

Sebastian Kempgen

Zur Zeitoptimierung der russischen Verbalmorphologie

In einem früheren Artikel (KEMPGEN 1995) wurde anhand der russischen Präsensformen erstmals überprüft, ob das Zipfsche "Principle of Least Effort" (ZIPF 1949) auch auf morphologischer Ebene gilt. Als Resultat konnte festgehalten werden: Der Grad, mit dem die russischen Präsensmorpheme der Zeitoptimierung entsprechen, beträgt rund 84%. Die Ergebnisse ermutigen unbedingt dazu, diese Annahme auch für andere Bereiche der russischen Flexion zu überprüfen, was in dem vorliegenden Beitrag für weitere finite Verbformen geschehen soll.

Die grundlegenden Annahmen seien an dieser Stelle nur ganz kurz re-sümiert – sie finden sich vollständiger in dem genannten Artikel. Natürliche Sprachen sind Codes mit variabler sog. Wortlänge: lange und kurze Wörter belegen dies augenfällig. Ein Code, der auf Effektivität des Zeitaufwandes hin konstruiert ist, muß „den häufigsten Symbolen die kürzesten Codewörter und, umgekehrt, den seltensten die längsten Codewörter zuschreiben“ (PADUČEVA 1961, 112). Das gleiche Prinzip kann, aber muß nicht zwangsläufig auch auf der Ebene aller Morphemtypen gelten. Plausibel ist dies dennoch: Von den verschiedenen Formen eines Flexionsparadigmas werden nicht alle Formen gleich häufig benutzt. Häufige Flexionsformen sollten deshalb von der Tendenz her die kürzeren Endungen aufweisen. Dabei ist freilich besonders bei einer Sprache wie dem Russischen zu bedenken, daß der Stamm von Wortformen u.U. selbst Veränderungen seiner Länge unterliegen kann: in bestimmten Formen können Vokale eingeschoben werden oder ausfallen, das sog. 'epenthetische l' kann den Stamm verlängern, es können Alternationen eintreten, bei denen ein Konsonant durch zwei ersetzt wird usw. Mit anderen Worten: wenn wir hier nur die Flexionsendungen für sich betrachten, so ist dies eine gewisse Abstraktion von den individuellen Verhältnissen konkreter einzelner Verben, gilt aber auf jeden Fall für den systemischen Aspekt der Flexionsendungen als solcher.

Wenn man die Flexionsmorphologie des Russischen auf das Prinzip der Zeitoptimierung hin untersucht, dann sollte ferner daran erinnert werden, daß die Silbenstruktur russischer Morpheme ebenfalls bestimmten, lange bekannten Prinzipien folgt: der Auslaut nominaler wie verbaler Wortstämme ist beispielsweise meist konsonantisch, der Anlaut der Endungen umgekehrt und konsequenterweise meist vokalisch – konsequent ist dies deshalb, weil auf diese Weise die übliche Silbenstruktur mit Abfolge von V(C)CV(C) gewahrt bleibt bzw. realisiert werden kann. Weiter ist

bekannt, daß die Flexionsendungen im Russischen meist einsilbig, seltener zweisilbig sind. Wenn man diese Faktoren berücksichtigt, dann muß es umso markanter erscheinen, wenn sich selbst unter diesen einschränkenden Rahmenbedingungen das “Principle of least effort” nachweisen läßt, bemerkbar macht.

In diesem Beitrag wollen wir konkret zwei Paradigmen untersuchen: das *Präteritum* und den *Imperativ*. Das Präteritum weist im Russischen insofern eine weitere Besonderheit auf, als es aus einem ehemaligen Partizip hervorgegangen ist und damit nominale grammatische Kategorien aufweist: Der Singular hat drei nach dem Genus unterschiedene Formen, der Plural eine für alle Genera gültige Form. Der russische Imperativ weist heute nur noch zwei synthetische Formen auf: eine für den Singular, eine für den Plural. Dies ist der Rest eines früher umfangreicheren Paradigmas. Daß es einzelne isolierte Verwendungsweisen weiterer Imperativformen gibt, braucht an dieser Stelle nicht berücksichtigt zu werden.

Daten über die Häufigkeit der einzelnen Flexionsformen liefert uns – wie in dem früheren Artikel – wieder das Wörterbuch von ŠTEINFELDT (1963, 141–167), da es im Wörterverzeichnis zu jedem Verb genau angibt, mit welcher Häufigkeit seine einzelnen Formen in der zugrundegelegten Stichprobe aufgetreten sind. Eine Summierung aller dieser Einzelwerte führt zu folgendem Ergebnis:

	masc. Sg.	fem. Sg.	neutr. Sg.	Plural	Σ
Absolut	9896	4729	2701	5260	22586
Relativ	0,438	0,209	0,120	0,233	1,0

Tab. 1: Frequenzen der russischen Präteritumsformen

	Singular	Plural	Σ
Absolut	1281	969	2250
Relativ	0,569	0,431	1,0

Tab. 2: Frequenzen der russischen Imperativformen

Ordnet man die Inhalte nach der Häufigkeit, mit der sie verwendet werden, so ergeben sich aus den genannten Daten für das Präteritum die folgende *Ränge*:

Präteritum		Imperativ	
Rang	Form	Rang	Form
1	masc. Sg.	1	Imp. Sg.
2	Plural	2	Imp. Pl.
3	fem. Sg.		
4	neutr. Sg.		

Tab. 3: Ränge der russischen Präteritum- und Imperativformen

Daß der russische *Imperativ* damit das Zipfsche Prinzip des geringsten Aufwandes befolgt, ist offensichtlich: die typische Singularendung lautet *-i*, die Pluralendung *-ite*. Dies braucht nicht also nicht weiter überprüft zu werden: Hier ist eine Korrelation offensichtlich mit ihrem Maximalwert gegeben, wie immer man sie auch berechnet.

Betrachten wir nun also die tatsächliche Länge der *Präteritumsmorpheme* im Russischen. In Buchstaben gezählt, ergibt sich folgendes:

masc. Sg.:	<i>-l</i> (oder \emptyset)	1 (0)
fem. Sg.:	<i>-la</i>	2
neutr. Sg.:	<i>-lo</i>	2
Plural:	<i>-li</i>	2

oder, geordnet:

- 1) masc. Sg.
- 2) -4) fem. Sg., neutr. Sg., Plural

Für das Präteritum ergibt sich als erste augenfällige Beobachtung, daß das Masculinum des Präteritums, die kürzeste Form, in der Tat die häufigste Form ist, während die übrigen Formen, die alle gleich lang sind, auf den weiteren Rängen (nach der Häufigkeit) folgen.

Wir wollen prüfen, ob dieser Augenschein einer präzisen Bewertung standhält. Dazu stellen wir die Ränge eines jeden Elementes in Bezug auf Häufigkeit und Länge zusammen, wobei den drei Elementen gleichen Ranges der Durchschnitt der ihnen zukommenden Rangzahlen 2 bis 4, also 3, zugeschrieben wird:

Form	Rang Häufigkeit	Rang Länge	D_i	D_i^2
masc. Sg.	1	1	0	0
fem. Sg.	3	3	0	0
neutr. Sg.	4	3	1	1
Plural	2	3	1	1
				Σ^2

Tab. 4: Ränge der russischen Präteritumsmorphe

In den beiden rechten Spalten sind die einfachen und die quadrierten Differenzen zwischen den beiden Rangreihen aufgeführt, die wir zur Berechnung des SPEARMANSchen Rangkorrelationskoeffizienten benötigen. Er ist für den Fall, daß Elemente mit gleichen Rängen auftreten, folgendermaßen definiert (vgl. ALTMANN/LEHFELDT 1980, 201; umgeformt bei SIEGEL 1976, 197):

$$r_s = \frac{K^3 - K - 6(\Sigma T_x + \Sigma T_y) - 6\Sigma D^2}{\sqrt{(K^3 - K - 12\Sigma T_x)(K^3 - K - 12\Sigma T_y)}}$$

Hierbei ist

$$T = \frac{t^3 - t}{12}.$$

K steht für die Anzahl der Elemente (d.h. hier 4), t für die Anzahl der Elemente mit jeweils gleichen Rängen. Wir erhalten für unseren Fall:

$$\begin{aligned} r_s &= \frac{4^3 - 4 - 6(0 + 2) - 6(2)}{\sqrt{(4^3 - 4 - 12(0))(4^3 - 4 - 12(2))}} \\ &= \frac{64 - 4 - 12 - 12}{\sqrt{(64 - 4 - 0)(64 - 4 - 20)}} \\ &= \frac{36}{48,9898} = 0.7348 \end{aligned}$$

Die Werte von r_s liegen im Intervall $\langle -1;1 \rangle$; da wir einen positiven Wert erhalten haben, deutet dies auf eine Tendenz zu gleichen Rängen in beiden Bereichen. Transformiert man den Bereich $\langle -1;1 \rangle$ durch die einfache Umrechnung $(x+1)/2$ in das Einheitsintervall $\langle 0;1 \rangle$, so erhalten wir den Wert $(0,7348+1)/2 = 0.8674$.

Mit anderen Worten: Der Grad, mit dem die russischen Präteritums- morpheme der Zeitoptimierung entsprechen, beträgt rund 87% und ist damit sogar noch etwas höher als der Wert für das Präsens. Beim Bau der Präteritums- morpheme des Russischen spielt die Optimierung auf möglichst geringen Zeitaufwand hin also eine wichtige Rolle.

Interessant ist es aber auch, sich rein hypothetisch einmal zu überlegen, unter welchen Bedingungen sich ein noch prägnanteres Ergebnis hätte ergeben können. Ein Blick auf die Häufigkeitsverhältnisse der Endungen, die alle zwei Grapheme lang sind, also fem.Sg., neutr.Sg. und Plural (vgl. Tab. 1), zeigt, daß die Endung des neutr. Sg. ruhig drei Grapheme lang sein dürfte – dies ergäbe eine fast perfekte Korrelation zwischen Häufigkeit und Länge. Daß das Neutrum aber – sozusagen ‘dennoch’ – nur zwei Grapheme lang ist, zeigt einfach nur, daß die Sprache gegenüber unnötiger Kürze toleranter ist als gegenüber hinderlicher Länge: die Kürzung zu langer Elemente ist sprachgeschichtlich weitaus häufiger zu beobachten als die Längung zu kurzer Elemente.

Das führt uns dazu, die Verhältnisse wenigstens andeutungsweise auch einmal sprachgeschichtlich zu beleuchten. Vor dem Ausfall der sogenannten “Halbvokale” (ca. 10.-12. Jh.) war die Endung des masc. Sg. zwei Grapheme und auch zwei Laute lang – genauso lang also wie alle übrigen Formen des Paradigmas auch. Es gibt nun keinen Grund anzunehmen, daß das masc. Sg. damals nicht auch schon die häufigste Form des Paradigmas gewesen sei. Der Ausfall des Halbvokals hat also in diesem Fall unbedingt dazu geführt, daß die mit Abstand häufigste Form auch die eindeutig kürzeste ist. Auch unter diesem Aspekt ist es bemerkenswert, daß die Orthographie des Russischen die traditionelle, ‘zu lange’ Schreibung noch fast 1000 Jahre bewahrt hat – sie wurde ja erst nach der Oktoberrevolution von 1917 reformiert und damit die Zeitoptimierung in der Schrift der Zeitoptimierung in der Aussprache angepaßt und beide Bereiche parallel gestaltet. Daß die Schrift einen weniger optimalen Zustand noch ein Jahrtausend länger als die Aussprache bewahren konnte, wird im übrigen seinerseits verständlicher, wenn man weiß, daß das Analphabetentum in Rußland ja Anfang des 20. Jh. noch bis zu 95% betrug: der Anpassungsdruck war in der Schrift also viel geringer als in der Aussprache.

Der vorliegende Beitrag wollte die Frage nach der Zeitoptimierung des Russischen auf morphologischer Ebene an zwei weiteren Ausschnitten aus der Flexionsmorphologie überprüfen. Mit dem Präsens, dem Imperativ und dem Präteritum gilt die geprüfte Tendenz jetzt nachweislich für alle finiten Verbformen des Russischen. Weitere Sprachen und weitere Formenparadigmen (z.B. der Nominalbereich) sollten entsprechend getestet werden. Hierfür liefert leider das Wörterbuch von ŠTEINFELDT (1963)

keine geeigneten Zahlen, da es nur jeweils Gesamtsummen für Singular und Plurals sowie für alle Kasus liefert, aber nicht für die einzelnen Kombinationen. Die benötigten Daten müßten also anders erhoben werden, beispielsweise aus den elektronisch verfügbaren Korpora des Russischen.

Literatur

- Altmann, Gabriel, Lehfeldt, Werner
1980 *Einführung in die quantitative Phonologie* (Quantitative Linguistics, vol. 7). Bochum: Brockmeyer.
- Kempgen, Sebastian
1995 “Codierung natürlicher Sprache auf morphologischer Ebene”. In: *Die Welt der Slaven* 40, 1, 1995, 52–57.
- Padučeva, Elena V.
1961 “Vozmožnosti izučenija jazyka metodami informacii”. In: Achmanova, Ol’ga S.; Mel’čuk, Igor’ A.; Padučeva, Elena V.; Frumkina, Revekka M. (Hg.), *O točnych metodach issledovanija jazyka*. Moskva: Moskovskij gos. universitet, 98–149.
- Siegel, Sidney
1976 *Nichtparametrische statistische Methoden. Mit einem Vorwort und Flußdiagramm zur Deutschen Ausgabe von W. Schüle*. Frankfurt/M: Fachbuchhandlung für Psychologie.
- Šteinfeldt, Evi
1962 *Häufigkeitwörterbuch der russischen Sprache. 2500 meistgebrauchte Wörter der modernen russischen Schriftsprache. Handbuch für Russischlehrer*. Moskva: Progress.
- Zipf, George Kingsley:
1949 *Human Behavior and the Principle of Least Effort. An Introduction to Human Ecology*. Cambridge, Mass.: Addison-Wesley.