

Peter Grzybek

## Versuchen wir einmal, die Kräfte aus dem Gleichgewicht zu bringen...

Quantitative Aspekte von Puškins *Eugenij Onegin* und *Domik v Kolomne*

*Zusammenfassung:* Ausgehend von der Grundannahme, daß sich (nicht nur sprachliche) Texte als selbstorganisierende Systeme verstehen lassen, muß eine der vorrangigen Fragen der Untersuchung der diesen Systemen zugrundeliegenden Gesetzmäßigkeiten gelten, auf deren Grundlage sich die Frage nach der Individualität der realisierten Texte in völlig anderem Lichte darstellt. – Im Bereich der Sprache sind mittlerweile eine Reihe von Gesetzmäßigkeiten z.B. über Wort- und Satzlänge, deren Vorkommenshäufigkeiten und wechselseitige Abhängigkeit relativ gut untersucht, auch wenn es nach wie vor vollkommen offen ist, welche Faktoren interner und externer Natur modifizierend einwirken. Im vorliegenden Beitrag soll vor diesem allgemeinen Hintergrund eine spezifische Frage an einem konkreten Text beleuchtet werden, nämlich: was passiert, wenn ein mehr oder weniger "etabliertes" System aus dem Gleichgewicht geworfen werden soll? Beispielshalber werden zwei poetische Texte von A.S. Puškin detailliert analysiert – was im System kippt, was nicht?

*Summary:* Based on the assumption that (not only linguistic) texts may be conceived of as self-organizing systems, one of the most important questions in the study of these systems must focus on the underlying regularities. On this basis, the individuality of texts appears in a completely different light. – In the realm of language, we know by now quite a number of such regularities, e.g. as to word length, or sentence length, as well as to their frequencies and interrelations, although we do not yet know how particular factors modify these regularities. In the present contribution, a specific question shall be dealt with in detail: what happens, if one attempts to unbalance a more or less "settled" system? By way of an example, two texts by A.S. Puškin shall be analyzed in detail – what in the system remains stable, what doesn't?

## 0. Einleitung: Der Herbst von Boldino

Ende August 1830 reiste der damals 31jährige Aleksandr Sergeevič Puškin von Moskau nach Boldino (im Gouvernement Nižnij Novgorod), um dort ein kleines Dorf von 200 Seelen als Teil seines väterlichen Erbgutes zu übernehmen; dieses hatte er aus Anlaß seiner bevorstehenden Heirat als materielle Grundlage erhalten. In Paris brach zu dieser Zeit die Julirevolution aus, in Moskau die Cholera – mit diesem polemischen Vergleich beschreibt Lotman (1981: 242) die zeitlichen Umstände. Doch Puškin wußte zu diesem Zeitpunkt noch nicht, daß diese Epidemie zur Verhängung einer Quarantäne führen sollte, die ihn unerwartet drei Monate lang in Boldino festhalten sollte. Als Ergebnis dieser Umstände verbrachte Puškin den berühmten „Herbst von Boldino“, eine ungemein produktive Schaffenszeit: die Beendigung seines berühmten Versromans *Evgenij Onegin*, an dem er mehrere Jahre gearbeitet hatte, bedeutete den Endpunkt eines wichtigen Abschnitts seiner dichterischen Entwicklung – gleichzeitig markierten die sogenannten „kleinen Tragödien“ sowie die *Povesti Belkina* (Die Erzählungen Belkins) den Beginn einer neuen Periode. Ebenfalls in diese Zeit fällt das Entstehen eines Poems, das in der Literaturwissenschaft im Spektrum von burleskem Scherzpoem bis hin zu poetologischer Metafiktion eingeordnet worden ist, *Domik v Kolomne* (Das Häuschen in Kolomna). Nicht zufällig hat sich gerade dieses Poem später aus Sicht der russischen Formalisten als Prototyp für eine „erschwerete Form“, als „Desorganisation der vorangegangenen Form“ (Jakobson 1919/21: 21) und damit als ein Prototyp für die Entblößung des Verfahrens dargestellt, als ein Werk, „das beinahe ganz aus der Beschreibung der in ihm selbst angewendeten Verfahren besteht“ (Šklovskij 1923: 16f.). Insofern ist es auch kein Zufall, daß Tynjanov (1927: 449) in seiner Abhandlung „Über die Literarische Evolution“ gerade auch *Domik v Kolomne* als ein Werk anführte, um die „Veränderlichkeit der Funktionen eines bestimmten formalen Elements“ exemplarisch zu belegen.

In der Tat beginnt der Text mit einer meta-poetologischen Passage, die nicht nur eine ausgeprägte Auto-Reflexivität erkennen läßt, sondern auch einen radikalen poetologischen Umbruch signalisiert:

*Četyrestopnyj jamb mne nadoel:  
Im pišet vsjakoj. Mal'čikam v zabavu  
Pora b ego ostavit'. Ja chotel  
Davnym-davno prinjat'sja za oktavu.*

Vierfüßige Jamben reizen mich nicht mehr:  
So schreibt ein jeder. Knaben zum Spaß lassen  
sollte man sie. Schon länger wollt' ich sehr  
Poeme nur in achtzeilige Strophen fassen.

Der vierfüßige Jambus, den Puškin hier vorgibt, leid zu sein, war Anfang des 19. Jahrhunderts die typische Versform der russischen Poesie schlechthin: 55% der Verse Puškins, aber z.B. auch 53% der Verse Lermontovs, oder 68% der Verse Baratynskijs waren vierfüßige Jamben (Gasparov 1984: 107). Vor dem Hintergrund dieser vermeintlichen Absage an den vierfüßigen Jambus einerseits, und in Anbetracht der Tatsache, daß schon drei Jahre später, 1833, z.B. der *Mednyj vsadnik* (Der eherne Reiter) in genau diesem Versmaß entstand, andererseits, läßt sich diese Aussage selbst als ein literarischer Kunstgriff verstehen, und vieles spricht dafür, daß es Puškin insgesamt weniger um eine allgemeine Absage an den vierfüßigen Jambus ging denn um den Ausdruck des Überdresses bei der Arbeit an dem ja genau in diesem Versmaß geschriebenen *Evgenij Onegin* (Semjonow 1965: 113f.).

Dennoch war die Einführung des fünffüßigen Jambus „eines der zentralen Ereignisse dieser Epoche der russischen Versfüßung“ (Gasparov 1984: 115), das vor dem Hintergrund der Entwicklungen in der gesamten europäischen Poesie zu sehen ist: Während dieses Versmaß in der italienischen, der englischen und der in dieser Hinsicht stark von der englischen beeinflussten deutschen Poesie eine dominante Stellung einnahm, wurde es in der französischen Poesie als altmodisch empfunden. Erst als der Einfluß der französischen Poesie auf die russische zurückging und sich eine stärkere Orientierung (auch) an deutscher und englischer Poesie bemerkbar machte, stieg Anfang des 19. Jahrhunderts das Interesse am fünffüßigen Jambus. Bei der Verwendung dieses Versmaßes nicht nur in lyrischen, sondern auch in epischen Texten spielten deshalb vor allem das (strophenlose, hauptsächlich in Distychen geschriebene) englische sowie das italienische, überwiegend in oktavischen Strophen geschriebene Poem eine besondere Rolle.

Vor diesem Hintergrund gewinnt die in der oben zitierten Passage hervorgehobene metrisch-strophische Verbindung von fünffüßigem Jambus und achtzeiliger Oktave besondere Bedeutung: Sie ist insofern bemerkenswert, als sowohl die Oktave als strophische Form als auch die jambische Metrik nicht nur bei Puškin, sondern in der russischen Verskunst des frühen 19. Jahrhunderts schlechthin ein Novum waren. Dazu ist zu sagen, daß im Werk Puškins

durchaus auch schon vor dem sog. Herbst von Boldino sowohl Oktaven als auch Jamben vorgekommen waren – allerdings waren die Oktaven in diesem Fall ohne jambisches Metrum, und wenn es einen fünffüßigen Jambus gab, dann war dieser nicht in oktavischer Strophenform.

Doch nicht genug, daß Puškin sich mit dieser einleitenden Textstelle des *Domik v Kolonne* explizit von vierfüßigen Jamben (und damit auch und gerade von eigenen, in diesem Versmaß geschriebenen Texten) abgrenzt; im weiteren Verlauf thematisiert er auch noch die funktionale Bedeutung der Zäsur in fünffüßigen Jamben (vgl. Vickery 1971): letztere stellen sich ohne Zäsur, wie Jakobson (1919/21: 21) bemerkt, nur für spätere, mit dieser Form dann bereits vertraute Leser als glatt und leicht dar, während sie für Puškin selbst eine erschwerte Form darstellen:

*Priznat'sja vam, ja v pjatistopnoj stročke  
Ljublju cezuru na vtoroj stope.  
Inače stich to v jame, to na kočke,  
I chot' ležu teper' na kanape,  
Vse kažetsja mne, budto v trjaskom bege  
Po merzloj pašne mčus' ja na telege.*

Euch's zu gestehen, in der fünffüßigen Zeile  
Lieb ich die Zäsur nach dem zweiten Fuß.  
Sonst fährt der Vers über Stock und Stein,  
Und lieg ich jetzt auch auf dem Kanapee,  
Scheint mit doch immer, daß in holperigem Lauf  
Über einen gefrorenen Acker ich im Pferdewagen sause.

In der Tat ist für die fünffüßigen Jamben bei Puškin bis zum Jahre 1830 eine Zäsur nach dem zweiten Versfuß charakteristisch, und die Verse sind dabei männlich, so. z.B. in seinem berühmten Gedicht „Ja vas ljubil“ (1829). Jetzt wendet er sich explizit von ihnen ab, und gerade in der entscheidenden Zeile *Ljublju cezuru na vtoroj stope* fehlt nicht nur die Zäsur, zudem ist die Klausel hier auch noch weiblich – was die meta-poetische Distanzierung um einen weiteren Faktor erhöht.

Doch lassen wir weitere poetologische und/oder literaturhistorische Details beiseite. Die angesprochenen Punkte sollten ausreichend sein, auch ohne weitere literaturgeschichtliche Details der dichterischen Entwicklung Puškins den Herbst von Boldino als eine Phase des literarischen Umbruchs zu markie-

ren, dessen Texte schon des öfteren Gegenstand literaturwissenschaftlicher Untersuchungen waren.

Gerade deshalb bietet es sich an, diese Texte als exemplarisches Analyse-material zu verwenden, um den genannten Umbruch nicht nur in einem allgemeineren Kontext zu beleuchten, sondern dies auch durch einen systematischen (systemtheoretischen) Zugang zu gewährleisten. Dabei kann es im hier gegebenen Zusammenhang freilich nicht darum gehen, grundsätzlich den Begriff des Systems oder die Grundlagen der modernen Systemtheorie in den jeweiligen geschichtlichen Dimensionen zu entfalten, und dies kann erst recht nicht im Hinblick auf die Beschreibung poetischer Texte und/oder poetologischer Normen erfolgen.

Was hier wohl geleistet werden kann und soll, ist ein Versuch, den von Puškin selbst thematisierten Umbruch vor systemtheoretischem Hintergrund zu betrachten und dabei die entscheidende Frage zu fokussieren, ob und in welchem Maße die intendierte Abkehr von einem normativen poetischen System dieses System zum Kippen bringt, welche Bestandteile des Systems dabei ins Wanken geraten oder aus dem Gleichgewicht geraten, und welche Elemente und Funktionen systemerhaltend bzw. systemstabilisierend wirken.

## 1. Systemtheoretische Voraussetzungen

Der ‚klassische‘, von Saussure ausgehende Strukturalismus verstand Sprache als ein „System, das nur seine eigene Ordnung kennt“, bei dem es relativ leicht sei, Externes von Internem zu unterscheiden; in der Folge galt dasselbe dann auch für nicht-sprachliche Systeme. Sprache bzw. ein sprachlicher Text war demnach also ein geschlossenes System („un système serré“). Dem klassischen strukturalistischen Paradigma zufolge – das von Elementen, Relationen und Strukturen eines Systems ausging – war also Sprache als ein abstraktes Netz von Relationen (Hjelmslev) zu begreifen. Und ausgehend von Begriffen wie ‚Regelhaftigkeit‘ und ‚Norm‘ stellte sich ein System als Summe aller von den Regeln der Sprache erzeugbaren Strukturen dar, wohingegen eine Norm tatsächlich realisierte Strukturen beinhaltet (Coseriu).

Während sich somit ein System zunächst als ein mehr oder weniger stabiler Gleichgewichtszustand zwischen den Elementen und Relationen auf verschiedenen Ebenen darstellte, ging man später davon aus, daß wir es weder in

der Sprache allgemein noch bei sprachlichen Texten mit stabilen Gleichgewichtszuständen zu tun haben, sondern mit einer mehr oder weniger stabilen Abfolge von Zuständen (einem „Fließ-Gleichgewicht“). Diese dynamische Sichtweise führte auch zu einem veränderten Verständnis von *Regelabweichungen*: Während diese im Rahmen einer Konzeption von Sprache als einem stabilen System als Verstoß gegen die Norm und infolge dessen als *Systemstörung* ausgelegt werden mußten, verstand man sie im Rahmen von Konzeptionen der Sprache als einem dynamischen, evolvierenden System in gewissem Sinne als systemimmanent und erachtete sie als notwendig für die ständige Entwicklung. Vor allem auch vor dem Hintergrund chaostheoretischer Ansätze verschob sich durch den Rückgriff auf das Konzept der Selbst-Regulation der Fokus stärker auf sog. Attraktoren, postulierte (End-)Zustände, auf die hin evolvierte Systeme zusteuern, und die letztendlich für die (Wieder-)Herstellung des Fließ-Gleichgewichts eines sich ständig in Veränderung befindlichen Systems verantwortlich sind. Vor diesem allgemein-theoretischen Hintergrund präsentieren sich natürlich Fragen der Selbst-Organisation nicht nur in der Natur, sondern auch in der Literatur und Kultur auf ganz neue Art und Weise (vgl. Hayles 1990, Hayles 1991, Werner 1999). Daß sich in diesem Zusammenhang insbesondere auch die Frage der literarischen Evolution vollkommen anders stellt, liegt auf der Hand, und genau dies soll an einem abgegrenzten Bereich im Hinblick auf die oben dargestellte Problematik im folgenden diskutiert werden.

Eines sei allerdings einschränkend vorausgeschickt: Wenn hier von einem Bruch im Schaffen Puškins die Rede ist, so ist das ohne Frage zutreffend; dennoch aber gilt es, sich der Relativität dieser Entwicklung bewußt zu bleiben. Denn de facto wäre es kaum angemessen, hier einen etwaigen radikalen Bruch mit der gesamten poetischen Tradition postulieren zu wollen – schließlich und letzten Endes bleibt Puškin auch weiterhin dem Paradigma der normativen Poetik des frühen 19. Jahrhunderts verhaftet.

Ungeachtet dessen soll an einem abgegrenzten Bereich die Frage der Selbst-Organisation zweier offenbar aufeinander bezugnehmender Texte exemplarisch untersucht werden. Bei diesem abgegrenzten Bereich soll es sich um nicht mehr und nicht weniger handeln als um das Verhältnis von Vers- und Wortlänge, wie es sich aufgrund der veränderten Anzahl von Hebungen im vier- bzw. fünffüßigen Jambus bei Puškin darstellt; als Material sollen dabei genau die beiden oben erwähnten Texte dienen, an denen Puškin fast gleich-

zeitig arbeitete, also *Eugenij Onegin* (EO) einerseits, und *Domik v Kolonne* (DK) andererseits.

Zu diesem Vergleich gilt es, folgendes zu sagen: Wie oben bereits angedeutet, sind die einzelnen Kapitel über einen Zeitraum von mehreren Jahren entstanden. Genau genommen ist es also eigentlich nur das letzte Kapitel des EO, das in etwa zeitgleich mit DK entstand. Ungeachtet dessen liegt den im folgenden dargestellten Analysen der vollständige Text des EO zugrunde, der sich aus acht Kapiteln zusammensetzt. Insofern ist also die Frage überaus berechtigt, inwiefern der Text des EO als homogener Text angesehen werden kann, bzw. ob sich nicht die einzelnen Kapitel durch eine innere Heterogenität auszeichnen, welche womöglich die Behandlung des EO als eines homogenen Textes nicht zuläßt. Diese Frage wird allerdings im hier gegebenen Zusammenhang nicht im Detail betrachtet werden können, und ihr wird an anderer Stelle ausführlicher nachzugehen sein (Grzybek 2003d). Im Rahmen der vorliegenden Untersuchungen wird also EO zunächst einmal unter der Annahme eines homogenen Textes betrachtet.

## 2. Verlänge

### 2.1. *Verlänge, in der Anzahl der Silben pro Vers berechnet*

Die Strophik des EO zeichnet sich durch einige Besonderheiten aus, die in engem Zusammenhang mit der Metrik stehen und nicht ohne Auswirkung auf die angesprochene Frage bleiben: Denn die sog. Oneginsche Strophe weist aufgrund ihrer Reimstruktur *AbAbCCddEffEff* in jeder Strophe acht 8-silbige Verszeilen mit männlicher Klausel (also: — — — — — — — —) und sechs 9-silbige Verszeilen mit weiblicher Klausel (also: — — — — — — — — —) auf. Insgesamt sollten sich also, wenn das Schema ausnahmslos realisiert wäre, in jeder Strophe 118 Silben auf jeweils 14 Verszeilen verteilen, was im Durchschnitt einer mittleren Verlänge von 8,43 Silben pro Vers (bei einer Standardabweichung von  $s = 0,50$ ) entsprechen würde. Allerdings weisen von den acht Kapiteln des EO nur die Kapitel I, II, V, VI und VII die Oneginsche Strophik voll ausgebildet auf, und es gibt die folgenden Abweichung:

1. im Kapitel III sind in der 3. Strophe nur die ersten acht Zeilen realisiert, die ausgelassenen sechs Verse sind durch Punkte markiert; außerdem weist das

„Pesnja devušek“ insgesamt 18 Verszeilen mit jeweils sieben Silben auf; auch die 79 Verse von Tatjanas Brief an Evgenij entsprechen nicht der Oneginischen Strophe;

2. im Kapitel IV lautet der letzte Vers der 37. Strophe „I odevalsja...“, sodaß sich hier ein 5-silbiger Vers ergibt;
3. im Kapitel VIII weist die 2. Strophe nur die ersten vier, die 25. Strophe nur die ersten acht Zeilen auf, und auch der Brief von Evgenij an Tat'jana entspricht mit seinen jeweils 30 acht- und neun-silbigen Versen nicht der Oneginischen Strophe.

Aufgrund dieser Umstände ergeben sich in den einzelnen Kapiteln geringfügige Abweichungen. Diese sind vom Umfang her insgesamt jedoch so gering, daß sie den o.a. theoretischen Durchschnittswert nicht beeinträchtigen: Er beträgt im gesamten *EO* in der Tat  $\bar{x} = 8,43$  Silben pro Vers ( $s = 0,50$ ).

Bei *DK* handelt es sich im Prinzip durchgehend um Oktaven; von den insgesamt 40 Strophen besteht lediglich die Strophe XXXVI aus sieben Verszeilen. Auch diese Oktaven weisen ein strenges Reimschema auf, und zwar jeweils stropfenweise alternierend, entweder *aBaBaBcc* mit 83 Silben pro Strophe (10-11-10-11-10-11-10-10) oder *AbAbAbCC* mit 85 Silben pro Strophe (11-10-11-10-11-10-11-11). Von den insgesamt 319 Verszeilen weisen somit 161 (50,47%) jeweils 10 Silben auf, die anderen 158 (49,53%) Verszeilen bestehen aus jeweils 11 Silben. Dies entspricht einer mittleren Verslänge von  $\bar{x} = 10,50$  Silben pro Wort ( $s = 0,50$ ).

Dieses Ergebnis, daß die Verse in *DK* insgesamt ca. zwei Silben mehr pro Vers aufweisen als *EO*, ist in Anbetracht der Anzahl von vier vs. fünf Jamben pro Vers nicht sonderlich überraschend und sagt letztendlich auch nicht viel aus: Die durchschnittliche, in der Anzahl der Silben berechnete Länge der Verszeilen ist in *DK* signifikant höher als in *EO* ( $z = 71,79$ ;  $p < 0,001$ )<sup>1</sup>; ohnedies sagen die Mittelwerte insofern nicht viel aus, weil der jeweilige Anteil von längeren und kürzeren Verszeilen in beiden Texten unterschiedlich ausfällt (s.o.)

Interessanter ist hingegen die Frage, wie sich in diesen Texten die Verslänge darstellt, wenn diese nicht in der Anzahl der Silben pro Vers, sondern in der Anzahl der Wörter pro Vers berechnet wird; auch hier wird die Berechnung sowohl für die beiden Gesamtexte als auch in Abhängigkeit von der jeweiligen Verslänge gesondert anzustellen sein.

## 2.2. Verslänge, in der Anzahl der Worte pro Vers berechnet

Bei der Berechnung der Wortanzahl pro Vers ist ein besonderes Problem zu berücksichtigen, insofern man bei einer graphematischen Wortdefinition im Russischen (wie in anderen slawischen Sprachen auch) von Wörtern auszugehen hat – in den meisten dieser Fälle handelt es sich hierbei um Präpositionen –, welche kein silbenbildendes Element (im Russischen also kein Element mit Vokalqualität) aufweisen. Diese Wörter werden mitunter einer eigenständigen Wortklasse von sog. 0-silbigen Wörtern zugeordnet (vgl. hierzu: Antić/Kelih 2003). Um die Auswirkung einer solchen Entscheidung unter Kontrolle zu halten, scheint es sinnvoll, die jeweiligen (in der Anzahl der Worte pro Vers berechneten) Verslängen sowohl unter Berücksichtigung als auch unter Nicht-Berücksichtigung sog. 0-silbiger Wörter als einer eigenständigen Wortklasse zu berechnen.

- (a) In *Evgenij Onegin* beträgt die Verslänge bei Berücksichtigung der 0-silbigen Wörter als einer eigenen Wortklasse im Durchschnitt  $\bar{x} = 4,29$  Worte pro Vers bei einer Standardabweichung von  $s = 1,06$ . Die durchschnittliche, in der Anzahl der Worte pro Vers berechnete Verslänge beträgt dabei in den 8-silbigen Versen  $\bar{x} = 4,26$  ( $s = 1,04$ ), in den 9-silbigen Versen  $\bar{x} = 4,34$  ( $s = 1,07$ ). – Werden die 0-silbigen Wörter nicht als eigenständige Wortklasse gerechnet, ist die mittlere Verslänge natürlich niedriger. Unter dieser Voraussetzung beträgt sie für den gesamten Text  $\bar{x} = 4,08$  ( $s = 0,96$ ), in den 8-silbigen Versen beträgt sie  $\bar{x} = 4,06$  ( $s = 0,95$ ), in den 9-silbigen Versen  $\bar{x} = 4,12$  ( $s = 0,96$ ). – Der Befund, daß in den 8-silbigen Versen im Durchschnitt weniger Wörter vorkommen als in den 9-silbigen Versen, entspricht durchaus der Erwartung; der Unterschied der durchschnittlichen Wortanzahl pro Vers ist zwar insgesamt gering, erreicht aber – unabhängig von der Art der Behandlung der 0-silbigen Wörter – das statistische 5%-Signifikanzniveau signifikant ( $p < 0,05$ ).

		Silben pro Vers		
		8-silbige	9-silbige	gesamt
mit 0-Silbern	$\bar{x}$	4,26	4,34	4,29
	$s$	1,04	1,07	1,06
ohne 0-Silber	$\bar{x}$	4,06	4,12	4,08
	$s$	0,95	0,96	0,96

Tab. 1: Durchschnittliche Wortzahl pro Vers in *EO*

(b) In *Domik v Kolonne* beträgt die Verslänge bei Berücksichtigung der 0-silbigen Wörter als einer eigenen Wortklasse im Durchschnitt  $\bar{x} = 5,53$  Wörter pro Vers ( $s = 1,10$ ). Die Verslänge beträgt dabei in den 10-silbigen Versen  $\bar{x} = 5,52$  ( $s = 1,10$ ), in den 11-silbigen Versen  $\bar{x} = 5,54$  ( $s = 1,09$ ). – Werden die 0-silbigen Wörter nicht als eigenständige Wortklasse gerechnet, beträgt die mittlere Verslänge für den gesamten Text  $\bar{x} = 5,23$  ( $s = 1,01$ ); in den 10-silbigen Versen beträgt sie  $\bar{x} = 5,24$  ( $s = 1,00$ ), in den 11-silbigen Versen  $\bar{x} = 5,23$  ( $s = 1,02$ ). Der Unterschied in der Verslänge zwischen den 10- und 11-silbigen Verse erweist sich unter beiden Bedingungen als nicht signifikant ( $z = 0,55$ ;  $p = 0,58$  bzw.  $z = 0,17$ ;  $p = 0,87$ ).

		Silben pro Vers		gesamt
		10-silbige	11-silbige	
mit 0-Silbern	$\bar{x}$	5,22	5,54	5,53
	$s$	1,10	1,09	1,10
ohne 0-Silber	$\bar{x}$	5,24	5,23	5,23
	$s$	1,00	1,02	1,01

Tab. 2: Durchschnittliche Wortzahl pro Vers in DK

Vergleicht man die durchschnittliche Verslänge beider Texte miteinander, so stellt sich heraus, daß sich die durchschnittliche Anzahl der Worte pro Vers in den beiden Texten insgesamt hoch signifikant voneinander unterscheidet ( $z = 39,94$  bzw.  $z = 37,98$ ;  $p < 0,001$ ); dies ist in Anbetracht der Erweiterung von vier auf fünf Jamben pro Vers (s.o.) freilich keine sonderliche Überraschung. Interessant ist jedoch, daß es bei der durchschnittlichen Anzahl der Worte pro Vers innerhalb der jeweiligen Texte (d.h. zwischen den 8- und 9-silbigen Versen einerseits, den 10- und 11-silbigen Versen andererseits) entweder keine signifikanten Unterschiede gibt (so in DK) oder daß dieser Unterschied relativ schwach signifikant ist (so in EO). Dies läßt sich als ein Hinweis darauf werten, daß die Wortlänge innerhalb dieser jeweiligen Verslängen deutliche Unterschiede aufweisen könnte. Dies führt zur Notwendigkeit, die Wortlänge in beiden Texten detaillierter zu untersuchen (s.u.). Bleiben wir zunächst aber noch bei der Untersuchung der Verslänge und gehen die Frage auch von einer zweiten Richtung her an.

2.3. Verslängenhäufigkeiten

Auch die Frage nach der Verslängenhäufigkeit fokussiert die in der Anzahl der Wörter pro Vers gemessene Verslänge. Ausgehend von der Tatsache, daß ein und derselbe Mittelwert aufgrund von sehr unterschiedlichen Verteilungen zustandekommen kann, stellt sich die Frage nach der Verslänge jedoch anders als bei der durchschnittlichen Verslänge, nämlich: Wie viele Verse mit zwei, drei, vier usw. Wörtern gibt es in den beiden Texten? Natürlich unterliegt auch eine Antwort auf diese Frage allgemeinen Prinzipien der Wortdefinition, insofern sich die Anzahl der Wortlängenklassen je nach Berücksichtigung von 0-silbigen Wörtern als einer eigenständigen Wortklasse verändert. Aus diesem Grunde enthalten die Tab. 3a/b neben den jeweiligen (in der Anzahl der Worte pro Vers berechneten) Verslängen in der zweiten Spalte die entsprechenden Häufigkeitswerte  $f[i]$ , und zwar sowohl unter Berücksichtigung als auch unter Nicht-Berücksichtigung der sog. 0-silbigen Wörter als einer eigenständigen Wortklasse.

Worte pro Vers (EO)	f[i] (m. 0)	f[i] (o. 0)	Worte pro Vers (DK)	f[i] (m. 0)	f[i] (o. 0)
2	82	95	2	–	–
3	1149	1417	3	6	7
4	2009	2125	4	45	67
5	1318	1200	5	114	130
6	528	345	6	97	81
7	129	46	7	45	28
8	15	2	8	9	6
9	–	–	9	3	–

Tab. 3a/b: Häufigkeit der in Anzahl der Worte pro Vers berechneten Verslänge in EO und DK (ohne und mit Berücksichtigung 0-silbiger Wörter als eigener Klasse)

Interessant ist nun die Frage, ob die Häufigkeitsverteilungen beider Texte ein und demselben Modell folgen und – wenn ja – ob und wie sie sich theoretisch modellieren lassen. Als Basis für einschlägige Überlegungen können die Arbeiten dienen, die Wimmer et al. 1994 bzw. Wimmer/Altmann 1996 zunächst

im Hinblick auf eine Theorie der Wortlängenhäufigkeit vorgelegt haben, die sich jedoch ohne weiteres auf andere sprachliche Ebenen übertragen läßt; in diesen Arbeiten lautet eine Grundannahme, daß die Verteilung der Vorkommenshäufigkeiten der jeweiligen Längensklassen nicht chaotisch organisiert ist, sondern daß die einzelnen Längensklassen in einem Proportionalitätsverhältnis zueinander stehen. Übertragen auf die Verslänge lautet die Frage also, ob die Häufigkeit, mit der 2-Wort-Verser vorkommen, in einer spezifischen Beziehung zu der Häufigkeit steht, mit der 3-Wort-Verser vorkommen, und ob diese ihrerseits in einer spezifischen Beziehung zur Häufigkeit von 4-Wort-Versern steht, usw. Mathematisch umformuliert: Läßt sich die Häufigkeit einer gegebenen Klasse zur Häufigkeit der jeweiligen Nachbarklasse in Form einer Relation  $P_x - P_{x-1}$  (d.h. als dynamisches System) verstehen?

Das Proportionalitätsverhältnis ist dabei nicht als konstant aufzufassen, sondern läßt sich mit der Funktion  $g(x)$  beschreiben, sodaß sich der folgende allgemeine Grundansatz ergibt:

$$(1) P_x = g(x)P_{x-1}.$$

Je nach Art der Funktion ergeben sich so unterschiedliche Verteilungsmodelle. Als Basisform setzen Wimmer/Altmann (1994: 115) für  $g(x) = ax^{-b}$ ; diese führt aufgrund der Differenzgleichung (1) zu

$$(2) P_x = \frac{a^x}{(x!)^b} \cdot P_0 \quad \begin{matrix} x = 0, 1, 2, \dots \\ a, b > 0 \end{matrix}$$

$$P_0^{-1} = \sum_{i=0}^{\infty} \frac{a^i}{(i!)^b}$$

d.h. zur sog. Conway-Maxwell-Poisson-Verteilung. Interessanterweise lassen sich die in Tab. 2a/b dargestellten Häufigkeiten alle mit ebendemselben Verteilungsmodell theoretisch beschreiben. Mit anderen Worten: Die Häufigkeit, mit der Verse mit einer bestimmten Anzahl von Worten vorkommt, ist stochastisch reguliert. Da in unseren Daten die Verslängensklassen  $x = 0$  und  $x = 1$  beide nicht besetzt sind, lautet die Formel für die entsprechend 2-verschobene Verteilung in Modifikation der obigen Formel wie folgt:

$$(3) P_x = \frac{a^{x-2}}{(x-2!)^b} \cdot P_0 \quad x = 2, 3, \dots$$

Bei der Anpassung dieser Verteilung an empirische Daten geht es nun darum, die Parameter (in unserem Fall:  $a$  und  $b$ ) so zu bestimmen, daß die empirischen und theoretischen Werte möglichst gut übereinstimmen. Dazu gibt es entsprechende Spezialsoftware (i.e. Altmann-Fitter 2000), die in iterativen Prozessen die Werte optimiert. Für *Evgenij Onegin* bzw. *Domik v Kolomne* wären demnach die in Tab. 4a/b aufgeführten Werte einzusetzen. Aufgrund dieser Werte ergeben sich die in der dritten Spalte der Tab. 2a/b dargestellten theoretischen Werte  $NP[i]$ . Die Güte der Anpassung läßt sich am  $\chi^2$ -Wert des  $\chi^2$ -Anpassungstests ablesen. Im Falle großer Stichproben – bei denen der  $\chi^2$ -Wert eher die Signifikanzschwelle überschreitet (s.o.) – verwendet man auch den sog. Diskrepanzkoeffizienten  $C$ , der sich als  $C = \chi^2 / N$  berechnet, und der für  $C < 0,02$  auf gute, für  $C < 0,01$  auf sehr gute Anpassungsergebnisse hinweist. Da wir es bei unseren Analysen ausnahmslos mit großen Stichproben zu tun haben, werden wir die Anpassungsgüte im folgenden ausschließlich mit Bezugnahme auf den Diskrepanzkoeffizienten überprüfen.

Wie der Tab. 4a/b zu entnehmen ist, stellt die Conway-Maxwell-Poisson-Verteilung sowohl in EO als auch in DK – unabhängig von der Art der Berücksichtigung der 0-silbigen Wörter – ein ausgezeichnetes Modell dar, die Häufigkeit der in der Anzahl der Worte pro Vers berechneten Verslängen theoretisch zu beschreiben.

Worte pro Vers (EO)	NP[i] (m. 0)	NP[i] (o. 0)	Worte pro Vers (DK)	NP[i] (m. 0)	NP[i] (o. 0)
2	84,64	98,90	2	–	–
3	1011,31	1274,32	3	4,48	6,61
4	2048,69	2257,14	4	49,56	68,71
5	1469,71	1252,35	5	110,59	126,27
6	504,79	304,95	6	96,75	84,25
7	97,92	39,20	7	43,56	27,40
8	12,95	3,14	8	11,71	5,75
9	–	–	9	2,35	–
<i>a</i>	11,95	12,88	<i>a</i>	11,0617	10,3877
<i>b</i>	2,56	2,86	<i>b</i>	2,3095	2,4988
<i>C</i>	0,0089	0,0063	<i>C</i>	0,0059	0,0010

Tab. 4a/b: Häufigkeit der in Anzahl der Worte pro Vers berechneten Verslänge in EO und DK (ohne und mit Berücksichtigung 0-silbiger Wörter als eigener Klasse)

Abb. 1a/b stellt für EO und DK die beobachteten und theoretischen Häufigkeiten in der Variante ohne Berücksichtigung der 0-silbigen Wörter als eigenständiger Wortlängenklasse, Abb. 1c/d mit Berücksichtigung als eigenständiger Klasse in anschaulicher Form dar.

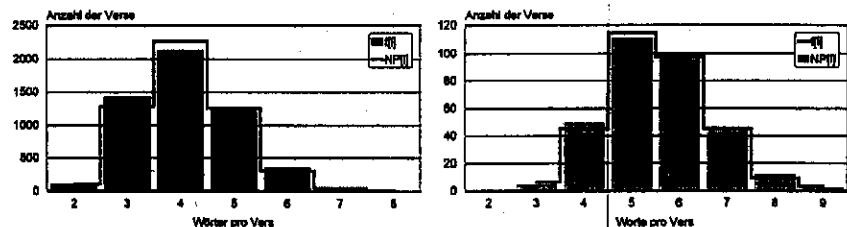


Abb. 1a: *Eugenij Onegin* (m. 0-Silber) Abb. 1b: *Domik v. Kolomne* (m. 0-Silbern)

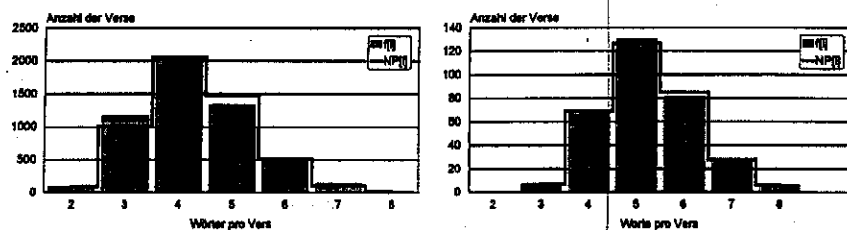


Abb. 1c: *Eugenij Onegin* (o. 0-Silber) Abb. 1d: *Domik v. Kolomne* (o. 0-Silber)

Abb. 1a-d: Verslängenhäufigkeiten

Die bisherigen Ergebnisse zusammenfassend können wir somit festhalten:

- (1) die Verse in DK werden im Vergleich zu EO im Durchschnitt deutlich länger;
- (2) die Verslängenhäufigkeiten verschieben sich in DK im Vergleich zu EO;
- (3) die Häufigkeiten der (in der Wortanzahl pro Vers berechneten) Verslängen sind sowohl in EO als auch in DK nicht chaotisch organisiert, sondern folgen einer bestimmten Regularität;
- (4) ungeachtet der Verschiebungen der Verslänge und ihrer Häufigkeiten in DK folgt die Organisation der Verslängenhäufigkeit(en) in beiden Texten ein und derselben Gesetzmäßigkeit.

Damit können wir uns als nächstes der Frage der Wortlänge zuwenden.

### 3. Wortlänge

#### 3.1. Wortlänge, in der Anzahl der Silben pro Wort berechnet

Die durchschnittliche Wortlänge berechnet sich im Grunde genommen als Quotient aus der Anzahl der im Text vorkommenden Wörter sowie der Anzahl der im Text vorkommenden Silben; auch die Berechnung der Wortlänge unterliegt somit Prinzipien der Wort- bzw. Silbendefinition. Dies ist im Hinblick auf die mittlere Wortlänge insofern von Bedeutung, als sich unter Berücksichtigung dieser 0-Silber als einer eigenständigen Wortklasse die Gesamtanzahl der Wörter erhöht (und folglich die durchschnittliche Wortlänge abnimmt), auch wenn der Umfang dieser Wortklasse in der Regel eher gering ist: in EO macht er mit 4,81%, in DK 5,39% des Wortbestandes.

Unter Berücksichtigung der 0-silbigen Wörter als eigener Wort(längen)klasse beträgt die mittlere Wortlänge in EO  $\bar{x} = 1,96$  ( $s = 1,06$ ). Im Hinblick auf die syllabotonische Struktur der beiden Texte scheint jedoch die Frage nach der durchschnittlichen (in Silben berechneten) Länge auf der Basis nur der zur Silbenstruktur beitragenden Wörter (also ohne Berücksichtigung der 0-silbigen Wörter als eigenständige Wortlängenklasse) die der Fragestellung angemessenere Vorgangsweise zu sein. Unter dieser Voraussetzung beträgt in EO die mittlere Wortlänge  $\bar{x} = 2,06$  Silben ( $s = 0,98$ ).

Im Vergleich dazu beträgt die mittlere Wortlänge unter Berücksichtigung der 0-silbigen Wörter als eigener Wort(längen)klasse in DK  $\bar{x} = 1,90$  Silben pro Wort ( $s = 1,02$ ), andernfalls  $\bar{x} = 2,01$  ( $s = 0,94$ ).

	m. 0-silbigen Wörtern		ohne 0-silbige Wörter	
	EO	DK	EO	DK
$\bar{x}$	1,96	1,90	2,06	2,01
s	1,06	1,02	0,98	0,94

Tab. 5: Wortlänge in EO und DK

Der Unterschied der Wortlänge im Vergleich zwischen EO und DK ist – unabhängig von der Art der Berücksichtigung der 0-silbigen Wörter – in beiden Fällen signifikant ( $z = 2,37$  bzw.  $z = 2,09$ ; jeweils  $p < 0,05$ ).

Damit läßt sich vorläufig resümieren, daß die poetologische Umstellung von vier- auf fünffüßige Jamben einerseits – wie zu erwarten war – dazu ge-



führt hat, daß die (in der Anzahl der Silben pro Vers berechnete) Verslänge sich in *DK* erhöht hat, daß aber die Verslängenhäufigkeiten ungeachtet dessen nach wie vor demselben Modell folgen; andererseits hat aber die (in der Anzahl der Silben pro Wort berechnete) Wortlänge in *DK* insgesamt abgenommen – eine Tendenz, die nicht ohne weiteres selbstverständlich ist und Anlaß zu weiterführenden Detailuntersuchungen gibt. Vor allem ergibt sich aus dieser Beobachtung die Notwendigkeit, die allgemeine Tendenz für die jeweiligen Verslängen gesondert zu überprüfen.

3.2. *Wortlänge, in Abhängigkeit von der (in Silben pro Vers gemessenen) Verslänge*

Anlaß zu dieser Überprüfung geben insbesondere aus der Quantitativen Linguistik bekannte Annahmen über den Zusammenhang von Konstrukt und (unmittelbaren) Konstituenten, so z.B. zwischen Satz- bzw. Teilsatzlänge und Wortlänge. So besagt das sog. Menzerath-Altmann-Gesetz in allgemeiner Form: „Je größer bzw. komplexer ein sprachliches Konstrukt ist, um so kleiner bzw. einfacher sind seine Konstituenten“ (Altmann 1980). Diesen Annahmen zufolge wäre zu erwarten, daß die Wortlänge bei zunehmender Verslänge abnimmt, wenn der Vers als Konstrukt verstanden und in Relation zum Wort gesetzt wird. Dabei wäre allerdings zu testen, welche Einheiten (also etwa Silben pro Vers oder Worte pro Vers) dem Konstrukt zugrunde zu legen sind. Hierzu fehlen bislang jegliche systematische Untersuchungen.<sup>2</sup>

Silben pro Vers	Eugenij Onegin				Domik v Kolonne					
	Anzahl der Verse	m.		o.		Anzahl der Verse	m.		o.	
		$\bar{x}$	<i>s</i>	$\bar{x}$	<i>s</i>		$\bar{x}$	<i>s</i>	$\bar{x}$	<i>s</i>
8	2965	1,88	1,00	1,97	0,94					
9	2246	2,07	1,11	2,18	1,03					
10						161	1,81	0,97	1,91	0,89
11						158	1,98	1,06	2,10	0,97

Tab. 6: Durchschnittliche Wortlänge in Abhängigkeit von der Verslänge

Schauen wir uns mit dieser Perspektive zunächst die mittlere Wortlänge an, und zwar in Abhängigkeit von den jeweiligen (in der Anzahl der Silben pro Vers berechneten) Verslängen. *Tab. 6* enthält die Mittelwerte und Standardabweichungen für beide Texte, aufgeschlüsselt nach den jeweiligen Verslängen.

Es ist deutlich erkennbar, daß eine allgemeine (d.h. text-überschreitende) Tendenz, der zufolge die Worte insgesamt bei zunehmender Verslänge kürzer würden, nicht existiert. In der Tat erreichen nicht einmal die Wortlängendifferenzen zwischen den 8-zeiligen Versen des *EO* und den 10-zeiligen Versen des *DK* ebenso wenig wie die zwischen den 9-zeiligen Versen des *EO* und den 11-zeiligen Versen des *DK* – unabhängig von der Art der Behandlung der 0-silbigen Wörter – ein statistisches Signifikanzniveau. Wohl aber zeichnet sich stattdessen eine Tendenz ab, die textintern für beide Texte in einem in ähnlichen Ergebnis resultiert: Demnach ist die mittlere Wortlänge innerhalb der jeweiligen Texte in den jeweils längeren Verszeilen, d.h. in den 9- bzw. 11-zeiligen Versen, länger als in den jeweils kürzeren, d.h. in den 8- bzw. 10-zeiligen Versen. *Abb. 2* stellt diese Tendenz in anschaulicher Form dar; zum Zwecke des Vergleichs zusätzlich eingezeichnet sind für beide Texte die mittleren Wortlängen, die sich bei Berücksichtigung der 0-silbigen Wörter als eigenständiger Wortklasse ergeben.

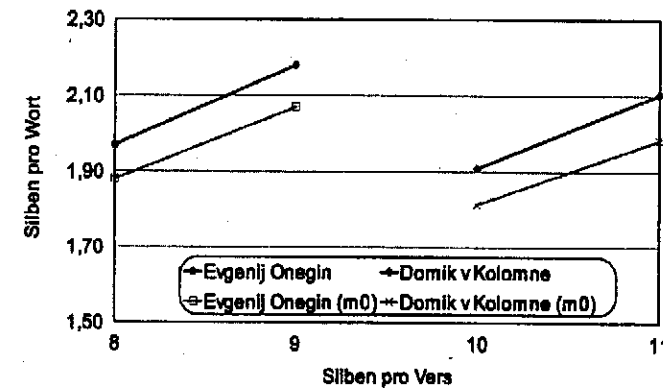


Abb. 2: Mittlere Wortlänge in Abhängigkeit von der Verslänge

In beiden Texten ist also ein und dieselbe Tendenz festzustellen, allerdings wird diese in *DK*, also in dem Text mit den längeren Versen, auf einem insgesamt niedrigeren Niveau realisiert. Zudem erreicht dieser Unterschied nur in

EO ein statistisches Signifikanzniveau ( $EO: z = 6,38; p < 0,001 - DK: z = 1,49; p = 0,14$ ); dennoch deckt sich dieser Befund mit der oben dargestellten Beobachtung, daß die mittlere Wortlänge in EO insgesamt etwas größer ist als in DK.

Im Hinblick auf die Erwartung abnehmender Wortlänge bei zunehmender Verslänge stellt sich also genau die entgegengesetzte Tendenz heraus, die offenbar innerhalb der jeweiligen Texte das Verhältnis von Wort- und Verslänge auf eine andere Art und Weise reguliert.

Es liegt nahe, daß die erwartete Tendenz deshalb nicht zum Ausdruck kommt, weil das Konstrukt ‚Vers‘ im Hinblick auf das Menzerath-Altmannsche Gesetz nicht zu seinen unmittelbaren (!) Konstituenten in Beziehung gesetzt wird, sodaß der obige Befund eher als Indiz für das sog. Arenssche Gesetz (Altmann 1983) zu werten wäre. Insofern stellt sich die Frage, ob sich die erwartete Tendenz unter der Bedingung bestätigen, daß das Konstrukt ‚Vers‘ nicht in der Anzahl der Silben, sondern in der Anzahl der Wörter pro Vers berechnet wird.

### 3.2. Wortlänge, in Abhängigkeit von der (in Worten pro Vers gemessenen) Verslänge

Vor dem Hintergrund der Tatsache, daß einerseits die Anzahl der Worte pro Vers gesetzmäßig organisiert ist, und daß andererseits eine einheitliche Tendenz in der Abhängigkeit der Silbenzahl pro Vers und der mittleren Silbenzahl pro Wort als gesichert gelten kann, stellt sich die Frage nach einem allfälligen Zusammenhang zwischen der Verslänge, definiert als Anzahl der Wörter pro Vers, und der mittleren (in der Anzahl der Silben gerechneten) Wortlänge. Auch bei dieser Betrachtung wirkt sich natürlich die Behandlung der 0-silbigen Wörter wesentlich aus, weshalb es sich anbietet, die Berechnungen auch im Hinblick auf diese Frage doppelt anzustellen.

Die Tab. 7a/b enthalten die durchschnittliche (in Silben gerechnete) Wortlänge in Abhängigkeit von der jeweiligen (in der Anzahl der Worte pro Vers berechneten) Verslänge.

Die Daten sind in Abb. 3 anschaulich dargestellt – deutlich zu sehen ist, daß die Art der Berücksichtigung 0-silbiger Wörter kaum einen Einfluß auf die zur Diskussion stehende Frage: In jedem Fall wird deutlich, daß in beiden Texten bei zunehmender Wortanzahl pro Vers die durchschnittliche, in der Anzahl der Silben pro Wort berechnete Wortlänge abnimmt.

Worte pro Vers (EO)	Silben pro Wort		Worte pro Vers (DK)	Silben pro Wort	
	m.0-Silbern	o. 0-Silber		m. 0-Silbern	o. 0-Silber
2	4,0732	4,0789			
3	2,8021	2,8064	3	3,444	3,4286
4	2,1056	2,1044	4	2,633	2,6381
5	1,6883	1,6910	5	2,098	2,1000
6	1,4104	1,4121	6	1,742	1,7346
7	1,2159	1,2050	7	1,514	1,5255
8	1,0417	1,0000	8	1,306	1,2917
9	–	–	9	1,148	–

Tab. 7a/b: Wortlänge in EO und D

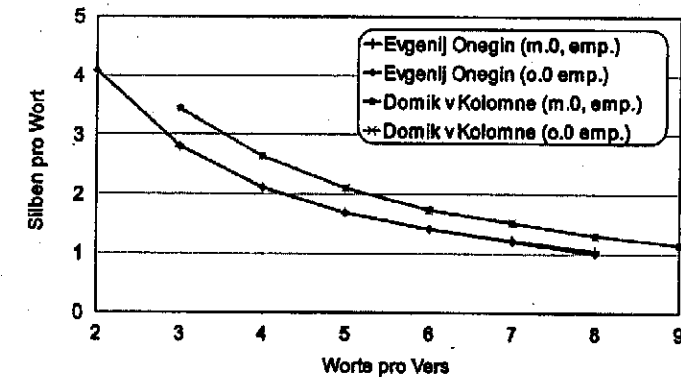


Abb. 3: Abhängigkeit der (in Silbenanzahl pro Wort berechneten) Wortlänge von der (in Wortanzahl pro Vers berechneten) Verslänge

Deutlich sind verschiedene Trends erkennbar:

1. die oben schon näher analysierte Verschiebung der Verslängen hin zu längeren Versen in DK;
2. im Vergleich zwischen EO und DK sind in den jeweiligen Verslängen die Wörter in DK durchgehend länger;
3. abnehmende Wortlänge bei zunehmender Verslänge in beiden Texten.

Gerade im Hinblick auf den letzten Punkt lautet die entscheidende Frage nunmehr: Läßt sich auch dieser Trend theoretisch modellieren und – wenn ja – folgen die beiden Texte ein und demselben Prinzip?

Wie oben bereits angesprochen, gibt es im Bereich der Quantitativen Linguistik allgemeine Grundsatzüberlegungen zur Lösung der hier zur Debatte stehenden Frage, die in dieser Form bislang kaum jemals auf Fragen der poetischen Textstruktur angewendet worden sind. In der Tat ist das Menzerath-Altman'sche Gesetz in der Regel nur auf den Zusammenhang von Satz- und/oder Teilsatz- und Wortlänge bezogen worden, und es wäre zu fragen, ob es sich auch im Hinblick auf den Zusammenhang von Vers- und Wortlänge als operational erweist.

Altmann hat eine mathematische Formalisierung dieser Tendenz in Form einer nicht-linearen Regressionsgleichung vorgeschlagen; in ihrer allgemeinsten Form (III) lautet die entsprechende Regressionsgleichung:

$$(4) \quad y = ax^b e^{-cx}$$

mit den beiden Spezialfällen (I) für  $b = 0, c \neq 0$  und (II) für  $b \neq 0, c = 0$ .

I	$b = 0$		$y = Ae^{-cx}$
II	$b \neq 0$	$c = 0$	$y = Ax^b$
III		$c \neq 0$	$y = Ax^b e^{-cx}$

Tab. 8: Das Menzerath'sche Gesetz in der Formalisierung von Altmann (1980)<sup>3</sup>

Ebenso wie im Fall der diskreten Häufigkeitsverteilungen geht es auch bei der nicht-linearen Regression darum, in den gegebenen Formeln die Parameter so zu optimieren, daß die sich aufgrund der Anpassung ergebenden theoretischen Werte möglichst gut mit den beobachteten übereinstimmen. Die Güte der Anpassung muß natürlich getestet werden; sie wird durch den sogenannten Determinationskoeffizienten  $R^2$  überprüft, ein Maß für den Zusammenhang zwischen den beobachteten und den theoretischen Werten im Intervall zwischen 0 und 1. Dabei betrachtet man in der Regel solche Anpassungen als gelungen, bei denen die theoretische Schätzung auf eine möglichst geringe Anzahl von Parametern reduziert werden kann. In diesem Sinne wird Formel (II) üblicherweise als „Standardfall“ für das Menzerath-Altman'sche Gesetz angesehen.

Für unsere beiden Texte erweist sich in der Tat der „einfache“ Fall (Formel II) als äußerst geeignet, die beschriebene Tendenz zu formalisieren; die in Tab. 6a/b aufgeführten theoretischen Werte ergeben sich auf der Grundlage der ebenfalls in der Tabelle aufgeführten Werte für die Parameter  $a$  und  $b$ .

Worte pro Vers (EO)	Silben pro Wort		Worte pro Vers (DK)	Silben pro Wort	
	m.0-Silbern	o.0-Silber		m.0-Silbern	o.0-Silber
2	4,0930	4,1051	2	–	–
3	2,7658	2,7642	3	3,4620	3,4482
4	2,0944	2,0879	4	2,6062	2,6029
5	1,6880	1,6795	5	2,0910	2,0927
6	1,4153	1,4059	6	1,7466	1,7511
7	1,2193	1,2097	7	1,5001	1,5061
8	1,0717	1,0620	8	1,3149	1,3218
9	–	–	9	1,1706	–
$a$	7,9989	8,0711	$a$	10,2392	10,0935
$b$	-0,9666	-0,9753	$b$	-0,9870	-0,9776
$R^2$	0,9996	0,9990	$R^2$	0,9995	0,9990

Tab. 9a/b: Theoretische Wortlängen in EO und DK

Wie zu sehen ist, liegt  $R^2$  in beiden Texten und unter beiden Bedingungen nur minimal unter dem Maximum von 1, nämlich bei  $R^2 \geq 0,999$  – das heißt, daß die durchschnittliche Wortlänge zu mehr als 99,9% durch die Verslänge determiniert ist! Abb. 4 veranschaulicht diesen Befund graphisch; aufgrund der hochgradigen Ähnlichkeit des Kurvenverlaufs unter der Bedingung von Berücksichtigung und Nicht-Berücksichtigung der 0-silbigen Wörter als eigenständiger Klasse bezieht sie sich lediglich auf die erste Bedingung, d.h. mit 0-Silbern als eigenständiger Wort(längen)klasse (siehe nächste Seite).

Es stellt sich somit im Ergebnis heraus, daß beide Texte – ungeachtet der beschriebenen Verschiebungen – ein und demselben Prinzip folgen, das sich schlagwortartig mit „Je länger die Verse, desto kürzer die Wörter“ beschreiben und theoretisch mit dem Menzerath-Altman'schen Gesetz modellieren läßt.

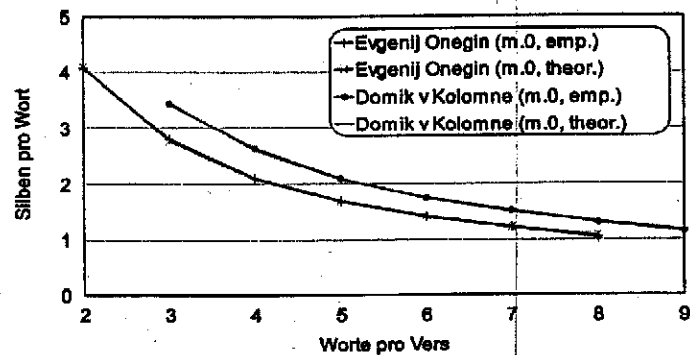


Abb. 4: Theoretische Modellierung der Abhängigkeit von (in Silbenanzahl pro Wort berechneter) Wortlänge und (in Wortanzahl pro Vers berechneter) Verslänge

### 3.3. Wortlängenhäufigkeiten

Aus dieser Beobachtung heraus stellt sich nunmehr eine nochmals weiterführende und abschließende Frage, nämlich die nach der silbischen Struktur des lexikalischen Bestandes in den jeweiligen Texten. Schauen wir uns aus diesem Grunde an, mit welcher Häufigkeit Wörter einer bestimmten Länge in den beiden Texten vorkommen, ob diese Vorkommenshäufigkeiten ebenfalls einer bestimmten Regularität folgen, und, wenn ja, ob diese Regularität in beiden Texten ein und dieselbe ist oder sich voneinander unterscheidet. Tab. 7 enthält die absoluten Werte ( $f_i$ ) und die entsprechenden Prozentwerte für beide Texte, zum Zwecke der Übersicht unter Berücksichtigung der 0-silbigen Wörter als eigenständiger Wortklasse.

Wie der Tab. 10 zu entnehmen ist, sind sich die Prozentsätze der Klassenbelegungen in beiden Texten relativ ähnlich. Dieser Eindruck wird recht deutlich durch die Abb. 5 bestätigt, welche die prozentualen Vorkommenshäufigkeiten der  $x$ -silbigen Wörter unter Berücksichtigung der 0-silbigen Wörter als eigenständiger Wort(längen)klasse für beide Texte in Form eines Balkendiagramms darstellt.

Silben pro Wort	Evgenij Onegin		Domik v Kolomne	
	$f(i)$	(in %)	$f(i)$	(in %)
0	1079	4,81	95	5,39
1	7210	32,15	605	34,30
2	7599	33,88	561	31,80
3	4894	21,82	404	22,90
4	1314	5,86	86	4,88
5	277	1,24	13	0,74
6	51	0,23		
7	4	0,02		
$\Sigma$	22428		1764	

Tab. 10: Vorkommenshäufigkeiten der Wortlängen

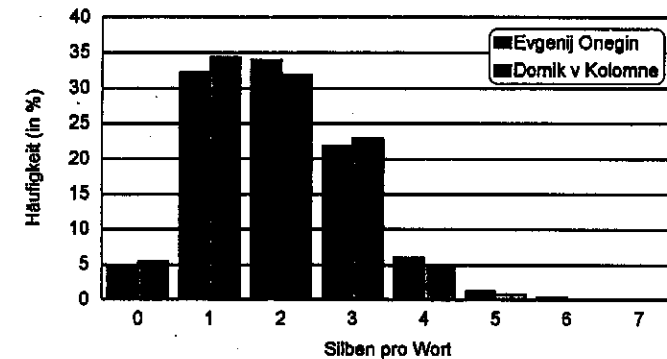


Abb. 5: Prozentuale Vorkommenshäufigkeiten der Wortlängen

Damit stellt sich nun wiederum die Frage, ob und wie sich die beiden Häufigkeitsverteilungen theoretisch beschreiben lassen. Nach Wimmer et al. (1994) bzw. Wimmer/Altmann (1996) lautet die Grundannahme, daß die Verteilung der Vorkommenshäufigkeiten der jeweiligen Längenklassen nicht chaotisch organisiert ist, sondern daß die einzelnen Längenklassen in einem Proportionalitätsverhältnis zueinander stehen. Setzt man, ausgehend von dem oben dar-

gestellten allgemeine Grundansatz (1)  $P_x = g(x)P_{x-1}$ , für  $g(x) = \frac{a-bx}{x}$ , so ergibt sich aufgrund der Differenzgleichung

$$(5) P_x = \frac{a-bx}{x} P_{x-1}$$

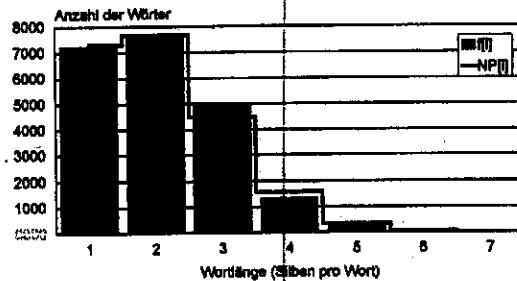
nach entsprechender Reparametrisierung die Binomialverteilung, und zwar für  $P_0 \neq 0$  deren Standardform (6), und für  $P_0 = 0$  die positive Binomialverteilung (7):

$$(6) P_x = \binom{n}{x} p^x q^{n-x} \quad x = 0, 1, \dots, n$$

$$(7) P_x = \frac{\binom{n}{x} p^x q^{n-x}}{1 - q^n} \quad x = 0, 1, \dots, n$$

Wie Grzybek (2003e) gezeigt ist, ist es genau dieses Verteilungsmodell (7), das sich im Hinblick auf die Wortlängenhäufigkeiten in EO als hervorragend geeignet erweist, wenn die 0-silbigen Wörter nicht als eigene Wort(längen)klasse berücksichtigt werden ( $C < 0,005$ ); die in Abb. 6 veranschaulichten Werte sind der Tab. 11 zu entnehmen.

x[i]	f[i]	NP[i]
1	7210	7292,41
2	7599	7657,41
3	4894	4467,04
4	1314	1563,54
5	277	328,36
6	51	38,31
7	4	1,92
n	7	
p	0,2593	
C	0,0045	



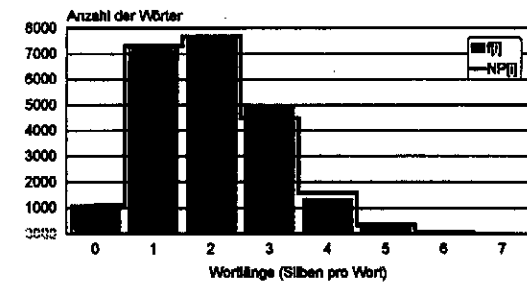
Tab. 11 / Abb. 6: Beobachtete und theoretische Wortlängenhäufigkeiten in EO (ohne 0-silbige Wörter)

Unter der Bedingung, daß die 0-silbigen Wörter als eigene Wort(längen)klasse berücksichtigt werden, stellte sich in der Analyse des EO von Grzybek (2003e) eine Modifikation der (positiven) Binomialverteilung als am besten geeignetes Modell heraus, nämlich die sog. erweiterte positive Binomialverteilung. Wie aus Formel (8) ersichtlich ist, beinhaltet diese die Erweiterung der gestutzten Verteilung um die Klasse  $x = 0$  und die anschließende Normierung aller Klassen:

$$(8) P_x = \begin{cases} 1 - \alpha & x = 0 & n \in \mathbb{N} \\ \alpha \binom{n}{x} p^x q^{n-x} & x = 1, 2, \dots, n & 0 \leq p \leq 1 \\ \frac{1 - q^n}{1 - q^n} & & q = 1 - p \\ & & 0 \leq \alpha \leq 1 \end{cases}$$

Bei der Anpassung dieses Verteilungsmodells an die Daten des EO ergeben sich die in Tab. 12 aufgeführten und in Abb. 7 veranschaulichten Werte:

x[i]	f[i]	NP[i]
0	1079	1079
1	7210	7292,70
2	7599	7657,43
3	4894	4466,89
4	1314	1563,43
5	277	328,32
6	51	38,31
7	4	1,92
n	7	
p	0,2593	
α	0,9519	
C	0,0045	

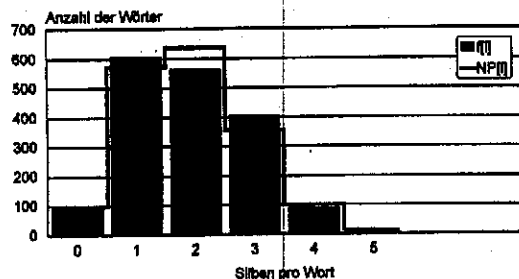


Tab. 12 / Abb. 7: Beobachtete und theoretische Wortlängenhäufigkeiten in EO (mit 0-silbigen Wörtern)

Die sich nun aufdrängende Frage ist, ob sich die beschriebenen Verteilungsmodelle auch für die Modellierung der Wortlängenhäufigkeit des DK als geeignet erweisen. Tab. 13 enthält die in Abb. 8 dargestellten Werte NP[i], die sich unter der Bedingung der Berücksichtigung 0-silbiger Wörter als eigener

Wort(längen)klasse bei der Anpassung der erweiterten positiven Binomialverteilung ergeben. Diese ist mit  $C < 0,02$  zwar als durchaus gut anzusehen, ohne Frage aber erkennbar schlechter als im Falle des EO – es erweist sich allerdings kein etwaiges anderes Modell als besser geeignet.

$x[i]$	$f[i]$	NP[i]
0	95	95
1	605	569,97
2	561	635,31
3	404	354,06
4	86	98,66
5	13	11,00
$n$		5
$p$		0,358
$\alpha$		0,946
$C$		0,0113*



Tab. 13 / Abb. 8: Empirische und theoretische Vorkommenshäufigkeiten der Wortlängen in DK (mit 0-silbigen Wörtern als eigenständiger Wortlängenklasse)

Anders als in EO – wo die Erweiterung der positiven Binomialverteilung (8) erst bei der Berücksichtigung der 0-silbigen Wörter als eigener Wort(längen)klasse notwendig wurde – muß diese bei DK auch zur Modellierung der Wortlängenhäufigkeiten herangezogen werden, wenn die 0-silbigen Wörter nicht als eigenständige Wortlängenklasse berücksichtigt werden. Es handelt sich hierbei also um die 1-verschobene erweiterte positive Binomialverteilung, deren Komplexität darin besteht, daß zunächst die Klasse  $x = 0$  gestutzt wird, dann die gesamte Verteilung um 1 verschoben, dann wiederum die Klasse  $x = 1$  modifiziert sowie anschließend das Ganze normiert wird, wie in (9) dargestellt:

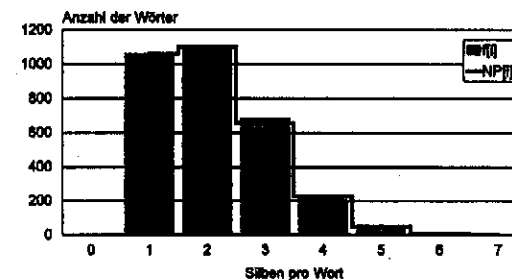
$$(9) P_x = \begin{cases} 1 - \alpha & x = 1 \\ \frac{\alpha \binom{n}{x-1} p^{x-1} q^{n-x+1}}{1 - q^n} & x = 2, 3, \dots, n+1 \end{cases}$$

$n \in \mathbb{N}$   
 $0 \leq p \leq 1$   
 $q = 1 - p$   
 $0 \leq \alpha \leq 1$

Wie dieser komplexe Umstand inhaltlich – d.h. poetologisch oder sprachhistorisch – zu interpretieren ist, kann an dieser Stelle nicht beantwortet werden. Es scheint auf den ersten Blick, auch vor dem Hintergrund des obigen Befundes zur Modellierung der Wortlängenhäufigkeiten unter Berücksichtigung 0-silbiger Wörter als eigener Wort(längen)klasse, nicht ausgeschlossen, daß der Text des DK ungeachtet der veränderten Versstruktur einem Modell der Wortlängenverteilung folgt, das einem anderen poetischen (bzw. versologischen) Modell (dem des EO) entlehnt ist und de facto dem neuen Versschema nicht mehr angemessen ist – daß es also einerseits obsolet ist, ohne daß sich andererseits ein anderes Modell herausgebildet hätte. Wenn diese Interpretation zuträfe, wäre das im Hinblick auf die Frage von Systemerhalt und Systemveränderung von weitreichender Bedeutung – doch wie gesagt, diese Interpretation ist aufgrund der geringen Materialbasis nur des einen Textes des DK nicht mehr als ein Denk-Spiel, und zukünftige Untersuchungen werden auf umfangreicherer Materialbasis sicherlich solidere Ergebnisse und Befunde bereitstellen können (Steinweiss 2003).

Auf jeden Fall aber ergeben sich die folgenden, in Tab. 9 dargestellten Anpassungsergebnisse, die alle ausgezeichnet sind:

$x[i]$	$f[i]$	NP[i]
1	605	605,00
2	561	574,92
3	404	371,14
4	86	106,48
5	13	11,46
$n$		4
$p$		0,301
$\alpha$		0,638
$C$		0,0044**



Tab. 14 / Abb. 9: Empirische und theoretische Vorkommenshäufigkeiten der Wortlängen in DK ohne 0-silbige Wörter als eigenständiger Wortlängenklasse)

## Zusammenfassung

Versuchen wir, die erhaltenen Ergebnisse in Form einer Zusammenfassung nochmals zusammenzuführen und im Hinblick auf die eingangs aufgeworfene Frage zu reflektieren.

Wie zu Beginn dieser Überlegungen dargelegt, stellt der sog. „Herbst von Boldino“ einen wichtigen Wendepunkt im Werk Puškins dar. Zwei der in dieser Zeit entstandenen Texte, *Evgenij Onegin* und *Domik v Kolonne*, werden im vorliegenden Aufsatz näher analysiert, und zwar – unter besonderer Berücksichtigung der metrischen Struktur – im Hinblick auf die Wort- und Verslänge.

Es stellt sich im Ergebnis heraus, daß beide Texte in bezug auf die folgenden Punkte exakt bestimmbarer Regularitäten folgen: (a) Häufigkeit der Verslängen, (b) Häufigkeit der Wortlängen, (c) Abhängigkeit von Wort- und Verslänge. Trotz unverkennbarer Unterschiede in der Textgestaltung handelt es sich bei beiden Texten um ein und dieselben Regularitäten. Besonders deutlich wird das bei der theoretischen Modellierung der Abhängigkeit von Wort- und Verslänge, die dem Menzerath-Altmannschen Gesetz folgt; dabei erweist sich die Einheit ‚Vers‘, wenn diese in der Anzahl der Wörter pro Vers berechnet wird, als eine linguistisch-poetische Einheit, die bislang als solche in einschlägigen quantitativen Untersuchungen nicht berücksichtigt wurde (vgl. Altmann 1996).

Während die Vorkommenshäufigkeiten der Verslängen in beiden Texten eindeutig ein und demselben Modell folgen, ist dies bei den Vorkommenshäufigkeiten der Wortlängen nicht, zumindest nicht so eindeutig, der Fall: Hier entsteht vielmehr der Eindruck, als ob ein durch die Veränderungen auf der strophischen und metrischen Ebene obsolet gewordenes Modell in *DK* nachwirkt, ohne daß sich ein neues Modell dieser Struktur angepaßt hätte. Ob dies tatsächlich der Fall ist, können nur umfangreichere Untersuchungen zeigen, deren Durchführung zu einer Einschätzung der obigen Befunde (vor allem auch im Hinblick auf eine Bewertung der Spezifik des Textes, der Poetik Puškins oder poetischer Texte allgemein) notwendig sind:

1. es gilt zunächst – vor allem in Anbetracht der mehrjährigen Entstehung des *EO* – die Homogenität des Gesamttextes bzw. dessen mögliche Heterogenität im Hinblick auf die einzelnen den Text konstituierenden Kapitel zu prüfen, d.h. der Frage nachzugehen, wie sich die beschriebenen Regularitäten in den einzelnen Kapiteln im Vergleich zum Gesamttext verhalten (Grzybek 2003d);

2. als nächstes gilt es, die sich hierbei einstellenden Befunde auf andere, vor allem auch in anderem Versmaß geschriebene Texte Puškins zu beziehen (Steinweiss 2003);
3. eine weitere Relativierung der obigen Befunde wäre durch eine Analysen von Übersetzungen des *EO* in andere Sprachen zu erwarten; hier versprechen Analysen kroatischer, serbischer und slowenischer Übersetzungen erste wichtige Anhaltspunkte (Grzybek 2003c);
4. schließlich wären entsprechende Analysen sowohl für andere Texte der russischen Poesie (vgl. Grzybek 2001) als auch poetische Texte anderer Sprachen durchzuführen; auch hier bieten Analysen kroatischer und slowenischer Texte von Mažuranić und Prešeren erste Einsichten (Grzybek 2003a,b).

Daß sich jedenfalls der aufgewiesene systemtheoretische Kontext zur Überprüfung dieser Annahmen eignet, scheint außer Frage.

## Anmerkungen

- 1 Zur Überprüfung des Mittelwertunterschieds bei kleineren Stichproben wird üblicherweise der sog. *t*-Test für unabhängige Stichproben verwendet. Da die dem *t*-Test zugrundeliegende Student-Verteilung gegen die Normalverteilung konvergiert und bei ca.  $N > 100$  hinreichend genau die Normalverteilung approximiert, führt in unseren Analysen ein *t*-Test zu demselben Ergebnis wie ein Normalverteilungstest; aus diesem Grunde wird im vorliegenden Aufsatz nur auf den *z*-Wert der Normalverteilungstests verwiesen, was die Angaben von Freiheitsgraden erspart.
- 2 Immerhin konnte an anderer Stelle eine solche Tendenz auch bereits an dem avantgardistischen Poem („Žuravi!“) von V. Chlebnikov aufgezeigt werden; demnach ergab sich eine Verallgemeinerung folgender Art: „Je länger die Verse, desto kürzer die Wörter“ (Grzybek 2001).
- 3 In einem noch weiter verallgemeinerten Ansatz gehen Wimmer et al. (2002) bzw. Wimmer/Altmann (2003) von der Formel

$$y = C e^{a_0 x} x^{a_1} e^{-a_2/x - a_3/(2x^2) - a_4/(3x^3) - \dots}$$

aus; für den Fall, daß  $a_1 < 0$ ,  $a_0 = a_2 = a_3 = \dots = 0$  erhält man die oben erwähnte Form des Menzerathschen Gesetzes.

## Literatur

- Altmann, Gabriel (1980). „Prolegomena to Menzerath's law“. In: Grotjahn, Rüdiger (Hg.). *Glottometrika 2*. Bochum: Brockmeyer, 1-10
- (1983). „H. Arens' ‚Verborgene Ordnung‘ und das Menzerathsche Gesetz“. In: Faust, Manfred (Hg.). *Allgemeine Sprachwissenschaft, Sprachtypologie und Textlinguistik. Festschrift für Peter Hartmann*. Tübingen: Narr, 31-39
- (1996). „The nature of linguistic units“. *Journal of Quantitative Linguistics* 3: 1-7
- Altmann-Fitter 2.1. (2000). RAM-Verlag Lüdenscheid / RST GmbH [Spezialsoftware]
- Antić, Gordana & Emmerich Kelih (2003). „On the question of so-called 0-syllable words in determining word length“. In: Grzybek, Peter (ed.). *Word Length Studies* [in print]
- Fomičev, S.A. (1984). „Oktavy ‚Domika v Kolomne‘ Puškina“. In: *Problemy teorii sticha*. Leningrad: Nauka, 125-131
- Gasparov, Michail Leonovič (1984). *Očerki istorii russkogo sticha. Metrika – ritmika – rifma – strofika*. Moskva: Nauka
- Grzybek, Peter (2001). „Zur Mikropoetik bei V. Chlebnikov“. Vortrag am Institut für Slawistik der Universität Salzburg
- Grzybek, Peter (2003a). „Anmerkungen zur Wort- und Versstruktur in France Prešerens *Krst pri Savici*“ [in Vorb.]
- (2003b). „Bemerkungen zur Wort- und Versstruktur in *Smrt Smail-age Cengića* von Ivan Mažuranić“ [in Vorb.]
- (2003c). „Evgenij Onegin in kroatischer, serbischer und slowenischer Übersetzung – Quantitative Aspekte“. Vortrag beim XIII. Internationalen Slawistenkongress 2003, Ljubljana
- (2003d). „Puškins Evgenij Onegin – Zur Frage der Homogenität des Textes und seiner Kapitel“. *Glottometrics* 5 [in Vorb.]
- (2003e). „Quantitative Aspekte slawischer Texte (am Beispiel von Puškins *Evgenij Onegin*)“. In: *Wiener Slawistisches Jahrbuch* 49 [im Druck]
- Harkins, William (1976). „The Place of ‚Domik v Kolomne‘ in Puškin's Creation“. In: Kodjak, Andrej & Kiril Taranovsky (eds.). *Alexander Puškin. A Symposium on the 175th Anniversary of His Birth*. New York: New York University Press, 196-205
- Hayles, Katherine N. (1990). *Chaos Bound. Orderly Disorder in Contemporary Literature and Science*. Ithaca–London: Cornell University Press
- (ed.) (1991). *Chaos and Order. Complex Dynamics in Literature and Science*. Chicago–London: University of Chicago Press
- Jakobson, Roman (1919/21). „Die neueste russische Poesie“. In: Stempel, Wolf-Dietrich (Hg.). *Texte der russischen Formalisten II*. München: Fink 1972: 19-135
- Lotman, Juri (1981). *Puschkin – Leben als Kunstwerk*. Leipzig: Reclam 1989 (21993)
- Semjonow, Jurij (1965). „Das Häuschen in Kolomna“ in der poetischen Erbschaft A.S. Puškins (=Acta Universitatis Upsaliensis. Studia Slavica Upsaliensis 3). Uppsala: Univ. Uppsala
- Steinweiss, Sabine (2003). *Wort- und Verlänge in russischen poetischen Texten*. Graz: Diplomarbeit [in Vorb.]
- Šklovskij, Viktor (1923). *Literatura i kinematograf*. Berlin
- Tynjanov, Jurij (1927). „Über die literarische Evolution“. In: Striedter, Jurij (Hg.). *Texte der russischen Formalisten I*. München: Fink 1969: 432-461
- Vickery, Walter (1971) s. Vikeri (1971)
- Vikeri, Val'ter (1971). „K voprosu o ritme cezurnogo pjatistopnogo jamba Puškina“. *International Journal of Slavic Linguistics and Poetics* 14: 134-175
- Werner, Hans C. (1999). *Literary Texts as Nonlinear Patterns: A Chaotic Reading of „Rainforest“, „Transparent Things“, „Travesty“, and „Tristram Shandy“* (=Gothenburg Studies in English 75). Göteborg: Acta Universitatis Gothoburgensis [zugleich Diss. Univ. Göteborg 1998]
- Wimmer, Gejza & Gabriel Altmann (1996). „The theory of word length distribution: Some results and generalizations“. In: Schmidt, Peter (Hg.). *Glottometrika 15*. Trier: wvt, 112-133
- (2002). „Unified derivation of some linguistic laws“. In: Grzybek, Peter (ed.). *Word Length Studies* [in print]
- Wimmer, Gejza; Köhler, Reinhard; Grotjahn, Rüdiger & Gabriel Altmann (1994). „Towards a theory of word length distribution“. *Journal of Quantitative Linguistics* 1: 98-106
- Wimmer, Gejza; Köhler, Reinhard & Gabriel Altmann (2002). „Unified derivation of some linguistic laws“. In: Köhler, Reinhard; Altmann, Gabriel & Raimund Piotrowskij (Hg.). *Handbuch der Quantitativen Linguistik*. Berlin: de Gruyter [im Druck]