezug gesetzt, zum Teil sogar miteinander identifiziert worden sind. In der Tat haben einschlägige Arbeiten aufgrund sich überschneidender Ansätze und Methoden in wesentlichen Teilbereichen miteinander kompatible Ergebnisse gezeigt – insbesondere dann, wenn empirische, textanalytische Verfahren zur Bestimmung der sprachlichen Komplexität eines Textes angewendet wurden, die im Weiteren dann als Schwierigkeitsindex interpretiert und als Erklärung für mangelnde Lesbarkeit/Verständlichkeit herangezogen wurden.

In der Lesbarkeitsforschung wurden zahlreiche Verfahren entwickelt, wobei man heutzutage von mehr als 200 Faktoren und auf diesen aufbauenden Formeln ausgeht, um die Komplexität bzw. Schwierigkeit von Texten zu beschreiben. Den Ausgangspunkt dieser Forschungen stellen amerikanische Untersuchungen der 20er Jahre dar, in denen Verfahren zur Bestimmung der *lexikalischen Last* eines Textes diskutiert wurden; entsprechende Analysen hatten gezeigt, dass die Verständlichkeit eines Textes eng mit dem Vorkommen von häufig gebrauchten bzw. allgemein bekannten Wörtern in Beziehung steht.⁴ So schlugen Lively & Pressey (1923) als erste ein Verfahren vor, durch Bezugnahme auf Wortfrequenzlisten die Vokabellast eines Textes zu berechnen. Nicht zu Unrecht sahen sie in ihrer Arbeit den Grundstein gelegt für den Nachweis der Effizienz statistischer Methoden zur Vorhersage von Textschwierigkeit. Interessant ist, dass sie in den Fällen, wo die lexikalische Schwierigkeit niedriger als erwartet war, auf den (von ihnen freilich nicht näher untersuchten) Faktor der syntaktischen Komplexität verwiesen, der offenbar den subjektiven Eindruck der Textkomplexität bedingte. Damit stand bereits ganz am Anfang der Forschung die Vermutung, dass Schwierigkeit eines Textes kein eindimensionaler, sondern vielschichtig verwobener Komplex ist, dessen Faktoren in den folgenden Jahren immer detaillierter unterschieden wurden. Als erste untersuchten Vogel & Washburne (1928) strukturelle Texteneigenschaften und setzten diese zu empirisch erhobenen Einstufungen von Textschwierigkeit in Bezug; dabei legten sie

⁴ Heute wissen wir genauer, dass Bekanntheit und Frequenz zwei verschiedene, wenn auch eng miteinander zusammenhängende Kategorien sind, sowohl auf der Ebene lexikalischer (Köhler & Rapp 2007) als auch sprichwörtlicher Einheiten (Grzybek 2008; Grzybek & Chlosta 2008).

sprachliche Faktoren wie z.B. Satztyp, Wortschwierigkeit, Satzlänge u.a. zugrunde. Da eine Reihe dieser Faktoren statistisch signifikant miteinander korrelierten, reduzierten sie aufgrund der multiplen Korrelation die Gesamtanzahl der eingangs untersuchten Variablen wieder und integrierten diese in einer ersten Regressionsgleichung. Dieses Vorgehen sollte nachhaltig die Weichen für die Lesbarkeitsforschung der folgenden Jahrzehnte stellen, insofern es einerseits um die Erhebung möglicher Einflussgrößen, andererseits um deren Reduktion auf eine möglichst überschaubare Anzahl und die Ableitung handhabbarer, praxistauglicher Berechnungsverfahren gehen sollte.

Von nachhaltigem Einfluss sollten sich in dieser Hinsicht die Forschungen des Wiener Exil-Juristen Rudolf Flesch erweisen, der in den 40er Jahren aufgrund umfangreicher Forschungen verschiedene Lesbarkeitsformeln entwickelte und popularisierte; die heute wohl bekannteste ist die Berechnung der sog. Reading Ease (*RE*), die mit zwei Faktoren, Wortlänge (*WOL*) und Satzlänge (*SAL*), auskommt:

$$RE = 206.835 - (1.015 * SAL) - (84.6 * WOL) \quad (1.1)$$

Flesch (1948) hatte diese Formel auf der Basis von Tests zum Textverstehen abgeleitet; deren Ergebnisse setzte er zu verschiedenen sprachlichen Indikatoren in Beziehung, die er schließlich auf *WOL* und *SAL* reduzierte. Nach Flesch, der die Formel (1.1) am amerikanischen Englisch erarbeitete, sollten die Werte von *RE* in der Regel zwischen 0 und 100 liegen, wobei ein kleiner Wert für *RE* als Indiz für maximal schwierige, ein großer Wert für leichte Texte anzusehen ist. Diese Formel wurde mittlerweile für viele Sprachen durch unterschiedliche Gewichtungen von *WOL* und *SAL* adaptiert (vgl. Rabin 1988); für das Deutsche dürfte die folgende Berechnung der Textschwierigkeit (*TS*) von Amstad (1978) die bekannteste sein:

$$TS = 180 - SAL - (58.5 * WOL) \quad (1.2)$$

Gegenüber diesen – und vielen anderen – Lesbarkeits- und/oder Verstehbarkeitsformeln ist immer wieder Skepsis vorgebracht worden (vgl. Grotjahn 2000). Unter anderem wurde eingewendet, dass auf diese Art und Weise die Komplexität von Textschwierigkeit bestenfalls partiell erfasst werden könne, dass die Formeln theorielos und rein pragmatisch definiert seien, dass ihre Verwendbarkeit aufgrund der jeweiligen Sprachspezifik eingeschränkt sei, u.v.a.m. Bei aller Berechtigung verschiedener Kritikpunkte muss allerdings auch gesagt werden, dass neueste Forschungen, die das Konzept der Textschwierigkeit mit Ansätzen aus der Quantitativen Linguistik verbinden, einige der eingewendeten kritischen Punkte relativieren und damit auch das Konzept der Textschwierigkeit in neuem Licht erscheinen lassen. Auf eine ausführliche Diskussion muss an dieser Stelle jedoch verzichtet werden.

Machen wir stattdessen einen Sprung in die aktuelle Praxis und Anwendung von „Lesbarkeitsformeln“ und unterziehen die obigen Dummy-Items einem Verständlichkeitstest, wie er im Internet unter Bezugnahme auf die obigen Formeln als freie Anwendung angeboten wird, so etwa bei <http://www.leichtlesbar.ch> (im Folgenden: LL) und <http://www.benutzerfreund.de/itext> (weiterhin: BF). Im Ergebnis zeigen sich einige Überraschungen, insofern die Basisberechnungen von Wortanzahl, Silbenzahl und Satzlänge nicht nur zwischen beiden Anwendungen deutlich voneinander abweichen, sondern zudem beide einer Überprüfung nicht standhalten.⁵

Tabelle 3 führt die tatsächlichen Längen ebenso wie die von beiden Applikationen ausgegebenen Werte an: *SAL1* ist die in der Anzahl der Worte, *SAL2* die in der Anzahl der Silben pro Satz berechnete Satzlänge, *WOL* ist die in der Anzahl der Silben pro Wort berechnete Wortlänge. In Ergänzung dazu enthält Tabelle 3 die sich aufgrund der Formel (1.1) und (1.2) ergebenden Flesch-Werte (*F*) und Amstad-Werte (*A*); zusätzlich angeführt ist der weniger bekannte Schwierigkeitsindex *T*, der von Tuldava (1993, 1994) eingeführt wurde; dieser berechnet sich als einfaches Produkt aus der Wortlänge und der logarithmierten Satzlänge:

$$T = WOL * LN(SAL) \quad (1.3)$$

Zwar ist auch dieser Index nicht theoretisch abgeleitet, hat allerdings eine Reihe potentieller Vorteile (insbesondere die Tatsache, dass er ohne sprachspezifische Konstanten auskommt), was ihn für weiterführende Überlegungen im Hinblick auf eine in der Quantitativen Linguistik fundierte theoretische Begründung prädestiniert.

⁵ Unterschiede ergeben sich nur zum Teil aufgrund der unterschiedlichen Behandlung des stimmten End-E in den beiden Applikationen; dies allein erklärt allerdings die Abweichungen und Fehler nicht, wobei exakte Angaben zur Silbenzählung entweder gar nicht angesprochen sind (<http://www.benutzerfreund.de/itext>, 17.4.2009) oder der entsprechende Link ins Leere führt (<http://www.leichtlesbar.ch/html/faq+9.html>, 17.4.2009).

Nr.	SW- SAL_1 SAL_2 WOL			F	A	T	SAL_1 SAL_2 WOL			SAL_1 SAL_2 WOL		
							BF			LL		
1	6	7	1,1667	102,05	105,75	2,09	7	9	1,2857	6	7	1,1667
2	6	8	1,3333	87,95	96,00	2,39	7	11	1,5714	6	8	1,3333
3	7	9	1,2857	90,96	97,79	2,50	8	11	1,3750	7	9	1,2857
4	7	9	1,2857	90,96	97,79	2,50	8	11	1,3750	7	7	1,0000
5	7	12	1,7143	54,70	72,71	3,34	8	13	1,6250	7	11	1,5714
6	7	12	1,7143	54,70	72,71	3,34	8	13	1,6250	7	11	1,5714
7	7	13	1,8571	42,62	64,36	3,61	8	14	1,7500	7	12	1,7143
8	8	12	1,5000	71,82	84,25	3,12	9	14	1,5556	8	12	1,5000
9	8	12	1,5000	71,82	84,25	3,12	9	12	1,3333	8	12	1,5000
10	9	12	1,3333	84,90	93,00	2,93	10	14	1,4000	9	11	1,2222
11	9	15	1,6667	56,70	73,50	3,66	10	18	1,8000	9	15	1,6667
12	10	16	1,6000	61,33	76,40	3,68	11	16	1,4545	10	15	1,5000
13	10	15	1,5000	69,79	82,25	3,45	11	17	1,5455	10	15	1,5000
14	11	15	1,3636	80,31	89,23	3,27	12	19	1,5833	11	13	1,1818
15	12	13	1,0833	103,01	104,63	2,69	13	16	1,2308	12	13	1,0833
Σ	124	180					139	208		124	171	
ξ	12,00	1,46	74,91	86,31	3,05	9,27	13,87	1,50	8,27	11,40		

Tabelle 3: Satz- und Wortlängen und Schwierigkeitsindices der 15 Dummy-Sprichwörter

Wie in den zusammenfassenden Berechnungen zu sehen ist, wird immerhin die in der Anzahl der Wörter pro Satz gemessenen Satzlängen einheitlich korrekt gerechnet, wenn man von dem leicht korrigierbaren Fehler absieht, dass die Anzahl bei BF in jedem Satz genau um 1 erhöht ist. Gibt man alle 15 Dummies in Form eines bedingt als „homogen“ zu bezeichnenden „Pseudotextes“ ein, beläuft sich deshalb die Summe der Wörter bei LL korrekt auf $N = 124$, bei BF hingegen auf $N = 139$. Deutlich abweichend hingegen werden allerdings die Silben gezählt, deren Anzahl gemäß LL bei $N = 171$ und BF zufolge bei $N = 208$ liegt – im Vergleich zu $N = 180$ der authentischen Zählung. Aufgrund letzterer ergeben sich durchschnittliche Flesch- und Amstad-Werte von $\Phi = 74.91$ bzw. $A = 86.31$. Abbildung 2a zeigt den linearen Zusammenhang zwischen beiden Werten für die 15 Sprichwort-Dummies ($r > 0.99, p < 0.001$). Während dieser Befund mehr oder weniger so zu erwarten war, weist Abbildung 2b interessanterweise auch den Zusammenhang zwischen den A-Werten und den

Tuldava'schen T-Werten als linear aus ($r = 0.90, p < 0.001$), wobei der Durchschnittswert für den Tuldava-Index bei $T = 3.05$ liegt.

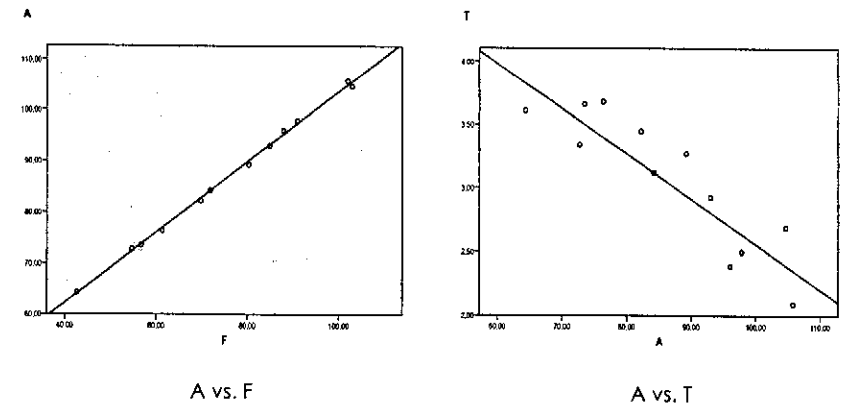


Abbildung 2a/b: Lineare Zusammenhänge zwischen Schwierigkeits-Indices

Diese Beobachtung öffnet weitreichende Perspektiven im Hinblick auf eine Neu-Begründung und theoretische Fundierung eines Index zur Einschätzung der Textschwierigkeit, die hier nicht weiter verfolgt werden können. Im vorliegenden Text wollen und müssen wir uns vielmehr auf die Schlussfolgerung beschränken, dass die Indexwerte der Textschwierigkeit sowohl für die authentischen als auch für die Dummy-Sprichwörter eindeutig in dem Bereich liegen, der üblicherweise als ‚sehr leicht‘ oder ‚extrem leicht‘ bezeichnet wird, und in dem Textsorten wie Werbetexte, Slogans u.ä. angesiedelt werden. Damit ist sichergestellt, dass die in das Experimental-Korpus zu Zwecken der Konsolidierung integrierten Dummy-Sprichwörter der sprachlichen Norm authentischer Sprichwörter entsprechen. Die weiteren Untersuchungen werden zeigen, ob auch diese Dummy-Sprichwörter auf einen gewissen Bekanntheitsgrad kommen; das sollte im Falle von Ganztext-Präsentationen durchaus zu erwarten sein – den Nageltest werden dann Teilttext-Präsentationen dieser Dummy-Sprichwörter darstellen.

Literaturverzeichnis

- Amstad, Toni (1978), *Wie verständlich sind unsere Zeitungen?* Diss., Zürich.
- Best, Karl-Heinz (2006), Sind Wort- und Satzlänge brauchbare Kriterien zur Bestimmung der Lesbarkeit von Texten? In: Wichter, Sigurd & Busch, Albert (Hrsg.) (2006), 21-31.
- Chlosta, Christoph & Ostermann, Thorsten (2003), Überlegungen zu einer Laiendefinition des Sprichworts. *Proverbium International Yearbook of Proverb Scholarship* Vol. 20, 51-70.
- Decker, Reinhold & Lenz, Hans-J. (eds.) (2007), *Advances in Data Analysis*. Berlin et al.: Springer.
- Flesch, Rudolf (1948), A New Readability Yardstick. *Journal of Applied Psychology* 32: 3, 221-233.
- Grotjahn, Rüdiger & Altmann, Gabriel (1993), Modelling the Distribution of Word Length: Some Methodological Problems. In: Köhler, Reinhard & Rieger, Burkhard (eds.) (1993), 141-153.
- Grotjahn, Rüdiger (1982), Ein statistisches Modell für die Verteilung der Wortlänge. *Zeitschrift für Sprachwissenschaft* 1, 44-75.
- Grotjahn, Rüdiger (2000), Determinanten der Schwierigkeit von Leseverstehensaufgaben: Theoretische Grundlagen und Konsequenzen für die Entwicklung des TESTDAF. In: Bolton, Sibylle (Hrsg.) (2000), *TESTDAF: Grundlagen für die Entwicklung eines neuen Sprachtests. Beiträge aus einem Expertenseminar*. Köln: VUB Gilde, 7-55.
- Grzybek, Peter (2006), History and Methodology of Word Length Studies. The State of the Art. In: Grzybek, Peter (ed.) (2006), *Contributions to the Science of Text and Language. Word Length Studies and Related Issues*. Dordrecht, NL: Springer [Paperback ed. 2007], 15-90.
- Grzybek, Peter (2008), Proverb Frequency and Proverb Familiarity: Two Faces of One Coin or False Friends? Paper presented at the International Congress of the European Society for Phraseology, Helsinki 2008 (To be published).
- Grzybek, Peter & Schlatte, Rudi (2002), Zur Satzlänge deutscher Sprichwörter – Ein Neu-Ansatz. In: Piirainen, Elisabeth & Piirainen, Ilpo (Hrsg.) (2002), *Phraseologie in Raum und Zeit, Akten der 10. Tagung des Westfälischen Arbeitskreises »Phraseologie / Parömiologie« (Münster 2001)*. Baltmannsweiler: Schneider, 287-305.
- Grzybek, Peter & Stadlober, Ernst (2007), Do We Have Problems With Arens' Law? A New Look at the Sentence-Word Relation. In: Grzybek, Peter & Köhler, Reinhard (eds.) (2007), 205-218.
- Grzybek, Peter & Köhler, Reinhard (eds.) (2007), *Exact Methods in the Study of Language and Text. Dedicated to Professor Gabriel Altmann on the Occasion of His 75th Birthday*. Berlin / New York: Mouton de Gruyter.
- Grzybek, Peter & Chlosta, Christoph (2009), Some Essentials on the Popularity of (American) Proverbs. In: McKenna, Kevin J. (eds.) (2009), 95-110.
- Grzybek, Peter; Stadlober, Ernst & Kelih, Emmerich (2007), The Relation of Word Length and Sentence Length: The Inter-Textual Perspective. In: Decker, Reinhold & Lenz, Hans-J. (eds.) (2007), 611-618.
- Grzybek, Peter; Kelih, Emmerich & Stadlober, Ernst (2008), The relation between word length and sentence length: an intra-systemic perspective in the core data structure. *Glottometrics* 16, 111-121.
- Grzybek, Peter; Chlosta, Christoph & Grotjahn, Rüdiger (2009), *Empirische Sprichwortforschung: Methoden – Status – Perspektiven* (In print).

- Hřebíček, Ludek & Altmann, Gabriel (eds.) (1993), *Quantitative Text Analysis*. Trier: wvt [= Quantitative Linguistics 52].
- Köhler, Reinhard & Rapp, Reinhard (2007), Familiarity and frequency: A psycholinguistic application of synergetic linguistics. *Glottometrics* 15, 62-70.
- Köhler, Reinhard & Rieger, Burkhard (eds.) (1993), *Contributions to Quantitative Linguistics*. Dordrecht, Boston, London: Kluwer.
- Lively, Bertha A. & Pressey, S.L. (1923), A Method for Measuring the 'Vocabulary Burden' of Textbooks. *Educational Administration and Supervision* 9: 7, 389-398.
- McKenna, Kevin J. (ed.), *The Proverbial „Pied Piper“: A Festschrift Volume of Essays in Honor of Wolfgang Mieder on the Occasion of His Sixty-Fifth Birthday*. New York et al.: Lang.
- Permjakov, Grigorij L. (1970), *Ot pogovorki to skazki. Zametki po obščej teorii kliše*. Moskva [Engl.: From proverb to folk-tale. Notes on the general theory of cliché. Moscow, 1979].
- Tuldava, Juvan (1993), The Statistical Structure of a Text and its Readability. In: Hřebíček, Ludek & Altmann, Gabriel (eds.) (1993), 215-227.
- Tuldava, Juvan (1994), Measuring Text Difficulty. *Glottometrika* 14, 69-81.
- Simrock, Karl (1846, 1988), *Die deutschen Sprichwörter*. Stuttgart: Reclam.
- Vogel, Mabel & Washburne, Carleton (1928), An Objective Method of Determining Grade Placement of Children's Reading Material. *The Elementary School Journal* 28, 373-381.
- Wichter, Sigurd, & Busch, Albert (Hrsg.) (2006), *Wissenstransfer – Erfolgskontrolle und Rückmeldungen aus der Praxis*. Frankfurt am Main: Lang.
- Wimmer, Gejza; Köhler, Reinhard; Grotjahn, Rüdiger & Altmann, Gabriel (1994), Towards a Theory of Word Length Distribution. *Journal of Quantitative Linguistics* 1: 1, 98-106.