

Jozef Mistrík

Exakte Typologie von Texten

Verlag Otto Sagner München · Berlin · Washington D.C.

Digitalisiert im Rahmen der Kooperation mit dem DFG-Projekt „Digi20“ der Bayerischen Staatsbibliothek, München. OCR-Bearbeitung und Erstellung des eBooks durch den Verlag Otto Sagner:

<http://verlag.kubon-sagner.de>

© bei Verlag Otto Sagner. Eine Verwertung oder Weitergabe der Texte und Abbildungen, insbesondere durch Vervielfältigung, ist ohne vorherige schriftliche Genehmigung des Verlages unzulässig.

«Verlag Otto Sagner» ist ein Imprint der Kubon & Sagner GmbH.

ARBEITEN UND TEXTE ZUR SLAVISTIK · 3
HERAUSGEGEBEN VON WOLFGANG KASACK

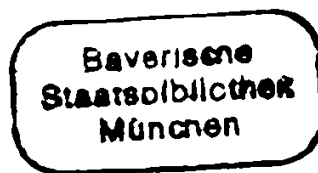
Jozef Mistrík

Exakte Typologie von Texten

1973

München · Verlag Otto Sagner in Kommission

Der Herausgeber dankt den Mitarbeitern des Slavischen Instituts der Universität zu Köln Frau Renate Menge-Verbeeck für die Übersetzung des Vorlesungsmanuskripts aus dem Slavischen und Frau Karin van Ackern für die Bearbeitung des Textes und die Herstellung der Druckvorlage. Aus technischen Gründen konnte die deutsche Fassung vom Autor vor der Drucklegung nicht durchgesehen werden.



Als Manuskript vervielfältigt
Alle Rechte vorbehalten

Gesamtherstellung Walter Kleikamp · Köln

7615622

VORBEMERKUNG

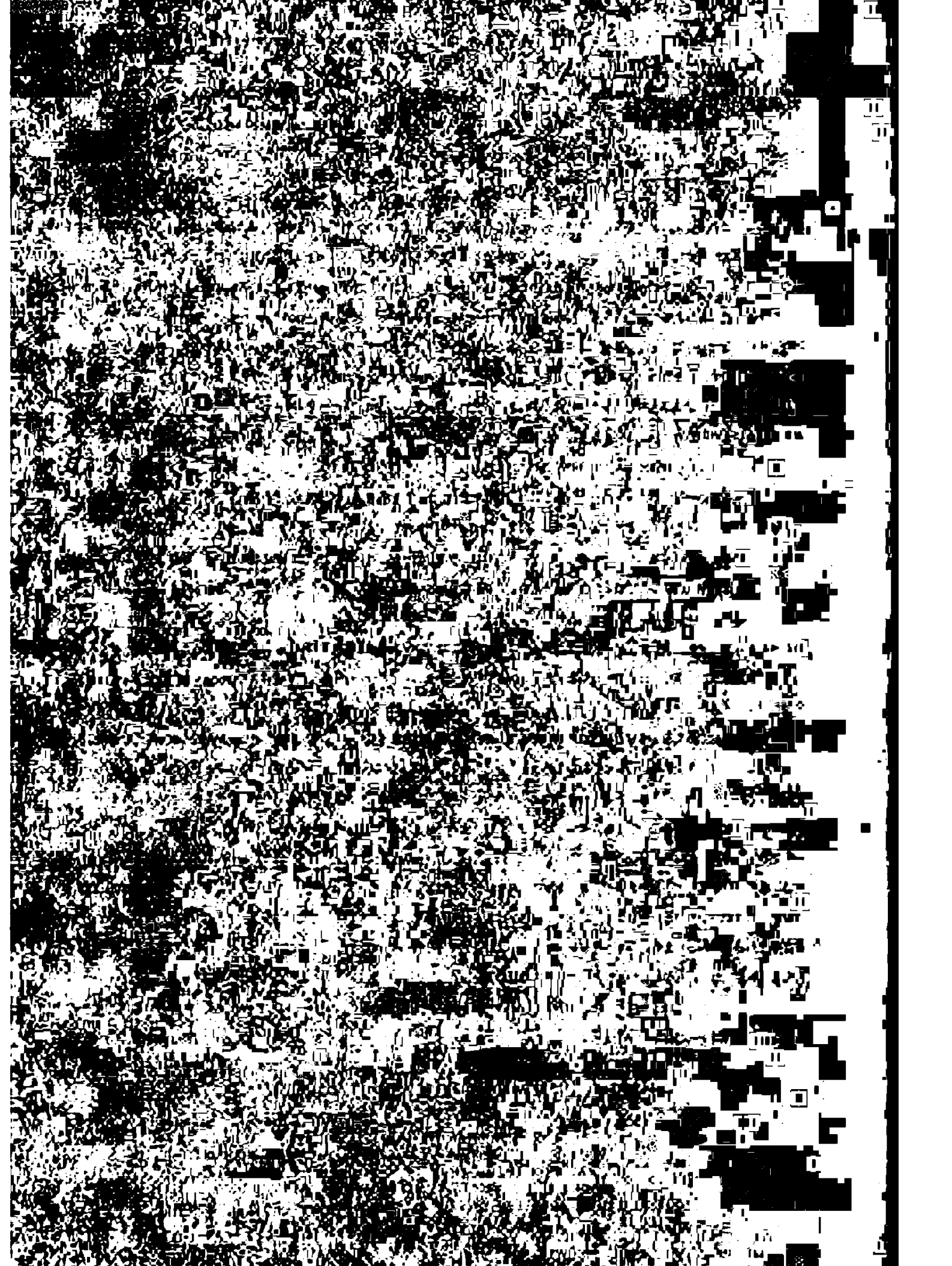
Die vorliegende Publikation von Professor Dr. Jozef Mistrík, Dr.scient., Bratislava, geht auf seine Vorlesungen und Vorträge über *Exaktné metódy pri tipológii textov* zurück, die er 1970 und 1971 an der Universität zu Köln und anderen Universitäten der Bundesrepublik Deutschland gehalten hat. Seine Lehr- und Forschungstätigkeit als Gastprofessor der Universität zu Köln stand im Rahmen der langjährigen wissenschaftlichen Beziehungen des Kölner Slavischen Instituts zu Universität und Slovakischer Akademie in Bratislava, die von Professor Dr. Dr.h.c. Reinhold Olesch aufgebaut worden sind.

Jozef Mistrík hat sich durch mehrere Monographien auf dem Gebiet der Stilistik und der mathematischen Linguistik ausgezeichnet. Wesentliche Voraussetzungen für das neue Buch, das vor allem methodische Anregungen gibt, enthalten seine Stilistik der slovakischen Sprache (*Štylistika slovenského jazyka*. Bratislava 1970. 415 S.) und sein Häufigkeitswörterbuch der slovakischen Sprache (*Frekvencia slov v slovenčine*. Bratislava 1969. 726 S.).

Die Herstellung der deutschen Fassung wurde dankenswerterweise durch eine Dachbeihilfe der Deutschen Forschungsgemeinschaft gefördert. Die Aufnahme in die Kölner Reihe der *Arbeiten und Texte zur Slavistik*, die durch einen Druckkostenzuschuß der Universität zu Köln ermöglicht wurde, ist ein Zeichen des Dankes an Jozef Mistrík für seine hiesige Lehrtätigkeit und ein Ausdruck der Verbundenheit der Universität zu Köln und der Universita Komenského in Bratislava, der Direktoren und Mitarbeiter des Kölner Slavischen Instituts und der slovakischen Gelehrten.

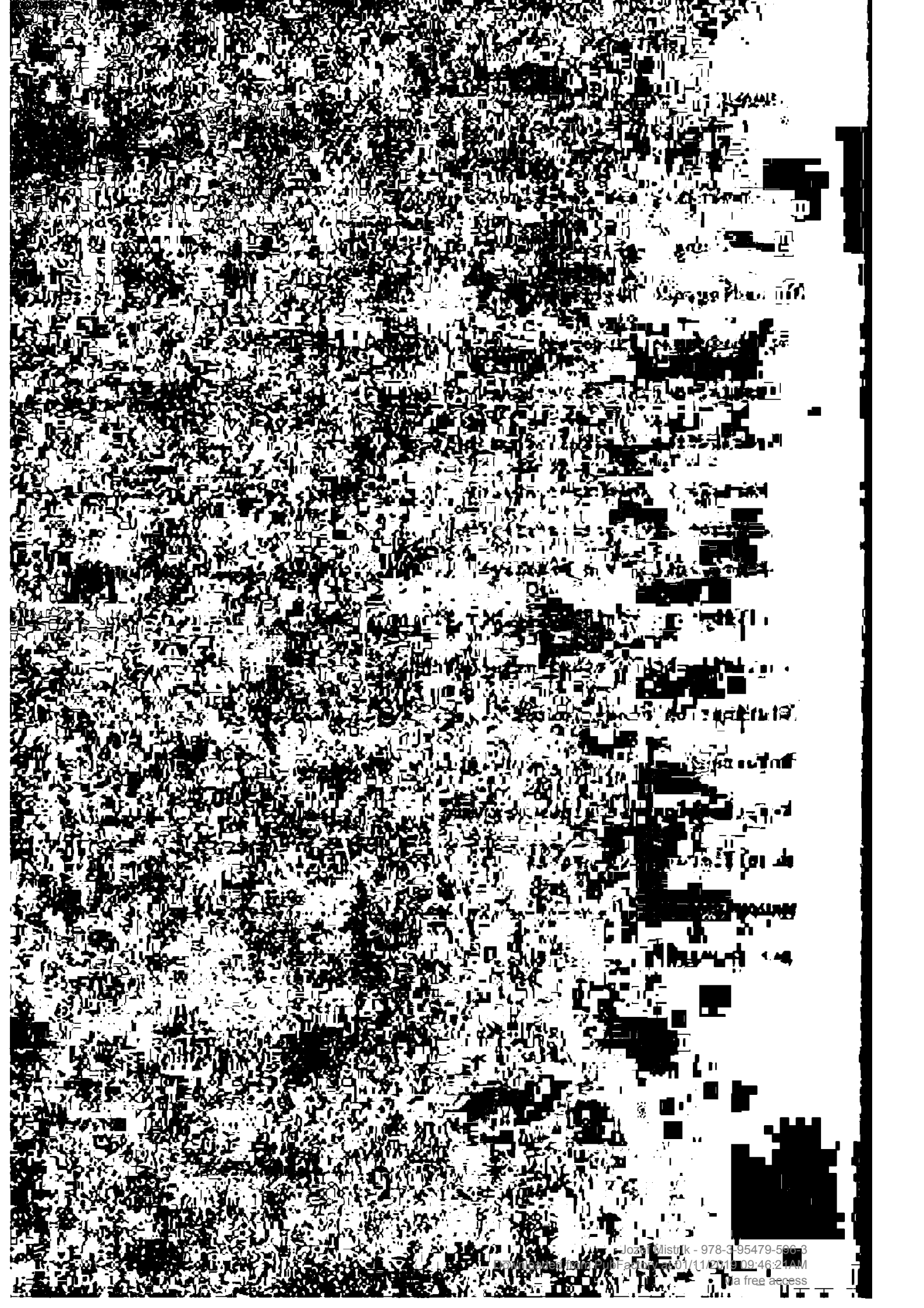
Köln, November 1973

Wolfgang Kasack



INHALT

<i>Vorbemerkung des Herausgebers</i>	5
Einleitung	9
Grundlegende Begriffe	10
Stilbildende Faktoren	15
Grundlagen exakter Methoden in der Typologie	26
Distribution lexikalischer Elemente	38
<i>Vertikale Wortdistribution</i>	38
<i>Horizontale Wortdistribution</i>	48
<i>Distribution von Präfixen</i>	65
Distribution morphologischer Elemente	72
Stilistische Relevanz der Anordnung syntaktischer Elemente	81
Glutination eines Textes	92
Rekurrenz	113
Verständlichkeit eines Textes	123
Modellierung der Genres	132
<i>Anmerkungen</i>	146
<i>Bibliographie</i>	151



EINLEITUNG

Wesen und Eigenschaften eines Textes waren schon immer Gegenstand des Interesses der Philologen. Es liegt ein besonderer Reiz darin, daß ein Text trotz der Kompliziertheit seiner Erscheinungen ein gewissen Gesetzmäßigkeiten folgendes Phänomen ist.

Traditionellerweise geht man beim Erkennen eines Textes intuitiv und empirisch vor. In den letzten zehn Jahren jedoch wurden bei der Typologisierung auch mathematische und quantitative Methoden angewandt. Auch heute werden diese exakten Methoden, die sich auf die meßbaren und sozusagen physisch feststellbaren Elemente eines Textes stützen, als modern und fortschrittlich empfunden, da sie noch nicht allgemein akzeptiert und verbreitet sind.

Die vorliegende Arbeit, die auf der Untersuchung vorwiegend deutscher und slovakischer Texte beruht, ist eine der ersten systematischen Beschreibungen dieser Methoden. Wir halten die Erforschung dieses Fragenkomplexes für noch nicht abgeschlossen. Unsere Untersuchungen sind vielmehr eine Einführung in die Arbeit mit exakten Methoden und illustrieren einige erprobte und bewiesene Ansätze.

GRUNDLEGENDE BEGRIFFE

• Die vorliegende Arbeit beschäftigt sich mit der Frage, wie sich ein Text typologisieren läßt. Der Begriff *Typologisierung* hängt eng mit den Begriffen *Analyse* und *Interpretation* bzw. mit dem Begriff *Modellierung* eines Textes zusammen. Diese enge Beziehung macht einleitend die Erklärung notwendig, in welchem Sinne wir die einzelnen Begriffe verstehen. Erläutern wir jedoch zunächst den Begriff *Text*.

Das Verständnis dieses Begriffes wird uns durch den Rückgriff auf die ursprüngliche Bedeutung des lateinischen Verbums *texo*, *texere* und des lateinischen Substantivums *textus* erleichtert: *texo*, *texere* heißt *weben*, *flechten*, *zusammenfügen*, *bauen*, *textus* heißt *Gewebe*, *Geflecht*, *Zusammenhang*, *Gefüge*. Ein Text ist also ein kompaktes Ganzes, dessen Inneres auf eine bestimmte Weise geflochten, d.h. aus der Fügung sprachlicher Elemente entstanden ist. Er ist eine in sich geschlossene sprachliche Äußerung im allgemeinen Sinne. Unter diesem Begriff fassen wir alle Arten und Typen künstlerischer und sachbezogener Äußerungen zusammen.¹ Mit dem Wort *Text* benennen wir die *unbekannte*, in keiner Weise analysierte oder interpretierte, die sprachlich *homogene* Masse, die als Ganzes einen bestimmten Sinn hat. Im Inneren des Textes kann eine gleichförmige, stereotype oder eine bunte, verschiedenförmige *Zeichnung* (d.h. *Gewebe*) vorliegen. Diese prägt den Charakter des Textes und ist die Grundlage zur Bestimmung der verschiedenen Textarten, die sich in erster Linie in künstlerische und sachbezogene Texte scheiden. Die künstlerischen weisen ein komplizierteres Gewebe als die sachbezogenen Texte auf. Berücksichtigt man das Gewebe nicht, so bleibt der Text eine zwar homogene, aber doch konturenlose Masse. Wir sprechen dann einfach von einem Text im weitesten Sinne des Wortes. Bevor man mit der Typologisierung, Analyse oder Interpretation eines Textes beginnt, muß man den Gegenstand tatsächlich als eine konturenlose, homogene Masse begreifen.²

Grundsätzlich entsteht ein Text dadurch, daß aus dem sprach-

lichen Inventar je nach Bedarf einige Elemente ausgewählt und nach bestimmten Gesetzen zusammengefügt werden. Wenn ein Text nicht zufällig, sondern aufgrund konkreter Absichten und Prinzipien entsteht, realisieren sich in ihm bestimmte Gesetzmäßigkeiten, die sich zwangsläufig mit Hilfe der entsprechenden Methoden aufspüren lassen. Wir behandeln die Texttypen und -arten in dem Kapitel über die stilbildenden Faktoren.

Die Typologisierung eines Textes. Der erste und grundlegende Schritt in dem Prozeß des Erkennens eines Textes ist die Typologisierung. Diese muß, bevor man mit der Analyse oder Interpretation beginnt, abgeschlossen sein. Als ein auf Erkennen und Enthüllen ausgerichteter Prozeß verläuft sie jedoch auch parallel zu Analyse und Interpretation. Auf ihre Ergebnisse stützen sich Kritiker, Literaturwissenschaftler und Leser, wenngleich die Typologisierung des Lesers im wesentlichen auf Intuition beruht. Unter dem Begriff *Typologisierung* verstehen wir das Aufspüren von Gesetzmäßigkeiten, das erste Erkennen, das erste Abtasten der Textstruktur. Bei der Typologisierung handelt es sich um die Identifizierung, um das Feststellen jener Elemente, die den Text eigenartig, individuell, selbständig zu dem machen, was er ist. Es handelt sich also um die Aufdeckung der grundlegenden, ihm eine Form verleihenden Züge und um das *Sameln* von Merkmalen, die eine Definition seiner Form ermöglichen.

Worauf gründet sich nun die Arbeitsweise des Typologen? Im Text sind auf verschiedene Weise Elemente verteilt und angeordnet, die der Autor mit dem Ziel, eine Aussage zu vermitteln, ausgewählt hat. Diese Elemente sind Punkte, *Codes*, Träger einer bestimmten Information. Sie bilden ein mehrdimensionales Netz und helfen - verfolgt man sie zurück - den Inhalt der vermittelten Aussage zu dechiffrieren. Sie sind *Eckpunkte*, von denen aus man den gesamten Raum, einfacher gesagt, die gesamte Ebene des Textes *sehen* kann. Sie dienen dem Typologen, Analytiker, Interpreten und Leser als Ausgangspunkte und Indikatoren und erlauben ihm, sich der Aussage, die in den Text chiffriert wurde, zu bemächtigen. Bei der Typologisierung werden solche Punkte gesucht und, wenn man sie entdeckt hat, aus verschiedenen

Blickwinkeln untersucht: innerhalb des Textes auf praktische und außerhalb auf theoretische Weise mit Hilfe von Vergleichen, Transposition und Transformation. Ein solcher Punkt kann ein Wort, ein Wortteil, die Art der Anordnung, ein Satz, ein Satzteil und seine Umgebung sein, d.h. alles das, was als Spurdient und zum Erkennen des Textes führt.³

Die Analyse eines Textes. Die bei der Typologisierung entdeckten Merkmale erweisen sich als Mittel, mit deren Hilfe man einen tieferen Einblick in das Textinnere gewinnen kann. Diesen Einblick gewährt die Analyse, die der Typologisierung um eine Stufe übergeordnete Erkenntnisform. Die Analyse kann aus didaktischen und erkenntnistheoretischen Gründen vorgenommen werden, je nachdem ob sich das Hauptaugenmerk auf die Bedeutung oder die formbildenden Elemente des Textes richtet. Im Unterschied zur Typologisierung versteht die Analyse die Textelemente nicht nur als Signale, sondern auch als Elemente, die sich zu einem Kontext verbinden können. Die Analyse ist eine Übergangsstufe zur Interpretation. In ihr können jeweils literatur- oder sprachwissenschaftliche Aspekte überwiegen. Die Typologisierung bleibt eng am Text, während sich die Analyse oft von ihm entfernt, um außerhalb befindliche Elemente zu untersuchen. Die Typologisierung gleicht einer Beschreibung. Sie antwortet auf die Fragen *was, wo, wieviel, wie*, die Analyse nur auf die Frage *warum*. Zwischen Typologisierung und Analyse läßt sich jedoch keine eindeutige Grenze ziehen. Beide Methoden beschäftigen sich hauptsächlich mit der Form eines Textes, beide können parallel zur Interpretation angewandt werden oder mit ihr interferieren. Charakteristisch für beide ist das objektive Konstatieren von Phänomenen, die Annahme, daß hinter diesen Methode, technisches Vorgehen und nicht eine subjektiv wertende Kraft - der Mensch - wirksam ist.

Die Interpretation eines Textes. Während Typologisierung und Analyse eng an die formale Seite eines Textes gebunden sind, stützt sich die Interpretation in beträchtlichem Maße auf die inhaltliche. Sie bedient sich dabei hauptsächlich subjektiver Momente und versteht die aufgrund der Typologisierung vorliegenden Informationen und die bei der Analyse aufgedeckten Be-

ziehungen nur als Signale und Ausgangspunkte für die Wertung und *Vervollständigung* des gegebenen Textes. So heißt Interpretieren nicht nur *Auslegen, Erläutern, Erhellern*, sondern auch anhand des Textes, der Gegenstand der Untersuchung ist, *einen neuen Gegenstand aufbauen*. Durch die Interpretation entsteht ein vollständiges, in sich geschlossenes Aussagegebilde, das mit anderen Worten direkt oder indirekt dasselbe sagt, was bereits in dem untersuchten Text gesagt wurde. Die Analyse macht Aussagen über den Text, die Interpretation anhand oder anstelle des Textes.

Nach Janusz Sławifski⁴ kann eine Interpretation dreierlei sein: 1. eine literaturhistorische Studie, 2. ein literarischer Essay, 3. eine literaturkritische Studie. Jede der genannten Formen hat ihren eigenen Blickwinkel (einen literaturhistorischen, einen persönlichen, einen literaturkritischen) und eine ihr angemessene Aussageform (Essay, Studie).

Eine wichtige Rolle bei der Interpretation spielt der Kontext, der sich mit ihr verbindet und auf den sie sich stützt. Normalerweise entstehen Interpretationen nur zu künstlerischen Texten, während die ihr untergeordneten Methoden auch zum Erkennen von Charakter und Eigenschaften sachbezogener Texte angewandt werden, denn hier läßt sich nur schwerlich eine Interpretation erstellen, die den Text sinnvoll deformieren könnte. Wenn die Typologisierung zuverlässig sein soll, so muß sie unabhängig vom Autor stets dieselben Informationen erbringen. Wenn dagegen die Analyse zuverlässig sein soll, so muß sie konsequent in der Methode und kompromißlos im gewählten Ansatz sein. Die Interpretation jedoch kann im Hinblick auf den Autor, auf die Zeit und andere Umstände immer eine neue sein, sie kann immer andere Informationen liefern, der Text jedes Mal anders verstanden und erklärt werden.⁵

Es muß noch erwähnt werden, daß sich der Autor einer Interpretation immer mit den ästhetischen Qualitäten eines Textes beschäftigt, über die die Typologisierung gar keinen Aufschluß, die Analyse nur dann, wenn sie direkt darauf abzielt, gibt. Diese Qualitäten sind nur potentiell im Text vorhanden und werden erst im Bewußtsein des Perzipierenden aktualisiert. Aus

diesem Grunde ist eine Interpretation, die sich nicht auf eine vorangehende Analyse oder Typologisierung anhand exakter Methoden stützt, eine rein intuitive Tätigkeit, deren Ergebnis ausschließlich vom Autor des interpretierten Textes abhängt.

Die Modellierung eines Textes.⁶ Modell heißt im weitesten Sinne des Wortes *Muster, das aufgrund der Nachahmung oder Beschreibung eines tatsächlichen Zustandes entsteht*. Von Modellierung spricht man in der Bildhauerei, in der Architektur, von Modellen in der Schneiderei, in der Sprachwissenschaft und überall dort, wo eine verallgemeinerte Form angestrebt wird. Während Modellierung die Abgrenzung der Elemente, die den Text zu dem machen, was er ist, meint, so entstehen Modelle auf der Grundlage sprachlich-kompositioneller Codes, die die Stützen des Textes sind. Von Modellierung spricht man ebenfalls im Zusammenhang mit der Typologisierung und der Analyse.

Für unsere Arbeit jedoch ist der Begriff der Typologisierung am wichtigsten. Die Typologisierung ist der erste Schritt zum Erkennen, gleichzeitig aber auch nur ein begleitender Vorgang. Analyse und Interpretation basieren auf der jeweils niedrigeren Stufe. Mit anderen Worten: eine Typologisierung ist nicht die einseitige Vorbereitung des Materials, sondern eine permanente Prüfung und Verifizierung der Informationen, die der sich mit dem Text Beschäftigende erhält. Sie stützt sich also nur auf das, was im Text tatsächlich vorhanden ist. Deshalb sind auch für sie exakte, in erster Linie mathematisch-quantitative und mathematisch-algebraische Methoden sinnvoll und ergiebig.

STILBILDENDE FAKTOREN

1. Man muß die Herausbildung eines Stils als einen Prozeß verstehen, in dem zwei Größen einander entgegenwirken. Die eine entspricht der Menge aller Ausdrucksmittel, die andere der aller Kräfte, die für die Wahl dieser Ausdrucksmittel richtungweisend sind. Die erste Größe korrespondiert mit dem passiven vorgefundenen Inventar an Wörtern, Konstruktionen und Kompositionsmitteln, über das der Autor verfügt und aus dem er nach Belieben dasjenige auswählt, welches die richtige, genaue und gegebenenfalls ästhetische Äußerung verlangt. Die zweite Größe basiert auf dem Zusammenspiel von Funktion, Milieu und Autor der Äußerung, d.h. der Kräfte, die auf die Wahl der Ausdrucksmittel einwirken. Sie entspricht der Menge der sogenannten stilbildenden Faktoren, die den Autor beeinflussen, d.h. ihn bei der Frage leiten, welche Ausdrucksmittel er dem Inventar entnehmen und welche Wörter und Sätze er verwenden soll, um seine Äußerung ihrer Funktion und ihrem Milieu anzupassen.⁷

Die stilbildenden Faktoren sind verschiedenartig.⁸ Sie wiederholen sich bei der Herausbildung des Stils und verlangen dabei die Wiederkehr stets derselben Elemente. Im Laufe der Zeit entstehen so bestimmte Gruppierungen, bestimmte Mengen, die in der Theorie der Stilistik *Stilschichten* genannt werden.⁹ Diese basieren auf Elementen aller sprachlicher und vieler außersprachlicher Ebenen. Sie sind keine geschlossenen Inventare mit scharf umrissenen Grenzen, sondern es gilt das Gegenteil.

Zur Herausbildung des Stils und der Stilschichten ist es nicht notwendig, daß die Elemente nur einer Schicht angehören. Es gibt vielmehr Elemente, die sich zwischen den Stilschichten bewegen, Elemente, die auf eine Schicht beschränkt sind, und solche, die gleichzeitig mehreren Stilschichten angehören, also polyvalent sind.

Wir stellen einen Vergleich zwischen der Lexik in Konversationen und der publizistischen Texte an um festzustellen, a) wieviele lexikalische Einheiten in den Texten identisch sind, d.h. welche ein doppeltes *Heimatrecht* besitzen, und b) welche Wörter

für einen Texttyp spezifisch sind. Dies geschieht mit Hilfe der Formel zur Berechnung des Diskrepanzgrades. Hierbei bezeichnen wir mit L_1 und L_2 die Anzahl der lexikalischen, mit L_x die der polyvalenten Einheiten in den verglichenen Texten und erhalten die folgende Gleichung:

$$\text{Dis} = 1 - \frac{L_x (L_1 + L_2)}{2 L_1 L_2}$$

Wir stellen fest, daß 54% der in den Konversationstexten vorkommenden lexikalischen Einheiten auch in den publizistischen Texten anzutreffen sind und 46% herausragen, d.h. nur in den Konversationstexten erscheinen. 22% aus den publizistischen Texten treten auch in den Konversationen auf, während 78% herausragen, also auf die publizistischen Texte beschränkt bleiben.

Ein Vergleich der Lexik publizistischer und wissenschaftlicher Texte bringt das folgende Ergebnis: 38% der lexikalischen Einheiten aus den publizistischen Texten kommen auch in den wissenschaftlichen vor, während 62% herausragen, 29% aus den wissenschaftlichen Werken auch in publizistischen Texten, während 71% herausragen. Mit Hilfe der o.a. Formel läßt sich nun errechnen, daß die Diskrepanz zwischen publizistischer und wissenschaftlicher Lexik 32% beträgt. In Analogie gelten diese Gesetzmäßigkeiten auch für das Vorkommen anderer, z.B. syntaktischer, Elemente.

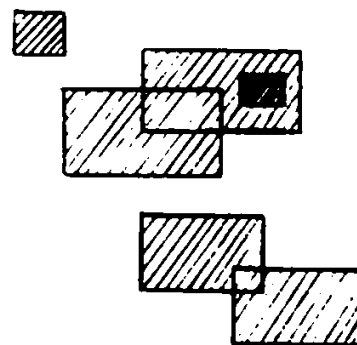
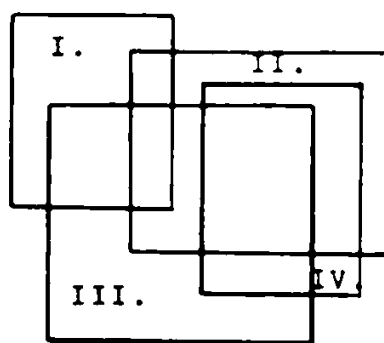
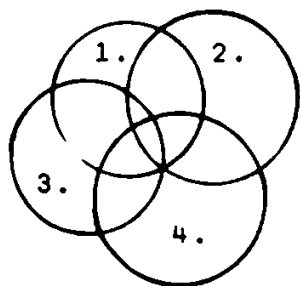
Bei der Herausbildung des Stils unterliegt der Autor drei Bedingungen, die ihn und damit seine Wahl der Ausdrucksmittel in eine bestimmte Richtung lenken. Zunächst muß in seiner Äußerung ein bestimmter Stil -z.B. der Konversations-, der wissenschaftliche, künstlerische oder administrative Stil - eingehalten werden. Also muß der Autor die sprachlichen Mittel in erster Linie im Hinblick auf die Kategorie *Stil* als allgemeinsten und höchsten Kategorie auswählen.¹⁰ Dann ist er an die Einhaltung der sogenannten stilistischen Methode gebunden. Mit diesem Begriff bezeichnen wir die Art der Anordnung der Elemente einer Äußerung sowie die der inneren Struktur und Komposition. Es kann sich dabei um eine einfache Information, eine Erzählung,

Beschreibung oder Erläuterung handeln. In jedem Falle ist der Autor bei der Wahl der sprachlichen Mittel auf diejenigen fixiert, die notwendig und angemessen sind, um - soweit möglich und erforderlich - eine bestimmte stilistische Methode zu bewahren. Schließlich unterliegt er noch der Forderung nach einer festgelegten stilistischen Form. Als solche bezeichnen wir Gebilde wie z.B. Vorlesungen, Erzählungen, Feuilletons und amtliche Schreiben. Der Autor muß also die sprachlichen Mittel auch danach auswählen, daß die Beibehaltung der einmal gefundenen stilistischen Form, falls notwendig, gesichert ist. Das Inventar der sprachlichen Mittel kann so Elemente enthalten, die jeweils die Funktionen A, B, C usw. erfüllen. Es überwiegt jedoch die Zahl der Mittel, die gleichzeitig mehreren Funktionen auf verschiedenen Ebenen (auf der Ebene des Stils, der stilistischen Methode und der stilistischen Form) gerecht werden.

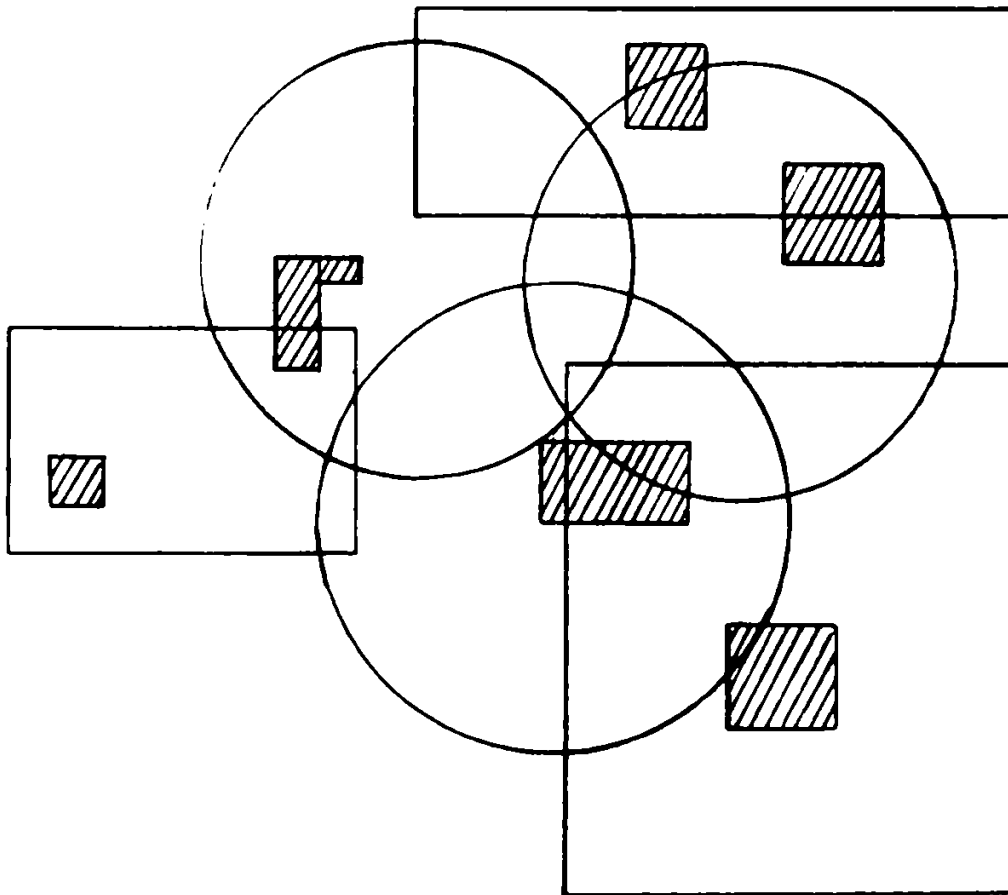
Anhand schematischer Darstellungen zeigen wir nun die Gliederung der sprachlichen Elemente in Mittel, die jeweils Funktionen auf der Ebene des Stils, der stilistischen Methode und der stilistischen Form erfüllen:

in Stilschichten:

entsprechend ihres Funktionierens
in der stil. Methode: in der stil. Form:



Zusammenfassende Darstellung der dreifachen Gliederung (aus Gründen der Übersichtlichkeit wurde die Zahl der Schichten verringert):



Fassen wir nun die Ergebnisse dieses Kapitelabschnitts zusammen. Ein und dieselbe Menge von Ausdrucksmitteln kann mehrfach gegliedert werden. Die schematischen Darstellungen zeigen, daß nicht jedes Mittel Träger einer einzigen Funktion und nicht alle Mittel Träger aller Funktionen sein müssen. Einige haben konstant, andere dagegen nur potentiell stilistischen Wert.

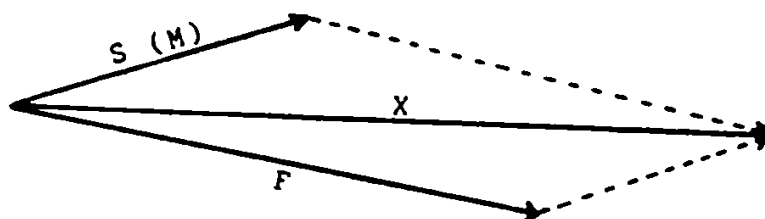
Bei der Typologisierung geht man von der Frage aus, welche Elemente in welchem Maß für den gegebenen Text gewählt wurden. Wenn diese auch im Text als eine homogene Masse erscheinen, so entstammen sie doch Gruppen, die verschiedenen Ebenen angehören.

2. Ein Vektor ist die aus der Physik bekannte Größe, die zugleich Richtung und Größe angibt und sich als Orientierungsabschnitt darstellen läßt. Jede Kraft ist eine Vektorengröße, ist

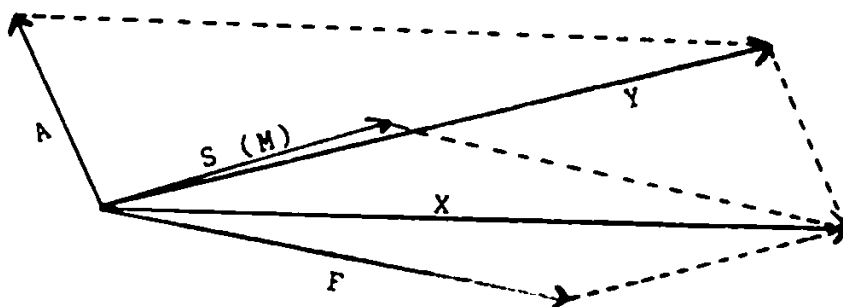
sie doch definiert durch Größe, Wirkungsrichtung und -bereich. Berücksichtigt man dies nicht, so kann man sie als Komponente einer Resultante verstehen.

Es wurde bereits festgestellt, daß die Grundlage einer sprachlichen Äußerung die Wirkung der verschiedenen stilbildenden Faktoren, d.h. Kräfte, auf die Wahl der sprachlichen Mittel ist. Diese wirken gleichzeitig und haben unterschiedliche Größe, Richtungen und Wirkungsbereiche. Im Endergebnis fügen sie sich jedoch zusammen und bilden eine Kraft, die die Wahl der geeigneten und angemessenen sprachlichen Mittel so ausrichtet, wie sie zum Aufbau der Äußerung notwendig sind. Wir geben zunächst ein einfaches Beispiel. Stellen wir uns die folgende Situation (S) vor: es existiert die Bedingung, daß die sprachliche Äußerung, mit der man auf sie reagiert, pejorativ (P) sein muß. Der Autor der Äußerung ist aber Optimist und hat ein solches Naturell, daß er meliorative Wörter (M) bevorzugt. Verlangt nun die Situation (S) z.B. die vierte Stufe der Kraft P und erreicht die Kraft M des Autors ebenfalls diese Stufe, dann befinden sich die beiden an sich gegensätzlichen Kräfte momentan im Gleichgewicht und neutralisieren einander. Wäre jedoch $P = 6$ und $M = 4$, so entstünde ein Ergebnis ($6-4=2$) zugunsten von P. Bei einem Ergebnis $P = 2$ müßte der Autor die sprachlichen Mittel so wählen, daß die Resultante = 2 wäre, d.h. daß sich eine mäßig pejorative Äußerung ergäbe. Bei einem Ergebnis $P = 0$ wäre die Äußerung neutral.

Stellen wir uns nun eine Situation vor, in der beim Aufbau der Äußerung in der Ebene Kräfte mit unterschiedlicher Größe und Richtung zur Wirkung kommen, die sich aus zwei stilbildenden Faktoren, z.B. Milieu (S) und Funktion (F), zusammensetzen. In diesem Falle muß der Autor sowohl S als auch F berücksichtigen, und zwar entsprechend ihrer Wichtigkeit und Richtung. Die Resultante korrespondiert mit der Kraft X, deren Größe und Richtung von den Vektoren S und F bestimmt wird:



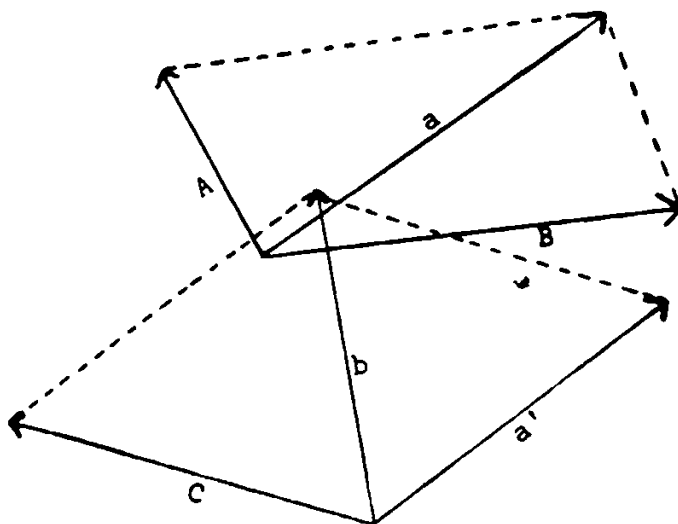
Berücksichtigen wir beim Aufbau der Äußerung auch die Kraft des Autoren-Ichs (A), die sich in Größe und Richtung von S und F unterscheidet, so entspricht die Resultante der Kraft Y:



Übertragen wir nun die Situation in den Bereich der Sprache, wo die resultierende Kraft, die sich aus der Kraft der einzelnen stilbildenden Faktoren zusammensetzt, über die Wahl der sprachlichen Mittel entscheidet. Gibt es auch dort zwei flächige, einander nebengeordnete, sich teilweise überlagernde Schichten M und F (wir wissen jedoch aus den bisherigen Ausführungen, daß sich die Elemente der jeweiligen Schichten in ihrer stilistischen Wertigkeit unterscheiden und diese nur potentiell vorhanden ist), so fällt die Wahl bei einer alleinigen Wirkung von Kraft A auf die Schicht A, bei gemeinsamer Wirkung der Kräfte S und F auf die Stelle, an der sich die Schichten S und F überlagern, also auf X usw. Wenn auch bei der letzten Resultante die Größe nicht so bedeutend ist wie ihre Richtung, so behält doch die Größe der einzelnen Komponenten ihr Gewicht. Zeigen die Informationen nämlich nur die Richtung, so gibt die Resultante immer den Durchschnittswert, die Mitte an. Es bestimmt also augenscheinlich die Größe der Kraft (gemeint ist die Kraft, die sich bei allen stilbildenden Faktoren unterscheidet) neben der Kraft auch die Richtung der Resultante.

Die Wirkung der stilbildenden Faktoren ist jedoch noch komplizierter, und zwar erstens deshalb, weil sie viel zahlreicher sind, als wir hier zeigen können, und zweitens, weil die Fragen der Auswahl sehr verwickelt sind. Die Gesamtheit der sprachlichen Mittel besteht nämlich aus vielschichtigen Gruppen (sagen wir der ersten Stufe), wozu noch die mit konstanten stilistischen Merkmalen (der zweiten Stufe) hinzutreten. Außerdem haben wir uns mit dem gewählten Beispiel nur auf der Geraden und in der Ebene bewegt, während die stilbildenden Faktoren im Raum

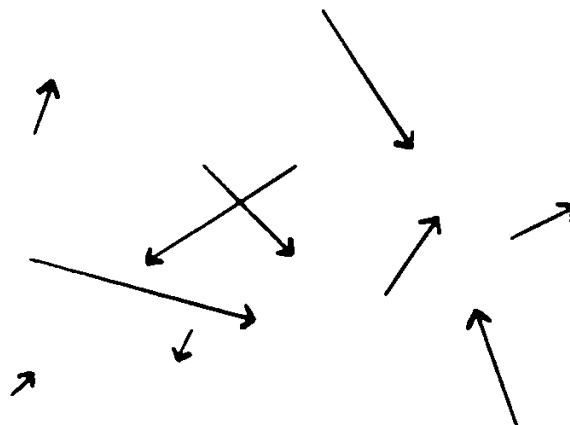
fungieren. Wir geben nun ein vereinfachtes Beispiel einer komplizierteren Situation, in der die Kräfte im Raum und nicht nur in einem Punkte wirksam sind. Der Aufbau ist deshalb verworrener als der der vorangehenden Beispiele, weil sein Ort transponiert wird.



Es gilt :

$$\begin{array}{r} A + B = a \\ a + C = b \\ \hline (A + B) + C = b \end{array}$$

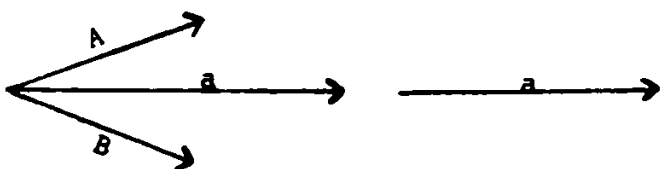
Die Wirkung der stilbildenden Faktoren muß man sich als die einer Menge von Kräften unterschiedlicher Art und Größe an verschiedenen Orten im Raum vorstellen:



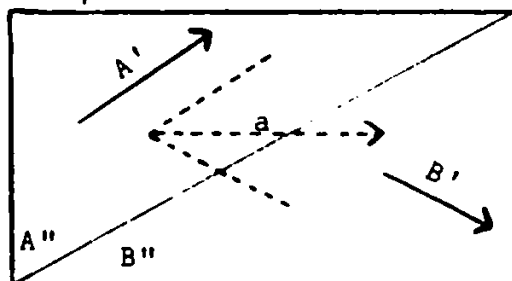
Die Resultante, die auf das sprachliche Material Druck ausübt, erhält man durch das - im wahrsten Sinne des Wortes - Zusammenlegen der Vektoren. Nach diesem Prozeß hat die auf das sprachliche Material Druck ausübende Kraft homogenen Charakter. Sie führt in eine Richtung und ist überall gleich groß. Diesem Wesen ist die nun entstandene Äußerung adäquat. Die Resultante übt ihre Kraft auf die Menge der sprachlichen Mittel nicht in

einem Punkt aus, sondern zerfällt im Augenblick ihrer Realisierung in verschiedene Komponenten, die wiederum in verschiedenen Richtungen an verschiedenen Orten wirksam sind.

Stilbildende Faktoren
und ihre Resultante:



Sprachliche Menge und in ihr die Kraft der stilbildenden Faktoren, zerlegt in einzelne Komponenten:



(A'' und B'' = verschiedene Stilschichten)

Die Komponenten der stilbildenden Faktoren verteilen sich also in der Weise, daß sie in denjenigen Stilschichten wirksam werden, auf die die stilbildenden Faktoren ausgerichtet sind. Jede stilbildende Kraft hat ihr Pendant in einer Menge von Ausdrucksmitteln. Ebenso bunt und vielfältig wie die mehrstufige, hierarchisch geordnete Gesamtheit der sprachlichen Elemente ist die der stilbildenden Kräfte, die in der schematischen Darstellung nur unzureichend, ja deformiert wiedergegeben ist.

Unter 1. haben wir die reiche Gliederung und komplizierte Einteilung der sprachlichen Mittel gezeigt. Verglichen mit der Wirklichkeit sind diese sehr vereinfacht und schematisch. Dennoch können wir uns mit ihrer Hilfe vorstellen, daß sich die Resultante der stilbildenden Faktoren theoretisch so zusammensetzt, daß ihre Komponenten alle diejenigen Teile der Menge umfassen, aus denen bei der Stilisierung eine Wahl getroffen wird. Die Resultante erscheint im Bewußtsein des Autors als ein Strahl, der in vielen Farben schillert. Auch sein Pendant, das Inventar der sprachlichen Mittel, gibt sich als eine polychromatische Menge von Elementen. Die Vektoren des stilbildenden Prozesses haben einerseits die Gestalt von stilbildenden Faktoren, andererseits von sprachlichen Mitteln.

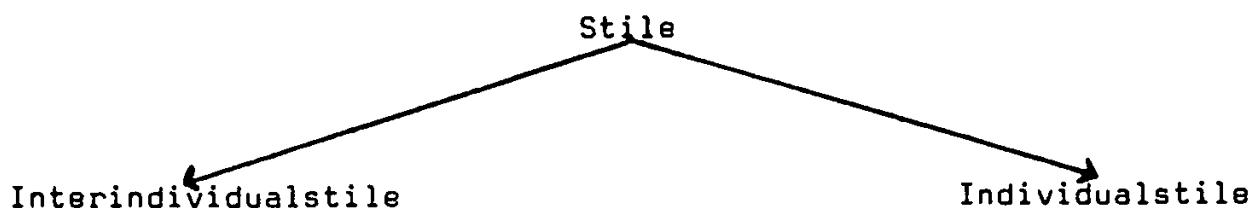
Es soll abschließend darauf hingewiesen werden, daß ein Autor nicht dazu in der Lage ist, diesen gesamten komplizierten Vorgang zu erfassen. Der stilbildende Prozeß verläuft in der Pra-

xis mehr oder weniger intuitiv. In demselben Maß ist das vorgeführte Modell eine Angelegenheit der Typologie und Diagnostik in sich geschlossener Textgebilde und ihrer Teile. Jeder Stil, jede stilistische Methode, jede stilistische Form und jedes geäußerte Ganze setzt ein auf sich ausgerichtetes Zusammenspiel von stilbildenden Faktoren voraus. Daraus folgt, daß der reine Stil, die reine stilistische Methode und die reine stilistische Form praktisch Angelegenheiten der Theorie und ihre Grenzen fließend sind.

Die stilbildenden Faktoren sind Kräfte, die - ebenso wie ihre Produkte - nach strengen Gesetzmäßigkeiten wirksam werden. Wenn dies zutrifft, muß ein Textgefüge in seine kleinsten Teilchen zerlegt und sein Aufbau nachvollzogen werden können. Aufgrund einer solchen Analyse lassen sich genaue Informationen gewinnen, die aber - zieht man die Kompliziertheit des stilbildenden Prozesses in Betracht - selbst bei größter Genauigkeit immer nur Signale bleiben, mit deren Hilfe sich der Text charakterisieren läßt.

3. Die Textarten. Als Funktionen bestimmter Mengen von stilbildenden Faktoren entstehen potentiell kompakte Texte, die sich aufgrund ihrer Eigenschaften zu einem hierarchisch geordneten System fügen. Man kann ein und denselben Text unter verschiedenen Aspekten beurteilen, so daß er gleichzeitig mehreren - wenn auch hierarchisch voneinander abgesetzten - Gruppen angehört.

Im Text manifestieren sich auch die persönlichen Neigungen und Eigenschaften des Autors. Sie bilden die Grundlage zur Bestimmung des Individualstils. Ein Text läßt sich ferner in die Gruppe der Interindividualstile einordnen. Eine weitergehende Klassifizierung macht den rein kommunikativen (sachbezogenen) oder künstlerischen Stil sichtbar. Auch diese kann - wie das folgende Schema zeigt - noch präzisiert werden.¹¹



Interindividualstile

Kommunikative (sachbezogene) Stile

1. umgangssprachlicher Stil
2. wissenschaftlicher Stil
3. publizistischer Stil
4. administrativer Stil

Künstlerische Stile

1. Epik
2. Lyrik
3. Drama

Der Stil der Umgangssprache ist hauptsächlich Konversationsstil, dessen man sich im privaten Bereich unter guten Bekannten bedient. Er ist typisch für die spontane, situationsbezogene Äußerung.

Der wissenschaftliche Stil ist vor allem Monologstil, der seine Anwendung im öffentlichen Leben findet. Er ist charakteristisch für Fachliteratur, Vorlesungen, Lehrbücher, Studien usw.

Der publizistische Stil kennzeichnet die Journalistik und die Zeitschriftenliteratur. Als enumerativ-benennender Stil operiert er mit parataktischen Satz- und Textmodellen.

Der administrative Stil ist typisch für administrativ-juristische Schriften, Urkunden, Verzeichnisse, Inventare, fakto-graphische Ankündigungen und Verlautbarungen.

Der künstlerische Stil erfüllt neben dem ästhetischen auch einen kommunikativen Zweck. Jedoch ist Kommunikation mit seiner Hilfe ein komplizierter Vorgang, da der künstlerische Stil im Unterschied zu den übrigen Stilen viele Dinge nur andeutet, signalisiert. Ihre "Ausbildung" bleibt dem Perzipierenden überlassen.¹²

Die Kategorien der Stile sind im Bereich der Stilistik die höchsten und abstraktesten. Neben dem Begriff *Stil* existiert auf einer anderen, hierarchisch tieferstehenden Ebene der Begriff *stilistische Methode*,¹³ womit wir die Art bezeichnen, in der die Elemente einer Äußerung angeordnet sind. Es gibt mindestens vier stilistische Methoden: die erzählende, die deskriptive, die informative und die erläuternde.

Die erzählende Methode wird zur Darstellung von Handlungen und Ereignissen angewandt. Ihr wichtigstes Moment ist die Orientierung an der Zeitachse, die es erlaubt, die Handlungsele-

mente so aufeinanderfolgen zu lassen, wie es den Aspekten der Zeit entspricht.

Die deskriptive Methode richtet ihre Aufmerksamkeit auf das Äußere der Dinge. Eigenschaften werden aufgezählt oder konstatiert, und zwar in der Reihenfolge, die der Autor selbst wählt.

Die informative Methode beschäftigt sich mit der einfachen Feststellung der Existenz einer Sache oder eines Faktums. Wichtig sind äußere Umstände: was, wo, wann geschehen ist.

Die erläuternde Methode erklärt die Beziehungen im Innern der Dinge oder ihr Verhalten zueinander. Sie zielt auf die Deutung von kausalen und anderen Zusammenhängen ab.

Die angeführten stilistischen Methoden überschneiden sich mit den einzelnen Stilen. Einige sind nur für bestimmte Stile typisch: z.B. die erläuternde Methode für den wissenschaftlichen, die informative für den publizistischen, die deskriptive für den künstlerischen und administrativen Stil.

Neben den Begriffen des Stils und der stilistischen Methode kennt die Stilistik noch den Begriff der Form. Eine Form ist ein konkreter Text, der zu einem bestimmten Stil oder einer bestimmten stilistischen Methode gehört und als geschlossene, ganzheitliche Einheit definiert ist. Hierzu zählen z.B. Urkunden, Briefe, Vorlesungen, Artikel, Erzählungen, Romane, Gedichte, Feuilletons, Nachrichten.

Wenn wir also bei der Typologisierung einen untersuchten Text einem bestimmten Stil zuordnen, so schließt das nicht aus, daß wir ihn gleichzeitig als stilistische Methode oder Form verstehen. Wichtig ist vor allem das Ziel der Typologie sowie das, worauf die aufgespürten Merkmale verweisen. Bei der Typologisierung geht es nicht darum, einen reinen Stil, eine reine stilistische Methode oder eine reine stilistische Form zu finden, sondern darum, daß gewisse grundlegende Tendenzen des betreffenden Textes erkannt werden.

GRUNDLAGEN EXAKTER METHODEN IN DER TYPOLOGIE

• Unter dem Begriff *exakt* versteht man im allgemeinen - und so auch in der vorliegenden Arbeit - *wissenschaftlich genau, auf mathematischen Berechnungen basierend*. Exakte Methoden sind Verfahren, mit deren Hilfe sich die Erscheinungen messen und objektiv darstellen lassen. Ihre Anwendung liefert genaue und meßbare Informationen über den Text.

Schon immer ging man bei der Identifizierung sprachlicher Gebilde so vor, daß man in ihnen bewußt oder unbewußt die räumliche und quantitative Distribution der Komponenten untersuchte. Dies war jedoch nicht genau und beruhte auf Schätzungen und Annäherungswerten. Lange Zeit wehrten sich Philologen gegen die Anwendung exakter Methoden, und teilweise geschieht dies auch heute noch dort, wo man den Zusammenhang zwischen Inhalt und physischer Gestalt eines Kommunikators nicht versteht. Im allgemeinen wird jedoch heute die Anwendung als sinnvoll anerkannt und das Inventar der exakten Methoden noch über den Bereich der mathematischen Verfahren hinaus erweitert.¹⁴

Allerdings sind die so erstellten Informationen klar, eindeutig und physisch begrenzt und müssen deshalb mit größerer Vorsicht gewertet und verwertet werden als die auf intuitiven Ansätzen beruhenden Ergebnisse. Eine exakte Information ist als Faktum bedeutend und gehört deshalb in Hände, die sie sorgfältig und abwägend werten.¹⁵ Wir wollen einige Prinzipien erörtern, auf denen die Anwendung exakter Methoden basiert und die unbedingt beachtet werden müssen.

1. Die exakte Information über die Existenz oder die Distribution eines Textelementes darf nicht zu einem komplexen Kriterium erhöht, sondern muß als Signal verstanden werden, das auf die potentielle Anwesenheit einer Textqualität hinweist. Die einzelne Information kann im Text polyvalent, mehrdeutig sein und ist als solche kein zuverlässiger, eindeutiger Indikator. Wir geben nun einige Beispiele für die Mehrdeutigkeit von Textphänomenen.

Mehrdeutig ist der Index der Wortwiederholung (IWW). Je bun-

ter und gesättigter ein Text in lexikalischer Hinsicht ist, je reicher also sein Wortschatz, desto schöner ist er, desto höher sein ästhetischer Wert, d.h. desto seltener sind Wortwiederholungen und desto niedriger der Index der Wortwiederholung. Dieser beträgt z.B. in lyrischen Texten 4,3 (die Angabe beruht auf Untersuchungen von Texten mit insgesamt 132 222 Wörtern), während er in Prosa 5,6, im Dialog 5,7 und in wissenschaftlichen Texten mehr als 7,5 lautet. Dies bedeutet, daß der journalistische Wortschatz bunter und reicher als der lyrische, oder simplifizierend gesagt, daß Lyrik = Journalistik ist. Obwohl diese Behauptung falsch ist, können wir den Index der Wortwiederholung nicht als fehlerhaften Indikator für den Wert eines Textes verurteilen, denn ein niedriger IWW verweist zwar auf starke lexikalische Variationen, aber nicht ad absurdum. Zu diesem einen Signal müssen noch weitere Signale hinzukommen, damit das abschließende Urteil Gültigkeit erlangen kann.

Ein ähnliches Signal ist das Vorkommen des verbalen Präsens (VVP). In sachbezogenen, deskriptiven und wissenschaftlichen Texten ist dieses irrelevant, da es hier ohnehin ausnahmslos verwendet wird. Sein Vorkommen hat also keinen stilistischen Wert. Anders verhält es sich bei seinem Gebrauch in erzählenden Texten, die eigentlich im Imperfekt stehen. In diesem Falle wird der entsprechende Teil der Erzählung durch die Anwendung des Präsens betont, hervorgehoben, dem Leser nähergebracht. Wie man sieht, kann also ein und dasselbe Element fast gegensätzliche Bedeutungen annehmen. Wiederum anders ist die Situation im Dialog, wo die Tempora Perfekt - Präsens - Futur in symmetrischer Anordnung erscheinen. Die statistische Information reicht auch im Hinblick auf das Vorkommen der Tempora nicht aus, um den stilistischen Wert eines Textes zu beurteilen.

Als ein weiteres Beispiel führen wir die grammatische Fügung der Sätze an. Ihre lockere Bindung ist charakteristisch für die expressiv-emotionale Ausdrucksweise. Dies resultiert aus der Tatsache, daß der Sprecher seine Emotionen schnell kundtun, sozusagen mit ihnen herausplatzen will und dabei Syntax und Satzzusammenhängen keine Beachtung schenkt. Ein lockeres syntaktisches Gefüge (das Fehlen verbindender Wörter) könnte also ein

Signal für Expressivität und Emotionalität eines Textes sein. Syntaktisch locker gefügt sind jedoch auch informative, journalistische Texte, die eine große Zahl von Aufzählungen enthalten. Journalistische Texte sind syntaktisch stereotype und zerbrechliche Konstruktionen, aber dennoch sachbezogen und nüchtern. Sie befinden sich so im Gegensatz zu den emotional-expressiven Texten. Die grammatische Fügung erweist sich folglich ebenfalls als ein mehrdeutiges Phänomen, das nicht absolut und ohne größeren Zusammenhang gesehen werden darf.

Die Beispiele zeigen, daß die sprachlichen Elemente zwar mit ihrer konventionellen Wertigkeit in den Text eingehen, potentiell aber polyvalent sind und ihre Funktion erst im Kontext des ganzen Gefüges erkennbar wird.

Wir vergleichen in diesem Zusammenhang die Diagnose des Wissenschaftlers mit der des Arztes. In dem wesentlichen Punkt stimmen beide überein: eine Information kann auf verschiedene Tatbestände verweisen. So z.B. läßt die Information, daß der Patient erhöhte Temperatur hat, mehrere Interpretationen zu, denn sie besagt zwar etwas, aber nicht eindeutig, nicht genau und nicht in jeder Situation dasselbe. Die erhöhte Temperatur verweist lediglich auf Krankheiten, zu deren Begleiterscheinungen Fieber gehört, besagt aber nicht, um welche Krankheit es sich handelt. Mit Hilfe der eliminierenden Methode muß der Arzt weiterforschen. Dasselbe gilt für die Typologisierung eines Textes: die einzelne Information kann nur eine Richtung angeben, d.h. sie ist nicht mehr als ein Signal für das potentielle Vorhandensein einer Eigenschaft und kann somit viele Bedeutungen haben.

2. Die Information über Existenz und Distribution eines Elementes darf nicht isoliert gewertet werden. Eine Zahlenangabe hat den Wert eines Ausgangspunktes, den Wert von Teilauskünften, die zueinander in Beziehung gesetzt werden müssen. Bezugspunkte können Informationen über ähnliche Situationen sowie theoretische oder parametrische Informationen sein, über die man aus umfangreichen Materialstudien verfügt.

Normalerweise kommen in künstlerischer Prosa 50% der Wörter je einmal, 30% je zweimal und 20% mehr als zweimal vor. Diese

Feststellung beinhaltet, daß ca. die Hälfte aller Wörter in künstlerischer Prosa als selten, okkasionell bezeichnet werden muß. Das sagt demjenigen, der sich eine klare Vorstellung von der Komposition künstlerischer Prosa machen will, jedoch recht wenig. Erst der folgende Vergleich gibt genauer darüber Aufschluß (die Informationen stammen aus einer Untersuchung von Texten mit insgesamt einer Million Wörtern):

(f = Frequenz)	Wörter mit		
	f 1	f 2	f 2<
künstlerische Prosa	50%	30%	20%
Poesie	55%	35%	10%
Journalistik	55%	35%	10%
wissenschaftliche Prosa	30%	30%	40%

Wir geben ein weiteres Beispiel. Der Index der Wortwiederholung beträgt in Konversationen 5,7, in journalistischen Texten 4,0, in der Poesie 4,3, in künstlerischer Prosa 5,6 und in wissenschaftlicher Prosa mehr als 7,5. In Volksliedern erreicht er die folgenden Werte:

im slovakischen Volkslied	4,81
im deutschen Volkslied	4,85
im russischen Volkslied	4,30
im tschechischen Volkslied	4,72
im polnischen Volkslied	4,40

Die Bedeutung dieser Informationen steigt durch den Vergleich der Volkslieder untereinander und mit anderen Textarten. Es zeigt sich, daß das Volkslied a) in allen Sprachen ein eigenes Genre darstellt, b) der künstlerischen Poesie am nächsten steht (das Volkslied hat einen Index der Wortwiederholung von 4,3 bis 4,85, die Poesie von durchschnittlich 4,3) und sich von ihr durch die von der Versform determinierte höhere Zahl der Wortwiederholungen unterscheidet und c) sein Stil sich wesentlich von den übrigen Stilen abhebt (eine Ausnahme bildet der journalistische Stil, der im vorangehenden Abschnitt diskutiert wurde).

In diesem Zusammenhang soll noch auf die folgende Tatsache hingewiesen werden. Ordnet man die in den Volksliedern vorkommenden Wörter nach ihrer Frequenz, so dominieren Interjektionen sowie Adjektive und Pronomina. Diese Information ist zwar aufschlußreich, hat aber nur dann wirkliche Bedeutung, wenn man

sie - anhand des Frequenzwörterbuchs - zu dem parametrischen Stand in Beziehung setzt. Im Frequenzwörterbuch werden an erster Stelle die Funktionswörter genannt, und erst auf diese folgen die im Volkslied dominierenden subjektiven und situationsbezogenen Wörter. Ein solcher Vergleich erlaubt jetzt die Feststellung, daß die Volkslieder in höchstem Maße subjektiv, expressiv-emotional und wertend sind. (Man vergleiche das häufige Auftreten adjektivischer Epitheta.)

Wir halten auch das folgende Beispiel für aufschlußreich. Die Komposition eines Dramas wird u.a. durch die Zahl der vorkommenden Repliken charakterisiert. So befinden sich z.B. in den Dramen des slowakischen Schriftstellers Š. Králik 860 (in *Hra bez lásky*) und 850 (in *Svätá Barbara*). Auch in anderen Dramen dieses Autors bleibt ihre Zahl auf weniger als 1 000 beschränkt. Dennoch lassen diese Angaben keinen Schluß auf die kompositionellen Eigenschaften der Dramengattung zu. Die Situation ändert sich nämlich, sobald man weitere Informationen sucht. Die Dramen Karvašovs *Pacient*, *Polnočná omša* und *Veľka parochňa* enthalten 1 300, 1 100 bzw. 1 200 Repliken. In den Dramen Laholas ist die Zahl doppelt so groß wie in denen Králik's (in *Atentát* z.B. beträgt sie 1 700), in den Dramen Barč-Ivans ist sie ungefähr gleich (in *Neznámy* und *Človek* beträgt sie jeweils 900, in *Matka* und *Veža* jeweils 800, in *Dvaja* etwas mehr als 700). Ein abendfüllendes Drama enthält ca. 1 000 Repliken, wobei eine geringere Anzahl durch szenarische oder epische Mittel ausgeglichen werden kann. Wir kommen zu dem Schluß, daß erst der Vergleich der gewonnenen Informationen mit parametrischen oder sonstigen Angaben es erlaubt, diese Informationen als Ausgangspunkt für die Typologisierung zu wählen. Die Wertung einer isolierten Information in ihrer absoluten Gestalt und das Unterlassen jeglicher Konfrontation mit anderen entspricht nicht der Arbeitsweise, die die sinnvolle Anwendung exakter Methoden verlangt.

3. Eine Information über Existenz oder Distribution eines Elementes kann als Voraussage für die Existenz bzw. Distribution eines anderen dienen, da zwischen den einzelnen Elementen jedes Textes ein direkter Zusammenhang besteht. Bei der Typologisie-

rung sollte man die eventuelle Kraft eines Elementes stets vor Augen haben, um so bestimmte Untersuchungen überflüssig machen zu können. Es existiert z.B. ein direkter Zusammenhang zwischen der Wort-, Satz- und Abschnittslänge. Je länger nämlich die Wörter eines Textes sind, desto länger sind auch die Sätze und - meistens - die Abschnitte, ebenso umgekehrt.

Wir übernehmen ein Beispiel aus der *Theorie der Texte* (1962) von Max Bense, der die Analogie von Wort- und Satzlängen nachweist. Er stützt sich dabei auf Zählungen von W. Fucks.

Texte	Durchschnittliche Länge	
	des Wortes	des Satzes
Rilke: <i>Cornet</i>	1,45	8,75
May: <i>Winnetou I</i>	1,61	14,10
Eyth: <i>Hinter Pflug und Schraubstock</i>	1,72	15,04
Bergengruen: <i>Der letzte Rittmeister</i>	1,85	15,26
Bismarck: Reden aus <i>Gesammelte Werke</i>	2,28	33,66
Heidegger: <i>Was ist Metaphysik?</i>	1,82	16,80
Freud: <i>Fragen der Laienanalyse</i>	1,89	19,14
Heisenberg: <i>Physik der Atomkerne</i>	1,91	20,53
Joos: <i>Theoretische Physik</i>	2,05	22,64
Kordt: <i>Die Wilhelmstraße</i>	2,11	26,00

Unter den 50 Titeln treten auch einige extreme Werte auf:

Eichendorff: <i>Aus dem Leben eines Taugenichts</i>	1,56	24,90
Fallada: <i>Kleiner Mann, was nun?</i>	1,53	10,68
Kästner: <i>Die verschwundene Miniatur</i>	1,73	8,43
Stifter: <i>Hochwald</i>	1,74	27,79
Hegel: aus <i>Sämtliche Werke</i>	1,84	31,38
Schliemann: <i>Trojanische Altertümer</i>	1,89	42,13
Du Bois-Reymond: <i>Über eine Akademie...</i>	1,94	11,18

Wir erwähnen diese exklusiven Beispiele deshalb, weil sie deutlich zeigen, wie notwendig eine vorsichtige Wertung und Behandlung der quantitativen Informationen ist. Sie widersprechen jedoch nicht der These von Zusammenhang und Prädiktabilität einzelner Phänomene. Eine falsche Information, die auf dieser basiert, wird aufgrund des unter 1. und 2. beschriebenen Mißtrauens korrigiert.

Eine gewisse Kongruenz zeigt sich auch in der Länge von Sätzen und Abschnitten. In Texten, die für Kinder geschrieben werden, z.B. sind Wörter, Sätze und Abschnitte durchweg kurz, während sich in wissenschaftlichen Texten das Gegenteil beobachten läßt. Je länger eine sprachliche Einheit ist, desto weniger

kommt jedoch das Gesetz von der Konkordanz zum Tragen.

Eine Übereinstimmung läßt sich ebenfalls zwischen Satzlänge, Sättigungsgrad an Substantiven und Adjektiven sowie dem Vorkommen bestimmter Kasus im Satz beobachten. Je mehr Nomen in einem Text auftreten und je länger die Sätze sind, desto größer ist auch die Zahl der obliquen Kasus, die wiederum eine Steigerung der Anzahl von Präpositionen bewirkt. Die Kongruenz erweist sich so als ein Mittel, einen Text unter verschiedenen Aspekten beurteilen zu können.

4. Ein Text ist ein kompliziertes Gebilde, das den Stempel des Autorensubjekts und der Entstehungssituation trägt. Die Elemente in seinem Innern stehen zueinander in Wechselbeziehung, nur wenige sind voneinander isoliert.

Man könnte diese Situation mit der des menschlichen Körpers vergleichen, denn auch dieser ist ein kompaktes, vollständiges und in sich geschlossenes Phänomen, das sich aus mehreren Komponenten zusammensetzt, von denen nur wenige voneinander unabhängig sind. Bezeichnen wir z.B. Herz und Lunge als voneinander abhängige und Augen und Füße als voneinander unabhängige Komponenten. Der Vergleich von Text und menschlichem Körper soll überspitzt zeigen, wie bei der Typologisierung vorzugehen ist. Zur Heilung bzw. Erstellung einer Diagnose untersucht der Arzt den Körper unter verschiedenen Gesichtspunkten. Er muß dabei genau über den Zusammenhang der einzelnen Teile informiert sein. So sind bei einem Beinbruch alle die Untersuchungen überflüssig, die bei einer Erkrankung der Verdauungsorgane unbedingt erforderlich wären, denn im vorliegenden Fall helfen sie nicht weiter. Dagegen werden bei einer Lungenkrankheit auch die Funktionen anderer Organe (z.B. des Herzens) untersucht, um auf diesem Umweg zur sicheren Erkennung von Grad und Charakter der Krankheit zu gelangen.

Diagnose und Typologisierung unterliegen ähnlichen Gesetzen. Es wird vorausgesetzt, daß der Untersuchende seine Arbeit nicht blind, ziellos, sondern zielbewußt, mit einer bestimmten Perspektive, ja sogar mit konkreten Vorstellungen über die Struktur des Textes verrichtet. Als weitere Voraussetzung gilt, daß sich im Verlauf der Typologisierung verschiedene Ansatzpunkte

und Vermutungen als fruchtbar erweisen und die Arbeit vorantreiben. Nennen wir einige konkrete Beispiele. Bei der Untersuchung der Satzlängen eines sachbezogenen Textes reicht die Berechnung des einfachen Durchschnitts aus, da sie in einem Text dieser Art nur wenig variiert werden. Völlig anders muß man jedoch bei essayistischen und künstlerischen Texten vorgehen, denn diese sind besonders reich an Variationsmöglichkeiten. In der modernen Prosa kann nämlich ein ganzer Roman oder auch eine ganze Erzählung in syntaktischer Hinsicht nicht gegliedert sein.

Es wäre unzweckmäßig, in wissenschaftlichen Texten die Frequenz der grammatischen Tempora zu untersuchen, da diese hier stilistisch irrelevant sind. Ebenso unzweckmäßig wäre es in diesem Falle, die inversive Anordnung der aktuellen Satzglieder statistisch aufzuschlüsseln, da die umgekehrte Folge psychologisches Subjekt - psychologisches Objekt nur in weniger als 1% der Fälle vorkommt. Überhaupt sind bei der Anwendung exakter Methoden Größen von weniger als 1% wertlos, von weniger als 5% nicht beweisfähig.

Je nach der Art des Textes ist für die Typologisierung entweder der Bereich der Lexik, Grammatik oder Komposition von Bedeutung. Es muß genau überlegt werden, für welche Phänomene in welchem Bereich eine Untersuchung sinnvoll ist. Vergleicht man z.B. die Lexik zweier Verfasser lyrischer Prosa, so kann man Konjunktionen oder Fragen der Konstruktion außer acht lassen. Die Typologisierung mit Hilfe exakter Methoden ist also kein schematischer Vorgang, der in vorgeschriebenen Etappen verläuft, sondern die im vorhinein gut durchdachte Arbeit der Suche nach numerischen Indikatoren.

5. Bei der Erläuterung der Grundbegriffe war von der Typologisierung nur als Vorstufe der Analyse und Interpretation die Rede. Dies bedeutet jedoch nicht, daß sie sich auf das bloße mechanisch-maschinelle Aufspüren von Signalen beschränkt, die dann später von Analyse und Interpretation bearbeitet werden. Im Gegenteil: da die quantitative Information nicht Ziel, sondern nur Mittel ist, heißt Typologisieren mehr als bloßes Sammeln von blinden Daten. Die echte Typologisierung verläuft nämlich in drei Etappen: 1. Aufstellen der Hypothese, 2. Aufspüren

und Sammeln von Daten, die die Hypothese stützen, modifizieren oder widerlegen, und 3. Abschluß durch Definition und Formulierung der These. Die Schritte 1 und 3 werden von Menschen getan, Schritt 2 kann irgendeine Maschine übernehmen. Die Typologisierung verläuft also über die Stationen Mensch - Maschine - Mensch. In diesem Zusammenhang soll auf den Einwand geantwortet werden, den die auf die Intuition vertrauenden Traditionalisten gegen die Anwendung exakter Methoden erheben. Dieser lautet: Worin besteht der Unterschied zwischen den durch die Anwendung exakter Methoden und den intuitiv gewonnenen Informationen, wenn man in beiden Fällen von der Intuition ausgeht und der Forscher von vornherein die Hypothese und das Beweisverfahren bestimmt? Antwort: Bei der Anwendung exakter Methoden wird ein Ergebnis vorausgesetzt. Im Verlauf der Bestimmung von Informationen, mit deren Hilfe man es erreichen will, ändert sich jedoch das Vorgehen und paßt sich der Situation an. Die Hypothese wird also korrigiert, während beim intuitiven Vorgehen eine subjektiv aufgestellte Hypothese nicht anhand objektiv meßbarer Informationen modifiziert wird, sondern gleich der These ist. So unterliegen die intuitiv erstellten Thesen nicht der strengen mechanisch-maschinellen Kontrolle.

Die Typologisierung ist also kein mechanisch-statistisches, sondern ein wissenschaftliches Vorgehen, bei dem der Mensch zeitweise und freiwillig von der Maschine kontrolliert werden kann.

Geben wir ein Beispiel. Bei der flüchtigen Betrachtung zweier Texte bemerkt der Interpret einen großen Unterschied in ihrem lexikalischen Umfang. Er stellt intuitiv eine Hypothese über die Verschiedenartigkeit auf und beginnt dann mit der maschinellen Untersuchung, um die Hypothese zu überprüfen und zu beweisen. Im Verlauf der Zählungen oder nach Erhalt des Endergebnisses zeigt sich jedoch, daß die Hypothese falsch war: Dem lexikalischen Umfang nach sind die Werke identisch. Bleibt dem Interpreten immer noch der Eindruck einer unterschiedlichen Lexik, so muß er eine andere Richtung wählen. Dabei bieten sich verschiedene Möglichkeiten an, z.B. kann der Charakter der lexikalischen Einheiten - d.h. die lexikalische Gliederung - un-

tersucht werden. Fachbegriffe rufen den Eindruck von lexikalischer Sättigung und Schwerfälligkeit hervor, Verben machen den Text leicht, Attribute belasten ihn usw. Der Interpret korrigiert seine ursprüngliche Meinung und formuliert die These, daß die betreffenden Texte nur quantitativ, nicht aber qualitativ gleich sind. Er kann ebenfalls einzelne Textpartien vergleichen und einen unregelmäßigen Rhythmus bei der Einführung neuer Wörter feststellen. Ferner kann er die Dichte in der Abfolge grammatischer Kategorien u.ä. überprüfen. Er muß also unaufhörlich den Wert der Intuition und die Gültigkeit der Hypothese, die als Ausgangspunkt für die exakte Untersuchung dient, kontrollieren.

6. Bei der Typologisierung ist ferner zu beachten, daß die bloße Existenz bzw. das häufige Vorkommen eines Elementes nichts über seine Funktion und seinen Wert aussagen. Wichtig sind immer seine Umgebung, seine Position in der Folge ähnlicher oder konträrer Elemente, die Möglichkeit einer Auswechselbarkeit usw. Es genügt also nicht, ein Element nur zu zählen, sondern es muß - besonders in künstlerischen Texten - unter verschiedenen Aspekten betrachtet werden. Wir wollen dies an einem Beispiel aus der Architektur verdeutlichen. Beim Bau eines Hauses wird eine bestimmte Menge Eisen benötigt, das verschiedene Funktionen ausübt. Ein Teil ist nach Bauabschluß sichtbar, ein anderer unsichtbar, ein Teil stützt die Konstruktion, ein anderer dient der Verschönerung. Spricht man also vom Eisen, das beim Hausbau verwendet wird, so versteht man darunter nicht eine homogene Masse, sondern die einzelnen Teile oder Gegenstände, die aus ihm hergestellt werden. Ein ähnlicher Fall ist die Konstruktion eines Gedichtes. Hier kann ebenfalls ein und dasselbe Element (Wort) Träger verschiedener Funktionen sein. Es kann z.B. in exponierten Stellungen wie in der Überschrift, am Zeilenende, als Reimwort und als Gedichtschluß, aber auch in Stellungen, die kein besonderes Gewicht oder keine auffällige Funktion haben, auftreten. An einigen Stellen ist es durch ein Synonym austauschbar, an anderen führt ein solcher Ersatz zur Zerstörung des Textsinnes. Tritt z.B. in einem Gedicht, in dem das Imperfekt dominiert, eine Präsensform - überdies in

exponierter Stellung - auf, so kommt diesem Tempus eine semantisch relevante Funktion zu. Eine ähnliche Wirkung erzielt z.B. ein einzelner Vulgarismus in einem Text gehobener Wortwahl, während er als ein Wort unter vielen keinerlei besonderen Stellenwert hat.

Das folgende Beispiel zeigt die Wichtigkeit der Gesetzmäßigkeiten, denen die Abfolge von Elementen unterliegt. In einem Text kommen die Wörter A und B je dreimal vor. Da das Wort A an der betreffenden Stelle gleich dreimal hintereinander, das Wort B jedoch an drei verschiedenen Stellen auftritt, erfüllen beide unterschiedliche Funktionen. Die isolierte Information besagt also nichts, wenn nicht noch weitere Informationen in Betracht gezogen werden.

Stellen wir einen Vergleich zwischen Roman und Erzählung an, so zeigt sich, daß die Konstruktionselemente im Roman frei, d.h. auf großem Raum, in der Erzählung dagegen eng, d.h. auf beschränktem Raum verteilt sind. Der Unterschied liegt also in der Dichte der informativen Elemente. Ähnlich verhält es sich aber auch bei einem Vergleich zwischen Romanrahmen und -innerem. Die Teile im Innern werden von den Reden der Personen beeinflusst, es wird mehr erklärt. Mit anderen Worten: obwohl Roman und Erzählung sich (normalerweise) deutlich voneinander unterscheiden, lassen sich doch in den Rahmenkomponenten gewisse Übereinstimmungen beobachten. Wird also für eine Untersuchung nur ein Teil ausgewählt, so muß deutlich gesagt werden, um welchen es sich handelt und warum man ihn wählt. Für die Schlußfolgerung ist dies besonders wichtig.

Bei der Untersuchung eines Textes spielt die vom Gewebe determinierte Länge eine hervorragende Rolle. Je kleiner der vorbestimmte Umfang ist, desto größer ist die Bedeutung der einzelnen Elemente. So ist z.B. eine Dialogreplik oder die Replikfolge einer Person anders konstruiert als ein dem Umfang nach identischer Monolog.

Die Anwendung exakter Methoden muß also auf der Voraussetzung beruhen, daß ein Element stets im Hinblick auf den Kontext untersucht wird. Durch die Eingliederung in einen Zusammenhang verändert sich die Wertigkeit eines Elementes. Es kann inner-

halb eines Textes sogar entgegengesetzte Funktionen und Bedeutungen haben.

Diese Ausführungen über den Anwendungs- und Geltungsbereich exakter Methoden könnten als Gebrauchsanweisung für eine bisher noch unbekannte, nicht konstruierte Maschine dienen. Würde diese imaginäre Maschine jedoch falsch gebraucht, so wäre sie überflüssig, ja sogar schädlich, denn auch sie könnte weder die bisher angewandten Methoden ersetzen, noch wäre sie zu allem fähig.

Da wir nicht glauben, daß man Typologisierung, Analyse oder Interpretation schematisch ausführen kann, sind wir auch jetzt nicht in der Lage, ein Muster für das Vorgehen anzubieten. Was mit Hilfe der exakten Methoden erreicht werden kann, bleibt allein der Fähigkeit und Erfindungsgabe des Forschers überlassen. Die exakten Methoden sollen und können nicht die Methoden ersetzen, die bis heute "unexakt arbeiten".

Eine fruchtbare Anwendung dieser Methoden hängt außerdem vom Typ des betreffenden Textes ab. Nicht immer lassen sich nämlich überzeugende Informationen gewinnen. Daher ist es notwendig, daß der Forscher ein hinreichend gebildeter Stilist ist und weiß, inwieweit die exakten Methoden sinnvoll und effektiv sind. Nur wenn ihn ihre Ergebnisse nicht überraschen, wendet er sie richtig an.

DISTRIBUTION LEXIKALISCHER ELEMENTE EINES TEXTES

Für den Aufbau eines Textes wird eine bestimmte Anzahl lexikalischer Einheiten (Wörter) gebraucht, von deren Verteilung die Textstruktur abhängt, d.h. der Autor wählt die Wörter direkt im Hinblick auf diese Struktur aus. Zerlegt man nun den Text in die ihn konstituierenden Elemente, so werden diese und ihre Zugehörigkeit zu bestimmten Systemen sichtbar. Strukturell ähnliche Texte ermöglichen die Herausbildung ähnlicher Systeme, in deren Gesetzmäßigkeiten sich dann die Absichten, mit denen die Elemente ausgewählt wurden, spiegeln. Diese gilt es zu untersuchen.

Wir geben ein Beispiel. Der Autor einer schnellen mündlichen Äußerung mißt der Wortwahl keine große Bedeutung zu, er will sich nur auf schnellstem Wege verständigen. Der Wortschatz ist deshalb sehr begrenzt, denn bestimmte Wörter wiederholen sich ständig. Dieses läßt sich jedoch erst nach der Analyse und Klassifizierung der gebrauchten Wörter genauer feststellen.

Bei der Typologisierung zerlegt man nun den Text zunächst in seine Wortkomponenten und weist diese dann verschiedenen Systemen zu, um so die Gesetzmäßigkeiten ihrer Distribution beschreiben zu können.

Die einzelnen nach ihrer Frequenz geordneten lexikalischen Einheiten lassen sich in zweifacher Hinsicht untersuchen: 1. im Hinblick auf ihre vertikale Wortdistribution (welche semantischen Gruppen bilden sie) und 2. auf ihre horizontale Wortdistribution (welche quantitativen Abweichungen und Schichten liegen vor).

Vertikale Wortdistribution

Ordnet man die Wörter zusammenhängender Texte nach ihrer Frequenz, so läßt sich eine Beziehung zwischen ihr und der lexikalischen Bedeutung feststellen. Bereits die empirische Beobachtung zeigt, daß die in der Rede dominierenden Einheiten die

Funktionswörter (Konjunktionen, Präpositionen, Partikel, Hilfsverben) sind, auf die die autosemantischen folgen. Unter diesem Gesichtspunkt teilt schon P. Guiraud die Wörter in vier Gruppen ein:¹⁶

1. dominierende - d.h. strukturelle und grammatische - Wörter (*semantiquement vides*), die ungefähr die Hälfte eines zusammenhängenden Textes ausmachen. Im Französischen gibt es davon ca. 100.
2. thematische Wörter. Ihre Anzahl beträgt ca. 50 und macht 9% eines zusammenhängenden Textes aus.
3. Basiswörter, d.h. Wörter, die zum Zentrum des Wortschatzes gehören. Es gibt davon ca. 4 500.
4. charakterisierende Wörter. Ihre Zahl überschreitet den Wert von 20 000.

Unser System der vertikalen Wortdistribution entstand auf Grund eines Vergleichs von fünf Frequenzwörterbüchern (slowakisch, tschechisch, russisch, deutsch, spanisch). Es bezieht sich nur auf die 110 dominierenden lexikalischen Einheiten und beruht darauf, daß diese in allen Sprachen ca. 50% eines beliebigen Textes ausmachen, während die restlichen Wörter entsprechend ihrer Bedeutung gewählt und verteilt sind. Man kann sie so als parametrische Menge verstehen, mit deren Hilfe sich der Wortschatz verschiedener Autoren, Stile und Genres untersuchen läßt. Da sie den Text außerdem am stärksten beeinflussen, ist es für die Typologisierung besonders wichtig, ihre bedeutungsmäßige Schichtung zu kennen.

Wir wenden das grammatisch-semantische Kriterium an und gliedern den Wortschatz folgendermaßen:

1. Konstruktionswörter K, zu denen grammatische Wörter wie Konjunktionen, Artikel, Hilfsverben usw. gehören. Sie nehmen in den Frequenzwörterbüchern die Plätze 1-20 ein.
2. Subjektive, von der Situation determinierte Wörter S, zu denen die Modalverben, Personalpronomina, einige Partikel und Interjektionen gehören. Sie nehmen die Plätze 21-70 ein.
3. Gnomische Wörter G, zu denen wir Ausdrücke mit allgemeiner, objektiver Bedeutung wie Numeralia und Nomina zählen. Diese nehmen die Plätze 71-110 ein.
4. Thematische Wörter T, die alle diejenigen Wörter enthalten, die in den oben genannten Schichten nicht vorkommen. Es handelt sich dabei vorwiegend um Substantive, Adjektive und Verben mit spezieller und terminologischer Bedeutung, die die Plätze 111 ff. einnehmen.

Diese Einteilung beruht primär auf dem Aspekt der Frequenz und erst sekundär auf semantischen Überlegungen. Prinzipiell geht sie auf die Notwendigkeit jeglicher Anwendung von exakten Methoden zurück: eine quantitative Information ist immer nur das Signal eines inhaltlichen Phänomens. Im vorliegenden Fall halten wir sie für einen zuverlässigen Indikator der semantischen Eigenschaften von Wörtern, die durch sie bestimmten Wortschichten zugeordnet werden. Der grammatische Aspekt der Frequenz ist deshalb vorrangig, weil er die exakte Einteilung in Schichten erlaubt, während der semantische die Grenzen fließend und daher schwer bestimmbar macht. Nur die zentralen Elemente der Wortschichten sind semantisch klar umrissen und müssen daher genau umschrieben werden.

Das Zentrum der ersten Schicht bilden die Präpositionen und Konjunktionen, d.h. die Einheiten, die in der Sprache als grammatisch-konstruktive Elemente dienen. Sie verweisen - neben einigen anderen Wörtern - auf die grammatischen Beziehungen. Deshalb gehören auch die Wörter *sein, haben, dieser, welcher* dieser Schicht an. Obwohl sie mehrdeutig sind, üben sie doch - lt. Statistik - vorwiegend grammatische Funktion aus.

Das Zentrum der zweiten Schicht nehmen die Wörter ein, mit deren Hilfe ihr Benutzer eine bestimmte Haltung (Partikel, Modalverben usw.), seine Einstellung zu Milieu (deiktische Pronomina und Adverbien) und Adressat (Personal- und Possessivpronomina) ausdrückt. Sie verleihen Aktualität und erhalten ihren Wert in der jeweiligen Situation, in der unmittelbaren Kommunikation.

Das Zentrum der dritten Schicht bilden die autosemantischen Wörter, mit denen allgemein und nicht-terminologisch die Phänomene (Gegenstände, Eigenschaften, Handlungen, Umstände) benannt werden. Zu ihnen gehören die Basiswörter wie z.B. *Mensch, Frau, Arbeit, Zeit, groß, alt, neu, gut, wissen, sehen, sagen, geben, machen*.

Während die Komponenten der beiden ersten Schichten größtenteils nur Relationen (sei es zwischen dem Benannten, zwischen dem Sprechenden und dem Milieu, sei es die Beziehung des Sprechers zum Benannten) angeben, sind die Wörter der dritten auto-

semantisch.

Wörter aus diesen drei Gruppen machen in einem beliebigen zusammenhängenden Text im Slowakischen und in anderen Sprachen 50% der lexikalischen Einheiten aus.

Die drei aus den 110 dominierenden Wörtern gebildeten Schichten sind in semantischer Hinsicht nicht scharf umrissen. In der Schicht der Konstruktionswörter sind nicht nur K-, sondern auch S-Wörter, in der Schicht der subjektiven Wörter auch K- oder G-Wörter vertreten.

Es werden also zwei Einteilungskriterien wirksam: a) der Frequenz, b) der Bedeutung. Anders ausgedrückt: zwei Mengen von Wörtern überschneiden sich. Die eine entsteht durch die Gliederung des Wortschatzes unter dem Aspekt der Frequenz (wir bezeichnen sie mit den Buchstaben K, S, G), die andere Menge durch die Einteilung unter semantischen Aspekten (sie bezeichnen wir mit den Buchstaben K, S, G).

Aus verständlichen Gründen kommt die nach dem Kriterium der Frequenz erstellte Schicht nicht mit der äquivalenten, unter semantischen Aspekten erstellten zur Deckung. Die Kongruenz der beiden läßt sich nur durch Approximation erzielen. Man kann allgemein das folgende bei der Typologisierung als Parameter dienende Gesetz aufstellen: in einem Text stehen an erster Stelle K-, an zweiter S- und an dritter G-Wörter.

Wendet man exakte Methoden zur Typologisierung und Interpretation an, so muß man in weit stärkerem Maße mit Annäherungswerten arbeiten als bei der Bestimmung grammatischer und lexikalischer Verhältnisse.

Im folgenden geben wir eine Liste der 110 dominierenden slowakischen Wörter der Gruppen K, S, G (gegebenenfalls ist eine Kollision mit K, S, G angezeigt):

1. *a* (K), *byt'* (K), *v* (K), *na*(K), *sa* (K), *ten* (K), *on* (K,S), *že* (K), *z* (K), *ako* (K), *s* (K), *do* (K), *mat'* (K), *ktorý* (K), *ja* (K,S), *čo* (K), *aj* (K), *o* (K), *za* (K), *ale* (K)
2. *po* (S,K), *k* (S,K), *ked'* (S,K), *len* (S), *už* (S), *všetok* (S), *tento* (S), *tak* (S,K), *svoj* (S), *i* (S,K), *môť* (S) *jeho* (S), *my* (S), *ešte* (S), *nie* (S), *od* (S,K), *pre* (S,K), *jeden* (S,G), *ty* (S), *ani* (S,K), *aby* (S,K), *vediet'* (S,G), *tu* (S), *nás* (S), *pri* (S,K), *fst'* (S,G), *alebo* (S,K), *človek* (S,G), *celý* (S), *musiet'* (S), *dat'* (S,G), *chciet'* (S), *vel'ký*(S,G),

povedat' (S,G), *taký* (S,K), *vy* (S), *pred* (S,K), *môj* (S), *dva* (S,G), *vidiet'* (S,G), *každý* (S), *rok* (S,G), *prvý* (S,G), *kde* (S), *potom* (S), *druhý* (S,G), *iný* (S), *by* (S,K), *čas* (S,G)

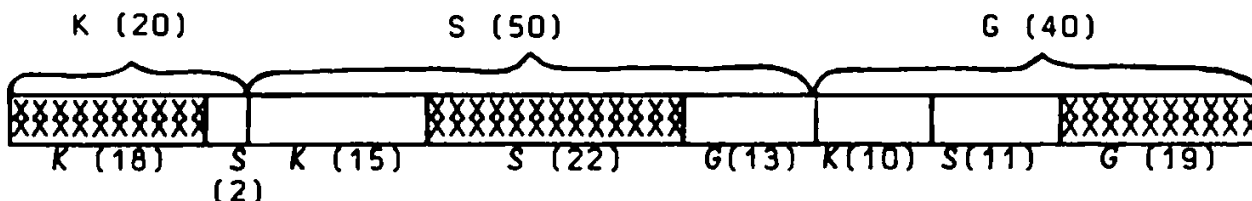
- 3. *lebo* (G), *deň* (G), *ak* (G,K), *veľ'a* (G,S), *hovorit'* (G), *až* (G,S), *nový* (G), *no* (G,K) *pášt'* (G), *preto* (G), *sám* (G,S), *bez* (G,K), *však* (G,K), *iba* (G,S), *svet* (G), *tam* (G,S), *medzi* (G,K), *pod* (G,K), *či* (G,K), *život* (G), *vec* (G), *ruka* (G), *dobre* (G), *kto* (G,), *teda* (G,K), *nič* (G,S), *voda* (G), *starý* (G), *aký* (G,S), *veľ'mi* (G,S), *dobrý* (G), *teraz* (G,S), *začat'* (G), *žena* (G), *u* (G,K), *malý* (G), *slovo* (G), *strana* (G), *raz* (G), *treba* (G,S), *oko*(G) .

Im slovakischen Frequenzwörterbuch (vgl. o.a. Liste) lassen sich die folgenden Überschneidungen beobachten: in der Menge K sind 2 Elemente, die semantisch zur Menge S gehören, in der Menge S - abgesehen von den 22 K-Elementen - 15 der Menge K und 13 der Menge G. Die Menge G enthält - außer den 19 G-Elementen- 10 der Menge K und 11 von S. Diese für das Slovakische typische Situation liegt analog auch in anderen Sprachen vor:

	Menge K			Menge S			Menge G		
	K	S	G	K	S	G	K	S	G
Slovakisch	18	2	0	15	22	13	10	11	19
Tschechisch	17	3	0	19	22	9	7	13	20
Russisch	11	9	0	14	19	17	7	14	19
Deutsch	18	2	0	18	19	13	6	17	17
Spanisch	16	4	0	13	20	17	5	9	26

In allen Sprachen ist also S die am wenigsten und K die am meisten homogene Menge.

Stellen wir nun die Äquivalenz der Mengen K, S, G und K, S, G im Slovakischen - gesehen unter dem Aspekt der Frequenz und Bedeutung - graphisch dar:



- K = Konstruktionswörter
- S = Subjektive, von der Situation determinierte Wörter
- G = Gnomische Wörter
- K = Menge der dominierenden Wörter, benannt nach der Menge K der Konstruktionswörter,
- S = Menge der Wörter, die nach der Menge S der subjektiven Wörter bezeichnet sind,
- G = Menge der Wörter, die nach der Menge G der gnomischen Wörter benannt sind.

Die einzelnen Mengen unterscheiden sich auch in formaler Hin-

sicht. So sind die zur Menge K gehörigen Wörter zum größten Teil indeklinabel mit einer Silbenlänge von durchschnittlich 1,15, während die Wörter der Menge S eine Silbenlänge von ca. 1,25 aufweisen. Diese sind in ihrer systemhaften, nicht durch den Kontext determinierten Gestalt einsilbig. Erscheinen sie jedoch im Zusammenhang, so sind sie aufgrund ihrer Endungen im Durchschnitt etwas länger. Die zur Menge G gehörenden Wörter sind zu 50% ein- und zweisilbig. Da ungefähr die Hälfte von ihnen deklinierbar ist, erreicht ihre Silbenlänge Werte bis zu 1,77.

Durch die vertikale Gliederung des Wortschatzes entsteht als vierte Menge T die der thematischen und terminologischen Wörter, deren Zentrum die terminologischen, die wegen ihrer durchschnittlich geringen Frequenz normalerweise nicht in den Frequenzwörterbüchern aufgeführt werden, bilden. Sie sind alle deklinierbar und haben im Durchschnitt mehr als zwei Silben.

Für die stilistische Typologisierung ist der folgende Fall relevant: in einem analysierten Text wird die Harmonie der äquivalenten Ebenen (der Frequenz und der Semantik) dadurch gestört, daß eine Menge K, S oder G unter dem hypertrophischen Einfluß fremder Elemente einen anderen semantischen Charakter bekommt, als die parametrische Konfiguration erwarten läßt. Wenn in der Frequenzliste eines konkreten Textes unter den ersten 20 Wörtern der Menge K z.B. mehr Elemente mit der semantischen Wertigkeit S, G, T erscheinen, so sind in der Menge S ebenfalls mehr Elemente mit der semantischen Wertigkeit G oder T enthalten, usw. Gelangen Elemente aus anderen Mengen in die Menge K, so führt dieses nicht zum Schwund eines Elementes, sondern zu seiner Verschiebung in Richtung nach hinten, und zwar um so viele Stellen, als fremde Elemente in die betreffende Menge eingedrungen sind. Dies gilt analog auch für die anderen Mengen.

Mit Hilfe mathematischer Terminologie und Symbole läßt sich dies darstellen (r gibt den Rang, d.h. die Stellung eines Wortes im Frequenzwörterbuch an). Es sei K die Menge der Wörter, die $r=1-20$ einnehmen, S die Menge derer mit $r=21-70$, G die Menge der Wörter mit $r=71-110$ und T die der Wörter mit $r=111\llcorner$.

Es sei K die Menge aller slovakischen Konstruktionswörter ohne Bezugnahme auf ihre absolute Frequenz, S die aller subjektiven, von der Situation determinierten Wörter, G die aller gnomischen und T die aller thematischen Wörter.

Es sei K_K eine Untergruppe in der Menge der Konstruktionswörter, die in die Menge r_{1-20} (K) gehören, K_S eine Untergruppe der subjektiven Wörter, die der Menge r_{1-20} (K) angehören, und S_K eine Untergruppe der Konstruktionswörter, die der Menge r_{21-70} (S) zugeordnet sind. Ähnliches gelte für S_S , S_G , S_T und G_K , G_S , G_T .

Das bedeutet: $K_K \cong K$, $K_S \cong K$, $S_K \cong S$, $S_S \cong S$, $S_G \cong S$, $G_K \cong G$, $G_S \cong G$, $G_G \cong G$, $G_T \cong G$, $T_K \cong T$, $T_S \cong T$, $T_G \cong T$, $T_T \cong T$.

Für alle Sprachen trifft allgemein und bei beträchtlicher Approximation die folgende Definition zu:

$$K \doteq K, S \doteq S, G \doteq G, T \doteq T.$$

Sie gilt allerdings nur unter der Voraussetzung, daß in der Menge der Konstruktionswörter $K_K > K_S$, in der der subjektiven $S_S > S_K$, $S_S > S_G$, $S_S > S_T$ und $G_G > G_K$, $G_G > G_S$, $G_G > G_T$ in der der gnomischen sowie $T_T > T_K$, $T_T > T_S$, $T_T > T_G$ in der der thematischen Wörter.

Die Gültigkeit von $K \doteq K$ und $T \doteq T$ in vielen Sprachen läßt sich durch auf mathematischem Wege erzielte Informationen bestätigen, während die von $S \doteq S$ und $G \doteq G$ nicht in dem Maße gesichert ist. Je umfangreicher der untersuchte Text ist, desto stärker macht sich der Sättigungsgrad der vorderen Felder mit Elementen der Gruppen K und S bemerkbar. In typologisch sehr verschiedenen Sprachen, z.B. in den agglutinierenden oder rein analytischen, kann eine etwas andere Situation vorliegen. Experimentell läßt sich jedoch auch die relative Gültigkeit der Definitionen 3 und 4 zeigen.

Die stilistische Charakteristik der Konfigurationen wollen wir im folgenden behandeln. Gelten in einem konkreten Text für die vertikale Wortdistribution die Definitionen $K \doteq K$, $S \doteq S$, $G \doteq G$ und $T \doteq T$, so kann die festgestellte absolute Frequenz kein stilistischer Indikator sein. Sie stellt vielmehr eine inhärent neutrale Folge dar, an der sich die stilistische Funkti-

on anderer Mittel manifestiert. Die Frequenz steht nur dann in Opposition, wenn noch weitere Faktoren den individuellen stilistischen Charakter der Äußerung herbeiführen. In einer solchen Disharmonie erlangt die vertikale Wortdistribution den Wert der Merkmalhaftigkeit und es entsteht die Situation, daß ein neutrales und passives Element positiven stilistischen Wert erhält.

Unterscheidet sich die Konfiguration der vertikalen Wortdistribution eines untersuchten Textes von den parametrischen Definitionen, so wird die Frequenz so oder anders merkmalshaft.

Gilt für den untersuchten Text $K_S > K_K$, so ist er potentiell expressiv, subjektiv oder Konversationstext, dynamisch, auf eine Situation bezogen. Für den Wortschatz aller schriftlichen Äußerungen Puškins¹⁷ z.B. gilt die Hypertrophie von S (die Untersuchung stützt sich auf 21 197 Wörter in der Form, wie sie im Wörterbuch erscheinen und die auf Texte zurückgehen, die 544 777 Wörter in der durch den Kontext determinierten Form enthalten): $K_S < K_K \approx K$, $S_K < S_S \approx S$, $S_G < S_S \approx S$. In Puškins Wortschatz dominieren also subjektive, auf die Situation bezogene, dynamische und expressive sowie Konversationswörter. Diese drängen die Konstruktionswörter weit zurück. Die Hypertrophie der Menge S ist für den Wortschatz romantischer Werke überhaupt typisch und gilt, wie empirisch festgestellt wurde, für slowakische, polnische und deutsche Werke in gleichem Maße.

Durch genaue Zählungen läßt sich zeigen, daß eine ähnliche Situation auch in den Werken zeitgenössischer slowakischer Dramatiker (Solovič, Karvaš, Stodola, Barč-Ivan, Bukovčan u.a.) vorliegt, womit der Beweis erbracht ist, daß in Dialogen generell die Elemente S überwiegen.

Ähnlich wie die - im Verhältnis zu der recht konstanten Zusammensetzung der Menge K - labile Zusammensetzung der Menge S signalisiert auch die der Menge S (besonders in Kontrast zu T) einen spezifischen Textcharakter. Die Fluktuation der Elemente aus den Mengen S und T ist offenbar ein signifikantes Indiz für den jeweiligen lexikalischen Aufbau einer Äußerung. Nicht zufällig erscheinen beide in komplementärer Verteilung: in einer Konversation haben die Elemente S das Übergewicht über die der

Menge T , während in einem wissenschaftlichen Text die Elemente der Menge T überwiegen.

Eine sehr bedeutende Erscheinung ist das Eindringen thematischer Wörter in die Mengen K und S . Je kompakter in begrifflicher und syntaktischer Hinsicht der betreffende Text ist, desto umfangreicher sind die Mengen T_T und S_T . Aus der theoretischen Darstellung *Dejiny robotnickeho hnutia* von M. Gosiorovský wählen und untersuchen wir einen 37 240 Wörter umfassenden Text. Die Mengen K und S sind darin folgendermaßen verteilt: $K = 12$ Elemente K , 2 Elemente S , \emptyset Elemente G und 6 Elemente T , $S = 14$ Elemente K , 5 Elemente S , 3 Elemente G und 28 Elemente T . In der Arbeit *Predhegelovská logika* von V. Filkorn sind $K = 12 : 3 : 5 : \emptyset$ und $S = 15 : 10 : 4 : 21$.

Statistische Angaben über die vertikale Wortdistribution in informativen, publizistischen und volkstümlichen Äußerungen zeigen, daß die komplementäre Verteilung der Elemente der Mengen S und T zugunsten der Elemente T entschieden wird. In der Menge S sind die folgenden Relationen wirksam: $S_S < S_K$, in G ist $G_T < G_K$, weiterhin $G_T < G_S$ und $G_T > G_G$, während in der Menge K immer $K_K > K_S$ gilt. Vereinfacht gesagt: für den informativen, publizistischen und volkstümlichen Stil ist die merkmallose, normale Dispersion der K -Wörter, die verringerte Anzahl der S -Wörter und das verstärkte Vorkommen der G - und T -Wörter typisch.

Eine gegensätzliche Verteilung liegt in der Poesie vor. Wir untersuchen Gedichtsammlungen von J. Smrek, V. Mihálik, P. Horov und J. Kostra und stellen - ähnlich wie in Konversationen - eine Hypertrophietendenz zu S fest. So z.B. sieht das Verhältnis $K : S$ in den Mengen K folgendermaßen aus: J. Smrek 13 : 6, V. Mihálik 16 : 4, P. Horov 14 : 5, J. Kostra 14 : 6. Ziemlich stark weichen auch das Verhältnis $K \approx S$ und der Druck der Menge T auf die Mengen S und G , ja sogar K , vom theoretischen Durchschnitt ab. Die Elemente der Menge T werden jeweils auf folgende Weise ersetzt: bei Smrek 1, 3, 16, Mihálik \emptyset , 9, 18, Horov 1, 16, 22. Diese Angaben verweisen deutlich auf eine Atrophie von Konstruktionswörtern in poetischen Texten. Diese basiert auf dem geringen Vorkommen von Syntagmata und damit auf der Hy-

hypertrophie von autosemantischen, begrifflichen Wörtern.

In der Prosa ist der Charakter der vertikalen Wortdistribution eng an das Genre gebunden. In längeren Formen haben die Elemente G das Übergewicht über die Elemente T, während in kürzeren Formen das Gegenteil vorliegt.

Die Konfigurationen der 110 dominierenden Wörter unterscheiden sich auch in den einzelnen Teilen einer Form. So z.B. überwiegen in den Rahmenpartien eines Romans die Elemente K und T, während sie in den inneren Teilen nur schwach, G und S dagegen aber stark vertreten sind. Eine größere Symmetrie herrscht in der Erzählung. In den publizistischen Genres sieht die Verteilung so aus: feuilletonistische Äußerungen sind reich an S, Reportagen an G, informative Texte an T, analytische an K und G. Im publizistischen Stil werden - im Gegensatz zum wissenschaftlichen - die Elemente G und T betont. T ist außerdem durch faktographische Angaben bedingt.

Wir untersuchen die Distribution der 110 dominierenden Wörter im slovakischen Volkslied und stellen fest, daß sie 59% des gesamten Textes ausmachen. Als Material dazu dient uns die 102 Volkslieder umfassende Sammlung *Zaspievajte si* (1958) von J. Geryk. Die Zahl der Wörter in der im Wörterbuch erscheinenden Form beträgt 1 245, die derer in der durch den Kontext determinierten 5 987. Sie sind wie folgt verteilt: K 41, S 36 und G 33. Es besteht also eine Hypertrophie von S-Wörtern, ebenfalls von G- (bzw. auch T-) sowie eine Atrophie von K-Wörtern. Diese Situation läßt den Schluß zu, daß die Volkslieder a) in ihrer Lexik ungewöhnlich expressiv und b) in Bezug auf den Kontext inkohärent und akontinuierlich sind. Das erste Merkmal ist für die volkstümliche Dichtung, das zweite für die Poesie allgemein typisch.¹⁸

Die Angaben zeigen, daß das Eindringen von Wörtern aus den unteren Regionen in die Menge der 110 häufigsten Wörter stilistisch relevant ist. Die grundlegende parametrische Anordnung (1. K, 2. S, 3. G) ist somit nicht Ausdruck einer stilistischen Färbung nur des untersuchten Textes. Sie ist jedoch auch keine stilistisch neutrale Zusammenstellung, sondern rein theoretischer Natur, eine Konfiguration, die auf der Abstraktion von

Einheiten aus thematisch und stilistisch verschiedenartigen Texten beruht. Um die vertikale Wortdistribution als typologisches Mittel verwenden zu können, ist es nicht notwendig, die Wortfrequenz eines gesamten Textes zu untersuchen. Informationen hierüber können bereits anhand von kleineren Textabschnitten gewonnen werden. Es müssen auch nicht immer die 110 ersten Wörter herangezogen werden, da das Verhältnis 2 : 5 : 4, d.h. $2/11 : 5/11 : 4/11$ (oder auch abgerundet zu einer Harmonie der Prozente 2 : 4 : 4), in jedem kleineren Abschnitt als Maßstab dienen kann. Man muß jedoch bedenken, daß die Zuverlässigkeit der Information (d.h. des Ergebnisses) mit der Größe des untersuchten Textes konform geht.

Horizontale Wortdistribution

Die Wortdistribution kann auch ohne Bezugnahme auf die Bedeutung der Wörter und Schichten untersucht werden. In diesem Fall interessiert allein ihre quantitative Anordnung. Diese nennen wir *horizontal*, weil es sich um die Verteilung auf der Textfläche und nicht um vertikal angeordnete Schichten handelt.

Die horizontale Wortdistribution bezieht sich auf Wortwiederholungen, den Umfang des untersuchten Wortschatzes, das Vorkommen okkasioneller und thematischer Wörter, die Diffusion, Variabilität und Prädiktabilität eines Wortes u.ä. Die genannten Faktoren signalisieren die Qualitäten des Textes, des Autors und seiner Zeit.

Bei einer Untersuchung der horizontalen Wortdistribution muß man berücksichtigen, daß ein Wort in zwei Formen vorkommen kann: a) in der durch den Kontext determinierten, b) in der im Wörterbuch verzeichneten Form. Ein Wort erscheint in Texten nur selten in der im Wörterbuch (wo es normalerweise als Stichwort fungiert) auftretenden Form. Wir bezeichnen die Mengen solcher Wörter wie z.B. *Freude*, *Flasche* usw. daher mit dem Buchstaben L (= Einheiten der *langue*).

Das Wort in seiner durch den Kontext determinierten Form ist ein Teil des Textganzen. Diese Form wird durch einen bestimmten

Raum begrenzt. Von daher betrachten wir sie auch nicht als Stichwort, sondern als materielles Textelement. Das slovakische Wort *len* (dt. *nur, bloß, allein*) z.B. kann in einem konkreten Text als vom Kontext determinierte Form, als Materie, mehrmals auftreten. Ein Text ist immer reicher an Wörtern dieser Form, die wir mit dem Buchstaben N bezeichnen. Die horizontale Wortdistribution beschäftigt sich also mit N- und L-Wörtern, d.h. mit Wörtern der Wertigkeit N und L.

Die Begründung für eine separate Untersuchung der horizontalen Wortdistribution liegt in der Anordnung der Wörter nach einer pyramidenförmigen Gesetzmäßigkeit. Wir erläutern dies am Beispiel der Novellensammlung *Tmavý kút* von V. Mináč. Hier treten 52% der Wörter je 1mal, 33% 2-5mal, 9% 6-15mal, 4% 16-50mal, 1% 51-100mal und 1% mehr als 100mal auf. Die genaue Darstellung dieser Gesetzmäßigkeit sieht folgendermaßen aus:

Frequenz	Zahl der L-Wörter	Wörter in %
1	2726	52,1
2 - 5	1724	32,9
6 - 15	492	9,4
16 - 50	197	3,8
51 - 100	47	0,9
101 - 500	44	0,8
501 - 1581	6	0,1

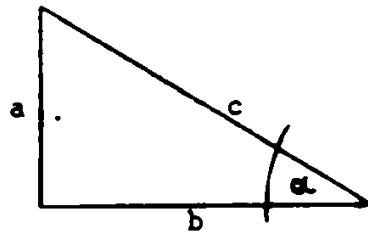
Wir untersuchen fünf Texte, die sich in stilistischer Hinsicht sehr unterscheiden, mit dem Ziel, die Differenzen in der distributiven Gesetzmäßigkeit zu zeigen. Die erwähnte Novellensammlung (N 37 029, L 5 236) steht dabei stellvertretend für die künstlerische Prosa, J. Kostras Gedichtsammlung *Báseň* (N 34 729, L 6 687) für die Poesie, Tajovskys Schauspiel *Zenský zákon* (N 10 088, L 1 273) für das Drama, das Lehrbuch *Prehl'ad stredoškolskej matematiky* (N 21 319, L 922) für die wissenschaftliche und die Tageszeitung *Práca* (N 14 404, L 3 593) für die journalistische Prosa. Es zeigen sich folgende Gesetzmäßigkeiten (ausgedrückt in %):

Frequenz	Künstl.	Prosa	Poesie	Drama	wissenschaftl.	u. journal. Prosa
1	52,1	53,6	51,8	28,2	55,8	
2 - 5	32,9	32,6	30,7	32,3	34,0	
6 - 15	9,4	9,1	11,0	16,8	7,2	
16 - 50	3,8	3,5	3,5	13,8	2,2	
51 - 100	0,9	0,7	1,7	4,7	0,4	
101 - 500	0,8	0,4	1,3	4,0	0,3	
500 u. mehr	0,1	0,1	0,1	0,3	0,0	

Die pyramidenförmige Gesetzmäßigkeit in der Wortdistribution ist für die Typologisierung eines Textes deshalb so wichtig, weil sie auf die deutliche Asymmetrie in der horizontalen Anordnung hinweist. Wörter, die 1-5mal vorkommen, nehmen 6/10 bis 9/10 der Textfläche ein. Die größte Fläche beanspruchen diese Wörter in publizistischen Texten (die o.a. Statistik spricht von 9/10). In künstlerischen Texten besetzen sie 8/10 der Textfläche, in wissenschaftlichen mit 6/10 die geringste Fläche.

Die pyramidenförmige Distribution verweist auf häufiges Vorkommen von Syntagmata und thematische Buntheit. Je größer die Zahl der Syntagmata, desto geringer ist auch die Zahl der Wörter mit hoher Frequenz. In der Journalistik kommen beide Tendenzen, in der Poesie nur die erste zur Wirkung. Im Gegensatz hierzu steht die wissenschaftliche Prosa, die zwar reich an Syntagmata, aber monothematisch ist.

Diese Art der Distribution läßt sich mit Hilfe des Tangensatzes beschreiben (a kennzeichnet die Zahl der Wörter mit der Frequenz 1 (fr_1) und b die Anzahl der L-Wörter): $tg \alpha = \frac{fr_1}{L}$:



Diese Angabe ist aber nur dann richtig, wenn Texte mit entweder gleichem L oder aber gleichem fr_1 verglichen werden. Je kleiner L, desto kleiner ist auch der Winkel α .

Der Index der Wortwiederholungen. Bereits im kleinsten Kontext, im einzelnen Satz, können sich einzelne Wörter wiederholen. Je umfangreicher ein Text ist, desto häufiger sind diese Wiederholungen und desto mehr Wörter werden wiederholt. Die Wortwiederholung hängt primär von der Textlänge und sekundär vom Umfang des Wortschatzes, mit dem im Text gearbeitet wird, ab.¹⁹ Wir bezeichnen die Textlänge mit dem Buchstaben N, den Wortschatzumfang mit dem Buchstaben L. Aus dem Verhältnis der beiden Größen zueinander läßt sich der Index der Wortwiederholung berechnen:

$$I_i = \frac{N}{L}$$

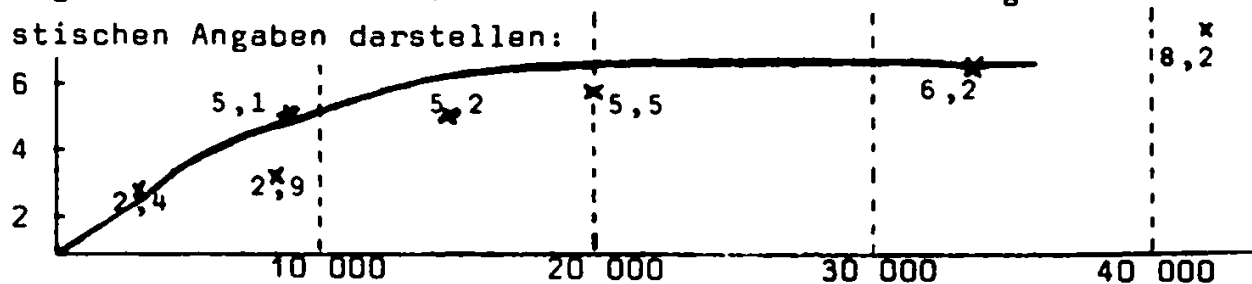
Die Abhängigkeit dieses Indexes (Ii) von der Textlänge kann durch den Vergleich mehrerer Texte eines Autors bewiesen werden. Wir wählen 18 Werke (5 slowakische Autoren und 5 slowakische Zeitschriften) und erhalten das folgende Ergebnis:

Quellen	N	L	Ii
Tajovský: <i>Matka</i>	3 937.	845	4,7
<i>Hriech</i>	5 210	1 081	4,8
<i>Ženský zákon</i>	10 086	1 273	8,3
Karvaš: <i>Diplomati</i>	12 529	2 627	4,8
<i>Jasva</i>	15 037	2 541	5,9
<i>Antigona a tí druhí</i>	17 321	2 671	8,8
Solovič: <i>Polnoc bude o päť minút</i>	8 704	1 703	5,1
<i>U nás také obyčaj</i>	13 021	2 198	5,9
Chudoba: <i>Kde pijú duhy</i> (Auszug)	2 600	1 075	2,4
<i>Kde pijú duhy</i> (gesamte Novelle)	22 551	4 615	4,9
Smrek: <i>Zrno</i>	3 658	1 332	2,7
<i>Božeké uzly</i>	3 935	1 397	2,8
<i>Iba oči</i>	8 171	2 096	2,9
Tageszeitungen: <i>Smena</i>	14 392	4 403	3,3
<i>Pravda</i>	14 404	3 593	4,0
<i>Práca</i>	15 468	3 602	4,3
Zeitschriften für Kinder: <i>Ohník</i>	5 527	1 874	2,9
<i>Pionierske noviny</i>	10 497	2 847	3,7

Wir vergleichen nun den Ii in Texten größerer Spannweite und verschiedener Autoren:

Chudoba: <i>Kde pijú duhy</i>	2 600	1 075	2,4
Zeitschrift: <i>Ohník</i>	5 527	1 874	2,9
Válek: <i>Dotyky, Prit'ažlivost'</i>	8 556	2 072	3,2
Solovič: <i>Polnoc bude o päť minút</i>	8 704	1 703	5,1
Sládkovič: <i>Marína</i>	14 573	2 816	5,2
Horov: <i>Poézia</i>	20 984	3 829	5,5
Karvaš: <i>Antigona a tí druhí</i>	32 358	5 212	6,2
Jašík: <i>Čierne a biele kruhy</i>	42 949	5 206	8,2
Puškins Werke	544 777	21 197	25,7

Die Indices verweisen deutlich auf die Abhängigkeit der Wiederholung vom Umfang des untersuchten Textes. Je größer nämlich der Umfang (N) ist, desto höher ist auch der Ii. Die Steigung verläuft jedoch nicht in einer Geraden, sondern in Form einer logarithmischen Kurve, die wir nun auf der Grundlage der statistischen Angaben darstellen:



Der I_i eines Textes kann nur dann als Signal dienen, wenn man ihn mit den Indices anderer Texte vergleicht. Dabei muß man auch das Verhältnis der Textlängen berücksichtigen. Sehen wir uns z.B. die Punkte oberhalb und unterhalb der Kurve an, so sind die oberen Angaben zu Prosa und Drama, die unteren, auf einen niedrigeren I_i verweisenden, Angaben zu Poesie und lyrischer Prosa (s. Chudoba), die bekanntlich reich an lexikalischen Einheiten ist. Je größer also der I_i , desto stärker ist die Rundung der logarithmischen Kurve und desto höher steigt sie, ebenso auch umgekehrt.

Wir stellen in 102 der bekanntesten Volkslieder (N 5 987, L 1 245) einen I_i von 4,81 fest. Obwohl die Volkslieder polythematisch sind, wiederholen sich in ihnen die für diese Gattung typischen Wörter ausgesprochen häufig.

Auch die objektiven und subjektiven stilbildenden Faktoren üben ihren Einfluß auf den I_i aus, wobei die Kraft der ersteren grundsätzlich stärker als die der zweiten ist. Die durch diese Faktoren hervorgerufenen Abweichungen sind jedoch verhältnismäßig gering.

Ein Vergleich der I_i im tschechischen²⁰ und slovakischen²¹ Frequenzwörterbuch soll ihre Abhängigkeit von den objektiven stilbildenden Faktoren verdeutlichen. Den kleinsten I_i im Tschechischen haben Poesie und Journalistik (3,2-4,9 bei einem Durchschnitt von 4,1). Eine ähnliche Situation beschreibt auch das slovakische Frequenzwörterbuch, indem es als durchschnittlichen I_i für die Poesie 4,3 und für die Journalistik 4,0 angibt. Mittlere Werte finden wir im tschechischen Frequenzwörterbuch für künstlerische Prosa (5,4), künstlerische Jugendliteratur (5,6) und Drama (5,7), was einen durchschnittlichen I_i von 5,56 für künstlerisch-belletristische Texte ergibt. Das slovakische Frequenzwörterbuch nennt für künstlerische Prosa den Wert 5,6 und für das Drama 5,7. Es liegt also wiederum eine ähnliche Situation wie im Tschechischen vor. Hohe Indexziffern weist das tschechische Frequenzwörterbuch für mündliche Äußerungen (6,3), wissenschaftliche Texte (6,3) und Fachliteratur (ebenfalls 6,3) auf. Da die genauen Größen zwischen 6,2 und 6,3 schwanken, erhalten wir einen Durchschnittswert von 6,25. Auch

das slovakische Frequenzwörterbuch ordnet der wissenschaftlichen Prosa die höchste Indexziffer zu. Hier beträgt sie allerdings durchschnittlich 7,5.

Auch die Indexziffern, die unter Berücksichtigung der stilistischen Methode und der stilistischen Form errechnet werden, beweisen die Abhängigkeit des *Ii* von den objektiven stilbildenden Faktoren. Wir vergleichen sie jeweils in zwei Texten, die ihrem Umfang nach gleich, dem Genre nach jedoch verschieden sind:

	Quellen	N	L	<i>Ii</i>
1	Zeitschrift: <i>Kultúrny život</i> (Publizistik)	38 642	9 218	4,5
	M. Urban: <i>Živý bič</i> - Teil 1- (künstlerische Prosa)	38 917	5 742	6,8
2	R. Moric: <i>Trikrát som ušiel</i> (Kinderprosa)	26 134	3 615	7,2
	Zeitschrift: <i>Pravda</i> (Publizistik)	26 721	5 499	4,9
3	D. Tatarka: <i>Rozhovory bez konca</i> (künstlerische Prosa)	32 175	4 364	7,3
	Ilkovič: <i>Fyzika</i> (wissenschaftliche Prosa)	32 982	2 336	14,1
4	Solovič: <i>U nás taká obyčaj</i> (Drama)	13 021	2 198	5,9
	Luby: <i>Základy všeobec. súkr. práva</i> (wissenschaftliche Prosa)	13 089	1 660	7,9
5	Ondruš: <i>Šialený mesiac</i> (Poesie)	3 805	1 265	3,0
	Tajovský: <i>Matka</i> (Drama)	3 937	845	4,7

Im ersten Fall ist - wie die Tabelle zeigt - bei fast völliger Identität der Umfänge der *Ii* der künstlerischen Prosa (Roman) um 50% höher als der des publizistischen Stils (für die statistischen Angaben über die Zeitschrift *Kultúrny život* wurden künstlerische Artikel nicht verwertet). Die stilistische Methode als Ergebnis der objektiven stilbildenden Faktoren beeinflusst den *Ii* in einem beträchtlichen Maße.

Auch im zweiten Fall unterscheiden sich die Indices trotz annähernder Identität der Umfänge erheblich. In Kinderprosa (Romanen) wiederholen sich die Wörter als Elemente eines funktio-

nell beschränkten Wortschatzes viel häufiger als in der - im Wortschatz verhältnismäßig reichen - *Pravda*. Ähnliche Unterschiede liegen auch den Beispielen 3 - 5 zugrunde. Der Einfluß der objektiven stilbildenden Faktoren, der stilistischen Methode und der stilistischen Form ist also augenscheinlich.

Wenn - wie bereits festgestellt - die subjektiven stilbildenden Faktoren die Indices in weit geringerem Maße als die objektiven beeinflussen, so ist doch auch ihre Wirkung unverkennbar. Als Beweis dafür dienen uns jeweils zwei Werke, die in Genre und Umfang identisch sind, aber von verschiedenen Autoren stammen, sich also durch die subjektiven stilbildenden Faktoren unterscheiden:

	Quellen	N	L	Ii
1	Solovič (Dramen): <i>Polnoc bude... , Kar na konoí... , U nás taká obyčej</i>	32 014	5 970	5,4
	Karvaš (Dramen): <i>Antigona... , Jazva</i>	32 358	5 212	6,2
2	Solovič (Drama): <i>Polnoc bude...</i>	8 704	1 703	5,1
	Králík (Drama): <i>Poeledná překážka</i>	9 130	1 429	6,4
3	Mihálik (Poesie): <i>Lyrika</i>	23 460	5 156	4,6
	Horov (Poesie): <i>Poésia</i>	20 984	3 829	5,5
4	Chudoba (künstlerische Prosa): <i>Kde pijú dýhy</i>	22 551	4 615	4,9
	Moric (künstlerische Prosa): <i>Trikrát som ušiel</i>	26 134	3 515	7,2
5	Mináč (künstlerische Prosa): <i>Tmavý kút</i>	37 029	5 236	7,1
	Urban (künstlerische Prosa): <i>Živý bič</i>	38 917	5 742	6,8

Wir haben nur künstlerische Werke herangezogen, da die subjektiven stilbildenden Faktoren ihre Wirkung in diesem Typ der Wortkunst am deutlichsten ausüben. In sachbezogenen und journalistischen Werken ist ihr Einfluß auf ein Minimum - ja Null - reduziert.

Die Abhängigkeit des Ii von den subjektiven stilbildenden Faktoren ist auch aus den Angaben des tschechischen Frequenzwörterbuchs (S. 31) ersichtlich. Auch hier werden unterschiedliche Indices für Texte mit annähernd identischem Umfang genannt.

Während sich der Ii in sachbezogener Prosa - unter Berücksichtigung von Umfang, Form und stilistischer Methode - ziem-

lich genau bestimmen läßt, unterliegt er in künstlerischer Prosa - selbst in Werken, die in Umfang, stilistischer Methode und Komposition übereinstimmen - größeren Schwankungen. In sachbezogener Prosa ist er ein Signal für die Werte des Textes, die durch die objektiven stilbildenden Faktoren - besonders auch durch die Funktion - bedingt sind. In künstlerischer Prosa verweist er auf Sprache und Stil des Autors, wenngleich auch dieser von den objektiv gegebenen Bedingungen beeinflusst wird. In beiden Texttypen jedoch wird der l_1 am stärksten von der Textlänge N modifiziert, erst sekundär kommen Funktion und übrige objektive stilbildende Faktoren zur Wirkung und an letzter Stelle macht sich die Persönlichkeit des Autors bemerkbar.

Gravität des Wortschatzes

Bei der Analyse der horizontalen Wortdistribution muß man diejenigen Wörter als eine Klasse ansehen, die jeweils nur einmal vorkommen, d.h. sich nicht wiederholen. Wir bezeichnen sie als okkasionelle Wörter.

Die Mehrzahl der wiederholten Wörter wird vom Autor souverän gebraucht. Sie bilden seinen aktiven Wortschatz, denn die Wiederholung eines Wortes setzt voraus, daß der Autor mit seiner Bedeutung bestens vertraut ist. Dagegen sind die nur einmal gebrauchten Wörter potentielle Zitate, die kaum zum aktiven Wortschatz gehören, wenn man diese Möglichkeit auch nicht ganz ausschließen kann, weil bestimmte Themenstellungen bestimmte Wörter nicht erfordern. (Wir denken hierbei an die Typologisierung umfangreicher Werke wie Romane und wissenschaftliche Monographien.) Auf jeden Fall besteht im Wortschatz des Autors ein beträchtlicher Unterschied zwischen den okkasionellen Wörtern f_1 und den wiederholten $f_{1<}$. Die Zahl der Wörter $f_{1<}$ ist normalerweise etwas höher als die Differenz $\frac{L}{2}$. Sehen wir uns das Verhältnis von L zu den Wörtern $f_{1<}$ in einigen Texten an:

Eine Ausgabe der *Zbierky zákonov ČSSR* 1 747 : 941,
Rol'nióke noviny 3 571 : 1 629,
Marína von Sládkovič 2 816 : 1 294,
Čítanka pre 1. ročník základnej školy 1 267 : 567.

Um zu zeigen, daß jeder Autor die Wörter f_1 - im Gegensatz zu $f_{1<}$ - mit einer gewissen Gesetzmäßigkeit benutzt, nennen wir

ihr prozentuales Vorkommen in je drei verschiedenen Werken der folgenden Autoren:

Smrek :	Solovič :	Tajovský :
<i>Božské uzly</i> 36%	<i>U nás také obyčej</i> 45%	<i>Matka</i> 42%
<i>Iba oči</i> 36%	<i>Polnoc bude...</i> 40%	<i>Hriech</i> 42%
<i>Zrno</i> 36%	<i>Kar na konci...</i> 42%	<i>Ženský zákon</i> 48%

P. Guiraud spricht in diesem Zusammenhang von der Konzentration des Wortschatzes. Er versteht darunter das Verhältnis der am häufigsten gebrauchten Wörter zum Umfang des Textes (N). Zur Bestimmung des theoretischen Wertes der Konzentration (C), die nach seiner Berechnung 0,09 beträgt, geht er von der Summe der 50 ersten dominierenden autosemantischen Wörter aus und wendet die folgende Formel an:

$$C = \frac{\sum \frac{1}{50}}{N}$$

Aufgrund dieser Formel beträgt die Konzentration in den Werken *Prútené kreslá* und *Rozhovory bez konca* von D. Tatarka 0,09 und 0,12, in der Gedichtsammlung von Smrek 0,08.

G.U. Yule legt zur Bestimmung der sogenannten charakteristischen Konstante (K) die folgende Formel fest (n_r bezeichnet hier die Zahl der Wörter mit der Frequenz r)²²:

$$K = \frac{\sum_{r=1}^s r^2 n_r}{N^2} - \frac{1}{N} \quad \text{d.h. vereinfacht:} \quad K = \frac{\sum r^2 n_r}{N^2}$$

Guirauds Konstante benennt den Grad der Explizität eines Textes. Je größer die Zahl der grammatischen Wörter ist, desto mehr Beschreibungen und Erklärungen kommen vor und desto kleiner ist C, je kompakter ein Text in lexikalischer Hinsicht ist, desto größer ist C. Diese Größe (C) wird jedoch auch von der thematischen Breite des Textes beeinflusst: je bunter ein Text in thematischer Hinsicht ist, desto kleiner ist C (und umgekehrt). Am größten ist C in Monographien.

Wir bezeichnen die von Guiraud *Konzentration* genannte Eigenschaft als *Gravität* und berechnen sie auf andere Weise.

Unter der Gravität eines Wortschatzes verstehen wir den Grad der grammatischen Aktivität der Wörter, die die Grundlage eines Textes bilden. Mit ihrer Hilfe wird die Problematik, die Gegenstand der Äußerung ist, erhellt.

In der kursorischen, auf den Makrokontext gerichteten Kommunikation - wie z.B. in längeren künstlerisch epischen Texten - ist die Gravität des Wortschatzes relativ klein, die durchschnittliche dem einzelnen Wort auferlegte Belastung gering. Zwischen den Wörtern ist viel Raum, der von nicht wiederholten Wörtern ausgefüllt wird. Zur Berechnung des Indexes der Gravität (lg) muß man die folgenden Prämissen in Erwägung ziehen: Die Gravität hängt von der Zahl der wiederholten Wörter und der Länge des betreffenden Textes ab. Wichtig sind vor allem die thematischen Wörter. Deshalb ziehen wir - im Gegensatz zu Guiraud, der sich bei der Berechnung von K nur auf die ersten 50 dominierenden Wörter stützt - alle diejenigen in Betracht, die der Autor souverän benutzt, d.h. deren $l_i f_{1<}$ ist. Da die Funktionswörter annähernd 50% eines Textes ausmachen, multiplizieren wir die Wörter $f_{1<}$ mit 2 ($L_{f_{1<}}$ gibt die Zahl der Wörter in den im Wörterbuch erscheinenden Formen, die im Text eine Frequenz von mehr als 1 haben, an.). Um nun den Index in ganzen Zahlen ausdrücken zu können, rechnen wir im Nenner mit $1/10 N$. Unsere Formel zur Berechnung der Gravität lautet also:

$$I_g = \frac{20 L_{f_{1<}}}{N}$$

Wir führen den lg einiger Texte an:	N	L	L_{f_1}	lg
Kostra (Poesie): <i>Báseň</i>	34 729	6 667	3 505	1,8
Chudoba (künstlerische Prosa): <i>Kde pijú dúhy</i>	22 551	4 615	2 441	1,9
Journalistische Prosa: <i>Noviny Vpřed</i>	9 038	2 881	1 490	3,1
Moric (Kinderprosa): <i>Trikrát som ušiel</i>	26 134	3 615	1 820	1,4

Aufschlußreich ist die Gravität von Texten, die von paranoischen oder schizophrenen Personen stammen. Hier ist der Prozentsatz von L-Wörtern mit f_1 verhältnismäßig hoch: er beträgt bis zu 75% und steigt mit dem Grad der Krankheit. Dies hängt mit dem Verlust der Determinationsfähigkeit eng zusammen. Ein geringer lg kann also als Signal dieses Verlustes dienen.

Der Ig muß streng vom Ii unterschieden werden. Während zur Berechnung des Ii alle Wörter L in Bezug auf das Textganze N herangezogen werden, benötigt man für den Ig nur die Wörter mit $f_{1<}$ ($L_{f_{1<}}$) im Gegensatz zum Textganzen. Von daher ist es auch

verständlich, daß z.B. die journalistischen Texte normalerweise durch einen sehr niedrigen l_i gekennzeichnet sind, wohingegen ihr I_g sich meistens als verhältnismäßig hoch erweist. Ähnliches gilt auch für einen Vergleich zwischen Poesie und künstlerischer Prosa.

Faktoren, die die Höhe des I_g beeinflussen, sind die stilistische Methode, die Elemente der Mikrostruktur und der Autor. Bei der Typologisierung eines Textes dient der I_g als Signal für die symmetrische Wortdistribution. Je höher der I_g ist, desto geringer ist die Symmetrie und desto unregelmäßiger die Distribution.

Lexikalische Exklusivität eines Textes

Die Exklusivität (E) ist der der Gravität entgegengesetzte Faktor. Während sich die Gravität auf das Verhältnis der wiederholten Wörter $f_{1<}$ zur gesamten Fläche (N) bezieht, betrifft E das Verhältnis zwischen den Wörtern f_1 und der Fläche. Auch hier gilt die Voraussetzung, daß ca. 50% eines Textes aus Funktionswörtern bestehen und deshalb bei der Berechnung des Exklusivitätsindex (I_e) die Zahl der Wörter f_1 mit 2 multipliziert werden muß. In einem Wortschatz läßt sich also die Dispersion der Wörter f_1 minus Funktionswörtern bestimmen. Um den I_e in ganzen Zahlen angeben zu können, rechnen wir im Nenner mit 1/10 (L_{f_1} bezeichnet die Zahl der Wörter, die im Text die Frequenz 1 haben). Die Formel lautet also:

$$I_e = \frac{20 L_{f_1}}{N}$$

I_e und I_g stehen zueinander nicht in der direkten Beziehung, daß mit steigendem I_e der I_g fällt und umgekehrt. Wir geben ein Beispiel dafür. In zwei Texten mit identischem Umfang von 1 000 Wörtern in der durch den Kontext determinierten Form liegt folgende Situation vor: in dem einen Text werden 300 Wörter nicht, 200 jedoch wiederholt, L beträgt also 500. Im zweiten werden je 250 Wörter wiederholt bzw. nicht wiederholt (L beträgt also ebenfalls 500). In beiden Fällen ist bei N 1 000 und L 500 der l_i 2. Im ersten Text jedoch beträgt der I_g 4 und der I_e 6, während im zweiten beide den Wert 5 aufweisen. Wären aber in den

Texten L_{f_1} und $L_{f_1<}$ verschieden, so könnte auch die Summe von I_e und I_g (in unserem Beispiel = 10) nicht identisch sein.

Wir nennen den I_e für die Texte, bei denen wir bereits den I_g feststellten (vgl. S. 57):

	N	L	L_{f_1}	I_g	I_e
Kostra (Poesie)	34 729	6 687	3 505	1,8	2,0
Chudoba (künstlerische Prosa)	22 551	4 615	2 441	1,9	2,2
Journalistische Prosa	9 038	2 881	1 490	3,1	3,3
Moric (Kinderprosa)	26 134	3 615	1 820	1,4	1,4

Anhand der Tabelle sehen wir, daß I_g und I_e zwar darin übereinstimmen, daß sie gleichzeitig größer bzw. kleiner werden, jedoch nicht in gleichem Maße. Darauf übt der Umfang der Texte einen gewissen Einfluß aus.

Nennen wir den I_e noch weiterer Texte. In der wissenschaftlichen Arbeit Filkorns stellen wir einen Wert von 0,7, in zwei Dramen von Karvaš 1,3, in Urbans Prosa *Živý bič* 1,5, in der Erzählung Kukučins *Ked' báčik z Chodol'ava umrie* 2,6 und in einer Nummer der Provinzzeitung *Hlas l'udu* 4,4 fest.

Der I_e verweist auf die Diffusion von Wörtern f_1 in einem Textganzen. Es ist daher verständlich, daß er in der Poesie, in lyrisierter Prosa und in publizistischen, informativen Texten mit reicher Thematik sehr hoch ist. In theoretischen Erörterungen, spontanen mündlichen Äußerungen, im Dialog, in Kinderprosa und in Prosa überhaupt ist er aufgrund des Reichtums an Dialogen und Reden dagegen sehr niedrig. In Beschreibungen ist er höher als in Erörterungen und Erzählungen. Ein hoher I_e ist also ein Merkmal des essayistischen Stils und belletristisch-journalistischer Äußerungen wie Feuilletons, Reportagen usw.

Man muß sich eine Textfläche - um der Anschaulichkeit willen - so vorstellen: die Wörter $f_{1<}$ erscheinen auf ihr als dunkle, die mit f_1 als helle Stellen. Die Summe der dunklen Stellen entspricht der Gravität, die der hellen der Exklusivität des Textes. Während es sich bei der Gravität um die Intensität der dunklen Stellen handelt, bezieht sich die Exklusivität nur auf ihre Quantität. Ausrichtung und Konzentration auf dunkle Stellen zeugen von einem tiefen Einblick in die Sache, von Besinnung, Wiederholung und Retardierung des Tempos, während ihre Vernachlässigung, d.h. die Anreicherung der Fläche mit hellen Stellen, hohes Texttempo, Akzeleration, bedeutet. Daher haben

statistische Texte einen relativ hohen I_g , dynamische hingegen einen hohen I_e .

Die Summe von I_g und I_e entspricht der Extensität des Wortschatzes, wobei der erste Index als Signal für die Art der Textanordnung sowie als Information über die Zahl der wiederholten Wörter dient, während der zweite die Verteilung der peripheren lexikalischen Einheiten widerspiegelt.

So sind z.B. bei Kostra der I_g 1,8 und der I_e 2,0, was einen Index der Extensität (I_{ext}) von 3,8 ausmacht. Dieser beträgt bei Chudoba 4,1, bei Moric 2,8, in der Zeitschrift *Vpred* sogar bis zu 6,4. Diese Angaben erhält man mit Hilfe der folgenden Formel:

$$I_{ext} = \frac{20 L}{N} ,$$

wobei eigentlich:
$$I_{ext} = \frac{20 L_{f_{1<}}}{N} + \frac{20 L_{f_1}}{N} + \frac{20 L}{N}$$

Die Extensität des Wortschatzes verweist auf den lexikalischen Umfang, auf Ausdrucksvariationen und lexikalische Kompaktheit. Sie darf nicht unter ästhetischen Überlegungen, sondern muß unter dem Aspekt der funktionellen Angemessenheit gesehen werden, denn der ästhetische Wert eines Textes ist von der Höhe des I_{ext} unabhängig.

Lexikalische Variabilität eines Textes

Zur Berechnung von Gravität und Exklusivität eines Textes wurde das Verhältnis der Wörter f_1 und $f_{1<}$ unter Berücksichtigung der Textfläche, d.h. der Zahl N der Wörter in ihrer vom Kontext determinierten Form, untersucht. Für den Stil ist jedoch auch die Relation dieser Wörter zu den im Wörterbuch erscheinenden Formen wichtig. Es bleibt also zu untersuchen, wie sich die Bedeutung ändert, d.h. wie die Textvariabilität beschaffen ist.

Wir erklären den Begriff *Variabilität* auf dem Hintergrund der Exklusivität. Nehmen wir z.B. zwei Texte: im ersten sei N 1 000 und L_{f_1} 500, im zweiten N 1 000 und L_{f_1} 400. $L_{f_{1<}}$ sei in beiden Fällen 200. Daraus folgt, daß L im ersten 700, im zweiten Text 600 beträgt. In Bezug auf die Variabilität ist eben diese Dif-

ferenz von 100 Wörtern (700 - 600), um die der eine reicher ist, wichtig. Bei L_{f_1} 500 und L 700 bzw. L_{f_1} 400 und L 600 weist der erste Text eine größere Variabilität auf, da unter den 700 angewandten Wörtern 500 nur einmal vorkommen. Der Grad der Variabilität (V) läßt sich anhand der folgenden Formel berechnen:

$$V = \frac{L_{f_1} \cdot 100}{L}$$

Das bedeutet für den ersten Text eine Variabilität von 71, für den zweiten von 66. Da die Variabilität in Prozenten angegeben wird, könnte man also auch sagen $V_1 = 71\%$ und $V_2 = 66\%$. Mit anderen Worten: 71% des ersten Textes machen ihn "frisch", wirken regenerierend und belebend, während es im zweiten nur 66% sind.

Stellen wir im folgenden die Variabilitätsgrade einiger - in diesem Kapitel schon erwähnter - Texte fest:

Zeitung: <i>Hlas l'udu</i>	59%	Tajovský: <i>Matka</i>	58%
Kukučín: <i>Ked báčik...</i>	59%	Mináč: <i>Tmavý kút</i>	49%
Jašík: <i>Čierne a biele kruhy</i>	44%	Zeitung: <i>Pravda</i>	53%
Zeitschrift: <i>Ohnák</i>	65%	Gosiorovský: <i>Dejiny...</i>	48%

Der Vollständigkeit halber führen wir noch einige Werte aus tschechischem Material an. In der Erzählung *Bratr Zák* von Olbracht beträgt V 58%, in Pujmanovas *Předtucha* 61% und in dem Buch *Život o dílo* von Čapek 58%.

Welche Gesetzmäßigkeit läßt sich nun aus dem Variabilitätsgrad ableiten? Normalerweise haben polythematische Texte wie Reisebeschreibungen, informative Äußerungen und Referate eine höhere Variabilität als monothematische. Diesen Tendenzen steht aber der folgende wichtige Gesichtspunkt gegenüber: subjektiv gefärbte Äußerungen (Jašík, Mináč) haben eine geringere Variabilität als objektive (z.B. Zeitungen und Zeitschriften). Entscheidend ist jedoch auch hier die Länge der untersuchten Stellen. In einzelnen Teilen der Erzählung *Kde pijú dúhy* von Chudoba beträgt die Variabilität 65% (N 2 600, L 1 075 und L_{f_1} 700), während sie in der gesamten Erzählung (N 22 551, L 4 615 und L_{f_1} 2 441) nur 53% ausmacht. Es wirken hier offensichtlich verschiedene Faktoren zusammen. Deshalb sind auch nur die Informationen zulässig, die aus dem Vergleich von Texten mit übereinstimmenden objektiven stilbildenden Faktoren stammen.

Prädiktabilität einer lexikalischen Einheit

Als Prädiktabilität bezeichnen wir die Voraussage dessen, was folgt. In einem Text treten entweder unerwartete, überraschende oder aber bereits erwartete Phänomene auf, deren Existenz latend geahnt wurde. So können auch die lexikalischen Einheiten erwartet oder unerwartet sein. Der Grad der Prädiktabilität (P) hängt nun davon ab, wie oft eine lexikalische Einheit auf der Textfläche (N) als unerwartet erscheint. Wird ein Wort schon zum zweiten Male wiederholt, so kann man von einer gewissen Prädiktabilität sprechen.

Als unerwartet werden jedoch nicht nur die Wörter bezeichnet, die im Text nur einmal vorkommen, sondern auch diejenigen, die neu erscheinen, selbst wenn sie später noch mehrmals wieder aufgenommen werden. Zu den unerwarteten zählen also alle Wörter f_1 und die mit $N - L_{f_1}$, die das erste Mal erscheinen. Ziehen wir die Wörter mit f_1 von der Gesamtzahl ab, so bleiben die im Text wiederholten Wörter übrig. Es muß also die Zahl der wiederholten Wörter errechnet werden. Dies geschieht mit Hilfe der folgenden Formel:

$$P = 100 - \frac{(L_{f_1} + L_{f_{1<}}) 100}{N} = 100 - \frac{100 L}{N}$$

Als Beispiel nennen wir wieder zwei Texte mit $N = 1\,000$. Im ersten sei $L_{f_1} = 400$ und $L_{f_{1<}} = 100$, im zweiten $L_{f_1} = 600$ und $L_{f_{1<}} = 200$. Das ergibt für P_1 einen Wert von 50 und für P_2 von 20. Im ersten Text, in dem die Zahl der Wörter mit f_1 sowie die der wiederholten Wörter kleiner ist, ist P höher (es treten nur 500 Wörter ganz neu auf) als im zweiten, wo häufiger neue Wörter und mehr Fälle zu verzeichnen sind, in denen diese unerwartet sind.

Die Prädiktabilität kann man ebenfalls mit einer anderen Formel errechnen, wobei allerdings die Wörter mit $f_{1<}$ außer Betracht bleiben. Sie lautet:

$$P = 100 - \frac{L_{f_1} \cdot 100}{N}$$

Die erste Formel ist für die Typologisierung und Interpretation künstlerischer, die zweite für Untersuchungen sachbezoge-

ner Texte zweckmäßiger.

Der Rhythmus der Wortwiederholungen (R_i)

Man kann die Wörter mit $f_{1<}$ auch im Hinblick darauf untersuchen, in welchen Intervallen sie sich wiederholen. (Der Rhythmus der Wortwiederholung ist eine Ergänzung des Gravitätsindex.) Die Intervalllänge der durchschnittlichen Frequenz mehrfach aufgenommener Wörter berechnen wir auf folgende Art:

$$R_i = \frac{N - L}{L \cdot f_{1<}}$$

Wir stellen fest, daß der Wert des R_i in Provinzzeitungen 4,5 ist. Dieses Ergebnis kam so zustande: nachdem alle nicht wiederholten oder das erste Mal auftretenden Einheiten von N abgezogen worden waren, ergaben sich Intervalle mit einer Länge von 4,5 Wörtern. Es wurde dann jeweils ein und dasselbe Wort wiederholt. Aus diesem Grunde könnte man den R_i auch als Index der Stereotypie bezeichnen. Als positives stilistisches Element ist der R_i in Äußerungen, in denen das Gewicht nicht auf der Form, sondern auf dem Inhalt liegt - d.h. in spontanen, nicht stilisierten Äußerungen und überall dort, wo vom Adressaten kursorische Lektüre erwartet wird -, immer hoch. Er ist somit Kennzeichen wissenschaftlicher Texte und Konversationen, während ein niedriger R_i informative, künstlerische und belletristische Äußerungen charakterisiert.

P. Guiraud berechnet den Reichtum (R) eines konkreten Textwortschatzes mit Hilfe der Formel

$$R = \frac{L}{\sqrt{N}},$$

wobei N sich aber nur auf die autosemantischen Wörter bezieht. Er gibt als theoretischen Index den Wert 20,50 an. Seine Formel gilt allerdings nur für Werke mit einem Umfang von 15 000 bis 40 000 Wörtern. Sie muß deshalb zur Berechnung eines höheren oder niedrigeren R_i -wertes mit $\pm \log \frac{L}{2N}$ modifiziert werden:

$$R = \frac{L}{\sqrt{N}} \pm \log \frac{L}{2N}$$

Berücksichtigt man auch die Funktionswörter, so gilt:

$$R = \frac{L}{\sqrt{n}} + \log \frac{L}{N}$$

Wir wollen abschließend noch einige Hinweise zur Anwendung von Informationen über die Wortdistribution bei der Typologisierung eines Textes geben:

Die Distribution der lexikalischen Einheiten dient als Signal dafür, daß an der Entstehung einer Äußerung mehrere Faktoren unterschiedlicher Stärke beteiligt sind. Die Untersuchung der Distribution wird damit zum Mittel, typische, den Text prägende Gestaltphänomene aufzuspüren. Sie läßt auch Rückschlüsse auf die Persönlichkeit des Autors zu, ebenfalls auf die Zeit und andere Faktoren, auf die der Wortschatz direkt oder indirekt hinweist. Die den Text konstituierenden lexikalischen Einheiten können nach quantitativen oder semantischen - bzw. quantitativ-semantischen - Kriterien verschiedenen Systemen zugeordnet werden. Es ist dabei sinnvoll, die erstellten Systeme mit den traditionellen Systemen der Lexikologie zu konfrontieren.

Wir zeigten mehrere Methoden, mit deren Hilfe sich die im Text manifestierten Gesetzmäßigkeiten aufdecken lassen. Diese müssen vor ihrer Anwendung gegeneinander abgewogen und auf ihre Relevanz für den konkreten Text hin geprüft werden.²³ Für die Typologisierung sollen nur die Methoden Verwendung finden, die im Hinblick auf Ziel und Situation sinnvoll und ergiebig erscheinen. Wird die Untersuchung von quantitativen Aspekten bestimmt, so darf der Textumfang nicht außer acht gelassen werden. Dies zeigte sich sehr deutlich bei der Berechnung des I_i .

Ein Wort, das im Gesamtsystem der Sprache eine periphere Stellung einnimmt, kann in einem Text durchaus zentrale Bedeutung erlangen. Ebenso ist bei zentralen Wörtern auch das Gegenteil möglich, denn gerade diese Verschiebungen sind überzeugende Signale für die gesuchten Gesetzmäßigkeiten.

Die Untersuchung einer größeren Anzahl von Texten ermöglicht auch die Bestimmung parametrischer Informationen (wie z.B. bei der vertikalen Wortdistribution). Fehlen diese, so muß mit Vergleichen operiert werden.

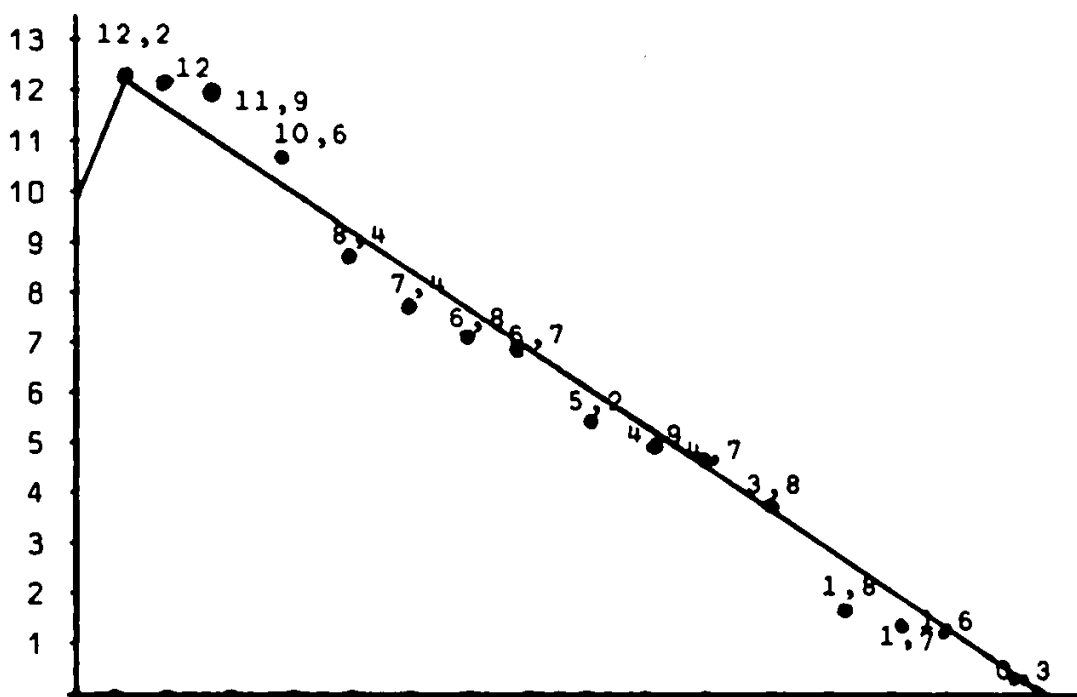
Wir betonen, daß die quantitativen Gesetzmäßigkeiten nicht

als direkte Aussagen zu Absicht und Ästhetik eines Textes verstanden werden können: Urteile dieser Art bleiben allein dem Interpreten überlassen. Die Gesetzmäßigkeiten sagen auch nichts über Kontextbedeutungen und Korrespondenz von Wörtern aus. Sie sind vielmehr tote Zahlen, die erst der Mensch mit Leben erfüllt. Dennoch - oder gerade deshalb - sind sie als exakte und vor allem objektive Indikatoren von großer Bedeutung. Für die Typologisierung als Vorstufe der Interpretation sind sie unersetzlich.

Distribution von Präfixen

Als indirekte Indikatoren des stilistischen Aufbaus einer Äußerung können auch lexikalische Morpheme dienen. Wir untersuchen im folgenden den Zusammenhang zwischen der Distribution von Präfixen und dem stilistischen Aufbau einer Äußerung. Die statistischen Angaben stammen aus Untersuchungen von fünf in stilistischer Hinsicht völlig verschiedenen Texten, die eine Million Wörter in der vom Kontext determinierten Form umfassen. Auf diesem Material basiert auch das slovakische Frequenzwörterbuch. Wir wollen die Gesetzmäßigkeiten feststellen, nach denen die häufigsten Präfixe der Umgangssprache, künstlerischer Prosa, Poesie, publizistischer und wissenschaftlicher Texte verteilt sind, und das systematische Vorkommen einzelner Präfixe in den Stilen - sie können bei der Diagnose und Typologisierung als indirekte Signale dienen - nachweisen. Das Zahlenmaterial sagt nichts über die absolute Frequenz dieser Präfixe in den einzelnen Kontexten aus, sondern bezieht sich nur auf ihr Erscheinen in dem von diesen erfaßten Wortschatz. So z.B. wird das Präfix *na-* für jedes Wort nur einmal und ohne Rücksichtnahme darauf, wie oft dieses Wort im Text auftritt, gezählt.

Die durchschnittliche Frequenz der 16 dominierenden slovakischen Präfixe ohne Berücksichtigung ihrer stilistischen Wertigkeit stellen wir anhand einer parametrischen Kurve (Angaben in %) dar:



za po vy e o na pre u roz od pri do pred v pod nad

Es zeigt sich, daß zwischen der bedeutungsmäßigen Vielfalt und der Anwendung eines Präfixes eine direkte Abhängigkeit besteht: je umfangreicher die Bedeutungsskala ist, desto häufiger kommt das Präfix vor, und umgekehrt. Es dominieren offensichtlich die Präfixe *za, po, vy, e* (1. Frequenzklasse), gefolgt von *o, na, pre, u, roz, od, pri, do* (2. Frequenzklasse), während *pred, v, pod, und nad* (3. Frequenzklasse) seltener sind.²⁴

Die quantitative Analyse beliebiger sprachlicher Einheiten zeigt, daß sich die funktionellen Stile in zwei homogene, deutlich voneinander abgesetzte Gruppen gliedern: 1. Stile, die umgangssprachliche Elemente verwenden (umgangssprachlicher und künstlerischer Stil), 2. Stile, die Elemente der exakten terminologischen Ausdrucksweise benutzen (wissenschaftlicher, publizistischer, administrativer Stil). Auf diese Dichotomie verweisen ihrer Frequenz nach die Präfixe *za-*, *po-*, *pred-*, *pod-*. Anhand der folgenden Tabelle veranschaulichen wir ihr Vorkommen in den fünf untersuchten Bereichen:

	<i>za-</i>	<i>po-</i>	<i>pre-</i>	<i>pod-</i>
umgangssprachlicher Stil	13,8	11,7	5,7	1
künstlerische Prosa	13,3	14,2	8,5	1,8
Poesie	13	12,5	6,9	1,3
publizistischer Stil	10,8	10	7,8	1,9
wissenschaftlicher Stil	10,5	11,1	7	1,9

Bereits hier zeigt sich deutlich, daß ein Präfix ein klarer und recht sicherer Indikator der Zugehörigkeit eines Textes zu einem stilistischen Bereich sein kann.

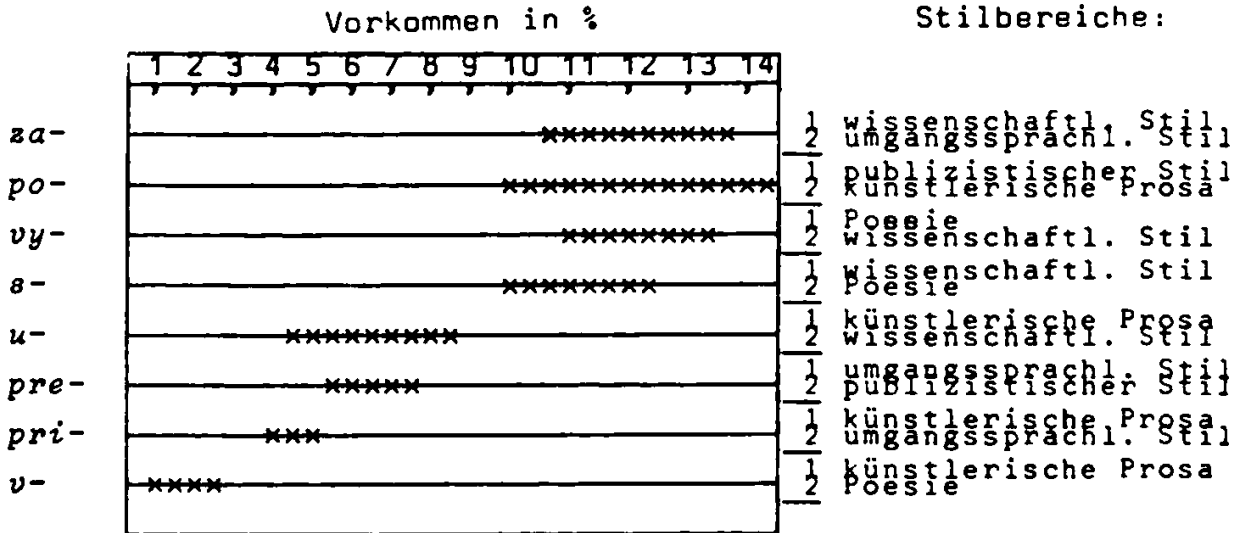
Präfixe, die durch große Häufigkeit oder relative Abwesenheit auffallen, können als Grundlage für die Bestimmung von Stilen dienen. In der 1. Frequenzklasse sind alle Präfixe stilistisch relevant, in der 2. sind es *u-*, *pre-*, *pri-* und in der 3. nur *v-*. Für die stilistische Beschreibung irrelevant sind *o-*, *na-*, *roz-*, *od-*, *do-*, *pred-*, *pod-*, *nad-*, die durch große Häufigkeit (+) oder relative Abwesenheit (-) signifikanten Präfixe:

	<i>za-</i>	<i>po-</i>	<i>vy-</i>	<i>s-</i>	<i>u-</i>	<i>pre-</i>	<i>pri-</i>	<i>v-</i>
umgangssprachlicher Stil	+		+		-	-	+	
künstlerische Prosa	+	+		-	-			
Poesie	+		-	+	-			+
publizistischer Stil	-	-			+	+		
wissenschaftlicher Stil	-	-	+		+			

Wie diese Einteilung zeigt, läßt sich jeder Stil mit Hilfe relevanter Präfixe charakterisieren. Der umgangssprachliche Stil ist durch die Hypertrophie (+) und Atrophie (-) folgender Präfixe gekennzeichnet: *+za-*, *+vy-*, *-u-*, *-pre-*, *+pri-*, die künstlerische Prosa: *+za-*, *+po-*, *-s-*, *-u-*, die Poesie: *+za-*, *-vy-*, *+s-*, *-u-*, *+v-*, der publizistische Stil: *-za-*, *-po-*, *+u-*, *+pre-* und der wissenschaftliche Stil: *-za-*, *-po-*, *+vy-*, *+u-*. Lassen wir die Poesie als eine in gewissem Maße künstliche Sprache einmal außer acht, so manifestiert sich im Vorkommen der Präfixe die Tendenz, daß die (polysemantischen) Präfixe der 1. Frequenzklasse in der 1. Stilklasse (umgangssprachlicher und künstlerischer Stil) auftreten und die der niederen Frequenzklassen für die 2. Stilklasse (publizistischer und wissenschaftlicher Stil) typisch sind. Dagegen ist das häufige Erscheinen unsilbischer Präfixe für die Poesie bezeichnend. Diese Situation ist dadurch bedingt, daß - da in einem Gedicht Wörter mit niedriger Silbenzahl bevorzugt werden - die genannten Präfixe die Silbenzahl nicht erhöhen.

Sehr aufschlußreich ist die Frequenzspanne, die sich zwischen den signifikanten Präfixen beobachten läßt. Wir veranschaulichen sie in der folgenden Tabelle und nennen für jedes Präfix die Antipoden, d.h. die Stilbereiche, in denen das Präfix am

häufigsten (2) bzw. am seltensten (1) vorkommt.



Die Präfixe *po-* (4,2), *u-* (4,1), *za-* (3,3), *vy-* (2,5), *s-* (2,3) und *pre-* (1,9) unterliegen den größten Schwankungen. Unter dem Aspekt der Frequenz gesehen stehen sich jeweils ein Stil der 1. (umgangssprachlicher Stil, künstlerische Prosa, Poesie) und der 2. Klasse (publizistischer und wissenschaftlicher Stil) gegenüber.

Wir geben nun einen vollständigen Überblick über das prozentuale Vorkommen der häufigsten Präfixe in den Stilen der slovakischen Schriftsprache und fassen unsere Darstellung in Thesen.

Präfixe	Durchschnittl. Vorkommen	Umgangssprachl. Stil	künstler. Prosa	Poesie	publiz. Stil	wissenschaftl. Stil
<i>za-</i>	12,2	13,8	13,3	13	10,8	10,5
<i>po-</i>	12	11,7	14,2	12,5	10	11,1
<i>vy-</i>	11,9	12,6	11,3	10,7	12	13,2
<i>s-/z-</i>	10,6	10,8	9,9	12,2	10,5	9,9
<i>o-</i>	8,4	7,9	8,6	8,4	8,3	8,5
<i>na-</i>	7,4	7	7,7	6,7	7,9	7,4
<i>pre-</i>	6,8	5,7	6,5	6,9	7,6	7
<i>u-</i>	6,7	7	4,6	5,6	7,9	8,7
<i>roz-</i>	5,2	4,7	5,9	5,7	4,8	4,6
<i>od-</i>	4,9	4,4	5,1	4,3	5,2	5,1
<i>pri-</i>	4,7	5,4	4,2	4,4	4,8	5
<i>do-</i>	3,8	4,4	4,2	3,2	3,8	3,7
<i>pred-</i>	1,8	1,8	1,2	2,3	2,4	1,6
<i>v-</i>	1,7	1,5	1,2	2,6	1,7	1,4
<i>pod-</i>	1,6	1	1,8	1,3	1,9	1,9
<i>nad-</i>	0,3	0,3	0,2	0,2	0,4	0,4
	100 %	100 %	100 %	100 %	100 %	100 %

Präfixe sind zuverlässige Indikatoren für die stilistische Ausrichtung eines Textes. Wir stützen uns auf das Vorkommen der vier häufig gebrauchten *za-*, *po-*, *vy-*, *s-* und der vier selten verwendeten Präfixe *u-*, *pre-*, *pri-*, *v-*, da die übrigen ihrer Frequenz nach unauffällig sind. Polysemantische Präfixe erscheinen häufig und sind typisch für den umgangssprachlichen und den künstlerischen Stil, während in bedeutungsmäßiger Hinsicht unkomplizierte Präfixe charakteristisch für Fachstile sind. Im wissenschaftlichen und künstlerischen Stil ist die Zahl präfigierter Wörter am höchsten. Poesie und künstlerische Prosa weichen der Präfixfrequenz nach am stärksten, der wissenschaftliche Stil am wenigsten von den parametrischen Informationen ab.

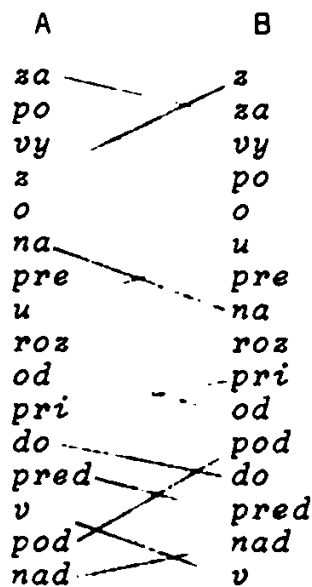
Ein und dasselbe Präfix kann in ein und demselben Text abwechselnd mehrere Funktionen (eine benennende, expressive und grammatische Funktion) erfüllen. Dies ist also in dem Frequenzsignal enthalten, auf dessen Grundlage dann später die quantitative stilistische Beschreibung des Textes entsteht.

Für das slovakische rückläufige Wörterbuch wurden sämtliche präfigierten Wörter gezählt, wobei sich die Größe 140 000 ergab. Wir vergleichen nun die Situation, die in einem Kontext von 1 000 000 und 21 823 Wörtern in der im Wörterbuch erscheinenden Form vorliegt, mit den o.a. 140 000 Einheiten. Wir ordnen die Präfixe nach ihrer Frequenz in 5 Spalten. Spalte A zeigt die im Kontext vorliegende Situation, die Spalten B - E stellen die Verteilung der 140 000 Wörter auf Substantive (B), Adjektive (C), Adverbien (D) und Verben (E) dar.

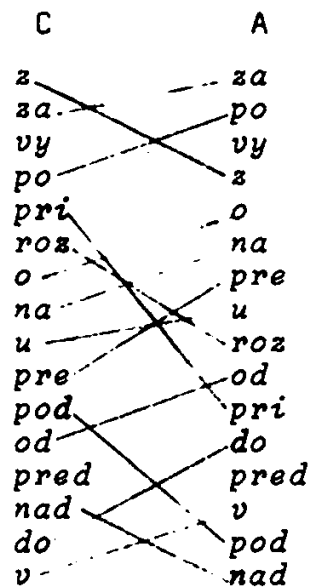
A	B	C	D	E
<i>za</i>	<i>z</i>	<i>z</i>	<i>z</i>	<i>z</i>
<i>po</i>	<i>za</i>	<i>za</i>	<i>na</i>	<i>po</i>
<i>vy</i>	<i>vy</i>	<i>vy</i>	<i>po</i>	<i>za</i>
<i>z</i>	<i>po</i>	<i>po</i>	<i>pri</i>	<i>vy</i>
<i>o</i>	<i>o</i>	<i>pri</i>	<i>za</i>	<i>o</i>
<i>na</i>	<i>u</i>	<i>roz</i>	<i>u</i>	<i>roz</i>
<i>pre</i>	<i>pre</i>	<i>o</i>	<i>o</i>	<i>na</i>
<i>u</i>	<i>na</i>	<i>na</i>	<i>roz</i>	<i>pre</i>
<i>roz</i>	<i>roz</i>	<i>u</i>	<i>pre</i>	<i>od</i>
<i>od</i>	<i>pri</i>	<i>pre</i>	<i>vy</i>	<i>pri</i>
<i>pri</i>	<i>od</i>	<i>pod</i>	<i>od</i>	<i>u</i>
<i>do</i>	<i>pod</i>	<i>od</i>	<i>do</i>	<i>do</i>
<i>pred</i>	<i>do</i>	<i>pred</i>	<i>pod</i>	<i>v</i>

Forts.	A	B	C	D	E
	v	pred	nad	v	pod
	pod	nad	do	pred	pred
	nad	v	v	nad	nad

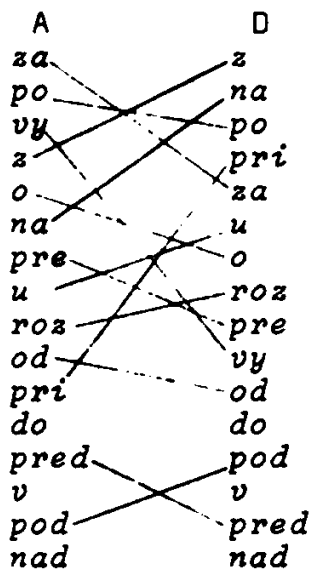
Wir vergleichen nun, wie sich die Anordnung der Präfixe unter A von den anderen unterscheidet:



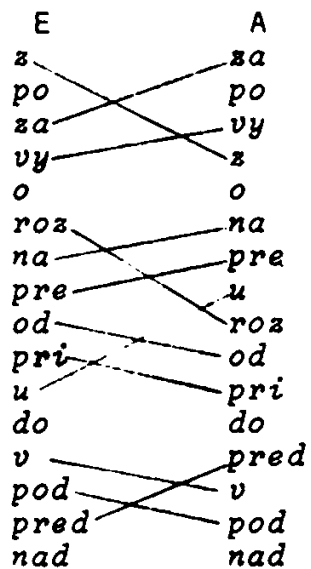
Differenz: 20



36



Differenz: 38



20

Am stärksten unterscheidet sich die Anordnung der Präfixe unter A von der der adjektivischen und adverbialen. Die Differenz zwischen A und C beträgt 36, zwischen A und D 38. Zwischen den Spalten A, B und E liegen nur geringe Unterschiede vor. Wir sehen dies als Beweis für die Verschiedenheit der Zahl von Substantiven und Verben in Kontext und Wörterbuch - verursacht

0046895

durch die asymmetrische Verteilung der Adjektive und Adverbien in den Texten - an. Diese Gesetzmäßigkeit stützt die These, daß auch Präfixe Signale für stilistische Qualitäten eines Textes sein können.

DISTRIBUTION MORPHOLOGISCHER ELEMENTE

Das Wort ist das einfachste und auffälligste stilistische Mittel. Es hebt sich im Text deutlich ab und erscheint dem Betrachter als verhältnismäßig autonom. Auch in semantischer Hinsicht es es oft auf eine Fläche, eine Stelle begrenzt. Die grammatischen Kategorien dagegen sind, obwohl sie sich in der Wortform manifestieren, latente stilistische Mittel, deren man sich - im Gegensatz zu den lexikalischen Bedeutungen - erst auf einer größeren Fläche bewußt wird. Sie sind daher relativ abstrakt und treten im Durchschnitt öfter als ein Wort auf. Obwohl der Häufigkeitsgrad eines sprachlichen Elements im indirekten Verhältnis zur stilistischen Ausdrucksfähigkeit steht, wird eine Vielzahl der grammatischen Kategorien als stilistisch neutral empfunden.

Die grammatischen Kategorien und die Wortform. Es ist nicht Aufgabe der vorliegenden Arbeit, die Beziehungen der grammatischen Kategorien zu den philosophischen zu erläutern. Dennoch wollen wir erwähnen, daß mit ihrer Hilfe einige - wenn auch nur wenige - philosophische Kategorien ausgedrückt werden können und daß diese nicht in allen Sprachen die gleichen sind, obwohl auch eine Übereinstimmung nicht auszuschließen ist.

Als Beispiel nennen wir die Kategorien *Genus* und *Numerus*. Wir wählen das slovakische Substantiv *kniha* mit der deutschen Entsprechung *Buch*, der französischen *libre*, der italienischen und spanischen *libro* und der englischen *book*. Dieses Wort hat im Slovakischen feminines, im Deutschen neutrales und in den romanischen Sprachen maskulines Geschlecht, während es im Englischen in Bezug auf das Genus merkmalllos ist. Das slovakische Wort *raňajki* zeigt Pluralmorphem mit femininem Genus, wohingegen seine deutsche (*Frühstück*) und spanische Übersetzung (*desayuno*) die Merkmale des Neutrum/Singular aufweisen. Die slovakische Entsprechung für *Mund* (mask./sing.) wird mit Substantiven von unterschiedlichem Genus und Numerus benannt: *tie ústa* (neutr./plur.), *tá papuľka* (fem./sing.) und *ten pysk* (mask./sing.), d.h. Wörter gleicher Bedeutung unterliegen in

ein und derselben Sprache oft unterschiedlichen grammatischen Kategorien. Dasselbe läßt sich bei einem Vergleich verschiedener Sprachen beobachten: Slowakisch: *tá tuberkulóza* (fem./sg.), *tie suchoty* (fem./pl.); Russisch: *туберкулёз* (mask./sg.) und Deutsch: *Tuberkulose* (fem./sg.). Ähnliches gilt auch für das Verhalten der Kasus in den einzelnen Sprachen, wohingegen sich bei anderen grammatischen Kategorien eine relative Gleichwertigkeit feststellen läßt.

Die Art, wie diese Kategorien ausgedrückt werden, hängt vom Typ der betreffenden Sprache ab. So z.B. ist im Englischen das Zeichen *s* das einzige differenzierende Merkmal zur Darstellung so verschiedener Kategorien wie des Plurals, Genitivs, der Possessivität verschiedener Pronomina (*its, his, yours*) und Substantiva (*father's, mother's*) sowie des Präsens der 3. Person singularis (*gives, puts, does*). Eine besondere Art, die Possessivität wiederzugeben, existiert im Ungarischen. Die Behandlung der Kasus erfolgt also in den einzelnen Sprachen auf unterschiedliche Weise, z.B. stimmt ihre Anzahl nicht immer überein. Ebenso verhält es sich mit den Tempora. Es gibt aber auch grammatische Kategorien, die sich nur mittelbar (z.B. das verbale Genus) oder nur auf dem Hintergrund eines Kontextes (z.B. psychologisches Subjekt und Prädikat) manifestieren. Hier liegt einer der Gründe dafür, daß ihre quantitative Distribution schwieriger zu erkennen als die der Wörter ist, was wiederum die These vom Signalcharakter der erzielten Informationen bestätigt. Dennoch kann aber die Untersuchung der grammatischen Kategorien eines Textes für die Typologisierung sinnvoll sein.

Die stilistische Signifikanz der Wortartdistribution. Die Zugehörigkeit eines Wortes zu einer bestimmten Wortart ist die Grundlage für seine Einordnung in einen Satz.²⁵ Wenn eine Wortart innerhalb eines Satzes auch mehrere Funktionen ausüben kann, so sind doch - wie die Frequenzangaben zeigen - bestimmte Wortarten für bestimmte Satzteile typisch: z.B. wird das Subjekt meistens durch ein Substantiv, das Attribut durch ein Adjektiv, das Prädikat durch ein Verb und das Objekt durch ein Substantiv repräsentiert. Dies trifft auf die Mehrzahl der untersuchten Fälle zu. Die Zugehörigkeit eines Wortes zu einer

bestimmten Wortart ist also auch ein Indikator für seine Funktion als Satzteil. Damit wird die Wortartdistribution zu einem stilistischen Signal, das auf die Basis der lexikalischen Bedeutung verweist. Diese Grundlage ist zwar nur latent vorhanden, aber sie stellt dennoch eine wichtige stilistische Komponente der Äußerungen dar.²⁶

In einem Text stehen sich Substantive und Verben diametral gegenüber. Als grundlegende Wortarten bestimmen sie den Rhythmus des Textganzen. Die Verteilung der übrigen Wortarten ist natürlich ebenfalls wichtig, wenn diese auch dem diametralen Verhältnis von Substantiv und Verb untergeordnet sind.²⁷ Als Beispiel führen wir die Wortartdistribution einiger Texte an:

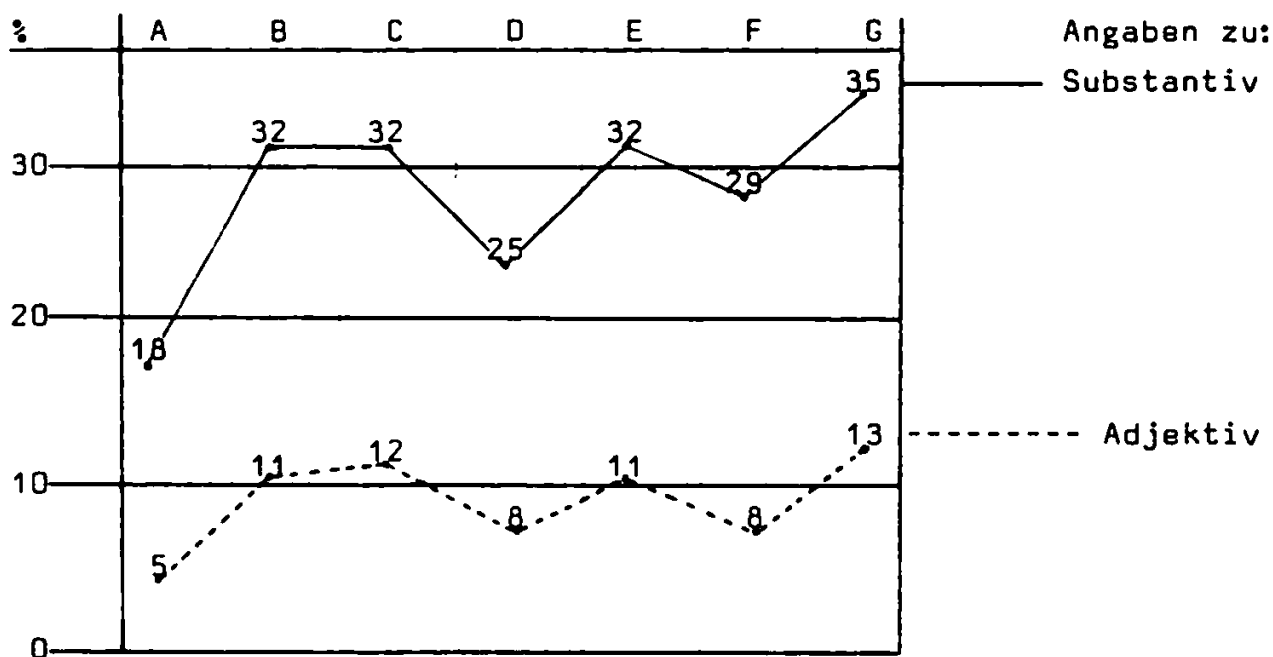
- A - Dialog im Drama *Posledná prekážka* von Š. Králik
 B - Publizistischer Text aus der Zeitung *Smena* (XIX/125)
 C - Gedichte von P. Horov: *Poézia*
 D - Künstlerische Prosa von A. Chudoba: *Kde pijú dúhy*
 E - Wissenschaftliche Prosa von V. Filkorn: *Predheglovská logika*
 F - Kinderprosa aus der Zeitschrift *Ohník* (XVIII/4)
 G - Administrativer Text aus *Písomný styk* von J. Kyndl

	A	B	C	D	E	F	G
Substantive	18	32	32	25	32	29	35
Adjektive	5	11	12	8	11	8	13
Pronomina	17	9	10	13	12	10	7
Numeralia	2	3	2	1	1	2	4
Verben	25	14	16	24	18	22	14
Adverbien	13	5	9	5	8	9	4
Präpositionen	7	10	10	10	8	10	10
Konjunktionen	10	7	4	8	11	7	7
Partikel	3	4	3	2	2	2	2
Interjektionen	1	-	1	-	-	-	-
Abkürzungen	-	4	-	-	-	-	-

Die Angaben auf der horizontalen Achse lassen sich folgendermaßen interpretieren:

a) Im Dialog (A) treten die Substantive prozentual hinter die Verben zurück. In sachbezogenen Texten (B, E, G), in Gedichten (C) und in der Kinderprosa (F) liegt das umgekehrte Verhältnis vor, während es in künstlerischer Prosa (D) ausgeglichen ist.

b) In direkter Beziehung zu den Substantiven stehen die Adjektive mit attributiver Funktion: je höher der Prozentsatz an Substantiven ist, desto größer ist auch die Anzahl der Adjektive. Zwei Kurven sollen dies veranschaulichen:



c) In Konversationen (A) - aber auch in künstlerischer Prosa (D) - kommen sehr häufig Pronomina vor, in publizistischen und administrativen Texten (B, G) dagegen nur selten.

d) Die geringste Zahl der Verben ist in der Poesie (C), d.h. in ungebundener Rede, sowie in publizistischen und administrativen Texten (B,G) festzustellen, wobei sich der administrative Stil (G) außerdem durch einen hohen Prozentsatz an enumerativen Konstruktionen, Zitaten, Namen und Fakten auszeichnet.

e) Die Adverbien stellen eine - vom stilistischen Standpunkt aus betrachtet - disparate und konturenlose Wortart dar, die sowohl dem Nomen (als Erläuterung zu Adjektiven) als auch den Verben dient.

f) Numeralia unterliegen einer besonderen Distribution. In sachbezogenen, besonders publizistischen und administrativen Texten (B,G) treten sie häufig, in künstlerischen Texten (D), in Konversationen (A) und Kinderprosa (F) selten auf, während ihr Vorkommen in wissenschaftlicher Prosa (E) themenabhängig ist.

g) Interjektionen sind nur in Konversationen (A) und Poesie (C) - vereinzelt auch in Kinderprosa (F) - anzutreffen.

h) Besondere Bedeutung für den Stil einer Äußerung hat die Distribution der funktionellen Wortarten (Präpositionen, Konjunktionen, Partikel), da sie an den Knotenpunkten der syntak-

tischen Konstruktionen stehen, wo sich autosemantische Wörter oder syntaktische Einheiten verbinden.

Nur wenigen Schwankungen unterliegt die Verteilung der Präpositionen, die in durch die Atrophie von Nomen und somit auch deklinierbarer Einheiten gekennzeichneten Konversationen (A) am seltensten sind.

Die wichtigste funktionelle Wortart stellen die Konjunktionen dar. Sie verweisen auf die Syndetizität von Sätzen. Ihren niedrigsten Prozentsatz finden wir in Gedichten (C), für die eine gewisse syntaktische Abstraktheit typisch ist, den größten dagegen in Konversationen (A) und wissenschaftlicher Prosa (E), die sich jedoch im Hinblick auf die ko- bzw. subordinierenden Konjunktionen unterscheiden: erstere überwiegen in Konversationen (A), die zweiten in wissenschaftlicher Prosa (E). Eine hohe Zahl von Konjunktionen macht einen Text eindeutig und verhindert das Entstehen nebensächlicher Assoziationen. In dem wissenschaftlichen Text (E) beträgt das Verhältnis der autosemantischen Wörter zu den Konjunktionen 7,4 und in der Gedichtsammlung (C) bis zu 22,1, d.h. auf eine Konjunktion kommen 22,1 autosemantische Wörter. In künstlerischen Texten sind die Verbindungen relativ locker gefügt und ermöglichen eine freiere Auslegung des Gesagten, reichhaltige Assoziationen und das Ausdeuten der Metaphern. Die die Syndetizität eines Textes signalisierende Dispersion der Konjunktionen läßt sich durch die Division der Zahl der autosemantischen Wörter (Nauts) durch die der Konjunktionen (Nconj) ermitteln:

$$D = \frac{Nauts}{Nconj}$$

Subordinierende Konjunktionen signalisieren die inhaltliche Geschlossenheit, die Dichte eines Textes: je höher ihre Zahl im Verhältnis zu den autosemantischen Wörtern ist, desto dichter und inhaltlich geschlossener ist der Text, und umgekehrt. Der aus den Angaben über ihre Distribution berechnete Index dient weniger als Signal für den Stil als für die stilistische Methode. Er ist z.B. in Zeitungen, die - bei geringer inhaltlicher Dichte - vorwiegend die informative stilistische Methode anwenden, relativ hoch. Die Syndetizität der Äußerung, signalisiert

durch die Dispersion der Konjunktionen allgemein, und die inhaltliche Geschlossenheit, angedeutet durch die Verteilung der subordinierenden Konjunktionen allein, sind voneinander unabhängige Phänomene.

Die Zusammenhänge zwischen dem Vorkommen einiger Wortarten führen - entsprechend ihrer Funktionen im Aufbau des Textes - zu bestimmten Gruppierungen, von denen die folgenden den Textcharakter am deutlichsten signalisieren und daher für die Typologisierung besonders wichtig sind:

- a) Nomina (Substantive und Adjektive)
- b) Konnektoren (Hilfswörter und einige Pronomina)
- c) Verben

Das Übergewicht der Nomina bewirkt Begrifflichkeit, statische Ausrichtung und Verdichtung des Textes, während das der Konnektoren den Eindruck von Erläuterung und innerer Kompliziertheit der Textstruktur hervorruft und das der Verben Aktion, Dynamik und Übersichtlichkeit signalisiert.

Die Angaben auf der vertikalen Achse der Tabelle (S. 74) lassen sich folgendermaßen deuten:

a) Der Dialog (A) ist - ähnlich wie die künstlerische Prosa (D) - durch einen hohen Prozentsatz an Verben, ferner durch häufiges Auftreten der Personalpronomina und geringes Erscheinen von adjektivischen Attributen und Präpositionen gekennzeichnet. Die übrigen Wortarten sind in stilistischer Hinsicht weniger markant.

b) Ein Merkmal des publizistischen Stils (B) ist die hohe Zahl der Substantive. Im Hinblick darauf kommt er dem administrativen Stil (G), dem Gedicht (C) und der wissenschaftlichen Prosa (E) nahe. Verben sind hier nur schwach vertreten. Der publizistische und administrative Stil sind arm an Pronomina.

c) Gedichte (C) als in hohem Maße abstrakte Texte fallen durch die Atrophie von Konjunktionen auf. Substantive und Adjektive dagegen sind hier relativ häufig vertreten.

d) Die künstlerische Prosa (D) ähnelt in der Zahl der Verben dem Dialog (A), in der der Substantive, Adjektive und Verben hingegen der Kinderprosa (F). Sie unterscheidet sich wesentlich vom administrativen Stil (G).

e) Die wissenschaftliche Prosa (E) ist - was das Vorkommen der Substantive und Adjektive betrifft - mit der publizistischen (B) vergleichbar. Ähnlich verhält es sich auch bei den Verben.

f) Kinderprosa (F) ist wenig abstrakt, leicht und dynamisch, die Zahl der Substantive daher gering, die der Verben und Präpositionen jedoch hoch.

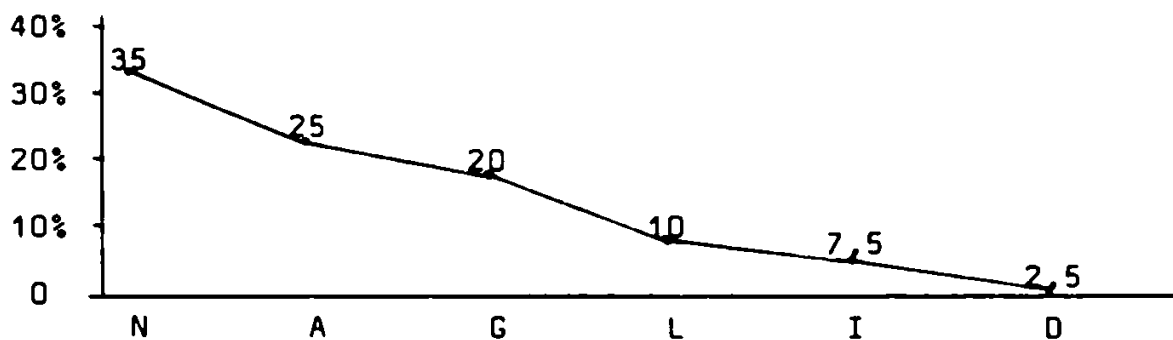
g) Der administrative Stil (G) ist ein typischer Buchstil mit einem hohen Prozentsatz an Personalpronomina, Adverbien und Partikeln.

Die Wortartdistribution verleiht dem Text einen bestimmten Rhythmus, dessen Impuls von der Textdynamik, die wiederum auf der Zahl der von den einzelnen Wörtern ausgedrückten Kategorien basiert, ausgeht. So z.B. ist das Substantiv Träger der Kategorien *Numerus* und *Kaesus*. Das Genus innerhalb der einzelnen Sprachen ist - wie bereits gezeigt wurde - allein eine Frage der Form. Das Verb fungiert als Träger der Kategorien *Tempus*, *Modus* und *Person*, in einigen Sprachen auch der des Aspekts, Genus u.ä. Es sei das Substantiv (S) Träger von 2, das Verb (V) von 4 Kategorien. Sehen wir diese als Valenzen an und fügen wir zu S und V je eine lexikalische Valenz hinzu, so hat das Substantiv drei, das Verb fünf Valenzen. Hätte ein Satz die Form SSVSS, so ergäbe die Summe der Valenzen $3 + 3 + 5 + 3 + 3$. Unterläge ein anderer Satz - bei identischer Anzahl der Satzteile - der Distribution SVVSV, so wäre die Summe $3 + 5 + 5 + 3 + 5$. Der zweite Satz ist dynamischer als der erste, denn er ist - bei gleicher Fläche - um vier Valenzen reicher. Der unterschiedliche Wechsel von Substantiv und Verb verweist also auf einen dynamischen Rhythmus. Dieses Problem ist in Wirklichkeit jedoch komplizierter, als es hier dargestellt wird, da sich die Kategorien in Bezug auf Dynamik und Schwierigkeit ihrer Perzeption sehr voneinander unterscheiden und wesentlich von ihrer allgemeinen Frequenz im Sprachsystem abhängen.

Wir beschäftigten uns bisher nur mit Substantiv und Verb. Die übrigen Wortarten, deren Kategorien von ihnen abhängig sind, komplizieren die Situation.

Rhythmus und Charakter eines Textes werden in starkem Maße

vom Vorkommen der Kasus beeinflusst, wobei Nominativ und Akkusativ (Kasus mit hoher Frequenz) keine besonderen Ansprüche an die Perzeption stellen und für die übrigen das Gegenteil gilt. Anhand slovakischer Texte mit einem Umfang von 50 000 Wörtern erstellen wir eine parametrische Kurve, die das Vorkommen der Kasus im Slovakischen veranschaulichen soll:



Es zeigt sich, daß die Kasus folgendermaßen verteilt sind: Nominativ (N) 35%, Akkusativ (A) 25%, Genitiv (G) 20%, Lokativ (L) 10%, Instrumental (I) 7,5 % und Dativ (D) 2,5 %.

Im Deutschen liegt eine ähnliche Situation vor:



Der geringe Umfang der untersuchten Texte (10 000 Wörter) läßt genauere Angaben nicht zu. Zuverlässig sind jedoch die Distribution der Kasus und die auf Annäherungswerten beruhende graphische Gestalt des Schaubildes.

Je häufiger im allgemeinen ein Kasus auftritt, desto schwächer als stilistisches Mittel ist er. Je weniger er im Text erscheint, desto auffallender ist er. So z.B. sind in den slavischen Sprachen die Texte am anspruchsvollsten, in denen zwischen der Zahl der Nominative und Akkusative einerseits und der übrigen Kasus andererseits nur ein geringer Unterschied besteht. Im Deutschen nehmen Nominativ und Akkusativ eine zentrale, Genitiv und Dativ eine periphere Stellung ein. Im publizistischen und wissenschaftlichen Stil hat der Genitiv, in Kon-

versationen haben Nominativ und Akkusativ hohe Frequenz. In Kinderprosa liegt eine ähnliche Verteilung wie in Konversationen und im künstlerischen Stil vor.

Die Untersuchung der Kasusfrequenz und ein Vergleich mit den parametrischen Angaben können Aufschluß über den Grad der Spontaneität einer Äußerung geben. In spontanen Äußerungen wird den Elementen im allgemeinen keine Aufmerksamkeit zuteil, da hier nur bekannte und automatisierte Einheiten (Nominativ und Akkusativ) verwendet werden. Die relativ stärkste Symmetrie in der Auswahl der Kasus und der sprachlichen Elemente überhaupt läßt sich in schriftlichen Monologen beobachten, bei deren Abfassung viel Zeit zur Verfügung stand.

Die stilistisch aufschlußreichsten Verbkategorien sind die der Zeit und der Person. Es gibt Texte, in denen diese stereotyp und ohne Wechsel gebraucht werden, aber auch solche, auf die das Gegenteil zutrifft. Für die wissenschaftliche Prosa z.B. ist das Fehlen der 1. und 2. Person singularis sowie der 2. Person pluralis ebenso typisch wie das alleinige Vorkommen des Präsens. Eine Änderung dieser Verhältnisse käme einem besonderen stilistischen Mittel gleich. In künstlerischer Prosa ist das Imperfekt merkmalllos, das Präsens dagegen merkmalshaft. Der Poesie fehlt die Kategorie der Zeit, da hier meistens allein das Präsens gebraucht wird. Im epischen Gedicht liegt eine ähnliche Situation wie in der künstlerischen Prosa vor.

Grammatische Kategorien, deren Vorkommen weniger als 10% ausmacht, sind als quantitative Indikatoren für die Typologisierung unbrauchbar. Für die Analyse und Interpretation hingegen können sie viel besagen.

Wir haben in diesem Kapitel gezeigt, daß morphologische Elemente - obwohl sie in den Texten kaum auffallen und oft nur indirekt festgestellt werden können - dennoch wesentliche, den Text formende Elemente sein können. Sie üben ihre Funktion fast automatisch aus und sind daher für die Typologisierung besonders wichtig.

STILISTISCHE RELEVANZ DER ANORDNUNG SYNTAKTISCHER ELEMENTE

Ein deutliches und objektiv meßbares, den Texttyp signalisierendes Element ist die quantitative Satzproportion, die im stilbildenden Prozeß unbewußt angewandt wird. Die Mehrzahl der Forscher, die einen Text mit Hilfe mathematischer Methoden diagnostizieren, versteht die Satzlänge als wichtiges Signal für den Stil.²⁸

Es soll gleich zu Anfang gesagt werden, daß die Information über die durchschnittliche Satzlänge allein nicht immer ausreicht. In belletristischen und künstlerischen Texten kann sie sich oft als falscher Indikator erweisen, wohingegen sie in Bezug auf nicht-künstlerische Texte zuverlässiger ist. Wir geben ein Beispiel. Der Autor eines Textes verwendet in 9 Sätzen je 3 und im 10. Satz 63 Wörter. Die durchschnittliche Satzlänge beträgt also 9. Obwohl diese Angabe genau ist, besagt sie dennoch nichts über den stilistischen Aufbau des Textes, denn sie trafe ebenfalls auf jede andere Äußerung mit 10 Sätzen von insgesamt 90 Wörtern zu. Hier wären die Intervalle zwischen den Sätzen aufschlußreicher. Wir führen deshalb ein weiteres Beispiel an. Der letzte Satz des Prosatextes *Možno si postavim bungalow* von V. Šikula umfaßt allein 1 400 Wörter. Da es in der modernen Prosa keine Ausnahme ist, daß eine ganze Erzählung oder Novelle in syntaktischer Hinsicht nicht gegliedert wird, kann also - in Bezug auf den künstlerischen Stil - die bloße Angabe der Satzlänge den Interpreten irreführen. In nicht-künstlerischen Äußerungen (z.B. im erörternden theoretischen Stil) ist eine ähnliche Situation möglich: durch längere Aufzählungen wird der Wert der durchschnittlichen Satzlänge mehr oder weniger verfälscht. Die Wahrscheinlichkeit eines Irrtums wird hier jedoch verringert, wenn sich die Angaben an längeren Texten orientieren. Die Gefahr einer Verfälschung ist also vom Umfang des untersuchten Textes abhängig.

Zahlreiche statistische Untersuchungen ergaben, daß Informationen über die Satzlänge durchaus objektive Indikatoren sein können. Dabei sind jedoch verschiedene Umstände zu bedenken.

Wir untersuchen die Beziehung der Satzgestalt zu den Eigenschaften der Textstruktur, die auf objektiven und subjektiven stilbildenden Faktoren beruhen, und sehen uns zunächst die mathematisch-statistische Satzproportion in dem essayistischen Buch *Od včerajška k dnešku* von Matuška an. Es enthält auf 223 Seiten insgesamt 1 326 Sätze. Die anhand der Formel

$$\bar{x} = \frac{\sum x_i \cdot n_i}{\sum n_i}$$

errechnete durchschnittliche Satzlänge beträgt 28,5, wobei 9 Sätze aus nur 1 Wort bestehen, der längste Satz 163 Wörter umfaßt und Sätze (41) mit 18 Wörtern dominieren. 98% weisen weniger, 2% mehr als 85 Wörter auf.

Für die stilistische Charakteristik ist die durchschnittliche Abweichung der Satzlengthen voneinander wichtig. Diese steigt mit der Veränderlichkeit der Äußerung. Der extreme Variabilitätswert, der auf der Differenz zwischen der maximalen und minimalen Satzlengthe beruht, beträgt bei Matuška 162 ($x_{\max} - x_{\min} = 163 - 1$) und ist ein für das Slovakische exklusiver Wert. Die durchschnittliche, anhand der Formel

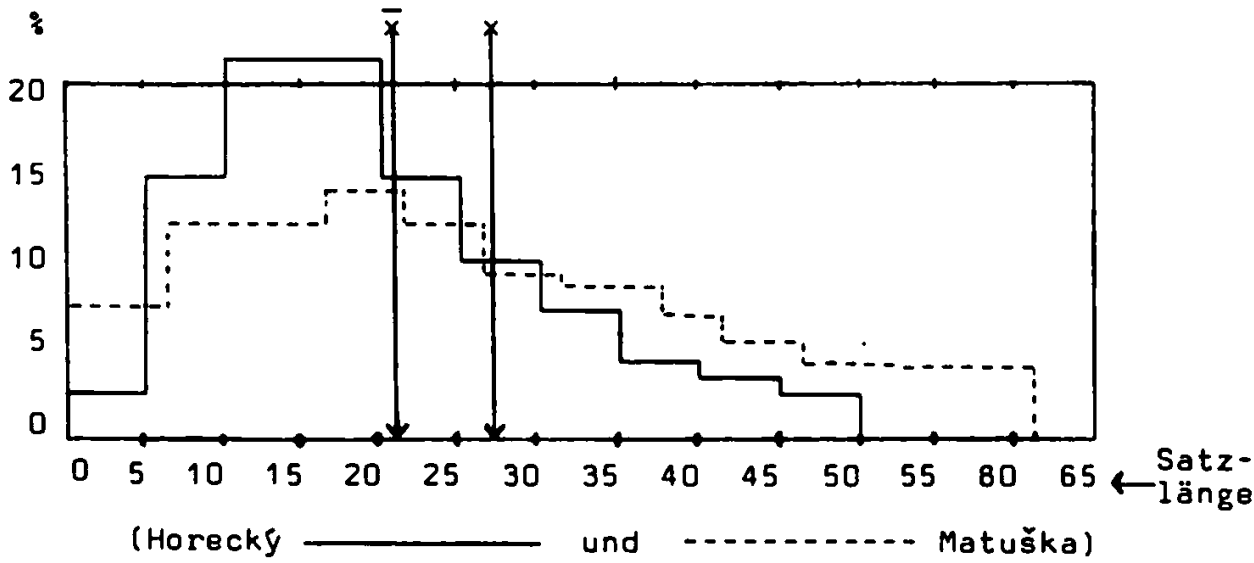
$$\bar{d} = \frac{\sum |x_i - \bar{x}| \cdot n_i}{\sum n_i}$$

errechnete Abweichung dagegen erreicht die Größe 85,20.

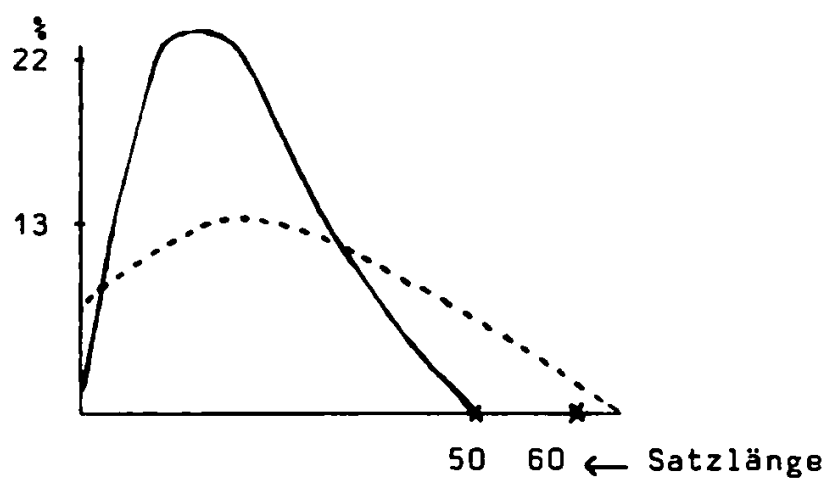
Zum Vergleich ziehen wir das Buch *Morfematická štruktúra slovenčiny* von J. Horecký heran, das 1 036 Sätze mit einer durchschnittlichen Länge von 20,1 enthält. Hier besteht der kleinste Satz mit der Frequenz 1 aus 2, der längste aus 69 Wörtern. Es dominieren Sätze (56) mit 18 Wörtern, während solche mit mehr als 45 Wörtern nur sporadisch vorkommen. Der extreme Variabilitätswert (V_r) beträgt also 68, die durchschnittliche Abweichung (\bar{d}) 16,82. Diese Satzlengthenwerte sind im Vergleich zu den bei Matuška berechneten konzentrierter und signalisieren eine neutrale Äußerung. Das Vorkommen einiger längerer Sätze ist durch Enumerationen bedingt. Der Schwankungsgrad steht so in indirektem Verhältnis zur Begrifflichkeit der Äußerung: je höher \bar{d} , desto weniger begrifflich ist der Text, und umgekehrt.

Die Angaben über die durchschnittliche Satzlengthe bei Matuška und Horecký lassen sich auch graphisch darstellen. Die Ver-

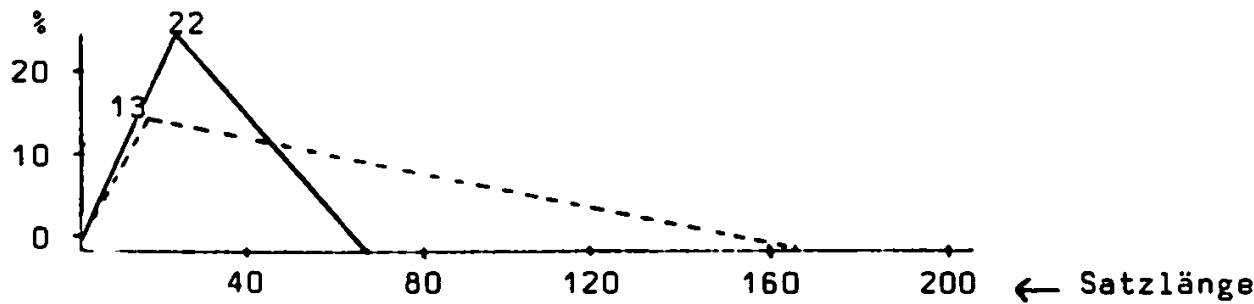
schiebungen im Bereich des regelmäßigen Erscheinens jeder Satz-
länge, ausgedrückt in %, sehen folgendermaßen aus:



Diese Tendenz könnte vereinfacht auch so dargestellt werden:



Um die recht beträchtlichen Unterschiede der durchschnittli-
chen Abweichung bei beiden ($\bar{\sigma}_M$ 85,20; $\bar{\sigma}_H$ 16,82) erkennbar zu
machen, müssen alle vorkommenden Satzlengthen berücksichtigt wer-
den. So zeigt sich ihre große Spannweite bei Matuška und die
relativ geringe bei Horecký:



Die durchschnittliche Abweichung ($\bar{\sigma}$) steht in Zusammenhang

damit, daß sich die Satzlänge bei Matuška zwischen 1 und 183, bei Horecký zwischen 1 und 69 bewegt, d.h. daß die variierte Spannweite beim ersten unverhältnismäßig höher als beim zweiten ist, was wiederum die Größe von \bar{d} beeinflusst.

Es wäre falsch, die Satzlänge und die durchschnittliche Abweichung als einzige Signale stilistischer Merkmalhaftigkeit zu verstehen: zu einer genauen Untersuchung ist es wichtig, noch weitere Informationen zu berücksichtigen, denn erst ihr Zusammenspiel ermöglicht einen zuverlässigen Indikator.

Als weitere Indikatoren für die stilistische Charakteristik sehen wir die Streuung (*Dispersion*) und die auf ihrer Grundlage errechnete *maßgebende* (oder richtungweisende) Abweichung an.

Den Wert der Dispersion erhalten wir mit Hilfe der folgenden Formel:

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2 \cdot n_i}{\sum n_i}$$

Er beträgt bei Matuška 4 887,63, bei Horecký dagegen nur 438,2. Die Dispersion bezieht sich auf den Formenreichtum, im vorliegenden Fall auf den Reichtum an Längen, die auf einer bestimmten Fläche vorkommen. Ihre Höhe sagt die Ausdehnung des Bereiches mit regelmäßigem Erscheinen der Längen voraus: bei Matuška 85, bei Horecký 45. Eine leichter feststellbare Information hinsichtlich des Formenreichtums ist jedoch die Variation der Konstruktionen, die bei Matuška 101, bei Horecký 56 beträgt.

Die maßgebende Abweichung läßt sich aus der Dispersion als positive Wurzel einer zweiten Wurzel errechnen:

$$\sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2 \cdot n_i}{\sum n_i}}$$

Sie bezieht so die vielfältigen Möglichkeiten der Satzkonstruktionen in größerem Maße mit ein als die durchschnittliche Abweichung, die sich allein aus der Differenz von $x_i - \bar{x}$ ergibt. Überspitzt könnte man sagen, daß die maßgebende Abweichung den Reichtum an Satzkonstruktionen ähnlich avisiert, wie es in Bezug auf den Wortschatz von der Information über die Dispersion der Elemente der *langue* geschieht (nach P. Guiraud):

$$R = \frac{R \sqrt{L}}{\sqrt{L}}$$

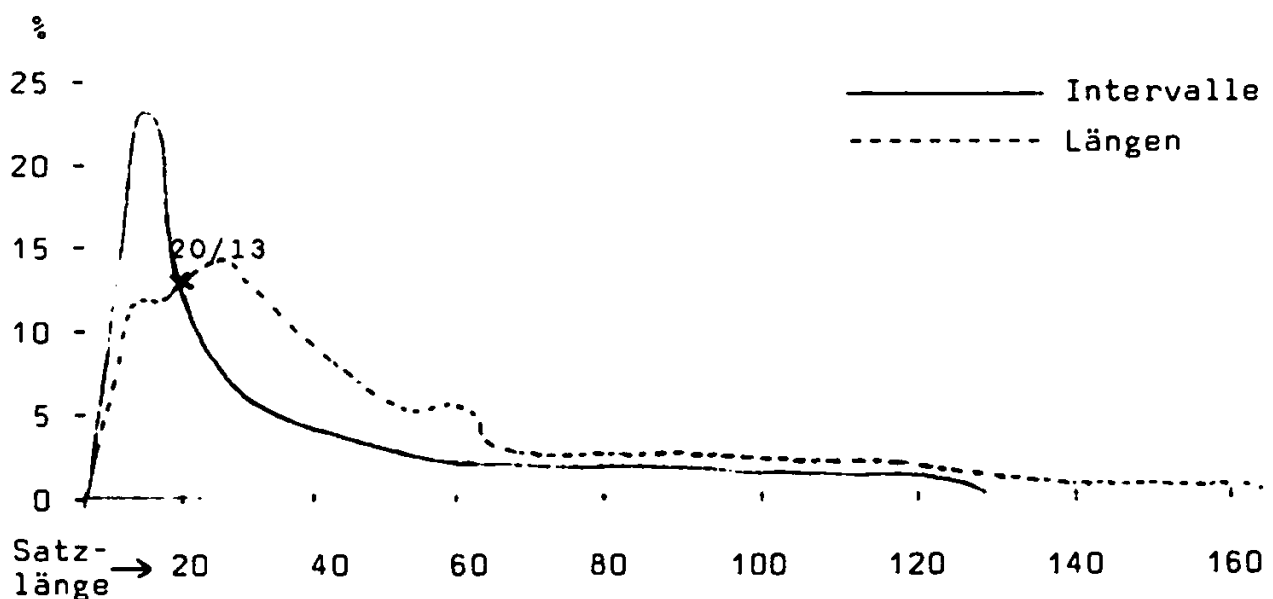
Je größer σ ist, desto vielfältiger sind die syntaktischen Konstruktionen eines Textes.

Statistische Informationen können auch mit Hilfe des variierenden Koeffizienten, der ein relatives Maß der Variationen ist und in % angegeben wird, verglichen werden. Der festgestellte Prozentsatz sagt den Grad von Dynamik und Expressivität des Kontextes, bedingt durch variierende Satztlängen, voraus. Es handelt sich dabei um eine potentielle Expressivität, deren Feststellung die Existenz eines weiteren Indexes voraussetzt. Der variierende Koeffizient (V_k), der anhand der Formel

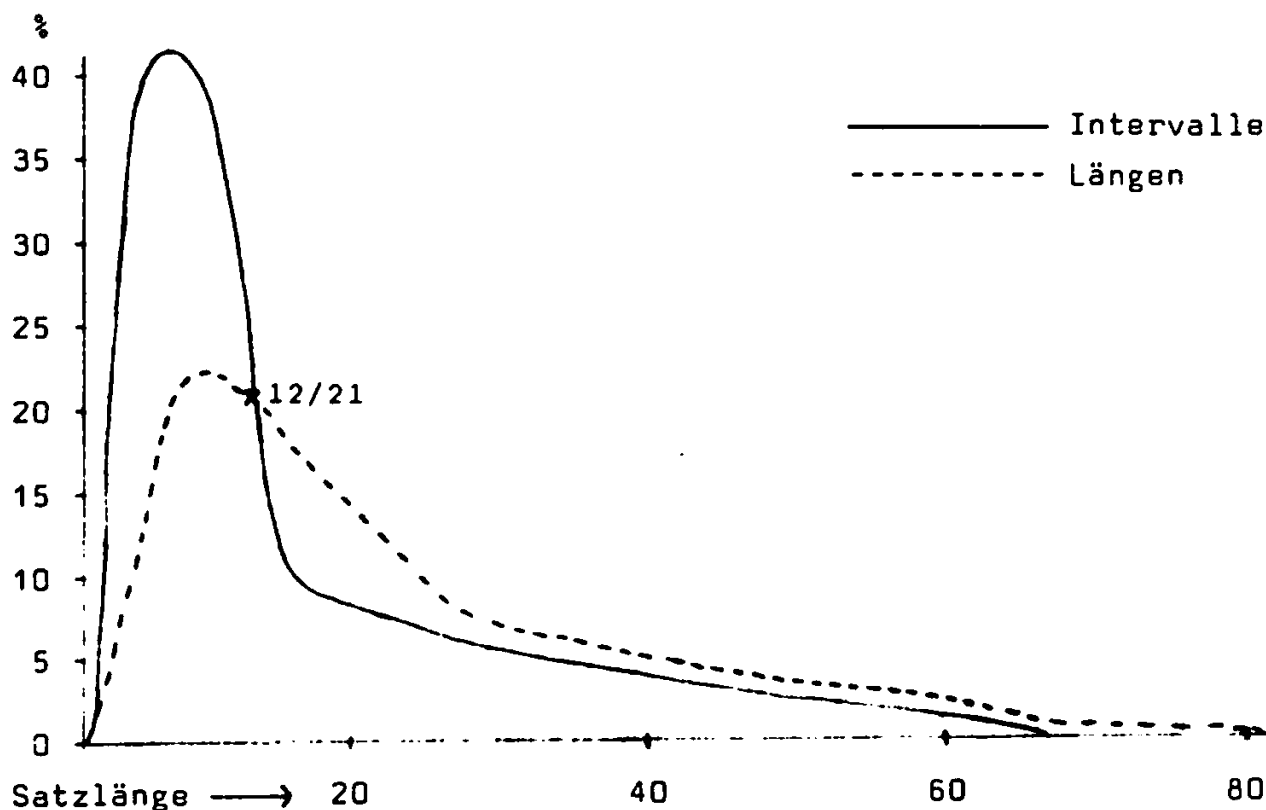
$$V_k = \frac{\sigma}{\bar{x}} \cdot 100$$

errechnet wird, beträgt bei Matuška 245,26, bei Horecký hingegen nur 126.

Ein Ausdruck der Expressivität sind auch die Intervalle, die sich im individuellen Stil zwischen den Sätzen bemerkbar machen. Je länger diese Intervalle sind, desto expressiver ist der Text. In einem sachbezogenen Text ist ihre Länge verhältnismäßig ausgeglichen, da es hier keine abrupten Änderungen gibt. Verglichen mit den Satztlängen sind die Intervalle konzentrierter und insgesamt auf eine kleinere Fläche verteilt. Das folgende Schaubild zeigt ihr prozentuales Erscheinen bei Matuška.

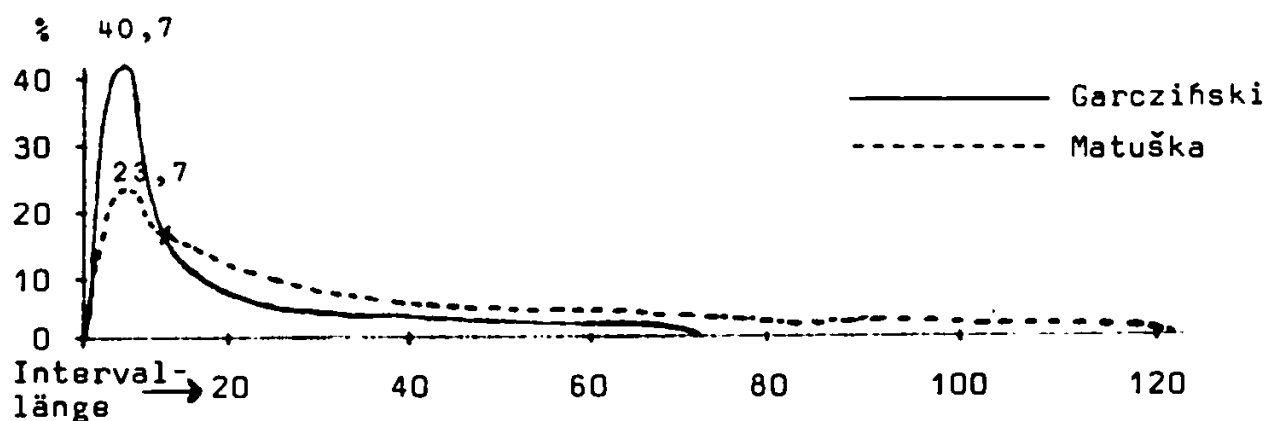


Wir untersuchen nun Satzlänge und Intervalle in dem populären polnischen Buch *Umenie zapamätat' si* von Š. Garcziński. Der Stil ist sachlicher und begrifflicher als bei Matuška, die Sätze sind kürzer und die Intervalle im Verhältnis noch kleiner. Ein Schaubild soll dieses deutlich machen:



Ein Vergleich der Längen und Intervalle zeigt, daß beide Texte in der Tendenz übereinstimmen und sich hauptsächlich in den Schnittpunkten der Kurven unterscheiden. Diese liegen in Matuškas expressiver Äußerung bei 20/13, in Garczińskis populärem Text bei 12/21.

Mit Hilfe der folgenden Darstellung wollen wir die Expressivität beider Texte vergleichen:



Die Intervalllänge bei Matuška ist vielfältiger, bei Garcziński dagegen schematischer. Matuškas Text ist stilisierter - das Subjekt dringt in den Text ein -, während Garczińskis weniger Variationen enthält, objektiver im Ausdruck und nur auf den Inhalt, nicht auf die Form ausgerichtet ist.

Was die Intervalle betrifft, so ist nicht ihre absolute Länge, sondern ihre Variation von Interesse. Diese ist umso höher, je niedriger der Prozentsatz der häufigsten Intervalle, d.h. ihre Konzentration ist.

Konkrete Beispiele sollen nun die Wirkung von \bar{x} , \bar{d} , σ^2 , σ und V_k als Indikatoren für die stilistische Textcharakterisierung zeigen.

Zunächst untersuchen wir die folgenden Texte von Karvaš: das Drama *Zmrŕtvychvstanie deduška Kolomana*, die dem Umfang nach fast identische Prosa *Toto pokolenie* (hier werden Autoren- und Personenrede getrennt behandelt), die theoretische Studie *O dramaturgiokom kontrapunkte* (in: *Slovenské pohl'ady* 6, 1965) und den volkstümlichen Artikel *Pre koho pišeme* (in: *Kultúrny život* 4, 1963).

Obwohl in dem Drama Sätze mit einem Wort (f 486 = 15% von 3 292 Sätzen) dominieren, beträgt die durchschnittliche Satzlänge dennoch 5,1. Diese Angabe ist zuverlässig, da V_r 35, d.h. relativ niedrig ist und die Variation nur 27 beträgt.

Wir geben im folgenden eine Übersicht über die bei Karvaš festgestellten Werte:

	Künstlerische Prosa		Erörterung	
	Drama	Personenreden	Autorenrede	populäre theoretische
\bar{x}	005,10	003,20	011,00	0 046,40
\bar{d}	017,10	003,36	017,46	0 028,73
σ^2	173,55	054,29	485,17	1 137,90
σ	013,10	007,30	022,00	0 033,70
V_k	256,86	228,10	100,00	0 072,63
Variation	027,00	015,00	048,00	0 086,00
V_r	035,00	017,00	063,00	0 200,00

Aus den Angaben wird der diametrale Unterschied zwischen Autoren- und Personenrede in ein und demselben Werk sichtbar.

Die Sätze der Personen sind kürzer als die des Autors, Abweichung und Variabilität in der Autorenrede stärker. Die Reden der Personen sind stereotyper konstruiert. Auch die Dispersion und die maßgebende Abweichung beweisen eine größere Vielfalt der Konstruktionen in der Autorenrede, die sich außerdem durch den höheren Wert des variierenden Koeffizienten auszeichnet. Aus diesen Gründen müssen Autoren- und Personenreden getrennt untersucht werden.

Für die Novelle *Kde pijú dúhy* von A. Chudoba lassen sich die folgenden Angaben machen: \bar{x} = 11,8 (Autorenrede) und 5,8 (Personenrede), σ^2 = 312,9 (Autorenrede) und 157,94 (Personenrede), σ = 17,7 (Autorenrede) und 12,5 (Personenrede), V_k = 150 (Autorenrede) und 215,5 (Personenrede).

Die Dialoge im Drama und in der Prosa unterscheiden sich dadurch, daß die dramatischen Reden auch Erzählungen des Autors enthalten, während die Personenreden der Prosa häufiger die Funktion einer Detailvermittlung erfüllen. Dabei ist es natürlich, daß die Repliken beider Typen verschieden sind. Bemerkenswert ist z.B. die Differenz im Hinblick auf \bar{x} und σ^2 .

Ein Unterschied besteht ebenfalls zwischen dem volkstümlichen und dem theoretischen Stil, deren variierender Koeffizient jedoch fast gleich ist (der Grad von Dynamik und potentieller Expressivität stimmen überein). Die variierenden Koeffizienten der Personen- und Autorenreden weichen stark von denen der nicht-künstlerischen Stile ab.

Wir untersuchen die Proportionen in zwei stilistisch und thematisch verwandten Texten: E. Pauliny: *Začiatky kultúrneho jazyka slovenskej národnosti* (in: *Jazykovedné štúdie VI*, S. 5-39) und J. Ružička: *Základné sporné otázky slovenskej skladby* (in: *Jazykovedné štúdie IV*, S. 7-34). Die durchschnittliche Satzlänge ist identisch (18,4) und kommt dem Wert bei Horecký (20,1) nahe. Bemerkenswert sind jedoch die Unterschiede in der Variation der Spannweite (Pauliny 89, Ružička 58) und in der durchschnittlichen Abweichung (Pauliny 19,45, Ružička 13,8). Hier spielen die zahlreichen Enumerationen, die Pauliny verwendet, eine erhebliche Rolle. Die übrigen Unterschiede fallen nicht so sehr ins Gewicht und können bei einer gewissen Annäherung

außer Betracht bleiben: $\sigma^2 = 629,23$, $\sigma = 25$, $V_k = 135,9$ (Pauliny); $\sigma^2 = 686,66$, $\sigma = 26,2$, $V_k = 141$ (Ružička).

Besonders aufschlußreich in Bezug auf die Satzlänge ist ein Vergleich von Prosatexten und Poesie. Wir wählen zu diesem Zweck Chudobas Prosa und seine Gedichtsammlung *Letokruh erdca*. Die deutlichsten Unterschiede betreffen die extremen Längen (weniger als 5 und mehr als 15 Wörter), während die mittleren Werte übereinstimmen. In der Prosa ist die Konzentration größer, die Intervalle und Extremwerte kleiner, da die an den Satz gestellten Ansprüche geringer sind. Der Satz der Poesie ist im wahrsten Sinne des Wortes künstlich, besonders in der gebundenen Rede, wo die Form fast unabhängig von der Syntax entsteht.

Sätze aus 4 bis 10 Wörtern kommen in allen Stilen, Formen und bei allen Autoren vor, mit mehr als 10 oder weniger als 4 dagegen sind in ihrer stilistischen Anwendungsmöglichkeit beschränkt und können nicht überall erscheinen.

Für die stilistische Typologisierung ist die Beachtung der Spannweite wichtig, die den Bereich des sporadischen Vorkommens von Sätzen einer bestimmten Länge anzeigt. Dieser beginnt z.B. im Konversationsstil bei 10 Wörtern und hat eine äußerst geringe Spannweite, in der Erörterung hingegen ist er breiter.

Wir zeigen nun anhand konkreter Beispiele das sporadische Vorkommen von Satztypen in verschiedenen Texten:

Quelle	Sporadisches Vorkommen der Satz­längen	Längen­spannung	
Karvaš: <i>Toto pokolenie</i> (Personenrede)	12 - 16	6	} Konver­ sation
Chudoba: <i>Kde pijú dúhy</i> (Personenrede)	25 - 32	7	
Karvaš: <i>Toto pokolenie</i> (erzählte Rede)	14 - 22	8	
Hykisch: <i>Nad'a</i> (Personenrede)	27 - 37	10	
Ďuríčková: <i>Danka a Janka</i>	15 - 27	12	
Karvaš: <i>Zmŕtvychvstanie</i>	19 - 36	17	
Švantner: <i>3 poviedky</i>	20 - 38	18	
Chudoba: <i>Kde pijú dúhy</i> (Autorenrede)	38 - 58	20	
Ďuríčková: <i>Danka a Janka</i> (erzählte Rede)	17 - 38	21	
Horecký: <i>Morfematická štruk­ túra slovenčiny</i>	46 - 69	23	

Forts.	Quelle	Sporadisches Vorkommen der Satzlängen	Längenspannung
	Karvaš: <i>Toto pokolenie</i> (Autorenrede)	36 - 64	28
	Garcziński: <i>Umenie zapamätat' si</i>	48 - 78	30
	Hykisch: <i>Nad'a</i> (Autorenrede)	68 - 123	55
	Chudoba: <i>Letokruh srdca</i>	23 - 95	72
	Matuška: <i>Od všerajška k dnešku</i>	65 - 163	98
	Švantner: <i>3 poviedky</i> (Autorenrede)	38 - 138	100

Auf den ersten sieben Plätzen stehen die sehr konzentrierten Texte des Konversationsstils, in denen die Längenspannung, d.h. die Differenz zwischen den Extremen, gering ist. Es folgen die Erörterungen von Horecký, Garcziński und Matuška, zwischen denen die künstlerischen Texte von Chudoba, Ďuríčková, Karvaš, Hykisch und Švantner verteilt sind. Der künstlerische Stil steht also auch hier außerhalb des Systems der übrigen Stile.

Wir sagten bereits, daß ein Replikensatz im Drama länger als in der Prosa ist. Wie eine statistische Untersuchung ergab, müssen mindestens zwei Replikentypen voneinander unterschieden werden: - die indirekte Replik, die aus mehreren Sätzen besteht, wobei der erste Satz einer direkten Replik entspricht und die übrigen Sätze die Situation ausmalen, beschreiben u.ä.

- die direkte Replik, die aus einem Satz besteht und in der nur kurz geantwortet wird.

Direkte Repliken haben durchschnittlich kürzere Sätze und überwiegen in der Prosa, indirekte dagegen im Drama.

In dem Drama *Zmŕtvychvstanie* von Karvaš sind 659 direkte Repliken zu verzeichnen, die übrigen sind indirekt. Die durchschnittliche Satzlänge der direkten Repliken beträgt 4,7, die der Repliken überhaupt 5,1, was für die indirekten einen Wert von 5,2 bedeutet.

Die Satzlänge ist ein exakt beschreibbares Element. Wie wir feststellten, kann man bei ihrer Untersuchung die durchschnittliche Satzlänge, die Längen der Intervalle, die Dispersion von Satztypen, die durchschnittliche und maßgebende Abweichung, den variierenden Koeffizienten, die Variabilität und andere Proportionen auf ihre Beziehung zum Ganzen überprüfen. Die gewonnenen

Daten signalisieren bestimmte Textqualitäten, die wiederum durch weitere Angaben - besonders durch aus anderen sprachlichen Ebenen stammende Informationen - ergänzt werden müssen.²⁹

Die festgestellten Werte \bar{x} , \bar{d} , σ^2 , σ , V_k und V_r beziehen sich auf die Merkmale einer Äußerung wie Dynamik - Statik, Variabilität - Stereotypie, Expressivität - Begrifflichkeit, Dialog - Monolog usw.

Die Analyse der Satzlänge verweist darauf, daß - auch bei deutlich feststellbarem Einfluß der subjektiven stilbildenden Faktoren - die auf der Einwirkung der objektiven stilbildenden Faktoren beruhende Komponente überwiegt. Dies gilt auch für künstlerische Äußerungen: die Funktion (ausgedrückt z.B. im Genre) steht über dem Autor. Die Satzlänge ist meistens kein auffallendes, sondern nur latentes stilistisches Signal.

GLUTINATION EINES TEXTES

Ähnlich wie ein Satz aus Wörtern besteht, so baut sich ein Text aus Sätzen auf. Zwischen den Phänomenen der Satz- und Textkonstruktion besteht eine gewisse Analogie. Der Unterschied liegt allein darin, daß der Text um eine Stufe höher steht als der Satz.

Wir nennen ein Beispiel: *Wir wissen ganz gut, daß Wissenschaft amüsant sein kann, aber nicht alles, was amüsant ist, gehört auf das Theater.* Dieser Satz von B. Brecht besteht aus 19 Wörtern, zwischen denen Grenzen verlaufen, die in der graphischen Darstellung als Raum erscheinen. Diese Räume sind alle gleich groß. Bei der phonetischen Realisierung jedoch verhält sich dies anders, da sich hier nämlich der Eindruck der Gleichförmigkeit modifiziert oder sogar verliert. Einige Wörter sind recht eng aneinander gebunden, während andere autonom, selbständig sind. Wenn wir die Wörter des Brechtschen Satzes mit den Ziffern 1 bis 19 belegen, so ist die Bindung der Wörter 1 - 2, 3 - 4, 18 - 19 am engsten, während die der Wörter 9 - 10 und 15 - 16 besonders frei ist. Zwischen diesen Extremen existiert aber noch eine ganze Skala enger und lockerer Bindungen, deren Grad von grammatischen Faktoren, semantischen Wertigkeiten u.a. abhängig ist. Im allgemeinen versteht man den Satzbau als Teil der Grammatik, und zwar der Syntax.

Auch zwischen den Sätzen verlaufen Grenzen, die formal - wie die Wortgrenzen durch leeren Raum - durch Interpunktionszeichen markiert sind. Wenn wir auch zugeben, daß die Konstruktion eines Textes komplizierter als die eines Satzes ist, so müssen wir dennoch eine gewisse Analogie anerkennen, denn auch zwischen den Sätzen befinden sich Grenzen verschiedenen Grades, von sehr abrupten bis zu fast unmerklichen. Wenn wir nun das Modell eines Textaufbaus beschreiben, geschieht dies auf vereinfachte Weise. Aufgrund der Überzeugung, daß die Form ein Signal für den Inhalt darstellt, stützen wir uns in dieser Erörterung auf formale Indikatoren, denen wir den Wert von Signalen zuerkennen.

Wir bezeichnen die Kraft, die die Sätze zu einem größeren Ganzen verbindet, als *Glutination*, deren Grad sich - wie wir glauben - messen läßt. Die Grundlage solcher Messungen bildet die grammatische Form im weitesten Sinne, d.h. auch eine latente, die manchmal nur auf dem Hintergrund des Kontextes erkennbar wird. So z.B. hat in den im folgenden zitierten Sätzen von G. Gebhardt ein und dieselbe morphologische Form *Dorit* in Bezug auf das Satzganze verschiedene grammatische Formen: a) *Dorit heißt das Mädchen, das da steil aufgerichtet, gespannt auf dem Stuhl sitzt.* b) *Mittelgroß, zierlich, graziös in der Bewegung, ist Dorit mit sicherem Geschmack gekleidet.* Dies trifft auch für das slovakische Wort *vlast'* (Vaterland) in den folgenden Sätzen zu: a) *Vlast' (Nominativ) je konečné slobodná.* b) *Sami si musíme ubránit' vlast' (Akkusativ).* Hier ist nicht einmal die Wortfolge entscheidend, da ihre Umstellung keine Änderung der grammatischen Form bewirkt: *Konečné je vlast' slobodná. Vlast' si musíme ubránit' sami.* Die grammatische Form, die hier eigentlich durch die Funktion des Wortes im Satz gegeben ist, geht vom Verb aus.

Wie bereits erwähnt, wollen wir die Kraft der Glutination mit Hilfe der grammatischen Formen der Wörter messen. Es bleibt noch hinzuzufügen, daß es sich dabei um die Initialwörter (die am Satzanfang stehenden Wörter) handelt. Sie sind die Mittel, die den Grad der Bindung unmittelbar aufeinanderfolgender Sätze signalisieren.

Bevor wir die vielfältigen Ausprägungen der Glutination beschreiben, geben wir als Beispiel zwei Textabschnitte von W. Borchert. Im ersten sind die Sätze eng, im zweiten nur locker aneinander gebunden: 1. *Drei waren es. In drei alten Uniformen. Einer hatte einen Pappkarton, einer einen Sack. Und der dritte hatte keine Hände.* 2. *Er tappte durch die dunkle Vorstadt. Die Häuser standen abgebrochen gegen den Himmel. Der Mond fehlte und das Pflaster war erschrocken über den späten Schritt. Dann fand er eine alte Planke.*

Je nach der Art der Bindung bekommt der Kontext den Charakter einer entweder flüssigen, geschlossenen, gegebenenfalls auch abrupten, abgehackten oder unregelmäßigen, ungleichen Äußerung,

die stellenweise vorwärtsdrängt oder aber sich dahinschleppt. Diese Unterschiede können funktionell, d.h. vom Autor beabsichtigt sein, doch es gelangen auch gewisse charakteristische Merkmale in den Text, ohne daß sich der Autor ihrer bewußt ist.

Wir zeigen nun, wie sich die Glutination eines Textes analysieren und bestimmen läßt. Die Klassifizierung der Wörter nach Wortarten ist eine lexikalisch-morphologische Angelegenheit. Die Zugehörigkeit eines Wortes zu einer Wortart zeigt, welche Funktionen das betreffende Wort im Satz und - als Satzteil - im Kontext ausüben kann. Ein Substantiv z.B. kann Subjekt, Objekt, Prädikat, Attribut und adverbelle Bestimmung sein. Ähnliche Funktionen können auch Adjektive, Pronomina und Numeralia erfüllen. Auf der Satzebene spricht man nicht mehr von Wortarten, sondern von Satzgliedern und Konstruktionswörtern, mit deren Hilfe sich die Sätze in einen Kontext eingliedern. Vom Standpunkt des Satzes und des Kontextes aus gesehen werten wir autosemantische Wörter und ihre Verbindungen als Satzglieder. Deshalb müssen auch die Initialwörter, die diese Bindungen erzeugen, im Hinblick auf ihre Zugehörigkeit zu bestimmten Satzgliedern beurteilt werden. Wir besprechen im folgenden die verschiedenen Satzglieder und syntaktischen Konstruktionen, die am Satzanfang erscheinen können.

Das Subjekt. Der häufigste Ausgangspunkt einer neuen Aussage ist das grammatische Subjekt eines Satzes. Es ist in den meisten Fällen auch das neue Thema (das psychologische Subjekt des neuen Satzes). Wie die Statistik zeigt, erscheinen im allgemeinen 99% aller Sätze in der sogenannten objektiven Anordnung der Aussageelemente, der psychologischen Teile, d.h. in der Reihenfolge, in der an erster Stelle der Ausgangspunkt steht. Wenn also ein neues Thema erwähnt wird, so tritt es fast immer am Satzanfang und in sachbezogenen Texten oft in der Funktion des Subjekts auf. Deshalb muß die Grenze vor einem mit dem Subjekt beginnenden Satz als abrupt, das neue Thema - oder Subthema innerhalb eines kleineren Kontextes - vom vorangehenden trennend angesehen werden. Wir führen hierzu einige Beispiele aus dem Deutschen und Slovakischen an:

Die jungen Wanderer, die hier einkehren, können gleich in der

Eingangshalle ihr Gepäck abstellen. Der Raum ist geschmückt mit Grünpflanzen, Bildern und Wanderkarten unter Glas. Ein Helfer hinter dem Anmeldeschalter überschaut in der Halle das Kommen und Gehen. Der Speiseraum zur Linken ist durch Sitzecken unterteilt. (H. Griesbach) Das Märchen führt aus der wirklich vorstellbaren Welt unversehens und bruchlos in die magische Welt. Unwirklichkeit wirkt glaubhaft, u. a. durch knappe, nicht schildernde Benennung, formelhafte Wendungen und Verse. Satzbau bevorzugt Nebenordnung. (I. Braak)

Hlad mu pažravo zavtja v útrobach. Krok čoraz stáva sa pomalším a slabším. Kolená ukladajú sa do trojuholníka. Prsia celkom preborila únava. Hlava mu kváčká a l'ahučká, páperová brada vidí sa mu st'a cent. (M. Figuli) Cesta bola rozrytá drobnými jarčekmi, ktoré unášali konský trus a triesky. Zem, čo sa len nedávno vynorila spod snehu, dýchala opojnou vôňou, kým domy, les i voda mali kalný výzor zotavujúcich sa l'udí, čo po dlhej chorobe vyšli na slnko... Zem volala po srne. Les vzdychal tajomným šumením, pramienky veselo zurčali... (M. Urban) Z Bernoláka možno uviesť podstatné meno dlažec... Fándly má napríklad: prekupec, sudec, ochranec. Loos vo svojom slovníku uvádza i podstatné mená kolec, lomec, metec... (Horecký)

Am Satzanfang stehende Subjekte sind typisch für deskriptiv-informative und erörternde Texte. Abschnitte, auf die dieses sehr häufig zutrifft, wirken statisch. Ist ein Satz mit einem subjektiven Anfang Teil einer Beschreibung oder Information, dann ist er so autonom, daß man ihn - von Ausnahmen abgesehen - austauschen oder ohne Störung der Kontinuität oder Kompaktheit streichen könnte. Vor mit Subjekten beginnenden Sätzen ist ein Text am zerbrechlichsten.

Subjekte können sich gegenseitig mehr oder weniger stark überschneiden. Der Grad dieser Überschneidung hängt von ihrer Unerwartetheit ab, die in Bezug auf die Glutination eine der grundlegenden Eigenschaften eines Subjekts ist. Wenn sich Subjekte vollkommen überschneiden, dann liegt Wiederholung vor. In einem solchen Fall ist die Kraft, den Satz vom vorangehenden Kontext zu trennen, nur sehr gering: *Das Gesicht lebt von den Augen. Die Traurigkeit darin macht betroffen. Eine Traurigkeit, die keinen Ausweg kennt.* (G. Gebhardt) Zwischen dem zweiten und dritten Satz dieses Beispiels verläuft eine kaum merkliche, eine Null-Grenze, da das Subjekt wiederholt wird. Es ist wichtig, diese Fälle, die in künstlerischen Texten den Wert poetischer Figuren haben, zu untersuchen.

Die Situation erweist sich bei einem Vergleich der flektie-

renden und analytischen Sprachen jedoch als komplizierter. In den flektierenden Sprachen liegt nämlich die grammatische Form des Verbs im Suffix (*vidíme, vidia ...*), in den analytischen dagegen in der explizierenden Anwesenheit der Personalpronomina (*we see, they see ...*). Soll die Person betont werden, so verwenden die flektierenden Sprachen erklärende Mittel, während in den analytischen (mit Ausnahme des Französischen *je - moi*) die Lautform die gleiche bleibt. Es entsteht die Frage, ob die in den analytischen Sprachen am Anfang stehenden explizierenden Personalpronomina als Subjekt oder nur als morphologisches Inkrement zu werten sind. Dazu einige Beispiele:

Sie wurde nicht schlecht empfangen, aber ein Blick auf ihre neue Schwägerin veranlaßte sie, das Kind sogleich als ihr eigenes vorzustellen. Erst nachdem sie erzählt hatte, daß ihr Mann in einem entfernten Dorf eine Stellung in einer Mühle hatte und sie dort mit dem Kind in ein paar Wochen erwartete, taute die Bäuerin auf und das Kind wurde gebührend bewundert. Nachmittags begleitete sie ihren Bruder ins Gehölz, Holz sammeln. Sie setzte sich auf Baumstümpfe, und Anna schenkte ihm reinen Wein ein. Sie konnte sehen, daß ihm nicht wohl in seiner Haut war. Seine Stellung auf dem Hof war noch nicht gefestigt, und er lobte Anna sehr, daß sie seiner Frau gegenüber den Mund gehalten hatte. Es war klar, daß er seiner Frau keine besonders großzügige Haltung gegenüber dem Protestantenkind zutraute. Er wollte, daß die Täuschung aufrechterhalten wurde. (B. Brecht)

Das Textbeispiel zeigt deutlich, daß die Pronomina nur zur Hervorhebung des neuen Themas verwendet werden, und zwar dort, wo das Verb zur Darstellung der Beziehung zwischen Person und Sache allein nicht ausreicht. Die Pronomina erscheinen hier also in synonymischer Funktion, als *pro-nomina*. In manchen Texten werden diese Synonyme im Hinblick auf den Kontext auch postponiert oder ganz weggelassen: *Nach Arbeitsschluß ging ich immer viel spazieren, sah mir Auslagen an, trank einmal in der Woche ein Glas Bier oder sah mir einen billigen Film an. Setzte ich einmal mit dem Bier oder dem Kino aus, dann konnte ich in der folgenden Woche ein Mädchen mit einladen, aber das war auch alles. (W. Toman)* Wir werten also die in analytischen Sprachen am Satzanfang stehenden Pronomina primär als Synonyme für die Nomen. Als Subjekte stehen sie deshalb im Nominativ. Der Artikel wird als Teil des Wortes verstanden.

Unter den Satzgliedern und Konnektoren hat das Subjekt die

geringste Fähigkeit, Sätze zu größeren Äußerungen zu verbinden, d.h. einen Text zu bilden. In der Skala der zur Glutination fähigen Satzglieder steht es daher mit dem Grad Ø an erster Stelle.

Das Objekt. Es ist eine Ergänzung zum Verb und der typisch prädikative Satzteil, den es vor allem aufgrund seiner lexikalischen Bedeutung, aber auch seiner Position, darstellt. Es steht meistens hinter dem prädikativen Verb: *kupujem knihu, rozumiem každému, hovorme so sestrou, čítam o nešt'astí* (ich kaufe ein Buch, ...). In vielen Sprachen ist die Folge Verb - Objekt so automatisiert, daß ein vorgezogenes Objekt - besonders ein am Satzanfang stehendes - als merkmalshaft empfunden wird. Nicht zuletzt dadurch, daß es meistens an zweiter Stelle und am Satzende erscheint, wurde auch seine intonatorische Form automatisiert: am Satzanfang steht es normalerweise nur bei Betonung. Auch ist es in der Regel kein Träger eines neuen Themas. Trifft dies dennoch zu, so entsteht vor ihm eine ebenso deutliche Pause wie vor dem Subjekt als neuem Thema. Wir geben wiederum einige Beispiele aus dem Deutschen und Slovakischen:

Am nächsten Tag verfehlt er, nachdem ihm der Junge mitgeteilt hat, daß er festgenommen worden sei, eine Stufe und fällt die Treppe hinunter. Mit Prellungen und einer Rückenstauchung kommt er glimpflich davon. (S.L. Englehardt) Unter den unzähligen Gästen, die aus fremden Landen herbeiströmten, die Wunderwerke Salomos zu sehen, war auch die Königin von Saba. Ihren Namen enthält uns die Bibel vor. (E.O. Hauser) Zu ihrem Freundeskreis gehörten auch viele damals führende Leute wie Victor Hugo und Emile Zola. Bei ihr zu Gast waren ferner Ferdinand de Lesseps und Oscar Wilde. (A. Maurois) Der Kaiser hat dir, dem einzelnen, dem jämmerlichen Untertanen, dem winzig vor der kaiserlichen Sonne in die fernste Ferne geflüchteten Schatten, gerade dir hat der Kaiser von seinem Sterbebett aus eine Botschaft gesendet. Den Boten hat er beim Bett niederknien lassen... (F. Kafka) Dabei scheuerte sie immer die Füße aneinander, weil die Kacheln so kalt waren. Schuhe zog sie nachts nie an. (W. Borchert)

Tento posledný spôsob prirovnávania často používal Hviezdoslav. Prirovnaniami vyniká napríklad jazyk Timravinej umeleckej prózy. (S. Šmatlák) Nechcel, aby sa jednotka, čo aj náhodne pozbieraná, rozpadla. Chlapcom bolo treba dať sa najest' a odpočinúť. (D. Tatarka) Šturtina učiteľ'ov matematiky nemá požadovanú kvalifikáciu. Na zvýšenie počtu učiteľ'ov na Slovensku komisia odporučila viaceré opatrení. (Tlač) Kým prvé vystúpenie sa hodnotí kladne, zatiaľ k včerajsiemu koncertu je veľá výhrad. Podrobnosti prinesieme v recenzii, ktorú zverejníme zaj-

tra. (Tlač)

Die Beispiele könnten kein deutlicherer und überzeugender Beweis dafür sein, daß mit einem Objekt eingeleitete Sätze enger an den vorangehenden Kontext gebunden sind als mit einem Subjekt beginnende. Ein weiteres (indirektes) Argument für diese Behauptung ist, daß das Thema als Ausgangspunkt einer neuen Aussage meistens im Nominativ steht. Ein Objekt führt nur auf bestimmte Weise das aus, was bereits gesagt wurde. Es ist daher fest im Text verankert und kann einen hohen Glutationsgrad haben. Es steht eine Stufe höher als das Subjekt, also auf Stufe 1.

Die adverbelle Bestimmung. Sie kann als inhärente oder adhärenente Bestimmung des Verbs auftreten. Die inhärente Bestimmung modifiziert qualifizierende Wörter und die durch das Prädikat ausgedrückte Handlung. Hierzu gehören z.B. Wörter vom Typ *vel'mi, trochu, pekne, dobre, pohodle* (*sehr, ein wenig, schön* usw.) und die modalen Partikel *len, sotva, asi, až, aspon* (*nur, kaum, etwa* usw.). Beide Typen sind vor allem subjektiv wertend und mit dem durch sie determinierten Ausdruck innerlich verbunden. Adhärenente Bestimmungen benennen die äußeren Umstände der Handlung wie Ort, Zeit und Ursache. Zu ihnen zählen Ausdrücke vom Typ *doma, tam, večera, dnes, do školy, začiatkom mesiaca, na posteli, v ruke, z rozličných dôvodov, na zemi, o hodinu* (*zu Hause, dort, gestern* usw.). Während inhärente Bestimmungen fast immer vor dem von ihnen determinierten Ausdruck erscheinen, sind adhärenente freier in ihrer Position. Wir diskutieren die inhärenten Bestimmungen im Zusammenhang mit den Attributen und beschränken uns jetzt auf die adhärenenten, objektiven Umstandsbestimmungen, die meistens durch einen besonderen Kasus realisiert werden.

Da adverbelle Bestimmungen den Verben zugeordnet sind und ihnen meistens vorangehen, wird in den mit ihrer Hilfe beginnenden Sätzen das Subjekt ignoriert: es ist entweder aus dem vorausgehenden Text oder aber der Situation schon bekannt. Der Satz braucht also kein Subjekt, kein neues Thema. Ein mit einer adverbellen Bestimmung eingeleiteter Satz ist daher enger an den vorangehenden Kontext gebunden als ein mit einem Subjekt

beginnender. Dazu einige Beispiele aus dem Deutschen und Slowakischen:

Sie leiht das Geld nur gegen Eintragung von Hypotheken aus. Auf diese Weise sind alle ausgeliehenen Gelder gesichert. (Deutsche Fachtexte aus Recht und Wirtschaft) *Die Fenster können bei Bedarf einzeln oder reihenweise durch Hebeldruck gekippt werden. In den unteren Reihen spiegeln sich je nach dem Neigungsgrad Straßenverkehr oder Fassadenausschnitte der gegenüberliegenden Geschäftshäuser. In den Scheiben der oberen Stockwerke spiegeln sich Himmel, Wolkentreiben, Flugzeuge, kurvende Taubenschwärme und Nacht.* (R. Rehmann) *Höher und heller als im Binnenland ist der Himmel. Glanz überstrahlt das Meer.* (Tlač) *Wir haben es damit nicht genug sein lassen, sondern die Probe aufs Exempel gemacht. Auf einer Strecke, die durch das Eis der Hohen Tatra führte, durch den Schnee der Karpaten und den Frühjahrschlamm der ungarischen Puszta.* (Zeitschrift Ford) *Wie die Leute vom Film behaupten, läuft an jedem Tag des Jahres irgendwo in der Welt ein Tarzanfilm. In Ägypten hat ein Kino sogar zwanzig Jahre lang denselben Tarzanfilm gezeigt.* (Das Beste aus Reader's Digest)

Udalosti rokov 1866-1867 prispeli k zmene stavu, ktorý existoval po vojenskej porážke revolúcie medzi vlastným Rakúskom a jeho uhorskou šľachtou. V júli 1866 sa dostalo Rakúsko do vojny s Pruskom. (M. Gosiorovský) *Matka vedela, že deti už nezahynú a dcéra bola tak preniknutá citom povinnosti, že už všetko dohovárание bolo zbytočné. V slávnostnej hodine pri smrti matkinej posvätila svoj mladý život t'ažkej úlohe, a tej potom aj žila.* (T. Vansová) *Sedliaci z horného konca pokladali ho za somára. V ich očiach klesol už tým, že si Doru vzal.* (F. Hečko) *Tvrдый život bežal. Celý deň pracoval mladý. Vilo vo výčape a večer sedel pri otcovi, ktorému bolo za tmy najhoršie.* (Tlač)

Sätze, die mit einer adverbialen Bestimmung eingeleitet werden, schließen sich enger an den Kontext an als solche, die mit einem Objekt beginnen. Dafür spricht, daß adverbialen Bestimmungen enger zum Prädikat gehören als Objekte. Die adverbialen Bestimmung ist eine nähere Erläuterung des Verbs, während das Objekt nur seine Ergänzung ist. Die adverbialen Bestimmung kann also in gewissem Sinne als Zubehör des Verbs betrachtet werden. Das verbale Prädikat fügt sich von allen Satzteilen am engsten an das ihm Vorangehende an. Wenn der Glutinationsgrad des Subjekts 0 und des Objekts 1 ist, dann muß der der adverbialen Bestimmung folgerichtig 2 sein.

Das verbale Prädikat (Verbum finitum). Aus der Theorie der aktuellen Satzgliederung ist bekannt, daß das Verbum finitum einer Aussage als verbindender Satzteil zwischen Ausgangspunkt

und Kern dient, genauer gesagt: das Verbum finitum steht am Beginn des Kerns einer Aussage. Tritt es am Anfang oder Ende eines Satzes auf, so ist seine Position merkmalshaft, in Bezug auf die Glutination besonders seine Anfangsposition. Wenn wir den Ausgangspunkt mit A, den Kern mit B und das Verbum finitum mit T bezeichnen, so erhalten wir als grundlegende Satzformel ATB. Alle anderen Anordnungen (TAB, TBA, ABT, BTA) sind ungewöhnlich, nicht automatisiert und auffällig. Es sind jedoch auch Aussagen möglich, in denen einer der erwähnten Satzteile fehlt. Dies trifft sehr oft für A und T zu. Fehlt A, so hat die Aussage die Form TB. Sie kommt neben der Form ATB - besonders in belletristischen Texten - am häufigsten vor.

Was beinhaltet nun die Form TB, d.h. die Aussage, in der der Ausgangspunkt fehlt und die mit einem verbindenden Satzteil, anders gesagt, mit dem Verbum finitum, beginnt? Sie verweist auf die Tatsache, daß der Autor die Wiedergabe des Kerns für vorrangig hält und deshalb den Ausgangspunkt wegläßt oder ihn aus dem Kontext für bereits bekannt hält. In beiden Fällen wird der der Aussage TB vorausgehende Raum auf ein Minimum reduziert. Mit einem Verb beginnende Aussagen sind entweder Zeichen der Expressivität oder eines Erzählvorgangs, in dem das Gewicht auf der Handlung liegt. Einige Beispiele aus dem Deutschen und Slovakischen sollen dies verdeutlichen:

Elbe: Was willst du hier? Beckmann: Pennen. Da oben halte ich nicht mehr aus. Das mache ich nicht mehr mit. Pennen will ich. Tot sein. Mein ganzes Leben lang tot sein. Und pennen. Endlich in Ruhe pennen. Zehntausend Nächte pennen. (W. Borchert) Nach Arbeitsschluß ging ich immer viel spazieren, sah mir Auslagen an, trank einmal in der Woche ein Glas Bier oder sah mir einen billigen Film an. Setzte ich einmal mit dem Bier oder dem Kino aus, dann konnte ich in der folgenden Woche ein Mädchen mit einladen, aber das war auch alles. (W. Toman) Entspricht dem nicht in völlig klarer Zuordnung der Begriff "Erholung"? Erholt man sich nicht sowohl von der Arbeit wie für die Arbeit? (J. Pieper) Er hielt sich ein Stück davon unter die Nase. Riecht beinahe wie Kuchen, lachte er leise. (W. Borchert)

Ešte toho istého večera, keď odznel výrok otcov, ešte toho večera vyhl'adala Samka. Vyvolaala ho. Padla mu do náručia, kľčovite ho objala. Čakala spásu a žalovala sa na osud, prekľnala život a žialila. (J. Čajak) Zaklopal som u nich, ale mali zamknuté. Rozhodol som sa, že pōjdem na prednášky. Vyšiel som skōr a pobral som sa radšej peši. (D. Kužel) Najprv zastal pri mne. Tvárim sa, že o ňom neviem. Ukazujem mu chrbát a robím sa,

že práve čítam správu. (V. Šikula) Alžírsko vyšle svoje jednotky na Stredný východ. Oznámila to v noci na dnešok alžírská tlačová kancelária. (Tlač) Dôležitou úlohou kybernetiky je pomáhať vedeckým disciplinám, ktoré sa zaoberajú psychikou, aby sa stali exaktnými a výkonnejšími. Máme na mysli predovšetkým psychológiu, psychiatriu, pedagogiku, estetiku a etiku. (J. Bober)

Die vom Verbum finitum eines Satzes geschaffene Situation zeigt, daß der Satzanfang nur eins von vielen Mitteln ist, die die Glutination signalisieren. Offensichtlich spielt auch die Länge der hypotaktischen Satzgefüge, gemessen an der Zahl der *clauses*, eine Rolle. (Mit dem englischen Ausdruck *clause* bezeichnen wir die einzelnen Sätze - Haupt- und Nebensätze - innerhalb eines Satzgefüges.) Die Teile eines hypotaktischen Satzgefüges sind immer enger aneinander gebunden als die autonomen Sätze eines Textes. Ein wichtiges Argument zur Stützung dieser These ist das formale Faktum, daß im Innern eines Satzgefüges viele Teile mit einem finiten Verb beginnen können. Beispiele aus dem Deutschen und Slovakischen:

Nun da der Arzt gekommen war, fühlte sich der Kranke gleich besser. (T. Mann) *Träfe ich mich morgen nicht mit Hans, hätte ich dich für morgen eingeladen.* (F. Dürrenmatt) *Als ich gestern abend nach Hause kam, brannte bei mir im Wohnzimmer Licht.* (S. Zweig) *"Und jetzt sagen Sie uns Ihren werten Namen", schnalzte die Stimme unwiderstehlich weiter.* (H. Malecha)

Len čo sa stretli, pustili sa do nezvyčajnej hádky. (J. Jenský) *Keby si skutočne nemal času, nemusíš ho navštíviť.* (J. G. Tajovský) *"Sed'te", povedal som jej, a ona si poslušne sadla.* (P. Jaroš)

Wenn die Glutinationskraft von Subjekt, Objekt und adverbialer Bestimmung mit 0, 1 bzw. 2 angegeben wird, so ist die des Verbs mindestens 3. Während der Unterschied zwischen den Graden 1 und 2 nicht besonders groß ist, tritt er zwischen 2 und 3 jedoch sehr deutlich hervor. Dies hängt damit zusammen, daß zwischen einem Objekt und einer durch einen Kasus ausgedrückten adverbialen Bestimmung nur ein fließender Übergang besteht, der Unterschied zwischen Objekt und adverbialer Bestimmung einerseits und dem verbalen Prädikat andererseits jedoch auffällig ist.

Das Attribut. Die mit Hilfe eines autosemantischen Wortes ausgedrückten Satzteile haben oft ein Attribut. Bei Subjekten,

Objekten und Teilen der adverbialen Bestimmung liegt meistens ein adjektivisches Attribut, bei verbalem Prädikat eine inhärente adverbialen Bestimmung vor. Beide Attributstypen gehen normalerweise dem durch sie erläuterten und determinierten Wort voran. Sie können auch am Satzanfang stehen. Da sie in ihrer o.a. Funktion hierarchisch eine Stufe tiefer als die bisher analysierten Satzteile auftreten, werden sie bei der Bestimmung der Glutination oft zugunsten der ihnen übergeordneten Satzteile vernachlässigt.

Die Konnektoren. Präpositionen und Konjunktionen gehören aus grammatischer Sicht zu derselben Klasse, da sie synsemantische Wörter sind. Für die Bestimmung der Glutination sind die Präpositionen irrelevant, weil sie als Zubehör der Wortform diese bilden und präzisieren: *na lavici, pod lavicou, s lavicou, o lavici* (auf der Bank, unter der Bank, mit der Bank, an der Bank). Wir lassen sie daher bei der Untersuchung der Glutination außer acht.

Die durch ihre grammatische Funktion der Wortverbindung integrierten Konjunktionen stellen eine disparate Wortart dar. Disparat ist sie vom Standpunkt der Textstruktur, bilden doch die koordinierenden Konjunktionen lineare und die subordinierenden stufenförmige Kontextteile. Verbindungen vom Typ *Solange er studiert, verdient er kein Geld* (*kým študuje, nezarába*) unterscheiden sich wesentlich von dem folgenden *Er studiert und verdient* (*študuje i zarába*). Durch koordinierende Konjunktionen (*und, oder*) verbundene Satzteile sind syntaktisch gleichwertig, da die Konjunktionen symmetrisch nach beiden Richtungen wirken. Subordinierende Konjunktionen (*daß, weil, wenn, als*) dagegen verweisen nur in eine Richtung. Daher sind die durch sie verbundenen Satzteile syntaktisch ungleichwertig. Wir stellen dieses graphisch dar:

<i>že</i> →		← <i>a</i> →
<i>aby</i> →		← <i>alebo</i> →
<i>keď</i> →		← <i>ale</i> →
<i>čím</i> →	← <i>tým</i>	← <i>avšak</i> →
<i>ako</i> →	← <i>tak</i>	

Steht eine subordinierende Konjunktion am Satzanfang, so muß man annehmen, daß sie auf das Satzinnere verweist und endokli-

tisch ist. Steht dagegen eine koordinierende Konjunktion am Satzanfang, so ist zu vermuten, daß sie auf den Text, in den sie sich einfügt, gerichtet, d.h. ihre linke Seite offen ist, und exoklitischen Charakter hat. Ein mit einer solchen Konjunktion eingeleiteter Satz ruft den Eindruck hervor, als ob die Seite des Satzes, die vorangehen sollte, amputiert oder vom vorausstehenden Text suppliert worden sei. Sätze dieser Art sind eng an den Text, an den sie den neuen Satz anschließen, gebunden.

Wir geben im folgenden Beispiele für die mit subordinierenden Konjunktionen beginnenden Sätze:

Die Verehelichung fand Ende derselben Woche statt. Während der Pfarrer die Trauungsformel murmelte, wandte der Kranke nicht ein einziges Mal den glasigen Blick auf Anna. (B. Brecht) Norderney ist leicht und bequem zu erreichen. Wenn die Eilzüge aus Hannover oder aus dem Rheinland in Nordeich ankommen, braucht der Reisende vom Bahnsteig bis zur Mole nur hundert Schritte zurückzulegen. (R. Rehmann) So haben sich, verlockt durch die hohen Preise, viele italienische Bauern ans Ausgraben gemacht. Als der Bürgermeister eines Dorfes gefragt wurde, wieviele Einwohner sich als illegale Altertumsforscher betätigten, antwortete er kurz und bündig: "Alle!" (G. Gaskill) Er rief seinen Freunden fröhlich zu, er wolle noch ein wenig weiterfahren, und sie riefen fröhlich zurück, das solle er nur. Als er zum drittenmal kam, rief er, er wolle nur seine Kinder holen... (J. Aichinger)

Schôdzka sa už skončila. Ak nie ste o jej priebehu informovaní, prečítajte si zápisnicu. (Ilač) Vo všetkých uvedených prácach je d'alej nedôslednosť v triedení slov, odvodených a zložených. Kým odvodené slová sa triedia podl'a významových kritérií, zložené slová sa vo všeobecnosti triedia podl'a formálnych kritérií. (J. Horecký) V lese bolo absolútne ticho. Kde sa len človek obzrel, všade holé konáre. (M. Urban)

Erscheint eine subordinierende Konjunktion, die den Nebensatz an den vorangehenden Satz oder Kontext anschließt, am Satzanfang, so ist sie ein merkmalshaftes Element. Sie erreicht dann die Glutinationskraft einer koordinierenden Konjunktion: *Zalovala sa. Ze je nešťastná. Ze stále plače. Ze je celá zúfalá. Ze... Ze...*

Es folgen Beispiele für die mit koordinierenden Konjunktionen eingeleiteten Sätze:

Sie hat weiter keinen Wert, meinte er entschuldigend, das weiß ich auch. Und sie ist auch nicht so besonders schön... Aber die blauen Zahlen sehen doch ganz hübsch aus, finde ich.

(W. Borchert) *Am ersten Tag ging es so langsam, daß eine Schnecke sein Vorreiter hätte sein können. Und wer ihn grüßte, dem dankte er nicht. Und wo ein Würmchen auf der Erde kroch, da zertrat er es. Aber schon am dritten und vierten Morgen kam es ihm vor, als wenn die Vögel schon lange nicht mehr so lieblich gesungen hätten wie heute. Und der Tau schien ihm so frisch und die Kornrosen im Feld so rot. Und alle Leute, die ihm begegneten, sahen so freundlich aus, - und er auch. Und jeden Morgen, wenn er die Herberge verließ, war's schöner, und er ging leichter und munterer dahin.* (J.P. Hebel) *Die Kohlrüben schlugen gut an und gaben Viehfutter für einige Monate. Aber die Bronzevasen der Königin blieben von nun an leer.* (V. Auburtin)

Pozrime sa, povedal kapitán Labuda. A spomenul si v okamihu na všetko. (V. Mináč) *Igora zbolelo, tón jeho reči bol láskyplný. Ale pocit'oval aj zadost'učinenie.* (R. Jašík) *V tomto texte sa všetky znaky tzv. kultúrnej západnej slovenčiny, ako sme ich vypočítali vyššie, objavujú v úplnosti. Ba sú tu aj početné stredoslovakizmy hláskové i tvaroslovné.* (E. Pauliny) *No dosiaľ nemáme dostatok experimentálneho materiálu, ktorý by nám úplne osvetl'oval úlohu vnútorného útlmu vo vývoji a pri jednotlivých životných výkonoch človeka. Avšak o jeho význame v živote človeka dá sa usudzovať z analýzy správania sa človeka.* (T. Pardel) *Nedávno som bol vojakom, mal som milióny kamarátov na svete, a neviem spočítat', čo stratili. Ani oni nevedia, čo stratili, lebo už nie sú, oni len boli, sú len ich matere.* (L. Ťašký)

Erscheinen am Satzanfang beide Konjunktionen nebeneinander, so steht die koordinierende an erster und die subordinierende an zweiter Stelle. Auch diese Gesetzmäßigkeit beweist, daß die koordinierende Konjunktion nach hinten, nach links und auf den bekannten Text hin orientiert ist, während die subordinierende auf das Satzinnere verweist:

Das Fernsehen hatte schon im Augenblick seines ersten Funktionierens die Erfindungskraft seiner Installateure überrundet. Denn als es darum ging, der Maschine einen Sinn zu geben, waren die Ingenieure nicht mehr zuständig, ...daß... (H. Naber) *Zu seiner Zeit herrschte in Rom ein reicher und mächtiger Kaiser und hatte eine Tochter mit dem Namen Anima, welche allen Augen angenehm war und den seinen auch. Weil sie aber eine zärtliche Freude an allem Geschaffenen hatte, es mochte nun sichtbar oder unsichtbar sein, so liebte sie auch die Tiere und die Pflanzen.* (W. Bergengruen)

Partikel vom Typ *no, pravda, teda, áno, also, denn, nämlich* haben am Satzanfang die gleiche Glutinationskraft wie koordinierende Konjunktionen und sind in dieser Position ebenfalls exoklitisch:

Tieto najvšeobecnejšie zákony, ako uvidíme, stotožňujú sa s hlavnými metodologickými direktívami. No najvšeobecnejšie záko-

ny jednotlivých disciplín neuplývajú na tvorbu filozofie v rovnakej miere. (V. Filkorn)

Wir hatten die Präpositionen außer acht gelassen, weil sie als Teile nominaler Konstruktionen keine Kontexteinheiten bilden können. Ebenso verhielten wir uns bei den determinativen oder sogenannten endoklitischen Konjunktionen, da sie gewöhnlich zum Satz, nicht aber zum Kontext gehören. Erscheinen am Satzanfang koordinierende Konjunktionen oder in merkmalthaften Kontexten auch subordinierende Konjunktionen mit exoklitischer Funktion, dann liegt der höchste Glutationsgrad von 4 vor.

Die Glutationsgrade 3 und 4, d.h. Sätze, die mit Verben oder exoklitischen Konnektoren beginnen, kommen hauptsächlich in künstlerischen und belletristischen Texten vor, 0, 1 und 2 dagegen sind typisch für sachbezogene Texte.

Sätze können auch mit einem nominalen Prädikat, einer Interjektion oder einem Replikenwort beginnen. Situationen dieser Art sind kennzeichnend für Dialoge, merkmalthaft in Monologen. Am Aufbau des Dialogs sind mehrere Autoren beteiligt. Von einem flüssigen Kontext kann hier keine Rede sein. Um aber die Situation im Dialog bewerten zu können, schreibt man dem nominalen Prädikat denselben Wert wie dem verbalen, den Interjektionen und Replikenwörtern als expressiven und emotionalen Ausdrücken den Wert 5 zu.

Wir können in unserer Analyse nicht alle möglichen Satzanfänge berücksichtigen. Zur Modellierung und Berechnung der Glutation als einem unter vielen Signalen reichen die hier angesprochenen Möglichkeiten jedoch aus.

Es muß betont werden, daß - während zwischen mit Subjekten und mit koordinierenden Konjunktionen beginnenden Sätzen auffällige und evidente Unterschiede bestehen - zwischen Sätzen, die mit Subjekten, Objekten und adverbialen Bestimmungen einerseits und mit Prädikaten, Konnektoren und Replikenwörtern andererseits eingeleitet werden, weniger deutliche Unterschiede vorhanden sind. Man muß eine bestimmte, vom Initialwort ausgehende Skala der Glutationswerte festlegen, von denen die folgenden besonders wichtig sind:

Glutinationsgrad

Ø
1
2
3
4
(5

Initialwort

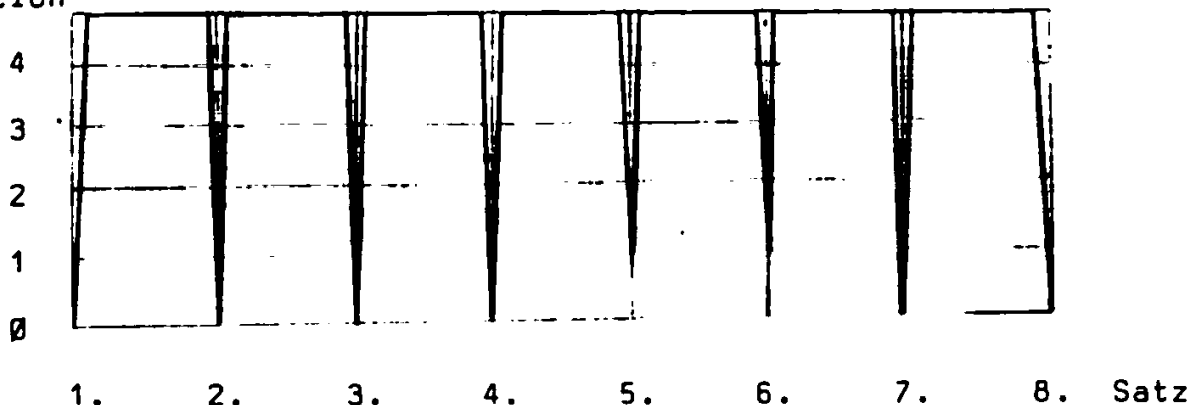
Subjekt
Objekt
adverbielle Bestimmung
verbales Prädikat
exoklitischer Konnektor
Replikenwort)

Wir geben hierzu einige konkrete Beispiele, wobei die zwischen den Sätzen angeführten Ziffern den Glutinationsgrad bezeichnen:

Ø- Die beiden Hauptbestandteile eines Reißverschlusses sind die Zahnketten, die an einem Streifen aus Stoff befestigt sind, und der Schieber, der den Öffnungs- bzw. Schließvorgang besorgt.
-Ø- Die Zahnkette besteht meist aus Metallzähnen, die auf ihrer Oberseite kleine buckelartige Erhöhungen und auf ihrer Unterseite entsprechend geformte Vertiefungen tragen. -Ø- Der Buckel des unteren Zahnes paßt dabei genau in die Vertiefungen des darüberliegenden Zahnes. -Ø- Die beiden Zahnketten, die in geöffnetem Zustand einander gegenüberliegen, sind gegeneinander versetzt. -1- Um den Verschluß zu schließen, müssen nur die beiden Zahnketten so aufeinander zugeführt werden, daß sich jeweils die beiden gegenüberliegenden Zähne ineinander verzahnen können. -1- Diese Arbeit besorgt der Schieber, der oben durch ein keilförmiges Metallstück in zwei auseinanderlaufende Kandle geteilt ist. -Ø- Diese Kandle werden im unteren Teil des Schiebers zusammengeführt. -Ø- Die Ausbildung des Schiebers führt die beiden Zahnketten genau in dem Winkel aufeinander zu.

Der zitierte Text hat terminologisch-wissenschaftlichen Charakter. Die Konfiguration Ø-Ø-Ø-Ø-1-1-Ø-Ø zeigt, daß er keine großen Ansprüche an den Leser stellt, da jeder Satz einen neuen Gedanken, d.h. ein neues Thema, enthält. Die Übergänge zwischen den Sätzen sind abrupt, die Konzentration des Lesers unterliegt Schwankungen. Wir veranschaulichen dies in einer graphischen Darstellung:

Gluti-
nation



Anhand der folgenden Formel berechnen wir die durchschnittliche Glutination:

$$\bar{G} = \frac{\sum x}{n}$$

Das bedeutet für unser Beispiel:

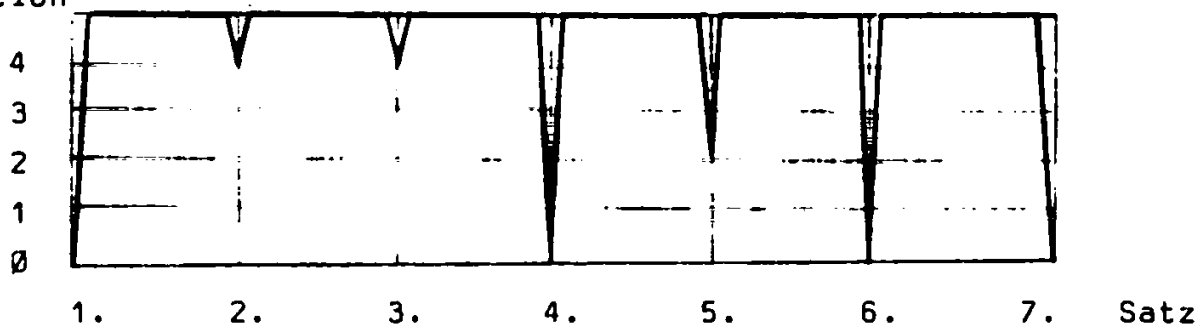
$$\frac{2}{8} = 0,25 \approx 0$$

Wir zitieren nun eine Textstelle von B. Brecht, deren Tempo nicht so regelmäßig wie das der vorangehenden ist. Brechts theoretische und künstlerische Texte sind überhaupt durch unregelmäßiges Tempo, expressiven und astereotypen Text- und Satzbau gekennzeichnet.

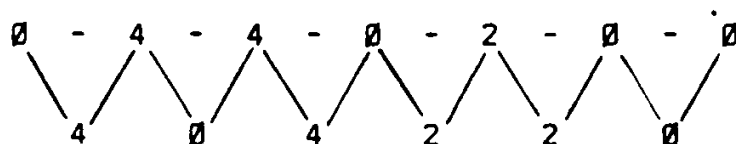
0 - Anna hörte ihm still zu und versprach ihm, nichts Unvernünftiges zu tun. -4- Jedoch wollte sie unbedingt noch schnell in die Gerberei schauen, ob dem Kind nichts fehle. -4- Und sie wollte allein gehen. -0- Sie setzte ihren Willen durch. -2- Mitten in der zerstörten Halle lag das Kind ruhig in seiner Wiege und schlief. -0- Anna setzte sich müde zu ihm und betrachtete es. -0- Sie hatte nicht gewagt, ein Licht anzuzünden...

Die Konfiguration der Glutination ist 0-4-4-0-2-0-0. Die durchschnittliche Glutination beträgt $10/7 = 1,4$, d.h. abgerundet 1. Wir stellen auch dieses graphisch dar:

Glutination



Die Glutination von 1,4 verweist auf Tempo und Kompaktheit des Textes. Zur Charakterisierung der Komposition dagegen ist es wichtig, auch die Intervalle zwischen den Glutinationsgraden festzustellen:



Die durch 6 dividierte Summe 12 verweist auf Intervallstufen von 2.

Wir zitieren im folgenden einige slovakische Texte und geben

die Glutinationswerte an:

3- Málo vravel. -3- Obzeral sa hore úbočou, tu i tam zastal, načúval. -3- Nepovedal, prečo. -3- Zaránky bolo počut' od rúba-
ne zhora ženskej ujúkat'. -Ø- Krstný otec prestal rúbat', zah-
l'adel sa hore a započúval sa, akoby bol chcel poznať dievky
podl'a hlasov. -4- Aj Gažo Chmúrny aj Jakabčič sa na chvíľku
zapozerali hore, no potom, akoby nič, d'alej si robili svoju
robotu. -4- Ale on odložil sekeru... (Ondrejov)

Die durchschnittliche Glutination dieses Textes beträgt 20/7,
also ca. 3, die Intervalllänge 7/6, d.h. ca. 1.

Ø- Mikuláš Aleš vždy obdivoval bohatú umeleckú trofej, ktorú
si odniesol zo Slovenska jeho slávny predchodca Jozef Mánes.
-3- Bol nadšeným obdivovateľom jeho znamenitých diel so slo-
venskou tematikou, medzi ktorými boli aj obrazy z okolia Tren-
čína. -Ø- Kresliarsky génius Jozefa Mánesa bol mu žiarivým prí-
kladom a dozaista túžil poznať kraj, ktorý tak verne a výstiž-
ne zobrazil český maliar. -4- Ved' ako neupadnúť do vytrženia
pri pohľade na veľkolepý obraz "Slovenská rodina", ako neob-
divovať Mánesove postavy slovenských diev a mužov. -Ø- Keď
Mánes robil potulky po trančianskom námestí, po rázovitých na-
šich jarmokoch, Aleš mal iba štyri roky. -3- Narodil sa 18. no-
vembra 1852, zomrel 10. júlia 1913. -4- Iba neskoro dosiahol
uznanie a úctu svojho národa. -3- Spočiatku bol prevažne ilu-
strátorom a vynikajúcim kresliarom. (G. Medvecký)

Die Glutination dieses Textes hat den Wert 17/8, also ca. 2,
die Intervalllänge 19/7, i.e. ca. 3.

Wir vergleichen nun Glutination und Intervalllänge der voran-
gehenden Beispiele:

Texte	1	2	3	4
Glutination	Ø	1	3	2
Intervalle	Ø	2	1	3

Wie die Tabelle zeigt, sind Glutinationsgrad und Intervalllän-
ge voneinander unabhängig. Je höher der Glutinationsgrad ist,
desto kompakter und geschlossener ist der Text. Je größer die
Intervalllänge ist, desto komplizierter und unregelmäßiger ist
der Textrhythmus und desto niedriger der Grad der Prädiktabili-
tät. Stile mit niedrigem Glutinationsgrad sind flüssig, mit ho-
hem dagegen langsam und abgerissen. Texte von Autoren des er-
sten Typs sind monothematisch, während die des zweiten eher po-
lythematischen Charakter aufweisen. Stile mit niedrigem Gluti-
nationsgrad sind stereotyp, rhythmisch einfach, mit hohem Glu-
tinationsgrad hingegen rhythmisch unregelmäßig und kompliziert.
Glutination und Glutationsintervalle signalisieren die Art,
wie der Autor die Realität als Gegenstand seiner Darstellung,

Erörterung oder Beschreibung versteht. Expressive und emotionale Texte haben einen hohen, rationale und sachbezogene dagegen einen niedrigen Index.

Erzählvorgängen kommt ein hoher, Beschreibungen und Informationen ein niedriger und Erörterungen ein mittlerer Glutinationsgrad zu. Die Intervalle jedoch sind in Erörterungen am größten, während sich in Beschreibungen und Informationen das Gegenteil feststellen läßt.

Der Glutinationsgrad hängt mit der eventuellen Textgliederung zusammen. Unterscheiden wir z.B. künstlerische und sachbezogene Texte, so gelten folgende Anordnungen: sachbezogene Texte sind streng logisch, nach Themen oder entsprechend einer Methode gegliedert. Künstlerische Texte sind zwar ebenfalls nach logischem Prinzip aufgebaut, jedoch können hier auch unlogische Anordnungen in ästhetischer Funktion erscheinen.

Wir geben ein Beispiel: Ist die Konfiguration der Glutination 1-4-4-1-3-4-2-Ø-3-3-1-4-2-Ø-3, so könnte man an der Stelle, wo sie den Wert Ø erreicht, einen Abschnitt machen. Für eine genauere Gliederung wäre dies auch nach Grad 1, dann 2, usw. möglich. Diese hierarchische Einteilung kann in künstlerischen Texten gestört sein und ist dann stilistisch merkmalshaft. Sie könnte als Enjambement (dabei wird der Text unerwartet an einer Stelle mit hohem Glutinationsgrad gegliedert) oder als akzentuierende Stelle (hier liegt an einer Stelle, an der es die Satzform erwarten ließe, keine Gliederung vor) gemeint sein.

Wir geben weitere Beispiele: Wäre ein Text mit der Konfiguration Ø-3-3-2-4-4-Ø-3-3 vor dem 5. und dem 8. Satz in Abschnitte eingeteilt, so läge hier ein Enjambement vor. Gäbe es z.B. vor dem 4. und 7. Satz keine Gliederung, so müßten die betreffenden Sätze (4 und 7) als akzentuiert bezeichnet werden.

Texte, deren einleitende Sätze einen hohen Glutinationsgrad aufweisen, signalisieren - ähnlich wie Titel und Überschriften - Dynamik. Die Abschnitte erzählender Texte beginnen oft mit Glutinationsgrad 2 (adverbielle Bestimmung). Dieses wäre dagegen in einer Erzählung, die Aussagen in eine Situation eingliedert, als merkmalshaft zu werten. Titel und sachbezogene,

notionale Texte und Abschnitte beginnen mit einem Glutinationsgrad von \emptyset .

Um die grundlegenden Thesen zu überprüfen und die Glutination von Sätzen und Abschnitten zu vergleichen, untersuchen wir umfangreiche informative, erörternde und deskriptiv-erzählende Texte. Wir wählen als informativen Text eine Nummer der *Pravda* von ca. 15 Seiten, als erörternden das Buch *Dejiny spisovnej slovenčiny* von E. Pauliny und als deskriptiv-erzählenden drei Novellen von M. Urban (*Pred dražbou, Drevený chlieb, Človek, ktorý hl'adal št'astie*). Die erste Untersuchung bringt das folgende Ergebnis (Angabe in %):

Art des Satzanfangs	Inf. Text		Erört. Text		Deskr.-erz. Text	
	Abschnittsbeginn	inneres	Abschnittsbeginn	inneres	Abschnittsbeginn	inneres
Subjekt Objekt Adv. Best. } -----	84	68	72	60	45	36
Vb. finitum Konnektor } -----	16	32	28	40	55	64
Insgesamt	100	100	100	100	100	100

Aus der Tabelle ist ersichtlich, daß in allen Fällen harte Anfänge mit Subjekt und adverbialer Bestimmung am Beginn eines Abschnittes häufiger sind als am Satzanfang im Abschnittsinnen. Im informativen Text sind es mehr als 16%, im erörternden mehr als 12% und im deskriptiv-erzählenden mehr als 9%.

Ferner zeigt sich, daß in allen Texttypen weiche Anfänge mit Verb und Konnektor innerhalb eines Abschnitts häufiger als am Abschnittsbeginn auftreten. Es liegt also eine vollkommen entgegengesetzte Situation vor. Im informativen Text handelt es sich um weniger als 16%, im erörternden um weniger als 12% und im deskriptiv-erzählenden um weniger als 9%.

Die Statistik zeigt, daß zwischen Abschnitten stärkere Brüche (niedriger Glutinationsgrad) als zwischen Sätzen (hoher Glutinationsgrad) erscheinen.

Die unterschiedlichen Glutinationsgrade lassen sich anhand der Formel

$$\bar{G} = \frac{\sum x}{n}$$

errechnen.

Der Glutationsgrad ist offensichtlich am geringsten zwischen den Sätzen und Abschnitten eines informativen Textes (1,5, 1,1). Zwischen denen eines erörternden Textes hat er höheren (1,8, 1,5), in deskriptiv-erzählenden Texten den höchsten Wert (2,3, 2,0).

	<u>Inf. Text</u>	<u>Erört. Text</u>	<u>Deskr.-erz. Text</u>
Glutationsgrad zwischen Sätzen	1,5	1,8	2,3
Glutationsgrad zwischen Abschnitten	1,1	1,5	2,0

Bemerkenswert ist auch, daß der Glutationsgrad der Sätze des informativen Textes nur so hoch ist wie der der Abschnitte des erörternden und daß die Sätze des informativen und erörternden Textes durchschnittlich freier gefügt sind als die Abschnitte des deskriptiv-erzählenden. Deskriptiv-erzählende Texte wurden hier als ein Typ der deskriptiven und erzählenden Texte aufgefaßt, da sich in der Belletristik die Beschreibung nur schwer von der Erzählung trennen läßt.

Aus diesen Ausführungen kann man für die Typologisierung folgende Schlüsse ziehen: hat man den Glutationsgrad von Abschnitten festgestellt, so kann man ihn annähernd auch für den gesamten Text bestimmen. Kennt man die Glutation von Abschnitten, so läßt sich ebenfalls indirekt festlegen, in welchem Maße welches Vorgehen im Text wirksam wird.

Wir weisen abschließend darauf hin, daß die Glutation nicht ad absurdum berechnet werden darf, da sie nur eines unter vielen Signalen darstellt. Ein Text ist ein zu kompliziertes Gewebe, als daß man ihn an bestimmten Stellen mit den Ziffern 1 bis 4 (ggf. auch 5) belegen könnte. Eine Wertung erfolgt vielmehr durch Schätzungen anhand von aufgespürten Gesetzmäßigkeiten. Erst wenn sie durch weitere Indikatoren ergänzt und verifiziert wird, d.h. durch Indikatoren auch anderer Ebenen, kann man ihr glauben und mit Hilfe der gewonnenen Informationen argumentieren. Als Korrektiv kann dabei die bereits erwähnte Kompliziertheit des Satzgefüges, insbesondere die Zahl hypotaktischer clauses dienen. Zwischen ihr und dem Glutationsgrad besteht der folgende Zusammenhang: je niedriger der Glutationsgrad

ist, desto höher ist die Zahl hypotaktischer clauses. Der korrigierte Glutationsgrad könnte durch die Addition der Anzahl von clauses ausgedrückt werden. So z.B. kommen in dem deutschen Text (S. 108) auf 8 Satzgefüge (Glutation 0) 7 hypotaktische clauses, während in dem Brechtschen Text (S. 107) bei 7 Satzgefügen (Glutation 1) 3 hypotaktische clauses auftreten. In dem Text von Ondrejov (S. 108) finden wir bei 7 Satzgefügen (Glutation 3) nur einen clause, in dem von Medvecký (S. 108) bei 8 Gefügen (Glutation 2) 4 hypotaktische clauses.

Der Glutationsgrad und die Intervalllänge sind Gesetzmäßigkeiten, die sich exakt feststellen lassen. Sie dürfen bei der Typologisierung jedoch nicht mit der Hyperkorrektheit gewertet werden, die das exakte Messen assoziieren könnte.

REKURRENZ

Im vorangehenden Kapitel sprachen wir davon, daß sich ein Satz hauptsächlich durch seine Initialwörter in einen Text eingliedert und daß sich anhand des am Satzanfang befindlichen Satzteilens der Grad seiner Eingliederung bestimmen läßt. Dies ist jedoch nur teilweise richtig. Ein Satz ist nämlich durch mehr als nur seine grammatische Form an den Text gebunden. Wir nennen ein Beispiel:

1. *Vychovávala ho hlavne matka. Otec nemal na to nikdy čas.*
2. *Vychovávala ho hlavne matka. Ona formovala jeho svetonázor.*

Im Hinblick auf die Glutination sind diese Sätze identisch, da sie beide mit einem Subjekt beginnen. Sie unterscheiden sich dennoch wesentlich voneinander. Im ersten Satz wird das Thema wirklich verändert (*matka* → *otec* = *Mutter* → *Vater*), während es im zweiten nur synonymisch, d.h. durch ein anderes lexikalisches Element (*matka* → *ona* = *Mutter* → *sie* = *Substantiv* → *Pronomen*), ersetzt wird.

Dieses Beispiel ist einfach und zugleich auffällig, doch im Verlauf des weiteren Textes kommen noch zahlreiche Variationen vor, mit denen sich sogar die Thesen über die Glutination zurückweisen lassen.

Wir wählen die Symbole A, B, C, D, E, X und denken uns einen Text mit 5 Sätzen aus. In jedem Satz erscheinen hauptsächlich die Wörter X, aber auch A, B, C, D, E, während ein neuer Satz nur mit Wörtern zwischen A und E beginnt.

1. Satz: X A₁ X X X B₁ X C₁
2. Satz: A₂ X X X C₂ X X
3. Satz: B₂ X X D₁ X E₁
4. Satz: B₃ C₃ E₂ X X
5. Satz: A₃ X X D₂ X X X

*In dem fingierten Text schließt sich der 2. Satz mit den Wörtern A₂ C₂ an den 1., der 3. mit B₂ an den 1., der 4. mit B₃ an den 3., mit C₃ an den 2. und mit E₂ an den 3., der 5. mit A₃ an den 2. und mit D₂ an den 3. Satz an. Die Wörter X sind für die

Verknüpfung irrelevant. Der fingierte Text ist offensichtlich so gebaut, daß sich die neuen Sätze durch Synonyme, Antonyme, Pronomina u.ä. nicht nur an den jeweils vorangehenden, sondern auch an weiter zurückliegende Sätze anschließen. Die Frage, wie ein neuer Satz den den Kontext eingliedert wird, ist also sehr kompliziert. Wir wollen sie im vorliegenden Kapitel untersuchen. Die Fähigkeit der Sätze, nach hinten zu sehen und sich mit bereits Bekanntem zu überschneiden, bezeichnen wir als *Rekurrenz*.

Die Theorie der Rekurrenz stützt sich auf die der funktionalen Satzperspektive, der sogenannten aktuellen Satzgliederung.

Jeder Satz hat zwei bedeutungstragende Teile: einen Ausgangspunkt und einen Kern (in der Fachliteratur Thema und Rhema). Durch das Thema fügt sich der Satz in den Text ein und verbindet sich mit bereits Bekanntem. Das Thema ist so eigentlich nicht Träger einer neuen Information. Durch das Thema korrespondiert der neue Satz nicht nur mit dem unmittelbar vorausgehenden, sondern es wird auch eine Verbindung zu weiter zurückliegenden Sätzen hergestellt. Das Thema stellt ein ganzes Netz von Bindungen dar und hat die Funktion, Sätze zu größeren Bedeutungseinheiten zu fügen, einen Text zu bilden. Deshalb kann z.B. in den Fällen, wo uns aus dem Kontext oder der Situation bekannt ist, woran sich die neue Aussage anschließt, das Thema auch weggelassen werden und von der gesamten Aussage *expressis verbis* nur das Rhema erscheinen. Das Thema fehlt häufig in linear gefügten Texten - z.B. in Erzählungen -, selten dagegen in erörternden Texten, die flächig oder räumlich gefügt sind.

Das Rhema bringt die neue Information. Es ist der Teil, um dessenthalb die Aussage überhaupt formuliert wird. Es treibt den Text voran und kann daher niemals fehlen.

Wie sind nun aber Thema und Rhema im Satz angeordnet? Wo sind ihre Grenzen? Das Thema steht - von einigen expressiven Aussagen abgesehen - am Anfang, d.h. auf der linken Seite, das Rhema am Ende, d.h. auf der rechten Seite des Satzes. Die Grenze bildet - wieder von Ausnahmefällen abgesehen - das Verb, das die Funktion des transitiven, die linke mit der rechten Satzseite verbindenden Satzteil ausübt. Manchmal - so etwa in kurzen

Sätzen, die nur aus Subjekt und Prädikat bestehen - erscheint das Verb als Rhema. In längeren Sätzen, in denen es transitive Funktion erfüllt, gehört es zum Rhema der Aussage. Aufgrund seiner aktualisierenden Kategorien kommt es dem Thema zu, auf Grund seiner lexikalischen Bedeutung und Rektion dem Rhema der Aussage.

Mit der linken Seite sieht der Satz nach hinten. Sie enthält die rekurrenten Elemente und heißt deshalb auch *rekurrenter* Satzteil. Der rechten Seite gehören die prokurrenten Elemente (sie treiben den Satz nach vorn) an, die folglich den *prokurrenten* Satzteil bilden.

Wie verhält es sich nun mit dem Sättigungsgrad der ihrer Funktion nach so unterschiedlichen Satzteile? Man kann den Satz nicht in der Praxis, wohl aber in der Theorie in seine grammatischen und lexikalischen Komponenten zerlegen. Diese erscheinen gerade beim Funktionieren in der Aussage als direkt voneinander unabhängig. Der erste Teil des Satzes ist in grammatischer Hinsicht gesättigter, da hier mit Hilfe der grammatischen Kategorien eine Verbindung zum Vorangehenden hergestellt wird, in lexikalischer Hinsicht dagegen weniger gesättigt, da in ihm bereits Bekanntes gesagt wird. Im zweiten Teil verhält es sich umgekehrt, da dieser das neue Thema einführt. Man kann so von einer doppelten Asymmetrie der sprachlichen Komponenten eines Satzes sprechen: in grammatischer Hinsicht überwiegt die linke, in lexikalischer die rechte Seite, erst die Summe aller Komponenten macht den Satz als Ganzes symmetrisch. Wenn wir zur Bezeichnung der lexikalischen Elemente L und der grammatischen G verwenden, so lautet die Satzgleichung

$$G + L = L + G,$$

wobei G der prokurrenten und rekurrenten Seite im Verhältnis $G_p < G_r$ und L auf beiden Seiten im entgegengesetzten Verhältnis, also $L_p > L_r$ steht.

Bei einer empirischen Untersuchung der Wortanzahl zeigte sich, daß auf der linken im allgemeinen weniger Wörter als auf der rechten Seite - und zwar im Verhältnis 1 : 2 - stehen. Dies gilt jedoch nicht für alle Texttypen. Die Relationen betragen

1 : 3 (in Konversationen), 2 : 5 (in erörternden Texten), 1 : 2 (in deskriptiv-erzählenden Texten) und 2 : 3 (in informativen Texten). Wie diese Angaben zeigen, enthält z.B. die rechte Seite in Konversationen, in denen im Hinblick auf das bereits Bekannte das Thema gekürzt oder weggelassen wird, mehr Wörter, wohingegen in informativen Texten eine annähernd gleichmäßige Verteilung vorliegt.

Das bisher Gesagte gilt für den einfachen Satz, der ein finites Verb sowie eine rechte und linke Seite aufweist. In Satzgefügen dagegen kommt das Problem der clauses hinzu, die zueinander in parataktischer oder hypotaktischer Beziehung stehen können. Parataktische clauses werden einzeln gesehen, wobei eine eventuelle Konjunktion stets zu dem sich anfügenden Satz gehört. Hypotaktische clauses dagegen gelten als erweiterte Satzteile, die dem Hauptsatz untergeordnet sind. Das Zentrum oder transitive Glied eines auch Nebensätze enthaltenden Satzgefüges bildet das finite Verb des Hauptsatzes. Alles links von ihm Befindliche bezeichnen wir, selbst wenn es sich um vollständige Nebensätze handelt, als den rekurrenten Teil, die rechte Seite als den prokurrenten Satzteil. Dazu einige Beispiele:

Die alte Oper // schließt die Diskussion des Inhaltlichen absolut aus. Das // sind Fortschritte, welche nur anzeigen, daß etwas zurückgeblieben ist. Da // war das Haarband an einen hohen Halm geknüpft /// und der Mann // erschrak bis ins Herz hinein.

Včera popoludní // sme ho stretli na lúke so zvesenou hlavou. Keď sa vrátilš, zaiste nám porozprávaš //, ako sa ti vodilo. Knihovnička // sme získali pre náš plán /// a ostatné už // išlo skutočne ľahko.

Mit Hilfe von Experimenten stellten wir fest, daß die Messung der Rekurrenz anhand der bloßen Wortanzahl für die Zwecke der Typologisierung keine überzeugenden Ergebnisse bringt, da sie sich zu sehr auf die grammatische Komponente der Aussage stützt. Da diese jedoch schon bei der Bestimmung der Glutination in Betracht gezogen wurde, wäre es hier sinnvoller, Informationen über die lexikalische Bedeutung zu verwenden.

Bei der quantitativen Charakterisierung der Wortbedeutung geht man von der These aus, daß die Bedeutung eines Wortes desto allgemeiner ist, je häufiger es im Text vorkommt. Je selte-

ner es dagegen auftritt, desto konkreter, enger und spezifischer ist seine lexikalische Bedeutung. Wörter mit geringem informativen Wert erscheinen auf der linken, mit hohem auf der rechten Seite. Es handelt sich also um den Unterschied zwischen der allgemeinen Bedeutung einerseits und der spezifischen andererseits.

Der Zusammenhang zwischen Häufigkeit und informativer Kraft eines Wortes läßt sich anhand von Gesetzmäßigkeiten erkennen, die in den Wortlisten, in denen die Wörter im Hinblick auf ihre Frequenz geordnet sind, zu beobachten sind: je häufiger ein Wort gebraucht wird, desto allgemeiner ist sein Inhalt, und umgekehrt. Wir führen dazu einige Beispiele aus dem Slovakischen, Russischen, Deutschen, Französischen, Englischen und Spanischen an.

Das Verb *robit'* (*machen, tun*) mit einer Frequenz von 598 steht im Slovakischen auf Platz 115 und hat verhältnismäßig allgemeine Bedeutung. Das Verb *používat'* (*gebrauchen, verwenden*) mit der Frequenz 77 und einer eindeutigeren, spezifischeren Bedeutung nimmt Platz 356, das Verb *oplyvnit'* (*Überfluß haben*) mit der Frequenz 10 Platz 422 und das Verb *zrkadlit'* (*widerspiegeln*) mit der Frequenz 3 Platz 429 ein.

Das Substantiv *človek* (*Mensch*) steht auf Platz 48 (es kommt in Texten mit einem Umfang von 1 Million Wörtern 1 634mal vor), während das Substantiv *mládenec* (*Jüngling, Junggeselle*) - ebenfalls *človek* - mit der Frequenz 64 Platz 388 und das Substantiv *mileneč* (*Geliebter, Liebhaber*) - ebenfalls *človek* und *mládenec* - mit der Frequenz 11 Platz 421 einnimmt.

Das Verb *сказано* (*sagen*) steht im Russischen auf Platz 47, das Verb *ответано* (*antworten*) auf Platz 308 und das Verb *наименовано* (*bezeichnen*) auf Platz 450.

Das Verb *machen* mit seiner recht allgemeinen Bedeutung hat im Deutschen Platz 127, das etwas spezifischere *halten* (ebenfalls *machen*) Platz 321, *führen* Platz 463 und *folgen* Platz 499.

Im Französischen steht das Verb *faire* (*machen*) mit seiner allgemeinen lexikalischen Bedeutung auf dem 22sten Platz, das Verb *dire* (*sagen*) auf dem 25sten, das in semantischer Hinsicht recht eindeutige Verb *expliquer* (*erklären*) auf dem 466sten und das

Verb *étonner* (in *Erstaunen* setzen) auf dem 471. Platz.

Im Englischen nehmen die Adjektive *good* (*gut*) und *deep* (*tief*) die Plätze 113 und 421 ein.

Im Spanischen hält das recht allgemeine Adverb *dónde* (*wo*) Platz 48, die genauere adverbelle Bestimmung *dentro* (*darin*) Platz 190 und das Adverb *moderno* (*modern*) Platz 451.

Die Beispiele zeigen deutlich, daß die Frequenz eines Wortes mit der steigenden Allgemeinheit der lexikalischen Bedeutung konform geht. Wie bereits erwähnt, ist die rekurrente Satzseite in lexikalischer Hinsicht weniger gesättigt ($L_r < L_p$). Wenn wir nun die auf die Frequenz verweisenden Werte dieser Seite addieren, so ist die Summe höher als der entsprechende Wert auf der prokurrenten Satzseite. Dies betrifft natürlich nur autosemantische, nicht aber synsemantische Wörter wie Konjunktionen, Präpositionen und Partikel, die für den grammatischen Bau eines Satzes von Bedeutung sind, denn die Rekurrenz bezieht sich auf die Wörter, die als Träger einer lexikalischen Bedeutung fungieren.

Wie man durch Addition der Wortfrequenzen auf der rechten und der linken Seite das Gewicht der rekurrenten und prokurrenten Satzteile feststellt, wollen wir anhand von fünf Sätzen aus einem Text von M. Figuli darstellen:

Hlad mu pažravo zavija v útrobach. Krok čoraz stáva sa pomalším a slabším. Kolená ukladajú sa do trojuholníka. Prsia celkom preborila únava. Hlava mu kväčká a l'ahučká, páperová brada vidí sa mu st'a cent.

Sätze	Frequenz der Wörter
<i>Hlad mu // pažravo zavija v útrobach.</i>	36 + 12 509 // 1 + 1 + 3
<i>Krok // čoraz stáva sa pomalším a slabším.</i>	207 // 44 + 104 + 140 + 97
<i>Kolená // ukladajú sa do trojuholníka.</i>	51 // 26 + 8
<i>Prsia // celkom preborila únava.</i>	58 // 390 + 3 + 19
<i>Hlava mu // kväčká a l'ahučká, páperová brada // vidí sa mu st'a cent.</i>	573 + 12 509 // 1 3 + 1 + 27 // 1 296 + 12 509 + 29 + 5
Summe der Frequenzen auf der rekurrenten Seite:	25 974
Summe der Frequenzen auf der prokurrenten Seite:	14 676
Division der Frequenzen (25 974 : 14 676):	≈ 2

Wir müssen ergänzen, daß den im Frequenzwörterbuch nicht auf-

geführten Wörtern der niedrigste Wert ($F = 1$) zugeordnet wird. In deskriptiven Texten ist das Verhältnis $L_r : L_p$ normalerweise ziemlich hoch, da auf der prokurrenten Seite seltene Wörter, d. h. Wörter mit geringer Frequenz, vorkommen (Unser Beispiel stammt aus einem deskriptiven Text). In Erörterungen, die auf der prokurrenten Seite weniger begrifflich als beschreibend sind, erscheinen auf der linken Seite viele Wörter mit hoher Frequenz.

Die Division der Frequenzangaben der linken und rechten Seite ergibt eine Zahl, die umso höher ist, je kleiner die prokurrente Seite ist. Wenn wir jedoch bedenken, daß die Summen \bar{F}_r und \bar{F}_p auf der Addition von Frequenzen beruhen, so können wir nicht sagen, daß ein Text umso geschlossener und kompakter ist, je größer die rekurrente Seite, je größer der Quotient $\bar{F}_r : \bar{F}_p$ ist. Im Gegenteil, ein neuer Satz fügt sich umso enger in den Text ein, je kleiner die Summe der Frequenzen, je kleiner also der Quotient ist. Damit das Ergebnis mit der Information über die Glutination übereinstimmt (je größer die Geschlossenheit eines Textes, desto höher der Index der Glutination), muß man die bei der Division gewonnenen Werte umdrehen. Dies geschieht durch die Subtraktion des Quotienten von einer Standardzahl. Wir wählen die Zahl 5 und berechnen den Rekurrenzquotienten (RQ) anhand der Formel:

$$RQ = 5 - \frac{\bar{F}_r}{\bar{F}_p} .$$

Das bedeutet für unser Beispiel von M. Figuli den Wert 3 (5-2).

Das Rekurrenzergebnis wird wesentlich beeinflusst von der Zahl der Pronomina und Hilfsverben, da diese im Gegensatz zu den anderen Wörtern unverhältnismäßig häufig auftreten. Die Rekurrenz muß also anhand von mindestens 20 Sätzen berechnet werden, damit das Endergebnis Zuverlässigkeit erlangt.

Die offene linke Seite wie z.B. in Sätzen, die mit einem finiten Verb beginnen oder eingliedrig sind, wird durch den Kern des vorangehenden Satzes ersetzt. Sie kann in längeren Sätzen sogar ganz entfallen. In Sätzen mit nominalem Prädikat gilt das Nomen als Kern des Prädikats, da das finite Verb hier nur als Träger grammatischer Kategorien fungiert. Ist der Text sehr ex-

pressiv und enthält er viele Sätze mit inversiver Anordnung von Thema und Rhema, so können die linke und rechte Seite nicht mechanisch als rekurrent oder prokurrent angesehen werden, da das Ergebnis verfälscht wäre. Derartige Fälle sind jedoch selten. In größeren Texten spielen sie keine Rolle und üben keinen Einfluß auf das Berechnungsergebnis aus.

Um die Richtigkeit der dargelegten Thesen zu beweisen, stellen wir die Rekurrenz dreier Texte dar: Erzählungen aus *Osmev bolesti* von Barč-Ivan (Bratislava 1968. S. 37-47), die Rubriken *Zvesti* der Bratislaver Abendzeitung *Večerník* (Jg. 14. S. 23-27) vom 28.1. und 1.2.1962 und der wissenschaftlich-erörternde Text *Logika* (Bratislava 1951) von M.S. Strogovič - in der Übersetzung von J. Vavro):

<u>Text</u>	<u>RQ</u>
Erzählender Text	2,5
Informativer Text	3,7
Erörternder Text	4,7

Die Ergebnisse verweisen auf eine Interferenz der am lockersten gefügten Sätze des erzählenden und der am engsten gefügten des erörternden Textes. Wenn wir bedenken, daß die Erzählung linear fortschreitet, so erscheint uns die geringe Rekurrenz als natürlich und begründet: die Sätze stehen nebeneinander, ohne weit nach hinten oder vorn zu weisen. In der Erörterung liegt die entgegengesetzte Tendenz vor: die Strahlen eines neuen Satzes dringen tief nach hinten. Der informative Text ähnelt im Hinblick auf seine lineare Struktur der Erzählung, hinsichtlich seiner informativen Sättigung jedoch der Erörterung. Wie die Zahlenangabe (RQ = 3,7) zeigt, steht er logisch und faktisch zwischen Erzählung und Erörterung.

Im Dialog ist die Situation komplizierter, da hier mehrere einzelne Texte interferieren. Mit Hilfe der Rekurrenz kann man aber auch dort den Kontextzusammenhang der Repliken oder den Grad der Selbständigkeit der einzelnen Sprecher exakt feststellen und das Maß der Abhängigkeit des Dialogs von Situation oder Kontext bestimmen.

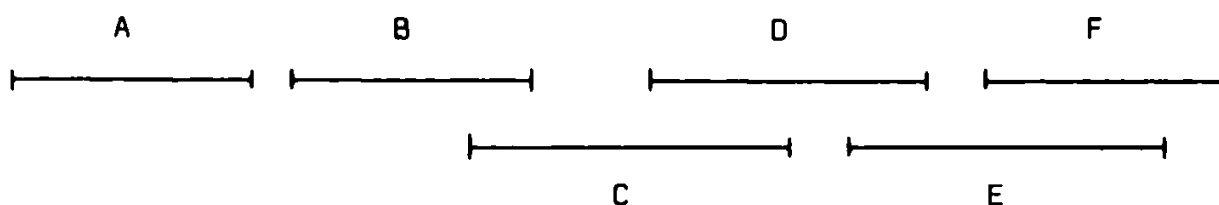
Die Korrelation des RQ und einiger physischer Qualitäten sprachlicher Einheiten. Für die Texte, in denen wir die Rekur-

renz untersuchten, wollen wir nun auch die durchschnittliche Satzlänge in Wörtern, die durchschnittliche Wortlänge in Silben und den Prozentsatz von Sätzen mit fehlendem Thema angeben:

<u>Text</u>	<u>Satzlänge</u>	<u>Wortlänge</u>	<u>Sätze ohne Thema</u>	<u>RQ</u>
Erzählung	13	2	30	2,5
Information	16	2,2	25	3,7
Erörterung	29	2,3	15	4,7

Die Tabelle zeigt, daß der RQ in direktem Verhältnis zu Satz- und Wortlänge und in indirektem zu Sätzen mit fehlendem Thema steht.

Der Rekurrenzgrad gibt aber nicht nur Auskunft über die angewandte stilistische Methode, sondern auch über Tempo, Tempowechsel und Textrhythmus. Wir stellen dies in einem Schaubild dar:



Die Geraden A bis F bezeichnen ganze Sätze, die sich entweder durch direkte Abfolge oder sich überschneidende Synonyme zusammenfügen. Von A bis C herrscht ein hohes Tempo, das sich zwischen D und F verringert. Diese Verlangsamung resultiert nicht nur aus der synonymischen Wiederholung unmittelbar vorangehender, sondern auch entfernterer Sätze oder Satzteile. Je niedriger $\bar{f}_r : \bar{f}_p$, je höher also der RQ ist, desto verhaltener ist das Texttempo, und umgekehrt.

Es gilt nahezu als Gesetz, daß die Rekurrenz in umfangreichen Texten wie Romanen, wissenschaftlichen Monographien und publizistischen Analysen klein ist. In kurzen stilisierten Formen wie Erzählungen und wissenschaftlichen Studien verhält es sich umgekehrt. In den Rahmenteilern umfangreicher Texte ist der RQ höher als in ihrem Innern.

Der RQ allein reicht zur Textcharakterisierung jedoch nicht aus. So ist er z.B. in einer kurzen, expressiv-explosiven Aussage, in der das Rhema betont wird, niedrig. Ebenso aber verhält es sich auch in sachbezogenen Texten mit symmetrischen Satzseiten. Seine unerwartete Veränderung kann in künstlerischen

schen Texten als stilistisch retardierendes oder aber auch akzelerierendes Mittel dienen. Durch seine Erhöhung in epischen Werken erreicht man die sogenannte epische Breite. In zeitgenössischen publizistischen Texten sinkt sein Wert, da sich hier enumerative und faktographische Elemente mit wenigen syntagmatischen Konstruktionen durchsetzen.

VERSTÄNDLICHKEIT EINES TEXTES

Der Grad der Verständlichkeit und Lesbarkeit eines Textes ist ebenfalls Untersuchungsgegenstand der Typologisierung, da man von ihm auf Typ und Charakter von Text und Autor schließen kann. Er ist ein klarer Indikator, der zur Auffindung von Qualitäten und Gesetzmäßigkeiten führt.³⁰

Läßt sich der Grad der Verständlichkeit einer sprachlichen Aussage objektiv bestimmen, wenn neben der Sprache Adressant und Adressat als zwei weitere wichtige Faktoren am Kommunikationsprozeß teilhaben? Sicher ist die Bereitschaft zur Aufnahme einer Äußerung individuell verschieden und hängt von der Kenntnis der Sache, der Situation und der intellektuellen Reife des Adressanten und Adressaten ab. Das ist jedoch kein Hindernis dafür, daß man die Texte als solche von der Situation isoliert und objektiv im Hinblick darauf, welche Ansprüche sie an den Leser stellen, untersuchen kann.

Jede Äußerung läßt sich als relativ schwer oder leicht, anspruchsvoll oder anspruchslos klassifizieren. Solche intuitiv gewonnenen Erkenntnisse lassen sich wiederum mit exakten Informationen untermauern, allerdings nur dann, wenn man sich der Möglichkeit von Abweichungen bewußt ist und die Zahlenangabe nur als Beweis oder Signal, nicht als einzigen und ausschließlichen Qualifikator einer Erscheinung ansieht.

Untersuchungen der Verständlichkeit (bei gelesenen Texten der Lesbarkeit) sind so alt wie die Frequenzwörterbücher. Der Autor des ersten Frequenzwörterbuchs F.W. Kaeding³¹ versuchte als erster, Kriterien für die Verständlichkeit eines Textes zu finden. Größte Aufmerksamkeit widmeten auch die Amerikaner dieser Frage. Zu den bekanntesten von ihnen zählt M.V. Bear³², der - ähnlich wie Kaeding - Länge und Frequenz eines Wortes als grundlegende Faktoren der Textverständlichkeit ansah. Dies geschah in der Zeit, als das älteste englische Frequenzwörterbuch von L.P. Ayers³³, das auf die Bedürfnisse von Pädagogen und Sprachwissenschaftlern abgestellt war, erschien. Von den damaligen Bemühungen, Parameter für die Verständlichkeit zu finden,

zeugt der Titel von Ayers' Arbeit *Measuring Scale for Ability in Spelling*. Jedoch nicht das auf einer geringen Zahl von Wörtern (368 000) beruhende Buch von Ayers, sondern das populäre und grundlegende Wörterbuch von E.T. Thorndike³⁴ diente als Ausgangspunkt für genauere Abhandlungen über die Verständlichkeit. Auf diesem basieren z.B. die Arbeiten von B.A. Lively und L.S. Pressey sowie E. Dalle und J. Challe (1923 bzw. 1948). Die ausführlichste Publikation auf diesem Gebiet stammt von D.G. Peterson und M.A. Tinker³⁵, die neben der Wortlänge auch die Satzlänge, die Zahl der Affixe und - in Bezug auf die schriftliche Äußerung - auch die graphische Gestalt in Betracht ziehen. W.S. Gray³⁶ untersucht als Pädagoge die Lesbarkeit von Texten. Er widmet sich vor allem pädagogisch-psychologischen Aspekten. In einer monographischen Darstellung beschäftigt sich auch der Amerikaner G.R. Klare mit dem Problem der Verständlichkeit.³⁷ Er verzeichnet 482 Arbeiten zu diesem Thema. Wie seine Vorgänger beschränkt auch er sich auf die Form von Wort und Satz als den Grundlagen der Verständlichkeit.

Auf die Arbeit von R. Flesch³⁸ baut der polnische Sprachwissenschaftler W. Pisarek³⁹ auf. Er stützt sich auf sprachlich-stilistische Untersuchungen der polnischen Presse.

Die Autoren der zitierten Arbeiten stimmen in der Ansicht überein, daß Länge und Frequenz von Wörtern und Sätzen grundlegende Faktoren sind. Nach unseren Feststellungen wurde niemals der Einfluß des Kontextes in Betracht gezogen. Daher wollen wir im folgenden zeigen, daß die Verständlichkeit auf drei der Größe nach verschiedenen Ebenen von Wort, Satz und Kontext beruht.

Für die Theorie der Stilistik sind die Faktoren wichtig, anhand derer man einen konkreten Text auf die Werte zwischen den Extremen Verständlichkeit - Unverständlichkeit prüfen kann. Aufgrund vieler Experimente erstellen wir zur Berechnung der Verständlichkeit eine Formel (wir schließen uns dem Usus, die Verständlichkeit mit $R = \text{Readability}$ zu bezeichnen, an) und versuchen, ihren Inhalt mit dem des stilbildenden Aktes sowie die Technik des Rechnens mit Umfang und Anspruch des Ziels in Übereinstimmung zu bringen. Den relativen Grad der Verständlichkeit kann man objektiv für jede fixierte Äußerung bestimmen.

Künstlerische Texte stellen eine Besonderheit dar, da hier 1. die sprachlichen Mittel, die die einzigen Anhaltspunkte für die Berechnung der Verständlichkeit sind, zugleich auch als ästhetische Mittel fungieren, sich aber im Kontext nicht als informative oder ästhetische Mittel unterscheiden lassen, und 2. der Grad der Verständlichkeit eines künstlerischen Textes nichts über den künstlerischen Wert aussagt. Wenn wir auch die Möglichkeit, die Verständlichkeit künstlerischer Texte zu bestimmen, nicht ausschließen, so ist diese Methode doch im allgemeinen hauptsächlich für rein kommunikative Texte typisch.

Die Berechnung der Verständlichkeit auf großen Flächen ist nicht nur praktisch unergiebig, sondern sogar überflüssig. Die Beschränkung auf eine Teilfläche reicht aus. Wir stellten experimentell fest, daß es ohne Belang ist, welchen Teil aus dem Innern einer Äußerung man wählt, da Abweichungen nie mehr als 5% betragen. Für die Rahmenteile gilt dies nicht, da sie in syntaktischer Hinsicht kürzer, in lexikalischer Hinsicht abstrakter und in thematischer Hinsicht reichhaltiger als die inneren Teile sind. Jedoch auch in den Rahmen- und inneren Teilen eines Textes betragen die Abweichungen nicht mehr als $\pm 10\%$. Es gilt allgemein, daß die Lesbarkeit in direktem Verhältnis zum Textumfang steht: je umfangreicher ein Text ist, desto höher ist auch der Grad seiner Lesbarkeit, und umgekehrt.

Die Verständlichkeit eines Textes hängt von mehreren Faktoren ab, z.B. die Kompliziertheit des ausgedrückten Gedankens. Je komplizierter dieser ist, desto schwieriger und umfangreicher sind auch die zur Kommunikation verwandten Sätze. Die Kompliziertheit eines Gedankens steht also in indirektem Verhältnis zur Verständlichkeit. Ein formales Signal ist dabei die Satzlänge, deren Durchschnitt wir zur Berechnung der Lesbarkeit in Betracht ziehen. Wir benötigen dazu folgende Formeln:

$$\bar{x} = \frac{\sum x}{n} \quad \text{oder:} \quad \bar{x} = \frac{\sum x_i n_i}{n_i}$$

Wie Versuche zeigten, reicht die erste (einfache) Formel zur Bestimmung der Lesbarkeit völlig aus. Ihre Angabe kann zur vollen Zahl aufgerundet werden. Wir werden allerdings im folgenden die durchschnittliche Satzlänge (\bar{x}) durch das Zeichen λ_v erset-

zen.

Ein weiterer, die Verständlichkeit beeinflussender Faktor ist die begrifflich-terminologische Schwierigkeit. Je exklusiver ein Thema ist, desto schwieriger ist auch der Wortschatz der Äußerung. Ausgefallene, seltene und begrifflich anspruchsvolle Wörter haben eine gemeinsame physische Qualität: sie sind durchschnittlich länger. Die Wortlänge ist also ein Signal für die begriffliche Kompliziertheit eines Textes. Wir wollen nun diese These mit Angaben aus der Arbeit *Frekvencia slov v slovenčine* belegen. Die Silbenlänge der 30 dominierenden slovakischen Wörter beträgt 1,16, die der folgenden dreißig 1,53, im 145. Dreißigstel 2,5 und im 280. mehr als 3,0. Die durchschnittliche Silbenlänge eines slovakischen Wortes ist 2,3. In Kaedings Wörterbuch beträgt sie 1,03 (im ersten Dreißigstel), 1,33 (im nächsten Dreißigstel) und 2,86 (im 145. Dreißigstel. Je länger also ein Wort ist, desto seltener wird es gebraucht. Den Zusammenhang zwischen der begrifflichen Schwierigkeit und der Länge der sprachlichen Einheiten haben wir bereits an anderer Stelle bewiesen. Wir geben nun die Wortlänge in Silben an und berechnen sie auf ähnliche Weise wie die Satzlänge. Da das Endergebnis von der Auffassung, ob unsilbische Präpositionen als einsilbige Wörter gewertet werden oder nicht, unabhängig ist, d.h. nicht verfälscht wird, rechnen wir sie zu den einsilbigen Wörtern.

Die Wortlänge bezieht sich jedoch nicht nur auf die Begrifflichkeit oder Exklusivität eines Themas, sondern auch auf die Textstruktur. Theoretische Texte fallen z.B. durch Hypertrophie von Nomen und Atrophie von Funktionswörtern und Verben auf. In leichten Konversationen sowie in volkstümlichen und kindlichen Äußerungen dagegen herrscht die entgegengesetzte Tendenz. Funktionswörter und Verben sind durchschnittlich kürzer als Substantive und Adjektive. Die Information über die Wortlänge impliziert also in zweifacher Hinsicht eine Aussage über die Textkompliziertheit und ist daher für die Berechnung der Verständlichkeit besonders wichtig.

Zwischen der durchschnittlichen Wort- und Satzlänge besteht eine offensichtliche Konkordanz. Zur Bestimmung der Lesbarkeit

werden beide Längen nicht verwendet, da sie keinen gemeinsamen Nenner haben. Man kann aber auch keine von ihnen außer acht lassen, denn es handelt sich hier mehr um eine Konkordanz von Tendenzen als um gegenseitige Anhängigkeit. Deshalb muß auch ihre Multiplikation, nicht aber ihre Summe oder Division - wie es in den Arbeiten von E. Dalle und J. Challe geschieht - untersucht werden. Wir bezeichnen in unserer Formel zur Berechnung der Lesbarkeit die durchschnittliche Wortlänge mit λ_s . Versuche ergaben, daß der errechnete Wert auf ein Zehntel aufgerundet werden kann. Die Größen R und λ_s stehen ähnlich wie R und λ_v in indirektem Verhältnis zueinander.

Ein weiterer, die Textverständlichkeit beeinflussender Faktor ist der Index der Wortwiederholung (I_i). Er wird in den o.a. Arbeiten nicht behandelt. Wir denken dabei nicht an die allgemeine, im Frequenzwörterbuch verzeichnete Häufigkeit eines Wortes, wie es die Amerikaner tun, sondern an seine Frequenz in einem konkreten Text. In dem Sinne, wie ihn amerikanische Forscher verwenden (unter ihnen Lively, Dalle, Challe, Bear), fassen wir ihn mit der Angabe über die Wortlänge zusammen, ihn ein zweites Mal in die Berechnungen einzubeziehen, halten wir für falsch.

Der I_i ist ein Indikator für den lexikalischen Reichtum eines konkreten Textes. Er verweist auf die durchschnittliche Wortwiederholung, die von Stil, stilistischer Methode, stilistischer Form usw. abhängt. Der I_i eines Wortes im konkreten, thematisch begrenzten Text muß nicht mit der im Frequenzwörterbuch verzeichneten, auf verschiedenartigem Material basierenden, absoluten Frequenz übereinstimmen. So z.B. erscheint das Wort *transistor* im slovakischen Frequenzwörterbuch mit der Frequenz 3 auf Platz 9 181. In einem von Transistoren handelnden Text könnte es mit hohem I_i Platz 1 belegen. Der I_i verweist außerdem auf die Prädiktabilität, die eine Steigerung der Verständlichkeit bewirkt. Trifft man zum erstenmal auf ein seltenes Wort, so verlangsamt sich das Tempo der Rezeption, das bei weiteren Begegnungen wieder ansteigt. Wie schon erwähnt, wird der I_i anhand der Relation zwischen der Gesamtzahl der im Kontext erscheinenden

den Einheiten (N) und dem benutzten Wortschatz (L) berechnet:

$$I_i = \frac{N}{L}$$

Er hängt in beträchtlichem Maße von der Textlänge ab: je länger der Text ist, desto öfter erscheint in ihm ein und dasselbe Wort. Die Kurve, die die Tendenz der Wortwiederholungen anzeigt, hat logarithmische Form. Je öfter ein Wort wiederholt wird, desto leichter und verständlicher ist der Text.

Die Größen I_i und R stehen zueinander in direktem Verhältnis, R und λ_v sowie R und λ_s in indirektem. Man muß ihren Wert deshalb von einer Standardzahl subtrahieren. Wir untersuchten Extremfälle und stellten fest, daß sich die Division der Multiplikation der λ -Werte durch I_i zwischen 5 und 40 bewegt. Unsere Formel zur Berechnung der Textverständlichkeit lautet:

$$R = 50 - \frac{\lambda_v \cdot \lambda_s}{I_i}$$

Versuche ergaben, daß der zu untersuchende Text nicht mehr als 300 Einheiten in ihrer durch den Kontext determinierten Form zu enthalten braucht. Sind es mehr, so steigt der I_i entsprechend. Dabei ist das Endergebnis logischerweise etwas höher - mit dem Umfang des perzipierten Textes steigt auch seine Verständlichkeit -, bleibt aber nach einer Lektüre von 20 000 Wörtern (ca. 100 maschinenschriftliche Seiten) konstant.

Bei der Berechnung des Verständlichkeitsindex muß man mit Abweichungen von $\pm 10\%$ rechnen und den Wert R auf ganze Zahlen aufrunden. Es zeigt sich dann, daß die Werte zufällig gewählter Texte aus jeweils gleichen stilistischen Bereichen identisch sind.

Die Ergebnisse können durch Berücksichtigung weiterer Faktoren modifiziert werden. Wir denken hierbei an das Vorkommen faktographischer Angaben wie Zahlen und Eigennamen sowie an die genauere Untersuchung syntaktischer Konstruktionen (Subtraktion der Zahl der koordinierenden Syntagmata von der der determinierenden). Modifizierungen könnten auch durch Multiplikation der I_i solcher Elemente entstehen. Die faktographischen Angaben (F) könnten dann durch Addition sämtlicher F und die syntaktischen

Strukturen (C) durch Addition sämtlicher subordinierender Konjunktionen eines Textes berechnet werden. Die Formel sähe also folgendermaßen aus:

$$R = 50 - \left[\left(\frac{\lambda_v \cdot \lambda_s}{I_i} \right) \left(1 + \frac{F}{N} \right) \left(1 + \frac{C}{N} \right) \right]$$

Dieses könnte z.B. bei der Untersuchung von R in informativen (publizistischen) Texten sinnvoll sein. Das Gesamtergebnis würde - wie wir an konkreten Beispielen feststellten - um weniger als den erlaubten Prozentsatz der Abweichung modifiziert.

Wir stellten empirisch fest, daß sich Texte mit R 0 bis 10 an der Grenze der Verständlichkeit bewegen. Theoretisch sind sogar Ergebnisse von weniger als 0 möglich. Hier lägen jedoch keine Kontexte im echten Sinne, sondern in syntaktisch und syntagmatischer Hinsicht künstlich ungeordnete Texte vor. Texte mit R 10 bis 20 sind stilisiert, schwer verständlich und lassen sich nur studieren, nicht aber lesen. Texte mit R 30 bis 40 sind leicht verständlich und lesen sich flüssig, mit R 40 bis 50 besonders leicht. Es handelt sich dabei meistens um Konversationen und Erzählungen.

Wir benutzten im vorliegenden Kapitel abwechselnd die Termini *Verständlichkeit* und *Lesbarkeit*. In amerikanischen Arbeiten wird der Terminus *readability* gebraucht. Wir ziehen den Begriff der Verständlichkeit jedoch vor, da wir ihn in einem weiteren Sinne verstehen und ihn auf schriftliche und mündliche Äußerungen anwenden können (*readability* bezieht sich nur auf schriftliche Äußerungen). In gewissen Fällen geben aber auch wir dem amerikanischen Terminus den Vorzug. Obwohl wir bemerkten, daß sich die Berechnung der Verständlichkeit nur auf rein kommunikative Texte erstreckt, wollen wir nun noch auf die Möglichkeit ihrer Bestimmung in künstlerischen Texten ohne exklusiven Aufbau hinweisen. Ein Roman z.B., in dem eine Sache breit geschildert wird, ist besser lesbar als kurze Formen wie Erzählungen und Novellen. Ebenso verhält es sich mit Kriminalromanen, die einen größeren Grad der Lesbarkeit aufweisen als philosophische Romane und lyrisierte Prosa. Dramen sind leichter lesbar als Gedichte, wobei epische Gedichte am leichtesten zu lesen sind.

Der Grad der Lesbarkeit kann also auch oft bei der Auswahl künstlerischer Texte entscheidend sein. Als unterhaltsam gelten Texte mit R 40 bis 50.

Wir stützten uns bei der Ausarbeitung der Methode zur Berechnung der Verständlichkeit zwar wesentlich auf slovakische Texte, überprüften sie jedoch auch an Texten aus anderen Sprachen. Die slavischen Sprachen sind flektierend. Daher brachte die Anwendung unserer Formel hier auch keinerlei Abweichungen. In den analytischen Sprachen wie auch in den agglutinierenden sind Wort-, Satzlänge und I_i ebenfalls Indikatoren der Verständlichkeit. In den agglutinierenden allerdings ist der Wert der Wortlänge (die durchschnittliche Wortlänge im Ungarischen z.B. beträgt 2,3) höher als in den analytischen (die durchschnittliche Wortlänge im Englischen beträgt 1,43). In Bezug auf die Satzlänge läßt sich die umgekehrte Situation feststellen. Was in den agglutinierenden Sprachen mit Hilfe von Affixen, in den flektierenden anhand von Flexion (beide Methoden tragen zur Wortverlängerung bei) ausgedrückt wird, wird in den analytischen Sprachen durch Wortkombinationen, die den Satz verlängern, erreicht. Die Gesetzmäßigkeiten und der Einfluß des I_i sind - aufgrund ihrer mehr psychologischen als sprachwissenschaftlichen Bedingtheit - in allen Sprachen gültig. Dasselbe läßt sich auf unsere Formel übertragen.

Die Verständlichkeit muß ein immanenter Bestandteil des informativ-kommunikativen Mediums sein, das ohne sie seine grundlegenden Aufgaben nicht erfüllen könnte. Verständlichkeit und Lesbarkeit sind Eigenschaften mit reichem, intensivem Spektrum. Isolieren wir eine Äußerung von der Situation, von Autor und Adressat, so können wir in ihr Elemente finden, die den Grad der Verständlichkeit objektiv signalisieren und es ermöglichen, ihn in Zahlen anzugeben. Die grundlegenden Signale sind die physische Gestalt von Wort, Satz und Text (Variabilität durch Wiederholung der Elemente).

Die obere Grenze der Textverständlichkeit hängt von den objektiv stilbildenden Faktoren, die durch konkrete Funktion und Situation gegeben sind, ab. Es müßte das Ziel jedes Autors sein, sich dem höchstmöglichen Punkt anzunähern. Der stilistische Wert

ist nicht zuletzt von der Länge der zur Perzeption benötigten Zeit abhängig.

MODELLIERUNG DER GENRES

In der Stilistik wird neben den Begriffen des Stils, der stilistischen Methode und der stilistischen Form der Begriff *Genre* gebraucht. Die Kategorie *Stil* ist die abstrakteste und hierarchisch am höchsten stehende Kategorie. Wir verstehen den Stil als Art der zielbewußten Auswahl und gesetzmäßigen Anordnung der sprachlichen Mittel im Hinblick auf die Situation, die Funktion, die Absicht des Autors und die inhaltlichen Komponenten der Äußerung. Die stilistische Methode bezeichnet die Art, wie diese Komponenten angeordnet sind, das kompositionelle Vorgehen, das sich im Innern der Äußerung manifestiert.

Die Begriffe *Genre* und *Form* werden im Wechsel gebraucht. Besonders die Bezeichnung *Genre* wird für zwei hierarchisch verschiedene Ebenen verwendet. Als Genres versteht man einerseits *Naturformen der Poesie* (Goethe), d.h. Epik, Lyrik und Dramatik (nach E. Staiger *das Epische, Dramatische, Lyrische*), andererseits einzelne Formen wie Roman, Erzählung, Tragödie, Nachricht und wissenschaftliche Studie. Wir sehen diese Gebilde als stilistische Formen an, denen die Genres übergeordnet sind. Eine stilistische Form kann also ein konkreter Roman, eine Novelle, eine Erzählung, Legende, Reportage, ein Feuilleton, eine Dissertation, Vorlesung, Rede, ein Vortrag, eine Notiz, Anmerkung, Einladung oder Quittung sein. Unter Genre verstehen wir Epik, Lyrik und Dramatik als künstlerische Stile, das erörternde und beschreibende Genre als wissenschaftliche Stile und das analytische und belletristische Genre als journalistische Stile. Als Genre bezeichnen wir auch den Roman im weitesten Sinne, der die Eigenschaften aller Romane in sich trägt. Ähnliches gilt für die Novelle, Erzählung und alle variablen Formen.

Unter den disparaten Begriff *Genre* fassen wir viele verschiedenartige Texte, die über Texte mit vorgeschriebener Form - wie z.B. administrative Texte - bis zu Texten mit freier Form - wie z.B. künstlerische Prosa - reichen. Das integrierende Element ist die Sprache als ihr grundlegendes, sie aufbauendes Material. Neben den sprachlichen Formen und Genres existieren auch

nicht-sprachliche. Wir finden sie in der Musik, im Theater, in der Geologie, Soziologie, Geometrie, Mineralogie oder Astronomie. Von den sprachlichen Formen und Genres muß man wiederum sprachlich-musikalische und theatralische abgrenzen.

Die systematische Beschreibung des Genres wird im allgemeinen skeptisch beurteilt, weil man sie als spekulatives Schema oder Rezept versteht, das zwar einen Überblick über den Formenreichtum ermöglicht, den Autor aber bindet und durch nicht immer gesicherte Normen hemmt. Wir betonen, daß das vorliegende Kapitel kein Atlas meßbarer und petrifizierter Formen sein soll. Wir wollen vielmehr die Dynamik dieser Phänomene und die Potenz der Sprache als Ausdrucksmittel beschreiben. Wir werden dabei auch zeigen, wie einige Formen erstarrten, während andere sich deutlich veränderten.

Was heißt Modellierung der Genres? Welches Ziel verfolgt sie? Wir sprachen von der Modellierung bereits auf Seite 14, wo sie als Herausbildung der grundlegenden und charakteristischen Merkmale eines Phänomens definiert wurde. Eine sprachliche Äußerung ist ein als Ganzes nicht wiederholbares Phänomen, das stets eine neue, andere Gestalt annimmt. Nur einzelne Elemente können sich wiederholen. Die Zahl derer ist aber sehr verschieden und hängt von Autor, Situation, Funktion und Ziel der Äußerung ab. Wenn sich für ein bestimmtes Genre ein Modell, in dem sich nur wenige Elemente wiederholen, herausbildet, ist es als Genre variabel. Das Modell ist sehr abstrakt. Möglich ist jedoch auch das Gegenteil. Als abstrakt bezeichnen wir die künstlerischen, als konkret die sachbezogenen Genres, wobei die administrativen schon erstarrt und streng vorgeschrieben sein können.

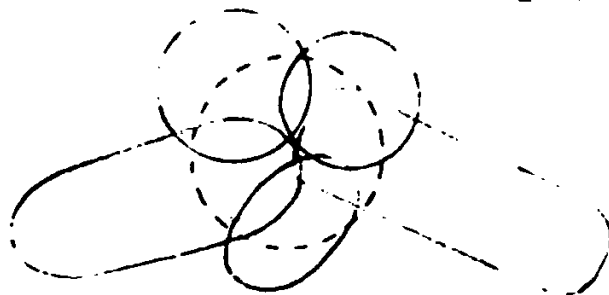
Die Modellierung ist für Theorie und Praxis von Bedeutung. Durch sie werden die relevanten Merkmale erklärt, Begriffe konkretisiert, Gesetzmäßigkeiten des Prozesses und die Herausbildung aufgezeigt.

Es wurde bereits erwähnt, daß die sprachliche Äußerung ein nicht wiederholbares Phänomen ist und daß einige Formen eine langsame, andere dagegen eine schnelle Entwicklung nehmen. Genres und Formen sind historisch bedingte Phänomene, die sich

synchronisch und diachronisch untersuchen lassen. Vom diachronischen Standpunkt ist in der Entwicklung der künstlerischen Formen eine gewisse Divergenz festzustellen, während sich in der der sachbezogenen Genres und Formen eine gewisse Konvergenz abzeichnet. Dies ist hauptsächlich durch den Grad der Freiheit des Autors bei der Konzipierung eines Kontextes bedingt. Die permanente Variabilität steht jedoch nicht unbedingt einer als Prozeß der wissenschaftlichen Abstraktion verstandenen Modellierung entgegen.

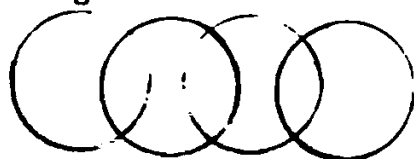
Der älteste Entwicklungs- und Veränderungsprozeß der Genres ist die Spaltung. Sie ist ein überall zu beobachtender natürlicher Prozeß. Durch Spaltung entstanden die Zwillinge Essay und Feuilleton, Roman und Novelle, Notiz und Protokoll, Studie und Dissertation.

In der künstlerischen Literatur ist als Entwicklungsprozeß häufig Dissipation, d.h. Auseinanderentwicklung von Formen und Genres zu beobachten. Durch Dissipation entstanden vor allem die künstlerischen Genres. Dies läßt sich graphisch so darstellen:

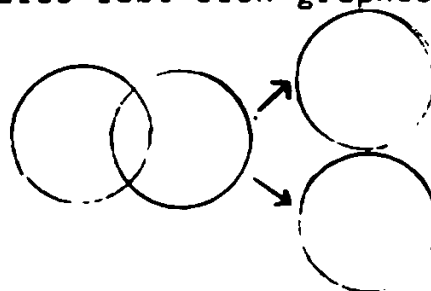


In der Rosette unterscheiden wir fünf verschiedene Formen, die zu einem Genre gehören. Sie entstanden stufenweise und übernahmen von den unmittelbar vorangehenden einige Elemente. Sie bilden dann weitere, neue Elemente aus und unterscheiden sich so wesentlich voneinander, haben nur noch wenig Gemeinsames. Der punktierte Kreis zeigt die Wirksamkeit der abstrakten Modelle dieses Genres an.

Die Entwicklung der Genres und Formen ging nicht selten sukzessiv, kettenförmig vor sich. Auch dies läßt sich graphisch auf folgende Weise wiedergeben:



oder:



So ergab sich aus dem Epos der Roman, aus ihm der Anti-Roman, aus dem Essay die Charakteristik, aus ihr die Kritik und daraus dann die Rezension.

Genres entstehen auch durch Assimilation, die in Fusion gipfelt. Als Assimilation bezeichnen wir die einfache Angleichung, die in sachbezogenen Stilen am häufigsten ist.

Die Ausführungen zeigen, daß die sogenannten reinen Genres in der Theorie nur ad hoc und unter synchronischen Gesichtspunkten existieren. Dieses dürfen wir bei der Modellierung der Genres und Formen nicht außer acht lassen.

Die sprachliche Form ist eine geschlossene Einheit, deren Fläche zwar mit Hilfe sprachlicher Mittel entsteht, die aber in Bezug auf den Kontext, zu dessen Entstehung die sprachlichen Mittel allein nicht ausreichen, von autonomem Wert ist. Sie hat ihre eigene Struktur, in der neben den rein sprachlichen auch noch andere Ausdrucksmittel zur Wirkung gelangen.

Wir verstehen die Genres und Formen vor allem als Phänomene mit einem bestimmten Inhalt. Das Zusammenspiel von Form und Inhalt ist hier - wie bei allen sprachlichen Äußerungen - sehr stark. In den Genres läßt sich eine Symbiose bestimmter inhaltlicher und formaler Elemente beobachten. Bei der Modellierung wird vor allem das Formale - mit Hilfe verschiedener Qualifikatoren -, aber auch das Engagement des Autors erfaßt. Diese Qualifikatoren sind a) morphologischer, b) stilistischer und c) modaler Natur.

a) Die morphologischen formbildenden Elemente eines Genres sind physische Elemente, Elemente der *parole* und komplementäre Elemente. Potentiell sind sie alle an der Herausbildung des Genres beteiligt. Faktisch jedoch existiert eine Parität nicht, ihre Relevanz schwankt, d.h. ihre Wirksamkeit ist veränderlich.

Die morphologischen, sich mit Hilfe physischer Ausdrucksmittel manifestierenden Elemente weisen als wirksamstes unter ihnen die Extensität (E) auf. Diese wird auch dem Perzipierenden als erstes bewußt. Roman, Dissertation und Reportage z.B. sind große, Fabel und wissenschaftliche Studie sowie Artikel dagegen sind kleine Formen. Die Festrede ist in Bezug auf ihren Umfang nicht fixiert. Es ist allgemein üblich, die Größe E als Bestandteil

einer Definition zu verwenden.⁴⁰ Untersuchungen von künstlerischen Texten mit einem Umfang von weniger als 50 000 Wörtern zeigten, daß die Grenze zwischen Novelle und Erzählung bei 10 000 Wörtern liegt. Die Genres Novelle, Erzählung und Roman unterscheiden sich also durch ihre Extensität. Ihnen kommen die folgenden Werte zu: 10 000 (Erzählung), 50 000 (Novelle) und mehr als 50 000 Wörter (Roman). Obwohl die quantitative Charakteristik in der Theorie der künstlerischen Stile stillschweigend abgelehnt wird, verstehen wir sie - natürlich neben anderen Merkmalen - als die deutlichste formale Orientierungshilfe. So ist das Ziel des Romans, *einen breiten Ausschnitt aus dem Leben und das Leben in seiner ganzen Breite und Kompliziertheit* zu erfassen, nur auf großem Raum bei größter Extensität möglich. In der Novelle wird nur *ein kleiner Ausschnitt von Aktion, Leben, Gestalten und Raum* dargestellt, wozu eine geringere Extensität ausreicht, während die Erzählung auf kleinstem Raum *nur ein Bild* erfaßt.

In den übrigen Stilen fehlen diese Extreme, obwohl man auch hier von langen, mittleren und kurzen Formen sprechen kann. So gibt es z.B. in der Journalistik Leitartikel, Feuilleton und Nachricht, in der wissenschaftlichen Literatur Dissertation, wissenschaftliche Studie und Resumé. Die einzelnen Größen bezeichnen wir mit E_3 (für den umfangreichen Text), E_2 (für den mittleren) und E_1 (für den kleinen Text).

Ein weiteres morphologisch formbildendes Element ist die Art der Textgliederung, die sogenannte horizontale Textsegmentation (H). Hierbei haben die Formen, die nicht in Abschnitte geteilt sind oder nicht die trichotomische Gliederung in Einleitung, Hauptteil und Schluß aufweisen, den niedrigsten Wert H_1 . Dazu gehören Zeitungsnachrichten, Fabeln, Begrüßungen u.ä. Formen mit der o.a. Dreiteilung (Artikel, Vorlesung, Leitartikel) kommt der Wert H_2 zu. Polychotomisch gegliederte Formen (Romane, Erzählungen, Märchen, Dissertationen) haben den Wert H_3 , während sich Referate, Rezensionen, Kommuniqués u.ä. nur schwer unter dem Aspekt der Gliederung beurteilen lassen.

Als drittes, empirisch leicht feststellbares morphologisch formbildendes Element verstehen wir die vertikale Textsegmenta-

tion (V). In einem geschlossenen und teleologisch ausgerichteten Kontext existiert eine Hierarchie bestimmter Mikrotex-te. Diese wird durch die folgenden Merkmale realisiert: in sachbe-zogenen Texten durch Letterntypen (Größen etc.), Einschübe, in künstlerischen Texten durch Setzung von Anführungszeichen, durch Abschnitte usw.

Vom Standpunkt der stilistischen Semantik liegen hier zwei diametral entgegengesetzte Gliederungsprinzipien vor, die beide zum Ausdruck von Tiefe, Plastizität und Wichtigkeit beitragen. Ungegliederte Texte (Nachrichten, Einladungen, künstlerische Erzählungen ohne direkte Rede) haben den Wert V_1 , Texte mit zweierlei Letterntypen (in der künstlerischen Prosa mit direk-ter Rede) V_2 und Texte mit mehr als zwei Letterntypen V_3 .

Unter den morphologischen, sich mit Hilfe sprachlicher Mittel der *parole* manifestierenden Elementen gelten als überzeugendste die formbildenden Elemente der Lexik (L). Als distinktives Merkmal zur Modellierung der Formen fungiert der Grad der po-tentiellen Polyvalenz von Wörtern. Man geht dabei von den fol-genden Prämissen aus: a) Es gibt in der Sprache klare, d.h. eindeutige Wörter wie z.B. Termini, faktographische Angaben und Eigennamen, die durchschnittlich sehr lang sind. b) Es existie-ren mehrdeutige Wörter, die erst im Kontext ihre eindeutige Be-stimmung erlangen, z.B. einige häufig gebrauchte Substantive, Verben, Adjektive, Pronomina und Adverbien. Sie sind im Durch-schnitt kürzer als die unter a) beschriebenen Wörter.⁴¹ c) Es gibt außerdem - besonders in künstlerischen Texten - Metaphern. In dieser Funktion können alle, besonders aber die unter b) er-wähnten Wörter auftreten. d) Die durchschnittliche, an der Sil-benzahl gemessene Wortlänge beträgt im Deutschen und in den ro-manischen Sprachen 1,9, im Englischen 1,43, im Slovakischen 2,3, im Tschechischen 2,2, im Russischen 2,2 und im Bulgari-schen 2,3.⁴²

Auf dem Hintergrund dieser Prämissen lassen sich - vom lexi-kalischen Standpunkt gesehen - drei Wortschatzgruppen beschrei-ben: Wörter, die länger als der Durchschnitt und eindeutig sind, bezeichnen wir mit L_1 , Wörter mit durchschnittlicher Län-ge als L_2 , mögliche Metaphern als L_3 . Zur Bestimmung der Wort-

länge genügt die Untersuchung eines kürzeren Textabschnittes. Manchmal ist sie auch empirisch feststellbar. Der Wortschatz L_1 ist typisch für wissenschaftliche Arbeiten, Zeitungsnachrichten und administrative Äußerungen, L_2 für Texte mit allgemeiner Lexik (Populärwissenschaftliche Artikel, Begrüßungen, Erzählungen und Leitartikel), L_3 für künstlerische Texte.

Jede Kontexteinheit, die größer als ein Satz ist und Anfang und Ende hat, muß als Form angesehen werden. Daran kann auch die Tatsache, daß manche dieser Formen nur auf dem Prinzip der Juxtaposition beruhen, nichts ändern. Außerdem gibt es Genres mit einem höchst kompakten, hermetisch abgeschlossenen Innern. Die Syndetizität (S) läßt sich in Graden angeben. Für die Modellierung berücksichtigen wir $S_1 - S_3$. Die Syndetizität manifestiert sich in vielen sprachlichen Mitteln, am deutlichsten jedoch in den Konjunktionen. Mit mathematischen Mitteln wurde festgestellt, daß in slovakischen Texten 5 - 10% Konjunktionen vorkommen.⁴³ In anderen Sprachen wurde empirisch ein ähnlicher Wert bestimmt. Texte mit weniger als 5% Konjunktionen bezeichnen wir als S_1 , mit 5 - 10% als S_2 und mit mehr als 10% als S_3 . Auch diese Werte lassen sich in kleinen Abschnitten oder empirisch feststellen: S_1 in Listen, Thesen, Informationen, S_2 in künstlerischen Beschreibungen, populären Artikeln, Essays, Konversationen, populär-umgangssprachlichen Texten, S_3 in wissenschaftlichen Texten.

Für die Charakterisierung eines Genres ist die Bestimmung der ihm zugrundeliegenden stilistischen Methode (M) wichtig. Als einfachste gelten die deskriptive und informative Methode (in Nachrichten, Beschreibungen, Reportagen), die wir als M_1 bezeichnen. Komplizierter ist die erzählende stilistische Methode, die auf dem Prinzip der Zeit aufbaut und in Erzählungen, Legenden, Märchen, d.h. in künstlerischen und belletristisch epischen Texten erscheint. Wir schreiben ihr den Wert M_2 zu. Am kompliziertesten ist die erörternde stilistische Methode, die in wissenschaftlichen und populär-wissenschaftlichen Texten wie in Vorlesungen und Polemiken zu finden ist. Sie erhält den Wert M_3 . Viele Texte gehen auf eine Kombination der erwähnten stilistischen Methoden zurück. Wir können die M-Werte daher nur em-

pirisch bestimmen.

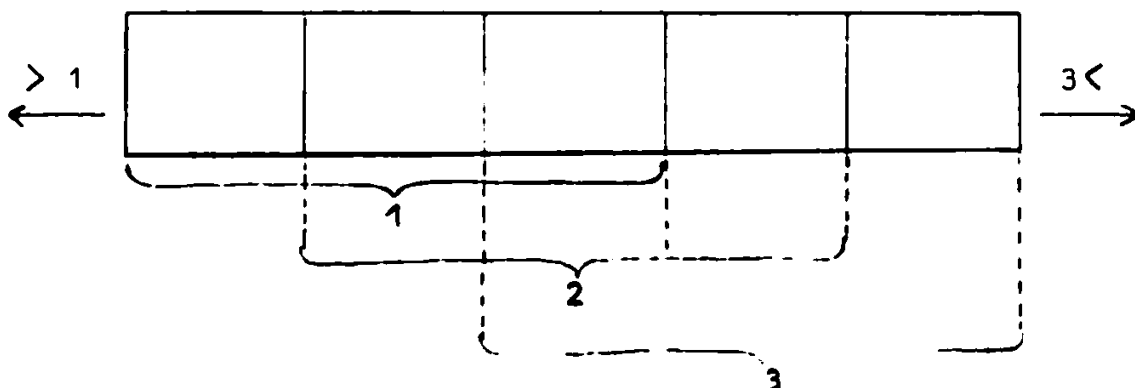
Bei der Entstehung der Formen und Genres spielen neben den direkt im Text verankerten Elementen auch außerhalb befindliche Elemente (die morphologischen, sich mit Hilfe komplementärer Ausdrucksmittel manifestierenden Mittel) eine Rolle. Diese fremden Elemente, zu denen Titel und kinetische, akustische und visuelle Elemente zählen, schieben sich auf bestimmte Weise in den Text. Wir bezeichnen sie als okkasionell (O). Sie kommen vorwiegend in mündlichen und in einigen graphisch attraktiv gestalteten Äußerungen vor. In publizistischen und administrativen Texten ist ihr Anteil unbedeutend. In Begrüßungen, feierlichen Reden, Vorlesungen, Gedichten und volkstümlichen, betont bildlichen Texten muß man mit O_2 rechnen. Der Maximalwert O_3 erscheint in Dramen, Szenarien, Liedern und spontanen Äußerungen wie z.B. Interviews.

Als komplementäres Element fungiert auch der Titel (T). Texte wie z.B. Nachrichten und volkstümliche Artikel, deren T den gesamten Inhalt verkündet, bezeichnen wir als T_3 , Texte mit nur informierenden und andeutenden Titeln als T_2 . Beide informieren und stellen eine Verbindung zwischen Text und Perzipienten her. Es gibt jedoch auch nichtssagende, funktionell verheimlichende Titel (sie kommen in künstlerischen Texten vor), denen wir den Wert T_1 zuschreiben. Sie können niemals fehlen, da sie eigentlich Bestandteil des Textes sind.

Die bei der Modellierung wirksam werdenden Qualifikatoren sind also: E, H, V, L, S, M, O, T.

Jeder Qualifikator kann einen dreifachen Intensitätsgrad haben. Zur Bezeichnung einer sehr geringen oder sehr großen Intensität wählt man die Zeichen $<$ und $>$. So z.B. könnte E in einem Roman und einer Dissertation als $E_{3<}$ und in einer Zeitungsnachricht als $E_{1>}$ bezeichnet werden. Ist ein Intensitätsgrad unklar oder irrelevant, so verzichtet man auf die Zahlenangabe und verwendet den Buchstaben n. So z.B. ist E in einer Reportage E_n . An den Grenzen zwischen 1 und 2 sowie 2 und 3 gehen die Werte spektralförmig ineinander über. Zwischen 1 und 3 gibt es keine scharfen Grenzen, sondern vielmehr deutliche Überlagerungen. Die Intensitätsgrade haben also nur relativen

Wert. Wir veranschaulichen dies in einer graphischen Darstellung:



Wie man sieht, überschneidet sich der Wert 1 nicht nur mit 2, sondern reicht sogar bis an 3 heran. Dies gilt analog für die übrigen Werte.

b) Da die morphologischen Merkmale zur Modellierung des Genres nicht ausreichen, gehen wir nun auf die stilistischen und modalen Merkmale ein. Im konkreten Text erscheinen viele Phänomene, die nur auf der inneren Beziehung von Elementen (z.B. der Sequenz) und auf der Grundlage der enormen Kombinationsmöglichkeiten beruhen, die wiederum mit den außerhalb des Textes befindlichen Ideen und Realitäten zusammenhängen. Im Text existieren zwar Signale, die Bedeutungen jedoch liegen außerhalb des Textes. Man muß also - neben den morphologischen Merkmalen - weitere Indikatoren und Koeffizienten in Betracht ziehen, die auf diese Bedeutungen verweisen.

Im Hinblick auf die stilistischen Qualifikatoren ist der Bedeutungsunterschied von V und M zwischen sachbezogenen und künstlerischen Stilen wesentlich. In sachbezogenen Texten werden in vollem Maße logische Momente und der Einfluß der sprachlichen Norm, in künstlerischen Texten dagegen ästhetische Momente und der Einfluß der ästhetischen Norm wirksam. Auch innerhalb des Systems der sachbezogenen Stile gibt es unterschiedliche Auffassungen der formalen Mittel. So wird z.B. O in einem rhetorischen Text anders gewertet als in einem publizistischen oder wissenschaftlichen Text. Zur genaueren Bestimmung der Modelle müssen also auch die stilistischen Merkmale in Betracht gezogen werden:

- A - administrativer Stil
- W - wissenschaftlicher Stil
- J - journalistischer Stil
- R - rhetorischer Stil
- P - künstlerischer Stil
(= poetischer Stil)

Wird z.B. eine Reportage, die künstlerischer oder journalistischer Art sein kann, modelliert, so müssen dabei in erster Linie die Merkmale P oder J genannt werden. Bei einem Referat sind es A und F. Die Anwesenheit der stilistischen Qualifikatoren ist in den künstlerischen Genres unbedingt erforderlich.

In einem Text können auch in schwächer oder stärkerem Maße Haltung und Absicht des Autors wirksam werden. Die von uns als modales Merkmal bezeichnete Haltung ist so polyvalent, wie es unendlich viele Typen von Menschen und Autoren gibt, die sich in einem Text engagieren. Aus einem Text können Humor, Ironie, Trauer, Zorn und Überzeugung herausgeföhlt werden. Dies alles sind Bestandteile, die sich mit bloßen formalen Mitteln und abstrakten Merkmalen nicht erfassen lassen. Als modale Merkmale kommen nur x und y in Frage, wobei x das Fehlen einer Autorenhaltung bezeichnet. Texte dieses Typs sind objektiv (administrative Ankündigungen, Referate, Notizen, wissenschaftliche Studien, Dissertationen, Reportagen, Nachrichten), Texte vom Typ y dagegen sind subjektiv, ihr Autor ist mehr oder weniger stark engagiert, und tendenziös wie z.B. alle Dichtung oder die Mehrzahl journalistischer Texte, Essays und Erwägungen. Stärke und Färbung von y lassen sich nicht ausdrücken, da seine Semantik so umfassend und reich wie die der Texte ist. Man muß also y ähnlich wie den Wert des Symbols einer lexikalischen Einheit bei der Erzeugung eines Satzes sehen.

c) Die Modelle von Genres und Formen werden in ähnlichen Formeln wie z.B. Molekülverbindungen ausgedrückt, d.h. fixiert. Auf diese Weise wird betont, daß die strukturellen Elemente nicht in additiver, sondern in multiplikativer Folge nebeneinander stehen. Durch Fortfall oder Veränderung eines Elementes ändert sich das Wesen der gesamten Fläche. Die Migration der Textelemente ist ähnlich kompliziert wie die der Moleküle und organischen Zellen. Als Beispiel zeigen wir die Fixierung eines

Romanmodells. Wir wählen als charakterisierende Elemente die Werte $E_3, H_3, V_3, L_3, S_3, M_3, O_1, T_1$. Wir nehmen ferner an, daß der Roman zum stilistischen Bereich P gehört und sich in ihm die Autorenhaltung y realisiert. Das Modell sieht dann folgendermaßen aus:

$$\text{Rom} = P (E_3 H_3 V_3) (L_3 S_3 M_3) (OT) ^y$$

In der Formel werden die physischen, komplementären und die Elemente der *parole* einzeln angegeben. Das Modell darf nicht dogmatisch und als Vorschrift, die die strikte Einhaltung verlangt, gewertet werden. Man könnte überspitzt sagen, daß jeder existierende Roman - wie auch jeder zukünftige - seine eigene Gestalt hat. Ca. 95% aller Romane entsprechen jedoch dem Modell

$$P (E_3 H_3 V_3) (L_3 S_3 M_3) (OT) ^y$$

Die übrigen 5% weichen auf folgende Weise von dieser Formel ab: es kann z.B. V_2 oder V_1 an die Stelle von V_3 , L_2 an die von L_3 , S_2 an die von S_3 oder M_2 an die von M_3 treten. Romane dieser Art sind jedoch selten. Unser Modell beschreibt also den Status quo und gilt - bei einer gewissen Verabsolutierung - für den Roman schlechthin. Da es das Gerüst der Mehrzahl aller Romane darstellt, kann es als Parameter angesehen werden. Diese Ausführungen gelten analog auch für die übrigen Genres und Formen.

Wir stellen nun das Modell der Erzählung vor:

$$\text{Erz} = P (E H_3 V_3) (L_3 S_2 M_2) (OT) ^y$$

Der Unterschied im Vergleich mit dem Romanmodell liegt in den Größen E, S, M.

Die Modelle der Dissertation, der wissenschaftlichen Studie und des populär-wissenschaftlichen Artikels haben die folgenden Formen:

$$\text{Diss} = F (E_3 H_2 V_3) (L_2 S_3 M_3) (OT_2) ^x$$

$$\text{Stud} = W (E_2 H_2 V_2) (L S_3 M_3) (OT_2) ^x$$

$$\text{Art} = W (E H_2 V_2) (L_2 S_2 M_2) (O_3 T_2) ^x$$

Ein Vergleich der Modelle für Roman und Dissertation zeigt Unterschiede in den Größen H, L, T, d.h. im morphologischen Be-

reich. Wichtiger sind hier die stilistischen und modalen Signale. Dissertation und Studie unterscheiden sich in E, V, L, Studie und Artikel in E, L, S, M, O.

Wenn Roman und Dissertation in den Werten H, L und T divergieren, so heißt das jedoch nicht, daß man durch quantitative Korrekturen aus der Dissertation einen Roman machen kann. Die Qualifikatoren P und y stehen qualitativ höher als die Signale V und M. Eine Umformung wäre daher nur theoretisch möglich.

Wir geben das Modell einer kurzen Zeitungsnachricht:

$$\text{Nach} = J \quad (\text{EHV}) \quad (\text{LS}_2\text{M}) \quad (\text{OT}_3)^x$$

Die Symmetrie der morphologischen formbildenden Mittel wird durch den für die Nachricht entscheidenden Titel (T) gestört.

Als kleine Form gilt der Diskussionsbeitrag, in dem die Betonung auf O, L, M liegt. Sein Modell lautet:

$$\text{Disk} = F \quad (\text{EHV}) \quad (\text{L}_2\text{S}_2\text{M}_2) \quad (\text{O}_3\text{T}_2)^x$$

Ein polemischer Diskussionsbeitrag hat anstelle von x den Wert y und anstatt T₂ den Wert T₃:

$$\text{Pol} = F \quad (\text{EHV}) \quad (\text{L}_2\text{S}_2\text{M}_2) \quad (\text{O}_3\text{T}_3)^y$$

Anhand der Modelle läßt sich die Flexibilität eines Genres gegenüber der petrifizierten Form oder der konkreten Definition eines einzelnen und einzigartigen Genres beweisen.

In der Literatur wird von ca. 100 Formen gesprochen. Wir halten diese Vielzahl und Verselbständigung der Formen für nicht gerechtfertigt, da die Modelle zeigen, daß zwischen den einzelnen konstanten Formen nicht unbedingt große Unterschiede bestehen. Wenig überzeugend ist deshalb die Differenzierung von Diskussion und polemischem Diskussionsbeitrag, von Erwägung und Essay, u.ä.

$$\text{Erwä} = F \quad (\text{E}_2\text{H}_2\text{V}_2) \quad (\text{L}_2\text{S}_2\text{M}_3) \quad (\text{OT}_2)^y$$

$$\text{Ess} = F \quad (\text{E}_2\text{H}_3\text{V}) \quad (\text{L}_2\text{S}_2\text{M}_3) \quad (\text{OT}_2)^y$$

Eine ähnliche Situation liegt auch bei Artikel und Vorlesung vor:

$$\text{Art} = F \quad (\text{EH}_2\text{V}_2) \quad (\text{L}_2\text{S}_2\text{M}_2) \quad (\text{O}_3\text{T}_2)^x$$

$$\text{Vorl} = F \quad (\text{EH}_2\text{V}_3) \quad (\text{L}_2\text{S}_3\text{M}_3) \quad (\text{O}_3\text{T}_2)^x$$

Unschärf ist auch die Grenze zwischen Roman und Novelle:

$$\text{Rom} = P (E_3 H_3 V_3) (L_3 S_3 M_3) (OT)_3^y$$

$$\text{Nov} = P (E_2 H_3 V_3) (L_3 S_3 M_3) (OT)_3^y$$

Die angeführten Zahlen sind oft sehr vage, aber die Modellierung ist ein Signal der praktischen Situation. Ein intuitives Vorgehen läßt in der Praxis manchen Unterschied nicht bewußt werden, da die Genres und Formen nicht unbedingt einer bestimmten Vorschrift entsprechen müssen.

Sind also die Modelle in ihrer hier gezeigten Gestalt Modelle von Genres und ggf. Formen? Falls wir vermuten, daß das eine oder andere Modell für ein Genre stehen kann, so geschieht das nur auf der Grundlage der Empirie, die sich auf statistische Angaben stützt: das dominierende Gebilde kann als ideal, als parametrisch verstanden werden, aber nur *hic et nunc*. Die hier vorgeführten Modelle beziehen sich also auf ganz konkrete Gebilde, die metonymisch (*pars pro toto*) für die Genres gelten.

Wir nennen noch weitere Modelle (Feuilleton, Reportage, Referat, Rezept: Arbeitsanleitung und Gutachten):

$$\text{Feu} = J (E_2 H_2 V) (L_2 S_2 M) (OT)_2^y$$

$$\text{Rep} = J (E_2 H V) (L S_2 M) (OT)_2^x$$

$$\text{Ref} = F (E_2 H_2 V_3) (L_2 S_3 M_3) (O_2 T_2)^x$$

$$\text{Rez} = F (E H V) (L S M) (O_2 T_2)^x$$

$$\text{Gtch} = F (E_2 H_2 V_2) (L S_2 M) (OT)_2^x$$

Analog lassen sich noch weitere, ggf. alte, untergegangene oder noch entstehende Gebilde fixieren.

In Bezug auf die Modellierung ist eine Anmerkung erforderlich. Die literarische Kritik wertet künstlerische Literatur auch unter den Aspekten der Zeit, des Milieus, der Situation. Durch den Roman und die außerhalb seines Textes befindlichen Realitäten können wichtige Kontexte entstehen, die ästhetischen oder gesellschaftlichen Wert haben. Die Kritik muß diese Kon-

texte sehen, wenn sie auch im isolierten Text nicht enthalten sind. Die Modellierung stützt sich nur auf den materiellen Text und hat so gnomischen, überzeitlichen, außerhalb der Situation liegenden Wert. Dies ist jedoch bei sachbezogenen Texten anders, da hier aus prinzipiellen Gründen nicht mit Assoziationen, zwischen den Zeilen, zwischen dem Text und der außertextlichen Realität befindlichen Zusammenhängen gerechnet werden kann.

Lassen sich auf die vorgeführte Art Gedichte modellieren? Gedichte haben eine andere Grammatik als Prosa aufzuweisen. Kriterien wie z.B. E, H, M sind daher irrelevant. Stattdessen ist hier das Kriterium O bedeutender. Ein Gedicht läßt sich nicht auf die dargelegte Art modellieren, denn es ist eigentlich als ein künstliches Gebilde, dessen Form ein direkter Bestandteil seines Inhalts ist, anzusehen.

Abschließend wollen wir wiederum betonen, daß die hier besprochenen exakten Methoden nur bedingt angewandt werden können. Die sprachlichen Aussagen sind zu komplizierte Einheiten, die sich nicht in ihrer ganzen Breite beschreiben lassen. Die Modellierung darf nicht zu einer Bremse in der Entwicklung der Genres und Formen werden. Sie ist vielmehr ein Weg in dem Prozeß, ihre Genealogie und Texteigenschaften sowie die Potenz der Sprache zu erkennen.

ANMERKUNGEN

¹Der Begriff Text reicht auch ästhetisch weiter als der Begriff Literatur. Natürlich ist Literatur immer Text und Text nicht immer Literatur, aber Text liegt tiefer im Horizont des Menschen als Literatur, er verwischt nicht so leicht die Spur der Herstellung, er macht die Halb- und Zwischenformen sichtbar, er beweist die vielfältigen Stufen der Übergänge und genau auf diesem Umstand beruht seine Funktion der Erweiterung des Begriffs Literatur.

Text bezieht immer Literatur ein, reflektiert beständig auf Literatur und was auf Text zutrifft, wird auf Literatur zutreffen: in diesem Ausmaß kann der Begriff der Literatur am Begriff des Textes erwogen werden und ein Phänomen an einer Realität gespiegelt.

Was gemacht werden soll, ist Literatur; was gemacht wird, ist Text.

Aus: M. Bense: *Programmierung des Schönen* (1960) S. 51.

²Literatur hat es mit einem verschlossenen, Text mit einem offenen Sein zu tun. Darauf beruht es, daß Literatur einen ideellen, Text einen materialen Charakter besitzt und daß Literatur einen Zustand des Bewußtseins, Text jedoch einen Zustand außerhalb des Bewußtseins festlegt. Es handelt sich ... um die Differenz zwischen einer Idee und einer Struktur.

Siehe M. Bense: aaO 52.

³ebd. Ein Text kann exakt sein, es bietet keine Schwierigkeit, sich vorzustellen, er sei in endlich vielen Schritten entstanden.

⁴J. Sławiński: *Funkcje krytyki literackiej* (1963)

⁵Siehe auch E. Staiger: *Kunst der Interpretation* (1957)

⁶Siehe auch J. Horecký: *Úvod do matematickej jazykovedy* (1969) S. 9-10.

⁷Cicero sagt: Man kann nicht in jeder Lage, für jeden Dienst, unter jedem Einfluß, in jedem Alter, an jedem Ort und in jeder Situation denselben Stil anwenden; in jeder sprachlichen Situation muß man das Angemessene in Betracht ziehen; der Stil muß der betreffenden Sache, dem Adressaten und dem Adressanten angepaßt sein.

Aus: Cicero: *De oratore. Antičnyje teorii jazyka i stilja*. Moskva - Leningrad 1936. S. 196.

⁸Vgl. F. Trávniček: *O jazykovém slohu* (1953) und K. Hausenblas: *K základním pojmům jazykové stylistiky* (1955) S. 1.

⁹Vgl. E. Pauliny: *O funkčnom rozvrstvení spisovného jazyka* (1955) S. 17.

¹⁰Es muß erklärend hinzugefügt werden, daß der Autor bei der Entstehung eines Werkes zweimal in Aktion tritt: 1. als Subjekt, das - entsprechend den subjektiven stilbildenden Faktoren (Temperament, Erziehung, Fähigkeit usw.) - unbewußt den Charakter der Äußerung bestimmt, und 2. als Resultante der Kräfte der

objektiven stilbildenden Faktoren (Milieu, Funktion der Äußerung, sprachliches Material, Kontakt zum Adressaten usw.).

¹¹Siehe J. Mistrík: *Stylistika slovenského jazyka* (1970) S. 84-92.

¹²Siehe L. Doležel: *Verojatnostnyj podchod k teorii chudožestvennogo stilja* (1964) S. 19-29.

¹³Siehe J. Mistrík: *Kompozícia jazykového prejavu* (1968) S. 107-163.

¹⁴Auf dem Gebiet der exakten Methoden sehen wir die Werke von M. Bense, W. Fucks, P. Guiraud, G. Herdan, G.U. Yule (vgl. Bibliographie) und ihrer Anhänger als grundlegend an.

¹⁵W. Fucks sagt in seiner Abhandlung *On the Mathematical Analysis of Style* (1952) S. 122-129: *Every significant text of a grammatical exposition consists of a certain material, the vocabulary, and some structural properties, the style, of its author. The passive vocabulary is formed by the totality of all words of that language, s, the author writes in, the active vocabulary is formed by a certain set, s', of that totality, the selection of which is determined essentially by the sort of literature the text belongs to and depends only in a lower degree on the peculiarity of the author. Style, however, is a characteristic of the author at a certain period of his personal development. The aim of the following investigation is to formulate mathematically some of the properties of structure constituting style, so that for a given text the application of a simple mathematical criterion allows its attribution to a particular author at a certain period of his mental development.*

¹⁶In seiner Arbeit *Les caractères statistiques du vocabulaire* (1954) analysiert P. Guiraud die Lexik der französischen Symbolisten Baudelaire, Rimbaud, Mallarmé, Apollinaire, Valéry und Claudel.

¹⁷Siehe F.M. Frumkina: *Materialy k častotnomu slovarju jazyka Puškina* (1963)

¹⁸Vgl. J. Mistrík: *Z kvantitativného výskumu slovenskej ľudovej piesne* (1967)

¹⁹Diese These vertreten auch G. Herdan (*The Advanced Theory of Language as Choice and Chance*. 1956. S. 75) und - gestützt auf Untersuchungen - F. Pap (*Količestvennyj analiz slovarnoj struktury russkich tekstov*. In: *Vja* VI. 1961. S. 93-110)

²⁰Vgl. J. Jelínek, J. Bečka, M. Těšitelová: *Frekvence slov, slovních druhů a tvarů v českém jazyce* (1961)

²¹Vgl. J. Mistrík: *Frekvencia slov v slovenčine* (1969)

²²G.U. Yule: *The Statistical Study of Literary Vocabulary* (1944)

²³Siehe auch H. Meier: *Deutsche Sprachstatistik* (1964) und J. Berckel, J.A.Th. Van, H. Brandt Corstius, R.J. Mokken, A. van Wijngarden: *Formal Properties of Newspaper Dutch* (1965)

²⁴Im Deutschen liegt eine etwas andere Situation vor, da es

hier vom Verb trennbare und nicht-trennbare Präfixe gibt. Empirisch wurde folgende Situation festgestellt: 1. Verben mit nicht-trennbaren Präfixen sind älter als Verben mit trennbaren. 2. Die deutschen Verben mit nicht-trennbarem Präfix sind den slovakischen Wörtern mit den häufigsten Präfixen (*za-*, *po-*, *vy-*, *s-*, *o-*, *na-*, *pre-*, *u-*, *roz-*) äquivalent. 3. Die deutschen Verben mit trennbaren Präfixen stimmen mit den slovakischen, aus seltenen Präfixen gebildeten (*od-*, *pri-*, *do-*, *pred-*, *v-*, *pod-*, *nad-*) überein. 4. Für den deutschen umgangssprachlichen und künstlerischen Stil ist eine Hypertrophie von Verben mit nicht-trennbaren Präfixen, für den wissenschaftlichen und publizistischen Stil dagegen von Verben mit trennbarem Präfix typisch. 5. Trennbare und wenig gebrauchte Präfixe sind gegenseitig einfach zu übersetzen: *nad-* = *über*, *pod-* = *unter*, *v-/do-* = *in*, *pred-* = *vor*, *pri-* = *bei*, *od-* = *von...* Dagegen sind die deutschen nicht-trennbaren und häufig gebrauchten Präfixe wie ihre Äquivalente im Slovakischen polysemantisch: *be-*, *emp-*, *ent-*, *er-*, *ge-*, *ver-*, *zer-*, *miß-*; *za-*, *po-*, *vy-*, *s-/z-*, *o-*, *na-*, *pre-*, *u-*, *roz-*.

²⁵ In dem Aufsatz *A Linguistic Analysis of Written and Oral Style* (1962) untersucht J. Blankenship die prozentuale Distribution der Wortarten, besonders des Verbs, im Satz. Der Autor schreibt den syntaktischen Strukturen individueller Stile eine größere distinktive Kraft als den Strukturen interindividueller Stile zu.

²⁶ G. Herdan untersucht in *The Advanced Theory of Language as Choice and Chance* die Distribution der Wortarten an russischem Material. Er stützt sich dabei auf 11 Texte, die 46 896 Wörter in der durch den Kontext determinierten Form enthalten: Griboedov: *Gore ot uma*, Dostoevskij: *Unižennye i oskorblennye*, Gončarov: *Obryv*, Saltykov-Šcedrin: *Skazki*, Garšin: *Rasskazy*, Belinskij: *Kritičeskie stat'i*, Amfiteatrov: *Vos'midesjatniki*, Gusev-Orenburgskij: *Gorjaščaja t'ma*, Erenburg: *Burja*, Simonov: *Dni i noči*, Ostrovskij: *Kak zakaljalas' stal'*. Er teilt die Wortarten in 4 Gruppen ein: 1. Substantive, Adjektive, Pronomina, Numeralia, 2. Verben, 3. Partizipien, 4. Adverbien, Partikel, Präpositionen, Konjunktionen. Diese sind prozentual folgendermaßen verteilt:

	<u>allgemein</u>	<u>im Monolog</u>	<u>im Dialog</u>
1. Gruppe	55,1	57,1	51,9
2. Gruppe	26,6	24,8	29,7
3. Gruppe	4,4	6,0	1,6
4. Gruppe	13,9	12,1	16,9

Der Autor beschäftigt sich ferner mit der Wortartdistribution bei den einzelnen Autoren.

²⁷ In seiner Studie *Entwicklung und Beurteilung des Stils* (1965) untersucht H. Fischer anhand von Schüleraufsätzen und 10 literarischen Werken das Vorkommen der Adjektive und Verben und bestimmt auf diese Weise die verschiedenen Stile.

²⁸ Studien über die Satzlänge sind schon aus den Jahren 1888 (L.A. Sherman), 1894 (G.W. Gerwig), 1896 (L. Pound) und 1897 (C. Hildreath) bekannt. Besondere Aufmerksamkeit widmet G. Herdan - *Type-Token Mathematics* (1960) - der Satzlänge als stilistischem Signal. G.A. Leskikis untersucht in *O zavisimosti meždu*

razmerom predloženija i karakterom teksta (1963) die Unterschiede der Satzlangen von Autoren-, Personenrede und erzählter Sache. Vgl. dazu auch seinen Aufsatz *O zavisimosti meždu predloženija i ego strukturoj v raznych vidach tekstov* (1964)

²⁹W. Winter kommt in seiner Arbeit *Relative Häufigkeit syntaktischer Erscheinungen als Mittel zur Abgrenzung von Stilarten* (1961) zu demselben Resultat (nach R.W. Bailey und L. Doležel: *An Annotated Bibliography of Statistical Stylistics*): *On the basis of rich data from two centuries of German letters, an attempt is made to use the relative frequency of certain syntactic features as a criterion for the delimitation of styles; German primarily spoken and German primarily written are isolated as basic categories.*

³⁰In der Fachliteratur wird häufiger der Terminus *Lesbarkeit* (*Readability*) anstelle von *Verständlichkeit* gebraucht. In dieser Arbeit wollen wir beide Begriffe voneinander unterscheiden: Die Verständlichkeit richtet sich auf den Prozeß des Verstehens einer Äußerung, wohingegen die Lesbarkeit allein die Wahrnehmung bzw. Aufnahme eines Textes meint. Die Verständlichkeit befindet sich also auf einer höheren Stufe der Perzeption.

Sachbezogene Texte operieren mit Daten und Termini, die bisweilen selbst Anhaltspunkte sind, anhand derer sich der Leser im Text orientieren und den ganzen Text oder wenigstens einzelne Teile verstehen kann. In künstlerischen Texten dagegen muß er - wenn er nicht nur der reinen Fabel, sondern auch der Ausdrucksform folgt - seine Aufmerksamkeit auf jedes Detail lenken und jeden Punkt genau durchdenken. Geht er nicht so vor, dann nimmt er sich das künstlerische Erlebnis, den Genuß, und zieht vom Text keinen Nutzen, auch wenn er die Fabel versteht und begreift. Die Lesbarkeit eines Textes betrifft also Sujet, Komposition und Sprache. In künstlerischen Texten ist der Prozeß des Lesens anspruchsvoller, der des Verstehens leichter. Für sachbezogene Texte läßt sich das Gegenteil behaupten.

Lesbarkeit und Verständlichkeit stehen einander nicht wie sich widersprechende Erscheinungen gegenüber, sind aber dennoch zwei verschiedene Dinge.

³¹F.W. Kaeding: *Häufigkeitswörterbuch der deutschen Sprache* (1898)

³²M.V. Bear: *The Length of Word as an Index of Difficulty in Silent Reading* (1927)

³³L.P. Ayres: *A Measuring Scale for Ability in Spelling* (1915)

³⁴E.T. Thorndike: *The Teacher's Word Book* (1921)

³⁵D.G. Paterson und M.A. Tinker: *How to Make Type Readable* (1940)

³⁶W.S. Gray: *Progress in the Study of Readability* (1947)

³⁷G.R. Klare: *The Measurement of Readability* (1963)

³⁸R. Flesch: *The Art of Readable Writing* (1962) und *How to Write, Speak and Think More Effectively* (1960)

³⁹W. Pisarek: *Recepty na zrozumiałość wypowiedzi* (1966)

⁴⁰E.M. Forster gibt in seinem Buch *Aspects of the Novel* (1940) die untere Grenze des Romanumfangs mit 50 000 Wörtern an. Er stimmt damit mit den traditionellen Ansichten überein, wenngleich diese expressis verbis nicht verzeichnet sind.

⁴¹Dies stimmt mit dem sogenannten Zipfschen Gesetz überein: Wörter werden umso häufiger gebraucht, je kürzer sie sind, und umgekehrt. Siehe dazu auch G. Herdan: *Type-Token Mathematics* (aaO 33), ferner P. Guiraud: *A propos des caractères statistiques du vocabulaire et l'équation de Zipf* (1955) und B. Mandelbrot: *An Information Theory of the Statistical Structure of Language* (1953)

⁴²Siehe dazu J. Mistrík: *Dĺžka slova a štylistická štruktúra* (1966), vgl. auch G. Herdan: *The Advanced Theory...* (1966) 284.

⁴³Siehe J. Mistrík: *Štylistika slovenského jazyka* (1970) S. 240-251.

BIBLIOGRAPHIE

- Albernathy, Robert: *Mathematical Linguistics and Poetics*. In: *Poetics*. Warszawa und 's-Gravenhage 1961. S. 563-69.
- Altmann, Gabriel: *Kvantitativne metódy v literárnej vede*. In: *Slovenská literatúra X* (1963) S. 432-41.
- Ayres, L.P.: *A Measuring Scale for Ability in Spelling*. New York 1915.
- Bailey, Richard W.: *Statistics and Style. A Historical Survey*. In: *Statistics and Style*. New York 1969.
- Bailey, Richard W. und Lubomír Doležel: *An Annotated Bibliography of Statistical Stylistics*. Michigan 1968.
- Bally, Charles: *Traité de Stylistique Française*. Heidelberg 1909.
- Barth, Gilbert: *Récherches sur la fréquence et la valeur des parties du discours en français, en anglais, et en espagnol*. Paris 1961.
- Bartkowiakowa, A. und B. Gleichgewicht: *O długości sylabicznej wyrazów w tekstach autorów polskich*. In: *Zastosowania matematyki VI* (1962) S. 309-19.
- : *O rozkładach długości sylabicznej wyrazów w różnych tekstach*. In: *Poetyka i matematyka*. Warszawa 1965. S. 164-172.
- Bear, M.V.: *The Length of Word as an Index of Difficulty in Silent Reading*. Chicago 1927.
- Belonogov, G.G.: *O nekotorych statističeskich zakonomernostjach v russkoj pis'mennoj reči*. In: *Voprosy jazykoznanija XI* (1962) 1. S. 100 f.
- Bense, Max: *Aesthetica*. Stuttgart 1954.
- : *Programmierung des Schönen. Allgemeine Texttheorie und Textästhetik*. Baden-Baden 1960.
- : *Theorie der Texte. Eine Einführung in neuere Auffassungen und Methoden*. Köln-Berlin 1962.
- : *Zusammenfassende Grundlegung moderner Ästhetik*. In: *Mathematik und Dichtung*. München 1965. S. 313-32.
- Berckel, J.A., Th.M. Van, H. Brandt Corstius, R.J. Mokken und A. van Wijngaarden: *Formal Properties of Newspaper Dutch*. Amsterdam 1965.
- Blankenship, Jane: *A Linguistic Analysis of Oral and Written Style*. In: *Quarterly Journal of Speech XLVIII* (1962) S. 419-422.
- Boldrini, Marcello: *Le statistiche letterarie e i fonemi elementari nella poesia*. In: *Statistiche Letterarie ed altri*

- Saggi: Contributi des Laboratorio di Statistica*. 6th series. Milano: Società Editrice Vita e Pensiero 1948. S. 1-15.
- Bordelois, I.: *Perspectiva de la estilistica*. Universidad de Buenos Aires 1962.
- Bosanquet, C.H.: *Distribution of Word Frequencies*. In: *Nature* CLXXIX (März 16, 1957) 4 559. S. 596.
- Burwick, Frederick: *Associationist Rhetoric and Scottish Prose Style*. In: *Speech Monographs* XXXIV (1967) S. 21-34.
- Carroll, John B.: *Vectors of Prose Style*. In: *Style in Language*. Cambridge, Mass. und New York 1960. S. 283-92.
- Chatman, Seymour: *Stylistics: Quantitative or Qualitative*. In: *Style* I (Arkansas 1967) S. 29-43.
- Condon, E.U.: *Statistics of Vocabulary*. In: *Science* LXVII (1928) 1 733. S. 300.
- Delbouille, Paul: *A propos de la définition du fait de style*. In: *Cahiers d'Analyse Textuelle* II (1960) S. 94-104.
- Doležel, Lubomír: *Model stylistické složky jazykového kódování*. In: *Slovo a slovesnost* XXVI (1965) S. 223-35.
- : *O stylu moderní české prózy*. Praha 1960.
- : *Verojatosťnyj podchod k teorii chudožestvennogo stilja*. In: *Voprosy jazykoznanija* XIII (1964) 2. S. 19-29.
- und Richard W. Bailey: *Statistics and Style*. New York-London-Amsterdam 1969.
- Fischer, Hardi: *Entwicklung und Beurteilung des Stils*. In: *Mathematik und Dichtung*. München 1965. S. 171-83.
- Flesch, Rudolf: *How to Test Readability*. New York 1951.
- : *How to Write, Speak and Think More Effectively*. New York 1960.
- : *The Art of Readable Writing*. New York 1962.
- Frumkina, F.M.: *Materialy k častotnomu slovarju jazyka Puškina*. Moskva 1963.
- : *Statističeskije metody izučenija leksiki*. Moskva 1964.
- Fucks, Wilhelm: *Mathematische Analyse von Sprachelementen, Sprachstil und Sprachen*. Köln 1955.
- : *On the Mathematical Analysis of Style*. In: *Biometrika* XXXIX (1952) S. 122-29.
- : *Unterschied des Prosastils von Dichtern und anderen Schriftstellern*. In: *Sprachforum* I (1955) S. 234-44.
- und Josef Lauter: *Mathematische Analyse des literarischen Stils*. In: *Mathematik und Dichtung*. München 1985. S. 107-22.
- Garth, Thomas R.: *A Statistical Study of the Newspapers*. In: *School and Society* III (1916) S. 140-44.

- Gornfeld, A.G.: *Chudožestvennoe slovo i naučnaja cifra*. In: *Literaturnaja mysl'* I (1922)
- Gray, W.S.: *Progress in the Study of Readability*. In: *Elementary School Journal* (1947) S. 491-99.
- und Bernice E. Leary: *What Makes a Book Readable*. Chicago 1935.
- Guiraud, Pierre: *A propos des caractères statistiques du vocabulaire et l'équation de Zipf*. In: *Bulletin de la Société de linguistique de Paris* LI (1955) S. 236-39.
- : *La Stylistique*. Paris 1954.
- : *Les caractères statistiques du vocabulaire*. Paris 1954.
- : *Problèmes et méthodes de la statistique linguistique*. Paris 1960.
- Hällberg, Peter: *Statistik i den litterära analysens tjänst*. In: *Litteraturvetenskap: Nya mål och metoder*. Stockholm: Natur och Kultur 1966. S. 9-36.
- Hausenblas, Karel: *K základním pojmům jazykové stylistiky*. In: *Slovo a slovesnost* XVI (1955)
- Havránek, Bohumil: *Studie o spisovném jazyce*. Praha 1963.
- Herdan, Gustav: *A New Derivation and Interpretation of Yule's Characteristic K*. In: *Journal of Applied Mathematics and Physics* VI (1955) S. 332.
- : *Language as Choice and Chance*. Groningen 1956.
- : *Quantitative Linguistics*. London 1964.
- : *The Advanced Theory of Language as Choice and Chance*. Berlin-Heidelberg-New York 1966.
- : *The Mathematical Analysis of Linguistics Behavior (thesis)*. London: Library of the Imperial College of Science 1940.
- : *The Relation between the Dictionary Distribution and the Occurrence Distribution of Word Length and its Importance for the Study of Quantitative Linguistics*. In: *Biometrika* 45 (1958) S. 222-28.
- : *Type-Token Mathematics: A Textbook of Mathematical Linguistics*. 's-Gravenhage 1960.
- Holstein, Arthur P.: *A Statistical Analysis of Schizophrenic Language: Preliminaries to a Study*. In: *Statistical Methods in Linguistics* IV (1965) S. 10-14.
- Horecký, Ján: *Úvod do matematickej jazykovedy*. Bratislava 1969.
- : *Pokus o štatistickú charakteristiku štýlu*. In: *Otázky žurnalistiky* 1963 (Martin 1964) S. 164.
- Jakobson, Roman: *Grammar of Poetry and Poetry of Grammar*. The Hague 1970.

- Janakiev, M.: *Zapiski po stilistika na b'lgarskija jezik*. Sofia 1965.
- Janson, Tore: *The Problems of Measuring Sentence-Length in Classical Texts*. In: *Studia Linguistica XVIII* (1964) 1 (1965) S. 26-36.
- Jelínek, J., J. Bečka und M. Těšitelová: *Frekvence slov, slovních druhů a tvarů v českém jazyce*. Praha 1961.
- Jelínek, Milan: *K charakteristice funkčního stylu uměleckého*. In: *Česká literatura* 15 (1967)
- Kaeding, F.W.: *Häufigkeitwörterbuch der Deutschen Sprache*. Berlin-Steglitz 1898.
- Kalinin, V.M.: *O statistike literaturnogo teksta*. In: *Voprosy jazykoznanija XIII* (1964) 1. S. 123-27.
- Kayser, Wolfgang: *Das sprachliche Kunstwerk*. Bern-München 1959.
- Klare, G.R.: *The Measurement of Readability*. The Iowa State: University Press 1963.
- Kniaginina, M.: *Próba zastosowania metod statystycznych w badaniach stylistycznych i składniowych*. In: *Język Polski XLII* (1962) S. 92-116.
- Kondratov, A.M.: *Matematika i poésija*. Moskva 1962.
- Kožina, M.N.: *O specifikie chudožestvennoj i naučnoj reči v aspekcie funkcional'noj stilistiki*. Perm 1966.
- Krallmann, Dieter: *Statistische Methoden in der stilistischen Textanalyse*. Bonn: Universitätsbibliothek 1966.
- Kraus, Jiří und Josef Polák: *Text Factors and Characteristics*. In: *Prague Studies in Mathematical Linguistics II* (1967) S. 155-71.
- Kraus, Jiří und Pavel Vašák: *Popytka količestvennoj tipologii tekstov*. In: *Prague Studies in Mathematical Linguistics II* (1967) S. 77-86.
- Kuraszkiewicz, W.: *La richesse du vocabulaire dans quelques grands textes polonais en vers*. Wrocław, Warszawa, Kraków 1963.
- Kurkowska, Halina und S. Skorupka: *Stylistyka polska*. Warszawa 1966.
- Kvaracchelija, G.Š.: *Statističeskij analiz stilja reči (na materiale gruzinskogo jazyka)*. Tbilisi 1968.
- Lauková, Júlia: *Exaktná rozbor románu M. Figuli. Tri gaštanové kone*. Bratislava: Filozofická fakulta University Komenského (Dizertačná práca) 1971.
- Leed, Jacob (Ed.): *The Computer and Literary Style: Introductory Essays and Studies*. Kent, Ohio 1966.
- Lesskis, G.A.: *O zavisimosti meždu razmerom predložénija i charakterom teksta*. In: *Voprosy jazykoznanija XIII* (1963) 3. S. 92-112.

- Lesskis, G.A.: *O zavisimosti meždu razmerom predloženiya i ego strukturoj v raznyh vidach teksta*. In: *Voprosy jazykoznanija* XIV (1964) 4. S. 99-123.
- Líšková, Zora: *K jazykovej morfológii titulku*. In: *Otázky žurnalistiky*. Bratislava 1968.
- Mandelbrot, Benoit: *An Information Theory of the Statistical Structure of Language*. In: *Communication Theory*. Butterworth 1953.
- Mayenowa, M.R. (Ed.): *Poetyka i matematyka*. Warszawa 1965.
- Meier, Helmut: *Deutsche Sprachstatistik*. Hildesheim 1964.
- Menzerath, P.: *Die Architektonik des Deutschen Wortschatzes*. In: *Phonetische Studien* 3 (1954)
- Miko, František: *Estetika výrazu*. Bratislava 1969.
- Mistrík, Jozef: *Dĺžka slova a štylistická štruktúra textu*. In: *Jazykovedný časopis* XVII (1966) 2. S. 114 ff.
- : *Exakte Methoden einer Typologie von Texten*. In: *Zeitschrift für slavische Philologie* 1971.
- : *Frekvencia slov v slovenčine*. Bratislava 1969.
- : *Kompozícia jazykového prejavu*. Bratislava 1968.
- : *Matematikostatističeské metody v stilistike*. In: *Voprosy jazykoznanija* XVI (1967) 3. S. 42-52.
- : *Štylistika slovenského jazyka*. Bratislava 1970.
- : *Z kvantitatívneho výskumu slovenskej ľudovej piesne*. In: *Slovenská reč* XXXII (1967) S. 328-35.
- Mlacek, Jozef: *O frekvencii jazykových prostriedkov a o jej štylistickej platnosti*. In: *Slovenská reč* (1965) S. 330-35.
- Montenegro, T.H.: *A análise matemática do estilo: a aplicação do processo de Fucks a obras literárias em português e espanhol*. In: *Coleção Teixeira de Freitas* 3 (Rio de Janeiro 1956)
- Moreau, R.: *Sur la distribution des formes verbales dans le français écrit*. In: *Etudes de linguistique appliquée* II (1963) S. 65-88.
- Mukařovský, Jan: *Kapitoly z české poetiky I-III*. Praha 1948.
- Muller, Charles: *Analyse statistique et analyse syntarique*. In: *Statistique et Analyse Linguistique*. Paris 1966. S. 123 f.
- : *Le Mot, unité de texte et unité de lexique en statistique lexicologique*. In: *Travaux de Linguistique et de Littérature*. Strasbourg 1963.
- Osgood, Charles E.: *Some effects of Motivation on Style of Encoding*. In: *Style in Language*. Cambridge-Mass.-New York 1960. S. 293-306.
- Pap, F.: *Količestvennyj analiz slovarnoj struktury russkich tekstov*. In: *Voprosy jazykoznanija* VI (1961) S. 93-100.

- Paterson, D.G. und M.A. Tinker: *How to Make Type Readable*. New York 1940.
- Pauliny, Eugen: *La phrase et l'énonciation*. In: *Recueil linguistique de Bratislava I* (1948)
- : *O funkčnom rozvrstvení spisovného jazyka*. In: *Slovo a slovesnost XVI* (1955)
- Perebijnis, V.S. u.a.: *Statistični parametri stiliv*. Kijev 1967.
- Pisarek, Walery: *Recepty na zrozumiałość wypowiedzi*. In: *Zeszyty prasoznawcze*. Nr. 2-3. Kraków 1966.
- Poetics*. Warszawa und 's-Gravenhage 1961.
- Poetics II*. Warszawa und The Hague 1966.
- Rickert, Edith: *New Methods for the Study of Literature*. Chicago 1927.
- Riffaterre, M.: *Criteria for Style Analysis*. In: *Word XV* (1959)
- Schlismann, Annemarie: *Die Sprachstilkurve: Eine graphische Darstellungsmethode im Dienst der Stilvergleichung*. In: *Innsbrucker Beiträge zur Kulturwissenschaft III* (1955) S. 51-58.
- : *Sprach- und Stilanalyse mit einem vereinfachten Quotienten*. In: *Wiener Zeitschrift für Philosophie, Psychologie und Pädagogik II* (1948)
- : *Wege exakter Stilanalyse*. In: *Schule und Psychologie*. 1956.
- Schröder, Hartwig: *Quantitative Stilanalyse: Versuch einer Analyse quantitativer Stilmerkmale unter psychologischem Aspekt*. Würzburg 1960.
- Sebeok, Thomas: *Style in Language*. Cambridge, Mass., New York und London 1960.
- Seidler, H.: *Allgemeine Stilistik*. Göttingen 1963.
- Sgall, P. u.a.: *Cesty moderní jazykovědy*. Praha 1964.
- Sherman, L.A.: *Analytics of Literature*. Boston 1893.
- : *Some Observations upon the Sentence-Length in English Prose*. In: *University of Nebraska Studies I* (1868) 2. S. 119-130.
- Skwarczyńska, Stefania: *Wstęp do nauki o literaturze I-III*. Warszawa 1954.
- Sławiński, Janusz: *Funkcje krytyki literackiej*. In: *2 teorii historii literatury*. Wrocław 1963.
- Somers, H.H.: *Analyse statistique du style: différences individuelles et facteurs psychologiques*. Louvain 1960.
- Spang-Hanssen, Henning: *Sentence Length and Statistical Linguistics*. In: *Structures and Quanta*. Copenhagen und New York 1963. S. 58-73.
- Spitzer, Leo: *Linguistics and Literary History: Essays in Stylistics*. Princeton 1948.

- Staiger, Emil: *Die Kunst der Interpretation*. Zürich 1957.
- Thavenius, Jan: *Kvantitativa metoder i stilistiken*. In: *Litteraturvetenskap: Nya mål och metoder*. Stockholm: Natur och Kultur 1966. S. 37-62.
- Thorndike, Edward L.: *The Teacher's Word Book*. New York 1921.
- Tomaševskij, B.V.: *Teorija literatury, poëtika*. Leningrad 1925.
- Trávníček, František: *O jazykovém slohu*. Praha 1953.
- Vinogradov, V.V.: *Stilistika, teorija poëtičeskoj reči, poëtika*. Moskva 1963.
- Wellek, R. und A. Warren: *Theory of Literature*. New York 1961.
- Wenstrand, Thomas E.: *An Analysis of Style: The Application of Sector Analysis to Examples of American Prose Fiction*. In: *Dissertation Abstracts XXVIII (1967) 1799 A*.
- Williams, C.B.: *A Note on the Statistical Analysis of Sentence Length as a Criterion of Literary Style*. In: *Biometrika* 31 (1940) S. 356-61.
- : *Statistics as an Aid to Literary Studies*. In: *Penguin Science News* 24 (1952) S. 99-106.
- Winter, Werner: *Relative Häufigkeit syntaktischer Erscheinungen als Mittel zur Abgrenzung von Stilarten*. In: *Phonetics VII* (1961) S. 193-216.
- Yule, George Udny: *On Sentence-Length as a Statistical Characteristic of Style in Prose, with Applications to Two Cases of Disputed Authorship*. In: *Biometrika* XXX (1936) S. 363-390.
- : *The Statistical Study of Literary Vocabulary*. Cambridge 1944.
- Zipf, G.K.: *Selected Studies of the Principle of Relative Frequencies in Language*. Cambridge, Mass. 1932.

ARBEITEN UND TEXTE ZUR SLAVISTIK
HERAUSGEGEBEN VON WOLFGANG KASACK

- 1 Sabine Appel: Jurij Oleša. "Zavist'" und "Zagovor čuvstv".
Ein Vergleich des Romans mit seiner dramatisierten
Fassung. 1973.
- 2 Renate Menge-Verbeeck: Nullsuffix und Nullsuffigierung
im Russischen. Zur Theorie der Wortbildung. 1973.
- 3 Jozef Mistrík: Exakte Typologie von Texten. 1973.
- 4 Andrea Hermann: Zum Deutschlandbild der nichtmarxistischen
russischen Sozialisten. Analyse der Zeitschrift
"Russkoe Bogatstvo" von 1880 bis 1904. 1974.

München · Verlag Otto Sagner in Kommission

Bayerische
Staatsbibliothek
München