

ПОЛУТРОПОН

к 70-летию
Владимира
Николаевича
Топорова

 ИЗДАТЕЛЬСТВО
«ИНДРИК»
Москва 1998

P. Grzybek
(Graz)

Explorative Untersuchungen zur Wort- und Satzlänge kroatischer Sprichwörter

(Am Beispiel der *Poslovice* von Đuro Daničić [1871])

0. Einleitung

Wie moderne semiotische Untersuchungen der letzten Jahre gezeigt haben, nimmt das Sprichwort im Hinblick auf seinen semiotischen Status einen ambivalenten bzw. doppelten Stellenwert ein, der mit seinem Wesen als «Minimaleinheit der übersprachlichen semiotischen Ebene» (Čerkasskij 1968) verbunden ist: Auf der einen Seite gehört es als komplexes sprachliches Zeichen zur Ebene der Sprache, auf der anderen Seite läßt es sich als 'sekundäres modellbildendes System' der Ebene der Folklore zuordnen (vgl. Permjakov 1979).

Bemerkenswert ist, daß ungeachtet der semiotischen Betrachtungen der Sprichwörter, die in der jüngsten Vergangenheit angestellt wurden (vgl. Grzybek 1984), ihr linguistischer Status vergleichsweise unberücksichtigt geblieben ist (vgl. Cram 1983). Charakteristisch ist in dieser Hinsicht zum Beispiel der Umstand, daß der synoptische Überblick von Röhrich/Mieder (1977) zum Thema 'Sprichwort' zwar ein Kapitel zur 'Sprachgeschichte', aber keines zur Sprachwissenschaft bzw. Linguistik enthält. In gewisser Hinsicht spiegelt sich hierin der Umstand wider, daß das Sprichwort als Forschungsgegenstand lange Zeit von solchen Disziplinen wie der Ethnographie, der Folkloristik u. a. dominiert wurde. Ungeachtet dessen hat es allerdings in der Vergangenheit zahlreiche linguistische Untersuchungen gegeben, angefangen von Untersuchungen zur Syntax, zur Semantik u. a. m. (vgl. z. B. Norrick 1985, Peukes 1977, Tarlanov 1982). All diese Untersuchungen haben jedoch stets einen mehr oder weniger exemplarischen Charakter gehabt, insofern sie in der Regel einzelsprachliche Aussagen getroffen haben und in den seltensten Fällen in Hypothesen über allgemeine Gesetzmäßigkeiten gemündet sind.

Es liegt nahe, daß insbesondere Ansätze der quantitativen Linguistik (QL) bzw. der Linguostatistik das offenkundige Defizit an Modellen zu allgemeinen sprachlichen Gesetzmäßigkeiten in Sprichwörtern zu überwinden in der Lage wären; denn Grundanliegen der

QL ist ja die Beschreibung und Erklärung sprachlicher Erscheinungen auf der Basis gesammelter (empirischer) Daten, die zur Überprüfung bestimmter Hypothesen mittels verschiedener (statistischer) Testverfahren geeignet sind (vgl. Kempgen 1995). Allerdings haben die methodischen Ansätze dieser Disziplinen die Sprichwortforschung bislang bestenfalls peripher berührt.

1. Parömiologie und Quantitative Linguistik

Die folgenden Ausführungen, die einen Versuch in diese Richtung darstellen sollen, haben aufgrund der geringen Anzahl von Vorarbeiten in diesem Bereich mitunter eine starke methodologische Ausrichtung und weisen einen eher explorativen oder hypothesenbildenden Charakter auf — sie können und sollen im Grund genommen nicht mehr und nicht weniger, als die Fruchtbarkeit quantitativer Ansätze im Bereich der Parömiologie zu belegen.

Quantitative Ansätze im Bereich der Parömiologie setzen die Existenz eines bestimmten Sprichwortkorpus voraus, unabhängig davon, ob generelle Aussagen über dieses Korpus als Ganzes oder spezifische Aussagen über die Eigenschaften einzelner Sprichwörter in Relation zum gesamten Korpus getroffen werden sollen. Die Ergebnisse dieser Untersuchungen sind abhängig von der Beschaffenheit der einzelnen Texte bzw. des Textkorpus und können von daher nicht ohne weiteres als repräsentativ für die betreffende Sprache schlechthin angesehen werden. Üblicherweise zeichnen sich Textkorpora, die quantitativen Analysen unterworfen werden, durch eine kontinuierliche Dynamik aus, d. h. daß es sich in der Regel um fortlaufende Fließtexte handelt; im Gegensatz dazu weist ein Sprichwortkorpus einen eher «additiven» Charakter auf, insofern die einzelnen das Korpus konstituierenden Einheiten (die einzelnen Sprichwörter) als jeweils in sich abgeschlossene Texte angesehen werden können. Insofern können Eigenschaften, die sich an einem Sprichwortkorpus beobachten bzw. quantifizieren lassen, in dreierlei Hinsicht von Relevanz sein, nämlich im Hinblick auf Eigenschaften:

- (a) eines Sprichwortkorpus als Ganzes
- (b) bestimmter Elemente innerhalb eines einzelnen Sprichworttextes
- (c) eines einzelnen Sprichworts als Ganzes

Entsprechend läßt sich — auch wenn sich die verschiedenen Betrachtungsebenen überlagern und nicht wechselseitig ausschließen — heuristisch zwischen einer vertikalen (a) und zwei horizon-

talen (b & c) Perspektiven auf das zu untersuchende Material unterscheiden:

Sw₁ a x x x b₁ x x b₁ x x

Sw₂ a x x x x x x

Sw₃ x b₂ a x x b₂ x x x

Sw₄ x x x x x x x x

...

Sw_n x x x x a b_n x b_n x x

Konkrete Beispiele für die verschiedenen Herangehensweisen sind an anderer Stelle diskutiert worden (vgl. Grzybek 1997d) und können deshalb an dieser Stelle ausgespart bleiben. Wir können uns deshalb unmittelbar der Fragestellung zuwenden, um die es in der vorliegenden Darstellung gehen soll, nämlich um die Frage der Wort- und Satzlänge von Sprichwörtern unter der Berücksichtigung des Faktors ihrer Bekanntheit. Über die «einfache» Untersuchung der Wort- und Satzlänge hinaus wollen wir dabei Überlegungen anstellen, die sich dem Wechsel- bzw. Abhängigkeitsverhältnis von Wort- und Satzlänge, ebenfalls unter Berücksichtigung des Faktors der Bekanntheit, anstellen. Die die folgenden Überlegungen leitenden Hypothesen können resümierend in folgender Form formuliert werden:

- (1) Der Faktor der Bekanntheit von Sprichwörtern wirkt sich auf die durchschnittliche Wortlänge aus.
- (2) Der Faktor der Bekanntheit von Sprichwörtern wirkt sich auf die durchschnittliche Satzlänge aus.
- (3) Die Wort- und Satzlänge von Sprichwörtern stehen in einem bestimmten Wechsel- bzw. Abhängigkeitsverhältnis.
- (4) Der Faktor der Bekanntheit von Sprichwörtern wirkt sich auf das Wechsel- bzw. Abhängigkeitsverhältnis von Wort- und Satzlänge der Sprichwörter aus.

2. Das Untersuchungsmaterial: Die *Poslovice* von Đuro Daničić (1871)

Wir wollen im folgenden diese Hypothesen der Reihe nach bearbeiten. Als Ausgangs- und Untersuchungsmaterial soll uns dabei die Sammlung kroatischer Sprichwörter *Poslovice* von Đuro Daničić dienen, die 1871 in Zagreb erschien. Dieser Sammlung ist innerhalb der kroatischen Parömiographie stets ein besonderer Stellenwert beigemessen worden, was sich unter anderem darin äußert, daß die in ihr enthaltenen Belege später komplett in die grundlegende Sammlung kroatischer Sprichwörter *Hrvatske poslovice* von Vicko Skarpa (1909)

integriert wurden. So bezeichnet auch Čubelić (1981: 321) noch in neuerer Zeit die Sammlung von Daničić als «eines der bedeutendsten Werke der parömiologischen Fachliteratur».

Die Sammlung von Daničić basiert auf vier Quellen:

1. einer Dubrovniker Handschrift aus dem Jahre (1697); diese unter dem Titel *Dubrovački rukopis* bekannte (heute in der Zagreber Universitätsbibliothek befindliche) Dubrovniker Handschrift — welcher der Jesuit Ivan Matija Matijašević (1714–1791) die Bezeichnung *Proričja slovinška* verlieh — war nur kurz vor dem Erscheinen der Sammlung von Daničić unter dem Titel *Starinnyj sbornik serbskich poslovic* (unter dem sie auch heute noch bekannt ist) von A. Gil'ferding (1868) herausgegeben worden. Daničić hatte sich in einer Rezension kritisch mit dieser Ausgabe auseinandergesetzt und dabei auch als Kenner der zeitgenössischen Parömiologie ausgewiesen (vgl. Daničić 1870).
2. einer Handschrift aus der ersten Hälfte des 18. Jahrhunderts, die mitunter als *gajuša* bezeichnet wird und die ihren Namen der Tatsache verdankt, daß ein Großteil der in ihr enthaltenen Sprichwörter in den Jahren 1864–66 von Ljudevit Gaj in der *Danica ilirska* publiziert wurde. Nach Kekez (1984: 18f., 1986: 15) geht diese Sammlung allerdings auf den Dubrovniker Dichter und Historiographen Ivo Aletina (1670–1743) zurück und nicht, wie von Gaj angenommen, auf Ignjat Đorpić [Đurpević].
3. Sprichwörter aus der Fabelsammlung *Fabulae ab illyricis adagiis desumptae* von Puro Ferić (Dubrovnik 1794).¹
4. Sprichwörter aus dem *Dizionario italiano-latino-illirico* von Ardelio Della Bella (Venedig 1728 / Dubrovnik 1785).

Die (zum Teil korrigierten) Sprichwörter dieser Sammlungen sind in der Sammlung von Daničić alphabetisch angeordnet. Dabei kommt es im Gesamtbestand nahezu selbstverständlich zu einer Reihe von Überschneidungen, d. h. zahlreiche Sprichwörter kommen in identischer Form in mehr als einer Quelle vor. Die Belegquellen der einzelnen Sprichwörter sind jeweils mit Kürzeln gekennzeichnet, so daß teilweise mehrere Quellen bei einem Sprichwort angegeben sind. Die Sprichwörter aus der Dubrovniker und der Zagreber Handschrift stellen dabei den Großteil der Sammlung dar: 3.598 Sprichwörter sind in der Dubrovniker Sammlung belegt, 2.237 Sprichwörter in der Handschrift

¹ Insgesamt finden sich 111 Sprichwörter jeweils zu Anfang der Fabeln; Daničić hat allerdings — abgesehen von einigen fehlerhaften Quellenangaben — die Sprichwörter zu den Fabeln 13–25 ausgelassen, so daß sich bei in seiner Sammlung nur 98 Sprichwörter finden.

von Aletina.² Im Endergebnis macht die Summe dieser Sprichwörter, 5.935 Sprichworttexte aus, die uns im folgenden als Untersuchungsmaterial dienen sollen.

3. Der Faktor der Bekanntheit

Im Rahmen eines Forschungsprojekts zur Bekanntheit kroatischer Sprichwörter wurde unter anderem die Sammlung von Daničić an insgesamt 15 Informant(inn)en im gesamten kroatischen Sprachbereich verteilt. Aufgabe dieser Personen war es, all diejenigen Sprichwörter zu kennzeichnen, die ihnen mit Sicherheit unbekannt waren, damit das heutzutage unbekanntes Material aus weiteren Untersuchungsschritten (mit Teiltex-Präsentationen (vgl. Grzybek/Chlosta 1993) ausgeschlossen werden kann. Der detaillierte Hintergrund des Projekts (vgl. Grzybek et al. 1993, Гржибек 1994) und die Einzelergebnisse zur Sammlung von Daničić (Grzybek 1997a) sowie zu anderen kroatischen Quellen (Grzybek 1997b, c) müssen an dieser Stelle nicht im einzelnen dargestellt werden. Es sei allerdings hier deutlich hervorgehoben, daß aufgrund der bislang verfügbaren Informationen noch keine allgemeinverbindlichen Daten über die Bekanntheit kroatischer Sprichwörter vorliegen. Dies betrifft nicht nur Aussagen über den tatsächlichen Bekanntheitsgrad der einzelnen Sprichwörter, sondern auch Frequenzanalysen zu den einzelnen Varianten der Sprichwörter. Beide Punkte sind für die unten folgenden quantitativen Analysen von besonderer Bedeutung, insofern einerseits Aussagen über die Bekanntheit von Sprichwörtern unbedingt auf eine solidere Basis gestellt werden müssen, und insofern andererseits die heute usuelle konkrete sprachliche Oberflächenform nicht unbedingt der bei Daničić kodifizierten entsprechen muß. — Ungeachtet dessen scheint es allerdings auch zum jetzigen Zeitpunkt bereits durchaus berechtigt, im Hinblick auf weitere Untersuchungen gesicherte Hypothesen über die Menge der eher bekannten Sprichwörter und ihnen gegebenenfalls eigene sprachliche Besonderheiten aufzustellen. — Im Rahmen der vorliegenden Untersuchung sind deshalb von den insgesamt 5.935 Sprichwörtern der Daničić-Sammlung diejenigen 625 Einheiten in

² Die Angabe von 3.598 Einheiten der Dubrovniker Handschrift geht auf die Kennzeichnungen von Daničić zurück und wurden nicht überprüft. — Was die Sammlung von Aletina angeht, so handelt es sich de facto um 2.237 Einträge, von denen 2.210 bei Daničić gekennzeichnet sind, eine weitere, bei Daničić gar nicht gekennzeichnete kommt hinzu. Kekez (1984:18, 1986: 15) hingegen geht davon aus, daß sich in der Zagreber Sammlung 2.226 Einheiten befinden.

die Teilmenge der «bekannteren» Sprichwörter aufgenommen worden, die im ersten, oben beschriebenen Untersuchungsschritt von mehr als sieben der fünfzehn Informant(inn)en als «nicht unbekannt» eingestuft wurden.

4. Durchschnittliche Satzlänge

Die durchschnittliche Satzlänge wird im folgenden nach der Anzahl der die einzelnen Sätze konstituierenden Wörter berechnet. Dabei gilt es zu berücksichtigen, daß einige Sprichwörter sich aus mehr als einem Satz zusammenfügen, so daß wir es insgesamt beim Daničić-Korpus mit 5.957 Sätzen zu tun haben. Die durchschnittliche Länge dieser Sätze beträgt $\bar{x} = 6.21$ Wörter bei einer Standardabweichung von $s = 2.61$. Allfällige weiterführende Aussagen oder gar Bewertungen — wie etwa im Hinblick auf die in der Parömiologie immer wieder als Definitionskriterium des Sprichworts vermutete Kürze der Sätze o. a. — müssen ohne einen Vergleich mit der kroatischen Norm als unzulässig angesehen werden; dies würde allerdings den Rahmen der vorliegenden Arbeit aufgrund der spärlichen Vorarbeiten zu dieser Frage überschreiten.

Fig. 1 veranschaulicht dabei sehr deutlich, daß die Vorkommenshäufigkeit der Sätze mit einer bestimmten Wortanzahl sich nicht gleichmäßig verteilt, sondern daß ca. die Hälfte der Sätze 4–6 Wörter enthält, und daß ca. 80% der Sätze 3–8 Wörter enthalten; die genauen Daten sind der Tab. 1 (s. u.) zu entnehmen.

Die Ergebnisse lassen sich anschaulich in Form der Fig. 1 darstellen.

Die nächste Frage richtet sich auf die Untersuchung der Satzlänge der 625 «bekannteren» Sprichwörter; die entsprechenden Analysen führen zu dem Ergebnis, daß die in diesem Korpus enthaltenen 626 Sätze eine durchschnittliche Satzlänge von $\bar{x} = 5.57$ Wörter pro Satz bei einer Standardabweichung von $s = 1.97$ aufweisen. Der Unterschied zwischen der Gesamtmenge und der Teilmenge der «bekannteren» Sprichwörter erweist sich als hoch signifikant ($t = 5.96$; $p < 0.0001$). Der Median des Gesamtkorpus ($Md = 4.14$) unterscheidet sich dabei weniger ausgeprägt von dem der «bekannteren» Sprichwörter ($Md = 4.00$); so läßt auch die graphische Darstellung deutlich werden, daß die Kurven sich in ihrem Verlauf nicht wesentlich unterscheiden, daß aber die Kurve der «bekannteren» Sprichwörter sich deutlich spitzer ausnimmt; in der Tat machen hier die Sätze mit 4–6 Wörtern fast zwei Drittel aus, die Sätze mit 3–8 Wörtern betragen über 90% (vgl. Fig. 2).

Wörter / Satz	Dan-Ges		Dan (> 50)	
	(f / abs)	(f / %)	(f / abs)	(f / %)
1	10	0,17	1	0,16
2	115	1,93	12	1,92
3	471	7,91	51	8,15
4	1022	17,16	129	20,61
5	1120	18,8	154	24,6
6	954	16,01	113	18,05
7	751	12,61	76	12,14
8	577	9,69	47	7,51
9	354	5,94	20	3,19
10	226	3,79	13	2,08
11	146	2,45	3	0,48
12	76	1,28	4	0,64
13	37	0,62	1	0,16
14	37	0,62	0	0
15	17	0,29	1	0,16
16	17	0,29	0	0
17	9	0,15	0	0
18	7	0,12	0	0
19	0	0	0	0
20	3	0,05	1	0,16
21	1	0,02		
22	1	0,02		
23	0	0		
24	1	0,02		
25	1	0,02		
26	1	0,02		
27	1	0,02		
28	0	0		
29	0	0		
30	1	0,02		
31	1	0,02		

Tab. 1: Durchschnittliche Satzlänge (in Wörtern)

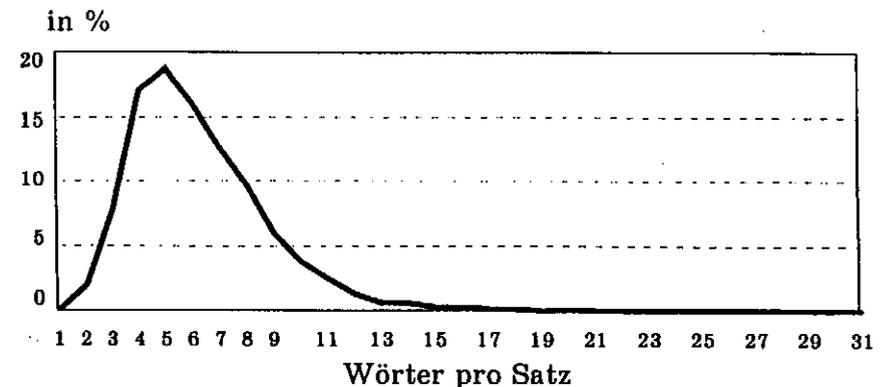


Fig. 1: Anzahl der Sätze (in %) in Relation zu Wörtern pro Satz

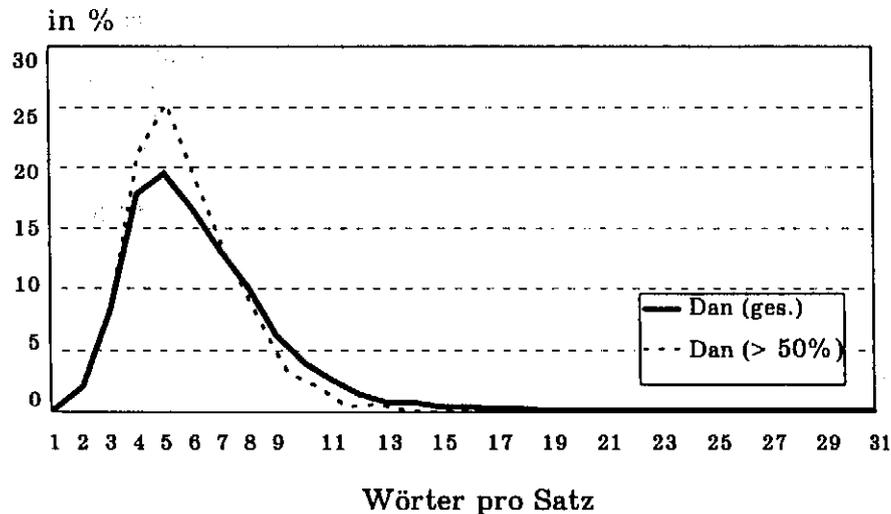


Fig. 2: Anzahl der Sätze (in %) in Relation zu Wörtern pro Satz unter Berücksichtigung der Bekanntheit der Sprichwörter

Wir können damit die Hypothese (1) als nachgewiesen betrachten: Der Faktor der Bekanntheit wirkt sich auf die durchschnittliche Satzlänge aus — die Menge der weniger bekannten Sprichwörter weist im Durchschnitt eine signifikant größere Satzlänge auf als die Menge der «bekannteren» Sprichwörter. Damit bestätigen sich am kroatischen Material die Ergebnisse zur Satzlänge deutscher Sprichwörter, an denen ebenfalls der positive Einfluß der Bekanntheit auf die durchschnittliche Satzlänge nachgewiesen werden konnte (vgl. Grzybek 1995).

5. Durchschnittliche Wortlänge

Die Wortlänge wird im folgenden nach der Anzahl der die einzelnen Wörter konstituierenden Grapheme berechnet. Die folgende Tab. 2 gibt die Daten sowohl für die Gesamtmenge der Sprichwörter als auch für die Teilmenge der 625 «bekannteren» Sprichwörter wider:

Wortlänge (G)	Dan-Ges (f / abs)	Dan (>50) (f / abs)	Dan-Ges (f / %)	Dan (>50) (f / %)
1	2711	234	7.33	6.70
2	7803	759	21.10	21.75
3	4471	448	12.09	12.84
4	6897	710	18.65	20.34
5	6319	624	17.09	17.88
6	3955	391	10.70	11.20
7	2427	188	6.56	5.39
8	1348	79	3.65	2.26
9	627	32	1.70	0.92
10	253	18	0.68	0.52
11	104	4	0.28	0.11
12	40	2	0.11	0.06
13	12		0.03	
14	3		0.01	
15	3		0.01	
16	1		0.00	
17	0			
18	0			
19	0			
20	0			

Tab. 2: Wortlänge

In der graphischen Darstellung (Fig. 3) zeigt sich die Ähnlichkeit des Kurvenverlaufs der «bekannteren» Sprichwörter und der Gesamtmenge:

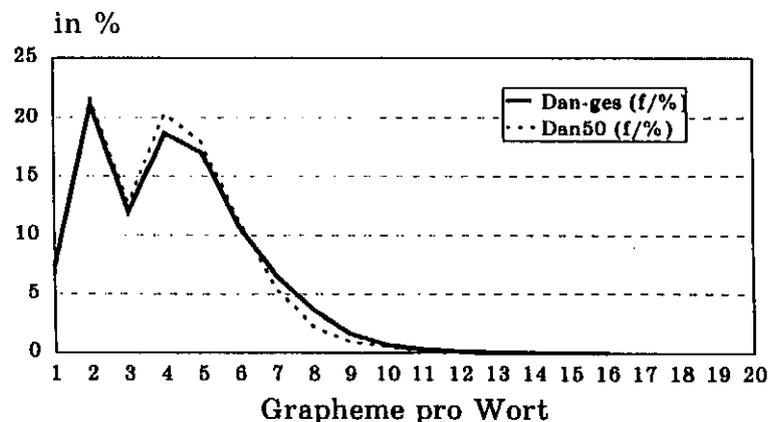


Fig. 3: Die Wortlänge

In der graphischen Darstellung täuscht allerdings der Einbruch bei der Anzahl der Wörter mit 3 Graphemen darüber hinweg, daß erstens die Spitze der «bekannteren» Sprichwörter mit fünf Graphemen ausgeprägter ist als beim Gesamtkorpus, und daß die Kurve der «bekannteren» ab sechs Graphemen steiler abfällt als beim Gesamtkorpus. In der Tat weicht die durchschnittliche Wortlänge von Gesamt- und Teilmenge voneinander ab: So beträgt die durchschnittliche Länge der insgesamt 36.974 Wörter der Daničić-Sammlung $\bar{x} = 4.12$ Grapheme pro Wort bei einer Standardabweichung von $s = 2.08$; die entsprechende Analyse der 625 «bekannteren» Sprichwörter zeigt, daß hier die insgesamt 3.489 Wörter eine durchschnittliche Länge von $\bar{x} = 3.98$ Graphemen pro Wort bei einer Standardabweichung von $s = 1.89$ aufweisen. Dieser Unterschied erweist sich als signifikant ($t = 3.83$, $p < 0.001$).

Damit kann auch die zweite oben aufgestellte Hypothese als nachgewiesen gelten: Der Faktor der Bekanntheit von Sprichwörtern wirkt sich positiv auf die durchschnittliche Satzlänge aus — die einzelnen Wörter der eher unbekannteren Sprichwörter sind im Durchschnitt länger als diejenigen der bekannteren Sprichwörter. Dieses Ergebnis deckt sich ebenfalls mit den Befunden am deutschsprachigen Material (Grzybek 1997d).

6. Die Beziehung zwischen Wort- und Satzlänge

Die folgenden Berechnungen beziehen sich auf die oben vorgetragene Annahme, daß es einen Zusammenhang zwischen der durchschnittlichen Wort- und Satzlänge gibt. Die Daten sind im einzelnen der Tab. 3 zu entnehmen: In der ersten Spalte ist die jeweilige Anzahl von Wörtern pro Satz angegeben, in der zweiten und dritten die absolute und relative Anzahl der Sätze, welche die jeweilige Wortanzahl aufweisen; in der vierten Spalte findet sich die durchschnittliche Anzahl von Graphemen pro Wort in den jeweiligen Sätzen.

Wie zu sehen ist, weist die durchschnittliche Anzahl der Grapheme pro Wort eine Tendenz auf, bei zunehmender Wortanzahl pro Satz abzunehmen. Diese Tendenz läßt sich anschaulich in Fig. 3 darstellen. Wir schränken dabei die graphische Darstellung auf den Bereich der Sätze mit $2 \leq n \leq 14$ ein, da nur in diesem Bereich eine aussagenkräftige Anzahl von Beobachtungen vorliegt.

Wörter / Satz	n (Sätze)	Sätze in %	\bar{x} (Grapheme)
1	10	0,17	6,50
2	115	1,93	6,04
3	471	7,91	5,10
4	1022	17,16	4,53
5	1120	18,80	4,34
6	954	16,01	4,16
7	751	12,61	4,03
8	577	9,69	3,93
9	354	5,94	3,84
10	226	3,79	3,77
11	146	2,45	3,78
12	76	1,28	3,78
13	37	0,62	3,71
14	37	0,62	3,65
15	17	0,29	3,73
16	17	0,29	3,79
17	9	0,15	3,93
18	7	0,12	3,9
19	0		
20	3	0,05	3,65
21	1	0,02	3,19
22	1	0,02	4,09
23	0		
24	1	0,02	5,17
25	1	0,02	3,84
26	1	0,02	4,08
27	1	0,02	3,63
28	0		
29	0		
30	1	0,02	3,20
31	1	0,02	4,00

Tab. 3: Durchschnittliche Wortlänge in Abhängigkeit von der Satzlänge

Grapheme pro Wort

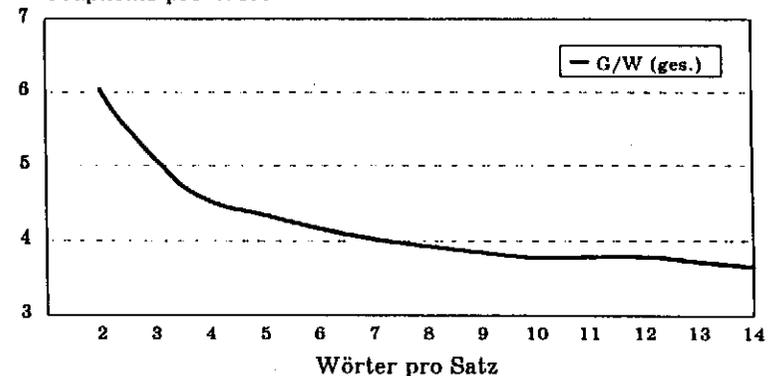


Fig. 4: Zusammenhang von Satz- und Wortlänge

Fig. 4 präsentiert die Daten graphisch als Kurvenverlauf: Der optische Eindruck bestätigt die Beobachtung, daß die durchschnittliche Anzahl der Grapheme pro Wort bei zunehmender Wortanzahl pro Satz abnimmt. Eine ähnliche Feststellung hat Krikmann (1967) bereits vor 30 Jahren an estnischem Sprichwortmaterial machen können, ohne daß seine Analysen allerdings die Parömiologie zu ähnlichen gelagerten Studien inspiriert hätten. Krikmann maß allerdings die Wortlänge auf der Basis der Silbenanzahl: dabei stellte sich heraus, daß die Wörter im Durchschnitt um so kürzer waren, je länger die jeweiligen Sätze waren. Es läßt sich somit die begründete Vermutung äußern, daß wir es hier mit einer gesetzmäßigen Erscheinung zu tun haben; sollte dies der Fall sein, müßte sich die Tendenz bzw. das ihr zugrundeliegende Abhängigkeitsverhältnis auch quantifizieren lassen. Eine einfache Korrelationsanalyse ergibt allerdings eine nicht signifikante negative Korrelation von $r = -0.38$; nach einer Logarithmisierung der Daten beträgt der Korrelationskoeffizient $r = -0.42$. Es ist deshalb davon auszugehen, daß keine lineare Beziehung zwischen der Wort- und Satzlänge vorliegt, sondern daß wir es mit einer komplexeren Erscheinung zu tun haben.

Theoretischen Hintergrund für weiterführende Annahmen und Analysen könnte dabei das sogenannte «Menzerath'sche Gesetz» darstellen. In der von Altmann (1980) vorgeschlagenen allgemeinsten Formulierung besagt dieses, daß die ein sprachliches Konstrukt konstituierenden Einheiten um so kleiner sind, je größer das betreffende sprachliche Konstrukt ist, wobei sich die Länge der Komponenten als Funktion der Länge der Sprachkonstrukte verstehen läßt. Im Hinblick auf das von Altmann zunächst angesprochene Problem, inwiefern die von ihm vorgeschlagenen Funktionen nur zwischen den jeweils nächsthöheren bzw. nächstniedrigeren Sprachebenen oder über diese hinweg vorliegen, gibt es seither zahlreiche Einzelstudien, u. a. auch recht früh zur Satzebene (vgl. Köhler 1982), die an den verschiedensten Sprachen folgende Gesetzmäßigkeiten festgestellt haben: Das Menzerath'sche Gesetz trifft demnach auf das Verhältnis der Einheiten unmittelbar benachbarter Ebenen (Satz — Clause — Wort — Morphem / Silbe — Graphem / Phonem) zu, während bei Überschreitung einer Ebene die mitunter auch als «Arens'sches Gesetz» bezeichnete Regelmäßigkeit gilt, derzufolge sich das Verhältnis in diesem Fall genau umkehrt, so daß hier längere Konstrukte der höheren Ebene auch längere Konstituenten auf der niedrigeren Ebene aufweisen. So hatte Arens (1965: 7), ohne freilich die Komplexität der verschiedenen Ebenen zu berühren, z. B. explizit festgehalten, «daß bei wachsender durchschnittlicher Satzlänge auch die durchschnittliche Wortlänge wuchs.»

Auf der Grundlage dieser theoretischen Überlegungen wäre also zu erwarten, daß in unseren Sprichwörtern mit zunehmender Satzlänge auch die Wortlänge zunimmt. Unsere Daten können jedoch — ebenso wenig wie diejenigen von Krikmann (s. o.) — diese Regelmäßigkeit nicht bestätigen, wofür wir zum gegenwärtigen Zeitpunkt keine verbindliche Erklärung haben. Interessant ist dabei, daß sich diese abweichenden Tendenzen offenbar sowohl ergeben, wenn man der Wortlänge die Silbenanzahl als auch die Graphemanzahl als Maßstab zugrundelegt. Abgesehen von einem unbedingt notwendigen Vergleich mit der kroatischen Norm ist jedoch nicht auszuschließen, daß Sprichwörtern eigene Gesetzmäßigkeiten eigen sind, die sich aus ihrer Eigenschaft des abgeschlossenen Textes ergeben: So geht Köhler (1984) in seinen über den formalen Ansatz hinausgehenden Interpretationen des Menzerath'schen Gesetzes von Sätzen als unmittelbaren Komponenten einer Texteinheit aus, wobei die transphrastischen Bezüge für ihn den Gegenstand einer über die Komponenten hinausgehenden 'Strukturinformation' darstellen. Die von uns untersuchten Sprichwörter hingegen lassen sich ihrerseits als sprachlich-abgeschlossene Texte³ betrachten, auch wenn sie als solche in der Regel in einen verbalen und situativen Kontext eingebettet werden. Eine begründete Erklärung für unsere Beobachtungen zu finden, muß allerdings zukünftigen Studien vorbehalten bleiben, welche u. a. Analysen zur Satzlängenverteilung und Kurvenanpassungen, aus denen sich Aussagen über die genaue Art und Signifikanz des tatsächlichen Zusammenhangs ableiten lassen, beinhalten.

Wenden wir uns stattdessen abschließend der Untersuchung des Verhältnisses von Wort- und Satzlänge unter Berücksichtigung des Faktors der Bekanntheit der Sprichwörter zu.

6. Die Beziehung zwischen Wort- und Satzlänge unter Berücksichtigung der Bekanntheit der Sprichwörter

Wie oben bereits herausgearbeitet wurde, unterscheidet sich sowohl die Wort- als auch die Satzlänge der «bekannteren» Sprichwörter von denen des gesamten Korpus. Mit anderen Worten, der Faktor der Be-

³ Wie an anderer Stelle argumentiert wurde, ist es im Hinblick auf eine Abgrenzung der Sprichwörter von sprichwörtlichen Redensarten (und damit einhergehende Fragen der Definition des Sprichworts) sinnvoll, die textuelle Abgeschlossenheit der Sprichwörter nicht auf die sprachlich-grammatische Abgeschlossenheit zu beschränken, sondern auf der Grundlage der ihr zugrundeliegenden Abgeschlossenheit der logischen Struktur zu beurteilen (vgl. Eismann/Grzybek 1994).

kanntheit von Sprichwörtern wirkt sich sowohl auf die durchschnittliche Wort- als auch Satzlänge aus. Es liegt von daher nahe, sich auch das Abhängigkeitsverhältnis von Wort- und Satzlänge unter dem Gesichtspunkt der Bekanntheit der Sprichwörter differenzierter anzuschauen. Die entsprechenden Daten sind der Tab. 4 zu entnehmen.

Wörter / Satz	n (Sätze)	Sätze in %	\bar{x} (Grapheme)
1	1	0,16	8,00
2	12	1,92	5,75
3	51	8,15	4,75
4	129	20,61	4,16
5	154	24,60	4,24
6	113	18,05	3,92
7	76	12,14	3,76
8	47	7,51	3,65
9	20	3,19	3,59
10	13	2,08	3,51
11	3	0,48	3,97
12	4	0,64	3,52
13	1	0,16	3,62
14	0	0,00	
15	1	0,16	3,33
16	0	0,00	
17	0	0,00	
18	0	0,00	
19	0	0,00	
20	1	0,16	4,25

Tab. 4: Durchschnittliche Wortlänge der «bekannteren» Sprichwörter in Abhängigkeit von der Satzlänge (n = 626)

Die Daten sind anschaulich in Fig. 4 präsentiert, die einen Vergleich zwischen der Gesamtmenge und der Teilmenge der «bekannteren» Sprichwörter bietet. Dabei beschränkt sich die Darstellung bei den «bekannteren» Sprichwörtern auf die Wörter mit $3 \leq n \leq 9$ Graphemen pro Wort, da nur in diesem Bereich ausreichend Fälle vorliegen, um zuverlässige Aussagen zu treffen.

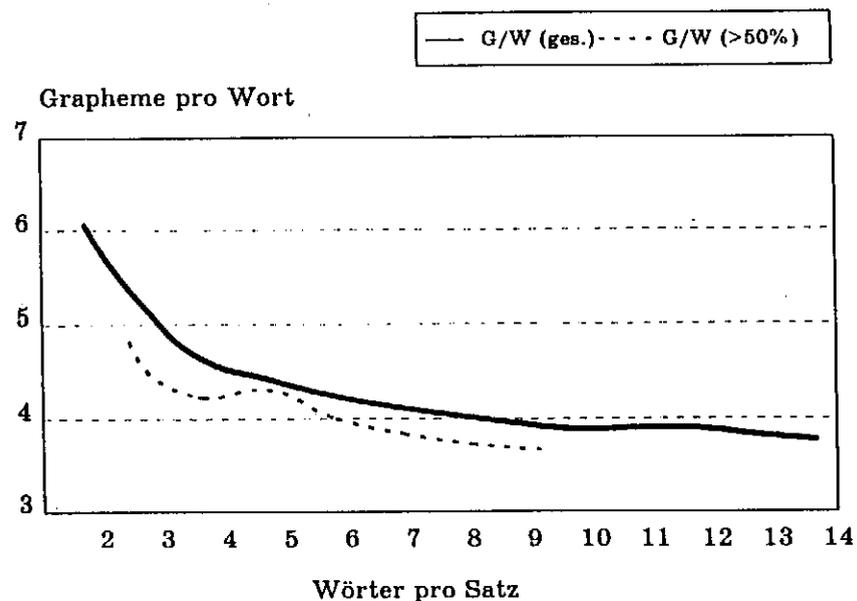


Fig. 5: Zusammenhang von Satz- und Wortlänge bei den «bekannteren» Sprichwörtern

Die Graphik veranschaulicht die oben bereits angesprochene Tendenz: Auch die Teilmenge der «bekannteren» Sprichwörter für sich genommen zeichnen sich von der Tendenz her bei zunehmender Satzlänge durch abnehmende Wortlänge aus. Auffällig ist dabei vom optischen Eindruck her die Aussetzung dieser Tendenz bei den fünf Wörtern umfassenden Sätzen. Entsprechende *t*-Tests, deren Ergebnisse in Tab. 5 dargestellt sind, bestätigen diesen Eindruck; die durchschnittliche Wortlänge in den Sätzen mit drei Wörtern — ebenso wie in denen mit fünf Wörtern — unterscheidet sich in der Gesamtmenge und der Teilmenge der «bekannteren» Sprichwörter nicht signifikant voneinander ($p > 0.05$), während sich in allen anderen Fällen die durchschnittliche Wortlänge signifikant voneinander unterscheidet.

Wörter / Satz	Grapheme/Wort (Daničić gesamt)		Grapheme/Wort (Daničić > 50%)		t	p
	x	s	x	s		
3	5,10	1,13	4,75	3,34	1,59	> 0.05
4	4,53	0,92	4,16	2,92	3,04	< 0.005
5	4,34	0,83	4,24	0,76	1,42	> 0.05
6	4,16	0,72	3,92	0,60	3,41	< 0.0005
7	4,03	0,69	3,76	0,59	3,29	< 0.005
8	3,93	0,62	3,65	0,46	3,03	< 0.005
9	3,84	0,62	3,59	0,46	1,77	< 0.05

Tab. 5: Wort- und Satzlänge in Abhängigkeit von der Bekanntheit

Ungeachtet dieser auffälligen Besonderheiten, die es an anderer Stelle auch an anderem kroatischen Sprachmaterial als an Sprichwörtern zu untersuchen gilt, bestätigt sich jedoch die oben bereits aufgezeigte Tendenz — dies macht Fig. 6 deutlich, die die Originaldaten in eine Trendberechnung transformiert: Es zeigt sich nicht nur die Tendenz der abnehmenden Wortlänge bei zunehmender Satzlänge, sondern auch die Tendenz ihrer signifikant stärkeren Ausprägung bei den «bekannteren» Sprichwörtern:

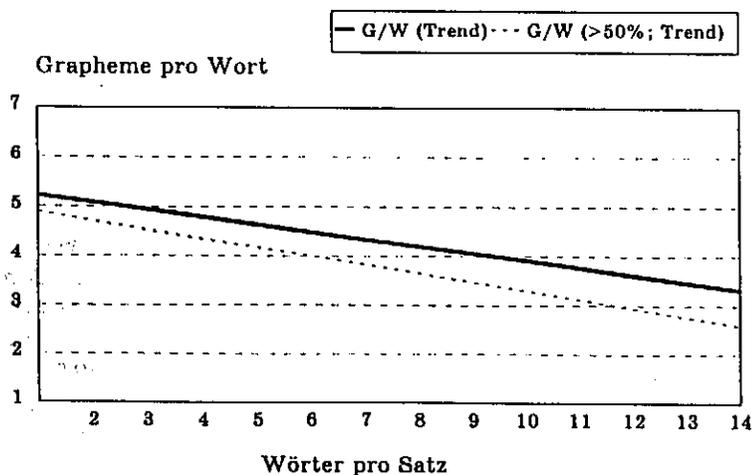


Fig. 6: Abhängigkeit von Wort- und Satzlänge im Trend

7. Resümee

Was eingangs bereits betont wurde, muß an dieser Stelle abschließend noch einmal hervorgehoben werden: der stark explorative Charakter der oben durchgeführten Analysen, der verschiedene Unsicherheitsfaktoren in sich birgt: die unzureichende empirische Absicherung des Faktors der Bekanntheit und die nicht empirisch standardisierte konkrete sprachliche Form der Sprichwörter einerseits, die Berechnung der Wortlänge auf der Basis von Graphemen andererseits.

Ungeachtet dieser «Risikofaktoren» scheinen jedoch für zukünftige Arbeiten einige nicht unwichtige Arbeitshypothesen durchaus begründbar. Dies bezieht sich nicht nur auf die Frage eines evtl. für Sprichwörter spezifischen Abhängigkeitsverhältnisses von Wort- und Satzlänge; dies bezieht sich auch auf die Annahme, daß der Faktor der Bekanntheit von Sprichwörtern bei allfälligen linguostatistischen Untersuchungen nicht vernachlässigt werden darf. Inwiefern hierbei in den Bereich der Psycholinguistik verweisende Hypothesen über die Richtung der Abhängigkeit von Bekanntheit der Sprichwörter einerseits und Wort- bzw. Satzlänge andererseits aufzustellen wären, geht allerdings über den eigentlichen Bereich der quantitativen Linguistik hinaus, auf deren Nutzbarkeit für den Bereich der Parömiologie ja eigentlich in den vorliegenden Zeilen aufmerksam gemacht werden sollte.

Literatur

- Arens, H. (1965): *Verborgene Ordnung. Die Beziehungen zwischen Satzlänge und Wortlänge in deutscher Erzählprosa vom Barock bis heute.* Düsseldorf.
- Altmann, G. (1980): «Prolegomena to Menzerath's Law.» In: Grotjahn, R. (Hg.), *Glottometrika 2.* Bochum. (1–10).
- Cram, D. (1983): «The Linguistic Status of the Proverb.» In: Mieder, W. (Hg.), *Wise Words. Essays on the Proverb.* New York/London, 1994. (73–97).
- Čerkasskij, M. A. (1968): s. Черкасский
- Čubelić, Tv. (1981): «Đuro Daničić — istraživač narodne književnosti.» In: *Zbornik o Đuri Daničiću.* Beograd/Zagreb. (313–328).
- Daničić, Đ. (1870): «Starinski sbornik srpskih poslovice A. Gilferding, S. Peterburg 1868.» In: *Rad JAZU*, knj. XII; 201–209.
- Daničić, Đ. (1871): *Poslovice.* Zagreb.
- Eismann, W.; Grzybek, P. (1994): «Sprichwort, sprichwörtliche Redensart, Phrasologismus: Vom Mythos der Nicht-Trennbarkeit.» In: Chlost, Ch.;

- Grzybek, P.; Piirainen, E. (eds.), *Sprachbilder zwischen Theorie und Praxis. Akten des Westfälischen Arbeitskreises „Phraseologie/Parömiologie“ (1991/1992)*. Bochum. (89–132).
- Gil'ferding, A. (1868): «Starinnyj sbornik serbskich poslovic.» In: *Zapiski Imperatorskogo Russkogo Geografičeskogo Obščestva (po otdeleniju tnografii), tom II*. Sankt Peterburg. (115–224).
- Grzybek, P. (1984): «Foundations of Semiotic Proverb Study.» In: Mieder, W. (Hg.), *Wise Words. Essays on the Proverb*. New York/London, 1994. (73–97).
- Гржибек, П. (1994): «Эмпирическая семиотика культуры на примере исследования пословиц с использованием пробного опытного изучения известности хорватских пословиц.» In: *Знаки Балкан. Часть 2*. Москва. (312–338).
- Grzybek, P. (1995): «Zur Frage der Satzlänge von Sprichwörtern (unter besonderer Berücksichtigung deutscher Sprichwörter).» In: Baur, R. S.; Chlosta, Ch. (Hg.), *Von der Einwortmetapher zur Satzmetapher. Akten des Westfälischen Arbeitskreises „Phraseologie/Parömiologie“*. Bochum. (203–217).
- Grzybek, P. (1997a): «Anmerkungen zu Obsoletheit und Bekanntheit traditioneller kroatischer Sprichwörter. I: Die „Poslovice“ von Gjuro Daničić (1870).» In: Damjanović, St. (Hg.), *Prvi hrvatski slavistički kongres*. Zagreb. (Im Druck).
- Grzybek, P. (1997b): «Anmerkungen zu Obsoletheit und Bekanntheit traditioneller kroatischer Sprichwörter. II: Die „Hrvatske narodne poslovice“ von Vicko Juraj Skarpa (1909).» In: *Suvremena lingvistika* (39). (Im Druck).
- Grzybek, P. (1997c): «Anmerkungen zu Obsoletheit und Bekanntheit traditioneller kroatischer Sprichwörter. III: Die „Sbirka narodnih posloviceh, rieči i izrazah“ von Mijat Stojanović (1866).» In: *Narodna umjetnost*, (34/1); 201–223.
- Grzybek, P. (1997d): «Parömiologie und Quantitative Linguistik.» In: Durčo, P. (Hg.) *Europhras '97*. Bratislava. (Im Druck)
- Grzybek, P.; Chlosta, Ch. (1993): «Grundlagen der empirischen Sprichwortforschung.» In: *Proverbium. Yearbook of International Proverb Scholarship. Vol. 10*. The University of Vermont. (89–128).
- Grzybek, P.; Škara, D.; Heyken, Z. (1993): «Der Weisheit der Gasse auf der Spur. — Eine empirische Pilotstudie zur Bekanntheit kroatischer Sprichwörter.» In: *Zeitschrift für Balkanologie*, 29/2; 85–98.
- Kekez, J. (1984): *Poslovice i njima srodni oblici*. Zagreb.
- Kekez, J. (1986): *Svaki je kamen da se kuća gradi. Hrvatske poslovice sakupljene u naše dane po književnim i jezikoslovnim djelima nastalima od 12. do 19. stoljeća*. Zagreb. [Osijek, I, 1990].

- Kempgen, S. (1995): *Russische Sprachstatistik. Systematischer Überblick und Bibliographie*. München.
- Köhler, R. (1982): «Das Menzerathsche Gesetz auf Satzebene.» In: Lehfeldt, W.; Strauss, U. (eds.), *Glottometrika 4*. Bochum. (103–113).
- Köhler, R. (1984): «Zur Interpretation des Menzerathschen Gesetzes.» In: Boy, J.; Köhler, R. (eds.), *Glottometrika 6*. Bochum. (177–183).
- Krikmann, A. (1967): «Keelestatistikat Eesti vanasxnadest», in: *Emakeele seltsi aamstaraamat* (13); 127–154.
- Röhrich, L.; Mieder, W. (1977): *Sprichwort*. Stuttgart.
- Norrick, N. R. (1985): *How Proverbs Mean. Semantic Studies in English Proverbs*. Berlin etc.
- Пермяков, Г. Л. (1979): «Грамматика пословичной мудрости.» In: Пермяков, Г. Л., *Пословицы и поговорки народов Востока*. Москва. (7–68). [Deutsch: «Die Grammatik der Sprichwörterweisheit.» In: *Kodikas Code / Ars Semiotica* (3/4), 1984; 295–344.]
- Peukes, G. (1977): *Untersuchungen zum Sprichwort im Deutschen*. Berlin u. a.
- Тарланов, З. К. (1982): *Очерки по синтаксису русских пословиц*. Ленинград.
- Черкасский, М. А. (1968): «Опыт построения функциональной модели одной частной семиотической системы (пословицы и афоризмы).» In: *Филологический сборник* (8/9); 474–485. [Deutsch: «Versuch der Konstruktion eines funktionalen Modells eines speziellen semiotischen Systems.» In: *Kodikas Code / Ars Semiotica* (3/4), 1984; 363–377.]