

**VORTRÄGE
UND ABHANDLUNGEN
ZUR SLAVISTIK**

herausgegeben von P.Thiergen

Band 26

Sebastian Kempgen

RUSSISCHE SPRACHSTATISTIK

**SYSTEMATISCHER ÜBERBLICK
UND BIBLIOGRAPHIE**

1995

Verlag Otto Sagner * München

**VORTRÄGE
UND ABHANDLUNGEN
ZUR SLAVISTIK**

herausgegeben von P. Thiergen

Band 26

Sebastian Kempgen

RUSSISCHE SPRACHSTATISTIK

**SYSTEMATISCHER ÜBERBLICK
UND BIBLIOGRAPHIE**

1995

Verlag Otto Sagner * München

Vorträge und Abhandlungen zur Slavistik
Bd. 26

ISBN 3–87690–617–2

©

by Verlag Otto Sagner, München 1995.
Abteilung der Firma Kubon und Sagner,
Buchexport/import GmbH, München
Offsetdruck: Kurt Urlaub, Bamberg

© Zweite, elektronische Ausgabe
Sebastian Kempgen 2007

Vorträge und Abhandlungen
zur
Slavistik

herausgegeben von Peter Thiergen (Bamberg)

Band 26

1995

VERLAG OTTO SAGNER * MÜNCHEN

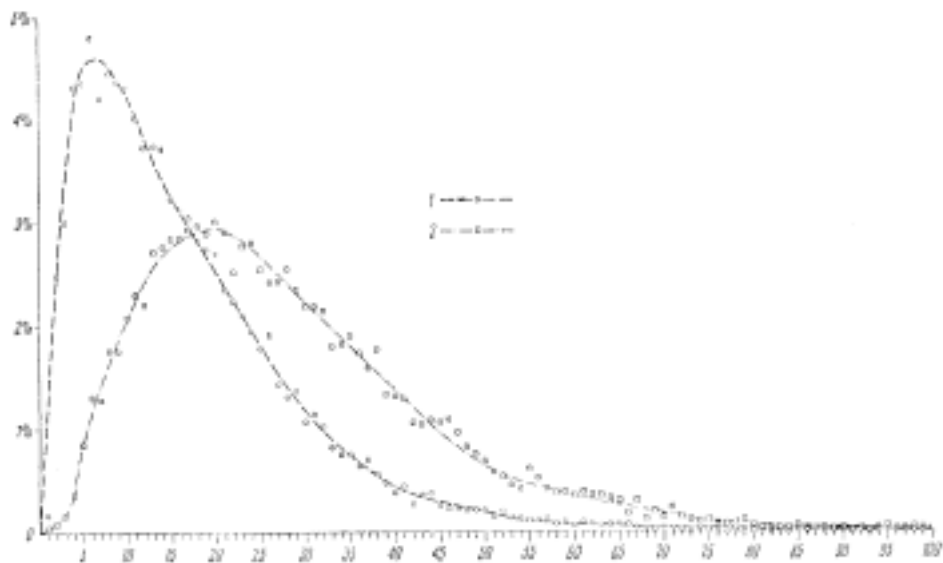
Sebastian Kempgen

Russische Sprachstatistik

Systematischer Überblick

und

Bibliographie



Vorwort

Erstaunlich viele Eigenschaften und Bereiche des Russischen sind bereits mit quantitativen Methoden untersucht und beschrieben worden, doch gab es bisher weder eine systematische Übersicht darüber, was bereits geleistet worden ist und in welchem Kontext diese Leistungen stehen, noch gab es eine umfassende bibliographische Erfassung einschlägiger Studien. Diese Lücke möchte die vorliegende Arbeit wenigstens teilweise schließen und demjenigen eine Hilfe sein, der gezielt nach Daten – oder nicht bearbeiteten Bereichen! – sucht oder sich überhaupt erst einmal in die sprachstatistische Erforschung des Russischen einlesen will. Eine "Einführung in die quantitative Linguistik für Russisten" will und kann die Arbeit selbstverständlich nicht sein.

Um einen möglichst großen Leserkreis anzusprechen, wurde die Darstellung bewußt nichtmathematisch gehalten, so daß in dieser Hinsicht für die Lektüre keinerlei Kenntnisse vorausgesetzt werden.

Mein Dank gilt an dieser Stelle dem Herausgeber der Reihe für seine Bereitschaft, die vorliegende Studie zu publizieren, ferner Kathrin Kraus, Ingrid Wagenknecht und Irina Werth für ihre Mithilfe bei der Materialbeschaffung, bei der Endredaktion des Textes sowie beim Erstellen der Bibliographie.

Für die vorliegende zweite, elektronische Auflage des Buchleins, das bei seinen Rezensenten eine erfreulich positive Aufnahme gefunden hat, wurden einige Tippfehler korrigiert.

Bamberg, im Okt. 1995 bzw. Dez. 2007 *Sebastian Kempgen*

Inhaltsverzeichnis

1.	Einleitung	7
1.1.	Zur Begründung der Quantitativen Linguistik (QL).....	7
1.2.	QL und Mathematische Linguistik	8
1.3.	QL und Grammatikmodell.....	9
1.4.	Aufbau und Ziel der vorliegenden Arbeit.....	10
2.	Zur Geschichte der Sprachstatistik in Rußland	13
3.	Synchrone Untersuchungen zum Russischen	17
3.1.	Graphematik	17
3.1.1.	Entzifferung.....	17
3.1.2.	Effektivität	17
3.1.3.	Interne Graphemstruktur	18
3.1.4.	Graphemfrequenzen.....	19
3.1.5.	Informationstheorie.....	22
3.2.	Phonologie	24
3.2.1.	Distinktive Merkmale.....	25
3.2.2.	Phonemzahl.....	25
3.2.3.	Phonemfrequenz.....	25
3.2.4.	Phonemkombinatorik.....	26
3.2.5.	Silbenstruktur, Silbenfrequenz	27
3.3.	Morphologie.....	29
3.3.1.	Morphem und Wort	29
3.3.2.	Flexionsmorphologie	36
3.3.3.	Derivationsmorphologie.....	44
3.4.	Lexikologie	46
3.4.1.	Allgemeine Frequenzwörterbücher	46
3.4.2.	Grundwortschatz und Lehrbücher.....	51
3.4.3.	Fachsprachliche Frequenzwörterbücher.....	52
3.4.4.	Mehrsprachige Häufigkeitwörterbücher	54
3.4.5.	Einzelne Werke bzw. Autoren.....	54
3.4.6.	Zusammenfassung	56
3.4.7.	Weitere Aspekte.....	58
3.5.	Syntax.....	58
3.5.1.	Syntaktische Relationen und Phrasenstruktur.....	59
3.5.2.	Eigenschaften ganzer Sätze	61

3.6.	Textlinguistik, Type-Token-Ratio	66
3.7.	Lexikalische Semantik, Onomasiologie.....	67
3.8.	Soziolinguistik, Sprache und Norm	72
3.9.	Psycholinguistik.....	75
3.10.	Dialektologie.....	75
3.11.	Statistik und Stil, Poetische Sprache.....	76
3.12.	Sprachtypologie und Universalien	80
4.	Diachrone Untersuchungen zum Russischen	83
4.1.	Grapheme	83
4.2.	Phoneme.....	83
4.3.	Morphologie.....	85
4.4.	Lexematik.....	88
4.5.	Syntax.....	92
4.6.	Semantik, Onomastik.....	93
4.7.	Dialektologie.....	93
5.	Bibliographien und Lehrbücher	95
6.	Bibliographie	97

1. Einleitung

1.1. Zur Begründung der Quantitativen Linguistik (QL)

Als Teil des menschlichen Verhaltens ist die Sprache eine Massenerscheinung. Wie die Soziologie und die Psychologie kann deshalb auch die Sprachwissenschaft empirische Methoden verwenden, um ihren Gegenstandsbereich zu untersuchen. Der QL (russ: *kvantitativnaja lingvistika, vyčislitel'naja lingvistika, lingvostatistika, statistika reči*) geht es dabei um eine Beschreibung wie um eine Erklärung der Daten. Hierzu leitet sie idealerweise Hypothesen über Zusammenhänge aus allgemeinen sprachtheoretischen Annahmen ab, sammelt empirische Daten, die zur Überprüfung der Hypothese geeignet sind, unterwirft sie statistischen Testverfahren, um die Signifikanz der Daten abzuschätzen, und entscheidet dann über die sprachwissenschaftliche Ausgangshypothese. Die genannten Schritte zeigen zugleich, daß am Anfang immer eine linguistische Fragestellung steht und am Ende eine Entscheidung über das untersuchte sprachwissenschaftliche Problem. Die Anwendung quantitativer Methoden ist der empirischen Arbeit nur Mittel zum Ziel, nicht Selbstzweck.

Zu einer höheren Stufe der Forschung gehört bereits das Modellieren eines sprachlichen Bereiches. Dies kann die Herleitung einer bestimmten Verteilung, mit der eine einzelne Variable beschrieben wird, sein. Beispielsweise läßt sich die Zahl der Phoneme in den Sprachen der Welt mit einer Gleichung theoretisch darstellen. Dies kann aber auch die Herleitung oder das Aufstellen von Sprachgesetzen bedeuten: Das sog. "Menzerathsche Gesetz" setzt z.B. die Länge einer sprachlichen Einheit in Abhängigkeit von ihren Konstituenten: je länger ein Wort ist, desto kürzer sind seine Silben. Dies mündet auf der höchsten Stufe schließlich in der Ausarbeitung einer Sprachtheorie oder mindestens von Teilfragmenten. Sie ist dadurch gekennzeichnet, daß hier nicht mehr einzelne Zusammenhänge isoliert nebeneinander stehen, sondern zueinander in Beziehung gesetzt werden, um komplexe Systeme abzubilden. Beispielsweise "reagiert" eine Sprache auf Störungen des Gleichgewichtes, auf Kommunikationsanforderungen etc., so daß eine Veränderung immer andere Veränderungen (gegebenenfalls Gegenreaktionen) nach sich zieht. Die Linguistik profitiert hier von den theoretischen Erkenntnissen anderer Disziplinen wie Systemtheorie, Synergetik u.a.

Die Notwendigkeit der Anwendung quantitativer Methoden kann auch auf andere Weise überzeugend hergeleitet werden: 1) Gälten alle Regeln einer Sprache hundertprozentig, so hätten sie einen absolut deterministischen Charakter, es gäbe weder Variation noch Veränderung. 2) Sprache variiert und verändert sich aber. 3) Daraus folgt: Sprachliche Regeln können nicht nur deterministischen Charakter haben. Heute geht man sogar davon aus, daß die ausnahmslosen Regeln sozusagen nur die Spitze des Eisberges sind, während der weitaus größere Teil der Sprache durch stochastische Tendenzen geprägt ist, die freilich nicht in gleicher Weise und mit den gleichen Mitteln wie die an der Oberfläche liegenden beobachtet und erfaßt

werden können. Ausnahmslos gültige Regeln sind in dieser Sicht nichts weiter als die Extremfälle auf der Wahrscheinlichkeitsskala, deren weitaus größter Bereich durch eine mehr oder weniger starke Tendenz gekennzeichnet ist.

Die erkenntnistheoretische Grundlage für die QL findet sich in den Aussagen moderner Wissenschaftstheoretiker wie CARNAP, HEMPEL u.a.: Sie betonen, daß der Unterschied zwischen Qualität und Quantität ein Unterschied in unserem Denken ist, ein Unterschied, den wir erst mit unseren Begriffen machen, nicht aber ein Unterschied in der Natur: "Wir können nicht in die Natur hinausschauen und fragen: »Sind die Erscheinungen, die ich hier sehe, quantitativ oder qualitativ?« Das ist nicht die richtige Frage. Wir können nur, wenn jemand diese Erscheinungen mit gewissen Ausdrücken beschreibt, diese Ausdrücke definiert und uns Regeln für ihren Gebrauch gibt, fragen: »Sind dies Ausdrücke einer quantitativen Sprache oder Ausdrücke einer vorquantitativen, qualitativen Sprache?«" (CARNAP 1974, 67).

Zu den allgemeinen Grundlagen der QL vgl. ALTMANN (1972a, 1972b, 1973, 1977, 1993), HOFFMANN (1975b); zur sowjetischen Sicht programmatisch TOPOROV (1959), FRUMKINA (1960b, 1961a, 1976), JAGLOM/DOBRUŠIN/JAGLOM (1960), PIOTROVSKIJ (1961), ANDREEV/ZINDER (1963), referierend PAPP (1966, 62ff.), als "Klassiker" seien genannt ACHMANOVA et al. (1961).

1.2. QL und Mathematische Linguistik

Die QL ist neben der algebraischen Linguistik, die mit Logik, Kalkülen und Mengentheorie arbeitet, Teil der mathematischen Linguistik. Beide Bereiche ergänzen einander, schließen einander aber nicht aus. Den algebraischen Teil der mathematischen Linguistik verbindet dabei mit der nichtmathematischen Linguistik jedoch die Gemeinsamkeit, daß sie mit ausnahmslosen Regeln arbeitet. ("Unscharfe Mengen", wie sie heute teilweise auch in der Mengentheorie verwendet werden, sind im Grunde nur ein "Reparaturversuch" an dieser fundamentalen Beschränkung.) In einem etwas ungenauen, jedoch üblichen Sprachgebrauch heißt die quantitative mathematische Linguistik einfach quantitative Linguistik, die algebraische mathematische Linguistik einfach mathematische Linguistik, was eigentlich nicht korrekt ist.

Die QL scheint sich somit in erster Linie über ihre Methoden zu definieren, nicht über den Gegenstand, den sie untersucht: Im Prinzip können alle bekannten Bereiche der Sprache, von den Allophenen bis zur Textstruktur und dem Lexikon, statistisch untersucht werden. Die Definition über spezielle Methoden ist jedoch nur auf den ersten Blick richtig. Tatsächlich veranlassen die quantitativen Methoden den Forscher auch dazu, neue Fragestellungen und Probleme aufzuwerfen, die man mit nichtquantitativen Methoden gar nicht gesehen hätte. So erweitert die QL nicht nur das Methodenspektrum der Linguistik, sondern auch ihr Erkenntnisinteresse, ihre Fragestellungen und damit ihren Gegenstandsbereich.

Gegenstand des vorliegenden Beitrages ist die quantitative (russistische) Linguistik, nicht die algebraische. Nur der Vollständigkeit halber sei deshalb der folgende Titel genannt: Das Buch von GLADKIJ/MEL'ČUK (1969) ist zu einem Standardwerk der algebraischen Linguistik geworden (auch in deutscher Übersetzung 1973). Von den nicht sehr zahlreichen konkreten Untersuchungen zum Russischen sei wenigstens auf LOMTEV (1972) verwiesen, der eine mengentheoretische Beschreibung der russischen Phonemebene vorgelegt hat. Programmatisch äußert sich MAČAVARIANI (1963).

Die QL als Disziplin hat von Autoren aus dem slawisch-sprachigen Raum wesentliche Impulse erfahren; neben Rußland seien hier besonders die Traditionen der QL in Polen und in der Tschechoslowakei erwähnt – siehe die Reihe "Prague Studies in Mathematical Linguistics", die separate Sektionen für algebraische und für statistische Arbeiten aufweist, und die Bände der in Deutschland herausgegebenen Reihe "Quantitative Linguistics" (Bd. 1–51: Bochum 1978–92, Bd. 52ff. Trier 1993ff.), die für viele osteuropäische Autoren eine Vermittlerfunktion ausübt und auch eine "Current Bibliography" fortschreibt. Auch deshalb mag es gerechtfertigt sein, die russische (bzw. russistische) QL in der vorliegenden Weise darzustellen.

1.3. QL und Grammatikmodell

Die Anwendung quantitativer Methoden ist an sich nicht an ein bestimmtes Grammatikmodell gekoppelt, sei es die traditionelle Grammatik, der Strukturalismus oder die Generative Grammatik. Es läßt sich jedoch nicht leugnen, daß sich die strukturalistische Sprachauffassung, die Einheiten und deren Relationen zueinander sowie zur nächsthöheren bzw. nächstniedereren Ebene und ihren Einheiten ermittelt und zugrundelegt, für quantitative Untersuchungen besonders anbietet und deshalb implizit oder explizit in den meisten Arbeiten verwendet wird (zu den wenigen Autoren, die dieses Problem ausdrücklich ansprechen, gehört KRYSIN). Es läßt sich ferner nicht wegdiskutieren, daß in nicht wenigen sprachstatistischen Arbeiten die linguistische Fundierung des eigenen Tuns ungenügend bedacht wird.

Gezählt und ausgewertet werden jedenfalls vornehmlich Phoneme, Silben, Morpheme, Wörter (Wortformen, Lexeme), Syntagmen, Sätze und Texte, d.h. Einheiten der Ausdrucksseite, und zwar mit synchroner wie diachroner Blickrichtung. Dabei kann man grundsätzlich zwischen einer *pragmatischen* Stichprobe, also Daten, die aus Texten gewonnen werden, und einer *systemischen* Stichprobe, das sind Daten, die aus einem Lexikon gewonnen werden, unterscheiden. Im ersten Fall spielen ganz offensichtlich die Gebrauchshäufigkeiten der einzelnen Wörter eine entscheidende Rolle – aus dem gesamten Wortschatz der Sprache wird ja bei Textuntersuchungen immer nur ein spezieller Ausschnitt angetroffen, und zwar in Form von *tokens*. Im zweiten Fall dagegen stehen der Untersuchung potentiell alle Lexeme der Sprache zur Verfügung (*types*) – entweder für eine vollständige Auswertung oder für eine Stichprobe; die Perspektive ist hier auf die Sprachstruktur als solche, das Sy-

stem, gerichtet. Beide Vorgehensweisen haben ihre Berechtigung und ergänzen sich wechselseitig.

Auf einer metalinguistischen Ebene sind die Vorschläge von MEL'ČUK (1963) angesiedelt, der verschiedene Indizes vorschlägt (z.B. Kompaktheit und Einfachheit), wie linguistische Modelle bewertet und verglichen werden können.

1.4. Aufbau und Ziel der vorliegenden Arbeit

Gegenstand der vorliegenden Darstellung sind solche Arbeiten zum Russischen, die sich ihrem Gegenstand mit quantitativen Methoden nähern, die statistische Daten ermitteln und präsentieren. Ziel der Übersicht kann dabei keine Vollständigkeit in dem Sinne sein, daß jede russistische Arbeit, in der sich irgendeine Zahlenangabe oder Prozentzahl findet, berücksichtigt wird. Vielmehr geht es darum, die wichtigeren Studien zu den verschiedenen sprachlichen Ebenen vorzustellen, von denen manche allgemeint bekannt, viele hingegen nur dem Spezialisten geläufig sind.

Gegenstand der Darstellung sind dabei sprachstatistische und quantitative Arbeiten im eigentlichen Sinne; der Bereich der Automatischen Übersetzung, mit dem es naheliegende Überschneidungen gibt, wird im folgenden bewußt ausgeklammert. Ausgespart wird auch der große Bereich der quantitativen linguistischen Forschung in Rußland, soweit er sich nicht mit dem Russischen beschäftigt. Sofern sich in Arbeiten, die sich primär mit den Grundlagen und Verfahren der QL beschäftigen, in Form von Beispielen gelegentlich Daten zum Russischen finden, werden diese in der Regel nur berücksichtigt, wenn sie von systematischem Interesse auch außerhalb des betreffenden Kontextes sind. Berücksichtigt werden bei unserer Darstellung alle Ebenen der linguistischen Forschung; mit dieser Wortwahl soll bewußt die akustische Phonetik ausgeklammert werden, die natürlich ebenfalls mißt und zählt (z.B. Tonhöhen und Tondauer), aber nicht zu den sprachwissenschaftlich-philologischen Bereichen im engeren Sinne gehört, da ihr Gegenstand ja nicht linguistisch definierte Einheiten sind. Als Beispiel einer solchen Arbeit sei BONDARKO et al. (1986) oder IGNATOVA (1969) genannt.

Im Sinne der oben genannten Stufen der quantitativen Forschung geht es ferner nicht um eine Darstellung der im einzelnen verwendeten statistischen Methoden als solche, sondern darum, die empirischen Daten zu charakterisieren, die sprachwissenschaftlichen Fragen und, wo vorhanden, möglich und sinnvoll, die erzielten Ergebnisse in knappster Form darzustellen und dabei auf Probleme, offene Punkte und Defizite hinzuweisen, wodurch gleichzeitig auch deutlich werden kann, welcher Raum sich der zukünftigen Forschung noch bietet. Der Beitrag will also über bereits erhobene Daten informieren, über die Fragen, die schon quantitativ bearbeitet sind, Auskunft geben, will in diesem Sinne demjenigen eine Orientierungshilfe bieten, der nach bestimmten Daten oder Antworten sucht. Aufgabe dieses Überblickes kann es dabei natürlich nicht sein, auf alle Zahlen die erforderli-

chen statistischen Tests anzuwenden und so eventuelle Versäumnisse der referierten Literatur auszugleichen. Ausgeschlossen ist es selbstverständlich auch, sämtliche relevanten Daten und Ergebnisse in die vorliegende Übersicht aufzunehmen. Wo Zahlen und statistische Angaben verwendet wurden, galt das Bemühen, Vergleiche anzustellen und die Resultate zu kommentieren, also nicht nur bereits vorhandenes Material wiederzugeben. Die Darstellung ist ferner so angelegt, daß keinerlei Kenntnisse der QL oder der Statistik vorausgesetzt werden; sie wendet sich also nicht so sehr an die Fachleute des Gebietes, sondern an jeden empirisch arbeitenden Russisten; diesem Ziel entsprechend ist die Darstellung vollkommen "unmathematisch" gehalten, verzichtet also weitestgehend auf Formeln.

Bei der Frage, wie diese Übersicht angelegt werden sollte, kam im wesentlichen zwei Prinzipien in Betracht (wenn wir eine biographische Behandlung des Themas einmal ausschließen): eine Orientierung an dem klassischen strukturalistischen Schichtenmodell der Sprache samt den Beziehungen zwischen den verschiedenen Ebenen und Elementen und eine Orientierung an den Publikationen (die ja z.T. zu mehr als einer Ebene Daten liefern). Wir haben uns dafür entschieden, der Darstellung im Prinzip eine sprachbezogene Systematik zugrunde zu legen, da dies die praktische Benutzung wesentlich erleichtert. So finden sich unter der "Phonologie" konsequent die Arbeiten, in denen phonologische Fragen behandelt werden. Wenn eine Publikation jedoch für mehrere Ebenen einschlägig ist, so wurde versucht, die Behandlung dieses Werkes nicht zu sehr zu stückeln, da sonst viele Wiederholungen notwendig wären.

Für die abschließende Bibliographie galt das Bemühen, die einschlägigen Arbeiten zum Russischen möglichst umfassend zusammenzustellen; daß sie dennoch keine Vollständigkeit beanspruchen kann, dürfte sich von selbst verstehen, zumal es an der Peripherie des Gebietes keine klaren Kriterien gibt, was aufzunehmen wäre und was nicht. Entlegenere Literatur konnte gelegentlich leider nicht im Original eingesehen werden; in solchen Fällen sind u.U. auch die bibliographischen Angaben, etwa in bezug auf die Seitenzahlen (bei Aufsätzen), unvollständig. Hinweise auf weitere Literatur und zur Vervollständigung schon vorliegender Einträge nimmt der Autor gerne entgegen.

2. Zur Geschichte der Sprachstatistik in Rußland

Im Rückblick wird deutlich, daß eine erste Blütezeit statistischer Sprachuntersuchungen in der Sowjetunion in den sechziger Jahren – beginnend Ende der fünfziger Jahre und endend mit Ausläufern in den siebziger Jahren – liegt. Diese Zeit ist die Grundlage für die gegenwärtige Forschung, und die während dieser Zeit erschienenen Arbeiten bilden auch den Schwerpunkt unserer systematischen Darstellung. Einen fundierten und kenntnisreichen Abriß der Wissenschaftsgeschichte bis zu diesem Zeitpunkt gibt PAPP (1966), auf den wir uns auch in einzelnen Ausführungen stützen.

Eine vereinzelt Anwendung statistischer Methoden hat es in Rußland bereits relativ früh gegeben (vgl. STRUVE 1918). So findet sich bei BUDILOVIČ (1883) bereits eine Graphemstatistik, zehn Jahre später bei DIKAREV (1891, 9) eine an (regionalen) Texten gewonnene Lautstatistik. Sind diese Autoren methodisch noch relativ unbedarft, so gilt dies nicht für KUDRJAVSKIJ (1909, 1911, 1912), der mustergültige sprachgeschichtliche Analysen zum altrussischen Verbsystem (Entwicklung des Tempusgebrauches) unternommen hat und dabei die Anwendung statistischer Methoden propagierte. MOROZOV (1915) geht davon aus, daß alle sprachlichen Elemente konstante Proportionen in ihrem jeweiligen Bezugsrahmen aufweisen, was er u.a. für den Autorschaftsnachweis nutzen will.

Begünstigend für die Entwicklung der statistischen Linguistik war aber auch die Tatsache, daß es in Rußland eine große Tradition an herausragenden Mathematikern und Statistikern gegeben hat, die nicht selten auch an sprachwissenschaftlichen Problemen Interesse zeigten. Erinnert sei nur an die sogenannten MARKOV-Ketten, die der Mathematiker A.A. Markov 1913 an den ersten 20.000 Buchstaben des "Evgenij Onegin" demonstrierte (MARKOV 1913, 1916, 1924), oder an die Tatsache, daß noch heute klassische statistische Tests international die Namen ihrer russischen Urheber tragen (wie der Kolmogorov-Smirnov-Test). MARKOV wurde mit seinen Arbeiten einer der Begründer der modernen Informationstheorie, deren Ausformulierung in dem klassischen Buch von SHANNON/WEAVER (1949) erfolgt.

In sowjetischer Zeit finden wir zunächst PEŠKOVSKIJ (1925) mit der ersten größeren Lautstatistik, JAKOVLEV (1928) mit seiner "mathematischen Formel" für die Konstruktion einer Phonemschrift sowie ČISTJAKOV/KRAMARENKO (1928) mit einem bemerkenswerten Büchlein, in dem die Anwendung statistischer Methoden begründet und am Russischen vergleichend demonstriert wird, ähnlich POLIVANOV (1931). Orthographiereform und Alphabetisierungskampagne bedingen eine Auseinandersetzung mit typographischen und polygraphischen Graphemstatistiken (PROSKURNIN 1933). PETERSON (1928, 1937) untersucht syntaktische Phänomene mit großer Präzision an älteren Texten.

Die weitere Entwicklung der Sprachstatistik in der Sowjetunion ist danach unmittelbar mit der Herausbildung der sprachwissenschaftlichen Schulen und ihres Schicksals in der Zeit des Marrismus unter Stalin verknüpft. Die "Moskauer" oder "formale Schule" FORTUNATOVs mit ihrer Konzentration auf die Ausdrucksseite bewies fast naturgemäß eine besonders große Affinität zur Anwendung exakter Methoden (FORTUNATOV selbst beschäftigte sich – nach PAPP 1966 – gerne und lange mit höherer Mathematik); ihre Methoden und Anhänger waren entsprechend einer besonderen Anfeindung ausgesetzt. Auch bei BAUDOIN de COURTENAY, dem Begründer der Kazaner (Petersburger) Schule und seinen Nachfolgern (v.a. BOGORODICKIJ) finden sich bereits vereinzelt Aussagen zur Nützlichkeit sprachstatistischer Untersuchungen und erste konkrete Untersuchungen (ERMOLAEV 1915), doch hatte diese Schule als solche weniger unter dem Marrismus zu leiden, was sie wohl besonders der Persönlichkeit ihres Hauptvertreters ŠČERBA verdankte (so PAPP 1966, 29ff.). Diese Schule setzte sich dann in den 50er Jahren in der Leningrader Gruppe "Statistika reči" fort (s.u.). Das wissenschaftliche Klima in der Sowjetunion änderte sich bekanntlich erst in den 50er Jahren wieder, als Stalin die linguistische Diskussion freigab. Mit der sog. "Strukturalismus-Diskussion" konnte sich auch die formale Moskauer Schule wieder zu Wort melden. Die ab den sechziger Jahren mit wichtigen Publikationen im Bereich der Sprachstatistik hervorgetretenen Autoren lassen sich entweder der Moskauer oder der Leningrader Schule zuordnen.

Für das Aufblühen der statistischen Linguistik in den 50er Jahren ist aber ein ganz anderes historisches Ereignis ursächlich gewesen: Der Zweite Weltkrieg. Er – und der in seiner Folge einsetzende Kalte Krieg – führten im Westen wie im Osten zu einem "Technologieschub", dessen Anfänge literarisch eindrucksvoll in Solženičyns "Im ersten Kreis" verarbeitet werden, wo Sprachwissenschaftler und Physiker gemeinsam an einem "Sprachzerhacker" für Stalin arbeiten. Neben Vorrichtungen zur Sprachverschlüsselung, -übertragung und -decodierung ist in diesem Zusammenhang natürlich besonders an die Erfindung und Entwicklung von Computern nach dem Zweiten Weltkrieg, zuerst in Gestalt von Großrechnern, zu erinnern, ohne die viele sprachstatistische Untersuchungen gar nicht möglich gewesen wären.

Viele sowjetische Arbeiten zur QL entsprangen einem Interesse an der durch die grundlegenden Arbeiten von C.E. SHANNON im Gefolge des Zweiten Weltkrieges entstandenen Informationstheorie (SHANNONs Arbeiten erscheinen 1963 in einer russischen Übersetzung in Moskau, wie viele klassische und neuere westliche Werke zur Statistik auch); auch die eher philologischen Arbeiten von G.K. ZIPF (1949) und sein "Prinzip der geringsten Anstrengung" wurden rezipiert.

Die Beschäftigung mit der Automatischen Übersetzung, die um die gleiche Zeit ihren ersten Höhepunkt erlebte, förderte ebenfalls statistische Sprachuntersuchungen, die als Grundlage bzw. vorbereitende Analysen benötigt wurden. Auch die Automatische Übersetzung hatte im übrigen mindestens einen Teil ihrer Wurzeln in militärischer Anwendung – auffällig ist z.B., daß die Sprache der Radioelektronik in der sowjetischen QL in den Anfangsjahren besonders gut untersucht wird.

Bereits 1956 (zum folgenden vgl. PAPP 1966, 44ff.) wurde in Moskau am Institut für Fremdsprachen eine Arbeitsgemeinschaft zur Maschinellen Übersetzung (unter der Leitung von ROZENCVEJG) ins Leben gerufen, in der sich auch I.I. REVZIN engagierte. An der MGU begannen erste Aktivitäten zeitgleich; sie führten 1960 zur Einrichtung einer Abteilung für Mathematische Linguistik unter Leitung von ZVEGINCEV. Hier waren u.a. P.S. KUZNECOV, V.V. IVANOV, V.A. USPENSKIJ, S.K. ŠAUMJAN, A.A. ZALIZNJAK und O.S. KULAGINA tätig. Ebenfalls seit 1956 waren Linguisten der Leningrader Universität unter Leitung von N.D. ANDREEV mit Problemen der maschinellen Übersetzung beschäftigt. Die mathematische Linguistik betreute dort ab 1958 L.R. ZINDER. 1959 wurde die Gruppe "Statistika reči" gebildet, deren "Motoren" R.G. PIOTROVSKIJ und P.M. ALEKSEEV waren (s.a. USPENSKIJ 1958, IVANOV 1958, ANDREEV 1960, PIOTROVSKIJ 1969a). Im Vordergrund stehen in dieser Gruppe die Untersuchung der Entropie, die Phonem- und Graphemstatistik sowie Untersuchungen von Texteigenschaften und die Automatische Übersetzung. Einen weiteren Schwerpunkt bilden alle Arten von Frequenzwörterbüchern, die Prinzipien ihrer Erstellung und die Kriterien ihrer Bewertung sowie eine statistische Modellierung der Flexionsmorphologie.

Relativ spät erst zog die Akademie der Wissenschaften nach. Hier wurde im Institut für Sprachwissenschaft 1959 eine Gruppe für Angewandte Linguistik eingerichtet, der neben R.R. REFORMATSKIJ zunächst nur I.A. MEL'ČUK angehörte, bevor dann L.N. IORDANSKAJA, R.M. FRUMKINA und einige andere hinzukamen. Am Institut für die Russische Sprache leitete S.K. ŠAUMJAN eine ähnliche Abteilung, zu der auch Ju.D. APRESJAN gehörte. In der Abteilung für Strukturelle Typologie des Institutes für Slawistik schließlich versammelte sich die größte Gruppe, zu der bedeutende Linguisten wie V.N. TOPOROV, V.V. IVANOV, I.I. REVZIN, T.N. MOLOŠNAJA, T.M. NIKOLAEVA, Z.M. VOLOCKAJA, Ju.K. LEKOMCEV, M.I. LEKOMCEVA und A.A. ZALIZNJAK u.a. gehör(t)en. Eine detaillierte Aufzählung weiterer Einzelereignisse, die in der Summe diese Neuorientierung der sowjetischen Linguistik bewirkten, findet sich bei PAPP (1964; 1966, 52ff.), zu Tagungsberichten vgl. GRIGOR'EV (1960) und ŠIROKOV (1961), zu einem allgemeineren Überblick vgl. ABERNATHY (1963b).

Nicht unerwähnt darf auch bleiben, daß die sowjetische Linguistik mit den 1952 gegründeten "Voprosy jazykoznanija" ihr wichtigstes Publikationsorgan bekam, das sich sprachstatistischen Arbeiten gegenüber eine Reihe von Jahren als ausgesprochen aufgeschlossen erwies. Jedenfalls ist eine ganze Reihe von heute klassischen Arbeiten in der ersten Hälfte der sechziger Jahre in dieser Zeitschrift erschienen – Ende der sechziger Jahre verschwand die Rubrik "Prikladnoe i matematičeskoe jazykoznanie" allerdings wieder aus der Zeitschrift (letztmalig findet sich die Rubrik in Heft 2/1969). Danach sank auch die Zahl der einschlägigen Arbeiten in dieser Zeitschrift beträchtlich. Als weitere Publikationsorgane speziellerer Ausrichtung kamen etwa "Mašinnyj perevod i prikladnaja lingvistika" (seit 1957) und die "Problemy kibernetiki" (seit 1958), die von den Amerikanern komplett ins Englische übersetzt wurden, hinzu.

Anerkennend sei an dieser Stelle auch vermerkt, daß es in der deutschen Slawistik besonders in der DDR Bemühungen um die QL gegeben hat, sei es durch das Erstellen spezieller Fachwortschatz-Häufigkeitswörterbücher (HOFFMANN 1970ff.), Übersetzungen (ALEXEJEW et al. 1973) oder durch gemeinsame Publikationen (HOFFMANN/PIOTROWSKI 1979).

Im weiteren slawischen und slawistischen Kontext sei im übrigen insbesondere auf Polen (KAMIŃSKA 1978, SAMBOR 1972) und auf die Prager Schule verwiesen, die in den "Prague Studies in Mathematical Linguistics" (PSML 1966ff.) einen bedeutenden Ausläufer gefunden hat, außerdem auf die zahlreichen Arbeiten von M. TĚŠITELOVÁ zum Tschechischen (1980, 1987, 1992 mit ausführlicher Bibliographie), schließlich außerdem auf die Tätigkeit von Gabriel ALTMANN, der, aus Bratislava stammend, in der Bundesrepublik die Reihe "Quantitative Linguistics" (Bochum 1978ff.) ins Leben gerufen hat. Diese Reihe hat für Autoren aus dem slawisch-sprachigen Raum bzw. dem ehemaligen Ostblock eine nicht zu unterschätzende Mittlerfunktion ausgeübt. Ohne Übertreibung läßt sich sagen, daß durch ALTMANNs Tätigkeit Deutschland zu einem der Zentren der QL geworden ist, weil hier inzwischen die organisatorische und editorische Verantwortung für einen Großteil der internationalen Forschung liegt (der "Erste Internationale Kongress zur Quantitativen Linguistik" [QUALICO] fand 1991 in Trier statt (s. KÖHLER/RIEGER [eds.] 1991), der zweite 1994 in Moskau; seit 1994 erscheint unter deutscher Leitung das "Journal of Quantitative Linguistics" als erste Zeitschrift überhaupt, die der QL gewidmet ist).

An einem Punkt sieht man im übrigen deutlich, wie schnell in den letzten Jahrzehnten bzw. Jahren die Entwicklung der Technik verlaufen ist: ältere Arbeiten zur Sprachstatistik behandeln oft den Einsatz von Computern für solche Fragestellungen und widmen sich nicht selten Algorithmen und Flußdiagrammen, die darstellen, wie bestimmte Fragen logisch zu strukturieren und dann zu programmieren seien (z.B. die Erstellung einer Häufigkeitsliste). Solche auf Großrechner gemünzte Problembehandlungen sind durch die massenhafte Verbreitung von Arbeitsplatzrechnern mit entsprechenden Statistik-Programmen mittlerweile natürlich längst überholt und nur noch Zeugnis der "Computer-Steinzeit".

Klassische sprachstatistische Arbeiten: FRUMKINA (1960b, 1964a), Sb. ĚNTROPIJA JAZYKA I STATISTIKA REČI (1966), Sb. STATISTIKA REČI (1968), Sb. STATISTIKA TEKSTA (1969), ŠTEJNFEL'DT (1963), PIOTROVSKAJA et al. (1962), ANDREEV (1965, 1967), ANDREEVA (1969), JAGLOM et al. (1960), ACHMANOVA et al. (1961). Referierend, mit etwas weiterem Spektrum, PAPP (1964, 1966), TĚŠITELOVÁ (1991).

3. Synchronische Untersuchungen zum Russischen

3.1. Graphematik

Der Bereich der Graphematik gehört seit Anfang an zu den empirisch gut bearbeiteten Bereichen des Russischen, da gedruckte Buchstaben klar segmentierbare und deshalb relativ unproblematische Einheiten sind, einer Zählung also entgegenkommen. Weniger gut bearbeitet ist in der Russistik dagegen die Frage, was man eigentlich mit entsprechenden Daten anfangen kann, wie sie auf Signifikanz zu testen sind, auf welche inhaltlichen Fragen sie eine Antwort geben usw. Hier bieten jetzt ALTMANN/LEHFELDT (1980, 112ff.) eine umfassende theoretische Behandlung des Themas.

3.1.1. Entzifferung

Die Grapheme des Russischen müssen natürlich im eigentlichen Sinne nicht "entziffert" werden. Trotzdem sei an dieser Stelle auf die Arbeiten von SUCHOTIN (1962, 1963) hingewiesen, der einen Algorithmus entwickelt hat, wie durch eine reine Frequenz- und Distributionsanalyse die Buchstaben einer (unbekannten) Alphabetschrift mindestens nach Vokalen und Konsonanten getrennt werden können, ein genügend großes Korpus vorausgesetzt. Der Algorithmus wird vom Autor am Russischen, Englischen und Französischen demonstriert; zu jeder Sprache findet sich entsprechend eine Tabelle mit der Frequenz aller Zweierkombinationen. Als Ergebnis liefert das Verfahren für das Russische: Vokale = {o, a, e, u, y, ъ, ь, я, ъ, э, ю}, Konsonanten = {б, в, г, д, ж, з, ь, к, л, м, н, п, с, т, ф, х, ц, ч, ш, щ}. Interessant ist dabei, daß ь und ъ aus kombinatorischer Sicht immer noch als die Vokale klassifiziert werden, die sie ja, historisch betrachtet, tatsächlich einmal waren. Die Anwendung des Algorithmus auf bekannte Sprachen dient natürlich in erster Linie der Überprüfung und Verfeinerung des Verfahrens, um es anschließend auch zur Entzifferung tatsächlich unbekannter Sprachen einsetzen zu können. Die Algorithmen sind später auch zur Analyse höherer Sprachebenen eingesetzt worden, vgl. SUCHOTIN (1976, 1984).

3.1.2. Effektivität

Die aus den praktischen Erfordernissen der Alphabetisierungskampagnen in der Sowjetunion erwachsene berühmte Arbeit von JAKOVLEV (1928) beschäftigt sich mit der Frage, in welchem Verhältnis die tatsächliche Graphemzahl und die minimal notwendige Zahl eines phonologischen Alphabetes in einer Sprache zueinander stehen. Zur Berechnung dieser unteren Grenze stellt der Autor seine bekannte Formel auf und erläutert sie am Russischen (hier in einer etwas modifizierten Notation):

$$A = (K + V) - (K' - G') + 1$$

Hier steht A für den Umfang des Alphabetes, K für die Zahl der "selbständigen Konsonanten" (33), V für die Zahl der "selbständigen Vokale" (5), K' steht für die Zahl der Phonempaare (12), G' für die Zahl der Vokalpaare (4). Unter den Phonempaaren werden am Beispiel des Russischen konkret nur die an der Palatalitätskorrelation beteiligten Paare berücksichtigt, unter den Vokalen wird /e/ ausgeschlossen, da es nur nach weichen Konsonanten auftritt. Die zusätzliche 1 soll ein Element berücksichtigen, das dem Ausdruck der Palatalität von Konsonanten vor anderen Konsonanten bzw. am Wortende dient, mit anderen Worten das Weichheitszeichen. So ergeben sich für das Russische die folgenden Werte:

$$A = (33 + 5) - (12 - 4) + 1 = 38 - 8 + 1 = 31.$$

Mit anderen Worten: Nach JAKOVLEV käme ein phonemisches Alphabet für die 38 russischen Phoneme durch Reduktion mit 31 Zeichen aus.

Diese Formel mag seinerzeit bei der Konstruktion neuer Alphabete für die Völker der Sowjetunion geholfen haben, aus heutiger Sicht ist aber ihre sprachliche wie linguistische Grundlage inkonsequent durchdacht und die mathematische Umsetzung wie die theoretische Fundierung naiv: die Formel ist nicht mehr als ein historisches Kuriosum. (Offensichtlich hält JAKOVLEV die Graphem- und die Lautebene nicht richtig auseinander und berücksichtigt nur die Palatalitätskorrelation, nicht aber die Stimmhaftigkeitskorrelation; ferner kann [e] auch nach harten Konsonanten auftreten. Der Autor berücksichtigt bei der Konstruktion seines idealen Codes auch nur *einen* Aspekt eines Alphabetes, nämlich seinen Umfang; die Tatsache, daß sich die Verringerung des Zeicheninventars auf die Wortlänge auswirkt, beachtet er nicht.)

Das Verhältnis der Zahl der minimal notwendigen Zeichen (Elemente) zur Zahl der tatsächlich verwendeten Zeichen wird in der heutigen Literatur die Effektivität eines Codes genannt, aber ganz anders berechnet, nämlich mithilfe einer logarithmischen Beziehung (zu einer systematischen Darstellung vgl. ALTMANN/LEHFELDT 1980, 66ff.).

3.1.3. Interne Graphemstruktur

Verschiedene statistische Auswertungen zur internen Struktur (russischer) kyrillischer Grapheme im Vergleich zu denen des lateinischen Alphabetes enthält die Arbeit von ARTEMOV (1933). Berührt werden Aspekte wie Verhältnis von Groß- und Kleinbuchstaben zueinander, Ähnlichkeit und Verwechselbarkeit von Buchstaben usw. Ähnlichen "schriftmorphologischen" Aspekten von Graphemsystemen ist der Beitrag von KEMPGEN (1993) gewidmet, in dem verschiedene typologische Indizes vorgeschlagen werden, um solche Systeme ganzheitlich zu charakterisieren.

3.1.4. Graphemfrequenzen

Einer- und Zweierkombinationen PROSKURNIN (1933) präsentiert eine Buchstabenstatistik getrennt für Groß- und Kleinbuchstaben, für alle Satzzeichen und die Ziffern – solche Daten waren für die sowjetische Druckindustrie im Zuge der Verschriftung bisher schriftloser Sprachen wichtige Grundlagen zur Anfertigung von Drucktypen. Ähnliche Daten finden sich zuerst bei MOROZOV (1915, 98), der die Zahl der Großbuchstaben “pro Pud Kleinbuchstaben vom Kegel Nr. 12a” auflistet – auch dies Erfahrungswerte, wie sie Schriftgießereien bei der Herstellung von Lettern aus offensichtlichen Gründen zusammengestellt haben (“Kegel” ist eine traditionelle Größenangabe für Schrifttypen).

Eine ausführliche *pragmatische* Graphemstatistik des modernen Russischen stammt von BELONOgov/FROLOV (1963), die die Frequenzen der Grapheme und zweigliedriger Graphemverbindungen aus Texten im Umfange von 30.000 Wörtern (entspricht ca. 200.000 Buchstaben) ermittelt haben. Sie schlüsseln ihre Daten vierfach auf: a) ohne Berücksichtigung der Position, b) wortinitial, c) wortmedial, d) wortfinal. Die einzige linguistische Auswertung ist dabei die bloße – aber naheliegende – Behauptung, das Auftreten hänge von der Position ab. Die behauptete Abhängigkeit auch von der Wortlänge wird nicht richtig mit Daten illustriert.

Eine nach Vokalen und Konsonanten geordnete Tabelle zur Häufigkeit aller Zweierverbindungen findet sich auch bei SUCHOTIN (1962, 202), der auf dieser Grundlage das Russische “dechiffriert” (s.o.).

Eine weitere Statistik zur Häufigkeit einzelner Grapheme stammt von GRIGOR’EV (1980a, vgl. auch 1980b), der Gegenwartstexte (Belletristik) mit einer Gesamtlänge von 50.000 Buchstaben ausgewertet hat. Bei KRYLOV (1982) findet sich ein diesen Daten angepaßtes mathematisches Modell. ŽURAVLEV (1970) vergleicht eigene Daten zur *razgovornaja reč’* mit denen von PEŠKOVSKIJ (1925). Die Buchstabenhäufigkeiten in der Fachsprache der Elektronik in einer Stichprobe von genau 100.000 Buchstaben findet man bei KALININA (1968a, 82). Die ermittelten Ränge werden denjenigen aus belletristischen Texten gegenübergestellt (die Verteilungen sind ziemlich ähnlich).

Angaben über die *systemische* Häufigkeit aller einzelnen Grapheme und der zweigliedriger Graphempaare am Wortende finden sich u.a. im OBRATNYJ SLOVAR’ (1974, 929ff.) samt einigen zusätzlichen Auswertungen. Von den dreigliedrigen auslautenden Kombinationen werden nur die häufigsten genannt (-*ться*, -*ать*, -*ный*, -*ние* usw.).

Für **Tab. 1** wurden die Ränge der russischen Grapheme in den einzelnen Positionen nach den Daten von BELONOgov/FROLOV ermittelt und der systemischen Stichprobe nach dem OBRATNYJ SLOVAR’ gegenübergestellt, wobei das nicht-positionsgebundene pragmatische Vorkommen die Grundlage abgibt. Die Ähnlichkeit bzw. Unterschiedlichkeit dieser Rangverteilung kann mit entsprechenden statisti-

schen Tests untersucht werden; ein Kommentar zu einigen offensichtlichen Dingen soll an dieser Stelle genügen.

Graphem	allg., pragm. (Peškov.)	allg., pragm. (Grigor'ev)	allg., pragm. (Bel./Frol.)	initial, pragm. (Bel./Frol.)	medial, pragm. (Bel./Frol.)	final, pragm. (Bel./Frol.)	final, syst. (Obr. Slov.)
О	1	1	1	2	1	6	7
Е	3	2	2	18	2	2	5
А	2	3	3	15	3	4	4
Н	5	5	4	6	4	16	10
И	6	4	5	5	7	1	11
Т	4	6	6	12	6	7	8
Р	10	10	7	8	5	20	9
В	8,5	9	8	3	8	12	18
С	7	7	9	4	9	22	15
Д	11	14	10	7	10	17	16
П	15	15	11	1	16	23	21
М	12	12	12	13	14	5	12
К	13	11	13	9	12	14	6
Я	16,5	17	14	26	21	3	3
Л	8,5	8	15	21	11	19	17
У	14	13	16	11	18	13	22
Ч	18	21	17	17	13	25	28
Ы	20	16	18	–	17	10	14
Э	21,5	22	19	14	15	18	20
Б	19	18	20	10	20	26	26
Ь	16,5	19	21	–	19	11	1
Й	23	23	22	–	24	8	2
Х	24,5	24	23	22	26	9	23
Г	21,5	20	24	16	23	21	19
Ж	27	24	25	25	22	24	24
Ю	26	27	26	27	27	15	29
Ш	24,5	26	27	24	25	28	27
Щ	29,5	28	28	28	28	29	30
Ц	29,5	30	29	23	29	27	13
Ф	31	31	30,5	20	32	31	25
Э	28	29	30,5	19	30,5	31	31
Ъ	32	32	32	–	30,5	31	–

Tab. 1: Rangverteilung der russ. Grapheme nach ihrer Frequenz

Es ist naheliegend, daß die allgemeine Rangverteilung (nach BELONOGOV/FROLOV) und die wortmediale Rangverteilung die größte Ähnlichkeit zueinander aufweisen. Einige interessante Abweichungen gibt es dennoch: die Häufigkeit von П beispielsweise ist medial geringer als allgemein, die allgemeine Häufigkeit wird nämlich durch das extrem häufige Vorkommen im Anlaut (Präfixe!) beeinflusst. Umgekehrt kommt Я initial und medial weniger häufig vor als allgemein, der gute allgemeine Rang geht also v.a. auf die extrem häufigen Endungen -Я zurück. Е, das den zweiten Rang in der allgemeinen Häufigkeit aufweist, ist initial extrem selten, was sich natürlich sprachgeschichtlich erklären läßt (Wandel zu о, vgl. *единъ > один*). Initial fehlen natürlich auch Й-, Ь-, Ы-, Ъ-, letzteres seit der Orthographiereform

von 1917 auch final. Am Wortende kommen pragmatisch überhaupt nur solche Grapheme in mehr als einem Prozent der Fälle vor, die Auslaut einer Flexionsendung sind; die Häufigkeit der Inhalte (v.a. der Kasus) bedingt die Häufigkeit der zugehörigen Endungen.

Zwischen der pragmatischen und der systemischen Stichprobe gibt es krasse Unterschiede, die aber ebenfalls leicht erklärlich sind: -И kommt in Texten am Wortende am häufigsten von allen Graphemen vor, wofür die vielen homonymen Kasus- und Verbendungen verantwortlich sind; im Lexikon hingegen hat nur etwas mehr als 1 Prozent aller Wörter einen Auslaut -И: so enden im Russischen neben einigen Sonderfällen nur wenige nichtflektierte Wörter (*ну, жюри* etc.). Umgekehrt schieben sich -б, -й, -я und -а in der systemischen Stichprobe klar nach vorne, weil die Nennformen praktisch aller Verben, Adjektive und vieler Substantive so enden – 70% aller russischen Lexeme enden auf einen dieser vier Buchstaben.

Die Ergebnisse von GRIGOR'EV entsprechen denen von BELONOGOV/FROLOV bei den obersten und bei den letzten Rängen recht gut, im Mittelbereich dagegen gibt es einige Unterschiede. Der auffälligste Unterschied betrifft die Frequenz von -л-, das in Prosatexten überproportional häufig auftritt; dies läßt sich mit dem in dieser Textsorte häufigen Vorkommen des finiten Präteritums leicht erklären.

Zweierkombinationen Unter den Clustern aus zwei Elementen dominieren am Wortende im Lexikon -ТЬ, -ИЙ und -СЯ unangefochten: fast 50% aller russischen Lexeme enden so, was nicht verwunderlich ist, da diese Kombinationen zufällig mit Suffixen bzw. Endungen zusammenfallen und praktisch sämtliche Verben und Adjektive sowie einen Teil der Substantive erfassen.

Insgesamt spielen demnach für die Graphemfrequenzen sowohl phonologische Restriktionen, Orthographieregeln, die Lautgeschichte wie auch die morphologische Struktur des Russischen eine gewichtige Rolle. Der morphologische Einfluß ist dabei am Wortende sowohl pragmatisch wie systemisch in jeder Hinsicht klar dominierend.

Die Kombinatorik der russischen Grapheme untersucht TOPOROV (1966) in einer grundlegenden Arbeit auf der Grundlage mehrerer Wörterbücher und stellt die Ergebnisse u.a. in Form ausführlicher Listen dar. Ein Kalkül ähnlich dem sog. Harary-Paperschen Modell dient dazu, das Kombinationsverhalten zu quantifizieren; hierbei geht es in erster Linie um den Grad der "Sättigung" der Mengen der Vorgänger- bzw. Nachfolgergrapheme eines gegebenen Graphems, um den Grad der Symmetrie zwischen beiden Mengen usw. KURBAKOV (1965) gibt gleichfalls ausführliche Daten zur Graphemkombinatorik an, ebenso DIETZE (1982b).

Dreierkombinationen Verbindungen von mehr als zwei Graphemen im Russischen sind im Vergleich zum isolierten Vorkommen und zu den Zweierkombinationen bislang weniger gut untersucht. LEBEDEV/GARMAŠ (1959) haben solche Verbindungen analysiert, wobei ihr Interesse von Anwendungen der Informationsthe-

orie für die Nachrichtenübertragung (Telegraph) bestimmt ist: In welchem Umfang könnte ein zu übertragender Text verkürzt werden, wenn nicht jeder Buchstabe für sich weitergegeben wird, sondern die Redundanzen berücksichtigt werden? Die Autoren kommen zu dem Ergebnis, daß eine Geschwindigkeitssteigerung von 5/3 möglich wäre.

Die Daten beruhen auf Stichproben im Umfang von 30.000 Buchstaben aus Tolstojs "Vojna i mir". Aufgeführt werden die 26 häufigsten Dreierkombinationen sowie die Entropiewerte (s.u.) für das Russische im Vergleich zum Englischen. Da der Leerraum auch als Zeichen gezählt wurde, ist es nicht verwunderlich, daß in der "Hitliste" der Dreierkombinationen u.a. folgende zu finden sind: *_u_, ть_, ла_, _он_, _в_, не_, что, ego, ся_*. Man erkennt hierin genau die häufigsten Präpositionen, Konjunktionen, Pronomen und Suffixe des Russischen (vgl. ZASORINA 1977). Mit anderen Worten: Die Häufigkeit, mit der diese Lexeme (und damit die *Inhalte*) gebraucht werden, bestimmt ganz wesentlich die extreme Häufigkeit der Buchstabenkombinationen.

KOLGUŠKIN (1970, 52) zeigt anschaulich, daß der "normale" Morse-Code der Buchstabenhäufigkeit im Russischen nicht gut angepaßt ist, und schlägt stattdessen eine optimierte Variante vor, von der er sagt, sie verringere die Zeitdauer zur Übertragung russischer Texte um mindestens 10% – in dem Kontext des Autors, nämlich der Sprachübertragung im militärischen Bereich, möglicherweise eine relevante Größe.

3.1.5. Informationstheorie

Die informationstheoretische Untersuchung der Sprache beruht zwar ebenfalls auf Häufigkeitszählungen, hat aber einen eigenen Begriffsapparat und entsprechende mathematische Formeln entwickelt, weshalb sie hier in einem separaten Abschnitt behandelt werden soll.

Dieser Ansatz geht von dem grundlegenden Axiom aus, daß die Sprache der Informationsübertragung dient. Sie untersucht Sprache als Kette von aufeinanderfolgenden Elementen (meist Buchstaben) und betrachtet jedes Auftreten eines Buchstabens als Ausgang eines "Experimentes". Dieses Experiment besteht einfach in der Vorhersage, welches der nächste Buchstabe sein wird, unter der Bedingung, daß die vorausgehenden Buchstaben bekannt sind. Offensichtlich hängt diese Voraussagemöglichkeit von verschiedenen Faktoren ab: von der (vermutlichen) Position im Wort, an der man sich befindet, von der Position des Wortes im Satz, auf der untersten Ebene aber ganz einfach davon, welche Grapheme sich häufig miteinander kombinieren, welche nicht. Nach einem *-v-* beispielsweise folgt im Russischen mit hoher Wahrscheinlichkeit ein *-m-* oder ein Vokal wie *-e-, -u-, -ю-*, kaum jedoch ein anderer Konsonant.

Je größer die Wahrscheinlichkeit für das Auftreten eines Elementes, desto geringer die Information, die es liefert, wenn es tatsächlich auftritt. Je unwahrscheinlicher das Auftreten eines Elementes, desto größer der "Überraschungseffekt", wenn

es tatsächlich auftritt, desto größer also die Information, die es überträgt. Maximale Unsicherheit besteht genau dann, wenn alle Buchstaben gleich wahrscheinlich wären. Minimale Unsicherheit besteht genau dann, wenn *ein* Element die Wahrscheinlichkeit 1 hat, alle anderen jedoch die Wahrscheinlichkeit 0. Von beiden Zuständen weichen natürliche Sprachen mehr oder weniger stark ab. Dies sichert die Übertragung gegen Störungen, indem Redundanz in das Signal eingebaut wird: wenn nicht alle Elemente gleich wahrscheinlich aufeinanderfolgen, läßt sich eine Störung u.U. rekonstruieren.

Die Wahrscheinlichkeit des Auftretens einzelner Elemente, auch die "bedingte Wahrscheinlichkeit" genannt, hängt nun ganz offensichtlich von der Länge der Kette ab: Nach zwei Konsonanten steigt etwa die Wahrscheinlichkeit, daß als nächstes ein Vokal folgt, ganz erheblich. Die Informationstheorie untersucht deshalb die Auftretenshäufigkeit von 1, 2, 3, 4, ... langen Kombinationen, was natürlich einen immer größeren empirischen Aufwand mit sich bringt. An einem Beispiel haben JAGLOM/DOBRUŠIN/JAGLOM (1960, 101) deutlich gemacht, welch' großen Einfluß die Berücksichtigung der Kombinatorik auf die Ergebnisse hat:

- F₀: оухеррохьдбщ яыхвшихйжтифвнарфенвштфрпхгпчькизряс
 F₁: т цыяь серв однг збя енвтша ьуемлолйк
 F₂: кая всванный рося ных ковкров
 F₃: покак поститвлennyй пот дурносакаа наконеппо зно стволонил
 F₄: весел вратся не сухом и непо и добре

Diese "Sätze" wurden zufällig erzeugt, wobei im ersten Fall keinerlei Kombinationseigenschaften und Häufigkeiten berücksichtigt wurden, im zweiten die Häufigkeit der Buchstaben, im dritten die Häufigkeit von Zweierkombinationen, dann die von Dreier- und Viererkombinationen. Die Wörter der letzten Zeile sehen "echten" Wörtern des Russischen bereits sehr ähnlich, auch wenn der "Satz" natürlich noch keinen Sinn ergibt.

Die "Entropie" ist nun einfach ein Maß dafür, wie stark die Häufigkeiten der Gleichverteilung angenähert sind, ist also ein Maß für das interne "Gleichgewicht", d.h. die "Unordnung", eines Systems. Die Maßeinheit ist dabei das "Bit", d.h. eine Ja/Nein-Entscheidung. Entropiewerte für Ketten der Länge 1, 2, 3, etc. sind für viele Sprachen berechnet worden. Für das Russische wurden die folgenden Werte ermittelt:

Entropie	schriftl.	mündl.
H ₀	5,00	5,38
H ₁	4,35	4,77
H ₂	3,52	3,62
H ₃	3,01	0,70
	N = 32	N = 42

Tab. 2: Entropie des Russischen

Die Werte zu den Graphemen stammen von LEBEDEV/GARMAŠ (1959, s.o), die Daten zu den Phonemen von CHERRY/HALLE/JAKOBSON (1953), die das Material von PEŠKOVSKIJ (1925) ausgewertet haben. Wegen des geringen Umfanges dieser Daten gilt jedoch insbesondere der extrem niedrige Wert von 0,7 für H_3 als korrekturbedürftig. Der Wert H_0 ist gleichzeitig der Maximalwert; er ist nur von dem Umfang K des Graphem- bzw. Phonemsystems abhängig: $H_0 = \lg 1/K$. Je länger die berücksichtigte Kette, desto kleiner wird notwendigerweise die durchschnittliche "Unsicherheit". Den Wert von 5,00 kann man sich so veranschaulichen: Wenn man aufgefordert würde, einen beliebigen Buchstaben des Russischen zu erraten, so würde man ihn mit 5 Ja/Nein-Fragen sicher herausbekommen (man zerlegt die Gesamtmenge jeweils in zwei Hälften und fragt danach, ob der Buchstabe in der einen Hälfte enthalten ist; je nach Antwort fährt man analog mit einer der beiden Hälften fort).

Wie gut Sprecher tatsächlich "raten", ist verschiedentlich mit entsprechenden Tests überprüft worden. Hierbei wird dem Probanden die Aufgabe gestellt, immer den jeweils nächsten Buchstaben eines Satzes zu erraten, wobei der Testleiter nur mit "ja" oder "nein" antwortet. Gezählt wird die Zahl der Versuche bis zum Erraten der richtigen Fortsetzung. Am Anfang eines unbekanntes Satzes ist diese Zahl natürlich normalerweise hoch, sinkt dann aber stark ab, um an der Grenze zum Endungsmorphem nochmals anzusteigen und dann wieder abzufallen. Dies zeigt: Im informationstheoretischen Sinne trägt der Wortanfang (und damit der Wortstamm) mehr Information als das Wortende, genauso der Morphemanfang mehr als das Morphemende. Dieses Ergebnis macht im übrigen plausibel, warum phonologische Veränderungen häufig am Wortende beginnen oder das Wortende betreffen. Die auf diese Weise gewonnenen empirischen Durchschnittswerte nennt man auch die "Informationsstruktur des Wortes" oder die "Informationsstruktur des Satzes". Konkrete Tests und Resultate finden sich vor allem bei PIOTROVSKIJ (1965, 1968) und PETROVA/PIOTROVSKIJ (1966), ferner bei ŽURAVLEV et al. (1968), NEVEL'SKIJ/ROZENBAUM (1971).

Die einschlägige Literatur zum Thema Entropie umfaßt außer den genannten Arbeiten insbesondere folgende: JAGLOM/DOBRUŠIN/JAGLOM (1960), PIOTROVSKIJ (1962, 1966), PIOTROVSKAJA/PIOTROVSKIJ/RASŽIVIN (1962), KUČERA (1963), GEORGIJEV et al. (1972). In diesen Arbeiten – wie natürlich auch in westlichen Darstellungen – findet man auch die grundlegenden Formeln zur Berechnung der Entropie.

3.2. Phonologie

Einzelne, aber durchaus nicht alle quantitativen Aspekte der russischen Phonologie sind gut bearbeitet. Zu grundlegenden theoretischen und praktischen Fragen vgl. u.a. SEGAL (1972), ALTMANN/LEHFELDT (1980). Die wichtigste Monographie zum Thema ist ohne Zweifel KUČERA/MONROE (1968). Weitere Teilbereiche werden in KEMPGEN (in Vorbereitung) behandelt.

3.2.1. Distinktive Merkmale

Die distinktiven akustischen Merkmale u.a. der russischen Vokale untersucht PIOTROVSKIJ (1960) und vergleicht die Ergebnisse mit Daten zu mehreren weiteren Sprachen. Zur Darstellung werden zahlreiche Diagramme benutzt. Diese – und ähnliche – Untersuchungen werden hier nicht im Detail dargestellt, da es sich um physikalische Messungen an Lauten handelt (Lautdauer, Tonhöhe, Formantien etc.). – TOLSTAJA (1968c) berechnet die Ähnlichkeit der russischen Konsonanten hinsichtlich ihrer Spezifikation nach den distinktiven Merkmalen und ermittelt die Zahl der am Wortanfang bzw. Wortende existierenden Konsonantenverbindungen für jede Phonemdistanz; hierzu s. auch KEMPGEN (1995b).

3.2.2. Phonemzahl

Das Russische ist – mit allen anderen slawischen Sprachen – typologisch gesehen eine Sprache mit “normal hoher” Phonemzahl. Dies sind alle Sprachen mit 19 bis 47 Phonemen. LEHFELDT (1975) zeigt dies anhand einer umfangreichen Stichprobe, wobei den empirischen Daten eine theoretische Verteilung angepaßt wird, aus der sich dann Grenzwerte für eine Klassenbildung ableiten lassen.

3.2.3. Phonemfrequenz

Die erste russische Lautstatistik (DIKAREV 1891, 9) ist eine pragmatische Stichprobe geringen Umfanges (7.500) regionaler Texte. Bekannter als diese Arbeit sind bereits die legendären “10.000 russischen Laute”, die PEŠKOVSKIJ (1925) aus Gesprächen erhob.

KUČERA/MONROE (1968) haben ein Korpus von genau 100.000 Phonemen aus Gegenwartstexten statistisch ausgewertet. Ihr Phoneminventar umfaßt 32 Konsonanten, /j/ und acht Vokale, nämlich fünf betonte und drei unbetonte (/i, a, u/). Eine Rang-Frequenztafel mit den Daten dieser Autoren enthält die **Tab. 3**.

Am häufigsten treten demnach die unbetonten Vokale /a/ und /i/ auf, was nicht erstaunlich ist, wenn man daran denkt, daß ja nur ein Vokal pro Wort betont ist und es demnach mehr unbetonte als betonte geben muß, ferner daran, daß den unbetonten /a/ und /i/ ja jeweils mindestens zwei graphische Realisierungen entsprechen, nämlich *a* und *o* bzw. *u* und *e*. An die Daten lassen sich auch typologisch relevante Beobachtungen anschließen: unter den paarigen Konsonanten sind die harten durchweg häufiger als die weichen (vgl. /t/ auf Rang 4, /t'/ auf Rang 20), außerdem die stimmhaften häufiger als die stimmlosen (vgl. nochmals /t/ auf Rang 4 mit /d/ auf Rang 22).

Rang	Freq. (%)	Phonem	Rang	Freq. (%)	Phonem
1	12,957	a	22	1,665	d
2	11,351	i	23	1,633	č
3	4,629	á	24	1,564	š
4	4,266	t	25	1,381	r'
5	4,137	j	26	1,367	z
6	4,100	n	27	1,361	ú
7	4,026	ó	28	1,304	g
8	3,175	k	29	1,096	b
9	3,093	s	30	1,038	d'
10	2,978	v	31	1,016	v'
11	2,907	r	32	0,987	x
12	2,781	í	33	0,951	ž
13	2,664	l	34	0,948	f
14	2,614	é	35	0,805	m'
15	2,317	m	36	0,562	c
16	2,309	p	37	0,534	k'
17	2,297	n'	38	0,475	p'
18	2,134	u	39	0,368	b'
19	2,081	l'	40	0,322	z'
20	1,885	t'	41	0,059	f'
21	1,863	s'			

Tab. 3: Rang-Frequenz-Verteilung der russischen Phoneme

Die Frage, *warum* ein gegebenes Phonem eine bestimmte zu beobachtende Häufigkeit aufweist, ist in der Literatur vielfach diskutiert worden. Zu einer Übersicht vgl. ALTMANN/LEHFELDT (1980, 112ff.). Auch die Frage, welcher allgemeinen Verteilung die einzelsprachlichen Daten angepaßt werden können, versucht man seit langem zu lösen. – NIKONOV (1963a) behauptet, daß die Phonemfrequenz in den slawischen Sprachen vor allem davon abhängt, wie stark ein Phonem an der Bildung von Derivationssuffixen und Flexionsendungen beteiligt ist – sicher ein wichtiger Einfluß.

Weitere Frequenzangaben finden sich u.a. bei KUČERA (1963), bei ELKINA/JUDINA (1964a).

3.2.4. Phonemkombinatorik

Pragmatisch NIKONOV (1962a), der dem Russischlehrer eine Grundlage dafür geben möchte, welche Lautverbindungen für das Russische typisch sind und deshalb im Fremdsprachenunterricht vorrangig geübt werden müssen, ermittelt die Häufigkeit zweigliedriger Konsonantenkombinationen in einer Stichprobe von 100.000 Lauten, wobei er gesondert auszählt nach (1) Autorenrede in der Belletristik

(60.000 Laute), (2) direkter Rede in der Belletristik (25.000 Laute) und (3) wissenschaftlicher und politischer Prosa (15.000 Laute). Die Daten der ersten Gruppe werden in dem vorliegenden Aufsatz in Form einer zweidimensionalen Matrix publiziert. Die genauen Quellen, denen die Stichprobe entnommen wurde, werden leider nicht genannt. Eine gewisse Einschränkung der Brauchbarkeit bedeutet vor allem aber die Tatsache, daß paarig hart-weiche Konsonanten zusammengefaßt werden, so daß es sich um ein Mittelding zwischen einer Buchstaben- und einer Lautzählung handelt. Auffällige Befunde werden anschließend kommentiert, eine statistische Analyse wird jedoch nicht durchgeführt.

Die Kombinatorik der Phonemklassen des Russischen dient ZINDER (1958) in seinem programmatischen Aufsatz als – kurzes – Beispiel, interpretiert als die Wahrscheinlichkeit, daß auf eine gegebene Klasse eine der zur Auswahl stehenden folgt.

Systemisch Bei KEMPGEN (1995b) findet sich eine tabellarische Übersicht über die zweigliedrigen Konsonantenkombinationen des Russischen im Wortanlaut bzw. Wortauslaut. Grundlage der Zählungen ist eine systemische Stichprobe, d.h. also aus Wörterbüchern. Die Konsonantenkombinationen werden nach dem statistischen Modell von Altmann/Lehfeldt (1980) eingeteilt in "präferierte", "aktuelle", "marginale", "virtuelle" und "unzulässige". In der genannten Arbeit dienen diese Frequenzdaten zur Überprüfung eines möglichen Zusammenhanges zwischen internen (distinktive Merkmale) und externen Phonemeigenschaften (Kombinatorik).

Weitere Arbeiten: NOVÁČEK (1963), BALDWIN (1969), OLIVERIUS (1970).

3.2.5. Silbenstruktur, Silbenfrequenz

Eine Untersuchung der Silbenstruktur russischer Wörter haben ELKINA/JUDINA (1964a, 1964b) in zwei Aufsätzen publiziert. Als Stichprobe dienten schriftliche Texte (50% Belletristik, 25% Publizistik [konkret: eine Rede von CHRUŠČEV], 25% Wissenschaft und Technik), die zunächst phonetisch transkribiert und dann nach den Prinzipien von ŠČERBA in Silben zerlegt wurden. Den 94.000 laufenden Silben lagen 3910 verschiedene Silben (*types*) zugrunde. Von ihnen werden die 359 häufigsten in Form einer Rang-Frequenz-Liste aufgeführt, die 1580 häufigsten außerdem in einer alphabetischen Liste. Zwei allgemeine Zusammenhänge werden zusätzlich tabelliert, nämlich die relative Häufigkeit der Silben verschiedener Länge und die Häufigkeit der Silbenstrukturen (1964a, 60). Diese Daten wollen wir hier in einer kombinierten Tabelle wiedergeben:

Länge	rel. Anteil	Struktur	rel. Anteil
1	0,105	V	0,0850
		K	0,0200
2	0,559	VK	0,0156
		KV	0,5434
3	0,277	KVK	0,1406
		KKV	0,1362
4	0,0522	KVKK	0,0039
		KKVK	0,0351
		KKKV	0,0132
5	0,0062	KKVKK	0,0062
≥6	0,0006	übrige	0,0008

Tab. 4: Silbenlänge und Silbenstruktur (pragmatisch)

Zu den von den beiden Autorinnen gewonnenen Daten muß allerdings unbedingt angemerkt werden, daß hier phonetische Einheiten gezählt werden, auch wenn von 'Phonemen' die Rede ist. Außerdem ist natürlich die Silbentrennung, die ŠČERBA vorgeschlagen hat, nicht allgemein akzeptiert und weicht z.T. erheblich von anderen Ansätzen, etwa dem von AVANESOV, ab. – Die zweite Untersuchung der beiden Autorinnen geht ganz ähnlich vor, nur werden diesmal Texte zur Radioelektronik ausgewertet. Ganz ähnliche Daten finden sich bei KUČERA/MONROE (1969, 47), die außerdem die Frequenz jedes Phonems in jeder Silbenposition angeben (1969, 57).

Die Struktur einsilbiger Wörter hat MOISEEV (1975) auf der Grundlage eines Wörterbuches statistisch ausgewertet und nennt die folgenden Zahlen für die 10 häufigsten Silbentypen, die fast 98% des Materials ausmachen (K = Konsonant, V = Vokal):

Struktur	abs. Häufigkeit	rel. Häufigkeit	Beispiel
KKVK	1003	0,318	stol
KVK	960	0,305	dom
KVKK	457	0,148	most
KKVKK	177	0,057	stolb
KKKVK	170	0,054	strach
KKV	120	0,038	dno
KV	71	0,023	na
VK	47	0,015	um
VKK	31	0,01	al't
KKKV	30	0,01	tekst
Summe	3066	0,978	

Tab. 5: Die Struktur einsilbiger Wörter (systemisch)

Es ist offensichtlich, daß die pragmatischen Häufigkeiten der gleichen Silbenstrukturen anders ausfallen würden, u.a. wegen der hohen Frequenz des Typs KV (*на, по, ты, мы, вы, да...*).

Daten zur Silbenstruktur des Russischen finden sich auch bei ZAPLATKINA; nach dieser Autorin liegt die durchschnittliche Silbenlänge bei 3,67 Buchstaben, was für die slawischen Sprachen im Vergleich ein ziemlich "normaler" Wert ist (vgl. 1980, 19).

Zu weiteren Arbeiten vgl. OLŠEVSKIJ/POSTOVALOVA (1968, 1969), APPEL (1957/58).

3.3. Morphologie

3.3.1. Morphem und Wort

Morphemfrequenz Die Häufigkeit der russischen Morpheme behandelt OLIVERIUS (1976) ausführlich. Diese Monographie besteht aus zwei großen Teilen: Zunächst wird die zugrundegelegte Morphemkonzeption (ausdrücklich anti-generativ und vielmehr *item and arrangement* verpflichtet) ausführlich erläutert, dann folgt das eigentliche Frequenzwörterbuch, das auf den Daten von ŠTEJNFEL'DT (1963) aufbaut: Die dort vorhandenen Wörter wurden phonologisch interpretiert und morphologisch analysiert. Ergebnis ist eine Rang-Frequenz-Liste von Morphemen (pp. 121–143), wobei die Häufigkeit jedes Morphems gegebenenfalls zusätzlich nach Allomorphen differenziert wird. So finden sich z.B. unter /mog/ mit seiner Gesamtfrequenz von 2221 Vorkommen Einzelangaben zu /moč'/ 1318, /mož/ 549, /mošč'/ 217 und /mog/ 138. Eine alphabetische Morphemliste (pp. 145–190) enthält zusätzliches Belegmaterial. Ausgeschlossen aus der gesamten Analyse sind Endungsmorpheme, da über sie aus der benutzten Quelle keine vollständigen Angaben zu entnehmen sind.

Zusammenfassende Daten über Morphemtypen enthält eine abschließende Tabelle (p. 191), die wir hier in etwas modifizierter und um die vierte Spalte erweiterter Form wiedergeben, vgl. **Tab. 6**.

Mit isolierten Morphemen sind hier Präfixe, Präpositionen, Konjunktionen, Interjektionen und Partikel gemeint. – Unter dem token/type-Verhältnis findet sich die durchschnittliche Häufigkeit eines jeden Morphemtyps. Daß es dabei zu ganz unterschiedlichen Werten kommt, ist in einigen Fällen offensichtlich und naheliegend: So werden z.B. die wenigen Wurzeln der geschlossenen Klasse der Pronomina extrem häufig benutzt, die vielen substantivischen Wurzeln hingegen unterdurchschnittlich häufig.

Morphemtyp	types	tokens	token/type
Isolierte Wurzeln	81	151371	1868,778
Σ	81	151371	1868,778
Verbale Wurzeln	273	82145	300,897
Substant. Wurzeln	425	51341	120,802
Nominale Wurzeln	142	26782	188,606
Pronom. Wurzeln	28	72342	2583,643
Σ	688	232900	338,517
Relationale Verbalsuffixe	6	44380	7396,667
Relationale Subst.suffixe	40	40439	1010,975
Relationale Nominalsuffixe	26	91218	3508,385
Relationale Pronom.suffixe	13	55379	4259,923
Gemischte Substant.suffixe	7	3884	554,857
Gemischte Nominalsuffixe	3	181	60,333
Modifizierende subst. Suffixe	6	3762	627,000
Modifiz. nominale Suffixe	4	3517	879,250
Σ	105	242750	2311,905
$\Sigma\Sigma$	1054	626831	594,716

Tab. 6: Morphemtypen im Russischen

Morphemlänge Der Frage, inwieweit die russischen Präsensmorpheme im Hinblick auf ihre Länge optimiert sind, geht KEMPGEN (1995a) nach. Wenn dies der Fall ist, so müßte dem häufigsten Inhalt der kürzeste Ausdruck entsprechen usw., um Ökonomie hinsichtlich des Aufwandes und der Übertragungsdauer zu erreichen. Tatsächlich läßt sich zeigen, daß das Russische dieses Prinzip durchaus befolgt, daß dies aber nicht der einzige Einfluß auf die Morphemlänge ist.

Länge	Anzahl	Länge	Anzahl
1	8	13	130
2	27	14	94
3	86	15	65
4	143	16	37
5	256	17	24
6	308	18	8
7	309	19	3
8	342	20	2
9	367	21	1
10	346	22	2
11	268	23	2
12	179	24	0

Tab. 7: Verteilung russischer Wortformen nach ihrer Länge

Wortlänge in Buchstaben Für diesen Zusammenhang gibt KURBAKOV (1965) eine Verteilung an, die aus Texten einer Gesamtlänge von 25.000 laufenden Wörtern gewonnen wurde (vgl. **Tab. 7**). Ausgewertet wurde dabei jedoch die Länge der *types*, jede vorkommende Wortform also nur einmal gezählt (N = 3007).

Diese Zahlen, bei KURBAKOV auch in Form einer Kurve wiedergegeben, zeigen u.a., daß die Mehrzahl der russischen Wörter zwischen 5 und 11 Buchstaben lang ist (73% aller Fälle).

BELONOGOV (1962) hat die durchschnittliche Länge (in Buchstaben) russischer Wörter in Abhängigkeit von ihrer Häufigkeit empirisch untersucht und ist dabei zu folgenden Werten gekommen:

x	10	50	100	200	600	1.800	12.000
Länge	2,16	3,31	3,96	4,51	5,33	6,11	6,83

Tab. 8: Durchschnittliche Länge der *x* häufigsten Wörter

Diese Zahlen zeigen sehr anschaulich den Zusammenhang von Häufigkeit und Wortlänge: Extrem häufige Wörter sind gewöhnlich kurz, was sprachliche Kodierung effektiv und ökonomisch macht. Mit abnehmender Häufigkeit wächst die durchschnittliche Wortlänge deshalb kontinuierlich an.

Wortlänge in Phonemen Aufgeschlüsselt nach verschiedenen Funktionalstilen (Umgangssprache, Belletristik, Wissenschaftl. Prosa, Publizistik) gibt NIKONOV (1978) folgende durchschnittliche Wortlänge an: 4,5 – 5,2 – 6,7 – 7,0. D.h.: schriftliche Texte zeichnen sich durch eine deutlich höhere Wortlänge aus als die gesprochene Sprache.

Wortlänge in Silben Eine Verteilung der Wortlänge in Silben findet sich im Anhang des OBRATNYJ SLOVAR' (1974, 936f.), wobei die Angaben völlig unnötigerweise nach den Endbuchstaben der Wörter aufgeschlüsselt werden. Mit anderen Worten: hier hat man ohne viel linguistischen "Input" das zur Verfügung stehende Material rein mechanisch zu Statistiken verarbeitet und die Tabellen unnötig aufgebläht. Ungeachtet dessen ist die Kenntnis der grundlegenden Verteilung typologisch, aber auch einzelsprachlich sehr interessant: was sind (im Russischen) lange, was sind kurze Wörter? Die einschlägigen Daten (relative Anteile) zeigt die **Tab. 9**.

Hiernach ist ganz klar: das Russische kennt zwar sehr lange Wörter (hier bis zu zwölf Silben), "normal" sind aber 3 bis 4 Silben (knapp 60% aller Fälle), typisch auch noch 2 und 5 Silben – 87% aller Lexeme fallen in diesen Bereich, alle anderen Wortlängen sind praktisch bedeutungslos.

Silbenzahl	rel. Anteil
0	0,0002
1	0,019
2	0,129
3	0,309
4	0,276
5	0,157
6	0,073
7	0,027
8	0,0072
9	0,0014
10	0,0004
11	0,0001
12	0,0000

Tab. 9: Wortlänge (in Silben), systemisch

Eine Graphik zur Wortlänge in Silben und in Morphemen im technischen Russisch findet sich bei MOSKOVIČ (1967), vgl. auch MANUČARJAN (1972) mit einem Vergleich zwischen dem Armenischen und dem Russischen.

Wortlänge und Betonungsstelle Der russische Akzent gilt als im Prinzip frei, weil im Prinzip jede Silbe (die erste, die zweite, die dritte eines Wortes usw.) betont sein kann; jedes einzelne Wort hat aber eine genau festgelegte Betonungsstelle (von einigen Variationen einmal abgesehen). Daß die Akzentstelle abgeleiteter Wörter durch die Derivationsuffixe weiter eingeschränkt wird, ist ebenfalls bekannt (Möglichkeiten: Beibehaltung der Akzentstelle des ableitenden Wortes oder Akzentstelle vor bzw. auf dem Suffix). Welche Regeln und Tendenzen aber die Akzentstelle der großen Masse der russischen Wörter bestimmen, das haben mehrere Autoren zu ergründen versucht.

Die grundlegende Hypothese geht dabei auf L.L. VASIL'EV zurück, der schon Anfang des 20. Jh.s als "Gesetz der russischen Betonung" die Behauptung aufstellte, die Akzentstelle tendiere zur Wortmitte. Als erster hat NIKONOV (1963b) diese Hypothese empirisch überprüft, und zwar an Texten und damit in Form einer pragmatischen Stichprobe.

MOISEEV (1969) führt die Fragestellung weiter und kommt zu einer Präzisierung der Antwort, führt aber auch Differenzierungen in das Material ein, die künstlich und irrelevant sind (gerade/ungerade Zahl von Silben) und die Suche nach Regeln in eine falsche Richtung lenken. Immerhin kommt er zu einigen empirischen Verallgemeinerungen, die er ihrerseits auf die Wirkung zweier Grundtendenzen zurückführt, 1) auf die Bevorzugung der Wortmitte gegenüber der Peripherie, 2) auf die Bevorzugung der zweiten Worthälfte gegenüber der ersten. Von beiden hält er die erste für grundlegender, obwohl ihr ein Teil der Daten widerspricht. Grundlage seiner Zählungen ist ein Wörterbuch (*Russkoe literaturnoe proiznošenie i*

udarenie, M. 1959), aus dem alle 49.483 mehrsilbigen Wörter ausgewertet wurden; es handelt sich also um eine systemische Stichprobe.

Auch das OBRATNYJ SLOVAR' (1974, 938ff.) beschäftigt sich mit dieser Frage, allerdings in unvollständiger und längst überholter Form und Fragestellung: es wird nur nach Anfangs- und Nichtanfangsbetonung unterschieden, dies dann aufgeschlüsselt nach den Endbuchstaben der Wörter (was erkennbar keinerlei Relevanz besitzt). So läßt sich den Zahlen nur entnehmen, daß die Nichtanfangsbetonung häufiger ist als die Anfangsbetonung – aber für die nichtinitiale Betonung gibt es ja auch bei drei und mehr Silben viel mehr Möglichkeiten.

Tab. 10 zeigt, wie häufig bei gegebener Wortlänge die einzelnen Silben betont werden (Zahlen korrigiert nach MOISEEV). Den in jeder Zeile höchsten Wert haben wir dabei unterstrichen. Es ist m.E. offensichtlich, daß eine statische Betrachtung (Anfangsbetonung vs. Endbetonung, erste Worthälfte vs. zweite Worthälfte), wie sie bisher angestrebt wurde, den Daten nicht gerecht wird. Augenscheinlich verlagert sich der "Gipfel" jeder Verteilung mit wachsender Wortlänge nach rechts, d.h. wächst parallel zur ihr. Das durchschnittliche Verhältnis von häufigst betonter Silbe zur Wortlänge beträgt 0,711. Mit einer einfachen Faustregel kann man hieraus die Gipfelsilbe aus der Wortlänge abschätzen: Man multipliziert die Wortlänge mit der Konstante 0,711 und rundet zur nächsten ganzen Zahl auf bzw. ab. Ein Beispiel: Bei einer Wortlänge von 6 Silben erhalten wir $6 * 0,711 = 4,266$. Die vierte Silbe sollte demnach die am häufigsten betonte Silbe sein, was tatsächlich zutrifft. Diese einfache lineare Abschätzung ergibt für 7 von 10 Fällen den richtigen Wert. Es ist allerdings offensichtlich, daß in einigen Fällen, vor allem bei den zweisilbigen Wörtern, andere Einflüsse "störend" wirken. Es kann jedenfalls festgehalten werden, daß die Akzentstelle im Russischen ganz offensichtlich eines der "Massenphänomene" ist, das statistischen Tendenzen unterliegt. Die Wortlänge als solche erlaubt bereits eine gute Abschätzung des Schwerpunktes der Verteilung, wohingegen die bloße Unterscheidung von Anfangsbetonung vs. Endbetonung ein zu starres Denkschema ist. Eine graphische Umsetzung solcher Zahlen ist im übrigen ein gutes heuristisches Hilfsmittel zur Hypothesenbildung. Im vorliegenden Fall sieht man bei einer Kurvendarstellung praktisch ein Gebirge, bei dem die Gipfel von rechts nach links "wandern".

Länge	Betont Silbe											Σ
	1	2	3	4	5	6	7	8	9	10	11	
2	3834	<u>5183</u>										9017
3	3138	<u>8224</u>	6189									17551
4	447	5241	<u>6139</u>	1440								13267
5	50	340	<u>3685</u>	1970	324							6369
6	2	59	332	<u>1592</u>	388	55						2428
7	–	1	24	199	<u>405</u>	66	6					701
8	–	–	1	7	50	<u>63</u>	3	–				124
9	–	–	–	–	3	<u>7</u>	<u>7</u>	–	–			17
10	–	–	–	–	1	1	2	<u>4</u>	–	–		8
11	–	–	–	–	–	–	–	<u>1</u>	–	–	–	1
Σ	7471	19048	16370	5208	1171	192	18	5	–	–	–	49483

Tab. 10: Zusammenhang von Wortlänge (in Silben) und Betonungsstelle

Flexionsakzent und Wortfrequenz Der Flexionsakzent des Russischen zeichnet sich bekanntlich dadurch aus, daß er beweglich ist, d.h. innerhalb eines Paradigmas seine Silbenposition verändern kann (vgl. N.Sg. *ruká* – Akk.Sg. *rúku* – N.Pl. *rúki*). Diese Eigenschaft kommt zu der oben genannten "Freiheit" des Akzentes hinzu, wobei jedoch vielfältige empirische Abhängigkeiten zwischen beiden Bereichen liegen. In der Literatur gibt es Hinweise, daß der Akzenttyp eines Wortes von seiner Häufigkeit abhängen könne; diesem Zusammenhang geht MUSTAJOKI (1981) nach. Er nimmt aus dem Häufigkeitswörterbuch von ZASORINA (1977) die 40, 400 und 4.000 häufigsten Wörter und aus ZALIZNJAK (1977) das Gesamtkorpus der Substantive, nämlich 40.000, und untersucht den Anteil der Wörter mit beweglichem Akzent in diesen Gruppen, getrennt nach stambetonten und endbetonten (bezogen auf die Nennform). Die sich ergebenden Werte lauten wie folgt:

x häufigste Wörter	Stambetonung mit bewegl. Akz.	Endbet. mit bewegl. Akzent
40	63,3	(88,9)
400	27,5	92,1
4.000	14,0	73,2
40.000	6,4	39,2

Tab. 11: Frequenzklasse und beweglicher Akzent (Anteile in %)

Diese Zahlen zeigen ganz eindeutig, daß unter den häufigsten Wörtern besonders viele mit beweglichem Akzent zu finden sind, ihr Anteil also stetig abnimmt, je größer die Gruppe der betrachteten Substantive ist. Bevor dies vorschnell als der gesuchte Beweis verstanden wird, sollte man sich jedoch einige Dinge klarmachen: 1) Ganz offensichtlich weisen alle endbetonten Wörter eine Tendenz zur Wechselbetonung auf – dies sieht man an dem sehr viel höheren Anteil in dieser zweiten

Gruppe. Die Betonungsstelle selbst hat also ebenfalls einen großen Einfluß auf das Betonungsmuster. 2) Es macht wenig Sinn, wie MUSTAJOKI dies tut, abgeleitete und nichtabgeleitete Wörter in einen Topf zu werfen – Akzentstelle und Betonungsmuster abgeleiteter Wörter unterliegen bekanntlich anderen, viel strikteren Regeln.

Da das Betonungsmuster auch an die Silbenzahl des Wortes gekoppelt ist, untersucht MUSTAJOKI auch diesen Zusammenhang und stellt fest, daß der durchschnittliche Anteil der beweglichen Betonung mit wachsender Wortlänge abnimmt. Auch diese Korrelation ist zwar formal richtig, muß aber durchdacht werden: Es gilt ja ganz allgemein, daß Häufigkeit und Wortlänge dergestalt voneinander abhängen, daß häufige Wörter besonders kurz sind und die Länge mit abnehmender Frequenz zunimmt. Also haben wir es hier in Wirklichkeit gar nicht mit einer neuen, unabhängigen Variablen zu tun. Zudem führt MUSTAJOKI als "Erklärung" für den genannten Zusammenhang das Vasil'evsche "Gesetz" (s.o.) an, das in dieser Form sicher nicht präzise genug ist. Ferner zeigen die (hier nicht wiedergegebenen) Zahlen, daß beweglicher Akzent überhaupt nur unter den bis zu dreisilbigen Wörtern relevant ist. Schließlich und endlich: die Wortlänge ist ein Merkmal, das unmittelbar mit der Frage zusammenhängt, ob das Wort abgeleitet ist oder nicht – Derivation macht ein Wort (im Russischen) fast immer länger. Damit aber sind wir wieder bei dem schon genannten Punkt angelangt, daß nach abgeleiteten und nichtabgeleiteten Wörtern zu trennen wäre.

Den Zusammenhang von Wortfrequenz und Betonungsmuster untersucht auch HENTSCHEL (1990), und zwar im Hinblick auf den von anderen behaupteten Zusammenhang von Markiertheit (Unmarkiertheit ist "natürlich[er]") und Sprachwandel (Markiertes, d.h. auch Irreguläres, weil "unnatürlich", tendiert zum Abbau). Konkret ermittelt er an 462 Substantiven, die in historischer Zeit ihr Akzentschema verändert haben, daß hiervon zwei Drittel zu beweglichem Akzent übergegangen sind, ein Drittel zu fester Stammbetonung. Die durchschnittliche Vorkommenshäufigkeit liegt in der ersten Gruppe signifikant höher als in der zweiten, und der Autor kommt so zu dem Schluß, daß hohe Frequenz die Annahme eines irregulären (markierten) Musters begünstigt. Unberücksichtigt bleibt dabei, von welchem Ausgangspunkt aus die Veränderungen erfolgten und wie groß die betreffenden Ausgangsgruppen überhaupt waren. Dies könnte das Bild möglicherweise beeinflussen. Jedenfalls sieht der Autor in hoher Frequenz ein "auslösendes Moment" für morphologischen Wandel, der dem genannten Natürlichkeitsprinzip durchaus widersprechen kann.

Wortauslaut und Wortart Eine Tabelle im OBRATNYJ SLOVAR' (1974, 940ff.) gibt Auskunft über die Verteilung der Wortarten und bestimmter Untergruppen auf den Auslaut (z.B. vollendete vs. unvollendete Verben bei den Auslauten *-b* und *-s*). Dieser Zusammenhang ist in einer flektierenden Sprache natürlich relevant, u.a. für die automatische Spracherkennung, aber auch als grundlegende Statistik zu den grammatischen Klassen. Ein Beispiel: Wenn ein Wort auf *-a* endet, kommen dann

außer den femininen Substantiven noch weitere grammat. Klassen in Betracht? Wenn ja, welche? Die Tabelle gibt dem, der nach ihr sucht, die Antwort: in 95% aller Fälle liegt tatsächlich der genannte Fall vor, aber auch Wörter anderer Wortarten können so auslauten (*дѣа, едѣа, когдѣа...*). Die weitergehende Frage, ob sich größere Eindeutigkeit ergibt, wenn die zwei, drei usw. letzten Grapheme berücksichtigt werden, kann die Tabelle natürlich nicht beantworten (*-ууа* läßt z.B. einen eindeutigen Schluß zu). Das Wörterbuch liefert in allen Fällen also nur die Rohdaten, statistische Tests und Interpretationen des Materials muß der Untersuchende selbst vornehmen.

3.3.2 Flexionsmorphologie

Flexion, Akzent Eine umfangreiche statistische Untersuchung aller möglichen Aspekte der Flexionsmorphologie des Russischen stammt von ILOLA/MUSTAJOKI (1989). Diese Arbeit basiert auf einer computergestützten Auswertung des bekannten "Grammatischen Wörterbuches" von A.A. Zaliznjak und enthält mehr als hundert Tabellen, denen jeweils illustrierende Beispiele für jeden belegten Fall beigegeben sind. In der Regel handelt es sich bei den Auswertungen um Kreuzklassifikationen, bei denen zwei oder gegebenenfalls drei Merkmale miteinander kombiniert werden. Eines der Merkmale gehört in der Regel der phonologischen Ebene an, z.B. Länge des Wortes in Silben, Auslaut des Wortes oder eine morphologische Alternation, das andere Merkmal ist in diesem Fall morphologischer Natur, z.B. ein bestimmtes Flexions- oder Betonungsmuster, die Akzentstelle usw. Im Prinzip findet sich jedenfalls zu jedem Kürzel, das in dem Wörterbuch verwendet wird, eine Auswertung zur Häufigkeit seines Auftretens in Abhängigkeit von weiteren Merkmalen. Alle Angaben werden in Form von absoluten Zahlen aufgeführt, statistische Tests irgendwelcher Art werden jedoch nicht durchgeführt, so daß das Material in dieser Hinsicht noch einer sorgfältigen Bearbeitung und Interpretation der Ergebnisse harret. Gelegentlich schießen die Autoren auch über das Ziel hinaus und schlüsseln Zusammenhänge im Detail nach einzelnen morphologischen Klassen auf, wo man sich zunächst einmal einfach eine Übersicht gewünscht hätte. Daß es in der Fülle des Materials auch einige Fehler gibt (so etwa S. 124f. in den Beispielen und Zahlen), ließ sich vielleicht nicht ganz vermeiden. Es versteht sich, daß es sich an dieser Stelle schon aus Platzgründen verbietet, auf die Auswertungen im einzelnen einzugehen. – Zur Häufigkeit der Akzentparadigmen im Grundwortschatz vgl. auch TORNOW (1984) mit vielen Einzelstatistiken.

Wortarten KEMPGEN (1981) hat ein empirisches Verfahren aufgezeigt, wie man von Inhaltsparadigmen und ihren grammatischen Kategorien zu Wortarten kommt und die resultierenden Klassen anschließend interpretiert (226ff.). Auf die Ergebnisse einer Ähnlichkeitsmessung wird hier ein hierarchisches Klassifikationsverfahren angewendet, das eine Wortartenklassifikation in Form eines Baumdia-

grammes liefert (239). Anschließend wird die Stärke des Zusammenhanges (Korrelation und prädiktive Assoziation) zwischen allen grammatischen Kategorien bestimmt und graphisch umgesetzt.

Die Häufigkeit der Wortarten des Russischen ist in einer Reihe von Arbeiten untersucht worden, so daß sich die Ergebnisse miteinander vergleichen und kontrastieren lassen. Daten aus Texten einer Gesamtlänge von 12.000 Wörtern hat MARKOV (1960b) ermittelt, wobei den Anstoß zu dieser Untersuchung die Beobachtung gab, daß entsprechende Angaben bei JOSSELSOHN offenbar nicht stimmen können. MARKOV führt die prozentualen Anteile für Autoren- und Personenrede getrennt auf, und seine Werte zeigen auch ohne Test, daß die Häufigkeit der Wortarten entscheidend von diesem Unterschied abhängt. **Tab. 12** stellt die Werte von MARKOV und die des OBRATNYJ SLOVAR' (1974, 944) zusammen. Die Werte zum Altrussischen stammen aus SANNIKOV (1978, 26); ihnen liegen allerdings nur juristische Texte zugrunde.

Wortart	Altrussisch (pragmatisch)	Autorenrede (pragmatisch)	Personenrede (pragmatisch)	Lexikon (systemisch)
Substantiv	47,12	28,0	19,4	46,35
Adjektiv	11,82	9,6	3,9	20,40
Numerale	1,71	1,4	0,8	0,10
Pronomen	2,29	10,6	20,2	0,08
Verb	26,59	15,8	20,0	30,71
Partizip	–	2,5	0,3	–
Adverbialpart.	–	1,2	0,0	–
Adverb	4,24	7,6	10,5	1,58
Präposition	} 6,24	12,4	7,5	0,12
Konjunktion		7,8	6,3	0,09
Partikel		3,1	10,1	0,12
Interjektion		0,0	1,0	0,28

Tab. 12: Anteile der Wortarten in Prozent

Bestimmte Unterschiede zwischen einer systemischen Stichprobe, wie sie das Lexikon darstellt, und dem Gebrauch in Texten sind naheliegend: Alle geschlossenen Wortarten wie Präpositionen, Konjunktionen und Pronomen umfassen verschwindend wenige Lexeme, die aber tatsächlich häufig gebraucht werden – bei diesen drei Wortarten ist der Anteil im Text um gut das Hundertfache höher als der Anteil am Lexikon! Umgekehrt werden die gewaltigen Mengen an Substantiven, Adjektiven und Verben, die das Lexikon bereitstellt, nicht in gleicher Weise auch ausgenutzt. – Zwischen den beiden pragmatischen Stichproben zum modernen Russisch lassen sich, wie schon angedeutet, ebenfalls wesentliche Unterschiede erkennen: Die Autorenrede ist durch einen deutlich höheren Anteil an Substantiven und Adjektiven gekennzeichnet als die Personenrede, was nicht überrascht (Schilderungen, Beschreibungen), umgekehrt verwendet die mono- oder dialogische Per-

sonenrede viel mehr Pronomen und Partikeln. Partizipien und Adverbialpartizipien gelten mit Recht als buchsprachlich, da sie mündlich fast nicht vorkommen. Die altrussischen fachsprachlichen Texte müssen am ehesten mit der Autorenrede verglichen werden. Sie zeigen einen extrem hohen Anteil an Substantiven und sehr wenigen Pronomen.

Den Einfluß der *razgovornaja reč'* auf den Gebrauch der Wortarten untersucht KLOČKOVA (1970), indem sie den schriftlichen und mündlichen Sprachgebrauch zweier Personen auswertet und die Ergebnisse in Tabellen zusammenstellt. Es ergibt sich zwar, daß tatsächlich Unterschiede zwischen beiden Bereichen bestehen, doch sind sie bei den Versuchspersonen uneinheitlich und z.T. sogar gegenläufig ausgeprägt, so daß es für Verallgemeinerungen noch zu früh ist, außer für Beobachtungen wie die, daß mündlich weniger Substantive und folglich Präpositionen gebraucht werden, jedoch mehr Partikeln. Ausführliche Statistiken zur Häufigkeit der Wortarten und bestimmter Unterklassen finden sich in einer weiteren Arbeit der gleichen Autorin (KLOČKOVA 1968). JIRÁKOVÁ (1976) analysiert in zahlreichen Tabellen und Graphiken, wie sich die Anteile der einzelnen Wortarten in der Rangfrequenzliste eines Wörterbuches allmählich verschieben, wozu sie mehrere Häufigkeitswörterbücher miteinander vergleicht. Die Häufigkeit der Wortarten bei verschiedenen russischen Autoren (Hercen, Turgenev, Gončarov) dient GOLOVIN (1974) als Beispielmateriale, RAECKE (1993) untersucht den Briefstil des 18. Jhs. u.a. in Bezug auf die Wortartenverteilung.

Flexionsparadigmen Die sowjetische Linguistik hat einen eigenständigen Ansatz entwickelt, um mithilfe eines algorithmischen Verfahrens die Flexionsparadigmen einer Sprache zu ermitteln. Dieser Ansatz ist vor allem mit den Namen von N.D. ANDREEV und L.D. ANDREEVA verknüpft. Das Verfahren benutzt Texte als "Input" und Angaben zur Häufigkeit der Grapheme, der Verteilung der Wort- und der Phrasenlänge, um auf dieser Grundlage eine morphologische Segmentierung des Materials vorzunehmen, aus der sich dann Klassen von Endungen und Stammmorphemen und schließlich sogar Wortarten ergeben sollen. Ausdrückliches Ziel ist es dabei, keinerlei Kenntnis der Inhaltsseite zu benutzen. In dieser Hinsicht ähnelt der Ansatz dem strengen Empirismus des amerikanischen taxonomischen Strukturalismus. Da die Ermittlung statistischer Daten zum Russischen nicht Ziel der genannten Arbeiten ist, sondern solche Daten vielmehr nur als Teil des Algorithmus benötigt werden, soll an dieser Stelle darauf verzichtet werden, diese Arbeiten ausführlich darzustellen. Die wichtigsten Publikationen sind ANDREEV (1965a, 1965b, 1967) sowie ANDREEVA (1969).

Kasusformen Angaben zur Häufigkeit der Kasus im Russischen lassen sich aus mehreren Quellen ersehen, wobei sich ein recht stabiles Bild ergibt. NIKOLAEV (1960) hat entsprechende Daten für fremdsprachendidaktische Zwecke ermittelt, wobei ihm als Material 10.146 zufällig erhobene Kasusformen aus zwei Pionier-Zei-

tungen der Jahre 1958/59 dienten. Die nachstehende Tabelle stellt die Ränge, die die Kasus in den Zählungen mehrerer Autoren einnehmen, vergleichend zusammen.

Kasus	Josselson	Nikolaev	Štejnfel'dt	Gurčenkova
Nominativ	1	1	33,6%; 1	2
Genitiv	3	2	24,6%; 2	1
Akkusativ	2	3	19,5%; 3	4
Präpositiv	4	4	9,4%; 4	3
Instrumentalis	5	5	7,8%; 5	5
Dativ	6	6	5,1%; 6	6
	N = 10.146			N = 5.846

Tab. 13: Rang-Verteilung der russischen Kasus

Die Gruppe der häufigen Kasus bilden Nominativ, Genitiv und Akkusativ mit zusammen mehr als 75% aller Fälle, wobei der Nominativ ca. 33% beisteuert, Genitiv und Akkusativ je ca. 20–25%. Präpositiv, Instrumentalis und Dativ folgen mit großem Abstand und tragen ca. 5–10% bei. Bei NIKOLAEV (1960) wie bei ŠTEJNFEL'DT (1963) stimmen die Daten praktisch überein. In der publizistisch-politischen Fachsprache hingegen (GURČENKOVA 1972) liegt der Genitiv mit über 60% an der Spitze, der Nominativ steuert nur weitere 17% bei. Diese Fachsprache weicht demnach eventuell signifikant von den allgemeinen Verhältnissen ab.

Die genannten Daten relativieren auch das Bild vom Russischen als einer flektierenden Sprache: Bei maskulinen und neutralen Substantiven (Nom. = Akk.) wird die Grundform des Lexems in über 55% aller Fälle benutzt (wenn wir von der Belebtheit einmal absehen).

Präpositionsloser und präpositionaler Kasusgebrauch treten nach NIKOLAEV im Verhältnis 7 : 3 auf, was hier den synthetischen Charakter des Russischen unterstreicht. Eine genauere Aufschlüsselung der Daten zeigt ferner folgende Korrelation: Je häufiger ein Kasus ist, desto häufiger wird er präpositionslos gebraucht – bei den selten gebrauchten Kasus ist das Verhältnis hingegen ausgeglichen (D, I) oder hundertprozentig präpositional (P). Wenn es also im Russischen ein Vordringen präpositionaler Wendungen und damit eine Veränderung zum Analytismus geben sollte, dann beginnt sie ohne Zweifel bei den seltener verwendeten Kasus.

NIKOLAEV schlüsselt den Kasus- und Präpositionsgebrauch auch nach den einzelnen Spezialbedeutungen (wie Genitiv des Subjektes/Objektes) auf, worauf hier jedoch nicht weiter eingegangen werden soll.

Daß der Kasusgebrauch nicht bei allen Substantiven gleichförmig ausgeprägt ist, sondern bestimmte Substantive vorwiegend in dem einen, andere Substantive hingegen eher in anderen Kasus gebraucht werden, das hat GREENBERG (1974) deutlich gemacht, indem er die Daten von ŠTEJNFEL'DT (1963) nach semantischen Klassen differenziert ausgewertet hat:

Gruppe	Nom.	Gen.	Dat.	Akk.	Instr.	Pröp.
Appelativa	28,3	26,0	5,0	21,8	8,6	10,3
Eigennamen	76,3	13,5	5,5	1,1	1,4	2,2
Personenbez.	54,1	22,5	6,9	7,0	8,0	1,5
Pers.kollektiva	23,9	48,0	4,2	9,6	6,2	8,3
Tierbez.	35,6	28,4	3,8	21,6	6,0	4,6
Körperteile	18,2	9,9	3,2	36,5	20,3	11,9
Gegenstände	23,0	20,7	4,3	32,0	9,4	10,5
Konkrete Materie	21,3	31,6	2,2	24,3	13,6	6,9
Zeitinstab. Obj.	34,5	19,0	4,1	21,5	10,8	8,8
Abstr. Eigensch.	33,3	24,9	3,8	17,4	12,3	9,0
Ortsbez.	11,8	30,6	6,0	24,4	3,3	23,8
Institutionen	13,0	40,9	2,3	17,8	1,8	24,1
Zeitangaben	12,8	37,5	2,0	36,0	3,4	8,3
Maßangaben	2,7	95,4	0,8	5,2	1,2	4,6
Pers.pron. 1./2.	51,8	21,8	14,5	9,5	2,2	0,2
Allgemein	33,6	24,6	5,1	19,5	7,8	9,4

Tab. 14: Kasushäufigkeit nach semantischen Gruppen

Zu beachten ist hier, daß sich die genannten Werte nur horizontal zu 100 aufsummieren, und daß die Klassen sich nicht wechselseitig ausschließen. Ein statistischer Test müßte zeigen, welche Daten signifikant von der allgemeinen Verteilung der Kasushäufigkeiten abweichen. Auffallend sind jedoch folgende Werte: Eigennamen werden primär im Nominativ gebraucht, offensichtlich um auf Personen zu referieren bzw. in der Funktion des Vokativs. Daneben kommt nur noch der Genitiv häufiger vor, wohl wegen seiner besitzanzeigenden Funktion. Körperteile kommen häufiger als normal im Akk. und im Instr. vor, offensichtlich als Objekt bzw. Instrument entsprechender Bewegungen und Handlungen. Orte und Institutionen werden häufig im Genitiv und – naheliegenderweise – im Präpositiv verwendet (Ortsangaben). Die Häufigkeit des Genitivs bzw. Akkusativs von Zeitangaben erklärt sich aus der Rektion der zusammen mit ihnen gebrauchten Präpositionen. Maßangaben stehen fast ausschließlich im Genitiv. Die Personalpronomen bevorzugen den Nominativ und, als einzige Klasse überhaupt, den Dativ.

Angaben zur Kasushäufigkeit von Substantiven und Adjektiven in Militärtexten finden sich bei KOLGUŠKIN (1970, 17).

Grammatische Kategorien des Verbs Die Häufigkeit, mit der die Verbkategorien Modus, Tempus und Aspekt in ihren einzelnen Bedeutungen benutzt werden, untersucht KOSAČEVA (1976) an Material aus der Umgangssprache (RR), wobei die Daten auf 19.200 Verbformen in einer Stichprobe von 120.000 laufenden Wörtern basieren. Neben Angaben zu einzelnen Verben stellt die Autorin ihre Daten zusammenfassend denjenigen aus ŠTEJNFEL'DT (1963) gegenüber, die ja auf der

Auswertung schriftlicher Quellen beruhen. Die Gegenüberstellung zeigt folgendes Bild:

Gramm. Kategorie	Gramm. Bedeutung	Umgangssprache	Schriftliche Texte
Modus	Indikativ	0,813	0,790
	Imperativ	0,051	0,040
	Konjunktiv	0,008	0,020
	Infinitiv	0,128	0,150
Tempus	Präsens	0,495	0,321
	Präteritum	0,405	0,565
	Futur	0,105	0,114
Aspekt	unvollendet	0,425	0,528
	vollendet	0,574	0,470
	zweiaspektig	0,001	0,002

Tab. 15: Relative Häufigkeit grammatischer Kategorien des Verbs

Es zeigen sich vermutlich signifikante Unterschiede in nur zwei Fällen: 1) Mündlich dominiert das Präsens über das Präteritum, schriftlich ist es – mit noch deutlicherem Abstand – genau umgekehrt. 2) Mündlich dominiert der vollendete Aspekt eindeutig, schriftlich ist der unvollendete etwas häufiger. Den ersten Punkt erklärt die Autorin mit dem häufigen Gebrauch einiger bestimmter Verben in Dialogen (*знать, говорить, думать, есть, мочь, понимать, работать, хотеть*). – Daten zum Tempusgebrauch in diversen Funktionalstilen des Russischen (Wissenschaftliche Prosa, Belletristik, Umgangssprache) finden sich auch bei SIROTININA (1974).

Verbale Konjugationsklassen Mit der Häufigkeit der Konjugationsklassen des russischen Verbs in einem Grundwortschatz der 2.500 häufigsten Wörter, darunter etwa 27% Verben, beschäftigt sich ausführlich KEIL (1965/66), der einleitend zugleich die Wichtigkeit statistischer Untersuchungen als Grundlage für die Ausarbeitung adäquaten Sprachlehrmaterials betont. Als Grundlage für die Zusammenstellung des Grundwortschatzes dienen KEIL dabei die Häufigkeitswörterbücher von JOSSELSOHN und ŠTEJNFEL'DT sowie dasjenige zu Puškin. Er ermittelt folgende Daten:

Konjugation	Anteil
e/o (-u, -eš'... nach Kons.)	0,21
je/jo (-ju, -eš'... nach Vok.)	0,40
i (-ju, -iš'...)	0,36
sonstige	0,03

Tab. 16: Relative Häufigkeit der Konjugationsklassen

Das heißt: die traditionelle e-Konjugation hat im Grundwortschatz einen Anteil von ca. 2 Dritteln, die i-Konjugation von einem Drittel. KEIL möchte freilich die e/o-Klasse von der je/jo-Klasse trennen und begründet dies mit den angegebenen Zahlen.

Aspekt Die Häufigkeit der Aspektformen im Grundwortschatz ist nach KEIL (1965/66) wie folgt gegeben:

Aspekt	Josselson	Štejnfel'dt	Puškin
imperfektiv	0,52	0,49	0,46
perfektiv	0,48	0,51	0,54

Tab. 17: Relative Aspekthäufigkeit

KEIL erwartet im Prinzip ein ausgeglichenes Verhältnis beider Aspekte zueinander. Das von den beiden übrigen Quellen abweichende Ergebnis bei Josselson erklärt er mit dessen Auswahl theoretischen Schrifttums und der Korrelation der je/jo-Klasse mit dem imperfektiven Aspekt. Die hier gegebenen Daten passen freilich nicht ganz zu den oben angeführten (s. **Tab. 15**): die eindeutige Dominanz des perfektiven Aspektes in der Umgangssprache findet nur in den Zahlen zu Puškin eine Bestätigung, während sonst tatsächlich ein eher ausgeglichenes Verhältnis gegeben ist.

Die Untersuchung des Zusammenhanges von Aspekt und Konjugationsklasse ergibt deutliche Zusammenhänge: unter den e/o-Verben sowie bei den Ausnahmen dominiert der perfektive Aspekte, umgekehrt unter den je/jo-Verben der imperfektive. Bei den Verben der i-Konjugation ist das Verhältnis eher ausgeglichen.

Aspekt	e/o-Konj.	je/jo-Konj.	i-Konjug.	Sonstige
imperfektiv	0,13	0,76	0,44	0,15
perfektiv	0,87	0,24	0,56	0,85

Tab. 18: Aspekt und Konjugationsklasse im Grundwortschatz

Ganz ähnliche Daten zu den Verbalkategorien finden sich auch bei GOLOVIN/URAMBAŠEV (1967); sie beruhen auf einer Diplomarbeit des Zweitautors. SEREBRJKOVA (1974) untersucht die Häufigkeit, mit der die verschiedenen Bedeutungen des unvollendeten Aspektes in den einzelnen Zeitstufen auftreten, wobei zusätzlich nach verschiedenen Funktionalstilen differenziert wird.

Mit den formalen Beziehungen zwischen Aspektpartnern beschäftigt sich JACHNOW (1971): er stellt ein Inventar der Mittel (Alternationen, Suffixgebrauch, Akzentverschiebungen) und ihrer Kombinationen auf, faßt diese zu Gruppen zusammen und belegt deren Ausnutzung im Russischen. Eine typologische Nutzung

der Daten wird ebenfalls kurz skizziert. DOLGOPOL'SKIJ (1963) systematisiert die Fälle, in denen der Gebrauch eines der beiden Aspekte vollständig bzw. nicht determiniert ist und kleidet seine Regeln in Form eines Algorithmus. Die Unmöglichkeit, in jedem Falle eindeutige Regeln zu finden, bezeichnet er als den "stochastischen Zusammenhang" zwischen Ausdruck und Inhalt. RATHMAYR (1976) untersucht Übersetzungsäquivalente anhand belletristischer Texte und gibt dabei den prozentualen Anteil der perfektiven Präsensform in jedem Strukturmuster an, vgl. auch die zusammenfassende Tabelle (1976, 174-176).

Imperativbildung Angaben zur Häufigkeit der Varianten der Imperativbildung in Abhängigkeit von den Stammklassen des Verbs finden sich bei BERGER (1981). Die Häufigkeiten werden statistisch bewertet, so daß erkennbar ist, zwischen welchen Variablen Assoziations- bzw. Dissoziationstendenzen bestehen.

Reflexive Verben Wovon das Auftreten reflexiver Verben, also von Formen auf *-ся*, in Texten abhängt, das untersucht BELOUSOVA (1969) auf der Grundlage von 37 Texten mit einer Gesamtlänge von 120.328 Wortformen, die zu 14.208 verschiedenen Lexemen gehören. Das Auftreten der Reflexivformen wird aufgeschlüsselt nach mehreren Faktoren: a) nach der Textlänge, b) nach der Gesamtzahl der Verbformen im Text, c) nach der Gesamtzahl der Verblexeme auf *-ся* im Text. Die Autorin interpretiert einzelne – natürlich schwankende – prozentuelle Anteile; statistische Tests werden nicht durchgeführt. Ihr – ziemlich banales – Ergebnis ist, daß der Gebrauch reflexiver Verbformen nach Autor und Textsorte verschieden ist (er schwankt zwischen 10 und 30% und liegt meist bei 20%, bezogen auf die Gesamtzahl der Verbformen). Die linguistische Fundierung der Datenerhebung bleibt fragwürdig: Wieso sollte z.B. das Auftreten reflexiver Verben überhaupt von der Textlänge beeinflusst werden? Bekannt ist, daß die Anzahl *verschiedener Lexeme* von der Textlänge unmittelbar abhängt. Eine Abhängigkeit der Anteile einzelner Wortarten und sogar grammatischer Klassen von dieser Größe ist jedoch schwer plausibel zu machen.

Verben der Fortbewegung Das Verhältnis der paarigen "Verben der Fortbewegung" zueinander untersucht BOBERTZ (1966) durch Auszählungen an modernen belletristischen Texten. Nach seinen Angaben dominieren die determinierten Verben generell mit einem Anteil von 3/4 aller Fälle.

Morphologische Variation Die ausführlichste Untersuchung und Darstellung zu Varianten der Norm in Morphologie (Flexion wie Wortbildung) und Morphosyntax (Kongruenz, Rektion) ist diejenige von GRAUDINA et al. (1976). Hier werden in Form eines Nachschlagewerkes häufige Variantentypen in standardisierter Form präsentiert: die Varianten (beispielsweise *директор пришла* – *директор пришел* [zur Bezeichnung von Frauen]), ihre absolute Häufigkeit (230/11) und ihr quantitatives

Verhältnis zueinander (95,43%/4,57%) bilden den "Kopf" des Eintrages. Darauf folgt ein Kommentar mit Angaben zum Gebrauch, zu erkennbaren Bedingungen und Regeln sowie, wo möglich, eine Liste aller Fälle. Den Eintrag schließen Literaturhinweise ab. Das Spektrum der untersuchten Varianten ist außerordentlich breit, die quantitativen Angaben wurden durch eigene statistische Auswertungen von Zeitungen im Gesamtumfang von 2 Mio. Wörtern gewonnen. Das praktische Vorgehen ist zudem durch theoretische Überlegungen abgesichert (GRAUDINA 1966, 1970). Das Wörterbuch ist damit ein unentbehrliches und in seiner Art für die Russistik innovatives Hilfsmittel bei der Einschätzung des aktuellen Sprachgebrauches der 60er und 70er Jahre, wobei sich das Spektrum der Varianten mit den von KRYSIN soziolinguistisch untersuchten oftmals deckt und so einen Vergleich der Ergebnisse erlaubt.

Komparativ MOROZOVA (1977) untersucht das Verhältnis von synthetischer zu analytischer Komparativbildung (*сильнее/сильней/посильнее* vs. *более сильный/более силен*). Ihre Stichprobe umfaßt 1150 Komparativvorkommen aus Gegenwartstexten (60er Jahre), die den drei Funktionalstilen Wissenschaft (Biologie), Publizistik und Belletristik im Verhältnis 500 – 400 – 250 entstammen. Die Daten zeigen einen krassen Unterschied in der Gebrauchshäufigkeit beider Bildungsweisen: Wissenschaft und Publizistik bevorzugen die analytischen Formen im Verhältnis 7:3; in der Belletristik dominiert die synthetische Form umgekehrt mit 8:2 (und in der Umgangssprache noch stärker). Die Wahl der einen oder anderen Form hängt dabei u.a. von der syntaktischen Position – attributiv (fast ausschließlich analytisch) oder prädikativ (eher synthetisch) – und deren Häufigkeit im Text ab, aber auch vom verwendeten Hilfsverb (wie *быть, казаться, стать, являться...*).

3.3.3. Derivationsmorphologie

Allgemein kann man sagen, daß Phänomene der Wortbildung und -ableitung bisher sehr viel seltener mit statistischen Methoden untersucht worden sind als andere Bereiche der russischen Morphologie, obwohl es gar nicht schwerfällt, "lohnende" Fragen zu formulieren (allgemein zum Thema vgl. TICHONOV 1983). Zu den bisher untersuchten Bereiche gehören u.a. die folgenden:

Ortsnamen АРАПОВ (1972) untersucht den Zusammenhang zwischen Ortsnamen und den davon abgeleiteten Bezeichnungen für die Bewohner des Ortes auf der Grundlage eines Lexikons, also in einer systemischen Stichprobe, und versucht, Regeln für diese Ableitungen zu ermitteln. Zur Modellierung der empirischen Zusammenhänge schlägt er ein recht komplexes mathematisches Modell vor.

Affixe Suffixe mit expressiver Funktion untersucht RAKUŠAN (1981). Die Autorin hat dem 17-bändigen Akademiewörterbuch eine (systemische) Stichprobe

von 2442 Basiswörtern und 3640 Ableitungen dazu entnommen, die in vier semantische Gruppen eingeteilt werden: (a) "neutral" in Bezug auf Emotionalität (*город – городок*), (b) "meliorativ" (*дядя – дяденка*), (c) "pejorativ" (*интеллигент – интеллигентик*), (d) "meliorativ und pejorativ" (*жених – женишок*). Sämtliche 3640 Ableitungen werden mithilfe von 66 Suffixen gebildet, deren Verwendungshäufigkeit und Zuordnung zu den genannten semantischen Gruppen genau aufgeschlüsselt wird. Einer Tabelle läßt sich entnehmen, welche semantische Funktion ein gegebenes Suffix vorzugsweise aufweist. So ist z.B. *-ишк-* fast immer pejorativ, *-ишч-* fast immer neutral; im Normalfall aber ist die Hauptfunktion "meliorativ". Unter den semantischen Gruppen ist die "meliorative" denn auch mit 65% die größte; ihr folgt die "neutrale" mit weiteren 20%.

Den Gebrauch der Präfixvarianten *о-* bzw. *об-/обо-* untersucht ROBERTS (1976) in Abhängigkeit vom Anlaut der Wurzel an Wörterbuchmaterial. DIETZE (1981) ermittelt aus einem Textkorpus die Verwendungshäufigkeit von Derivationsuffixen, die er – nach Wortarten und Funktion getrennt – mit ihrer Häufigkeit versehen alphabetisch auflistet.

Übersetzungsäquivalente Eine kontrastive Fragestellung liegt der Arbeit von NIKONOV (1962b) zugrunde. Er untersucht die polnischen Äquivalente russischer Adjektivsuffixe an je einer Stichprobe aus dem Lexikon und einer Turgenev-Übersetzung und stellt die Ergebnisse in Form einer zweidimensionalen Matrix dar. Ihr läßt sich entnehmen, wie häufig einem gegebenen russischen Suffix ein bestimmtes polnisches Suffix entspricht, welche Suffixe überhaupt infrage kommen usw. So entspricht z.B. im Lexikon einem russischen *-чек-* in 40 von 55 Fällen (d.h. in 73%) einem polnischen *-czn-*; im Text erhöht sich der Wert noch, und zwar auf 9 Fälle von 11, d.h. auf 82%. Kurz wird in dem Beitrag auch der Einsatz der Methode für historische und typologische Zwecke angedeutet; als Beispiel werden die Suffixhäufigkeiten dreier altrussischer Texte und diejenigen eines Turgenev-Textes und seiner Übersetzungen in vier slawische Sprachen zusammengestellt.

Neologismen Die Bildungsmuster russischer Neubildungen der Gegenwart untersucht KEMPGEN (1988) an empirischem Material. Derivationsstypen (wie Präfigierung und Suffigierung), Lehnvorgänge und Lehn Sprachen werden statistisch nach Wortarten aufgeschlüsselt und in ein zusammenhängendes Schema gebracht, aus dem hervorgeht, daß die großen Wortarten (Substantive, Adjektive, Verben) in erster Linie sich selbst reproduzieren, während die Derivationsbeziehungen, die mit einem Wechsel der Wortart verbunden sind, demgegenüber in den Hintergrund treten.

Adverbien Die Wortbildungsmuster von Adverbien untersucht GOLUBZOWA und gibt hierzu auch Daten zur Häufigkeit einzelner Derivationsmuster an (vgl. zusammenfassend 1991, 262–269); 80% aller Fälle machen dabei allein die erstarrten

Kasusformen mit und ohne Präposition aus (wenn man so will, zwar Wortbildung, aber nicht Derivation im engeren Sinne).

3.4. Lexikologie

Im Bereich von Lexikographie und Lexikologie ist zuallererst natürlich an die Ausarbeitung von Frequenzwörterbüchern zu denken. Solche Wörterbücher gehören inzwischen zu einem gut bearbeiteten Bereich der russischen Sprache, und zwar in praktischer wie in methodisch-theoretischer Hinsicht (FRUMKINA 1964a, 5ff., ALEKSEEV 1975, 1984). Manche Autoren verstehen unter einer "vyčislitel'naja leksikografija" allerdings in einem umfassenderen Sinne einfach eine computergestützte Lexikographie (vgl. etwa die Arbeiten von ANDRJUŠČENKO).

Die Ausarbeitung von Häufigkeitswörterbüchern war von Anfang an ausgesprochen anwendungsbezogenen motiviert, nämlich durch die spätestens seit dem Ende der 50er Jahre immer stärker empfundene Notwendigkeit, Wortschatzminima für den Russisch-Unterricht in den nicht muttersprachlichen Teilen der Sowjetunion bzw. im Ausland zusammenzustellen (s. FRUMKINA/ŠTEJNFEL'DT 1960, MARKOV 1960, 1964, 1965/66, MUCKOV 1964, SMIRNOVA 1958, VOLOCKAJA et al. 1958, KOSTOMAROV 1963).

Ironischerweise stammt das erste große Frequenzwörterbuch zum Russischen aus dem Land des Hauptgegners im "Kalten Krieg" nach dem Zweiten Weltkrieg (JOSSELSOHN 1953), nachdem eigene frühere Ansätze in Gestalt des Buches von ČISTJAKOV/KRAMARENKO (1929) weitgehend vergessen, jedenfalls aber folgenlos geblieben waren. Das erste moderne sowjetische Frequenzwörterbuch zum Russischen kam aus dem Baltikum (ŠTEJNFEL'DT 1963) und war durch fremdsprachendidaktische Erfordernisse motiviert. Erst mehr als zwei Jahrzehnte nach JOSSELSOHN konnte mit ZASORINA (1977) ein in Rußland erarbeitetes großes Frequenzwörterbuch zum Russischen vorgelegt werden. Dieser chronologischen Reihenfolge entspricht die nachfolgende Darstellung.

3.4.1. Allgemeine Frequenzwörterbücher

1. JOSSELSOHN (1953) stellt das erste große Häufigkeitswörterbuch zum Russischen dar; seine Charakteristika sind oft genug behandelt worden, weshalb die Darstellung hier knapp erfolgen kann.

Bei der Erhebung der Stichprobe stellte die statistische Zuverlässigkeit der Daten ein besonderes Ziel dar. Offenbar wurden hier erstmals in der Geschichte der Frequenzwörterbücher überhaupt Computer eingesetzt (so FRUMKINA 1964a, 14).

Insgesamt wurden aus 148 Prosa-Texten mit einer Gesamtlänge von ca. 1 Mio. Wörtern Stichproben im Umfang von 506.044 Wörtern ausgewertet, die sich zu 41.115 verschiedenen Lexemen gruppieren. Von diesen wiederum wurden 5230 in das vorliegende Wörterbuch, verteilt auf sechs Listen, aufgenommen, wobei die weniger häufigen zusätzlich von Russischlehrern auf ihre Relevanz für die aktive

bzw. passive Sprachkompetenz bewertet wurden (ein entsprechender Index zeigt dies bei den einzelnen Wörtern an).

Die Häufigkeitsangaben werden für mehrere Perioden (19. Jh., 1900-1918, seit 1918 bis zur Gegenwart), für dialogische wie monologische Rede und für verschiedene Genres (Publizistik, Belletristik, Wissenschaft und Sachtexte) getrennt aufgeführt. Die Gewichtung der drei Perioden (mit 25% – 25% – 50%) scheint plausibel, berücksichtigt die klassische russische Literatur aber zu wenig. Alle Quellen sowie weitere Charakteristika der Zusammensetzung des Materials werden genau benannt. Vor den Häufigkeitslisten findet sich eine längere Einleitung, in der es u.a. um die stilistischen Charakteristika einzelner Autoren sowie um die Häufigkeitsverteilung grammatischer Phänomene geht.

Was diesem Wörterbuch aus heutiger Sicht bemerkenswerterweise fehlt, ist eine einfache Rang-Frequenz-Liste. Da die sechs Teil-Listen in sich alphabetisch geordnet sind, wird ihnen eine alphabetische Liste aller Lexeme beigelegt, aus der bloß hervorgeht, in welcher Liste ein gegebenes Wort zu finden ist.

Die Aufnahme dieses ersten Frequenzwörterbuches zum Russischen durch die Rezensenten (DOBROSMYSLOV 1957, NOVAK/PIOTROVSKIJ 1958) war im Prinzip positiv, jedoch wurden auch die methodischen Mängel der Untersuchung deutlich benannt. Die Bedeutung dieses Wörterbuches (und die Schwierigkeit, es in der Sowjetunion zu bekommen) läßt sich u.a. auch daraus erkennen, daß ein Teil der Frequenzlisten ohne großen Kommentar in russischen Zeitschriften nachgedruckt wurde (vgl. SMIRNOVA 1958, MARKOV 1960c).

2. ŠTEJNFEL'DT (1963) stellt eine Pionierleistung auf dem Gebiet der Frequenzwörterbücher in der Sowjetunion dar und wurde auch in einer Übersetzung in mehrere westliche Sprachen publiziert (dt. ŠTEJNFEL'DT o.J. [ca. 1966], engl. ŠTEJNFEL'DT 1966).

Seine Erstellung sollte insbesondere die Ausarbeitung eines Grundwortschatzes zum Russischen für den Fremdsprachenunterricht in der Sowjetunion ermöglichen (vgl. ŠTEJNFEL'DT 1961; ROOVET/ŠTEJNFEL'DT 1965). Auch wenn mit ZASORINA (1977) inzwischen ein repräsentativeres Frequenzwörterbuch vorliegt, bleibt das ŠTEJNFEL'DTsche Wörterbuch aufgrund seiner vielen zusätzlichen Auswertungen ein unentbehrliches Werkzeug und nützliches Hilfsmittel, dessen Angaben vielen sprachstatistischen Analysen zugrundegelegt haben. Es hat auch deshalb nichts an Aktualität eingebüßt. Die Auswertung der Texte mit einem Gesamtumfang von 400.000 Wörtern erfolgte hier noch vollständig manuell. Die Repräsentativität des Wörterbuches, dessen Erarbeitung durch begleitende Studien eine theoretische Fundierung erfuhr (FRUMKINA 1959a, 1959b, 1963b, FRUMKINA/ŠTEJNFEL'DT 1960), wurde anschließend an nicht zur ursprünglichen Stichprobe gehörenden Texten überprüft, was zufriedenstellende Ergebnisse ergab (vgl. auch FRUMKINA 1964a, 21f.). Mehrere Textbeispiele, denen man die erreichte Textdeckung plastisch ansehen kann, werden zudem abgedruckt (vgl. 53–58).

Nach ihrer Häufigkeit geordnet erscheinen die Wörter in der "Allgemeinen Häufigkeitsliste" (63–93), in der zu jedem Wort seine Verbreitung (verstanden als die Zahl der Texte, in denen es vorkam) und seine Vorkommenshäufigkeit angegeben wird. Aufgeführt sind alle Lexeme, die 14 mal oder öfter vorkamen, insgesamt 2.500, also im Grunde ein etwas erweiterter Grundwortschatz. Die Angaben zu Häufigkeit und Verbreitung tauchen auch in den nachfolgenden Listen jeweils wieder auf. Die gleiche Liste erscheint später noch einmal, hier jedoch alphabetisch geordnet (177–209), was das Nachschlagen der Werte eines konkreten Wortes sehr erleichtert.

Von außerordentlicher Wichtigkeit sind auch die nach Wortarten getrennten Wortlisten (95–176), und zwar deswegen, weil hier zu jedem flektierten Wort Angaben über die Häufigkeit einzelner grammatischer Kategorien bzw. Bedeutungen angegeben sind. Leider sind diese Daten nicht immer detailliert genug aufgeschlüsselt: So findet man bei den Substantiven beispielsweise Angaben über die Häufigkeit von Singular, Plural und jedes Kasus, aber nicht über deren Kombination, also keine Angabe über Nominativ Singular, Genitiv Singular etc. Bei den Adjektiven und Partizipien wird entsprechend nur nach der Häufigkeit von Langformen und Kurzformen unterschieden. Bei den Verben hingegen kann man der Tabelle die Häufigkeit jeder einzelnen (finiten) Form entnehmen

Im Anhang des Wörterbuches (210–233) finden sich Angaben über die Rektion der Verben: zu jedem Verb ist seine Gesamthäufigkeit angegeben, die weiter aufgeschlüsselt wird nach dem Gebrauch mit konkreten Kasus bzw. Präpositionen, die jeweils einen bestimmten Kasus erfordern. Als Beispiel mag das Wort *вѣрять* dienen, dessen Häufigkeit 74 beträgt. 27 mal kam das Verb mit dem Dativ vor, 23 mal dagegen mit *с* + Akk. Intransitiv wurde es folglich 24 mal benutzt. Von großem Nutzen sind auch die vielen zusätzlichen Auswertungen zu Einzelaspekten der russischen Grammatik (21–52), auf die wir hier gar nicht im einzelnen eingehen können (einige Resultate finden sich jedoch an den entsprechenden Stellen des vorliegenden Überblickes).

Die Aufnahme des Wörterbuches würdigte seine Leistungen entsprechend (vgl. EŠAN/ALEKSEEV 1964), wies jedoch einschränkend besonders auf die Tatsache hin, daß in der Stichprobe die für Kinder gedachten Texte weit überrepräsentiert sind, was u.a. dazu führte, daß Wörter wie *ребята*, *пионер*, *мама*, *мальчик*, *девочка* hier zu den häufigsten Substantiven zählen.

3. Das Standardwerk unter den Häufigkeitswörterbüchern stellt das Lexikon von ZASORINA (1977) dar. Es umfaßt mit ca. 40.000 Wörtern knapp die Hälfte eines "normalen" Wörterbuches (ca. 100.000 Wörter) und hat zum Ziel, den "aktiven Wortschatz eines Gebildeten der Gegenwart" (6) zu erfassen. Insgesamt wurde Text im Umfang von über 1 Million Wörtern ausgezählt. Das Material hierzu wird in vier Funktionalstile gegliedert (Zeitungen, Dramen, Wissenschaft und Publizistik, belletristische Prosa), die im Unterschied zu den beiden anderen Frequenzlexika des

Russischen an der Stichprobe einen fast gleichen Anteil haben; die verwendeten Quellen werden bibliographisch exakt genannt.

Den Hauptteil des Wörterbuches nimmt eine alphabetische Liste der 39.268 angetroffenen Lexeme ein, in der zu jedem Eintrag seine allgemeine (absolute) Frequenz in der Gesamtstichprobe von 1.056.382 Wörtern aufgeführt ist, sowie seine absolute) Frequenz in jedem der vier Funktionalstile. Außerdem findet sich zu jedem Wort eine Angabe, in wievielen Texten jeder Sub-Stichprobe es zu finden ist. Mit diesen Angaben läßt sich bewerten, wie gleichmäßig verbreitet bzw. wie spezialisiert ein Wort in Bezug auf die Funktionalstile ist. Zur Bewertung der Gebräuchlichkeit (*употребительность*) eines Wortes ist die zweite Angabe wichtig: ein Wort kann häufig sein, im Extremfall aber nur in einem einzigen Text vorkommen. So hat z.B. *телефонить* eine Frequenz von immerhin 9, aber alle Vorkommen stammen aus einem einzigen Text! (Eine theoretische Untersuchung des Zusammenhanges von Häufigkeit und Gebräuchlichkeit unternehmen VANNIKOV/MASLOV [1969a, 1969b] an russischem Material.)

Die Frequenz-Liste führt alle Wörter geordnet nach ihrer Häufigkeit auf, die mindestens 10 mal vorkamen, wodurch sich der Umfang gewaltig reduziert, nämlich auf nur noch 9044 Wörter. Zusammenhänge wie der zwischen Häufigkeit und Länge (häufige Wörter sind kurz, seltene können ruhig lang sein) lassen sich aus dieser Liste durch bloßen Augenschein ablesen.

Aufzupassen gilt es bei den Pronomen: Hier werden Formen wie *я* und *мы* sowie *он*, *она*, *оно*, *они* als separate Lexeme aufgelistet. Die Einträge *его*, *ее* und *их* sind jedoch die Possessivpronomen, nicht die Kasusformen der Personalpronomen. Angaben zur Häufigkeit einzelner grammatischer Formen sind also nicht vorhanden; gelegentliche grammatische Doppeldeutigkeiten wie bei *его* (masc. und neutr.) werden nicht aufgelöst.

Der dritte Teil ist eine Rang-Frequenzverteilungstabelle. Hier werden die Frequenzen aller Wörter bis zu einem gegebenen Rang in absoluten und relativen Zahlen aufsummiert. Hieraus läßt sich ersehen, daß beispielsweise mit den 10 häufigsten Wörtern mehr als 18% aller Vorkommen in der Stichprobe erfaßt sind, mit den 230 häufigsten 50% aller Fälle, die 9000 häufigsten Wörter der zweiten Liste decken durchschnittlich 92% eines beliebigen Textes ab usw. Solche Auswertungen braucht man einerseits zur Erstellung von Wortschatzminima, das Anwachsen der "Textdeckung", die mit einer gegebenen Anzahl von x häufigsten Wörtern erreicht wird, wird von der QL jedoch auch als eigene Fragestellung untersucht und weist eine charakteristische Kurve auf.

Im Anhang des Wörterbuches finden sich noch verschiedene Statistiken und Graphiken, leider nicht immer in der wünschenswerten Form – wichtige Verteilungen, wie etwa die Verteilung der Wortlänge im Text und im Wörterbuch, werden nur als Kurven dargeboten, die numerischen Werte fehlen, können also auch nicht für weitere Überprüfungen herangezogen werden, z.B. für die Frage, ob sich die Funktionalstile signifikant voneinander unterscheiden oder nicht.

Zu Rezensionen dieses Wörterbuches vgl. u.a. WENZEL (1979) und TORNOW (1980).

4. Das neueste Häufigkeitswörterbuch zum Russischen stammt von LÖNNGREN (1993). Es beruht auf einem selbst erhobenen, umfangreichen Korpus, dem sog. "Uppsala-Korpus", das der Forschung in computerlesbarer Form zur Verfügung steht. Es umfaßt 1 Mio. laufende Wortformen und wurde in den Jahren 1987 bis 1991 erarbeitet. Die ca. 600 ausgewerteten Texte stammen aus Büchern, Zeitschriften und Zeitungen der 60er Jahre oder später. Nicht publiziert, aber ebenfalls vorhanden und zugänglich, ist eine vollständige Konkordanz, aus der sich der Kontext einer jeden Wortform genau ersehen läßt.

Das Wörterbuch enthält u.a. eine Rang-Frequenzliste der häufigsten Lexeme, der häufigsten Wortformen, der fachsprachlichen Lexik, der belletristischen Lexik, und, als wichtigsten Bestandteil, ein alphabetisches Wörterverzeichnis, in das alle 9.281 Lexeme, die öfter als neunmal auftraten, aufgenommen wurden.

Das "Wörterbuch" wirkt insgesamt eher als Basis-Dokumentation zu dem genannten Uppsala-Korpus, indem es in seinem einleitenden Teil ausführlich auf Fragen der Lemmatisierung eingeht, die Quellen benennt, die Computer-Dateien beschreibt usw., während die verschiedenen Wortlisten, die ja das eigentliche Frequenzwörterbuch darstellen, hier als bloßer "Anhang" deklariert sind.

Zu dem vorliegenden Frequenzwörterbuch vgl. auch die – fast durchgehend positive – Rezension von KEIL (1994), in der LÖNNGRENS Resultate stichprobenartig mit denjenigen von ZASORINA verglichen werden, u.a. um zu prüfen, ob die von dem neueren Wörterbuch angestrebte Aktualität irgendwelche spürbaren Auswirkungen auf die gewonnenen Resultate hat (möglicherweise).

Das Wörterbuch von LÖNNGREN kann sich vom Umfang des publizierten Materials her nicht mit demjenigen von ZASORINA messen, ersetzt es also nicht, sondern ergänzt es. Der eigentliche Wert liegt mehr in der Kombination von Häufigkeitswörterbuch und allgemein verfügbarem Korpus, zu dem es eine erste Auswertung darstellt.

5. Ein Frequenzwörterbuch ist auch in OLIVERIUS (1976) enthalten, wobei die Daten von ŠTEJNFEL'DT (1963) übernommen wurden. Ein gewichtiger Unterschied liegt allerdings in der Tatsache, daß bei OLIVERIUS alle Einträge phonologisch notiert, mit zusätzlichen Markern versehen und in Morpheme zerlegt sind. Das Material wird präsentiert in Form einer Rang-Frequenzliste der 2.500 häufigsten Wörter des Russischen (pp. 71–102) und als alphabetische Liste in orthographischer Notation (pp. 103–119) – diese Liste ist eine einfache Übernahme aus ŠTEJNFEL'DT (vgl. 1966, 179–209), wobei die Information über die Verbreitung sogar noch weggelassen wird. Die eigene Leistung OLIVERIUS' steckt in der Frequenzliste auf Morphemebene, die aus der ŠTEJNFEL'DTschen Liste gewonnen wurde (dazu s.o. unter Morphologie).

MORKOVKIN (1985, vgl. auch 1972) enthält im Anhang einen tabellarischen Vergleich der jeweils 2.500 häufigsten Wörter aus 8 verschiedenen Häufigkeitswörterbüchern des Russischen, wobei jeweils die Hundertergruppe angegeben ist, zu der das Wort seinem Rang nach gehört. Dies ermöglicht einen ungefähren Eindruck von der Stabilität der Ergebnisse. Das eigentliche Wörterbuch ist eine Kompilation der Ergebnisse von insgesamt 12 Frequenzlisten; das Material wird in verschiedenen Listen alphabetisch und mit grammatischen Informationen versehen präsentiert.

3.4.2. Grundwortschatz und Lehrbücher

Grundsätzliche methodische Fragen wie Textauswahl, Stichprobenlänge, notwendigen Umfang und lexikalische Systematik behandeln etwa FRUMKINA (1959a, 1959b), MARKOV (1960a) – mit illustrativem Material aus JOSSELSOHN und PUŠKIN – sowie FRUMKINA/ŠTEJNFEL'DT (1960), diese als Vorbereitung für ŠTEJNFEL'DT (1961, 1962, 1963). Unberücksichtigt bleiben hier alle Arbeiten, die, wie z.B. ŠANSKIJ (1982), lediglich einen Grundwortschatz ohne quantitative Angaben enthalten.

Spontane Kindersprache von Schülern der 1. bis 3. Klasse hat CHARAKOZ (1971) ausgewertet. Aufgezeichnet wurden – an russischen Schulen in Kirgisien – Gespräche im Umfange von 312.000 Wörtern, denen 8.568 Lexeme zugrundelagen. Von diesen werden in dem Wörterbuch nur die 2.830 gebräuchlichsten (die jedoch für 96,6% der Gesamtstichprobe stehen!) in Form einer Rang-Frequenzliste und alphabetisch aufgeführt.

MARKOV (1966) versucht der Frage nachzugehen, ob belletristischer und mündlicher Wortschatz übereinstimmen oder nicht, was bedeutende Konsequenzen für die Erstellung von Lehrmaterialien für das passive Lesen bzw. aktive Sprechen hätte. Zu diesem Zweck vergleicht er die Daten mehrerer Frequenzwörterbücher miteinander, in erster Linie MARKOV/VIŠNJAKOVA (1965) mit ŠTEJNFEL'DT (1963). Der Autor gibt eine Übereinstimmung von 2/3 des Grundwortschatzes an, wobei jedoch einzelne Wortarten in den verglichenen Quellen einen unterschiedlichen Anteil aufweisen können. Substantive etwa sind in einer Text-orientierten Zählung deutlich häufiger als in der gesprochenen Sprache. Der wichtigste Unterschied liegt jedoch in der Textdeckung, die mit den x häufigsten Wörtern erreicht wird: Hier kommt die Umgangssprache mit sehr viel weniger Wörtern aus als die Sprache der Literatur, weist also ein geringeres Vokabular auf. Um 80% eines Textes abzudecken, benötigt die Literatursprache ca. 2.500 Wörter, die Umgangssprache jedoch nur 1.200 Wörter. – Die nichtkanonische Lexik aus Alltagsbriefen ist Gegenstand der Arbeit von ALEKSEEV/GRIGOR'EVA (1981).

KELLER (1991) vergleicht die Daten von Frequenzwörterbüchern mit dem Vokabular, das in Lehrbüchern zu finden ist, und kommt zu dem Schluß, daß das "pädagogische Vokabular" oft genug nicht in Übereinstimmung mit den Häufigkeitszählungen steht. Seine konkrete Empfehlung an die Autoren von Lehrbüchern lautet, nach jeweils ca. 5 Lektionen eine themenorientierte Zusammenfassung des Vo-

kabulars zur Wiederholung und Systematisierung einzubauen. Zu einer ähnlichen Analyse vgl. auch ASCHENBRENNER (1964) und KEHRER (1965/66). Eine umfassende Analyse ausgewählter Russisch-Lehrbücher nehmen SCHMITZ/ZÜHLKE (1974) vor; zu den Kriterien, die sie verwenden, gehört auch der Grad der Übereinstimmung des Lehrbuchwortschatzes mit Häufigkeitswörterbüchern. Diverse Häufigkeitswörterbücher stellen auch v. FISENNE/WERTHER (1975) vergleichend vor.

3.4.3. Fachsprachliche Frequenzwörterbücher

Umgangssprache An der Lumumba-Universität in Moskau wurde ein Häufigkeitswörterbuch der Umgangssprache erarbeitet, das in verschiedenen Publikationen und unter verschiedenen Titeln auftaucht. Eine vorläufige Liste der 1.200 häufigsten Wörter veröffentlichten MARKOV/VIŠNJAKOVA (1965). Sie beruht auf 234 Gesprächen mit einer Gesamtlänge von 100.000 Wörtern, die 1964/65 in Moskau aufgezeichnet wurden. Die Lexeme werden alphabetisch unter Angabe ihrer Häufigkeit aufgelistet, bei Homonymen finden sich grammatische Hinweise. MARKOV (1967) und MARKOV/VIŠNJAKOVA (1968) umfassen die vollständigeren Listen von 1.500 bzw. 2.380 häufigsten Wörtern. Diese letzte Liste ist u.a. von NOWIKOWA (1972) nachgedruckt worden

TURKO (1968) führt in Form einer Rang-Frequenzliste die 1.172 Wortformen auf, die in einer Stichprobe von 50.000 Wörtern nicht weniger als fünf mal vorkamen. Als Material dienten Tonbandaufzeichnungen von Spontansprache (zwei Beispieltexte werden zur Illustration mit abgedruckt). Es sei betont, daß hier tatsächlich Wortformen, nicht Lexeme, ausgewertet wurden. Homonyme werden durch grammatische Angaben identifiziert, Polyseme tragen den gleichen Index wie im Wörterbuch von OŽEGOV.

Die Sprache der Privatkorrespondenz untersuchen ALEKSEEV et al. (1973) sowie GRIGOR'EVA (1980, 1983), wobei die letztgenannte Arbeit im wesentlichen eine Liste der 1.045 häufigsten Wörter dieses Genres ist (*я, и, в, не, ты, быть...*). Daß die Personalpronomen der 1./2. Person des Singulars in diesem Kontext besonders häufig sind, erstaunt natürlich nicht.

Elektronik Die 992 häufigsten Wortformen mit einer Frequenz von mindestens 30 in fachsprachlichen Texten zur Elektronik findet sich bei KALININA (1968b). Der kurze Artikel besteht nur aus einer Rang-Frequenz-Tabelle.

Presse, Zeitung Das erste Häufigkeitswörterbuch zum Russischen überhaupt wurde in der Tschechoslowakei unter Leitung von MALÍŘ (1951) publiziert, ist jedoch ausdrücklich auf das fremdsprachliche Übersetzen ausgerichtet. Ausgewertet wurde ein Korpus von 113.032 laufenden Wörtern von Zeitungstexten aus dem Jahre 1948. GURČENKOVA (1972) wertet ein Korpus von 400.000 Wörtern von Zeitungstexten der Jahre 1967-70 aus. Neben einer Rang-Frequenz-Liste mit den 500 häufigsten Wörtern finden sich verschiedene kleinere Auswertungen des Materials

(Häufigkeit der Wortarten – geringer Verbanteil, Verbformen ausschließlich in der 3. Person und vorwiegend im Präteritum, Kasus-Frequenz – Dominanz des Nominativs usw.).

Das umfangreichste Frequenzwörterbuch zur Zeitungssprache ist dasjenige von POLJAKOVA/SOLGANIK (1971). Ausgewertet wurden hier "Pravda"- und "Izvestija"-Texte der Jahre 1960–69 mit einem Gesamtumfang von 200.000 Wörtern. Von den fast 15.000 verschiedenen Lexemen, die in dieser Stichprobe auftraten, werden die 2.000 häufigsten in Form einer Rang-Frequenzliste und einer alphabetischen Liste publiziert. Außerdem enthält das Wörterbuch entsprechende separate Listen für jeden von zehn thematischen Bereichen (wie z.B. "Ideologie", "Moral", "Landwirtschaft", "Sport" usw.).

Wissenschaftliche und technische Lexik Das Wörterbuch von DENISOV et al. (1978) vereinigt die Daten und Ergebnisse zweier Vorläufer, nämlich STEPANOVA (1970) und SAF'JAN (1971), deren Zählungen jeweils auf 200 Stichproben à 2.000 laufenden Wörtern, mithin insgesamt 400.000 Wörtern, beruhen. In das Lexikon aufgenommen wurden alle Wörter, die mindestens 10 mal auftraten. In den beiden Quellen waren dies 2.123 bzw. 2.393 Lexeme, bei DENISOV sind es durch Vereinigung beider Mengen 3.047 Wörter. Die Publikation des "kombinierten Wörterbuches" wird vor allem mit fremdsprachendidaktischen Gründen motiviert.

Die ermittelten Lexeme werden in Form einer Rang-Frequenz-Tabelle sowie in einer alphabetischen und zweier rückläufiger Listen aufgeführt, wobei die alphabetische Darstellung zusätzlich Angaben zur Wortart und zum Genus (Subst.) bzw. zum Aspekt (Verben) enthält. Die Häufigkeiten werden ferner getrennt nach den beiden Quellen angegeben, zudem die Zahl der Texte, in denen diese Wörter vorkamen, woran sich die allgemeine Verbreitung eines Lexems beurteilen läßt.

An diesen Hauptteil des Lexikons schließt sich ein "synoptischer Teil" an, der in mehr als 30 Tabellen einzelne Aspekte zusammenfaßt und auswertet. So wird der Lexembestand aufgeschlüsselt nach den Wortarten und grammatischen Kategorien, nach den ersten drei Anfangs- bzw. Endbuchstaben, es findet sich eine Verteilung der Wortarten auf den Auslaut, eine Verteilung der Wörter nach ihrer Silbenzahl, Statistiken zum Stammauslaut in den einzelnen Wortarten usw. Alle Zahlen werden in absoluter Form und als Prozentzahlen angegeben, weitere Auswertungen oder Bewertungen erfolgen nicht. Die Daten wurden rein mechanisch ermittelt, ein linguistisches Interesse an einer einzelnen Statistik wird nicht formuliert und ist auch nicht immer einsichtig. Ein gewisser Teil der Zählungen ist aber von nachvollziehbarem Nutzen.

Lehrbücher 15 Lehrbücher für die Klassen I-IV des Jahres 1964 in einer Gesamtlänge von 408.466 Wörtern hat CHARAKOZ (1971) ausgewertet und zu Rang-Frequenzlisten sowie alphabetischen Listen verarbeitet. Der Wortschatz umfaßt hier 19.250 Lexeme, von denen die 5.025 gebräuchlichsten abgedruckt werden (sie

sind für 91,6% aller Wörter im Text verantwortlich). Dieser Lehrbuchwortschatz wird in getrennten Listen auch noch einmal nach zwei verschiedenen Sachgebieten aufgeschlüsselt (Arithmetik, Naturkunde).

Militärtexte In dem Büchlein von KOLGUŠKIN (1970) finden sich u.a. eine alphabetische Liste der 3.000 häufigsten Lexeme in Militäertexten sowie eine Rang-Frequenzliste der 1.000 häufigsten Wörter der gleichen Sub-Sprache (Rang 4: *рубеш*, Rang 7: *в направлении*, Rang 8: *противник*)

3.4.4. Mehrsprachige Häufigkeitswörterbücher

Diese Kategorie wird hauptsächlich von Lexika vertreten, die in den 70er Jahren unter der Leitung von L. HOFFMANN an der Universität Leipzig erarbeitet wurden und dort auch unter seiner Herausgeberschaft erschienen sind. Bearbeitet wurden die folgenden Bereiche: *Medizin* (1970), *Physik* (1970), *Chemie* (1973b), *Bauwesen* (1976), *Mathematik* (1976), *Tierproduktion Veterinärmedizin* (1978).

Diese Lexika enthalten die jeweils ca. 1.100 häufigsten Lexeme der untersuchten Fachsprachen, mit denen eine Textdeckung im fachwissenschaftlichen Bereich von durchschnittlich mindestens 80% erreicht werden soll. Den Wörterbüchern geht es also in erster Linie um eine praktische Anwendbarkeit bei der Lektüre von Fachtexten in den drei wichtigsten Fremdsprachen. Sie präsentieren ihr Material – für jede Sprache getrennt – in Form einer Rang-Frequenz-Liste mit Angabe der relativen Häufigkeiten sowie eines alphabetischen zweisprachigen Verzeichnisses Fremdsprache – Deutsch, in dem die Angaben über Rang und Frequenz wiederholt werden. Hinzu kommt eine kurze Liste häufiger Suffixe und Präfixe mit illustrierenden Beispielen, die vor allem dazu dienen soll, die Sprachstruktur einseitiger zu machen. – Vgl. auch LUK'JANENKOV/SERGEEVA (1978).

3.4.5. Einzelne Werke bzw. Autoren

Die echten Frequenzwörterbücher haben einen Vorläufer in Wortverzeichnissen (*slovoukazatel'*), bei denen bei jedem Wort seine Belegstellen in einem oder mehreren literarischen Werken genannt werden. Die Häufigkeit ist hier indirekt enthalten: sie kann über die Zahl der Belegstellen ermittelt werden. Für das Russische ist KUNICKIJ (1894) ein frühes Beispiel, der einen Griboedov-Text auf diese Weise aufschlüsselt.

Puškin Der Begründer der modernen russischen Literatursprache ist der einzige russische Autor, dessen Gesamtwerk vollständig lexikostatistisch erfaßt und ausgewertet ist. Die vier Bände des "Slovar' jazyka Puškina" (Moskva 1956-61) enthalten bereits zu jedem Wort seine Gebrauchshäufigkeit, bieten also das Rohmaterial für statistische Auswertungen. Diese wurden unter der Leitung von R.M. FRUMKINA vorgenommen und in mehreren Arbeiten publiziert (vgl. 1960a, 1963a, 1964a,

48ff., 94ff., SIJaP 1956-61). Die Daten zur statistischen Struktur der Puškischen Lexik werden seitdem in der Literatur vielfach zitiert.

Puškin hat in seinen Werken nach den genannten Quellen 544.777 laufende Wörter gebraucht, denen 21.197 verschiedene Lexeme zugrundeliegen, was einen sehr hohen aktiven Wortschatz darstellt. Die 2.000 häufigsten Lexeme sind bei FRUMKINA (1963b, Tab. 1) publiziert. Von allen Lexemen hat der Autor zwar 30% (genau: 6.388 Wörter) nur ein einziges Mal benutzt, aber genau solche Wörter sind es natürlich, die den Reichtum seines Wortschatzes ausmachen.

Umgekehrt stehen die 56 häufigsten Lexeme für 40% des Gesamttextes, etwas mehr als die Hälfte aller Fälle (52%) bestreiten die 200 häufigsten Lexeme – im übrigen nur damals stilistisch neutrale –, und mit dem "Puškin-Grundwortschatz" seiner 1.000 häufigsten Wörter kann man im Durchschnitt 70% eines Textes abdecken. Mit den publizierten 2.000 häufigsten Wörtern deckt man im Durchschnitt 80% ab, um aber eine Textdeckung von 95% zu erreichen, bedarf es bereits 8.000 Lexeme (vgl. FRUMKINA 1963a, Tab. 3).

Diese Zahlen sind denjenigen, die für das Russische insgesamt gelten (ZASORINA 1977), sehr ähnlich. Die Tatsache, daß sie in der Gesamtsprache um 3-4 Prozent niedriger liegen, bestätigt die naheliegende Vermutung, daß man in einem thematisch eingeschränkteren Bereich etwas schneller (heißt: mit weniger Lexemen) eine gleich große Textdeckung erreicht. Ob die Unterschiede zwischen diesen beiden Verteilungen jedoch signifikant sind oder nicht, müßte ein statistischer Test erst noch zeigen.

Als Beweis für die Behauptung, daß PUŠKIN zum modernen Russisch zu rechnen ist, dient die Tatsache, daß von seinen 500 häufigsten Wörtern heute nur knapp 10% als veraltet (im Sinne von "heute weniger gebräuchlich") gelten müssen, oft aufgrund äußerer gesellschaftlicher Veränderungen (*царь, граф, князь, государь*). PUŠKINs Sprache wirkt also "modern", weil dieses halbe Tausend häufigster Wörter, mit dem bereits 60% des Textes erfaßt wird, noch modern ist.

Ein Vergleich der Rangfrequenzverteilung bei PUŠKIN mit den Voraussagen nach dem ZIPFSchen Gesetz findet sich bei FRUMKINA (1964a, 31, 34, 39) – die Übereinstimmung ist groß. – Ausgewertet wurde der Puškische Wortschatz ferner nach den vier Bereichen Poesie, Prosa, Publizistik und Briefe; aus jedem Genre sind die 220 häufigsten Wörter publiziert (FRUMKINA 1963a, Tab. 4-7). Daten zur Häufigkeit der Wortarten bei Puškin finden sich bei ŠULJAK (1976).

Lenin Ein Kuriosum stellt das SLOVAR' JAZYKA V.I. LENINA (1990) dar: dieses Frequenzwörterbuch zu den Gesammelten Werken des Verfassers wurde seinerzeit für 1990 angekündigt und befand sich seit längerem in Bearbeitung (vgl. DENISOV 1985), wurde mit dem Ende der Sowjetunion aber offensichtlich als nicht mehr so dringlich empfunden. Die Arbeiten an diesem Wörterbuch wurden zwar nicht eingestellt, haben aber bisher auch nicht zu einer Publikation geführt.

Sonstige Weitere Häufigkeitswörterbücher einzelner literarischer Werke sind MINC et al. (1967; zu A. Blok), GENKEL' (1974; zu Mamin-Sibirjak), LJATINA (1968; zu Šolochov) und das "Častotnyj slovar' romana L.N. Tolstogo «Vojna i mir»" (1978).

Weitere Arbeiten zu diesem Komplex: SAJACHOVA (1964), NELJUBIN (1977), ALEKSEEV/TURYGINA (1974), KLIMOVIČ (1986), ALEKSEEV (1971), ROTAR' / ČIŽAKOV-SKIJ (1976), SADČIKOVA (1974), BIGAEV et al. (1967).

3.4.6 Zusammenfassung

Tab. 19 zeigt die 25 häufigsten Wörter des Russischen in einigen der oben vorgestellten Häufigkeitswörterbücher nur mit ihrem Rang, also ohne die Frequenz. Mit diesen 25 Wörtern werden durchschnittlich bereits 26-28% eines beliebigen Textes abgedeckt.

Diese Tabelle zeigt mehrere Dinge ganz anschaulich:

- In den Häufigkeitswörterbüchern, die für die russische Sprache insgesamt repräsentativ sind, befinden sich unter den 25 meistgebrauchten Wörtern praktisch nur Präpositionen, Konjunktionen, Pronomina (darunter alle Personalpronomina) sowie die Kopula. Substantive, Adjektive und Vollverben folgen erst auf den nächsten Rängen.

- Im Häufigkeitswörterbuch der Umgangssprache treten die typischen Dialogpartikel des Russischen in den Vordergrund (*ну, так, да, нет, вот*), die in schriftlichem Gebrauch unter den häufigsten Wörtern nicht zu finden sind.

- Die Zusammensetzung der häufigsten Wörter hängt ganz offensichtlich mit Textsorte und Stil zusammen, und es kann folgende Korrelation in Form einer Hypothese behauptet werden: Je enger gefaßt die Textsorte, desto mehr Autosemantika sind unter den häufigsten Wörtern zu finden. In der allgemeinwissenschaftlichen Fachsprache wie bei DENISOV treten bereits unter den 25 häufigsten Wörtern Autosemantika in Form von Substantiven und Verben auf, die Personalpronomen fehlen hier ganz, Präpositionen und Konjunktionen sind weniger umfangreich vertreten. In der gegenüber der allgemeinwissenschaftlichen Fachsprache thematisch noch engeren Fachsprache der Elektronik (KALININA) sind unter den 25 häufigsten Wörtern bereits etliche typische Lexeme dieser Fachsprache zu finden (*рисунок, ток, напряжение, электроны, поле*). In der Kindersprache (CHARAKOZ 1971) stehen naturgemäß die Personalpronomina, *ja/nein* und typische Handlungen (*sein, geben, gehen, spielen*) obenan.

Die Rang-Frequenzverteilung von Wörtern in ihrer theoretischen Form, die in solchen Häufigkeitslisten ihren Niederschlag findet, ist der Gegenstand des sog. Zipf-Mandelbrotschen Gesetzes, mit dem sich auch die russische QL ausführlich auseinandergesetzt hat (vgl. u.a. FRUMKINA 1961b, KALININ 1964a, 1964b).

Rang	Засорина Allgemein	Штейнф. Allg./Kind.	Лённгрен Allgemein	Турко Umggsspr.	Пол./Солг. Zeitung	Денисов Wissensch.	Калинина Elektronik	Харакоз Kinderspr.	Пушкин Belletr.
1	в	и	и ¹	не	в	в	в	я	и
2	и	в	в ¹	я	и	и	и	ты	в
3	не	на	на ¹	а ¹	на	на	при	а	я
4	на	не	не	в	он	с	на	не	он
5	я	он	с	у ²	с	при	для	он	не
6	быть	я	этот	и ¹	не	этот	с	в	быть
7	что	что	быть	ну	что	для	от	и	на
8	он	с	он	так	этот	по	по	что	с
9	с	этот	а ¹	нет	быть	тот	что	у	что
10	а	быть	весь ¹	он	по	из	рис.	быть	ты
11	как	а	что ²	что ²	весь	который	из	мы	мой
12	это	весь	тот	да ¹	к	что	как	вот	она
13	вы	они	как	на ¹	о	к	не	на	свой
14	ты	она	к	ты	год	быть	к	пойти	но
15	к	как	по	вы	за	не	а	ну	к
16	мы	мы	но ¹	как	который	или	тока	нет	весь
17	этот	к	они	это	тот	от	можно	идти	они
18	она	у	и ²	она	а	как	ток	дать	как
19	они	ты	что ¹	с	свой	а	напряжения	с	тот
20	но	за	она	что и	из	весь	так	как	мы
21	по	тот	из	вот	но	иметь	электронов	этот	о
22	весь	но	который	мне д	для	мочь	случае	сейчас	вы
23	за	вы	от	там	как	так	до	да	а
24	то	по	за	меня р	это	такой	или	весь	за
25	все	из	я	очень	мы	число	поля	играть	это

Tab. 19: Die 25 häufigsten Wörter des Russischen

3.4.7. Weitere Aspekte

Fremdwortgebrauch Durch Erhebung empirischer Daten aus Tageszeitungen zeigt KRYSIN (1965), daß der Gebrauch von Fremdwörtern in den 20er Jahren gegenüber dem ersten Jahrzehnt des Jahrhunderts statistisch signifikant zugenommen hat. Da es aber in diesem Zeitraum kaum neue Fremdwörter gegeben hat, interpretiert er dies als eine verstärkte Aneignung älterer Fremdwörter (aus dem 19.Jh.) durch breitere Sprecherschichten.

Kulturvergleich Die 1.000 häufigsten Wörter des Russischen vergleicht YOKOYAMA (1987) mit denen des Amerikanischen Englisch, und zwar im Hinblick auf ausgewählte semantische Gruppen von Adjektiven und Substantiven, nämlich Ideologie/Geschichte/Politik, Ökonomie, Kultur, Militär, Bildung und Religion. Die Daten zum Russischen sind dabei dem Wörterbuch von ZASORINA (1977) entnommen. Zu jedem Wort wird seine Häufigkeit in beiden Sprachen ermittelt und in Tabellenform dokumentiert. Die Ergebnisse werden im Hinblick auf einen Kulturvergleich interpretiert, wobei die Autorin zu dem Schluß kommt, daß in der UdSSR Themen wie Ideologie, Ökonomie, Kultur und Militär eine größere Rolle spiel(t)en als in den USA, während dort umgekehrt die Themenbereiche Religion und Bildung häufiger angesprochen werden. Abschließend stellt sie richtig fest, daß die Ergebnisse als solche einen Slawisten kaum überraschen, vielmehr die Tatsache, daß bereits eine relativ kleine Stichprobe diese Ergebnisse liefern kann.

Lexemkombinatorik Die Kombinatorik einer Gruppe von sieben Verben (*возбудить, вызвать ...*) und 50 substantivischen Gefühlsbezeichnungen (*любовь, подозрение ...*) untersucht KOPYLENKO (1965) aufgrund von Tests mit 20 Informanten. Ziel der Auswertung ist es, durch Punktevergabe zu einer abgestuften Aussage darüber zu kommen, welche Lexemkombinationen bevorzugt werden und welche den Sprechern unzulässig erscheinen. In die erste Gruppe gehört beispielsweise *возбудить интерес*, in die zweite dagegen *разбудить радость*.

3.5. Syntax

Die russische Syntax gehört zu den Bereichen, die bislang nicht umfassend mit sprachstatistischen Methoden untersucht wurden, was natürlich auch mit der Vielzahl der Aspekte zusammenhängt, die man am Satzbau untersuchen kann, und mit der Tatsache, daß die sprachlichen Einheiten auf dieser Ebene eine offene Menge darstellen, keine abzählbare. Damit entfällt natürlich auch die Möglichkeit von Häufigkeitszählungen wie auf der Ebene von Phonemen, Morphemen und Lexemen; in der Syntax müßte man hier stattdessen etwa auf Satzmuster zurückgreifen. Die klassische Fragestellung der QL in diesem Bereich betrifft die Häufigkeitsverteilung der Satzlänge (mit welcher Frequenz treten Sätze der Länge 1, 2, ..., n Wörter auf?). Generell kann man bei syntaktischen Untersuchungen *absolute* und *relative* Frage-

stellungen in folgendem Sinne unterscheiden: Das Auftreten von Wörtern, Wortarten oder dergl. in der ersten, zweiten, dritten ... bzw. letzten, vorletzten, drittletzten ... Position des Satzes wäre eine absolute Betrachtung in bezug auf die Rahmeneinheit Satz. Das Auftreten eines bestimmten Wortes (oder einer Wortart) vor oder nach einem anderen Wort (oder einer Wortart) ohne Rücksicht auf die Stellung im Satz wäre eine relative Perspektive.

3.5.1. Syntaktische Relationen und Phrasenstruktur

Rektion Ein Versuch, innerhalb der syntaktischen Relationen Kongruenz, Rektion und Adjunktion die oft benutzte (aber problematische) Unterscheidung von starker und schwacher Rektion mit den Mitteln der QL zu präzisieren, stammt von APRESJAN (1964). Sein Vorschlag geht dahin, die bedingte Wahrscheinlichkeit des Erscheinens einer abhängigen Wortform (bzw. eines Komplexes) nach einem bestimmten Lexem als Grundlage der Messung zu verwenden. Beispielsweise heißt es unter insgesamt 343 Vorkommen von *жить* 14 mal *жить с N_i* (= 0,04), 26 mal *жить на N_p* (= 0,08), jedoch 86 mal *жить в N_p* (= 0,25), wobei der Index von N (= Nomen) dessen Kasus angibt. Es ist offensichtlich, daß die Fügungspotenz dieses Verbs, also eine abstrakte Eigenschaft, ungleich in Bezug auf diese möglichen drei Fortsetzungen ausgeprägt ist; der in Klammern hinzugesetzte relative Anteil ist hier die Stärke der betreffenden Rektion. Zusammen haben wir in diesem Falle 14 + 26 + 86 = 126 Vorkommen mit abhängigem Wort, jedoch 343 – 126 = 217 ohne Ergänzung, d.h. einen relativen Anteil von 0,367 bzw. 0,633. APRESJAN entwickelt dieses Maß dann noch in zwei Schritten weiter. Die Abgrenzung zwischen starker und schwacher Rektion baut auf diesem komplexeren Maß auf und liegt nach APRESJANS Untersuchung bei 0,7. Abschließend teilt er eine Auswahl russischer Verben in vier Klassen mit schwacher und zwei Klassen mit starker Rektion ein. Grundlage der Untersuchung sind die Daten des Häufigkeitswörterbuches von ŠTEJNFEL'DT. (Zu APRESJANS Verfahren vgl. präzisierend auch SOLONICYN 1966.)

Der Rektion ist auch eine zweite Arbeit des gleichen Autors gewidmet (APRESJAN 1965). In ihr geht es darum, die Fügungspotenz von Verben zu quantifizieren, auf dieser Grundlage die Ähnlichkeit der Verben zu berechnen und so zu einer Klassenbildung zu kommen. Wiederum sind die Daten von ŠTEJNFEL'DT Grundlage der Untersuchung, denn die genannte Autorin gibt ja in einer eigenen Tabelle die Häufigkeit jedes Rektionsmusters bei ausgewählten Verben an.

Mit der Wahl zwischen dem Akkusativ und dem Genitiv nach verneinten Verben beschäftigt sich GREEN (1979) in einem längeren Artikel, wobei er versucht, die Wahl in Abhängigkeit vom Genus des Substantivs, das als Objekt gebraucht wird, vom Aspekt des Verbs, vom konkreten Lexem usw. genauer zu beschreiben.

Kongruenz Einige quantitative Daten zur Kongruenz (formal oder semantisch?) eines verbalen Prädikates mit Mengenangaben wie *большинство, пять, несколько, двое* etc. finden sich bei CORBETT (1981) im Anschluß an SUPRUN (1969,

185ff.). Er kommt zu dem Schluß, daß der von manchen Autoren behauptete Anstieg der semantischen Kongruenz tatsächlich auf die Kollektivzahlwörter beschränkt sei und nicht allgemein behauptet werden könne.

Syntagmen BACHMUTOVA (1970) untersucht die spezifischen syntaktischen Eigenschaften einiger besonders häufiger Verben in der *razgovornaja reč'* an empirischem Material. Dabei geht es ihr insbesondere darum, die tatsächlichen Verwendungsweisen der Verben gegen die Angaben, wie sie in Wörterbüchern erscheinen, abzugrenzen. Der prozentuelle Anteil der häufigsten Konstruktionen wird genannt und mit Beispielen belegt, allerdings nicht übersichtlich zusammengefaßt.

Einige quantitative Daten zu Umfang und innerer Struktur von Umstandsangaben im Russischen (gewonnen an mathematischen Texten) finden sich bei SANNIKOV (1963). Untersucht wird hier konkret die Stellung des Bezugswortes (Voranstellung vs. Nachstellung, vgl. *произвольно избранный vs. избранный нами*) und die Länge solcher Syntagmen. Die Ergebnisse lassen u.a. erkennen, daß die Stellung direkt vom Umfang der Phrase abhängt: nur Erweiterungen aus zwei Wörtern tendieren zur Voranstellung, alle anderen zur Nachstellung, von denen es in der Stichprobe auch insgesamt sehr viel mehr Fälle gab.

Wortverbindungen aus drei Wörtern untersucht NIKITINA (1971), indem sie zunächst in einer umfangreichen Stichprobe publizistischer Texte die 120 häufigsten Substantive bestimmt und dann alle Verbindungen dieser Substantive mit den beiden im Text vorangehenden Wörtern ermittelt. Tatsächlich erweisen sich diese rein mechanisch bestimmten Gruppen zu einem hohen Prozentsatz tatsächlich als Syntagmen. Abgedruckt wird eine Rang-Frequenzliste, die beispielsweise die Wendungen *в ц году [ц = цифра], Верховного Совета СССР, Совета Министров СССР, во всем мире* anführen. – Die 1.000 häufigsten Syntagmen aus Militärtexten lassen sich KOLGUŠKIN (1970, 161ff.) gleich zweifach entnehmen. Unangefochten nimmt den ersten Rang die Wendung *наступать в направлении* ein, der auf Rang drei *уничтожать противника* und auf Rang acht *ядерное нападение* folgen.

Eine Stichprobe von 50.000 Dreiwortverbindungen aus Texten zur Radioelektronik dient drei Aufsätzen als Grundlage für zwei ganz parallele Auswertungen, in denen freilich die linguistischen Grundlagen des Tuns, der Sinn der Fragestellung und eine Auswertung der Ergebnisse zu kurz kommen bzw. einfach fehlen, so daß die vorgelegten Daten praktisch wertlos sind. BULAŠEVA (1969) untersucht Dreiersyntagmen, bei denen das "Stützwort" das mittlere ist (*как видно из*), ISABEKOVA (1969) solche mit rechtem "Stützwort" (*с течением времени*), BELOCERKOVSKAJA (1969) solche mit verbalem Stützwort (*приведены в таблице*). Was jeweils als "Stützwort" betrachtet wird, wird in den ersten beiden Arbeiten aufgrund einer Liste der häufigsten Wörter zur Radioelektronik entschieden.

Die Unterschiedlichkeit von Hochsprache und *razgovornaja reč'* in bezug auf syntaktische Strukturen weisen VANNIKOV/ICHMAL'JAN (1969, 1970) mit statistischen Tests nach. Unterschiede in der Komplexität verbaler Wortfügungen in Um-

gangssprache und Hochsprache betont auch KORMILICYNA (1974) als Ergebnis ihrer Untersuchungen.

3.5.2. Eigenschaften ganzer Sätze

Wortfolge SCHALLER (1970) prüft die Anwendbarkeit informationstheoretischer Untersuchungen, die ja in erster Linie der Abfolge von Buchstaben gegolten haben, auf syntaktischer Ebene, d.h. im Bereich der Wortstellung, und kommt zu dem allgemeinen Ergebnis, die Wortstellung sei doch zu stark einzelsprachlichen Regeln unterworfen, als daß sie informationstheoretisch erfaßt werden könnte. Diesem Schluß liegt allerdings ein Mißverständnis zugrunde: Informationstheoretische Untersuchungen setzen mitnichten voraus, daß alle jeweils denkbaren Möglichkeiten auch gleich wahrscheinlich sind.

Den Einfluß der Wortstellung und eines Faktors wie Bejahung/Verneinung auf die Wahl des Kasus beim Prädikatsnomen (Nom. vs. Instr., Akk. vs. Gen.) untersucht HENTSCHEL (1992b). Er registriert ferner ganz allgemein eine klare Tendenz zur Initialstellung des Subjektes – durchschnittlich 85% aller Sätze seines Korpus weisen diese Struktur auf. Außerdem werden Subjekt und Prädikatsnomen in ca. 75% der Fälle durch das Prädikat (Verb) getrennt, d.h. die häufigste Struktur ist demnach S – V – O (wobei O hier sowohl für das direkte Objekt wie das Prädikatsnomen steht).

Die Vor- und Nachstellung bei Wörtern wie *рад, уверен, известно, жаль, должен* etc. untersuchen SIROTININA et al. (1968) und geben für beide Möglichkeiten jeweils Prozentzahlen an, die einer empirischen Stichprobe belletristischen Materials entstammen.

BUTTKE (1965) nimmt ausführlich zu den Grundlagen, Anwendungsmöglichkeiten und Vorzügen sprachstatistischer Untersuchungen Stellung, um dann konkret die Stellung der Personalpronomina in bezug auf das Prädikat (davor – danach) an belletristischen Texten zu untersuchen. Seine Daten zeigen, daß nur Puškin und Gogol' einen höheren Anteil der Nachstellung des Pronomens (mit ca. 13%) aufweisen, während er sonst zwischen 2% und 8% liegt. Die Stellung von Nomen–trans. Verb (79%) gegenüber trans. Verb–Nomen (21%), von Nomen–intrans. Verb (51%) zu intrans. Verb–Nomen (49%), von Nomen–(Kopula)–Adjektiv (16%) gegenüber Adjektiv–(Kopula)–Nomen (84%) und von Partzip–(Kopula)–Nomen (62%) gegenüber Nomen–(Kopula)–Partzip (38%) untersucht der gleiche Autor (BUTTKE 1968) aufgrund einer eigenen (aber nicht genauer spezifizierten) Stichprobe. – Zur Wortfolge vgl. auch MENŠIKOV (1974, 1978).

Wortfolge und Rhythmus In diesen Bereich gehört die Untersuchung von RESTAN (1981), der das Auftreten von Substantiven, Verben und sonstigen Wortarten am Satzende auf der Grundlage selbst erhobener Stichproben analysiert. Die-

se Verteilung unterscheidet die Funktionalstile *knižnaja reč'* und *razgovornaja reč'* eindeutig voneinander: in der ersteren dominieren am Wortende mit durchschnittlich 86% Substantive, während Verben mit nur 4% praktisch keine Rolle spielen; in der letzteren dagegen ist das Verhältnis der drei Klassen zueinander fast ausgeglichen (39 : 33 : 28), also offensichtlich signifikant anders. Auch der Rhythmus unterscheidet beide Stile: in der Umgangssprache ist ein "maskulines" Satzende (letzte Silbe betont) doppelt so häufig wie in der Schriftsprache (40 : 20), während umgekehrt dort ein daktylischer Satzauslaut (drittletzte Silbe betont) doppelt so häufig ist (35 : 17).

Satztypen Das Verhältnis der grundlegenden Satztypen Deklarativsatz, Interrogativsatz und Ausrufesatz zueinander untersuchte bereits ERMOLAEV (1915) an Texten (vgl. PAPP 1966, 18f.). Er stellte einen besonders hohen Anteil von Deklarativsätzen in der Prosa fest (84,2%), während er in der Lyrik bedeutend niedriger liegt (59,8%). Dramen zeichnen sich demgegenüber durch einen relativ hohen Anteil von Fragesätzen (14,9%) und einen relativ niedrigen Anteil von Aussagesätzen (58,2%) aus. Die Gültigkeit der Ergebnisse ist allerdings insofern einzuschränken, als hier Texte aus verschiedenen Epochen miteinander vermischt werden. – Die Häufigkeit der Satztypen des einfachen Satzes sowie seiner Bestandteile (Subjekt, Prädikat, Ergänzung etc.) in verschiedenen Funktionalstilen stellt RUSOVA (1974) dar, ähnlich MENŠIKOV (1974); GRECHNĚVA (1974) untersucht Nebensatztypen.

Satzstrukturen in Zeitungstexten sind der Gegenstand der Untersuchung von ŠVEC (1979). Die Autorin ermittelt das Auftreten einfacher und komplexer Sätze, jeweils nochmals unterteilt in mehrere Untertypen, in den verschiedenen Textsorten der Zeitung (Überschrift, Leitartikel, Kommentar...) für die Jahre 1924 und 1974 mit dem Ziel, bestimmte Veränderungstendenzen aufzuzeigen. Ihr Resümee lautet, die Entwicklung der Zeitungssprache lasse sich den allgemeinen Entwicklungstendenzen des Russischen in Richtung Demokratisierung, Vereinfachung und sprachlicher Ökonomie unterordnen und zeige sich im untersuchten Bereich in einer Zunahme einfacherer Satzstrukturen, sowohl innerhalb der komplexen Sätze wie auch in einer Gewichtsverschiebung zugunsten einfacher Sätze insgesamt, woran auch die Zunahme umgangssprachlicher Strukturen ihren Anteil habe. Die statistische Bearbeitung des Materials zeichnet sich durch Explizitheit aus (die verwendeten Formeln, Schwellenwerte etc. werden im Anhang tabellarisch aufgeführt). Allerdings ist auch die ideologisch geprägte Einleitung selbst für sowjetische Verhältnisse ausführlich zu nennen; in ihr geht es u.a. zum Beispiel um Lenins Verhältnis zu Überschriften.

Satzgefüge ŽURAVLEV (1967) zeigt aufgrund einer Untersuchung belletristischer, publizistischer und umgangssprachlicher Texte, daß der Anteil der Objektsätze (*слышать что...*; *желать, чтобы...*) an den komplexen Sätzen insgesamt einer Normalverteilung folgt. Das heißt: ein sehr geringer oder sehr hoher Anteil solcher

Sätze kann als ein Stilmerkmal der betreffenden Autoren verstanden werden. In den publizistischen Arbeiten der untersuchten Autoren (Gor'kij, A. Tolstoj, Šolochov, Ėrenburg, Simonov, Leonov) liegt der untersuchte Satztyp mit einem Anteil von 33% deutlich unter dem in belletristischen (47%) und umgangssprachlichen Texten (49%), in denen er das charakteristischste Satzgefüge überhaupt darstellt. In Bezug auf die Autoren gesehen, stellen Gor'kij mit einem Anteil von 65% solcher Sätze in der Belletristik und Leonov mit nur 30% die Extremfälle dar. Eine graphische Veranschaulichung ergibt folgendes Bild, wobei der Umfang der Kreise der Variationsbreite in der Anwendung des untersuchten Satztyps entspricht (nach ŽURAVLEV 1967, 107):

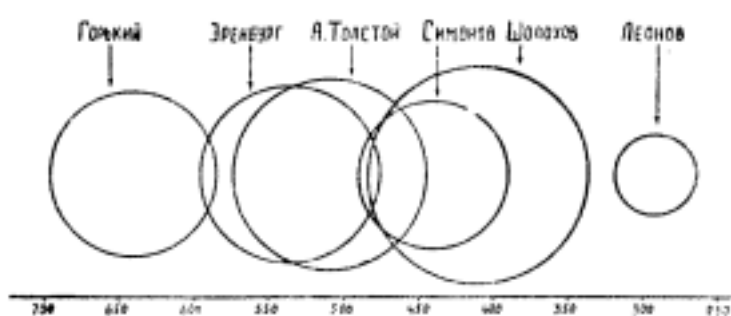


Fig. 1: Verteilung einiger Satztypen bei russischen Autoren

Die komplexen Sätze in der *razgovornaja reč'* sind der Untersuchungsgegenstand von ZIL'BERT (1970). Ihm geht es u.a. um den Anteil der subordinativen Sätze, den er aufgrund empirischer Stichproben bei etwa 10% ansiedelt, während er in verschiedenen Funktionalstilen der Literatursprache bei 30% liegt (Wissenschaft, Publizistik, Belletristik). Untersucht werden auch die konkreten Formen des Nebensatzanschlusses, mit einem ähnlich klaren Unterschied zur Hochsprache. Der Autor zeigt aber zugleich, daß die Personenrede in Prosa und Dramen die typischen Verhältnisse der RR aufweist, also durchaus als Teil der "echten" gesprochenen Sprache aufgefaßt und untersucht werden kann.

Satzlänge Die Länge von Sätzen ist eine ebenso offensichtliche wie charakteristische Eigenschaft und kann vornehmlich in Wörtern bzw. Teilsätzen gemessen werden. Untersuchungen, die einer Überprüfung der Hypothesen gewidmet wären, die sich aus dem bekannten "Menzerathschen Gesetz" für die syntaktische Ebene ableiten lassen (vgl. ALTMANN/SCHWIBBE 1989, 8–14), fehlen bisher für das Russische weitgehend. Das genannte Gesetz besagt allgemein: "Je größer (komplexer) ein sprachliches Konstrukt, desto kleiner (einfacher) seine Konstituenten". Die genannten Autoren leiten daraus für die syntaktische Ebene vor allem die folgenden speziellen Hypothesen ab: "Je länger eine Phrase (Wortverbindung, Syntagma), desto kürzer die Wörter" (Hypothese 11), "Je länger der Satz, desto kürzer die Teil-

sätze (Clauses)“ (Hypothese 12). Beide Zusammenhänge ließen sich relativ einfach durch Auswertung von Textkorpora überprüfen.

Ohne von den genannten Sprachgesetzen zu wissen, bestätigen allerdings KALAŠNIKOVA/AL'NIKOVA mit ihren Befunden indirekt Hypothese 12. Diese Autoren untersuchen mehrprädikative Sätze vom Typ *Эти времена в армии прошли, и нет признаков, что они могут вернуться* (K. Simonov) und stellen durch Betrachtung der graphischen Umsetzung ihrer Daten folgenden Zusammenhang fest: “Je länger ein komplexer Satz, desto mehr Gliederungsebenen weist er auf“ (1991, 85).

Die Satzlänge als solche ist eine Eigenschaft des Russischen, die bereits ausführlich untersucht wurde. Dies ist vor allem den Arbeiten von LESSKIS (1962, 1963a, 1963b, 1964) zu verdanken. Dieser Autor hat ein außerordentliches großes Textkorpus aus verschiedenen Funktionalstilen des Russischen untersucht, wobei die gewonnenen Daten primär graphisch präsentiert werden (m.a.W.: die eigentlichen Zahlen zur Verteilung der Satzlänge im Russischen werden *nicht* aufgeführt, was einigermaßen unverständlich erscheint). Deutlich ist dabei, daß die Aufsätze mit der Zeit immer “mathematischer” werden; offensichtlich hat der Autor als Reaktion auf den ersten Aufsatz, der noch keinerlei statistische Tests oder Kennzahlen enthält, entsprechende Hinweise erhalten, auch auf einschlägige westliche Arbeiten zur Satzlänge (G.U. Yule).

Die erste Arbeit untersucht die Satzlänge in Texten der zweiten Hälfte des 19.Jhs. an Romanen von Gončarov, Dostoevskij, Leskov, Saltykov-Ščedrin, Tolstoj, Turgenjev und Černyševskij (insges. 70.643 Sätze), an wissenschaftlicher Prosa der gleichen Zeit (insges. 17.302 Sätze) sowie an Briefen Turgenjovs, Dostoevskijs und Tolstojs (insges. 8186 Sätze). Die Daten zu Satz- und Wortzahl werden in absoluten und relativen Zahlen für jede Quelle getrennt angegeben, in der Belletristik außerdem nach Autorenrede, Personenrede und Einschüben getrennt. Diese Daten, die eine Berechnung der durchschnittlichen Satzlänge erlauben, können deshalb genauso gut auch als Untersuchung des Stils der betreffenden Autoren wie der betreffenden Zeit wie des zugrundegelegten Genres (Psychologische Romane) gelten.

Als Beispiel sei eine der Graphiken (**Fig. 2**) wiedergegeben. Sie zeigt die Verteilung der Satzlänge getrennt nach Belletristik (nur Autorenrede, Linie 1) und wissenschaftliche Prosa (Linie 2), die eine deutliche Tendenz zu längeren Sätzen erkennen läßt.

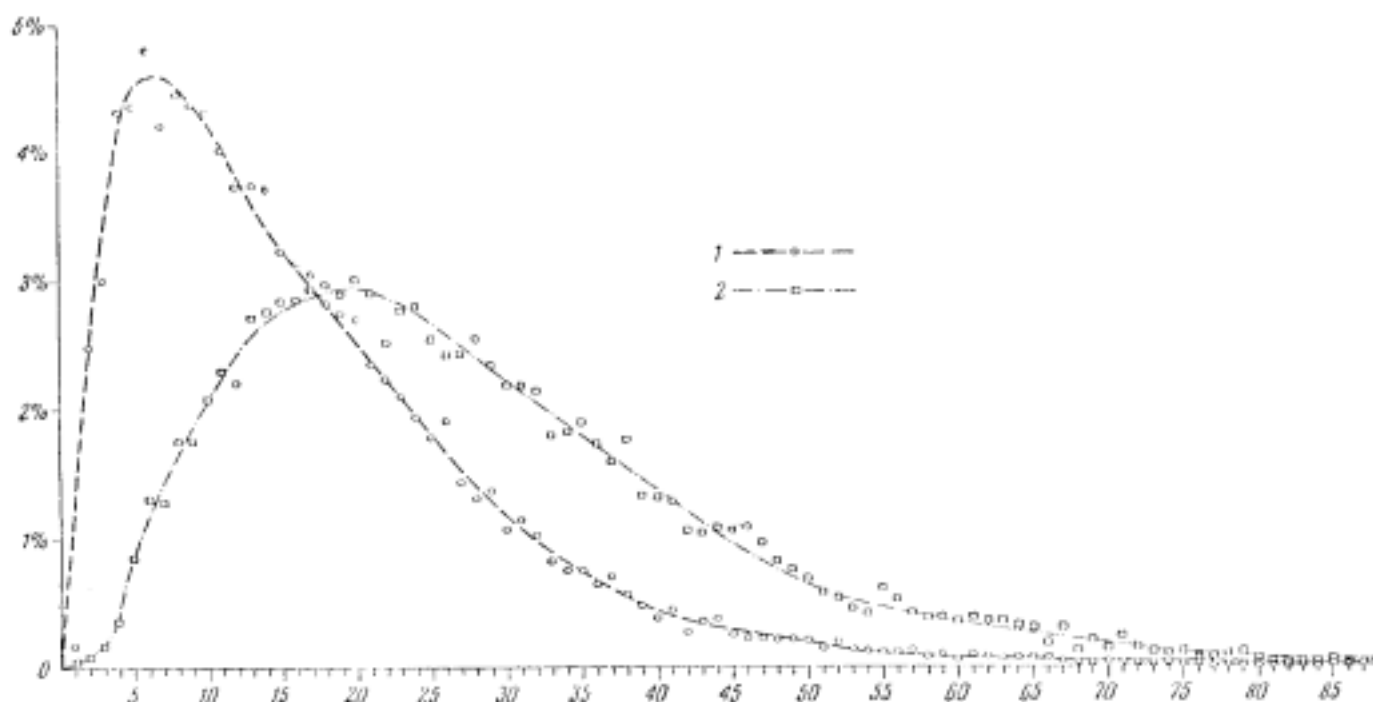


Fig. 2: Verteilung der Satzlänge im Russischen des 19. Jh.

In ähnlicher Weise untersucht die zweite Arbeit LESSKIS' (vgl. 1963a) die Satzlänge in altrussischen Texten, in Texten des 18. Jhs. sowie im 19.-20.Jh. Diese Angaben ließen sich folglich auch für diachrone Fragestellungen auswerten. Zusätzlich finden sich hier Angaben zu den Anteilen einfacher vs. komplexer Sätze, die in der Belletristik im Verhältnis 60 : 40 oder ausgeglichener zueinander stehen – bei manchen Autoren überwiegen die einfachen Sätze, bei anderen die komplexen (Tolstoj findet sich mit verschiedenen Werken an beiden Enden des Spektrums). Die wissenschaftliche Literatur zeigt ein konstantes Überwiegen komplexer Sätze (70 : 30 bis 80 : 20). Die übrigen Aufsätze des Autors wiederholen den gleichen Ansatz oder vertiefen die Analyse in Bezug auf einzelne Punkte. – Weitere Arbeiten: ARAPOV / ŠREJDER (1970), AKIMOVA (1973, 1977), MOSSNER (1969).

Entropie Den Informationsverlauf in ganzen Sätzen demonstrieren ŽURAVLEV et al. (1968) sowie PIOTROVSKIJ (1968) an einigen einfachen und komplexen russischen Sätzen. Die Resultate basieren auf einem Experiment, bei dem den Versuchspersonen das erste Wort des Satzes vorgegeben wurde, worauf dann das zweite und alle weiteren zu erraten waren. Hierzu siehe auch den Abschnitt 'Graphematik'.

3.6. Textlinguistik, Type-Token-Ratio

Sprachstatistische Untersuchungen basieren natürlich oft auf Zählungen an Texten; im eigentlichen Sinne textlinguistische Untersuchungen gibt es jedoch nur wenige. Auch MISTRÍK (1973) hat bisher kein Gegenstück im Russischen. Soweit Untersuchungen die Charakteristika eines individuellen Texten betreffen, gehen wir auf sie im Abschnitt "Statistik und Stil" (s.u.) ein.

Type-Token-Ratio Als Textcharakteristikum im engeren Sinne kann die sogenannte Type-Token-Ratio betrachtet werden, das heißt das Verhältnis von verschiedenen Lexemen (*types, different words*) zur Gesamtzahl der Wörter eines Textes (*tokens, running words*). Da sich Wörter in Texten zunehmend wiederholen, je länger ein Text ist, hängt diese Variable u.a. direkt von der Textlänge ab. Berechnet und vergleicht man die Type-Token-Ratio also für verschiedene Texte, so muß man zuallererst darauf achten, gleichlange Texte bzw. Stichproben zu untersuchen. Eine etwas andere Fragestellung zielt darauf ab zu untersuchen, wie sich das genannte Verhältnis mit fortschreitender Textlänge verändert bzw. wie groß die Zahl der neuen Lexeme pro Textstück ist. Durch einfaches Nachdenken kann man hier schon plausible Hypothesen bilden: Am Anfang eines Textes ist praktisch jedes Wort "neu", die Type-Token-Ratio also ausgeglichen. Je länger der Text ist, desto weniger neue Wörter kommen vor, d.h. desto öfter werden die bisher schon gebrauchten Wörter wiederholt (man spricht deshalb auch von der sog. "Wiederholungsrate"). Diese Abhängigkeiten numerisch genau zu ermitteln, graphisch darzustellen, allgemeine Funktionen für den genannten Zusammenhang aufzustellen oder abzuleiten, Grenzwerte für die Zahl der neuen Wörter zu finden usw., dies alles sind Fragen, mit denen sich die QL in diesem Bereich beschäftigt (vgl etwa ALTMANN/BURDINSKI 1982, die auch Daten zum Russischen berücksichtigen).

Eine Untersuchung dieses Bereiches hat etwa SMITH (1968) vorgelegt. Neben der Type-Token-Ratio untersucht er auch Phänomene wie die sog. *Hapax Legomena*, also die genau einmal gebrauchten Wörter, die nichts weiter als ein Spezialfall des allgemeineren Zusammenhanges sind (nämlich Wörter mit einer Wiederholungsrate gleich Null). Seine Auswertungen beruhen auf einem Textkorpus von mehr als 200.000 Wörtern, das mit Computerprogrammen, die vorher schon für ähnliche Untersuchungen zum Amerikanischen Englisch entwickelt worden waren, ausgewertet wurde. Das Korpus setzte sich aus 100 gedruckten Prosa-Texten zu je 2000 Wörtern aus den Jahren 1960 bis 1967 zusammen. Auf die Resultate im einzelnen einzugehen, ist im Rahmen der vorliegenden Übersicht nicht möglich; sie können der sehr übersichtlich gehaltenen Arbeit entnommen werden. – Russische Daten berücksichtigt auch MAAS (1972).

PAPP (1961) geht der Frage nach, inwieweit die Zahl der neu (= erstmals) auftauchenden Wörter von der Textprogression und damit von einzelnen Textfragmenten abhängt. Untersucht wird diese Frage, eine spezielle Variante der Type-Token-Ratio, an Texten von Puškin, Tolstoj, Šolochov und der Akademiegram-

matik. Jeder Text wird in Abschnitte zu 100 Wörtern unterteilt und in jedem Abschnitt der Anteil der neuen Lexeme (und der verschiedenen Lexeme insgesamt) am Text ermittelt. Die Ergebnisse zeigen, daß der Anteil der neuen Wörter von 67–77% unter den ersten hundert Wörtern bis auf 30% sinkt, wenn der Text auf eine Länge von 2.000 Wörtern wächst. Der Anteil an neuen Wörtern sinkt dabei am Anfang (bis zur Textlänge 400) besonders stark, danach flacht die Kurve deutlich ab. Die untersuchten Texte unterscheiden sich zwar im Einzelfall, die Grundtendenz ist jedoch einheitlich und damit überindividuell. – In einen größeren Zusammenhang eingebettet, findet man entsprechende Daten zum Russischen (Puškin, Evgenij Onegin, Kapitanskaja dočka) auch bei HERDAN (1965, 102–103).

Mit Wortwiederholungen beschäftigt sich auch – eher theoretisch – CHOVA NOV (1978) auf der Grundlage belletristischer Texte.

Textdeckung Hierunter versteht man die Frage, wieviel Prozent eines gegebenen bzw. eines beliebigen Textes durch die x häufigsten Wörter einer Häufigkeitsliste abgedeckt werden. Auch diese Abhängigkeit läßt sich in Form einer ganz typischen Kurve darstellen. Da diese Funktion eng mit Frequenzwörterbüchern zusammenhängt und gewöhnlich als eine der Standardauswertungen in ihnen wiedergegeben wird, sind wir hierauf schon im Abschnitt zur lexikalischen Statistik eingegangen (vgl. oben).

3.7. Lexikalische Semantik, Onomasiologie

Seme Ein Häufigkeitswörterbuch “semantischer Quantoren” oder Deskriptoren stellt KARAULOV (1980) dar, wobei damit die zur Erklärung eines Eintrages benutzten Wörter in einem einsprachigen Wörterbuch gemeint sind, die ja die Bedeutung des Lexems beschreiben sollen. Grundlage der Auswertung sind die Wörterbücher von OŽEGOV und UŠAKOV. Einträge des Wörterbuchs sind die um redundante Elemente verkürzten Wortstämme, z.B. СУМАЩ, ДЕНЬГ, СПУТН, insgesamt 5562 Elemente mit Häufigkeiten von 1 bis 69. Sie werden in einer Rang-Frequenz-Liste und (in einer etwas umfangreicheren Liste) alphabetisch aufgeführt.

Lexikalische Bedeutungen als “unscharfe Mengen” behandelt PIOTROVSKIJ (1979, vgl. 42-43) am Beispiel von *ночь – утро – день – вечер – ночь*.

Polysemie und Synonymie Es ist ein gut bekanntes Phänomen, daß die Häufigkeit eines Wortes mit der Zahl seiner Bedeutungen positiv korreliert ist: Je allgemeiner die Bedeutung eines Wortes ist, d.h. je mehr Bedeutungen es hat, desto häufiger wird das Wort verwendet. Gleichzeitig gilt aber auch folgende negative Korrelation: Es gibt im Lexikon einer Sprache viele eindeutige Wörter und nur wenige extrem vieldeutige. Eine Auswertung des Wörterbuches von OŽEGOV auf die Verteilung der Polysemie hin hat KRYLOV (1982, vgl. 100) unternommen; er schlägt auch ein gut angepaßtes mathematisches Modell für den Zusammenhang zwischen

den Variablen 'Zahl der Bedeutungen eines Wortes' und 'Zahl der Lexeme mit x Bedeutungen' vor ("Wieviele Wörter gibt es, die ein, zwei usw. Bedeutungen haben?"). Eine nach Substantiven und Verben getrennte Auswertung zeigt, daß Verben im Durchschnitt mehr Bedeutungen aufweisen als Substantive. Die grundlegenden Zahlen seien hier genannt:

Bed.zahl	1	2	3	4	5	6	7	8
Lexeme	23456	6288	1682	649	232	110	55	34
Bed.zahl	9	10	11	12	13	14	15	16
Lexeme	23	16	7	2	2	0	2	1

Tab. 20: Verteilung der Zahl der Bedeutungen nach Lexemen

PLOTNIKOV (1979) geht von der distributionalistischen Maxime aus, die Bedeutung eines Wortes sei sein Gebrauch, und untersucht dementsprechend die Zahl der Kontexte kleinerer Gruppen von Wörtern (beispielsweise *звезда*, *луна*, *небо*, *земля* und einigen anderen) in Abhängigkeit von der Textsorte. Unter den Kontexten wird hierbei die Zahl der verschiedenen Lexeme verstanden, mit denen sich die untersuchten Wörter verbinden, wobei die Annahme lautet, daß die Bedeutung eines Wortes umso breiter ist, je höher die Zahl dieser Kombinationspartner ist. Um dies nachzuweisen, wird die Lexemkombinatorik etwa in wissenschaftlichen Texten mit derjenigen in belletristischen Texten verglichen. Es zeigt sich tatsächlich, daß in fachsprachlichen Texten auf tausend Kombinationen nur ca. 200–250 verschiedene Lexeme kommen, während in der Belletristik diese Zahl bei ca. 500–600 liegt. Dieser Unterschied ist ohne Zweifel signifikant.

In ähnlicher Weise ermittelt PLOTNIKOV den Grad der Synonymie von Wörtern, indem er den Korrelationskoeffizienten für die Ähnlichkeit der Distribution in bestimmten syntaktisch-semantischen Kontexten berechnet – seit langem ist ja bekannt, daß auch traditionell als Synonyme bezeichnete Wörter eine durchaus verschiedene Distribution aufweisen, und daß es "echte Synonymie" danach gar nicht gibt. Man kann auch sagen: dies beweist nur, daß "Synonymie" kein kategorisches Merkmal, sondern eine graduelle Eigenschaft ist. Nach PLOTNIKOV (1986) weisen beispielsweise *мысль* und *идея* einen Synonymiegrad von 0,93 auf, *мнение* und *мышление* hingegen nur von 0,36, so daß auf dieser Grundlage zwischen "starker" und "schwacher" Synonymie unterschieden werden kann, deren Grenze der Autor bei 0,7 ansetzt. Dem gleichen Thema sind weitere Arbeiten des Autors gewidmet (vgl. PLOTNIKOV 1980, 1984, 1985). – Die Synonymieverhältnisse unter Farbwörtern untersuchen VASIL'EVICĀ/SKOKAN (1986) mit quantitativen Methoden am Beispiel von "gelb".

An englischen Lehnwörtern untersucht REINTON (1978) ein spezielles Synonymieverhältnis, nämlich dasjenige zwischen dem Lehnwort und einem einheimischen Äquivalent bzw. Äquivalenten. Seinen Angaben, dargestellt in einer Tabelle, läßt sich etwa entnehmen, daß *матч* – *встреча* – *игра* im Verhältnis 66 : 21 : 13 ge-

braucht werden, das Verhältnis weitgehend ausgeglichen ist zwischen *тайм-аут* und *минутный перерыв* (55 : 44), aber der russische *нападающий* den englischen *forward* mit 63 : 37 überflügelt. Die Frequenzangaben wurden vom Autor aus russischen Sportzeitungen der Jahre 1976–78 ermittelt.

Semantische Häufigkeitswörterbücher systematisiert ALEKSEEV (1973).

Häufigkeit und Bedeutungszahl Die Abhängigkeit der Bedeutungszahl von der Häufigkeit beschreibt das sog. "Zipf-Guitersche Gesetz": ZIPF hatte angenommen, die Zahl der Bedeutungen eines Wortes sei gleich der Wurzel aus seiner (relativen) Häufigkeit (was so nicht gilt, wie man inzwischen weiß). Im Russischen hat eine erste Näherung die folgende Formel ergeben, wobei x für eine Rang-Gruppe steht, \bar{y} für die durchschnittliche Bedeutungszahl der Wörter in dieser Gruppe, und \bar{y}^* für den theoretisch berechneten Wert (vgl. ALTMANN/SCHWIBBE 1989, 78–80):

$$\bar{y}^* = 6,0364 x^{-0,3022}$$

Die nach dieser Formel berechneten Werte stehen in der letzten Spalte der folgenden Tabelle:

Ränge	x	\bar{y}	\bar{y}^*
1-10	1	7,3	6,04
51-60	2	3,1	4,90
101-110	3	4,4	4,33
151-160	4	4,6	3,97
201-210	5	5,5	3,71
251-260	6	3,1	3,51
301-310	7	3,7	3,35
351-360	8	2,2	3,22
401-410	9	2,7	3,11
451-460	10	3,9	3,01

Tab. 21: Verteilung der Bedeutungszahl nach Rängen

Der Vergleich der theoretischen Werte mit den tatsächlich beobachteten steht nach den Autoren "an der Grenze der angenommenen Signifikanz und ist nicht ganz zufriedenstellend", weshalb weitere Wörterbücher zur Überprüfung untersucht werden sollten. Insbesondere fällt an den empirischen Daten des Russischen auf, daß die Bedeutungszahl nicht gleichmäßig abnimmt, sondern bei Rang 200 noch einmal ansteigt.

Farbwörter Ein Assoziationsexperiment zu den Farbwörtern des Russischen haben MORGAN/CORBETT (1989) nach bekannten Vorbildern unternommen. Sie ließen eine Gruppe von 31 Moskauer Probanden alle ihnen innerhalb von fünf Minuten einfallenden Farbwörter notieren, wobei nach jeder Minute jeweils eine Markie-

nung erfolgte. Die Ergebnisse, wie oft jedes Farbwort in jedem Zeitabschnitt genannt wurde, publizieren die Autoren in Form einer (unnötig umfangreichen) Liste. Die "Grundfarben" stehen erwartungsgemäß an der Spitze der Nennungen innerhalb der ersten Minute.

Die Struktur des semantischen Feldes der Farbwörter untersucht auch MOSKOVICĀ (1965, 1969) vergleichend in mehreren Sprachen (v.a. Russisch, Englisch, Französisch). Er geht u.a. davon aus, daß die unterschiedlichen Verwendungshäufigkeiten der Wörter eines Feldes die Struktur dieses Feldes erkennen lassen, indem sie eine – statistisch abgesicherte – Einteilung des Feldes in mehrere Gruppen erlauben. Mit der Häufigkeit sieht dieser Autor die sog. "Aktivität" des Wortes positiv korreliert; hierunter versteht er den Grad der Polysemie sowie die Fähigkeit, Derivate zu bilden und an Phraseologismen teilzunehmen. Semantische Ähnlichkeit (von Adjektiven) wird definiert als Ähnlichkeit der Kombinatorik mit gleichen Substantiven; untersucht wird in diesem Zusammenhang auch die Übersetzungsäquivalenz, z.B. russ. *синий* zu engl. *blue* (0,45) gegenüber *голубой* zu *blue* (0,41). Konkret wird auf der Grundlage einer empirischen Stichprobe ermittelt, mit welchen Substantiven sich die Farbadjektive (mindestens fünf mal) verbinden (vgl. 1969, 96ff.).

Auf der Grundlage der distributionellen Ähnlichkeit werden die Beziehungen folgendermaßen graphisch dargestellt (1969, 116):

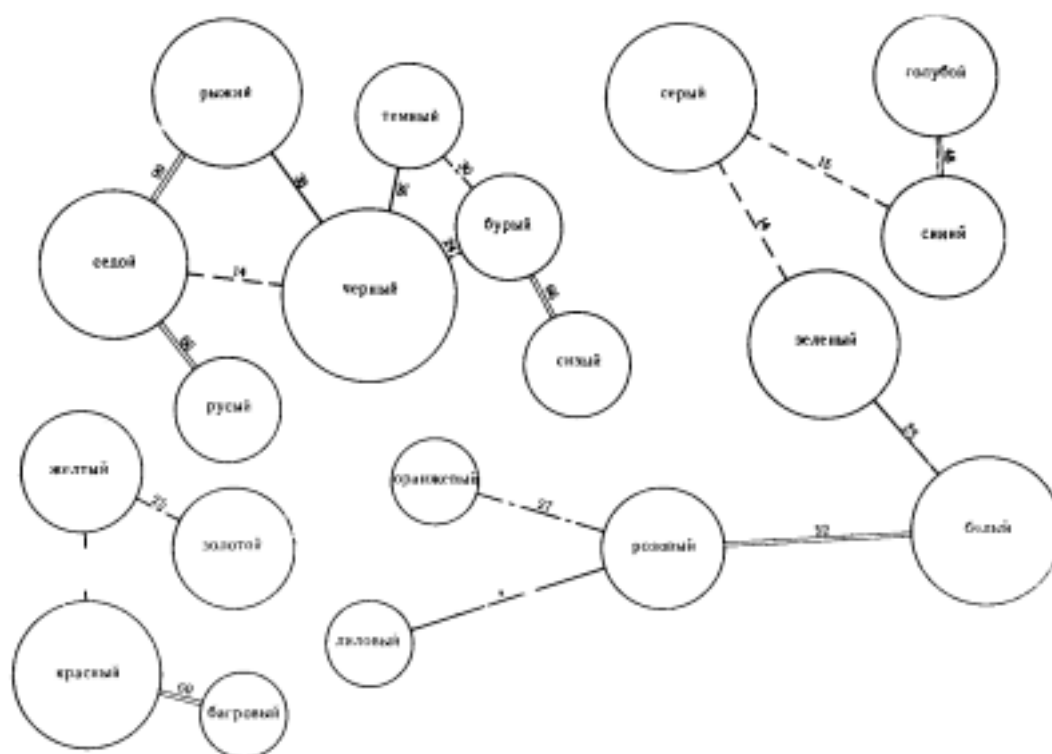


Fig. 3: Die Struktur des Farbwortfeldes im Russischen

Das heißt, die Farbwortbezeichnungen werden in vier Klassen eingeteilt:

a) *черный, рыжий, седой, русый, темный, бурый, сизый*;

- b) *красный, багровый, желтый, золотой;*
- c) *белый, зеленый, розовый, оранжевый, лиловый;*
- d) *серый, голубой, синий.*

Vornamen Eine "Hitparade" der beliebtesten Vornamen in Rußland findet sich bei NIKONOV (1967), getrennt nach männlichen und weiblichen Vornamen, vor allem aber genau aufgeschlüsselt nach städtischen und ländlichen Gebieten. *Елена, Марина* und *Ирина* sind danach in Städten viel beliebter als auf dem Land, umgekehrt sind *Галина* und *Татьяна* "ländliche" Vornamen. Auf dem Land wird der stereotype *Иван* bis zu 10 mal häufiger vergeben als in der Stadt, während die "warägischen" Vornamen *Игорь* und *Олег* klar "Stadtnamen" sind. Eine ähnliche Untersuchung stammt von BONDALETOV (1972, 1973), der in getrennten Arbeiten die Verbreitung männlicher und weiblicher Vornamen in der Stadt Penza betrachtet. Für die drei Perioden Vorrevolutionäre Zeit – Revolution bis Zweiter Weltkrieg – Nachkriegszeit gibt der Autor zu jedem Namen seine Verwendungshäufigkeit nach Ausweis der standesamtlichen Unterlagen an. Diese Daten zeigen beispielsweise, daß bereits die zehn häufigsten Vornamen in der Zeit vor der Revolution mehr als die Hälfte aller Namensgebungen abdecken; diese Zeit nennt der Autor deshalb "wenig beweglich und eintönig" (1973, 830). Nach der Revolution drängen andere Namen nach vorn, darunter etwa *Тамара* und *Надежда, Владимир, Борис* und *Виктор*, das Spektrum wird insgesamt weiter. Die vom Autor als solche bezeichnete "anthroponymische Flut" ließ jedoch bereits in den 30er Jahren wieder nach. Die Angaben der diversen Tabellen und Aufstellungen lassen das Auf und Ab in der Popularität einzelner Namen deutlich erkennen – der sprichwörtliche *Иван* beispielsweise ist nach diesen Daten von seinem Anteil von 8,5% am Ende des 19. Jhs. abgesunken auf einen bedeutungslosen Wert weit unter einem Prozent.

Flurnamen Strukturelle Typen von Flurnamen untersucht PLESKALOVÁ (1984) in mehreren slawischen Sprachen und gibt für jeden Typus seinen relativen Anteil an. Von insgesamt fünf aufgestellten Typen sind danach drei für das Russische (oder jedenfalls für die Stichprobe aus dem Kirover Gebiet) charakteristisch und dominierend (mit einem Anteil von einem Viertel bzw. einem Drittel), der Rest ist bedeutungslos. In den drei untersuchten westslaw. Sprachen sind immer die gleichen zwei Modelle relevant, das Bulgarische weist eine deutlich andere Verteilung auf.

Lexikalische Stilistik Die stilistische Färbung des russischen Wortschatzes untersuchen DENISOV/KOSTOMAROV (1969), indem sie das Wörterbuch von Ožegov komplett daraufhin auswerten, wie oft in diesem Material stilistische Markierungen benutzt werden. Es ergab sich, daß ziemlich genau ein Drittel des Wortschatzes in diesem Wörterbuch eine Markierung wie "hochsprachlich", "buchsprachlich", "umgangssprachlich", "veraltet" usw. trägt. Den "Löwenanteil" machen dabei ganz ein-

deutig die Markierungen "umgangssprachlich" und "vulgär" aus, Fachsprache, Buchsprachliches und Veraltetes folgen. Die gleichen Daten finden sich – in leicht abgewandelter Form – bei GOLUBZOWA (1991, vgl. 193-197). Die gleiche Autorin legt auch eigene Zahlen zu den von ihr untersuchten Adverbien vor (vgl. 262-269), zu denen sie ebenfalls die stilistische Markierung angibt, sie gleichzeitig aber noch nach weiteren Kriterien differenziert, vor allem nach dem Wortbildungsverfahren. In einigen Fällen scheint es in dieser Kreuzklassifikation tatsächlich signifikante Zusammenhänge zu geben – ein Adverb, das auf eine erstarrte Kasusform zurückgeht, ist z.B. niemals "buchsprachlich".

3.8. Soziolinguistik, Sprache und Norm

Sprache und Gesellschaft Soziolinguistische Untersuchungen zum Russischen sind primär mit dem Namen KRYSIN verknüpft. Dieser Autor hat – mit seinen Mitarbeiterinnen – die einzige große empirische soziolinguistische Untersuchung zum Russischen unternommen und die Ergebnisse publiziert, vor allem in den beiden Monographien KRYSIN (1974) und KRYSIN/ŠMELEV (1976); hinzu kommt eine größere Zahl von vorbereitenden, begleitenden und ergänzenden Aufsätzen (KRYSIN 1973, 1975). Zur (sowjetischen) Soziolinguistik vgl. u.a. auch GIRKE/JACHNOW/SCHRENK (1972), GIRKE/JACHNOW (1974), JACHNOW (1984), KRAUS (1974).

Die empirischen Daten wurden mittels Fragebögen in den Jahren 1963 bis 1966 erhoben; Probanden waren 3.000–4.000 muttersprachliche, gebildete Sprecher des Russischen aus größeren Städten (um Dialekteinflüsse auszuschließen). Sie waren aufgefordert, unvollständige Sätze zu vervollständigen (Lückentest), Auswahlen aus vorgegebenen Alternativen oder Bewertungen ("Was halten Sie für richtiger?") vorzunehmen. Die Auswertung der Ergebnisse erfolgt praktisch immer nach dem gleichen Schema: Die relativen Anteile der einzelnen Glieder einer Variationsreihe werden differenziert nach Alter, Beruf und Bildungsstand sowie der Herkunft der Sprecher, wobei diese Kriterien pragmatisch und nicht sozialistisch-dogmatisch gehandhabt werden, allerdings auch nicht so differenziert sind, wie dies im Westen üblich wäre (beispielsweise kommen keine "Schichten" vor). Neben der tabellarischen Wiedergabe wird vor allem das Liniendiagramm eingesetzt, das vor allem für die Darstellung zeitlicher Entwicklungen Sinn macht (vgl. KEMPGEN 1992 zu einigen methodischen Mängeln). Statistische Tests erfolgen selten, ihre Wichtigkeit wird aber in den grundlegenden Arbeiten erkannt und demonstriert (vgl. KRYSIN 1975). Die inhaltliche Auswertung kommentiert zunächst gewöhnlich die untersuchte Variation in synchroner wie diachroner Perspektive, um dann auf die Stichprobendaten einzugehen, wobei die subjektiv auffälligsten Veränderungen und Abhängigkeiten benannt und kommentiert werden. Die angebotenen "Erklärungen" für das Sprecherverhalten operieren mit der "Nähe zu den hochsprachlichen Zentren", der dialektalen wie ostslawischen Umgebung, dem Grad der Vertrautheit mit den und der Internalisierung der hochsprachlichen Normen des Russischen, um nur

die häufigsten Begründungen zu nennen. Generell läßt sich festhalten, daß das Alter der Sprecher fast immer einen sehr wichtigen Einfluß darstellt, wobei dieses Kriterium natürlich mit anderen positiv korreliert ist (je älter, desto höher und "philologisch" kann die Bildung sein). Das heißt aber auch: soziale Merkmale *sui generis* spielen – untersuchungsbedingt – oft eine eher untergeordnete Rolle. Geschlechtsspezifische Unterscheidungen werden selbst dort nicht gemacht, wo sie nahegelegen hätten (Berufsbezeichnungen für Frauen).

KRYSIN (1974) enthält konkret Daten zu Phonetik/Phonologie sowie Flexions- und Derivationsmorphologie. Im ersten Bereich finden sich Phänomene wie die assimilative Erweichung der Konsonanten (vgl. *снег* – [s'n'ek]), aufgeschlüsselt nach verschiedenen artikulatorischen Klassen, die Härte bzw. Weichheit langer Zischlaute (vgl. *позже* – [p'ož':e]) sowie die Aussprache unbetonter Vokale nach Zischlauten (vgl. *жара* – [žar'a]) und in Endungen (vgl. *поле* – [p'ol'ɪ]). Variationen im Bereich der Flexionsmorphologie betreffen den sog. "zweiten (partitiven) Genitiv" (vgl. N. *сахар*, G. *сахара* vs. *сахару*), den "zweiten Präpositiv" (vgl. N. *снег*, P. *в снеге* vs. *в снегу*), den Nom.Pl. auf *-а* (vgl. Sg. *лектор*, Pl. *лекторы* vs. *лектора*), die Flexion von Ortsnamen auf *-ино/-ово* (vgl. *к Иваново* vs. *к Иванову*), die Flexion ursprünglich ukrainischer Familiennamen auf *-ко*, die Flexion bestimmter Verbgruppen (vgl. Inf. *капать*, Präs. *каплет* vs. *капает*) sowie Akzentvarianten. Im Bereich der Wortbildung werden Derivationsvarianten bei Berufsbezeichnungen untersucht (*трамбовщик* – *трамбовальщик* – *трамбователь*), Bezeichnungen für Frauen (*лифтер* – *лифтерша*, *прыгун* – *прыгунья*) und einiges mehr.

Quasi als Fortsetzung hierzu können die Beiträge im Sammelband KRYSIN/ŠMELEV (1976) betrachtet werden. GLOVINSKAJA beschäftigt sich ausführlich mit der Aussprache von Fremdwörtern im Russischen (Erweichung vor [e], Länge von Doppelkonsonanten, Aussprache eines unbetonten [o]). KUZ'MINA untersucht die Vereinfachung von Konsonantengruppen (vgl. *студентка* – [nk]), VORONCOVA die Genitiv Plural-Endung bestimmter Ethnonyme, Obstnamen und Maßbezeichnungen (*граммов* vs. *грамм*). Die Konkurrenz zwischen formaler und semantischer Kongruenz bei Wörtern wie *врач*, *управдом* etc. untersucht KITAJGORODSKAJA, getrennt nach adjektivischem Attribut und verbalem Prädikat. Der Ausfall des Suffixes *-ну* im Präteritum ist Gegenstand von IL'INAS Beitrag, GOLANOVA behandelt Derivationsvarianten bei bestimmten Werkzeug- oder Gerätebezeichnungen. – Alle hier genannten Autorinnen sind auch die Verfasserinnen entsprechender Teile des Buches von KRYSIN (1974).

Auf die gleichen empirischen Daten greifen auch die von PANOV (1968) herausgegebenen vier Bände "*Russkij jazyk i sovetskoe obščestvo*" zurück. Allerdings stehen hier sprachstatistische Auswertungen nicht im Vordergrund; nur in den beiden Bänden zu Phonetik und Morphologie finden sich einige Tabellen.

Sprache und Norm Eng verwandt mit soziolinguistischen Arbeiten sind die Untersuchungen zum Verhältnis von Sprache und Norm, die, soweit sie statistisch

arbeiten, oft ebenfalls konkrete Analysen einzelner Variantenreihen vorlegen (vgl. den Sammelband JAZYKOWAJA NORMA I STATISTIKA 1977, GRAUDINA 1965, 1966, 1970, 1975, 1977). Auf OŽEGOV geht dabei die – durchaus plausible – Auffassung zurück, daß die Norm nicht einfach mit den statistischen Beobachtungen, wie tatsächlich gesprochen werde, gleichgesetzt werden können (hierbei handelt es sich ja in der Tat eher um den Usus). Dennoch betonen Autoren wie v.a. GRAUDINA, daß statistische Untersuchungen sehr wohl geeignet seien, Normfragen zu lösen, und daß auch der Begriff der “Ausnahme” quantitativ faßbar sei (nämlich als signifikante Abweichung von der normalen Häufigkeitsverteilung einer Erscheinung). Grundsätzliches vgl. bei PIOTROVSKIJ/TURYGINA (1971), BOGDANOV (1973).

Konkret gehen die quantitativen Untersuchungen zur Sprachvariation so vor, daß den üblichen Grammatiken des Russischen bekannte Variationsreihen entnommen werden (z.B. miteinander konkurrierende Kasusformen, Kongruenzprinzipien, Derivationsmuster u. dergl.) und empirisch durch Fragebögen mit Ergänzungsaufgaben ermittelt wird, welche Variante wie häufig von den Befragten benutzt wird.

In GRAUDINA (1966 = 1975) beispielsweise wird der Genitiv Singular nichtzählbarer mask. Substantive untersucht (*чашка чаю* vs. *чашка чая* “eine Tasse Tee”), wobei sich für *чая* ein Anteil von 58,4% und für *чаю* ein Anteil von 40,0% ergab. Da diese eine Stichprobe natürlich nur eine Annäherung an das “wirkliche” Verhältnis der beiden Formen zueinander sein kann, wird mit einem statistischen Test die Spannbreite des “wahren” Wertes ermittelt. Außerdem wird gezeigt, daß sich die Anteile der beiden Formen im Laufe der Zeit verschoben haben (*чая* nimmt zu) und auch von soziolinguistischen Faktoren abhängen. Auf diese Weise soll demonstriert werden, daß die Norm nicht einfach einem Stichprobenmittelwert gleichzusetzen ist, sondern daß er eine gewisse Spannbreite besitzt und auf der Zeitachse veränderlich ist. Der Beitrag von GRAUDINA (1965) untersucht den Genitiv Plural des Wortes *носки*, der *носков* (alte Norm) oder *носков* (Neuerung) lauten kann; die statistischen Daten zeigen die steigende Verbreitung der neuen Form; weitere Phänomene werden ebenfalls von der gleichen Autorin untersucht (vgl. 1964a, 1964b).

Ihren bedeutendsten Ausdruck haben diese quantitativen Normuntersuchungen in dem Nachschlagewerk von GRAUDINA/ICKOVIČ/KATLINSKAJA (1976) gefunden, das eine Fülle nicht zu unterschätzenden Materials für das Russische bereitstellt, und zwar vornehmlich für die Bereiche Flexion (Endungsvarianten) und Morphosyntax (Kongruenz). Eine ins einzelne gehende Behandlung der zahlreichen Themen dieses Wörterbuches ist an dieser Stelle jedoch aus Platzgründen leider nicht möglich.

3.9. Psycholinguistik

Das von LEONTEV (1977) herausgegebene Wörterbuch führt als seine – alphabetisch angeordneten – Einträge 196 von 500 insgesamt getesteten Stimuli auf, zu denen sämtliche Responses geordnet nach Häufigkeit angegeben werden. Einleitend behandelt eine Reihe von Aufsätzen einschlägige theoretische und praktische Fragen von Assoziationstests. Die häufigsten Reaktionen auf *русский* waren beispielsweise *язык* 40, *человек* 36, *народ* 12, *казах* 10, *немец* 8, *лес* 7 usw. Zu jedem Stimulus werden ferner sein Rang und seine Frequenz nach den drei wichtigsten Häufigkeitswörterbüchern angegeben, außerdem die Gesamtzahl der Antworten (meist ca. 600-700). Probanden waren erwachsene Muttersprachler (16-50 Jahre) mit höherer Bildung, die 1969-1972 befragt wurden; notiert wurde von den Testkandidaten jeweils nur das erste Wort (bzw. die erste Wortverbindung), das (die) ihnen zum gegebenen Stimulus in den Kopf kam. Das Wörterbuch soll die häufigsten Syntagmen und typische semantische Felder verdeutlichen, in die ein gegebenes Wort eingeht, und in erster Linie praktischen Zwecken dienen (Erstellung von Materialien für den landeskundlichen Fremdsprachenunterricht), nicht so sehr psycholinguistischen Fragestellungen. Das vorliegende Wörterbuch ist die erste größere Untersuchung dieser Art zum Russischen, während es zu den wichtigsten westlichen Sprachen schon seit langem entsprechende Daten gibt (die in dem Wörterbuch auch genannt werden). Das hier vorgelegte Material eröffnet eine ganze Reihe von interessanten Auswertungsmöglichkeiten, neben rein linguistischen auch theoretisch-methodische, vor allem aber speziell kulturwissenschaftliche.

Einem psycholinguistischen Experiment nach amerikanischem Vorbild sind zwei Arbeiten von FRUMKINA (1966b, 1971) gewidmet. Die Autorin untersucht an einigen russischen Beispielwörtern die theoretische Frage, ob Sprecher in der Lage sind, solche Stimuli nach ihrer vermutlichen Häufigkeit zu ordnen. Die Antwort lautet, daß diese "subjektive Wahrscheinlichkeit" sehr genau differenziert ist und der an Texten ermittelten Wahrscheinlichkeit gut entspricht. Die Schlußfolgerung freilich, daß solche subjektiven Wahrscheinlichkeiten anstelle zeitaufwendig zu ermittelnder Textdaten in der Linguistik verwendet werden können, ist wohl durch die technische Entwicklung überholt.

Methoden der Ähnlichkeitsmessung von Paraphrasen untersuchen ORLOV/ŽURAVLEV (1967) an einem russischen Beispiel. Als psycholinguistisches Experiment kann man auch die Farbwortuntersuchung von MORGAN/CORBETT (1989) verstehen. Sie wurde oben bereits vorgestellt.

3.10. Dialektologie

Die in Westeuropa seit längerem betriebene quantitative Analyse und Bearbeitung dialektologischer Daten ("Dialektometrie") hat in Rußland bislang nicht recht Fuß fassen können. Dies betrifft sowohl die Zählung sprachlicher Einheiten, wie man sie auch an der Hochsprache durchführen kann, vor allem aber den Bereich,

der spezifisch für die Dialektologie ist, nämlich die Dialektklassifizierung mit Clusterverfahren, die mehrdimensionale Darstellung dialektologischer Daten mit entsprechenden Verfahren, die maschinelle Erstellung von Sprachatlanten usw.

Eine leider wenig beachtete frühe Anwendung der Verfahren der sog. "Numerischen Taxonomie" stellt der Aufsatz von PŠENIČNOVA (1977a, vgl. auch 1973, 1977b, 1979a, 1979b) dar. In ihm werden die Dialekte im Gebiet Ladoga – Novgorod – Toržok – Pskov erst in Form einer Merkmalsmatrix beschrieben, auf dieser Grundlage ihre gegenseitige Ähnlichkeit berechnet und schließlich eine Klassifikation erarbeitet, die der "traditionellen" Dialektgliederung des gleichen Gebietes gegenübergestellt wird. (Die Numerische Taxonomie geht auf SOKAL/SNEATH – vgl. die klassische Arbeit von 1963 – zurück.)

Unter weiteren dialektologischen Arbeiten sei hier DIKAREV (1891) genannt, zugleich einer der Vorläufer der Sprachstatistik in Rußland überhaupt. Dieser Autor untersucht ausführlich die Lauthäufigkeiten in den Dialekten von Priluck und Vornež, deren Daten er häufig mit solchen zum Bulgarischen, Serbischen, Čechischem und Polnischen vergleicht. Er stellt aber nicht nur Einzellaute, die er auch nach artikulatorischen Klassen ordnet, in eine Rangreihenfolge, sondern publiziert auch ausführliche Tabellen zu den zweigliedrigen Lautverbindungen der beiden untersuchten Dialekte (vgl. 1891, 62–67). Daß in der damaligen Zeit Laute und Buchstaben nicht konsequent getrennt wurden und die Phonologie als Disziplin noch nicht entwickelt war, kann einem einzelnen Autor natürlich nicht zum Vorwurf gemacht werden, beeinträchtigt aber dennoch die ungeprüfte Übernahme und Weiterverwendung dieser Daten.

Aus der Gegenwart gibt es quantitative Untersuchungen mehrerer Dialekte auf dem Gebiet der damaligen UdSSR, nämlich der *deutschen* Dialekte von Najdorf (vgl. VEJLERT 1967, 1973, 1974) bzw. Kant (vgl. GOOGE 1973, 1974). Zu einer auch das Russische betreffenden Arbeit vgl. VEJLERT (1979).

3.11. Statistik und Stil, Poetische Sprache

Der Bereich "Statistik und Stil" gehört zu denjenigen, die besonders von der (modernen) Prager Schule und ihren Ausläufern bearbeitet wurden, vgl. die Bibliographie von BAILEY/DOLEŽEL (1968) und den Überblick von BAILEY (1969, 1979); Grundsätzliches bei DOLEŽEL (1964, 1965, 1967, 1972), TARANOVSKIJ (1966b), PEREBEJNOS (1974), im Sammelband DOLEŽEL/BAILEY (1969) mit vielen wichtigen Arbeiten, bei KREUZER/GUNZENHÄUSER (1965, mit Bibliographie), KOŽINA (1968), KAUFMAN (1970), HOFFMANN/PIOTROWSKI (1979, 148ff.) usw. In der Sowjetunion hat sich besonders der Lehrstuhl für Russische Sprache der Universität Gor'kij unter Leitung von B.M. Golovin mit dem Zusammenhang von Funktionalstil und (russischer) Grammatik beschäftigt; einige Ergebnisse einschlägiger Diplomarbeiten werden bei GOLOVIN (1967) vorgestellt. Der "Klassiker" dieses Bereiches liegt jedoch in ukra-

inischer Sprache vor und zieht ukrainisches Material als Beispiel heran (PEREBEJNOS 1967a).

Zum Thema Satzlänge, die ja auch ein wichtiges stilistisches Merkmal darstellt, vgl. den Abschnitt zur Syntax, Häufigkeitwörterbücher zu einzelnen Autoren (v.a. Puškin) wurden im Zusammenhang mit allgemeinen Häufigkeitwörterbüchern behandelt (s.o.).

Versstruktur Die Versstruktur hat bereits in den Anfängen der QL in Rußland große Aufmerksamkeit gefunden, vgl. auch die Überblicksartikel von LEVÝ (1965) und BAILEY (1979). TOMAŠEVSKIJ (1923, 1929), Mitglied des sog. "Moskauer linguistischen Zirkels", wendet die Theorie der Markov-Ketten in mustergültiger Weise (er war ausgebildeter Statistiker) auf Probleme der Prosodie, insbesondere des Versrhythmus, u.a. bei Puškin, an. Ein fundamentales Werk stammt von ŠENĀELI (1923, s.a. 1960), nachdem das grundlegende Programm einige Jahre zuvor von dem Symbolisten Andrej BELYJ (1910) formuliert – dann aber methodisch naiv umgesetzt – worden war (vgl. 1929).

Besondere Aufmerksamkeit hat dann in den 60er Jahren insbesondere das Schaffen von V.V. Majakovskij gefunden, und zwar insbesondere die Frage, ob er einen neuen Verstypus geschaffen oder "nur" die klassische syllabotonische Dichtung deformiert hat. Der Klärung dieser Frage ist der mit ausführlichen Tabellen versehene Artikel von KONDRATOV (1962b) gewidmet, dessen Ergebnisse bei KOLMOGOROV (1963) z.T. methodisch korrigiert werden. Eine gemeinsame Arbeit beider Autoren (KOLMOGOROV/KONDRATOV 1962) untersucht ebenfalls ein metrisches Muster bei Majakovskij, siehe auch KOLMOGOROV/PROCHOROV (1963, 1964), GASPAROV (1965, 1967b, 1972), KOLMOGOROV (1965). Eine populäre Darstellung der bis dato erzielten Ergebnisse findet sich in der Broschüre von KONDRATOV (1962a). Direkt von KOLMOGOROVs Ansatz ist der Beitrag von IVANOV (1966a) inspiriert, der sich des Rhythmus' eines weiteren Majakovskij-Gedichtes annimmt. Die Versstruktur dieses Autors ist auch Gegenstand der Untersuchung von GASPAROV (1967a).

ABERNATHY (1963b) vergleicht die statistische Abfolge betonter Vokale in einem Gedicht von Blok mit der Struktur der Fuge. BOBROV (1964) nimmt sich des Rhythmus' eines Puškin-Gedichtes an. GASPAROV (1963) ermittelt auf breiter empirischer Grundlage verschiedene Typen des russischen *dol'nik*. Die Verteilung der betonten Vokale in russischen Versen des 19. und des 20. Jh. sind Untersuchungsgegenstand von JONES (1965). Einen "Schwierigkeitskoeffizienten" für russische Reimschemata stellt KONDRATOV (1963a) auf, um so zu einer Klassifikation zu kommen; bei ihm finden sich auch ausführliche Tabellen zur Reimstatistik für Werke von Lomonosov bis ins 20. Jahrhundert. TARANOVSKIJ (1966, vgl. auch 1953, 1955-56) untersucht speziell den Jambus Andrej Belyjs, BEL'SKAJA (1974) denjenigen bei Esenin, LEVIN (1964, 116) gibt einige Werte zum gleichen Versmaß bei verschiedenen Autoren, um so deren Ähnlichkeit berechnen zu können. BAILEY (1970) verfolgt

den Gebrauch des 5+5-Trochäus im Zeitraum von 1790 bis 1920, eine weitere Arbeit (1973) ist Andrej Voznesenskij gewidmet. LEVINTON (1981) zeigt, daß vor einem Reimauslaut auf *-ю/ы*, der Struktur des Russischen entsprechend, hauptsächlich Sonanten (*j, л, н, р*) stehen. ŠAPIR (1992) untersucht den Jambus bei Griboedov und seinen Zeitgenossen, SMITH (1985) analysiert das metrische Repertoire speziell der neueren Emigrationsliteratur.

Die größte systematische Untersuchung, geplant als Beginn einer Serie, stellt bisher der Sammelband RUSSKOE STICHOSLOŽENIE (1979) dar. In ihm werden die Gedichte von V.A. Žukovskij, K.N. Batjuškov, A.Ch. Vostokov, A.S. Puškin, A.A. Del'vig, E.A. Baratynskij, A.V. Kol'cov, F.I. Tjutčev und Ja.P. Polonskij nach Versmaß, Reimschema, Umfang und historischer Entwicklung aufgeschlüsselt und statistisch ausgewertet. Zu den gleichen Autoren siehe auch PAPINA/IVANOVA-MARKOVA (1974).

Weitere Untersuchungen Mit der Entwicklung der Wortfolge in polnischen und russischen Versen des 18. und des 19. Jhs. beschäftigt sich WIERZBICKA (1966). HU (1968) ermittelt statistische Charakteristika in russischen Totenklagen (Versstruktur, Übereinstimmungen der Vokale), ähnlich ROBBLEE (1989). Eine Stiluntersuchung M. Šolochovs unternimmt LJATINA (1968). WEDEL (1993) präsentiert eine Statistik zur Struktur russischer Romantitel des 19.Jhs. (eingliedrig, zweigliedrig, mit/ohne Untertitel etc.). Ihr läßt sich u.a. entnehmen, daß die eingliedrigen Titel zugenommen haben. SCHREINER (1989) untersucht den Gedichtumfang bei A. Achmatova, B. Pasternak, O. Mandel'stam und M. Cvetaeva und belegt ihre Funde mit ausführlichen Tabellen, die die Frage nach verschiedenen zusätzlichen Kriterien aufschlüsseln.

KJETSAA (1972) gibt eine Statistik der Übersetzungen von Lermontov-Werken in skandinavische Sprachen, in einer weiteren Arbeit (KJETSAA 1973) finden sich Rang-Frequenzlisten zu Substantiven, Verben und Adjektiven in den Werken des gleichen Autors, der zudem mit Baratynskij und Tjutčev verglichen wird. Die Farbsymbolik bei Lermontov untersucht GIL (1973) und stellt eine Häufigkeitsliste der Farbwörter bei diesem und sechs weiteren Autoren zusammen. Eine ausführliche Lautstatistik bei Lermontov stammt wieder von KJETSAA (1974), während GASPAROV das von LAPŠINA/ROMANOVIC'/JARCHO gesammelte statistische Material zum Versmaß publiziert hat (1966).

Zur Illustration der Notwendigkeit statistischer Methoden in der Stilanalyse untersucht GOLOVIN (1968) die Häufigkeit von Pronomen, Adverbien, Konjunktionen und Partikel bei Dostoevskij und Tolstoj in jeweils 10 Stichproben zu 500 Wörtern. Dabei zeigt sich, daß Dostoevskij von allen diesen Wortarten beständig einen höheren Anteil verwendet als Tolstoj, wobei dieser Anteil bei Dostoevskij außerdem viel konstanter zu sein scheint. Als zweites Beispiel dient eine Analyse des Substantivgebrauches bei A. Tolstoj und K. Simonov, verglichen mit Alltags- und Zeitungstexten. Der gleiche Autor untersucht die Entwicklung der Wortartenverteilung

lung im 19. und 20. Jh., wiederum anhand belletristischer Stichproben (GOLOVIN 1965). Die Verteilung der Wortarten (und einiger grammatischer Kategorien des Verbs) bei Tjutčev untersucht BILOKUR (1970) und vergleicht die ermittelten Werte mit denen der üblichen Frequenzwörterbücher. AKIMOVA (1973, 1977) gibt einige Daten zur mittleren Satzlänge bei Lomonosov und Sumarokov, MARTYNENKO (1972, 1983) klassifiziert Autoren aufgrund syntaktischer Merkmale, wobei Grundlage des Vergleiches Eigenschaften wie Satzlänge und -komplexität sind. Die Klassifikation wird als Graphik dargestellt (1983, 69):



Fig. 4: Klassifikation russischer Autoren nach syntaktischen Merkmalen

Autorschaftsnachweis Das Unterfangen, einen Autorschaftsnachweis mit statistischen Methoden zu führen, gehört zu den klassischen Anwendungsfällen der QL; die zwei Aufsätze von KJETSAA (1976, 1978) und das Buch von KEJTSAA et al. (1984) beschäftigen sich mit dem bedeutendsten einschlägigen Fall in der russischen Literaturgeschichte, nämlich der Autorschaft des "Stillen Don", für den Michail Šolochov den Literaturnobelpreis erhalten hatte, dessen Autor aber in Wirklichkeit der Kosakenschriftsteller Fëdor Krjukov gewesen sein soll. Dies jedenfalls behauptete ein gewisser "D*" in einem 1974 im Westen erschienenen Buch, das seinerzeit viel Staub aufwirbelte, unter anderem deshalb, weil kein geringerer als Solženicyn das Vorwort zu dem Buch geschrieben hatte. Durch diese Auseinandersetzung veran-

laßt, machte sich daraufhin eine Gruppe skandinavischer Slawisten mit Computerhilfe an eine empirische Überprüfung dieser Anschuldigungen.

In der ersten Arbeit, als "Pilotstudie" deklariert, wird die Verteilung der Satz­längen aus dem "Stillen Don" mit unzweifelhaften Werken der beiden infrage kom­menden Autoren verglichen, jeweils auf der Grundlage von 1000 zufällig erhobe­nen Sätzen. Die resultierenden Verteilungen stimmen für Šolochov und den "Stillen Don" sehr gut überein, während Krjukov mit einem deutlich geringeren Anteil kur­zer Sätze und mehr längeren Sätzen von beiden abweicht. Eine zweite Analyse un­tersucht die Verteilung und die Kombination der Wortarten in Stichproben von je­weils 10.000 Wörtern bei den beiden Autoren und im "Stillen Don", und zwar all­gemein und bezogen auf bestimmte Positionen im Satz (Anfang, Ende). Ein statisti­scher Test ergibt auch hier signifikante Unterschiede zwischen Krjukov und dem "Stillen Don", während die Verteilungen von Šolochov und "Stillem Don" als nicht signifikant verschieden betrachtet werden können. In der zweiten Arbeit werden außerdem noch die Verteilung der Wortlängen (in Buchstaben) und der "lexikali­sche Reichtum" der Texte untersucht (d.h. das *type-token*-Verhältnis). Auch hier ist Šolochov dem "Tichij Don" immer näher als Krjukov es ist. Dies alles besagt aber nur folgendes: Die Wahrscheinlichkeit dafür, daß Krjukov den veröffentlichten Text des "Tichij Don" geschrieben hat, ist gering; er entspricht vielmehr Šolochovs Stil – jedenfalls in den hier untersuchten Punkten. Das bedeutet aber nicht, daß Krjukov nicht doch der ursprüngliche Verfasser eines Rohentwurfes gewesen sein kann, der durch Šolochov dann nur noch bearbeitet wurde; dies gesteht auch KEJTSAA zu, will Šolochov aber vom Vorwurf des Plagiariats freisprechen. – Ist es ein Zufall, daß keine von den zahlreichen sowjetischen sprachstatistischen Arbeiten sich jemals dem gleichen Problem gewidmet hat?

Dem Autorschaftsnachweis bei anonymen russischen Werken ist das Buch von ERMOLENKO (1988, vgl. auch 1977) gewidmet; den wohl frühesten Ansatz finden wir aber in der Arbeit von MOROZOV (1915). Dieser Autor versucht sich an Stilpa­rametern wie dem Präpositionsgebrauch u.a., die er bei Gogol', Puškin und Tolstoj untersucht (vgl. die ausführlichen Tabellen pp. 128–132). Seine Versuche, Indizes zu definieren bzw. empirische Daten graphisch umzusetzen, sind jedoch noch recht unbeholfen. Eine umfassende Darstellung unter Anwendung statistischer Methoden findet sich jetzt bei MILOVA (1994).

3.12. Sprachtypologie und Universalien

Arbeiten zur Sprachtypologie und Universalien, die das Russische berücksich­tigen, sollen hier nur kurz genannt werden, da sich in diesem Bereich in der Regel andere, weitergehende methodische Probleme stellen, deren Behandlung nicht Gegenstand der vorliegenden Arbeit ist. Die Entwicklung speziell der quantitativen morphologischen Sprachtypologie, die auf den Greenbergschen Indizes aufbaut, wird von LEHFELDT/KEMPGEN (im Druck) ausführlich dargestellt.

Distinktive Merkmale Die Ausnutzung und Relevanz der distinktiven Merkmale in den phonologischen Systemen der slawischen Sprachen machen LEKOMCEVA et al. (1963) zur Grundlage mehrerer Indizes. Die Ergebnisse werden benutzt, um die untersuchten Sprachen in eine Rang-Reihenfolge zu bringen. Den Anteil der Konsonanten bzw. Vokale am Phonemsystem macht ISAČENKO (1939/40) zur Grundlage einer Rangreihenfolge der slawischen Sprachen (vgl. hierzu auch KEMPGEN 1991). ALTMANN (1971) charakterisiert die slawischen Sprachen hinsichtlich der Ausnutzung der distinktiven Merkmale in Form eines sog. "Profilvektors", der die Grundlage für eine Berechnung der Ähnlichkeit der Sprachen zueinander abgibt, wobei auf deren Grundlage wiederum eine hierarchische Klassifikation aufbaut. In dem resultierenden Baumdiagramm bildet das Russische zusammen mit dem Ukrainischen eine Untergruppe. Nach LEHFELDT (1975) ist das Russische eine Sprache mit "normal" vielen Phonemen, wie alle anderen slawischen Sprachen auch.

Phonemkombinationen Das Russische wird – neben einigen anderen slawischen Sprachen – in der klassischen Arbeit von GREENBERG (1964) berücksichtigt, in der eine ganze Reihe von Implikationen bezüglich Zahl und Art der Konsonantencluster formuliert werden. TOLSTAJA (1968c, 1974) ermittelt die initialen und finalen Konsonantenverbindungen einer Reihe slawischer Sprachen, um dann zu ganzheitlichen Charakterisierungen der betreffenden Subsysteme zu kommen, etwa hinsichtlich ihrer internen Symmetrie, hinsichtlich des Ausnutzungsgrades der verschiedenen Clustertypen etc. LEHFELDT (1972) klassifiziert die Phonemverbindungen in 15 slawischen Sprachen bzw. Dialekten zunächst nach dem Grad ihrer Assoziations- bzw. Dissoziationskraft, ermittelt dann, wie stark die einzelnen Assoziationsklassen besetzt sind, um auf der Grundlage einer Ähnlichkeitsberechnung schließlich eine hierarchische Klassifikation der slawischen Sprachen zu erarbeiten, die als Baumdiagramm dargestellt wird. Das gleiche Vorgehen wird anschließend für die Ausnutzung der distinktiven Merkmale und schließlich kombiniert für beide Bereiche durchgeführt. In der kombinierten Klassifikation bildet das Russische zusammen mit dem Niedersorbischen und dem Polnischen eine Untergruppe. Die Phonemdistribution ist auch die Grundlage der Klassifikation von ALTMANN/LEHFELDT (1972) und KLEINLOGEL/LEHFELDT (1972); auch hier wird das Russische neben einigen anderen slawischen Sprachen berücksichtigt.

Genussystem Die Ähnlichkeit einiger slawischer Sprachen zueinander hinsichtlich der Struktur ihres Genussystems berechnet KARPINSKAJA (1964) auf der Grundlage eines entsprechenden Indexes. Die Ergebnisse stellt sie auch in Form eines Graphen dar.

Analytizität/Synthetizität VOLOCKAJA et al. (1963) untersuchen die Flexionsendungen der Substantive in den slawischen Sprachen und berechnen auf

deren Grundlage zwei Indizes: den "Grad der Abweichung von einer synthetisch einfachen Sprache" und den "Grad der Abweichung von einer morphologisch einfachen Sprache".

Silbenstruktur Die Struktur der Silbe als eines besonders für die slawischen Sprachen äußerst wichtigen Strukturelementes ist in mehreren Arbeiten untersucht worden. ZAPLATKINA (1980) ermittelt die Strukturtypen einsilbiger Wörter (d.h. die Arten der Abfolge von Vokalen und Konsonanten), deren durchschnittliche Länge, den Anteil der tatsächlichen an den theoretisch möglichen Silbentypen, und berechnet auf dieser Grundlage die Ähnlichkeit der slawischen Sprachen untereinander.

Farbwörter Das Feld der Farbwortbezeichnungen ist in vielen Sprachen nach der Häufigkeit ähnlich gestaffelt (vgl. ARAPOV et al. 1975, 6), im Russischen *Weiß – Rot – Schwarz – Blau – Grün – Grau – Gelb – Rosa*. Vgl. auch VASIL'EVICĀ/SKOKAN (1986) und MORGAN/CORBETT (1989). Die Farbwörter sind auch das semantische Feld, das MOSKOVIĀ (1965) vergleichend am Beispiel des Russischen, Englischen und Französischen untersucht (s. Abschnitt "Semantik").

Standardsprachen Insgesamt 15 Aspekte von Standardsprachen, ihrer dialektalen Grundlage, ihres Ausbauzustandes etc. formuliert BROZOVIĀ (1967) in Form von binären Merkmalen, erstellt eine Merkmalsmatrix für alle slawischen Sprachen und berechnet daraus die paarweise Ähnlichkeit der Sprachen sowie eine Reihenfolge nach einem Punktesystem.

Wortschatz ŹURAVLEV (1992) gibt einige Zahlenwerte zum Erbwortschatz der slawischen Sprachen an, verglichen mit einem rekonstruierten urslawischen Bestand. In einer weiteren Monographie (ŹURAVLEV 1994, vgl. auch 1988) modelliert der Autor die genetischen Verwandtschaftsverhältnisse der slawischen Sprachen auf lexikostatistischer Grundlage.

Syntax Die Ähnlichkeit der slawischen Sprachen untereinander und zum Kirchenslawischen hinsichtlich des Gebrauches von Syntagmen wie *дать совет* untersucht KOPYLENKO (1969), vgl. auch BORĀENKO (1977).

4. Diachrone Untersuchungen zum Russischen

Sprachgeschichtliche Untersuchungen fassen wir aus praktischen Gründen in einem eigenen Kapitel zusammen, wobei hierunter diachrone Untersuchungen im eigentlichen Sinne (d.h. im Sinne eines Längsschnittes) und synchrone Untersuchungen historischer Sprachzustände zu verstehen sind. Da solche Untersuchungen noch viel seltener als synchrone Studien sind, hätte es Kapitel 3 der vorliegenden Übersicht nur unnötig aufgebläht, wäre dort jedesmal die Unterscheidung nach Synchronie und Diachronie mit hineingenommen worden.

Den Verlauf von erfolgreichem wie abgebrochenem Sprachwandel beschreibt in theoretischer Form das sog. "Piotrowski-Gesetz" (vgl. ALTMANN et al. 1983, ALTMANN 1983). Es macht deutlich, daß die Durchsetzung einer sprachlichen Veränderung ähnlich wie eine Epidemie verläuft, graphisch darstellbar in Form einer "S-Kurve". Nach einem langsamen "Start" übernehmen immer mehr Sprecher eine Neuerung, deren Zuwachs sich deshalb beschleunigt, bis ein Wendepunkt erreicht wird, ab dem die Zahl der noch nicht "infizierten" Sprecher immer kleiner wird, also das Ausmaß des Zuwachses immer weiter abnimmt. Erst wenn schließlich alle Sprecher erfaßt sind, hat sich eine Veränderung vollkommen durchgesetzt. Leider ist dieses Gesetz bislang noch nicht an Daten zum Russischen überprüft worden. – Eine ganze Reihe von historischen Implikationen hat auch das sog. "Menzerathsche Gesetz", das in allgemeiner Form lautet: "Je länger ein sprachliches Konstrukt, desto kürzer seine Konstituenten". Welche Hypothesen sich hieraus – etwa über Lautwandel in langen bzw. kurzen Wörtern, über Derivation und Wurzellänge etc. ableiten lassen, dazu s. ALTMANN/SCHWIBBE (1989, 9ff.). Auch dieser Bereich ist für das Russische noch weithin unbearbeitet. – Zu mathematischen Modellen in der historischen Linguistik vgl. auch PIOTROVSKAJA/PIOTROVSKIJ (1974), die einzelne russische Beispiele zur Demonstration heranziehen.

4.1. Grapheme

Die erste Graphemstatistik in Rußland überhaupt, und zugleich der Beginn der QL in Rußland, ist eine Frequenzanalyse der kirchenslawischen Lettern bei BUDILOVIČ (1883, 67f., 97f. und 116f., in Anlehnung an SCHLEICHER). Die Ergebnisse seien hier wenigstens kurz aufgelistet:

Т	10,7%	Ш	5,1%	И	20,4%	Ь	6,0%
В	9,1%	Г	4,3%	Ъ	13,9%	Ѧ	4,3%
Н	8,7%	К	3,9%	Є	13,4%	Ж	4,3%
С	8,7%	Ж	3,9%	О	13,4%	Ы	3,0%
Й	6,9%	Б	3,5%	Ѧ	12,7%	ОУ	2,5%
Р	6,8%	П	3,2%	Ѣ	6,1%		
Д	6,4%	Ц	2,2%				
М	6,3%	З	1,8%				
Л	6,2%	Х	1,5%				
		У	0,8%				

Tab. 22: Häufigkeit kirchenslawischer Buchstaben

BUDILOVIČ ermittelt und kommentiert ferner den Anteil verschiedener artikulatorischer Klassen, das Verhältnis von Vokalen zu Konsonanten usw., oft im Vergleich zum Lateinischen und Griechischen.

4.2. Phoneme

Halbvokale Ein früher und eigenständiger Versuch, den Ausfall der sog. Halbvokale (Jers) quantitativ zu beschreiben (und womöglich zu erklären), stammt von ABERNATHY (1956). Der Autor untersucht zwar kirchenslawische Texte, deutet am Schluß aber an, daß die Ergebnisse mit leichten Einschränkungen wohl auch für das Russische gelten. An Auszügen aus dem *Codex Zographensis* (nach LESKIEN) zählt der Autor, wie oft die 7 verschiedenen Weiterentwicklungsmöglichkeiten der Halbvokale (von denen der Ausfall und die Aufwertung zu *o/e* zwei sind) im Vergleich zu einem hypothetischen "Urzustand" in starker bzw. schwacher Position auftreten. Für beide Fälle ergibt sich, daß sich die häufigsten Varianten durchgesetzt haben. In einer informationstheoretischen Betrachtung sind dies zugleich die Weiterentwicklungen, die die Entropie des Systems (also die durchschnittliche Unsicherheit) am stärksten verringern.

Mit den Ergebnissen von ABERNATHY beschäftigt sich auch KOLESOV (1964a), der seinerseits zusätzliche statistische Daten zu altrussischen Texten des 12.Jhs. anführt, um gewisse noch nicht befriedigend gelöste Probleme zu klären, v.a. die Frage, ob die Änderung in der Häufigkeit der einzelnen Vokale als Ursache oder nur Bedingung des Wandels zu werten ist. Sein Fazit lautet, daß hierin die Ursache der Veränderungen zu suchen sei.

MALKOVA gibt einige Zahlenwerte zur Vertretung der Halbvokale in einzelnen Wörtern in Verbindung mit der Entwicklung des Polnoglasie (1981) bzw. der Palatalisierung (1980).

Akan'e KOLESOV (1964b) untersucht auch Erscheinungen im Umfeld der Entstehung des russischen Akan'e mit statistischen Methoden, wiederum nach dem

Muster der Untersuchung ABERNATHYS, indem er insbesondere das Verhältnis der einzelnen Vokale (z.B. von *ě*) in betonter und unbetonter Stellung sowie vor/nach harten bzw. weichen Konsonanten kontrastiert, um zu zeigen, in welchen Positionen die bekannten Veränderungen im Phonembestand bzw. in der Phonemverteilung (*ě* > *e*, *e* > *o* usw.) ihren Ausgangspunkt genommen haben.

Phonemsystem Veränderungen im Phonemsystem und die Zahl der jeweils möglichen Distinktionen sowie Länge und Häufigkeit der Lexeme sind die Komponenten des Ansatzes von LOMTEV (1965), der seine Überlegungen mit einzelnen Beispielen aus der Geschichte des Russischen illustriert (Entwicklung der Palatalitätskorrelation). Mit den Reflexen der urslawischen Verbindungen *tj und *dj in älteren und neueren Texten beschäftigt sich ZACHAROVA (1980, 1981a, 1981b).

Morphologie Hinsichtlich charakteristischer Lautgruppen hat DIETZE (1978, 1983) den Lexembestand mehrerer Chroniken untersucht und ausgewertet. Ihm geht es dabei v.a. um die Gegenüberstellung südslawischer (kirchenslawischer) und ostslawischer Elemente, deren Häufigkeit er jeweils berechnet. Im einzelnen geht es um *č* – *šč*, *ž* – *žd*, #je- – #o-, das Polnoglasie sowie weitere Phänomene. Der durchschnittliche Anteil der "phonetischen Slawismen" liegt nach dieser Untersuchung bei 52% bzw. 63%, in Bezug auf einzelne Elemente jedoch gelegentlich auch deutlich höher, etwa bei dem Präfix *pre-/prě-*, das in dieser Form einen Anteil von 80% aufweist. Nach dem sog. "Zweiten südslawischen Einfluß" liegen die genannten Werte nochmals deutlich höher.

4.3. Morphologie

Nominalflexion Den Zustand der slawischen Nominalflexion im 11.-16. Jh., also vor der Herausbildung der ostslawischen Einzelsprachen, beleuchten die Bände von GERD et al. (1974, 1977), von denen der erste zunächst aus einer einzigen statistischen Tabelle besteht, die zeigt, wie häufig welche Flexionsendungen in den einzelnen Gebieten (Bulgarien – Serbien – Rus') und Texten aufgetreten sind. So läßt sich z.B. verfolgen, wie sich das Verhältnis der konkurrierenden Endungen -y und -ovu im Dat.Sg. mask. Subst. im Laufe der Zeit gewandelt hat. Ausgewertet wurden für die Untersuchung mehr als 50 Texte, dem Zeitraum entsprechend natürlich in erster Linie geistliche Literatur und Chroniken. Die Quellen werden bibliographisch vollständig genannt. Aus jedem Text wurde eine Stichprobe von 12.000 laufenden Wörtern ausgezählt, um zuverlässige Daten zu erhalten. Bei der Präsentation des Materials wäre allerdings ein Mehr an Übersichtlichkeit von großem Gewinn für die spätere Benutzung gewesen. Es werden z.B. die Frequenzen zu sämtlichen Flexionsendungen aller Deklinationen der drei Genera in einer einzigen riesigen Tabelle (pp. 18-109!) aufgeführt, statt deutlich nach den grammatischen Klassen zu trennen.

Die weiteren Kapitel des Buches versuchen, die erhobenen Daten nach der einen oder anderen Richtung auszuwerten, wobei sich die Auswertung in statistischer Hinsicht auf das Berechnen von Prozentzahlen beschränkt. Hier wird z.B. dargestellt, welche Endungsvarianten in vielen/allen bzw. nur in wenigen Texten belegt sind. Die Zahl der selten belegten Endungen steigt dabei im Laufe der Zeit deutlich, was den Umbau und die Vereinheitlichung des ganzen Flexionssystems zeigt. Die markantesten Tendenzen innerhalb der einzelnen Endungen werden abschließend verbal zusammengefaßt (164ff.). Eine reine Kasus- bzw. Numerusstatistik (174ff.) zählt das Vorkommen der Inhalte, nicht mehr das der Ausdrucksvarianten. Als globales Ergebnis zeigt sich dabei die eindeutige Dominanz der Kasus Nom. – Akk. – Gen. über die Kasus Dat. – Instr. – Lok. und genauso des Singulars über Plural/Dual etwa im Verhältnis 8 : 2 (vgl. 198, 200), worin natürlich der Kern der weiteren Entwicklung des bulgarischen Flexionssystems erkennbar wird. Die Verteilung der Genera nach Gebieten, Textgattungen und Zeiten scheint weitgehend konstant im Verhältnis 2 : 1 : 1 für mask. – neutr. – fem. (vgl. 204f.), was an sich nicht überrascht, wohingegen die Häufigkeit, mit der Substantive gebraucht werden, klar von der Textgattung abhängt (Bibel-Texte verwenden eher weniger, Chroniken eindeutig mehr Substantive, vgl. 206f.). Das Buch bietet dem, der danach sucht, ohne Zweifel viel wertvolles (Roh-)Material, das jedoch noch einer fundierten statistischen Auswertung harret. – Zu einer weiteren Arbeiten der gleichen Autoren vgl. GERD et al. (1978), in der die erhobenen empirischen Daten zum 15.Jh. – bei den Texten handelt es sich um "gramoty" geographisch breit gestreuter Provenienz (Dvina, Pskov, Moskva, Süden) eines Gesamtumfanges von über 43.000 Wörtern – mit einer Vielzahl von Kurven graphisch umgesetzt werden.

Ob und wie morphologischer Wandel bei den Endungen mit der Frequenz des betreffenden Wortes korreliert, untersucht HENTSCHEL (1992a), und damit auch die verbreitete Ansicht, daß häufige Wörter "besonders gerne" irreguläre Endungen konservieren bzw. als erste annehmen. Beides bestätigt er für den endungslosen Gen.Pl., den Nom.Pl. auf *-a*, den 2. Präpositiv und den 2. Genitiv (auf *-y*) maskuliner Substantive aufgrund einer empirischen Untersuchung, bei der er die Angaben zur Flexion nach dem "Grammatischen Wörterbuch" von ZALIZNJAK (1977) mit den Häufigkeitsangaben nach ZASORINA (1977) zusammenstellt. Alle Aussagen beruhen hier auf statistischen Tests, die für jeden untersuchten Fall einen signifikanten Zusammenhang ergeben. Die in der Literatur ebenfalls zu findende Behauptung, morphologischer Wandel beginne bei infrequenten Wörtern (so MANĆZAK), wird damit widerlegt. Diese Widerlegung läßt sich wohl auch theoretisch herleiten, wenn man daran denkt, daß ja sprachlicher Wandel generell einem epidemiologischen Muster folgt, das in der Sprachwissenschaft den Namen "Piotrowski-Gesetz" trägt: nur solche Veränderungen habe eine Chance, sich durchzusetzen, die bei häufigen Wörtern auftreten, weil nämlich nur solche "Mutationen" einer großen Zahl von Sprechern häufig genug begegnen. Seltene Wörter hingegen sind notwendigerweise schlechte "Wirte" für die Verbreitung von Neuerungen. Dies bedeutet im übr-

gen nicht, daß eine Neuerung *entsteht*, weil das Wort häufig ist; es bedeutet nur, daß sich eine Neuerung *eher durchsetzt*, wenn sie bei häufigen Wörtern auftritt.

– Weitere Arbeiten: KAPORULINA (1983).

Wortarten Die Häufigkeit der einzelnen Wortarten in mehreren russischen Texten des 18.Jhs. untersucht RAECKE (1993), dazu weitere Text-Eigenschaften wie die Wiederholungsrate, Textdeckung u.ä. Die Daten werden z.T. mit denen des modernen Russisch verglichen. Bei den Texten handelt es sich um private Korrespondenz wie um literatursprachliche Texte (Karamzin, Puškin), die miteinander kontrastiert werden.

Adjektivkurzformen VINOKUR (1959) hat die Entwicklung im Bereich des adjektivischen Formenbestandes durch einen Vergleich der Gebrauchshäufigkeiten der Adjektivkurzformen bei BATJUŠKOV und in zwei Perioden von PUŠKINs Schaffen untersucht und konnte dabei zeigen, daß diese Formen schon in PUŠKINs Frühwerk seltener als bei dem Vergleichsautor benutzt und in seinen späteren Werken noch viel stärker zurückgedrängt werden (vgl. auch FRUMKINA 1964a, 55ff.).

Verbformen Mit Veränderungen im Gebrauch bestimmter Verbalformen hat sich in drei Aufsätzen bereits KUDRJAVSKIJ (1909, 1911, 1912) beschäftigt, der als einer der Pioniere der russischen QL gelten muß (vgl. PAPP 1966, 14f.), u.a. auch deshalb, weil er mit großer Klarheit die Fragen stellt, um deren Beantwortung es ihm geht, nämlich nach dem Verlauf von Sprachwandel. Er zählt dazu den Anteil der Verbformen insgesamt sowie speziell der Vergangenheitstempora (Aorist, Imperfekt und I-Partizip) in der Laurentiuschronik aus und stellt die Ergebnisse in Form einer Graphik dar. Zum Aoristgebrauch stellt er fest, daß sein Anteil in den narrativen Textteilen höher liegt als in den nichtnarrativen, ebenso, daß der Anteil der Verbformen insgesamt im Vergleich zu anderen Wortarten in diesem Text nach 1110 (dem zweiten Teil der Chronik also) zurückgeht. Die von ihm ermittelten Passagen mit besonders hohem Anteil der an sich noch recht seltenen I-Partizipien zeigen die Richtung an, in der sich das russische Temporalsystem entwickeln wird (das I-Partizip ersetzt die ererbten Tempora und wird zu einer finiten Form). Die zweite Arbeit ist ganz der Entwicklung dieses I-Partizips gewidmet; hier wird auf der Basis von 30.000 untersuchten Verbformen aufgezeigt, wie der Übergang von einer Partizipialform zu einer finiten Form vor sich gegangen ist: Den Anfang machten hierbei die 1.Ps.Sg. und vor allem die 3.Ps.Pl., bei denen die Kopula am ehesten wegfällt, während im 12./13. Jh. die Formen der 2.Ps. (Sg., Pl., Du.) das Hilfsverb noch regelmäßig verlangen. Zugleich wird deutlich, daß der Gebrauch des I-Partizips auf den mündlichen Sprachgebrauch des Russischen zurückgeht: schon in der Laurentiuschronik sind es vor allem Passagen mit direkter Rede, die einen erhöhten Anteil dieser Form aufweisen. Bei der Wiedergabe fremder Rede habe der Schreiber weni-

ger Zwang gespürt, den traditionellen Chronik-Stil beizubehalten, so die Interpretation.

Derivation Den Gebrauch des Derivationssuffixes *-ива/-ыва-* untersucht OTTEN (1978) anhand einiger Chroniktexte. Ausführliche Statistiken geben Aufschluß über die absolute Zahl der Belege, das Auftreten in finiten bzw. nichtfiniten Verbalformen sowie in einzelnen Paradigmen (als typischer Anwendungsfall erweisen sich u.a. die Vergangenheitsformen sowie der Infinitiv) bis hin zu den einzelnen Wortformen eines einzelnen Stammes (*-казывати*).

Das Konkurrenzverhältnis von Verben mit den Präfixen *вы-* und *из-* im Altrussischen untersucht BELOZERCEV (1964, 1966) und gibt zum Auftreten dieser Bildungen in einer ganzen Reihe von Texten absolute Zahlen an. – VJALKINA (1964) beschäftigt sich mit der Bildung komplexer Wörter des Altrussischen, die auf griechische Vorlagen zurückgehen, und gibt eine Rangliste der 46 häufigsten Bildungstypen an. ULUCHANOV (1964) beleuchtet das Konkurrenzverhältnis der Präpositionsvarianten *передъ* und *непередъ* im Altrussischen und gibt Daten zum Auftreten in einzelnen Texten an: während im hohen Stil die kirchenslawische Form eindeutig dominiert, finden sich in anderen Texten beide Formen gemischt. AVERINA/GERD (1983) ermitteln die Häufigkeit substantivischer Derivationsuffixe in einer ganzen Anzahl altrussischer Texte aus dem XI. bis XVI. Jahrhundert. Die Zahlen werden in Tabellenform übersichtlich dargestellt, die Interpretation der Ergebnisse ist allerdings knapp.

4.4. Lexematik

Frequenzwörterbücher Ein Frequenzwörterbuch zum Altrussischen haben beispielsweise VJALKINA/LUKINA (1966) zusammengestellt, wobei ihre Auswertungen vor allem auf dem *Merilo pravednoe* (Mitte 14. Jh.) und einigen kürzeren zusätzlichen Texten (*Žitija Feodosija Pečerskogo*, *Skazanija o Borise i Glebe*, *Čtenija o Borise i Glebe*) beruhen. Die Angaben zur Textdeckung besagen beispielsweise, daß man 66% des genannten Textes mit nur 300 Lexemen erfaßt, 75% mit 500 und 93% mit 2.000 Lexemen. Insgesamt stehen den 31.262 Vorkommen 4.311 Lexeme gegenüber. Die Liste der häufigsten Lexeme wird neben der üblichen Form zusätzlich auch nach verschiedenen Wortarten getrennt aufgeführt. Dem ersten Aufsatz (VJALKINA/LUKINA 1964) liegen die gleichen Daten zugrunde; Thema ist hier der Vergleich von Funktionen wie der Rangfrequenz-Verteilung bzw. der Textdeckung im Altrussischen mit den bekannten Daten zum heutigen Russisch (der keine Auffälligkeiten ergibt, die sich nicht unmittelbar durch das Material erklären ließen).

Mehrere Häufigkeitwörterbücher zu altrussischen Texten hat DIETZE (19877, 1984a, 1984b) vorgelegt. Dieser Autor untersucht Novgoroder Chroniken (z.T. ausgewählte Bestandteile), deren Entstehungszeit zwischen der zweiten Hälfte des 13. Jahrhunderts und dem Anfang des 16. Jahrhunderts liegt. Die Lemmata werden in

Form mehrerer Listen präsentiert, vor allem alphabetisch und rückläufig, wobei in der Regel absolute und relative Häufigkeit angegeben werden, für die niederfrequenten Wörter z.T. auch die Fundstellen. Eine Rang-Frequenz-Tabelle, eigentlich Grundbestandteil jedes Häufigkeitswörterbuches, gibt es hier nur für Wortformen, nicht für Lexeme, was eine bedauerliche Entscheidung ist, der Grad der Textdeckung wird in einer unüblichen Weise dargestellt, wegen einzelner Fehler müßten etliche Zahlenangaben und Werte korrigiert werden. Für die quantitative Linguistik stellen diese Wörterbücher eine noch weitgehend unausgeschöpfte, lohnende Materialquelle dar. – Die häufigsten Wortformen einer dieser Untersuchungen (1984a) finden sich in der vergleichenden Tabelle (s.u., *1-ja Nov. L.*).

Zur Anwendung von Häufigkeitswörterbüchern in der historischen Lexikologie des Russischen und den damit verbundenen Schwierigkeiten nimmt auch TVOROGOV (1967) Stellung. So weist er z.B. auf den geringen Umfang vieler Texte hin, der Stichproben im erwünschten Umfang oft nicht zulasse. Eigene Daten legt er zur *Povest' vremennykh let* vor, und zwar getrennt nach Gesamttext und Gesamttext ohne Eigennamen. Er berechnet das Type-Token-Verhältnis (4684 Lexeme, 47.371 Wortformen), den Anteil der Wortarten am Text, geht auf die Zahl der selten gebrauchten Wörter ein, nennt die am häufigsten gebrauchten Verben, Lexeme und Eigennamen in getrennten Listen und legt darüber hinaus in einer Monographie (1984) ein vollständiges Wortverzeichnis und eine Rang-Frequenzliste zu diesem Text vor.

Die 25 häufigsten Lexeme (bzw. Wortformen) des Altrussischen – in einzelnen Texten – stellt **Tab. 23** zusammen, wobei die z.T. vereinfachten Schreibungen aus der Literatur übernommen werden.

Verglichen mit entsprechenden Angaben zum modernen Russischen (s. **Tab. 19**, S. 57), wird das eingeschränktere Vokabular dieser ja nicht für das Altrussische insgesamt repräsentativen Texte deutlich, in denen einige spezifische Lexeme natürlich überrepräsentiert sind.

Zu weiteren Untersuchungen vgl. u.a. GRUZBERG (1974), PALAGINA (1977, 1979), GRIGOR'EVA (1983).

Rang	Мер. прав.	Ж. Ф. П.	Ск. Б. и Г.	Чт. Б. и Г.	П. вр. лѣтъ	1-я Нов. Л.
1	и	и	и	и	и (союз)	и
2	не	же	быти	же	же	же
3	быти	быти	же	быти	въ	на
4	же	тъ	въ	святый	и (мест.)	в
5	въ	въ	святый	его	быти	а
6	отъ	съ	съ	тъ	на	въ
7	на	его	вьсь	на	сей	не
8	его	яко	его	не	не	с
9	тъ	на	не	ихъ	речи	по
10	да	отъ	на	съ	къ	от
11	о	не	яко	яко	свой	его
12	аще	вьсь	мои	свои	отъ	князь
13	ли	свои	отъ	вьсь	съ	того
14	свои	иже	къ	отъ	а	бысть
15	ихъ	блажєныи	по	бо	по	лѣта
16	вьсь	такo	язъ	язъ	яко	съ
17	яко	къ	о	иже	лѣто	то
18	съ	по	свои	къ	бо	новгор.
19	по	бо	тъи	възяти	тотъ	святаго
20	или	отъць	рѣчи	блажєныи	иже	къ
21	иже	о	еѣ	съ	прити	у
22	ни	богъ	ты	нъ	богъ	бо
23	съ	ихъ	господь	по	азъ	яко
24	рѣчи	съ	а	църкы	да	к
25	къ	глаголати	видѣти	еѣ	весь	их

Tab. 23: Die 25 häufigsten Wörter (Wortformen) des Altrussischen

Glottochronologie Eine besondere Fragestellung der QL ist die nach der "Lebensdauer" von Wörtern, deren Altersbestimmung und dergl. Ein möglicher Ansatz fragt beispielsweise nach der Wahrscheinlichkeit, daß ein bestimmtes Wort mit gegebenem Rang in der Rang-Frequenzliste von diesem Rang verdrängt und durch ein anderes ersetzt wird. Dies nennt man, auf das Lexikon insgesamt bezogen, die sogenannte "Ersetzungsrate". Konkret geht etwa АРАПОВ (1983) so vor, daß aus den Häufigkeitswörterbüchern zum modernen Russisch eine zufällige Stichprobe von 500 Wörtern gezogen wurde, woraufhin das (Mindest-)Alter jedes Wortes durch Überprüfung historischer Lexika ermittelt wurde. Legt man bei 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1725, 1750, 1775, 1800, 1825, 1850 jeweils einen Schnitt, kann man zu jedem dieser Zeitpunkte genau bestimmen, wieviele Wörter der modernen Stichprobe zum gegebenen Zeitpunkt schon belegt sind. Stellt man die Ersetzungsrate graphisch in Form einer Linie dar (vgl. 1983, 59), so hat sie einen Spitzenwert zur Zeit Peters des Großen: viele neue Wörter drängen in dieser Zeit in die russische Sprache ein. Ein Minimum liegt im 14.Jh., also nach dem Mongolen-

sturm. Der Autor warnt jedoch vor einer vorschnellen kulturellen Interpretation dieser Beobachtung, da das Fehlen einer genügend großen Zahl von schriftlichen Quellen aus dieser Zeit die Ergebnisse beeinflussen können.

Periodisierung SANNIKOV hat 773 altostslawische (bzw. altrussische, altukrainische und altweißrussische) Texte mit einer Gesamtlänge von fast 270.000 Wörtern untersucht und zu einem Frequenzwörterbuch (SANNIKOV 1974) verarbeitet und in Aufsätzen ausgewertet (SANNIKOV 1975, 1978, 1983; der zweite Aufsatz ist eine methodisch verbesserte Variante des ersten und weist einen erfreulich hohen Grad an Klarheit und Selbstreflexion auf).

Der erste wichtige Punkt seiner Untersuchung betrifft die lexikalischen Nähe der drei ostslawischen Sprachen und verschiedener Etappen ein und derselben Sprache (z.B. russischer Urkunden des 15. und des 16. Jhs.) zueinander. SANNIKOV untersucht die lexikalischen Übereinstimmungen dabei nicht nur auf der Ebene der *types* (wieviele gemeinsame Wörter kommen in den Vergleichstexten vor?), sondern berücksichtigt auch deren Frequenz, was einen höheren Grad an Übereinstimmung ergibt, bzw. berechnet die Rangkorrelation zwischen zwei verschiedenen Wortverzeichnissen. So stimmen zwar die genannten altrussischen Urkunden des XV. und des XVI. Jhs. nur zu 53,9% in ihrem Wortschatz überein, doch zu 73,6% in ihrem Wortgebrauch. Solche Werte lassen sich z.B. als Begründung dafür heranziehen, beide Jahrhunderte der gleichen Periode der russischen Sprachgeschichte zuzuordnen (jedenfalls was die lexikalische Entwicklung betrifft). Ein Vergleich gemeinostslawischer Texte (XI.–XIV. Jh.) mit altrussischen Texten (XI.–XVI. Jh.) ergibt eine hohe Rangkorrelation von $R = 0,686$, d.h. eine hohe Übereinstimmung in Bezug auf die Frage, welche Wörter im gegebenen Zeitraum jeweils vergleichbar häufig waren. Hingegen ist die Übereinstimmung zwischen den altostslawischen Texten und den altukrainischen bzw. altweißrussischen schon deutlich geringer. Ein Vergleich von Texten des XIV. Jh. mit denen der vorhergehenden (XII.–XIII. Jh.) und der nachfolgenden Zeit (XV.–XVI. Jh.) ergibt eine praktisch gleich hohe Übereinstimmung, so daß der übliche "Schnitt" in der Periodisierung des Russischen, der nach dem XIV. Jh. angesetzt wird, hier jedenfalls keine eindeutige Grundlage hat.

Der zweite Bereich, den SANNIKOV untersucht, betrifft die Frage, inwieweit die altrussische Lexik in späteren Jahrhunderten bzw. im modernen Russisch unverändert fortlebt. Als Kriterium für das Fortleben in der Gegenwartssprache wurde dabei das 17-bändige Akademiewörterbuch herangezogen. Von den 1.700 Wörtern des altrussischen Frequenzwörterbuches sind nach dieser Untersuchung 1.189 oder 69,9% erhalten, wobei von der Veränderung in der Lexik besonders Adverbien, Partikel und Konjunktionen betroffen sind, die jeweils die Hälfte ihres Bestandes verlieren (bzw. austauschen). Entscheidend hängt bei allen Wörtern die "Überlebensfähigkeit" von ihrer Frequenz ab: häufige Wörter bleiben besser erhalten als seltene. Dies zeigt SANNIKOV in Übernahme und Weiterentwicklung eines Modells aus АРАПОВ/CHERC (1974), das seinerseits auf G.K. ZIPF zurückgeht. – Weitere

Arbeiten: GOLOVIN (1965) untersucht die Entwicklung des russischen Wortschatzes im 19. und 20. Jh. und arbeitet dabei einige quantitative Stil-Differenzen heraus.

Lehnbeziehungen Zum Stellenwert der Lehnbeziehungen zwischen dem Russischen und den diversen gebenden Sprachen findet sich bei HAARMANN (1984, 489) eine zusammenfassende Tabelle, deren wichtigste Angaben hier wiederholt werden sollen, ohne auf die Quellen im einzelnen zu verweisen. Die Zahlenangaben beziehen sich auf den Bestand in der *heutigen* Literatursprache, nicht auf den Gesamtumfang der Lehnbeziehungen, und auch nicht auf dialektale Besonderheiten:

Kirchenslawisch	1/3 des Lexikons	Finnougrisch	ca. 80
Englisch	ca. 2700	Germanisch	ca. 55
Französisch	ca. 1000	Dt. vorpetrinisch	ca. 47
Griechisch	ca. 500	Skandinavisch	ca. 35
Dt. nachpetrin.	ca. 450	Italienisch	ca. 22
Holländisch	ca. 260	Ukrainisch	ca. 19
Turksprachen	ca. 230	Baltisch	ca. 12
Polnisch	ca. 180	Iranisch	ca. 5

Tab. 24: Lehnwortbestand im Russischen

Die Entwicklung der Genuszuordnung bei Lehnwörtern untersucht GIMPELEVIČ (1977) auf der Grundlage mehrerer Wörterbücher.

4.5. Syntax

Syntaktische Konstruktionen in verschiedenen Versionen der "Zadonščina" untersucht MATEJKA (1963) mit dem Ziel, Anhaltspunkte für eine zeitliche Abfolge der Abschriften zu gewinnen. LESSKIS (1963) legt u.a. Daten zur Satzlänge in einigen altrussischen Texten vor, die sich mit neuen Texten in dieser Hinsicht vergleichen ließen, um Aufschluß über eventuelle Veränderungstendenzen zu gewinnen (genauer zu den Arbeiten von LESSKIS siehe im Abschnitt zur Synchronie).

GÄLÄBOV (1978) untersucht die Stellung der Enklitika im Igorlied, und zwar sowohl relativ ("vor/nach dem Verbum") wie absolut ("am Satzanfang/nicht am Satzanfang"). – BOECK (1958) gibt genaue Daten zur Verwendung der finiten Tempora (Präsens vs. Präteritum) in Objekt- und Subjektsätzen in Abhängigkeit von Randbedingungen wie indirekter Rede, Anschluß mit Konjunktion *umo* oder *как* in Texten vom 17. Jh. bis zur Gegenwart. In einer weiteren Arbeit (1964) untersucht der gleiche Autor den Gebrauch postponierter Partizipien seit dem 18. Jahrhundert, wobei sich zeigt, daß die Vorliebe für derartige Konstruktionen offensichtlich in hohem Maße an den Individualstil eines Autors gebunden ist. Bei der Auswahl zwischen Partizip Präsens und Partizip Präteritum läßt sich jedoch eine allgemeine

Entwicklung feststellen, die in der Zurückdrängung des Präsenspartizips besteht. – MICHAJLOVSKAJA (1964) untersucht das Konkurrenzverhältnis zwischen Adjektiven mit der Bedeutung "stark" in Kombination mit verschiedenen Substantivklassen und legt hierzu auch einige wenige Zahlen vor.

BACHAREV (1981) legt eine Tabelle zur Häufigkeit von Verneinungsformen im Altrussischen vor, aus der sich für einzelne Lexeme deutliche Unterschiede im Auftreten ablesen lassen. KOTOVA (1981) beschäftigt sich unter glottometrischen Gesichtspunkten kontrastiv mit russischer und bulgarischer Syntax (unter "Glottometrie" versteht man Verfahren und Versuche, Erscheinungen sprachlicher Chronologie statistisch zu modellieren).

4.6. Semantik, Onomastik

Anthroponomastik Mit der Entwicklung der Struktur der russischen Personennamen im 18. Jh. beschäftigt sich ZININ (1969). Er stellt die Ergebnisse einer Auswertung von 500 juristischen Texten in Form zweier Tabellen dar, die Angaben zur Häufigkeit der einzelnen Benennungsmuster enthalten, und zwar für Männer wie Frauen getrennt. Den Daten läßt sich u.a. entnehmen, daß erst am Ende des betrachteten Zeitraums allmählich die heute üblichen dreigliedrigen Namensformen häufiger werden, wobei die genannte Differenzierung nach dem Sexus der Bezeichneten auffällige Unterschiede ergibt – bei der Bezeichnung von Frauen steht am Anfang ein viel umständlicheres System, nämlich mit Verweis auf Vater *und* Mann, dessen Überwindung entsprechend längere Zeit benötigt. Erst mit Beginn des XX. Jhs. haben sich sowohl für Männer wie für Frauen die dreigliedrigen Personennamen endgültig allgemein durchgesetzt. (Die Tatsache, daß die Entwicklung der Familiennamen auch vom Stand der Person abhängt – je höher, desto früher –, spricht der Autor zwar an, sie kommt in den Tabellen aber nicht zum Ausdruck.)

Toponomastik Einige Zahlen zur Struktur mittelalterlicher nordrussischer Ortsnamen im Gebiet am Onega-See legt VITOV (1967) vor, wobei er einen "lokativen Typ", einen "possessiven Typ" und einen "qualitativen Typ" unterscheidet.

4.7. Dialektologie

Novgorod Die genetische Verwandtschaft des Alt-Novogoroder Dialektes mit anderen slawischen Literatursprachen bzw. Dialekten untersucht ŽURAVLEV (1993), indem er die Angaben des "Etymologischen Wörterbuches der Slawischen Sprachen" mit einer Wortliste zum Alt-Novogoroder Dialekt vergleicht und auf dieser Grundlage Ähnlichkeitskoeffizienten für alle Sprachpaare berechnet.

Tomsk Ein historisches Häufigkeitswörterbuch zum russischen Dialekt von Tomsk findet sich bei PALAGINA (1979, vgl. auch 1977).

5. Bibliographien und Lehrbücher

Es gibt bisher keine Bibliographien speziell zu den quantitativen Untersuchungen zum Russischen bzw. zu den slawischen Sprachen insgesamt. Allerdings gibt es einige Spezialbibliographien zur Quantitativen Linguistik, die in diesem Zusammenhang von Nutzen sein können. Es sei hier verwiesen auf ARZIKULOV/SADČIKOVA (1985), BAILEY/DOLEŽEL (1968), BEKTAEV (1972), BILLMEIER/KRALLMANN (1969), ČIŽAKOVSKIJ/BEKTAEV (1986), ERMOLENKO (1967, 1970), IVANOV (1962), Kvantitativní lingvistika (1964–). Auch PAPP (1966), HOFFMANN/PIOTROWSKI (1979) und TĚŠITELOVÁ (1992) enthalten eine umfangreiche Bibliographie. Besondere Beachtung verdient auch auf die "Current Bibliography" jeweils am Ende der "Glottometrika"-Bände der Reihe "Quantitative Linguistics".

Von den russischen oder das Russische berücksichtigenden Lehrbüchern oder Übersichten zur linguistischen Statistik bzw. quantitativen Linguistik seien die folgenden Titel genannt: ALEKSEEV (1975), ARAPOV (1988), ARAPOV/CHERC (1974), BEKTAEV/PIOTROVSKIJ (1973-74), ERMOLENKO/MOGILEVSKIJ (1979), GOLOVIN (1971), HOFFMANN/PIOTROWSKI (1979), KAUFMAN (1970), NELJUBIN (1983), NOSENKO (1981), PIOTROVSKIJ et al. (1977), TĚŠITELOVÁ (1992).

6. Bibliographie

Abernathy, R.:

- 1951 The structure of Russian roots. Harvard (unpublished).
- 1956 The fall of the jers. A statistical interpretation. In: M. Halle et al. (eds.), *For Roman Jakobson. Essays on the occasion of his 60th birthday*, The Hague, 13–18.
- 1963a Mathematical Linguistics. In: Th. Sebeok (ed.), *Current trends in linguistics 1*, The Hague, 113–132.
- 1963b A Vowel Fugue in Blok. *International Journal of Slavic Linguistics and Poetics* VII, 88–107.

Achmanova, O.S., Mel'čuk, I.A., Padučeva, E.V., Frumkina, R.M.:

- 1961 O točnych metodach issledovanija jazyka. Moskva [Engl. u.d.T. "Exact methods in linguistics research", Los Angeles 1963].

Akimova, G.N.:

- 1973a Razmer predloženiya kak faktor stilistiki i grammatiki (Na materiale ruskogo literaturnogo jazyka XVIII v.). *Voprosy jazykoznanija* 2, 67–79.
- 1973b Širina sopodčinennogo kusta kak princip razvertyvaniya predloženiya (na materiale jazyka M.V. Lomonosova). *Československá rusistika* 5, 199–208.
- 1977 O nekotorych osobennostjach poëtičeskogo sintaksisa. *Voprosy jazykoznanija* 1, 96–108.

Alekseev, P.M.:

- 1969 Nekotorye voprosy teorii i praktiki statističeskoj leksikografii. In: *Statistika teksta*, t. I, Minsk, 12–37.
- 1971 Častotnyj anglo-ruskij slovar' - minimum po élektronike. Moskva.
- 1973 Semantičeskie častotnye slovari. In: *Statistika reči i avtomatičeskij analiz teksta 1972*, Leningrad, 20–36.
- 1975 Statističeskaja leksikografija (tipologija, sostavlenie i primenenie častotnyh slovarej). Leningrad.
- 1983 O novykh rabotach v slavjanskoj statističeskoj leksikografii. *Voprosy jazykoznanija* 4, 136–140.
- 1988 Kvantitativnaja tipologija teksta. Leningrad: LGPI.
- 1984 Statistische Lexikographie. Zur Typologie, Erstellung und Anwendung von Frequenzwörterbüchern. Ein Lehrbuch. Übersetzt von W. Lehfeldt (*Quantitative Linguistics*, vol. 22). Bochum.

Alekseev, P.M., Grigor'eva, A.S.:

- 1981 Učebnye materialy po ruskoi nekodificirovannoi reči. Leksika bytovykh pisem. Častotnyj minimum. Leningrad.

Alekseev, P.M., Grigor'eva, A.S., Kaširina, M.E.:

- 1973 Statističeskie issledovanija leksiki pisem. In: *Teorija jazyka i inženernaja lingvistika*, Leningrad, 74–81.

Alekseev, P.M., Kalinin, V.M., Černjad'evja, E.A.:

- 1963 O statističeskich zakonernostjach v sovremennyh ruskich i anglijskich tekstach po élektronike. *Voprosy radioélektroniki*, serija VIII, vyp. 3.

- Alekseev, P.M., Kaširina, M.E., Tarasova, E.M. (red.):
 1979 Častotnyj anglo-ruskij fizičeskij slovar'-minimum. Moskva: Voenizdat.
- Alekseev, P.M., Turygina, L.A.:
 1974 Častotnyj anglo-ruskij slovar'-minimum gazetnoj leksiki. Moskva: Voenizdat.
- Alexejew, P.M., Kalinin, W.M., Piotrowski, R.G. (Hgg.):
 1973 Sprachstatistik. Übers. v. e. Kollektiv u. L. v. L. Hoffmann. München–Salzburg: W. Fink
 [= *Statistika reči*, Leningrad 1968, deutsch].
- Altmann, G.:
 1971 Die phonologische Profilähnlichkeit. Ein Beitrag zur Typologie phonologischer Systeme der slawischen Sprachen. *Phonetica* 24, 9–22.
 1972a Status und Ziele der quantitativen Linguistik. In: S. Jäger, (Hg.), *Statistik und Linguistik*, Braunschweig, 1–9.
 1972b Zur linguistischen Unbestimmtheit. *Linguistische Berichte* 22, 74–79.
 1973 Mathematische Linguistik. In: W.A. Koch (Hg.), *Perspektiven der Linguistik I*, Stuttgart, 208–232.
 1977 Sprachregeln und Erklärung. *Linguistische Berichte* 50, 31–37.
 1983 Das Piotrowski-Gesetz und seine Verallgemeinerungen. In: K.H. Best, J. Kohlhasse (Hrsg.), *Exakte Sprachwandelforschung*, Göttingen, 59–90.
 1993 Science and linguistics. In: R. Köhler, B. Rieger (eds.), *Contributions to quantitative linguistics*, Dordrecht usw., 3–10.
- Altmann, G., Burdinski, V.:
 1982 Towards a Law of Word Repetition in Text-Blocks. In: W. Lehfeldt, U. Strauss (eds.), *Glottometrika* 4, 147–167.
- Altmann, G., Lehfeldt, W.:
 1972 Typologie der phonologischen Distributionsprofile. *Beiträge zur Linguistik und Informationsverarbeitung* 22, 8–32.
 1980 Einführung in die Quantitative Phonologie (*Quantitative Linguistics*, vol. 7). Bochum: Brockmeyer.
- Altmann, G., von Buttlar, H., Rott, W., Strauß, U.:
 1983 A law of language change. In: Brainerd, B. (ed.), *Historical Linguistics*, Bochum, 104–115.
- Altmann, G., Schwibbe, M.H.:
 1989 Das Menzerathsche Gesetz in informationsverarbeitenden Systemen. Mit Beiträgen von W. Kaufmanns, R. Köhler und J. Wilde. Hildesheim usw.: Georg Olms.
- Andreev, N.D.:
 1959 Modelirovanie jazyka na baze ego statističeskoj i teoretiko-množestvennoj struktury. In: *Tezisy Soveščanija po matematičeskoj lingvistik*, Leningrad, 16–22.
 1960 Soveščanie po matematičeskoj lingvistike. *Voprosy jazykoznanija* 1, 131–137.
 1963a Algoritmy statistiko-kombinatornogo modelirovanija morfologii, sintaksisa, slovoobrazovanija i semantiki. In: *Materialy po matematičeskoj lingvistik* i mašinomu perevodu, sb. II, Leningrad, 3–44.
 1963b Ob odnom eksperimente v oblasti ruskoj orfoëpii. In: *Voprosy kul'tury reči*, IV, Moskva, 49–52.
 1965a (red.) Statistiko-kombinatornoe modelirovanie jazykov. Moskva–Leningrad: Nauka.
 1965b Metody statistiko-kombinatornogo analiza jazyka v dejstvii i v perspektive. In: ders. (red.): *Statistiko-kombinatornoe modelirovanie jazykov*, Moskva–Leningrad, 5–20.

- 1965c Morfemnoe členenie slova i statistiko-kombinatornaja granica meždu slovoizmenenijem i slovoobrazovanijem. In: ders. (red.): *Statistiko-kombinatornoe modelirovanie jazykov*, Moskva–Leningrad, 21–26.
- 1965d O vozmožnosti gruppovogo vydelenija tipov pri statistiko-kombinatornom modelirovanii morfologii. In: ders. (red.): *Statistiko-kombinatornoe modelirovanie jazykov*, Moskva–Leningrad, 65–75.
- 1967 Statistiko-kombinatornye metody v teoretičeskom i prikladnom jazykovedenii. Leningrad: Nauka.
- Andreev, N.D., Zinder, L.R.:
- 1963 O ponjatijach rečevogo akta, reči, rečevoj verojatnosti i jazyka. *Voprosy jazykoznanija* 3, 15–21.
- Andreeva, L.D.
- 1963 Statistiko-kombinatornoe vydelenie paradigmy pervogo morfologičeskogo tipa v ruskom jazyke. In: *Materialy po matematičeskoj lingvistike i mašinomu perevodu*, sb. II, Leningrad, 45–60.
- 1965a Algoritmičeskoe formirovanie razrjada suščestvitel'nogo v ruskoj morfologii. In: Andreev, N.D. (red.): *Statistiko-kombinatornoe modelirovanie jazykov*, Moskva–Leningrad, 27–38.
- 1965b Vydelenie paradigmy vtorogo morfologičeskogo tipa pri statistiko-kombinatornom modelirovanii ruskoj morfologii. In: Andreev, N.D. (red.): *Statistiko-kombinatornoe modelirovanie jazykov*, Moskva–Leningrad, 229–235.
- 1965c Statistiko-kombinatornoe vydelenie paradigmy tret'ego tipa pri modelirovanii ruskoj morfologii. In: Andreev, N.D. (red.): *Statistiko-kombinatornoe modelirovanie jazykov*, Moskva–Leningrad, 286–291.
- 1965d Vydelenie paradigmy četvertogo morfologičeskogo tipa pri statistiko-kombinatornom modelirovanii ruskoj morfologii. In: Andreev, N.D. (red.): *Statistiko-kombinatornoe modelirovanie jazykov*, Moskva–Leningrad, 311–314.
- 1965e Vydelenie paradigmy pjatogo morfologičeskogo tipa pri statistiko-kombinatornom modelirovanii ruskoj morfologii. In: Andreev, N.D. (red.): *Statistiko-kombinatornoe modelirovanie jazykov*, Moskva–Leningrad, 315–317.
- 1969 Statistiko-kombinatornye tipy slovoizmenenija i razrjady slov v ruskoj morfologii. Leningrad: Nauka.
- Andreeva, L.D., et alii:
- 1965 Polučenie pervogo morfologičeskogo tipa russkogo jazyka v pod"jazyke radioelektroniki posredstvom algoritma statistiko-kombinatornogo modelirovanija. In: Andreev, N.D. (red.): *Statistiko-kombinatornoe modelirovanie jazykov*, Moskva–Leningrad, 49–64.
- Andrjuščenko, V.M.:
- 1986a Vyčislitel'naja leksikografija, ee vozmožnosti i perspektivy. *Voprosy jazykoznanija* 3, 42–53.
- 1968b Novye raboty v oblasti statističeskoj leksikografii. *Voprosy jazykoznanija* 5, 135–148.
- 1969 Nekotorye voprosy teorij i tehniki sostavlenija častotnych slovaroj. *Inostrannye jazyki v škole* 1.
- 1976 Nekotorye problemy avtomatizacii leksikostatističeskich rabot. In: *Voprosy lingvostatističeskogo analiza ruskoj razgovornoj reči*. Sb. statej, Moskva, 66–77.
- 1978 (red.) Issledovanija v oblasti vyčislitel'noj lingvistiki i lingvostatistiki. Moskva: Izd. Mosk. uni-ta.

Appel, W.:

- 1957/58 Energiebasis – Artikulationsbasis. Über die Abhängigkeit des Lautbestandes einer Sprache vom spezifischen Energieverlauf. *Wiener Slavistisches Jahrbuch* 6, 73–104.

Apresjan, Ju.D.:

- 1964 O sil'nom i slabom upravlenii (Opyt količestvennogo analiza). *Voprosy jazykoznanija* 3, 32–49.
- 1965 Opyt opisanija značenij glagolov po ich sintaksičeskim priznakam (tipam upravlenija). *Voprosy jazykoznanija* 5, 51–66.
- 1967 Èksperimental'noe issledovanie semantiki russkogo glagola. Moskva.

Arapov, M.V.:

- 1972 Sootnošenie suščestvitel'nych so značeniem «nazvanie mestnosti» i «nazvanie žitelej danoj mestnosti» v sovremennom russkom jazyke. In: *Problemy strukturnoj lingvistiki 1971*, Moskva, 469–503.
- 1978 Izmerenie leksičeskogo bogatstva tekstov. In: *Tekst – Język – Poetyka*, Wrocław etc., 185–210.
- 1983 Word Replacement Rates for Standard Russian (A.D. 1100–1850). In: B. Brainerd (ed.), *Historical Linguistics (Quantitative Linguistics, vol. 18)*, Bochum, 50–61.
- 1988 Kvantitativnaja lingvistika. Moskva: Nauka.

Arapov, M.V., Cherc, M.M.:

- 1974 Matematičeskie metody v istoričeskoj lingvistike. Moskva.
- 1983 Mathematische Methoden in der historischen Linguistik (*Quantitative Linguistics, vol. 17*). Bochum.

Arapov, M.V., Efimova, E.N., Šrejder, Ju.A.:

- 1975 Rangovyje raspredelenija v tekste i jazyke. *Naučno-techničeskaja informacija*, ser. 2, no. 2, 3–7.

Arapov, M.V., Šrejder, Ju.A.:

- 1970 O zakone raspredelenija dlin predloženiya v svjaznom tekste. *Naučno-techničeskaja informacija*, serija 2, č. 3, 11–15.

Artemov, V.A.:

- 1933 Technografičeskij analiz summarnych bukv novogo alfavita. In: *Pis'mennost' i revoljucija*, Sb. I, Moskva–Leningrad, 58–71.

Arzikulov, Ch.A., Arzikulova, Ž.Ch.:

- 1980 O dvujazyčnych častotnych slovarjach-minimumach. *Naučnye doklady vysšej školy. Filologičeskie nauki*, 3, 78–82.

Arzikulov, Ch.A., Sadčikova, P.V.:

- 1985 Statistika reči 1973–1985. Bibliografičeskij ukazatel'. Samarkand: Gos. uni-tet.

Aschenbrenner, M.:

- 1964 Die Vermittlung eines ausreichenden Wortschatzes im russischen Unterricht. *Die Neueren Sprachen* 3, 139–147.

Averina, S.A., Gerd, A.S.:

- 1983 Statističeskij kommentarij k drevneslavjanskim pamjatnikam XI–XVI vekov. In: *Strukturnaja i prikladnaja lingvistika. Mežvuzovskij sbornik. Vyp. 2*, Moskva, 108–118.

Bacharev, A.I.:

- 1981 Otricanie i sposoby ego vyraženiya v russkom jazyke XV–XVII vv. *Voprosy jazykoznanija* 2, 99–109.

- Bachmutova, N.I.:
1970 Úpotreblenie nekotorych naibolee častotnych glagolov v razgovornoj reči. In: *Russkaja razgovornaja reč', sb. naučnych trudov*, Saratov, 89–97.
- Baevskij, V.S.:
1966 O čislovoj ocenke sily slogov v stiche al'ternirujuščego ritma. *Voprosy jazykoznanija* 2, 84–89.
- Bailey, J.:
1969 Statistics and Style: A Historical Survey. In: Doležel/Bailey (eds.), 217–232 (mit Bibliographie).
1970 Literary Usage of a Russian Folk Song Meter. *The Slavic and East European Journal* XIV, 436–452.
1973 The Verse of Andrej Voznesenskij as an Example of Present-Day Russian Versification. *The Slavic and East European Journal* XVII, 155–173.
1979 The Russian Linguistic-Statistical Method for Studying Poetic Rhythm: a Review Article. *The Slavic and East European Journal* XXIII, 251–261.
- Bailey, R.W., Doležel, L.:
1968 An annotated bibliography of statistical stylistics. Ann Arbor.
- Baklušin, A.V.:
1965 Statistiko-kombinatornoe vydelenie pervogo morfofonematičeskogo tipa v ruskom jazyke. In: Andreev, N.D. (red.): *Statistiko-kombinatornoe modelirovanie jazykov*, Moskva–Leningrad, 39–48.
- Baldwin, J.R.:
1969 Syllable Division in Russian. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 22, 211–217.
- Barinova, G.A.:
1966 O proiznošenii [ž'] i [š']. In: *Razvitie fonetiki sovremennogo russkogo jazyka*, Moskva, 25–54.
- Bektaev, K.B.:
1972 Statistika reči 1957–1972 gg. (Bibliografičeskij ukazatel'). Alma-Ata.
1979 (red.) Častotnyj slovar' po romanu M.O. Auézova "Put' Abaja". Alma-Ata.
- Bektaev, K.B., Iščanov, K.I., Sadčikova, P.V.:
1970 Sravnitel'no-statističeskoe izučenie padežnyh form suščestvitel'nyh v ruskom i kazachskom jazykach. In: *Materialy naučnoj konferencii professorsko-prepodavatel'skogo sostava Čimkentskogo pedagog. in-ta, posv. 100-letiju so dnja roždenija V.I. Lenina i 50-letiju Kazachstana, Čimkent*.
- Bektaev, K.B., Belocerkovskaja, L.I., Piotrovskij, R.G.:
1977 Norma – situacija – tekst i lingvostatističeskije priemy ix issledovanija. In: *Jazykovaja norma i statistika*, Moskva, 5–42.
- Bektaev, K.B., Piotrovskij, R.G.:
1973 Matematičeskije metody v jazykoznanii. Č. I. Teorija verojatnostej i modelirovanie normy jazyka. Učebnoe posobie. Alma-Ata.
1974 Matematičeskije metody v jazykoznanii. Č. II. Matematičeskaja statistika i modelirovanie teksta. Učebnoe posobie. Alma-Ata.
- Bel'skaja, L.:
1974 Četyrechstopnyj jamb S. Esenina. In: *Voprosy statističeskoj stilistiki*, Kiev, 228–237.

Belocerkovskaja, L.I.:

- 1969 Statistika trechslownych sočetańij s opornoj glagol'noj slovoformoj ili slovoformoj, prinaležaščej k kategorii sostojanija (na materiale russkich tekstov po radioelektronike). In: *Statistika teksta*, t. I., Minsk, 389–394.

Belonogov, G.G.:

- 1960 Nekotorye statističeskie charakteristiki russkich pečatnych tekstov. In: *Pytannja prykladnoj lingvistyky*, Černivci.
- 1962 O nekotorych statističeskich zakonernostjach v russkoj pis'mennoj reči. *Voprosy jazykoznanija* 1, 100–101.
- 1964 Raspredelenie častot pojavlenija flektivnych klassov russkich slov. *Problemy kibernetiki* 2, 189–198.

Belonogov, G.G., Frolov, G.D.:

- 1963 Ėmpiričeskie dannye o raspredelenii bukv v russkoj pis'mennoj reči. *Problemy kibernetiki* 9, Moskva, 287–305.

Belozercev, G.I.:

- 1964 Sootnošenie glagol'nych obrazovanij s pristavkami *vy-* i *iz-* vydelitel'nogo značenija v drevnerusskich pamjatnikach XI–XIV vv. In: R.I. Avanesov (red.), *Issledovanija po istoričeskoj leksikologii russkogo jazyka*, Moskva, 161–217.
- 1966 O sootnošenii elementov knižnogo i narodnogo jazyka v pamjatnikach XV–XVII vv. (na materiale glagolov s pristavkami *vy-* i *iz-* vydelitel'nogo značenija). In: *Leksikologija i slovoobrazovanie drevnerusskogo jazyka*, Moskva, 68–122.

Belousova, E.A.:

- 1964 Statističeskij analiz glagol'nych form (Na materiale russkogo jazyka). In: *Materialy konferencii "Aktual'nye voprosy sovremennogo jazykoznanija i lingvističeskoe nasledie E.D. Polivanova"*, Tezisy dokladov i soobščeniĭ mežvuzovskoj konferencii 9.-15. sentjabrja 1964 g., Samarkand, 259–260.
- 1969 Statističeskij analiz nekotorych grammatičeskich javlenij v russkom literaturnom jazyke (Na materiale upotreblenija glagol'nych form na *-sja*). In: *Voprosy teorii i istorii jazyka. Sb. statej, posvojaščennyj pamjati B.A. Larina*. Leningrad, 284–291.

Belyj, A.:

- 1910 Simvolizm. Moskva.
- 1929 Ritm kak dialektika. Moskva.

Berger, T.:

- 1981 The Formation of the Imperative in Modern Russian. *Russian Linguistics* 6, 65–80.

Bigaev, R.I. et alii:

- 1967 Častotnyj russko-uzbekskij slovar'-minimum. Taškent.

Bilokur, B.:

- 1970 Statistical Observations on Tjutčev's Lexikon. *The Slavic and East European Journal* XIV, 3, 302–316.

Billmeier, G., Krallmann, D.:

- 1969 Bibliographie zur statistischen Linguistik (*Forschungsbericht 69/3 des Instituts für Kommunikationsforschung und Phonetik der Universität Bonn*). Hamburg: Helmut Buske.

Bobertz, J.:

- 1966 Die Doppelzeitwörter der Fortbewegung. Untersuchungen zu Fragen des Terminus, des Bestandes, der Klassifizierung und der Häufigkeit. *Zeitschrift für Slawistik* 11, 535–549.

Bobrov, P.S.:

- 1964 Opyt izučeniya vol'nogo sticha Puškinskich 'Pesen zapadnych Slavjan'. *Teorija verojatno-
stej i ee primenenija* IX, 262–272.

Boeck, W.:

- 1958 Der Tempusgebrauch in den russischen Objekt- und Subjektsätzen, seine historische Ent-
wicklung und sein stilistischer Wert. *Zeitschrift für Slawistik* 3, 209–234.
1964 Zum Gebrauch der postponierten imperfektiven Partizipien des Aktivs in der russischen
Literatursprache seit der zweiten Hälfte des 18. Jahrhunderts. *Zeitschrift für Slawistik* 9,
198–212.

Bogdanov, V.V.:

- 1973 Statističeskie koncepcii jazyka i reči. In: *Statistika reči i avtomatičeskij analiz teksta 1972*, Le-
ningrad, 9–19.

Bogoljubova, N.A.:

- 1965 Vydelenie statistiko-kombinatornym metodom pervogo morfoložičeskogo tipa russkogo
jazyka na materiale tekstov chudožestvennoj literatury (L.N. Tolstoj, «Vojna i mir»). In:
Andreev, N.D. (red.): *Statistiko-kombinatornoe modelirovanie jazykov*, Moskva–Leningrad,
76–84.

Bogomazov, G.M.:

- 1971 Učet statističeskich i fonetičeskich charakteristik pri rabote nad sočetanijami soglasnych.
In: *Teoretičeskie problemy fonetiki i obučenie proiznošeniju*, Moskva, 76–90.

Boguslavskij, V.M.:

- 1977 Prilagatel'nye vysokoj položitel'noj ocenki. In: *Jazykovaja norma i statistika*, Moskva, 76–94.

Bondaletov, V.D.:

- 1972 Der russische Vornamenschatz, seine Zusammensetzung, statistische Struktur und die
Besonderheiten seiner Veränderung. *Zeitschrift für Slawistik* 17, 3–18.
1973 Zusammensetzung und statistische Struktur des russischen Vornamenschatzes (Weibli-
che Vornamen). *Zeitschrift für Slawistik* 18, 825–837.

Bondarko, L.V., Egorov, S.K., Ljublinskaja, V.V., Nurumbetova, G.A.:

- 1986 O primenenii vyčislitel'noj tehniki v eksperimental'no-fonetičeskich issledovanijach.
Voprosy jazykoznanija 2, 90–100.

Borčenko, E.D.:

- 1977 Sintaksičeskij standart v gazete. In: *Jazykovaja norma i statistika*, Moskva, 102–113.

Borunova, S.N.:

- 1966 Sočetanija [š' č'] i [š'] na granicax morfem. In: *Razvitie fonetiki sovremennogo russkogo jazyka*,
Moskva, 55–71.
1977 O razvitii norm proiznošenija (assimiljativnaja mjagkost' soglasnych v sceničeskij reči).
In: *Jazykovaja norma i statistika*, Moskva, 244–265.

Breu, W.:

- 1987 Quantitative Beziehungen zwischen Form und Inhalt in der Flexionsmorphologie. In: G.
Freidhof, P. Kosta (Hgg.), *Slavistische Linguistik 1986*, München, 95–124.

Brozović, D.:

- 1967 Slavjanskije standartnye jazyki i sravnitel'nyj metod. *Voprosy jazykoznanija* 1, 3–33.

Budilovič, A.:

- 1883 Načertanie cerkovnoslavjanskoj grammatiki, primenitel'no k obščej teorii russkago i dru-
gich rodstvennych jazykov. Warszawa.

Bulaševa, N.S.:

- 1969 Statistika trechslownych sočetanj s opornym častotnym slovom iz tekstov po ruskoj radio-
elektronike. In: *Statistika teksta*, t. I, Minsk, 376–382.

Buttke, K.:

- 1965 Zur quantitativen Charakteristik der Wortfolge im Russischen. *Wissenschaftliche Zeitschrift der Karl-Marx-Universität Leipzig* 14, 53–68.
1968 Zur quantitativen Analyse der Beziehungen zwischen aktueller Satzgliederung und Wortfolge im Russischen. *Zeitschrift für Slawistik* 13, 209–217.

Carnap, R.:

- 1974 Einführung in die Philosophie der Naturwissenschaft. Hrsg. v. Martin Gardner. Aus dem Amerik. v. W. Hoering (*Sammlung Dialog Nr. 36*). Darmstadt: Wiss. Buchges.

Častnye voprosy (ČVAAT)

- 1972 Častnye voprosy avtomatičeskogo analiza tekstov. Minsk.

Častotnye slovari (ČSiAPLT)

- 1968 Častotnye slovari i avtomatičeskaja pererabotka lingvističeskich tekstov. Mežvuzovskaja konferencija 4–6 aprolja 1968 g. Minsk.

Častotnyj slovar' Tolstogo

- 1978 Častotnyj slovar' romana L.N. Tolstogo «Vojna i mir». Tula: Gosudarstvennyj pedagogičeskij institut.

Čeremisina, N.V.:

- 1974 Primenenie statističeskogo analiza pri issledovanii ritmointonacionnoj organizacii chudožestvennoj reči. In: *Voprosy statističeskoj stilistiki*, Kiev, 54–69.

Černova, E.I., Beljakova, G.A., Malinnikova, T.G.:

- 1986 Issledovanie vosprijatija kardinal'nych glasnych nositeljami russkogo jazyka. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 39, 472–483.

Charakoz, P.I.:

- 1971 Častotnyj slovar' sovremennogo russkogo jazyka. Frunze: Mektep.

Cherry, E.C., Halle, M., Jakobson, R.:

- 1953 Towards the logical description of languages in their phonemic aspect. *Language* 29, 34–46.

Chovanov, G.M.:

- 1978 Nekotorye voprosy količestvennogo analiza povtorenija slova v tekste. In: *Issledovanija v oblasti vyčislitel'noj lingvistiki i lingvostatistiki*, Moskva, 41–58.

Chretien, C.D.:

- 1949 Stress patterns in Russian noun inflection. *Studies in Linguistics* 7, 53–57.

Čistjakov, V.F., Kramarenko, B.K.:

- 1929 Opyt priloženija statističeskogo metoda k jazykoznaniju. Vyp. I. Krasnodar.

Čižakovskij, V.A., Bektaev, K.B.:

- 1986 Statistika reči 1957–1985. Bibliografičeskij ukazatel'. Kišinev.

Corbett, G.G.:

- 1981 Agreement with quantified subjects in Russian: a fictitious linguistic change? *Russian Linguistics* 5, 287–290.

Demidova, L.E.:

- 1976 Charakteristika materiala po temam i istočnikam zapisi. Svedenija ob informantach. In: *Voprosy lingvostatističeskogo analiza russkoj razgovornoj reči*. Sb. statej, Moskva, 9–13.

Denisov, P.N.:

- 1985 Slovar' jazyka V.I. Lenina kak novyj tip slovarja. *Voprosy jazykoznanija* 1, 3–12.
1972 (red.) Leksičeskie minimumy russkogo jazyka. Moskva.

Denisov, P.N., Kostomarov, V.G.:

- 1969 Stilističeskaja differenciacija leksiki i problema razgovornoj reči (po dannym «Slovarja russkogo jazyka» S. I. Ožegova, izd. 3, M., 1953). In: *Voprosy učebnoj leksikografii*, Moskva, 105–113.

Denisov, P.N., Morkovkin, V.V., Saŭ'jan, Ju.A.:

- 1978 Kompleksnyj častotnyj slovar' russkoj naučnoj i tehničeskoj leksiki. 3047 slov. Moskva: Russkij jazyk.

Derjagin, V.Ja.:

- 1980 Ob istoriko-stilističeskom issledovanii aktovyh tekstov. *Voprosy jazykoznanija* 4, 97–107.

Dietze, J.:

- 1977 Frequenzwörterbuch zur Synodalhandschrift der Ersten Novgoroder Chronik (*Univ. Halle-Wittenberg, Wiss. Beiträge*, F 11), Halle: Martin-Luther-Universität.
1978 Phonetische Slawismen in der Synodalhandschrift der Ersten Novgoroder Chronik. Eine frequenzstatistische Untersuchung. *Zeitschrift für Slawistik* 23, 586–596.
1981 Ableitungssuffixe der russischen Fachsprache. Eine morphostatistische Untersuchung. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 34, 322–333.
1982a Das einsprachige Frequenzwörterbuch für die diachronische Sprachbetrachtung – Aufbau und rechnergestützte Herstellung. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 35, 394ff.
1982b Grapheme und Graphemkombinatorik der russischen Fachsprache. Eine phonostatistische Untersuchung. In: Lehfeldt, W., Strauss, U. (eds.), *Glottometrika* 4, Bochum, 80–94.
1983 Phonetische Slawismen in der jüngeren Redaktion der Ersten Novgoroder Chronik. Eine frequenzstatistische Untersuchung. *Zeitschrift für Slawistik* 28, 845–853.
1984a Frequenzwörterbuch zur jüngeren Redaktion der ersten Novgoroder Chronik (SSS, Band 5). München: Otto Sagner.
1984b Frequenzwörterbuch zur Vierten Novgoroder Chronik (*Univ. Halle-Wittenberg, Wiss. Beiträge*, F. 49) Halle: Martin-Luther-Universität.
1984c Frequenzstatistische Möglichkeiten zur Nutzung der linguistischen Datenverarbeitung in der Diachronie der russischen Sprache. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 37, 355–360.

Dikarev, M.A.:

- 1891 Očerok voronežskogo meščanskogo govora sravnitel'no s ukraino-russkim narečiem (Opyt statističeskogo issledovanija zvukovyh javlenij). In: ders., *Voronežskij étnografičeskij sbornik. Izdanie Voronežskogo Gubernskogo Statističeskogo Komiteta*. Voronež, 1–68.

Dobrosmyslov, V.:

- 1957 Garri G. Žossel'son. Podščët russkich slov i častotnyj analiz grammatičeskikh kategorij obrazcovogo russkogo literaturnogo jazyka. SŠA, Detroit. *Russkij jazyk v nacional'noj škole* 2, 83–87.

Doležel, L.:

- 1964 Verojatnostnyj podchod k teorii chudožestvennogo stilja. *Voprosy jazykoznanija* 2, 19–29.

- 1965 Zur statistischen Theorie der Dichtersprache. In: H. Kreuzer, R. Gunzenhäuser (Hgg.), *Mathematik und Dichtung*, München, 275–293.
- 1967 The Prague school and the Statistical Theory of Poetic Language. In: *Prague Studies in Mathematical Linguistics* 2, 97–104.
- 1972 Ein Begriffsrahmen für die statistische Stilanalyse. In: Ihwe, J. (Hg.): *Literaturwissenschaft und Linguistik*, Bd. 1. Frankfurt a. M., 67–87.
- Doležel, L., Bailey, R.W. (eds):
- 1969 Statistics and Style (*Mathematical Linguistics and Automatic Language Processing*, ed. by D.G. Hays, vol. 6). New York: American Elsevier Publ. Co.
- Dolgopol'skij, A.B.:
- 1963 Kategorija vida v rusckom jazyke i verojatnostnyj charakter svjazi označaeomogo s označajuščim. In: *Problemy strukturnoj lingvistiki 1963*, Moskva, 266–281.
- Elkina, V.N., Judina, L.S.:
- 1964a Statistika slogov rusckoj reči. In: *Vyčislitel'nye sistemy. Sbornik trudov instituta matematiki SO AN SSSR*, vyp. 10, 58–78.
- 1964b Statistika otkrytych slogov rusckoj reči. In: *Sbornik trudov instituta matematiki SO AN SSSR*, 14, 55–91.
- Ėntropija jazyka i statistika reči (ĖJaSR)
- 1966 Ėntropija jazyka i statistika reči. Otv. red. R.G. Piotrovskij. Minsk.
- Ermolaev, V.:
- 1915 Sintaksičeskaja charakteristika žanrov slovesnych proizvedenij v statističeskom osveščeenii, č. I, vyp. I. Kazan'.
- Ermolenko, G.V.:
- 1965 Obratnyj častotnyj slovar' sovremennogo rusckogo literaturnogo jazyka. In: *Metodika prepodavanija rusckogo jazyka v kazachskich gruppach vuzov*, Alma-Ata, 31–73.
- 1967 Tematičeskaja bibliographija rabot po lingvističeskoj statistike na rusckom jazyke. Pod red. Ch.Ch. Mavhmudova. Alma-Ata.
- 1970 Lingvističeskaja statistika. Kratkij očerk i bibliografičeskij ukazatel'. Pod red. Ch.Ch. Machmudova. Izd. 2-e, dop. Alma-Ata.
- 1974 O prologe k komedii A.S. Griboedova «Gore ot uma». In: *Voprosy statističeskoj stilistiki*, Kiev, 251–262.
- 1977 Rol' rečevykh paralelej pri atribucii chudožestvennykh tekstov. In: *Jazykovaja norma i statistika*, Moskva, 94–101.
- 1988 Anonimnye proizvedenija i ich avtory. Na materiale rusckich tekstov vtoroj poloviny XIX – načala XX v. Minsk.
- Ermolenko, G.V., Mogilevskij, R.I.:
- 1979 Učebnoe posobie po lingvističeskoj statistike dlja studentov-filologov. Samarkand: Gos. un-t.
- Ešan, L.I., Alekseev, P.M.:
- 1964 [Rez.] Ė.A. Štejnfel'dt, Častotnyj slovar' sovremennogo rusckogo literaturnogo jazyka. Tallin, 1963. *Voprosy jazykoznanija* 6, 130–134.
- Fisenne, B. von, Werther, H.:
- 1975 Probleme der Auswahl und Verwendung von Wortschatzminima im Fremdsprachenunterricht (unter besonderer Berücksichtigung des Russischunterrichts). In: *Notizen und Materialien zur russistischen Linguistik, Unterlagen für die Seminararbeit Nr. 3*, Frankfurt, 5–47.

Fokina, G.V.:

- 1965 Leksičeskaja osnova kursa ruskogo jazyka. In: *Iz opyta prepodavanija ruskogo jazyka neruskim*, vyp. 3, Moskva, 36–59.

Frumkina, R.M.:

- 1959a Nekotorye voprosy metodiki sostavlenija častotnych slovarej. In: *Mašinnyj perevod i prikladnaja lingvistika* 2 (9), Moskva, 23–39.
- 1959b Metodika sostavlenija statističeskich slovarej. In: *Tezisy Konferencii po matematičeskoj lingvistike*, Leningrad, 22–23.
- 1960a Statističeskaja struktura leksiki Puškina. *Voprosy jazykoznanija* 3, 78–81.
- 1960b Primenenie statističeskich metodov v jazykoznanii. *Voprosy jazykoznanija* 4, 129–133.
- 1961a Primenenie statističeskich metodov v izučenii jazyka. In: *O točnyh metodach issledovanija jazyka*, Moskva, 67–90.
- 1961b K voprosu o tak nazываемom “zakone Cipfa”. *Voprosy jazykoznanija* 2, 117–122.
- 1962a Ot statističeskogo opisanija reči k statističkim modeljam jazyka. In: *Tezisy dokladov mežvuzovskoj konferencii na temu “Jazyk i reč” (27 nojabrja–1 dekabrja)*, Moskva, 76–77.
- 1962b O zakonach raspredelenija slov i klassov slov. In: *Strukturno-tipologičeskie issledovanija*, Moskva, 124–133.
- 1962c Statističeskie metody izučenija slovarnogo sostava. Kand. diss. Moskva.
- 1963a (red.) Materialy k častotnomu slovarju jazyka Puškina (prospekt). Moskva: AN SSSR, Institut jazykoznanija, Predvaritel’nye publikacii.
- 1963b Nekotorye praktičeskie rekomendacii po sostavleniju častotnych slovarej. *Russkij jazyk v nacional’noj škole* 5, 74–76.
- 1964a Statističeskie metody izučenija leksiki. Moskva: Nauka.
- 1964b Avtomatizacija issledovatel’skich rabot v leksikologii i leksikografii [obzor]. *Voprosy jazykoznanija* 2, 114–119.
- 1966a [Rez.] N.V. Stavroski, Lingvističeska statistika. Sofija 1964. *Voprosy jazykoznanija* 1, 126–128.
- 1966b Ob’jektivnye i sub’jektivnye ocenki verojatnostej slov. *Voprosy jazykoznanija* 2, 90–96.
- 1971 Verojatnost’ elementov teksta i rečevoe povedenie. Moskva: Nauka.
- 1976 On the place of Statistical Methods in Contemporary Linguistics. In: W. Girke, H. Jachnow (eds.), *Theoretische Linguistik in Osteuropa. Originalbeiträge und Erstübersetzungen*, Tübingen, 183–196.

Frumkina, R.M., Štejnfel’dt, E.:

- 1960 Statističeskie metody otbora leksiki dlja slovarja-minimuma po ruskomu jazyku. *Russkij jazyk v nacional’noj škole* 6, 17–25.

Frumkina, R.M., Vasilevič, A.P.:

- 1968 Izučenie “proiznositel’noj trudnosti” russkich trechbukvennych sočetanij metodom škalirovanija. In: *Materialy vtorogo simpoziuma po psiholingvistike*, Moskva.

Frumkina, R.M., Zolotarev, V.M.:

- 1959 K verojatnostnoj modeli predloženija. In: *Tezisy Konferencii po matematičeskoj lingvistike*, Leningrad, 29.

Gäläbov, I.:

- 1978 Eine bedeutsame Besonderheit der Sprache des Igorliedes. *Zeitschrift für Slavische Philologie* 40, 241–257.

Ganiev, Z.V.:

- 1966 O proiznošenii sočetanij [stk], [zdk], [ntk], [ndk]. In: *Razvitie fonetiki sovremennogo ruskogo jazyka*. Moskva, 85–95.

Gasparov, M.L.:

- 1963 Statističeskoe issledovanie russkogo trechudarnogo dol'nika. *Teorija verojatnostej i ee primenenija* VIII, 1, 102–108 [engl. u.d.T. "Statistical Investigation of Russian 'Dolnik' Trimeter" in: *Probability Theory and its Applications* VIII, 1, 1963, 96–102].
- 1965 Vol'nyj chorej i vol'nyj jamb Majakovskogo. *Voprosy jazykoznanija* 3, 76–88.
- 1967a Akcentnyj stich rannego Majakovskogo. In: *Trudy po znakovym sistemam III. Uč. zap. Tartuskogo gos. uni-ta, vyp. 198*, Tartu, 324–360.
- 1967b Jamb i chorej sovetskich poetov i problema evoljucii russkogo sticha. *Voprosy jazykoznanija* 3, 59–67.
- 1968 Taktovik v russkom stichosloženii XX v. *Voprosy jazykoznanija* 5, 79–90.
- 1972 Metričeskij repertuar russkoj liriki XVIII–XX vv. *Voprosy jazykoznanija* 1, 54–67.
- 1975 Russkij narodnyj stich v literaturnych imitacijach. *International Journal of Slavic Linguistics and Poetics* 19, 77–107.

Gavrilova, A.I.:

- 1979 Rezul'taty statističeskoj obrabotki leksiki pervogo urovnja raspreditel'nogo slovarja /RS/ russkogo jazyka i sopostavlenie dannyh RS russkogo i francuzskogo jazykov. In: *Vestnik Char'kovskogo universiteta* 185 (Inostrannyj jazyk 12), 7–11.

Genkel', M.A.:

- 1974 Častotnyj slovar' romana D.N. Mamina-Sibirjaka "Privalovskie milliony". Perm'.

Georgiev, Ch.C. et al.:

- 1972 Izmerenie smyslovoj informacii. In: *ČVAAT*, 6–16.

Gerd, A.S.:

- 1985 Drevneslavjanskij jazyk i ego tipy po lingvostatističeskim dannyh. In: *Kvantitativnaja lingvistika i avtomatičeskij analiz tekstov. Trudy po lingvostatistike*, Tartu, 9–22.
- 1987 Étalonnye tipy morfoložičeskich paradigm drevneslavjanskich tekstov. In: Tuldava, Ju.A. (red.), *Kvantitativnaja lingvistika i avtomatičeskij analiz tekstov*, Tartu, 55–72.

Gerd, A.S., Kaporulina, L.V., Kolesov, V.V., Ruskova, M.P., Čerepanova, O.A.:

- 1974 Imennoe sklonenie v slavjanskich jazykach XI–XIV vv. Lingvostatističeskij analiz po materialam pamjatnikov drevneslavjanskoj pis'mennosti. Leningrad: Izd. Leningr. uni-ta.
- 1977 Imennoe sklonenie v slavjanskich jazykach XV–XVI vv. Leningrad.

Gerd, A.S., Kaporulina, L.V., Kolesov, V.V.:

- 1978 Lingvostatističeskoe opisanie russkogo imennogo sklonenija (Imennoe sklonenie po pamjatnikam delovoj pis'mennosti XV v.). In: *Strukturnaja i prikladnaja lingvistika, Mežzovskij sbornik, vyp. 1*, Leningrad, 110–138.

Gil, St.:

- 1973 Cvetovye épitety v stichotvorenijach M.Ju. Lermontova. *Scando-Slavica* XIX, 63–73.

Gimpelevič, V.S.:

- 1977 Razvitie modelej oformlenija roda inojazyčnych suščestvitel'nyh v russkom jazyke (po dannyh statističeskogo analiza). In: *Jazykovaja norma i statistika*, Moskva, 189–207.

Girke, W., Jachnow, H.:

- 1974 Sowjetische Soziolinguistik. Probleme und Genese (*Scriptor Taschenbücher Linguistik und Kommunikationswissenschaft*, S 17). Kronberg/Taunus.

Girke, W., Jachnow, H., Schrenk, J.:

- 1972 Soziolinguistik in der Sowjetunion – eine referierte Bibliographie. *Zeitschrift für Literaturwissenschaft und Linguistik* 2, 131–160.

Gladkij, A.V., Mel'čuk, I.A.:

- 1969 Élementy matematičeskoj lingvistiki. Moskva (Deutsche Übersetzung u. d. Titel "Elemente der mathematischen Linguistik", Berlin, Dt. Verlag der Wiss. 1973).

Glovinskaja, M.Ja.:

- 1976 Variativnost' proiznošenija zaimstvovannyh slov v sovremennom rusском jazyke. In: L.P. Krysin, D.N. Šmelev (red.), *Social'no-lingvističeskie issledovanija*, Moskva, 100–122.

Golanova, E.I.:

- 1976 Slovoobrazovatel'naja variativnost' naimenovanij prisposoblenij i mehanizmov. In: L.P. Krysin, D.N. Šmelev (red.), *Social'no-lingvističeskie issledovanija*, Moskva, 167–180.

Golovin, B.N.:

- 1961 O vozmožnostjach količestvennoj charakteristiki rečevykh stilej. In: *Tezisy dokladov Mežvuzovskoj konferencii po stilistike chudožestvennoj literatury*, Moskva.
- 1964 O verojatnostno-statističeskom izučenii stilevoj differenciacii jazyka. Kiev.
- 1965 Opyt verojatnostno-statističeskogo izučenija nekotorych javlenij istorii russkogo literaturnogo jazyka XIX–XX vv. *Voprosy jazykoznanija* 3, 137–146.
- 1966 Iz kursa lekcii po lingvističeskoj statistike. Gor'kij.
- 1967 K voprosu o verojatnostno-statističeskom ponimanii stilja jazyka i stilja reči. In: *Uč. zap. NII PMK GGU. Prikladnaja matematika i kibernetika*, Gor'kij, 513–530.
- 1968 O stiljach jazyka i ich izučenii. *Russkij jazyk v škole* 4, 11–19.
- 1969 O statističeskom modelirovanii grammatičeskoj struktury vyskazyvanija. In: *Problemy prikladnoj lingvistiki. Tezisy mežvuzovskoj konferencii*, č. 1, Moskva.
- 1971 Jazyk i statistika. Moskva: Prosveščenie.
- 1974 Statističeskie metody i sfery ich primenenija v statističeskoj stilistike. In: *Voprosy statističeskoj stilistiki*, Kiev, 5–16.

Golovin, B.N., Urambašev, I.V.:

- 1967 O statističeskich priznakach stilevoj differenciacii glagol'nych form sovremennogo russkogo literaturnogo jazyka. In: *Uč. zap. NII PMK GGU. Prikladnaja matematika i kibernetika*, Gor'kij, 501–512.

Golubzowa, L.:

- 1991 Adverb und Sprachstil. Untersuchungen zur stilistischen Differenziertheit, insbesondere im lexikalischen Bereich (*Slavistische Beiträge*, Bd. 269). München.

Googe, D.Ja.:

- 1973 Sistema glagola nižnemeckogo govora sela Kant Kirgizskoj SSR. Avtoref. Kand. diss., Minsk.
- 1974 Častotnost' upotreblenija glagol'nach form (Na materiale nižnemeckogo govora s. Kant Kirgizskoj SSR). In: *Voprosy romano-germanskogo i obščego jazykoznanija*, Minsk, 446–456.

Graudina, L.K.:

- 1964a O nulevoj forme roditel'nogo množestvennogo u suščestvitel'nych mužskogo roda. In: *Razvitie grammatiki i leksiki sovremennogo russkogo jazyka*, Moskva, 181–209.
- 1964b Razvitie nulevoj formy roditel'nogo množestvennogo u suščestvitel'nych–edinic izmernenija. In: *Razvitie grammatiki i leksiki sovremennogo russkogo jazyka*, Moskva, 210–221.
- 1965 Noskov ili nosok? In: *Voprosy kul'tury reči* VI, Moskva, 137–146.
- 1966 Opyt količestvennoj ocenki normy. In: *Voprosy kul'tury reči* VII, Moskva, 75–88 [dt. Übersetzung siehe 1975].
- 1970 Norma i statistika. In: *Aktual'nye problemy kul'tury reči*, Moskva, 326–368.

- 1975 Ein Versuch der quantitativen Normbewertung, in: W. Girke/H. Jachnow (Hgg.), *Sprache und Gesellschaft in der Sowjetunion*, München, 164–175.
- 1977 Statističeskij kriterij grammatičeskoj normy. In: *Jazykovaja norma i statistika*, Moskva, 135–173.
- Graudina, L.K., Ickovič, V.A., Katlinskaja, L.P.:
- 1976 Grammatičeskaja pravil'nost' russkoj reči. Opyt častotno-stilističeskogo slovarja variantov. Moskva.
- Grechněva, G.M.:
- 1974 Stilevaja differenciacija pridatočnych predloženíj v sovremennom russkom jazyke. In: *Voprosy statističeskoj stilistiki*, Kiev, 164–177.
- Green, B.D.:
- 1979 Factors in the Choice of the Case of Direct Objects after Negated Transitive Verbs in Russian. *The Slavonic and East European Review* 57, 161–186.
- Greenberg, J.:
- 1964 Nekotorye obobščeniya, kasajuščiesja vozmožnyh načal'nyh i konečnyh posledovatel'nostej soglasnyh. *Voprosy jazykoznanija* 4, 41–65.
- 1974 The relation of frequency to semantic feature in a case language (Russian). In: *Working papers on Language Universals* 16, 21–45.
- Grigor'ev, V.I.:
- 1960 O razvitii strukturnykh i matematičeskikh metodov issledovanija jazyka. *Voprosy jazykoznanija* 4, 153–155.
- 1964 Statističeskoe raspoznavanie i diferencial'nye priznaki fonem. *Voprosy jazykoznanija* 3, 63–69.
- 1980a O dinamike raspredelenija bukv v tekste. In: *Aktual'nye voprosy strukturnoj i prikladnoj lingvistiki*, Moskva, 40–48.
- 1980b Frequency distribution of letters and their ranks in a running text. In: *Symposium: Computational Linguistics and Related Topics*, Tallinn, 43–47.
- Grigor'eva, A.S.:
- 1980 O častotnom slovare russkoj obichodnoj pis'mennoj reči. In: *Lingvostatistika i kvantitativnye zakonomernosti teksta. Trudy po lingvostatistike* 6, Tartu, 24–31.
- 1981a Statističeskaja struktura russkogo ėpistoljarnogo teksta (leksika častnyh pisem). Diss. kand. fil. nauk, Leningrad.
- 1981b O leksiko-morfologičeskoj statistike russkoj ėpistoljarnoj reči. In: R.G. Piotrovskij (red.), *Inženernaja lingvistika i prepodavanje inostrannykh jazykov s pomošč'ju TSO*, Leningrad, 140–148.
- 1983 Častotnyj slovar' russkoj ėpistoljarnoj reči. In: *Strukturnaja i prikladnaja lingvistika. Mežvuzovskij sbornik. Vyp. 2*, Moskva, 89–94.
- Gruzberg, A.A.:
- 1974 Častotnyj slovar' russkogo jazyka vtoroj poloviny XVI – načala XVII veka. Perm'.
- Gurčenkova, V.P.:
- 1972 Častotnyj slovar' pod' jazyka publicistiki (politika). In: *ČVAAT*, 131–138.
- Haarmann, H.:
- 1984 Zu den historischen und rezenten Sprachkontakten des Russischen. In: H. Jachnow (Hg.), *Handbuch des Russisten*, Wiesbaden, 482–515.

Hentschel, G.:

- 1990 Natürlichkeit vs. Frequenz. Überlegungen zu betonungsabhängiger Allomorphie in der Deklination russischer Substantive. In: J.M. Dosuna/C. Pensado (eds.), *Naturalists at Krams. Papers from the Workshop on Natural Phonology and Natural Morphology...*, Salamanca, 65–76.
- 1992a Verwendungshäufigkeit und Innovation im Flexionssystem – Beobachtungen zum morphologischen Wandel im Russischen und Polnischen. *Zeitschrift für Slawistik* 37, 50–59.
- 1992b Zum Einfluß der Konstituentenfolge auf die Kasuswahl im Russischen. *Lingua* 87, 231–255.

Herdan, G.:

- 1956 Language as choice and chance. Groningen.
- 1957 The numerical expression of selective variation in the vowel-consonant sequence in English and Russian. In: E. Pulgram (ed.), *Studies presented to J. Whatmough on his 60th birthday*, The Hague, 91–104.
- 1960 Type-token mathematics. A textbook of mathematical linguistics. The Hague.
- 1962 The calculus of linguistic observations. The Hague.
- 1964 Quantitative linguistics. London.
- 1965 Eine Gesetzmäßigkeit der Sprachenmischung. Mit einem Exkurs über Goethes »West-Östlichen Divan«. In: H. Kreuzer, R. Gunzenhäuser (Hgg.), *Mathematik und Dichtung*, München, 85–106.
- 1966 The advanced theory of language as choice and chance. Berlin etc.
- 1969 The Mathematical Theory of Verse. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 22, 225–234.

Hermenau, O.:

- 1960 Wortschatzminimum für den Russischunterricht in der zehnklassigen allgemeinbildenden polytechnischen Oberschule. Berlin.

Hoffmann, L.:

- 1966 Zur maschinellen Analyse der statistischen Struktur wissenschaftlicher Texte (Lexik und Morphologie des Russischen). Leipzig (Unveröff. Habil.).
- 1967 Zur quantitativen Charakteristik der Sprache wissenschaftlicher Texte. *Wissenschaftliche Zeitschrift der Karl-Marx-Universität Leipzig, Gesellschafts- und sprachwissenschaftliche Reihe* 16, 1/2, 77–82.
- 1969 Die Bedeutung sprachstatistischer Untersuchungen für den Fremdsprachenunterricht. *Glottodidactica* 3/4, 47–81.
- 1970a (Hg.) Fachwortschatz Medizin. Leipzig (7/1986).
- 1970b (Hg.) Fachwortschatz Physik. Leipzig (4/1982).
- 1970c Methoden zur quantitativen Charakterisierung der Wissenschaftssprache. In: *Actes du Xe Congrès International des Linguistes III*, Bukarest, 405–410.
- 1970d Linguo-statistische Untersuchungen zur wissenschaftlichen Prosa. In: *LAB Linguistische Arbeitsberichte* 1, 69–73.
- 1970e Zur maschinellen Bearbeitung sprachlicher Daten bei linguo-statistischen Untersuchungen. In: *LAB Linguistische Arbeitsberichte* 2, 52–68.
- 1970f [Rez.] V.A. Moskovič, Statistika i semantika, Moskva 1969. *Fremdsprachenunterricht* 14, 36–38.
- 1971 Zur maschinellen Bearbeitung sprachlicher Daten bei linguo-statistischen Untersuchungen. *Deutsch als Fremdsprache* 8, 77–87.
- 1973a (Hg.) Fachwortschatz Chemie. Häufigkeitswörterbuch Russisch – Englisch – Französisch. Leipzig.

- 1973b (Hg.) Sprachstatistik. Übersetzt aus dem Russischen von einem Kollektiv unter Leitung von L. Hoffmann (*Sammlung Akademie-Verlag*, Bd. 22). Berlin (s.a. Alexejew et al. 1973).
- 1973c Zur maschinellen Bearbeitung sprachlicher Daten bei der Zusammenstellung dreisprachiger fachbezogener Häufigkeitswörterbücher. In: *Beiträge zur maschinellen Sprachdatenverarbeitung*, Berlin, 53–72.
- 1974 Ein Weg zum Grundwortschatz von Fachtexten – fachsprachliche Häufigkeitswörterbücher. *Fremdsprachen* 18, 81–86.
- 1975a (Hg.) Fachsprachen und Sprachstatistik. Beiträge zur angewandten Sprachwissenschaft (*Sammlung Akademie-Verlag*, Bd. 41). Berlin.
- 1975b Zum Forschungsgegenstand der statistischen Linguistik. In: Hoffmann 1975a, 9–24.
- 1976a (Hg.) Fachwortschatz Bauwesen. Häufigkeitswörterbuch Russisch – Englisch – Französisch. Leipzig (2/1984).
- 1976b (Hg.) Fachwortschatz Mathematik. Häufigkeitswörterbuch Russisch – Englisch – Französisch. Leipzig (2/1980).
- 1978 (Hg.) Fachwortschatz Tierproduktion/Veterinärmedizin. Häufigkeitswörterbuch Russisch – Englisch – Französisch. Leipzig.
- Hoffmann, L., Piotrowski, R.G.:
1979 Beiträge zur Sprachstatistik (*Linguistische Studien*). Leipzig.
- Hu, J.Th.:
1968 Syntactic and phonological parallelisms in the nine-syllable and thirteen-syllable Russian funeral laments. *Slavic and East European Journal* 32, 574–584.
- Ignatova, V.N.:
1969 Primenenie kriterija x2 dlja statističeskoj obrabotki rezul'tatov vosprijatija zvukov reči. In: *Statistika teksta*, t. I, Minsk, 200–205.
- Il'ina, N.E.:
1976 Glagol'nye varianty s čeredovanijem -nu-/nul' suffiksa v osnove. In: L.P. Krysin, D.N. Šmelev (red.), *Social'no-lingvističeskie issledovanija*, Moskva, 156–167.
- Ilola, E. Mustajoki, A.:
1989 Report on Russian Morphology as it Appears in Zaliznyak's Grammatical Dictionary (*Slavica Helsingiensia*, 7). Helsinki.
- Isabekova, N.I.:
1969 Imennye slovosočetačija v russkich tekstach po radioelektronike. In: *Statistika teksta*, t. I, Minsk, 383–388.
- Isačenko, A.V.:
1939/40 Versuch einer Typologie der slavischen Sprachen. *Linguistica Slovaca* I/II, 64–76.
- Isengel'dina, A.A.:
1979 Tipologija sočetanij soglasnych kazachskogo i russkogo jazykov v ich forme i funkcii. In: *Voprosy kazachskoj fonetiki i fonologii*, Alma-Ata, 99–130.
- [IVLL]
1978 Issledovanija v oblasti vyčislitel'noj lingvistiki i lingvostatistiki. Moskva.
- Ivanov, V.V.:
1958 Komitet po prikladnoj lingvistike. *Voprosy jazykoznanija* 3, 136–137.
1962 Mašinnyj perevod. 1949–1960. Bibliografičeskij ukazatel'. Moskva.
1966a Ritm poëmy Majakovskogo 'Čelovek'. In: *Poetics II*, Warszawa–The Hague, 243–276.
1966b Ritmičeskoe stroenie 'Ballady o cirke' Mežirova. In: *Poetics II*, Warszawa–The Hague, 277–299.

Ivanova-Markova, L.P.:

- 1974 Sopostavitel'nyj analiz upotreblenija sistemy odnosostavnyh i dvusostavnyh predložennij v ruskoj i ukrajinskoj poëtičeskoj reči. In: *Voprosy statističeskoj stilistiki*, Kiev, 243–251.

Jachnow, H.:

- 1971 Zur Distinktion russischer Aspektkorrelate im Infinitiv. *Die Welt der Slaven* XVI, 350–359.
- 1984 Zur theoretischen und empirischen Soziolinguistik in der UdSSR. In: ders. (Hg.), *Handbuch des Russisten*, Wiesbaden, 790–819,

Jaglom, A.M., Jaglom, I.M.:

- 1960 Verojatnost' i informacija. Izd. 2. Moskva: Fizmatgiz (dt. u.d.T. "Wahrscheinlichkeit und Information", Berlin 1965).

Jaglom, I.M., Dobrušin, R.L., Jaglom, A.M.:

- 1960 Teorija informacii i lingvistika. *Voprosy jazykoznanija* 1, 100–110.

Jakovlev, N.:

- 1928 Matematičeskaja formula postroenija alfavita (Opyt praktičeskogo priložennija lingvističeskoj teorii). In: *Kul'tura i pis'mennost' Vostoka I*, Moskva, 41–64.

Jazyk i obščestvo

- 1967 Jazyk i obščestvo. Sb. Saratov.

Jazykovaja norma i statistika

- 1977 Jazykovaja norma i statistika. Red. kolegija: R.G. Piotrovskij, L.K. Graudina, V.A. Ickovič. Moskva: Nauka.

Jiráková, I.:

- 1965 Kvantitativní analýza současných ruských hutnických textů (dodatek). Ostrava.
- 1976 Zavisimost' količestvennogo sostava grammatičeskich kategorij porjadka častej reči ot ob'ema častotnyh slovarej ruskogo jazyka. *Prague Studies in Mathematical Linguistics* 5, 37–52.

Jones, L.G.:

- 1965 Tonality structure in Russian verse. *International Journal of Slavic Linguistics and Poetics* IX, 125–151 [abgedr. in Doležel/Bailey, eds., *Statistics and Style*, 122–143].

Josselson, H.H.:

- 1947 Most important words of the Russian Language. Detroit.
- 1948 Stress patterns in Russian noun inflection. *Studies in Linguistics* 6, 53–62.
- 1953 The Russian Word Count and Frequency Analysis of Grammatical Categories of Standard Literary Russian. Detroit (Reprint New York 1967).
- 1966 Podščet slov i častotnyj analiz grammatičeskich kategorij ruskogo literaturnogo jazyka. Perevod s anglijskogo L.N. Smirnovoj. In: L.N. Zsorina (red.), *Avtomatizacija v lingvističeskoj nauke*, Moskva–Leningrad, 105–131.

Issledovanija v oblasti vyčislitel'noj lingvistiki (IVLL)

- 1978 Issledovanija v oblasti vyčislitel'noj lingvistiki i lingvostatisti. Sbornik. Moskva.

Kadomcev, O.G.:

- 1975 Ob ispol'zovanii statistiki v fonetičeskom eksperimente. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 28, 3–7.

Kagarov, J.:

- 1922 O ritme ruskoj prozaičeskoj reči. *Nauka na Ukraine* IV, 324–332.

Kalašnikova, G.F., Al'nikova, V.Ju.:

- 1991 O strukturnoj organizacii russkogo polipredikativnogo složnosoičinennogo predloženiya. *Voprosy jazykoznanija* 6, 78–89.

Kalinin, V.M.:

- 1964a O statistike literaturnogo teksta. *Voprosy jazykoznanija* 1, 123–127.
1964b Nekotorye statističeskie zakony matematičeskoj lingvistiki. *Problemy kibernetiki*, vyp. II, Moskva, 245–255.

Kalinina, E.A.:

- 1968a Izučenie leksiko-statističeskich zakonomernostej na osnove verojatnostnoj modeli. In: *Statistika reči*, Leningrad, 64–107.
1968b Častotnyj slovar' russkogo pod"jazyka élektroniki. In: *Statistika reči*, Leningrad, 144–150.

Kamińska, I.:

- 1978 Bibliography of Quantitative Linguistics in Poland. In: Altmann, G. (ed.), *Glottometrika* 1, Bochum, 199–219.

Kaporulina, L.V.:

- 1983 Statistika padežnych form imeni suščestvitel'nogo v drevnerusskom jazyke. In: *Strukturalnaja i prikladnaja lingvistika. Mežvuzovskij sbornik. Vyp. 2*, Moskva, 99–108.

Karaulov, Ju.N.:

- 1980 Častotnyj slovar' semantičeskich množitelej russkogo jazyka. Moskva: Nauka.

Karpinskaja, O.G.:

- 1964 Tipologija roda v slavjanskich jazykach. *Voprosy jazykoznanija* 6, 61–76.

Kaufman, S.I.:

- 1970 Iz kursa po statističeskoj stilistike. Moskva.

Karpenko, M.B.:

- 1961 Razmery predloženiya v romane L.N. Tolstogo 'Vojna i mir'. In: *Lev Tolstoj*, Černovicy.

Kehrer, H.:

- 1965/66 Der Wortschatz in unseren Russischlehrbüchern. *Zeitschrift für den Russischunterricht* II, 2, 32–47.

Keil, R.-D.:

- 1965 Einheitliche Methoden in der Lexikometrie. *IRAL* 3.2, 95–112.
1965/66 Zur Häufigkeit russischer Verben und Verbformen. *Zeitschrift für den Russischunterricht* 1/2, 35–50.
1994 [Rez.] Lönngren, L. (red.), Častotnyj slovar' sovremennogo russkogo jazyka. Uppsala 1993. *Zeitschrift für Slavische Philologie* LIV, 212–214.

Keller, H.H.:

- 1991 Word Frequency and Pedagogical Value: Contrasting Textbook Vocabulary Lists with Word Frequency Counts. *The Slavic and East European Journal* 35, 228–244.

Kempgen, S.:

- 1976 "Zentrum" und "Peripherie". Zur Bewertung der phonotaktischen Wortstruktur. *Linguistische Berichte* 42, 29–35.
1981 "Wortarten" als klassifikatorisches Problem der deskriptiven Grammatik. Historische und systematische Untersuchungen am Beispiel des Russischen (*Slavistische Beiträge*, Bd. 143). München.
1988 Zur lexikalischen Entwicklung der "Weltsprache Russisch". In: J. Raecke (Hg.), *Slavistische Linguistik 1987*, München, 184–208.

- 1991 Isačenkos Typologie der slawischen Sprachen aus heutiger Sicht. In: K. Hartenstein, H. Jachnow (Hgg.), *Slavistische Linguistik 1990*, München, 146–163.
- 1992 Methodische Probleme der sowjetischen Soziolinguistik. In: T. Reuther (Hg.), *Slavistische Linguistik 1991*, München, 185–207.
- 1993 Spezifika slawischer Schriften. In: ders. (Hg.), *Slavistische Linguistik 1992*, München, 111–143.
- 1994 Bemerkungen zum russischen Flexionsakzent. In: H.R. Mehlig (Hg.), *Slavistische Linguistik 1993*, München, 103–126.
- 1995a Morphologische Codierung natürlicher Sprache. *Die Welt der Slaven XL*, 1, 52–57.
- 1995b Phonemcluster und Phonemdistanzen (im Russischen). In: D. Weiss (Hg.), *Slavistische Linguistik 1994*, München (im Druck).
- i.Vorb. Phonotaktische Wortstruktur des Russischen (Monographie).
- Kijak, T.R., Kostaš, Ju.P., Skorochod'ko, Ė.F.:
- 1980 Ocenka stepeni motivirovannosti nemeckoj i russkoj terminologičeskoj i obščepotrebitel'noj leksiki. In: *Strukturnaja i matematičeskaja lingvistika* 8, 37–44.
- Kitajgorodskaja, M.V.:
- 1976 Variativnost' v vyražeenii roda suščestvitel'nogo pri oboznačenii ženščin po professii. In: L.P. Krysin, D.N. Šmelev (red.), *Social'no-lingvističeskie issledovanija*, Moskva, 144–155.
- Kjetsaa, G.:
- 1972 Michail Lermontov im Norden. *Scando-Slavica XVIII*, 13–36.
- 1973 Neskol'ko zamečanij o leksike stichotvorenij M.Ju. Lermontova. *Scando-Slavica XIX*, 49–62.
- 1974 Lomonosov's Sound Characteristics. *Scando-Slavica XX*, 77–93.
- 1976 Storms on the Quiet Don. A Pilot Study. *Scando-Slavica XXII*, 5–24 (russ. Übers. u.d.T. "Burja vokrug 'Tichogo Dona'" in: *Tichij Don: Uroki romana*, Rostov-na-Donu, 1979).
- 1978 Problema avtorstva v romane *Tichij Don*. *Scando-Slavica XXIV*, 91–105.
- 1979 A quantitative norm for the use of epithets in the age of Puškin. In: *Russian Romanticism. Studies in the Poetic Codes. Stockholm Studies in Russian Literature* 10, 204–226.
- 1981 Stil' i norma. *Linguistica* 14, 48–67.
- Kjetsaa, G., Gustavsson, S., Beckman, B., Gil, S.:
- 1984 The authorship of the *Quiet Don*. Oslo, Solum Forlag (russ. Übers. u.d.T. "Kto napisal 'Tichij Don'. Problema avtorstva 'Tichogo Dona'", M. 1989).
- Kleinlogel, A., Lehfeldt, W.:
- 1972 Zur Problematik einer syntagmatisch-phonologischen Sprachklassifikation. In: S. Jäger (Hg.), *Linguistik und Statistik*, Braunschweig, 51–64.
- Klimenko, A., Suprun, A.:
- 1963 O častotnosti nekotorych grammatičeskich javlenij v slavjanskich jazykach. In: *Sinchronoe izučenie različnych jarusov struktury jazyka*, Alma-Ata.
- Klimovič, N.I.:
- 1986 Sopostavitel'nyj častotnyj slovar' pod''jazyka taktiki (russkij i anglijskij jazyki). Moskva.
- Kločkova, Ė.A.:
- 1968 O raspredelenii klassov slov v nekotorych funkcional'nych stiljach russkogo jazyka. In: *Voprosy slavjanskogo jazykoznanija*, Saratov, 109–118.
- 1970 O vlijanii formy razgovornoj reči na raspredelenie klassov slov. In: *Russkaja razgovornaja reč', sb. naučnych trudov*, Saratov, 126–134.

Knowles, F.:

- 1980 The quantitative syntagmatic analysis of the Russian and Polish phonological systems. In: Zampolli, A., Calzolari, N. (eds.), *Computational and Mathematical Linguistics. Proceedings of the International Conference on Computational Linguistics, Pisa, 27.8. to 1.9.1973*, vol. 1, Firenze, 79–81.

Kobrin, R.Ju., Pekarskaja, L.A.:

- 1977 Lingvostatističeskij analiz upotreblenija terminov normativnyh slovarej i GOSTov v real'nyh naučno-tehničeskich tekstach. In: *Jazykovaja norma i statistika*, Moskva, 265–279.

Köhler, R., Rieger, B.B. (eds):

- 1991 Proceedings of the first international conference on quantitative linguistics, QUALICO. Trier.

Kolesov, V.V.:

- 1964a Padenie reducirovannyh v statističeskoj interpretacii. *Voprosy jazykoznanija* 2, 30–44.
1964b O nekotoryh osobennostjach fonologičeskoj modeli, razvivajuščej akan'e. *Voprosy jazykoznanija* 4, 66–79.

Kolguškin, A.N.:

- 1970 Lingvistika v voennom dele (Razrabotka i ispol'zovanie častotnyh slovarej voennoj leksiki). Moskva: Voennoe izdatel'stvo Ministerstva Oborony SSSR.

Kolmogorov, A.N.:

- 1963 K izučeniju ritmiki Majakovskogo. *Voprosy jazykoznanija* 4, 64–71.
1965 Zamečanja po povodu analiza ritma 'Stichov o sovetskom pasporte' Majakovskogo. *Voprosy jazykoznanija* 3, 70–76.

Kolmogorov, A.N., Kondratov, A.M.:

- 1962 Ritmika poëm Majakovskogo. *Voprosy jazykoznanija* 3, 62–74.

Kolmogorov, A.N., Prochorov, A.V.:

- 1963 O dol'nike sovremennoj russkoj poëzii (Obščaja charakteristika). *Voprosy jazykoznanija* 6, 84–95.
1964 O dol'nike sovremennoj russkoj poëzii (Statističeskaja charakteristika dol'nika Majakovskogo, Bagrickogo, Achmatovoj). *Voprosy jazykoznanija* 1, 75–94.

Kondratov, A. M.:

- 1962a Matematika i poëzija. Moskva.
1962b Évoljucija ritmiki V.V. Majakovskogo. *Voprosy jazykoznanija* 5, 101–108.
1963a Statistika tipov russkoj rifmy. *Voprosy jazykoznanija* 6, 96–106.
1963b Teorija informacii i poetika. (Éntropija ritma russkoj reči.) *Problemy kibernetiki* 9, 279–286.
1969 Information Theory and Poetics: The Entropy of Russian Speech Rhythm. In: Doležel/Bailey (eds.), 113–121 (Übersetzung von 1963b).

Kopejkin, S.V., Ostapenko, V.E.:

- 1977 Zakon Cipfa i sopostavitel'nyj analiz častotnyh struktur anglijskogo, francuzskogo, rumynskogo i russkogo jazykov na baze matematičeskich modelej. In: *Naučnye trudy Kujbyševskogo pedagogičeskogo instituta* 193, 91–94.

Kopylenko, M.M.:

- 1965 K éksperimental'nomu izučeniju sočetaemosti leksem. *Voprosy jazykoznanija* 2, 85–91.
1969 Opyt sopostavitel'nogo izučenija frazeologičeskich edinic tipa *dat' sovet* v slavjanskich jazykach. *Voprosy jazykoznanija* 2, 46–53.

Kormilicyna, M.A.:

- 1974 Statističeskoe izučenie funkcionirovanija glagol'nych slovosočetanij v rusškoj razgovornoj reči. In: *Voprosy statističeskoj stilistiki*, Kiev, 178–188.

Kosačeva, L.G.:

- 1976 Leksiko-statističeskoe opisanie morfoložičeskich osobennostej razgovornoj reči (glagol). In: *Voprosy lingvostatističeskogo analiza rusškoj razgovornoj reči*, Sb. statej, Moskva, 77–86.

Kostomarov, V.:

- 1963 Principy otbora leksičeskogo minimuma. *Russkij jazyk v nacional'noj škole* 1, 29–35.

Kotov, R.G.:

- 1958 Lingvističeskaja statistika tekstov rusškogo jazyka. In: *Tezisy Konferencii po mašinomu perevodu*, Moskva, 48–49.

Kotova, N.V.:

- 1981 Glottometrija i sopostavitel'nyj sintaksis rodstvennych jazykov (Na materiale rusškogo i bolgarskogo jazykov). *Vestnik Moskovskogo universiteta*, Serija 9, Filologija, 4, 33–42.

Kotova, N.V., Janakiev, M.:

- 1968 Sopostavlenie nekotorych količestvennyh charakteristik slavjanskich jazykov. In: *Slavjanskaja filologija* 7, Moskva, 58–77.
- 1973 Glottometričeskij podchod k voprosu o blizosti slavjanskich jazykov. In: *Slavjanskaja filologija* 9, Moskva, 238–260.

Koubourlis, D.J.:

- 1967 A Statistical Analysis of the Russian Verbal Aspect. The University of Washington: Ph.D. diss.

Koutsoudas, A., Halpin, A.:

- 1958 Russian physics vocabulary with frequency count. Univ. of Michigan, *Willow Run Laboratories in MT Report 2144–312–T, AD–204453*.

Kozak, A.S.:

- 1962 High Frequency Words and Occurrence Forms in Russian Physics. *The Rand Corporation Memorandum RM–3383–PR*.

Kožina, M.N.:

- 1968 K osnovanijam funkcional'noj stilistiki. Perm'.

Kozlik, L.A.

- 1962 Some Frequency Characteristics of Russian Style (M.A. Thesis). Austin Texas.

Kraus, I.:

- 1974 K kvantitativnym sociolingvističeskim modeljam v teorii jazyka i stilja. In: *Voprosy statističeskoj stilistiki*, Kiev, 203–216.

Kreuzer, H., Gunzenhäuser, R. (Hgg.):

- 1965 *Mathematik und Dichtung. Versuche zur Frage einer exakten Literaturwissenschaft (sammlung dialog 3)*. München (4/1971).

Krylov, Ju.K.:

- 1982 Ob odnoj paradigme lingvističeskich raspredelenij. In: *Tartu Riikliku Ülikooli Toimetised*, Bd. 628, Tartu, 80–97.

Krysin, L.P.:

- 1965 Inojazyčnaja leksika v rusškoj literaturnoj reči 20-ch godov. In: E.A. Zemskaja, D.N. Šmelev (red.), *Razvitie leksiki sovremennogo rusškogo jazyka*, Moskva, 117–134.

- 1968 Russkij jazyk po dannym massovogo oprosa. Prospekt. Moskva.
- 1970 Jazykovye varianty i social'noe rassloenie govorjaščich. *Russkij jazyk v nacional'noj škole* 4, 10–16.
- 1973 K social'nym različijam v ispol'zovanii jazykovych variantov. *Voprosy jazykoznanija* 3, 38–49.
- 1974 (red.) Russkij jazyk po dannym massovogo obsledovanija. Opyt social'no-lingvističeskogo izučenija. Moskva: Nauka.
- 1975 Die russische Sprache nach Daten einer Massenbefragung. In: W. Girke, H. Jachnow (Hgg.), *Sprache und Gesellschaft in der Sowjetunion*, München, 209–241.
- Krysin, L.P., Šmelev, D.N.:
- 1976 (red.) Social'no-lingvističeskie issledovanija. Moskva: Nauka.
- Kučera, H.:
- 1963 Entropy, redundancy and functional load in Russian and Czech. In: *American Contr. to the 5th International Congress of Slavists*, Vol. 1, The Hague: Mouton, 191–218.
- 1968 Some Quantitative Lexical Analyses of Russian, Czech, and English. The Hague: Mouton.
- Kučera, H., Monroe, G.K.:
- 1968 A Comparative Quantitative Phonology of Russian, Czech, and German. (*Mathematical Linguistics and Automatic Language Processing*, vol. 4). New York: American Elsevier Publ.
- Kudrjavskij, D.N.:
- 1909 K statistike glagol'nych form v Lavrent'evskoj letopisi. In: *Izvestija Otdelenija russkogo jazyka i slovesnosti Imp. Akad. Nauk* XIV, 2, 48–54.
- 1911 K istorii russkogo prošeďšego vremeni. In: *Russkij filologičeskij vestnik* LXV, 1, 119–139.
- 1912 Drevne-russkie pričastija nastojaščego vremeni na -a. In: *Russkij filologičeskij vestnik* LXVIII, 4, 389–397.
- Kühlwein, M.:
- 1968 (Rez.) N.P. Vakar, A Word Count of Spoken Russian, The Soviet Usage, Ohio 1966. *International Review of Applied Linguistics* 6, 200–202.
- Kulešova, L.V.:
- 1981 Opyt opisaniya kornej russkogo jazyka s pomošč'ju statističeskich metodov. In: *Količestvennye metody v gumanitarnych naukach*, Moskva, 200–204.
- Kunickij, V.N.:
- 1894 Jazyk i slog komedii "Gore ot uma". Kiev.
- Kupčinskij, O.A.:
- 1979 Statistika i stratigrafija vostočno-slavjanskich toponimov na -iči i nekotorye voprosy istoričeskoj geografii zaselenija: na materiale USSR. *Voprosy geografii* 110, 89–103.
- Kurbakov, K.I.:
- 1965 Nekotorye charakteristiki raspredelenija simvolov russkogo alfavita. *Naučno-tehničeskaja informacija* 9, 41–46.
- Kuz'mina, S.M.:
- 1966 O fonetike zaudarnyx fleksij. In: *Razvitie fonetiki sovremennogo russkogo jazyka*, Moskva, 5–24.
- 1976 Variativnost' proiznošenija grupp soglasnych. In: L.P. Krysin, D.N. Šmelev (red.), *Social'no-lingvističeskie issledovanija*, Moskva, 123–129.

Kuznecova, A.I.:

- 1980 Periferijnye javlenija v morfologii russkogo jazyka. In: *Aktual'nye voprosy strukturnoj i prikladnoj lingvistiki*, Moskva, 128–141.

Kvantitativní lingvistika

- 1964ff. Kvantitativní lingvistika. Praha.

Lapšina, N.V., Romanovič, I.K., Jarcho, B.I.:

- 1934 Metričeskij spravočnik k stichotvorenijam A.S. Puškina. Moskva–Leningrad: Academia.
1966 Iz materialov "Metričeskogo spravočnika k stichotvorenijam M.Ju. Lermontova". *Voprosy jazykoznanija* 2, 125–137.

Lebedev, D.S., Garmaš, V.A.:

- 1959 Statističeskij analiz trechbukvennych sočetanj russkogo teksta. In: *Problemy peredači informacii* II, Moskva, 78–80.

Lehfeldt, W.:

- 1972 Phonologische Typologie der slavischen Sprachen. *Die Welt der Slaven* XVII, 2, 318–340.
1975 Die Verteilung der Phonemanzahl in den natürlichen Sprachen. *Phonetica* 31, 274–287.
1977 Eine Methode zum synchronen Vergleich der Formenbildung in den slavischen Sprachen. In: Girke, W., Jachnow, H. (Hg.): *Slavistische Linguistik 1976*, München, 109–128.
1978 Formenbildung des russischen Verbs. Versuch einer analytisch-synthetisch-funktionellen Beschreibung der Präsens- und der Präteritumflexion. München.

Lehfeldt, W., Kempgen, S.:

- (i. D.) Quantitative morphologische Typologie. In: *Morphology. An International Handbook on Inflection and Word Formation*. Ed. by G. Booij, Chr. Lehmann, J. Mugdan. Berlin.

Lekomceva, M.I. et al.:

- 1963 Opyt postroenija fonologičeskoj tipologii blizkorodstvennych jazykov. In: *Slavjanskoe jazykoznanie, Doklady sovetskoj delegacii, V Meždunarodnyj s'ezd slavistov (Sofija, 1963)*, Moskva, 423–476.

Leont'ev, A.A. (red.):

- 1977 Slovar' assoziativnych norm russkogo jazyka. Moskva: Izd. Mosk. uni-ta.

Lesskis, G.A.:

- 1962 O razmerach predloženj v russkoj naučnoj i chudožestvennoj proze 60-ch godov XIX v. *Voprosy jazykoznanija* 2, 78–95.
1963a O zavisimosti meždju razmerom predloženiya i charakterom teksta. *Voprosy jazykoznanija* 3, 92–112.
1963b O nekotorych različijach prostogo predloženiya v naučnoj i chudožestvennoj proze. *Russkij jazyk v nacional'noj škole* 6, 7–15.
1964 O zavisimosti meždju razmerom predloženiya i ego strukturoj v raznyh vidach teksta. *Voprosy jazykoznanija* 4, 99–123.
1966a Dva sposoba opisaniya vnejazykovych situacij. In: *Lingvističeskie issledovanija po obščej i slavjansknoj tipologii*, Moskva, 32–50.
1966b K voprosu o grammatičeskich različijach naučnoj i chudožestvennoj prozy. In: *Trudy po znakovym sistemam*, II, Tartu.
1968 Nekotorye statističeskie charakteristiki prostogo i složnogo predloženiya v russkoj naučnoj i chudožestvennoj proze XVIII–XX vv. *Russkij jazyk v nacional'noj škole* 2.

Levin, Ju.I.:

- 1964 Ob opisanii sistemy lingvističeskich ob'ektov, obladajuščich obščimi svojstvami. *Voprosy jazykoznanija* 4, 112–119.

- 1967 O količestvennyh karakteristikah raspredelenij simvolov v tekste. *Voprosy jazykoznanija* 6, 112–121.
- Levinton, G.A.:
1981 Kak naprimer na JU. *Russian Linguistics* 6, 81–102.
- Levitskij, V.V.:
1966 Častotnyj slovar' jazyka učebnyh posobij medinstituta. Moskva: Izd. Mosk. uni-ta.
1989 Statističeskoe izučenie leksičeskoj semantiki. Kiev.
- Levý, J.:
1965 Die Theorie des Verses – ihre mathematischen Aspekte. In: H. Kreuzer, R. Gunzenhäuser (Hgg.), *Mathematik und Dichtung*, München, 211–231.
- [Lingvostatistika]
1982 Lingvostatistika i vyčislitel'naja lingvistika (Uč. Zapiski Tartuskogo Gos. Uni-ta, vyp. 628). Tartu.
- Ljatina, A.M.:
1968 Opyt statističeskogo analiza jazyka pisatelja (po materialam "Podnjatoj celiny" Šolochova). Kand. diss., Leningrad: LGU.
- Lönngren, L. (red.):
1993 Častotnyj slovar' sovremennogo russkogo jazyka. With a Summary in English. A Frequency Dictionary of Modern Russian. (*Acta Universitatis Upsaliensis, Studia Slavica Upsaliensia*, vol. 32) Uppsala.
- Lomtev, T.P.:
1965 Ob odnoj vozmožnosti istolkovanija fonologičeskogo razvitija. *Voprosy jazykoznanija* 3, 89–101.
1972 Fonologija sovremennogo russkogo jazyka na osnove teorii množestv. Moskva: Vysšaja škola.
- Losev, A.F.:
1965 O metodach izloženiya matematičeskoj lingvistiki dlja lingvistov. *Voprosy jazykoznanija* 5, 13–30.
- Luk'janenkov, K.F., Sergeeva, V.N.:
1978 Častotnyj anglo-russkij slovar'-minimum po sudovoždeniju. Moskva: Voenizdat.
- Maas, H.-D.:
1972 Über den Zusammenhang zwischen Wortschatzumfang und Länge eines Textes. *Lili, Zeitschrift für Literaturwissenschaft und Linguistik* 2,8, 73–96.
- Mačavariani, M.V.:
1963 O vzaimootnošenii matematiki i lingvistiki. *Voprosy jazykoznanija* 3, 85–91.
- Mal'ceva, G.F.:
1969 Nekotorye količestvennye priemy opisaniya individual'nogo avtorskogo stilja. In: *Statistika teksta*, t. I, Minsk, 206–248.
- Malř, F. (Hg.):
1951 Rusko-český slovník mejdůležitějších slov pro četbu sovětského tisku. Praha (21952).
1962 Untersuchungen zur Festlegung eines Wortschatzminimums für den Russischunterricht an der tschechoslowakischen neunjährigen Grundschule. *Sprach-Praxis*.

Malkova, O.V.:

- 1980 Imelo li mesto vtoričnoe smjagčenie soglasnyh pered [e], [i] v dialektach južnoj zony drevnerusskogo jazyka? *Voprosy jazykoznanija* 6, 76–88.
- 1981 K probleme vtorogo polnoglasija. *Voprosy jazykoznanija* 6, 97–108.

Manučarjan, R.S.:

- 1972 Voprosy interpretacii i izmerenija glubiny slova. *Voprosy jazykoznanija* 1, 114–123.

Markov, A.A.:

- 1913 Primer statističeskogo issledovanija nad tekstem 'Evgenija Onegina', illjustrirujuščij svjaz' ispytanij v cep'. In: *Izvestija Ross. Akademii Nauk (Bulletin)*, serija VI, t. 7, no. 3, 153–162.
- 1916 Ob odnom primenenii statističeskogo metoda. In: *Izvestija Imp. Akademii Nauk (Bulletin)*, serija VI, t. 10, no. 4, 239–242.
- 1924 Isčislenie verojatnostej. Izd. 4-e. Moskva.

Markov, Ju.:

- 1960a Leksičeskaja statistika i otbor slovarnogo minimuma. *Russkij jazyk v nacional'noj škole* 3, 3–14.
- 1960b K voprosu o častotnosti grammatičeskich kategorij. *Russkij jazyk v nacional'noj škole*, 4, 19–20.
- 1960c Častotnyj spisok 3.000 slov russkogo jazyka (po G. Džosselsonu). *Russkij jazyk v nacional'noj škole* 5, 75–83.
- 1963 O roli leksičeskoj statistiki pri otbore slovarnogo minimuma. *Russkij jazyk v nacional'noj škole* 4, 12–23.
- 1965/66 Über die Rolle der Wortstatistik bei der Auswahl eines Wortschatzminimums. *Zeitschrift für den Russischunterricht* II, 1, 51–64.
- 1966 Nekotorye aspekty razgovornoj reči s točki zrenija leksičeskoj statistiki. *Russkij jazyk v nacional'noj škole* 5, 21–28.
- 1967 Spisok 1500 naibolee upotrebitel'nyh slov. *Russkij jazyk v nacional'noj škole* 3.

Markov, Ju., Višnjakova, T.:

- 1965 Russkaja razgovornaja reč'. 1200 naibolee upotrebitel'nyh slov. *Russkij jazyk v nacional'noj škole* 6, 27–34.
- 1968 (red.) 2380 slov, naibolee upotrebitel'nyh v russkoj razgovornoj reči. Moskva, Izd. Universiteta družby narodov im. P. Lumumby.

Martynenko, G.Ja.:

- 1971 Statističeskoe issledovanie sintaksičeskoj složnosti predloženiya. In: *Informacionnye voprosy semiotiki, lingvistiki i avtomatičeskogo perevoda*, vyp. 1, Moskva, 84–101.
- 1972 Količestvennyj analiz sintaksičeskoj složnosti predloženiya. Avtoref. kand. diss. Leningrad.
- 1978 Nekotorye zakonomernosti koncentracii i rassejanija èlementov v lingvističeskich i drugih složnyh sistemach. In: *Strukturnaja i prikladnaja lingvistika. Mežvuzovskij sbornik*. Vyp. 1, Leningrad, 63–79.
- 1983 Mnogomernyj sintaktiko-statističeskij analiz chudožestvennoj prozy. In: *Strukturnaja i prikladnaja lingvistika. Mežvuzovskij sbornik*. Vyp. 2, Leningrad, 58–72.

Mašinnyj perevod i prikladnaja lingvistika.

- 1959ff. Sb. Mašinnyj perevod i prikladnaja lingvistika. Moskva.

Maslenikov, O.A.:

- 1956 Rhythm patterns in the trisemic verse of Andrej Belyj 1900–1909. In: M. Halle, H.G. Lunt et al. (eds.), *For Roman Jakobson. Essays on the Occasion of his Sixtieth Birthday*, The Hague: Mouton, 322–327.

Matejka, L.:

- 1963 Comparative Analysis of Syntactic Constructions in the Zadonščina. In: *Amer. Contr. Fifth Intern. Congr. of Slavists*, The Hague, 221–241.

[Materialy k slovarju Puškina]

- 1963 Materialy k častotnomu slovarju jazyka Puškina (prospekt). Moskva: AN SSSR, Institut jazykoznanija, Predvaritel'nye publikacii. Darin: R.M. Frumkina, *O statističeskom issledovanii slovarnogo sostava tekstov Puškina*, 5–8; N.G. Michajlovskaja, R.M. Frumkina, *2000 naibolee častych slov v tekstach Puškina* (Tabl. 1), 9–27; R.M. Frumkina, *Raspredelenie slov, vstrečajuščichsja v tekstach Puškina s dannoj častotoj* (Tabl. 2), 28–30; R.M. Frumkina, *Ėmpiričeskaja funkcija raspredelenija častot slov v tekstach Puškina* (Tabl. 3), 31; I.M. Davydova, D.Z. Neroslavskaja, R.M. Frumkina, *220 naibolee častych slov v različnyh žanrach proizvedenij Puškina* (Tabl. 4–7), 32–51.

[Materialy po primeneniju]

- 1961 Materialy mežvuzovsk. konfer. po primeneniju strukturnych i statističeskich metodov issledovanija slovarnogo sostava jazyka. Moskva.

Mdivani, R.R.:

- 1966 O lingvostatistike. In: *Voprosy sovremennogo obščego i matematičeskogo jazykoznanija*, Moskva, 149–169.

Mel'čuk, I.A.:

- 1959 Vsesojuznoe soveščanie po matematičeskoj lingvistike. In: *Mašinnyj perevod i prikladnaja lingvistika* 3, 82–93.
- 1963 O standartnoj forme i količestvennyh karakteristikah nekotoryh lingvističeskich opisanij. *Voprosy jazykoznanija* 1, 113–123.

Men'šikov, I.I.:

- 1974 K voprosu o žanrovo-stilevoj obuslovlennosti sintaksičeskoj struktury frazy. In: *Voprosy statističeskoj stilistiki*, Kiev, 138–146.
- 1978 O nekotoryh zakonernostjach vzaimoraspoloženija komponentov frazy v sovremenom russkom jazyke (količestvennyj analiz). In: *Strukturnaja i prikladnaja lingvistika. Mežvuzovskij sbornik*, vyp. 1, Leningrad, 30–38.

Meredov, E.:

- 1974 Ierarchičeskaja sistema inicial'nyh dvoučlennyh konsonantnyh sočetanij v sovremenom russkom jazyke. *Vestnik Moskovskogo universiteta* 3, 49–55.

Mežvuzovskaja konferencija

- 1966 Mežvuzovskaja konferencija po voprosam častotnyh slovarej i avtomatizacija lingvističeskich rabot. Tezisy dokladov i soobščenij. Leningrad: Izd-vo LGU.

Michajlov, V.:

- 1992 Informacionnye i statističeskie charakteristiki parametrov ustnoj reči. Moskva: Izd. Mosk. uni-ta.

Michajlovskaja, N.G.:

- 1964 Sinonimičeskie prilagatel'nye so značeniem 'sil'nyj po karakteru svoego projavlenija' v drevnerusskom jazyke XI–XIV vv. In: R.I. Avanesov (red.), *Issledovanija po istoričeskoj leksikologii drevnerusskogo jazyka*, Moskva, 18–42.

Milova, L.V.:

- 1994 (red.) Ot Nestora do Fonvizina. Novye metody opredelenija avtorstva. Moskva: Progress.

Minc, Z.G., L.A. Aboldueva, Šiškina, O.A.:

- 1967 Častotnyj slovar' "Stichov o prekrasnoj dame" A. Bloka i nekotorye zamečanija o strukture cikla. In: *Trudy po znakovym sistemam III, Uč. zap. Tartuskogo gos. uni-ta*, vyp. 198, Tartu, 209–316.

Mis'kevič, G.I.:

- 1977 K voprosu o prinjatii normoj novych slov. In: *Jazykovaja norma i statistika*, Moskva, 208–218.

Mistrík, J.:

- 1967 Matematiko-statističeskie metody v stilistike. *Voprosy jazykoznanija* 3, 42–52.
1973 Exakte Typologie von Texten (*Arbeiten und Texte zur Slavistik*, 3), München.

Moiseev, A.:

- 1975 Tipologija slogov v sovremennom russkom literaturnom jazyke. *Voprosy jazykoznanija* 6, 109–115.
1976 Mesto slovesnogo udarenija v sovremennom russkom literaturnom jazyke. In: *Studia Rossica Posnaniensia VII*, 1976, 77–87.

Morgan, G., Corbett, G.:

- 1989 Russian Colour Term Salience. *Russian Linguistics* 13, 125–141.

Morkovkin, V.V. (red.):

- 1972 Sravnitel'nyj spisok naibolee upotrebitel'nych russkich slov. In: P.N. Denisov (red.), *Leksičeskie minimumy russkogo jazyka*, Moskva, 16–74.
1985 Leksičeskie minimumy sovremennogo russkogo jazyka. Moskva: Russkij jazyk.

Morozov, N.A.:

- 1915 Lingvističeskie spektry. Sredstvo dlja otličenija plagiatov ot istinnych proizvedenij togo ili drugogo izvestnogo avtora. Stilemetričeskij étjud. In: *Izvestija Otdelenija Russkogo Jazyka i Slovesnosti Imp. Akad. Nauk*, t. XX, kn. 4, 93–134.

Morozova, M.I.:

- 1977 Variantnye formy prilagatel'nych (formy sravnitel'noj stepeni). In: *Jazykovaja norma i statistika*, Moskva, 59–75.

Moskovič, V.A.:

- 1965 Opyt kvantitativnoj tipologii semantičeskogo polja. *Voprosy jazykoznanija* 4, 80–91.
1967 Glubina i dlina slov v estestvennyh jazykach. *Voprosy jazykoznanija* 6, 17–33.
1969 Statistika i semantika. Opyt statističeskogo analiza semantičeskogo polja. Moskva: Nauka.

Mossner, F.:

- 1969 Welchen Einblick gibt uns die Faktoranalyse in den Satzbau der natürlichen Sprachen? *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 22, 243–254.

Muckov, L.:

- 1964 Izpolzuvane na dannite na statističeskata lingvistika pri obučenieto po ruski ezik. *Narodna prosveta* 2, 46–59.

Muratova, V.M.:

- 1974 Statističeskoe opisanie stilevoj differenciacii aktivnoj leksiki (po materialam «Častotnogo slovarja sovremennogo russkogo jazyka»). In: *Voprosy statističeskoj stilistiki*, Kiev, 196–203.

- Mustajoki, A.:
1981 Russian stress: Word frequency as an indicator of mobile stress. *Scando-Slavica* 27, 199–211.
- Nalimov, V.V.:
1974 Verojatnostnaja model' jazyka. Moskva.
- Neljubin, L.L.:
1977 Častotnyj rusko-anglijskij voennyj slovar'-minimum. Moskva.
1983 Perevod i prikladnaja lingvistika. Moskva: Vysšaja škola.
- Nesterov, M.N.:
1979 Častotnye slovari russkogo jazyka. *Russkaja reč'* 3, 108–112.
- Nevel'skij, P.B., Rozenbaum, M.D.:
1971 Ugadyvanie professional'nogo teksta specialistami i nespecialistami. In: *Statistika reči i avtomatičeskij analiz teksta*, Leningrad, 134–148.
- Nikitina, L.S.:
1971 Imennye trechslolvnye sočetačija v russkich publicističeskich tekstach. In: *Statistika reči i avtomatičeskij analiz teksta*, Leningrad, 262–285.
- Nikolaev, V.:
1960 Nekotorye dannye o častotnosti upotreblenija padežnyh form v sovremennom ruskom literaturnom jazyke. *Russkij jazyk v nacional'noj škole* 5, 19–26.
- Nikolaeva, T.M.:
1961 O ruskom jazyke v zarubežnyh rabotach po mašinomu perevodu. *Voprosy jazykoznanija* 5, 122–128.
- Nikonov, V.A.:
1959 Statistika padežej russkogo jazyka. In: *Mašinnyj perevod i prikladnaja lingvistika* 3, 45–65.
1960 Konsonantnyj koëfficient. *Lingua Posnaniensis* VIII, 228–233.
1961 Bor'ba padežej. *International Journal of Slavic Linguistics and Poetics* IV, 13–34.
1962a Konsonantnye sočetačija. *Russkij jazyk v nacional'noj škole* 3, 24–29.
1962b Metod issledovanija suffiksov prilagatel'nyh. In: *Strukturno-tipologičeskie issledovanija*, Moskva, 103–118.
1963a Interpretacija fonetičeskich častot. *Učenyje zapiski Instituta Slavjanovedenija AN SSSR* 27, 259–270.
1963b Mesto udarenija v ruskom jazyke. *International Journal of Slavic Linguistics and Poetics* VI, 1–8.
1967 Ličnye imena v sovremennoj Rossii. *Voprosy jazykoznanija* 6, 102–111.
1978 Dlina slova. *Voprosy jazykoznanija* 6, 104–111.
- Nosenko, I.A.:
1974 Statističeskaja differenciacija funkcional'nyh stilej. In: *Struktura i soderžanie kompleksnogo otraslevogo učebnika po inostrannomu jazyku*, Moskva.
1981 Načala statistiki dlja lingvistov. Moskva.
- Nováček, Č.:
1963 Ke kombinatoricé ruských konsonantů. *Československa rusistiká*, 193–198.
- Novak, L.A., Piotrovskij, R.G.:
1958 [Rez.] H.H. Josselson, The Russian word count and frequency analysis of grammatical categories of standard literary Russian, Detroit 1953. *Voprosy jazykoznanija* 3, 133–135.

Nowikowa, I. (ed.):

- 1972 2380 slov. Die gebräuchlichsten Wörter der russischen Umgangssprache. Hamburg: Buske (= 2380 slov, Moskva 1968).

Obratnyj slovar'

- 1974 Obratnyj slovar' russkogo jazyka. Okolo 125000 slov. Leningrad.

Oettinger, A.G.:

- 1954 The distribution of word-length in technical Russian. In: *Mechanical Translation* vol. 1, no. 3, 38–40.

Oliverius, Zd.F.:

- 1966 K distribuci alomorfů současné ruštiny. *Československa rusistika* XI, 207–214.
1970 K distribuci fonémů současné ruštiny. *Filologické studie* II, Praha, 11–22.
1976 Morfemy russkogo jazyka. Častotnyj slovar'. Praha: Univerzita Karlova.

Ol'ševskij, V.V., Postovalova, V.I.:

- 1968 K voprosu o strukture sloga v russkom jazyke. *Naučnye doklady Vysšej školy, Filologičeskie nauki*, 2, 53–68.
1969 Statističeskij analiz svjazej differencial'nych priznakov fonem v sloge. In: *Sistema i urovni jazyka*, Moskva, 56–67.

Orlov, Ju.K.:

- 1982 Ein Modell der Häufigkeitsstruktur des Vokabulars. In: H. Guiter, M.V. Arapov (eds.), *Studies on Zipf's Law (Quantitative Linguistics, vol. 16)*, Bochum, 154–233.

Orlov, Ju.K., Žuravlev, A.P.:

- 1967 Opyt psichometričeskoj ocenki smyslovych rasstojanij meždu transformami. *Voprosy jazykoznanija* 5, 93–103.

Otten, F.:

- 1978 Materialien zur Vertretung des Verbalsuffixes *-iva/-yva-* in der russischen Hochsprache des 16./17. Jahrhunderts (an Hand ausgewählter russischer Chroniken). *Zeitschrift für slavische Philologie* 40, 9–51.

Ovsienko, Ju.G.:

- 1966 Slovar' russkoj razgovornoj reči. In: *Mežvuzovskaja konferencija po voprosam častotnyh slovarej i avtomatizacija lingvostatističeskich rabot. Tezisy dokladov i soobščeniij*. Leningrad, Izd. LGU, 23–25.
1967 (Rez.) N.P. Vakar, A Word Count of Spoken Russian. The Soviet Usage, Ohio 1966. *Voprosy jazykoznanija* 6, 142–144.
1976 Zadači opisanija razgovornoj reči i voprosy postroenija lingvostatističeskogo spravočnika. In: *Voprosy lingvostatističeskogo analiza russkoj razgovornoj reči*. Sb. statej, Moskva, 3–9.
1977 Častotnyj slovar'-spravočnik russkoj razgovornoj reči. In: *Aktual'nye problemy učebnoj leksikografii*, Moskva, 94–101.

[PSML]

- 1966– Prague Studies in Mathematical Linguistics. Vol. 1–. Praha: Academia.

Padučeva, E.V.:

- 1958 Statističeskoe issledovanie struktury sloga. In: *Voprosy statistiki reči*, Leningrad, 100–111.

Palagina, V.V.:

- 1977 Materialy dlja častotnogo slovarja tomskich delovych dokumentov XVII veka. In: *Russkoe slovo v jazyke i reči* 2, Kemerovo, 82–94.

- 1979 Častotnyj slovar' tamožennych knig Tomska 1624–1627 gg. In: *Russkij govory v Sibiri*, Tomsk, 78–83.
- Panov, M.V. (red.)
1968 Russkij jazyk i sovetskoe obščestvo. Soziologo-lingvističeskoe issledovanie. T. 1: Fonetika sovremennogo russkogo literaturnogo jazyka, Narodnye govoy. T. 2: Morfologija i sintaksis sovremennogo literaturnogo jazyka. Moskva.
- Papina, A.F., Ivanova-Markova, L.P.:
1974 Bor'ba kategorij kačestvennosti i dejstvennosti v poëtičeskoj reči pervoj poloviny XIX ve-ka. In: *Voprosy statističeskoj stilistiki*, Kiev, 237–243.
- Papp, F.:
1961 Količestvennyj analiz slovarnoj struktury nekotorych russkich tekstov. *Voprosy jazykozna-nija* 6, 93–100.
1964 Mathematische und strukturelle Methoden in der sowjetischen Sprachwissenschaft. *Acta Linguistica Academiae Scientiarum Hungaricae*, XIV, 119–137.
1966 Mathematical Linguistics in the Soviet Union (*Janua Linguarum, Series Minor No. XL*). The Hague: Mouton.
1967 O nekotorych količestvennyh charakteristikach slovarnogo sostava jazyka. *Slavica (De-brecen)* 7, 51–58.
- Perebejnos (Perebyjnis), V.I.:
1967a (red.) Statistyčni parametry styliv. Kiev: Naukova Dumka.
1967b Ispol'zovanie statističeskich metodov v tipologičeskich issledovanijach (na materiale sla-vjanskich i germanskich jazykov). In: Filin, F.P. (red.), *Problemy jazykoznanija. Doklady i soobščeniya sovetskich učenych na X Meždunarodnom Kongrese lingvistov*, Moskva, 230–234.
1974 Metody i urovni modelirovanija nulevogo stilja. In: *Voprosy statističeskoj stilistiki*, Kiev, 16–35.
1990 (red.) Statističeskaja leksikografija i učebnyj process. Kiev.
- Peškovskij, A.M.:
1925/26 Desjat' tysjač zvukov russkogo jazyka. In: *Metodika rodnogo jazyka, lingvistika, stilistika, po-ëtika*, Leningrad–Moskva, 167–191.
- Peterson, M.N.:
1928 Konstrukcii s predlogom 'z' u Lermontova. In: *Sbornik Otdelenija russkogo jazyka i sloves-nosti* 101, 3, 410–412.
1937 Sintaksis Slova. *Slavia* 14, 347–392.
- Petrenko, V.F.:
1965 Opyt statističeskogo issledovanija ključevych slov teksta. *Vestnik LGU* 20, 150–155.
- Petrova, N.V., Piotrovskij, R.G.:
1966 Slovo, kontekst, morfologija. *Voprosy jazykoznanija* 2, 111–124.
- Piotrovskaja, A.A., Piotrovskij, R.G.:
1974 Matematičeskie modeli diachronii i tekstoobrazovanii. In: *Statistika reči i avtomatičeskij analiz teksta 1974*, Leningrad, 361–400.
- Piotrovskaja, A.A., Piotrovskij, R.G., Razživin, K.A.:
1962 Ėntropija russkogo jazyka. *Voprosy jazykoznanija* 6, 115–130 [engl. Übers. in Foreign De-velopments in Machine Translation and Information Processing no. 121, 1963, 1–28].

Piotrovskij, R.G.:

- 1958 Nekotorye voprosy statističeskogo obsledovanija leksičeskich grupp. In: *Voprosy statistiki reči*, Leningrad, 85–92.
- 1960 Ešče raz o diferencial'nych priznakach fonem. *Voprosy jazykoznanija* 6, 24–38.
- 1961 O matematičeskom jazykoznanii. *Russkij jazyk v nacional'noj škole* 2, 9–15.
- 1962 O teoretiko-informacionnyh parametrach ustnoj i pis'mennoj form jazyka. In: S.K. Šaumjan (red.), *Problemy strukturnoj lingvistiki*, Moskva, 43–59.
- 1965 Teoretiko-informacionnaja struktura russkogo slova. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 18, 149–172.
- 1966 Informacionnye izmerenija pečatnogo teksta. In: *Ėntropija jazyka i statistika reči*, Minsk, 5–86.
- 1968 Informacionnye izmerenija jazyka. AN SSSR, Izd. Nauka, Len. Otd.
- 1969a Lingvostatističeskie issledovanija v gruppe "Statistika reči". In: *Statistika teksta*, t. 1, Minsk, 6–9.
- 1969b Entropy and redundancy in four European languages. *Statistical Methods in Linguistics* 5, 34–35.
- 1979 Inženernaja lingvistika i teorija jazyka. Leningrad: Nauka.

Piotrovskij, R.G., Bektaev, K.B., Piotrovskaja, A.A.:

- 1977 Matematičeskaja lingvistika. Posobie. Moskva: Vysšaja škola.
- 1985 Mathematische Linguistik. Übers. von A. Falk (*Quantitative Linguistics*, vol. 27). Bochum.

Piotrovskij, R.G., Popeskul, A.N., Čažinskaja, M.S., Rachubo, N.P.:

- 1985 Automatische Wortschatzanalyse. Übers. von A. Falk (*Quantitative Linguistics*, vol. 28). Bochum.

Piotrovskij, R.G., Turygina, L.A.:

- 1971 Antonimija «jazyk – reč'» i statističeskaja interpretacija normy jazyka. In: *Statistika reči i avtomatičeskij analiz teksta*, Leningrad, 5–46.

Pleskalová, J.:

- 1984 Zur Klassifizierung der slawischen Flunamen nach Modellen. *Zeitschrift für Slawistik* 29, 80–89.

Plotnikov, B.A.:

- 1979 Distributivno-statističeskij analiz leksičeskich značenij. Pod red. A.E. Suprunu. Minsk: Vyšejšaja škola.
- 1980 Leksičeskaja sočetaemost' i značenie slova v tekste. In: *Vosprijatie jazykovogo značeniija*, Kaliningrad, 64–71.
- 1984 O tipologii svjazej meždu formoj i značeniem slov. *The Prague Bulletin of Mathematical Linguistics* 42, 27–36.
- 1985 Sočetaemost' i semantika. In: *Metody izučeniija leksiki*, Minsk, 74–81.
- 1986 Meaning, Word Combination, and Statistics. In: *Wiener Slawistischer Almanach* 18, 251–260.

Polivanov, E.:

- 1934 I matematika byvaet poleznoj. In: *Za marksistskoe jazykoznanie. Sb*, Tbilisi.

Poljakova, G.P., Solganik, G.Ja.:

- 1971 Častotnyj slovar' jazyka gazety. Moskva: Izd. MGU.

Potapov, V.V.:

- 1993 Jazykovaja specifiika strukturno-komponentnoj aktualizacii ritma reči. *Voprosy jazykoznanija* 5, 83–97.

Problemy peredači informacii

- 1965 Problemy peredači informacii. T. 1, vyp. 1. Moskva.

Proskurnin, N.:

- 1933 Podsčety častoty liter i komplektovka šrifta. In: *Pis'mennost' i Revoljucija* Sb. I, Moskva–Leningrad, 72–82.

Pšeničnova, N.N.:

- 1973 Statističeskij analiz mnogočlennogo dialektnogo sootvetstvija (Gluchie dolgie šipjaščie v russkich govorach). *Voprosy jazykoznanija* 6, 80–91.
- 1977a Primenenie metoda taksonomičeskogo analiza k klassifikacii govorov. In: *Dialektologičeskie issledovanija po rusckomu jazyku*, Moskva, 3–14.
- 1977b O primenenii statistiki pri izučenii govorov. In: *Govory territorij pozdnego zaselenija*, Saratov, 49–57.
- 1979a Mera specifičnosti i nekotorye voprosy klassifikacii častnyh dialektnyh sistem. In: *Obščeslavjanskij lingvističeskij atlas. Materialy i issledovanija. 1977*, Moskva, 236–249.
- 1979b Nekotorye sposoby gruppirovanija ob'ektov, primenjaemye v dialektologii. In: *Problemy strukturnoj lingvistiki 1979*, Moskva, 278–286.

Raecke, J.:

- 1993 Zu den möglichen Quellen einer Geschichte der (modernen) russischen Literatursprache. In: S. Kempgen (Hg.), *Slavistische Linguistik 1992*, München, 197–224.

Rakušan, J.:

- 1981 On the Expressivity of Nominal Suffixes in Modern Russian. An Attempt at Quantitative Analysis. *Russian Linguistics* 6, 41–55.

Rathmayr, R.:

- 1976 Die perfekte Präsensform im Russischen (*Öst. Akad. d. Wiss., Phil.-hist. Klasse, Sitzungsber., 310. Bd., 1. Abhandlung*). Wien.

[Razvitie fonetiki]

- 1966 Razvitie fonetiki sovremennogo russkogo jazyka. Moskva: Nauka.

Reinton, J.E.

- 1978 The Relationship between English Loanwords and Their Synonyms in Russian Sport Terminology. *Scando-Slavica* XXIV, 217–237.

Restan, P.:

- 1981 Word Classes, Rhythm, Frequency – Stylistic Observations (Russian Material). *Scando-Slavica* 27, 225–232.

Rotar', A.S., Čižakovskij, V.A.:

- 1976 Častotnyj nemecko-russkij slovar'-minimum gazetnoj leksiki. Moskva: Voenizdat.

Roberts, C.B.:

- 1976 Lexical Differentiation of the Russian Prefixal Allomorphs O-, OB-, OBO-. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 29, 64–76.

Robblee, K.E.:

- 1989 Syntax and Colon Structure in the Russian Funeral Lament. *The Slavic and East European Journal* 33, 3, 411–421.

- Roovet, È., Štejnfel'dt, È.A.:
1965 Slovar'-minimum russkogo jazyka dlja 2-8 klassov èstonskich škol. Tallin.
- Rusova, N.Ju.:
1974 Opyt statističeskogo analiza stilevoj differenciacii struktury prostogo predloženiija. In: *Voprosy statističeskoj stilistiki*, Kiev, 125-138.
- Russkoe stichosloženie
1979 Russkoe stichosloženie XIX v. Materialy po metrike i strofike russkich počtov. Red. kolleģija M.L. Gasparov, M.M. Giršman, L.I. Timofeev. Moskva: Nauka.
- Rylov, S.A.:
1974 Iz nabljudenij nad grammatičeskim razvertyvaniem struktury vyskazyvanija. In: *Voprosy statističeskoj stilistiki*, Kiev, 147-156.
- Sadčikova, P.V.:
1974 Anglijskij i russkij častotnye slovari po chimii polimerov. In: *Statistika reči avtomatičeskij analiz teksta 1974*, Leningrad, 295-329.
- Saf'jan, Ju.A.:
1971 Častotnyj slovar' russkoj tehničeskoj leksiki. Erevan.
- Sajachova, L.:
1964 Opyt sostavlenija slovarja-minimuma dlja nerusskoj školy na osnove častotnyh slovarj. *Russkij jazyk v nacional'noj škole* 3, 21-30.
- Sambor, J.:
1972 Słowa i liczby. Zagadnienia językoznawstwa statystycznego. Wrocław etc.
- Sannikov, V.Z.:
1963 Mesto rasprostranennogo opredelenija po otnošeniju k opredeljaemomu slovu v russkoj fraze. *Voprosy jazykoznanija* 1, 124-130.
1974 Alfavitnyj, častotnyj i obratnyj slovari vostočnoslavjanskich juridičeskich tekstov XI-XVI vv. Vyp. 1-5 (*Predvaritel'nye publikacii Instituta russkogo jazyka AN SSSR, vyp. 53, 54, 82, 83, 84*). Moskva.
1975 Očerok vostočnoslavjanskoj sravnitel'no-istoričeskoj leksikologii (lingvostatističeskij aspekt) (*Predvaritel'nye publikacii Instituta russkogo jazyka AN SSSR, vyp. 69*). Moskva.
1978 Očerok vostočnoslavjanskoj sravnitel'no-istoričeskoj leksikologii (Lingvostatističeskij aspekt). *Russian Linguistics* 4, 13-40.
1983 Ein Versuch zur Anwendung statistischer Methoden in der ostslavischen historisch-vergleichenden Sprachwissenschaft. In: B. Brainerd (ed.), *Historical Linguistics (Quantitative Linguistics, vol. 18)*, Bochum, 62-103.
1985 O stepeni leksičeskoj blizosti drevnerusskoj, starorusskoj, staroukrainskoj i starobeloruskoj pis'mennoj reči. In: *Vostočnye slavjane. Jazyki, istorija, kul'tura. K 85-letiju ak. V.I. Borovskogo*, Moskva, 156-164.
- Šanskij, N.M. (red.):
1982 4000 naibolee upotrebiteľ'nych slov russkogo jazyka. Izd. 3-e, ispravl. Moskva: Russkij Jazyk.
- Šapir, M.I.:
1992 "Gore ot uma": semantika poëtičeskoj formy (Opyt praktičeskoj filosofii sticha). *Voprosy jazykoznanija* 5, 90-105.

Schaller, H.W.:

- 1970 Informationstheorie und Wortstellung (gezeigt am Beispiel des Russischen). *Beiträge zur Linguistik und Informationsverarbeitung* 19, 40–61.

Schenkowitz, G.:

- 1976 Der Inhalt der sowjetrussischen Vorlesestoffe für Vorschulkinder. Eine quantifizierende Corpusanalyse unter Benutzung eines Computers (*Slavistische Beiträge, Bd. 95*). München: Otto Sagner.

Schmitz, W., Zühlke, Chr.:

- 1974 Zur formalen Analyse von Lehrwerken (Lehrbüchern) für den Russischunterricht. In: *Notizen und Materialien zu russistischen Linguistik, Unterlagen für die Seminararbeit 2*, Frankfurt, 47–101.

Schreiner, J.:

- 1989 Die Kategorie des Gedichtumfangs und die Vierzeiler bei Anna Achmatova. *Die Welt der Slaven* XXXIV, 197–235.

Segal, D.M.:

- 1972 *Osnovy fonologičeskoj statistiki (na materiale pol'skogo jazyka)*. Moskva: Nauka.

Šeljachovskaja, L.A.:

- 1963 Opyt statističeskogo analiza upotreblenija abbreviativ v sovremennom russkom jazyke. In: *Sinchronnie izučenie različnych jarusov struktury jazyka*, Alma-Ata, 54–56.

Šengeli, G.:

- 1923 *Traktat o russkom stiche*. Moskva—Petrograd.
1960 *Technika sticha*. Moskva.

Serebrjakov, L.A.:

- 1974 O stilevoj variativnosti funkcionirovanija vido-vremennyh form sovremennogo russkogo glagola (vremennaja paradigma nesoveršennoho vida). In: *Voprosy statističeskoj stilistiki*, Kiev, 109–125.

Sevbo, I.P., Petuchin, Ju.I.:

- 1976 Issledovanie sintaksičeskoj struktury statističeskimi metodami. In: *Naučno-techničeskaja informacija*, Ser. 2, no. 2, 17–36.

Shannon, C.E., Weaver, W.:

- 1949 *The Mathematical Theory of Communication*. Urbana.

Širokov, O.S.:

- 1961 Konferencija po strukturnoj i matematičeskoj lingvistike. *Voprosy jazykoznanija* 1, 155–159.

Sirotinina, O.B.:

- 1965 Porjadok slov v russkom jazyke. Saratov.
1968 Nekotorye žanrovo-stilističeskie izmenenija sovschoj publicistiki. In: *Razvitie funkcional'nych stilej sovremennogo russkogo jazyka*, Moskva, 101–126.
1969a Izmenenija v jazyke publicistiki (na materiale meždunarodnyh obzorov). In: *Voprosy stilistiki*, Saratov, 5–37.
1969b Izmenenija v jazyke naučnoj prozy. In: *Voprosy stilistiki*, Saratov, 37–55.
1974 Formy vremeni v razgovornoj reči. In: *Voprosy statističeskoj stilistiki*, Kiev, 97–109.

Sirotinina, O.B., Žuk, R.T., Tokareva, L.I.:

- 1968 O leksiko-semantičeskoj obuslovlennosti porjadka slov (Na materiale pis'mennoj reči). In: *Voprosy slavjanskogo jazykoznanija*, Saratov, 88–100.

[Sistemnyj analiz]

1988 Sistemnyj analiz lingvističeskich javlenij v tekste. Vladivostok.

Šljachovskaja, L.L.:

1964 O častotnosti abreviatur v jazyke gazet. In: *Aktual'nye voprosy sovremennogo jazykoznanija i lingvističeskoe nasledie E.D. Polivanova. Materialy konferencii*, t. I, Samarkand.

Slovar' jazyka V.I. Lenina:

[1990] Slovar' jazyka V.I. Lenina. Alfavitno-častotnyj slovoukazatel' k Polnomu sobr. soč. Moskva: Nauka [angekündigt, aber bisher nicht erschienen].

Slovar' jazyka Puškina [SlJaP]

1956–61 Slovar' jazyka Puškina. Tt. I–IV. Moskva.

Smirnova, L.:

1958 K voprosu o častotnosti upotreblenija russkich slov. *Russkij jazyk v nacional'noj škole* 2, 69–76.

Smith, G.S.:

1985 The Metrical Repertoire of Russian Émigré Poetry, 1941–1970. *Slavonic and East European Review* 63, 210–227.

Smith, R.N.:

1968 A Quantitative Study of Russian Vocabulary. Ph.D. Thesis, Brown University.

Sokal, R.R., Sneath, P.H.A.:

1963 Principles of Numerical Taxonomy. San Francisco.

Solonicyn, Ju.V.:

1966 O cifrovych pokazateljach, karakterizujuščich «silu» upravljenija. *Voprosy jazykoznanija* 2, 80–83.

Sprachstatistik

1973 Sprachstatistik. Herausgegeben von P.M. Alexejew, W.M. Kalinin, R.G. Piotrowski. Übersetzt von einem Kollektiv unter Leitung von Lothar Hoffmann. München—Salzburg: W. Fink [Originaltitel: Statistika reči, L. 1968].

Statistika reči (StR)

1968 Statistika reči. Red. kolegija: P.M. Aleksejev, V.M. Kalinin, R.G. Piotrovskij. Leningrad: Nauka [dt. u.d.T. Sprachstatistik. München—Salzburg 1973].

Statistika reči i avtomatičeskij analiz teksta (StRAAT)

1971 Statistika reči i avtomatičeskij analiz teksta. Leningrad.

1973 Statistika reči i avtomatičeskij analiz teksta 1972. Leningrad.

1974 Statistika reči i avtomatičeskij analiz teksta 1974. Leningrad.

1979 Statistika reči i avtomatičeskij analiz teksta 1979. Moskva–Leningrad.

1980 Statistika reči i avtomatičeskij analiz teksta 1980. Leningrad.

Statistika teksta I (StT–I)

1969 Statistika teksta. Sb. statej. Tom I: Lingvostatističeskije issledovanija. Redkolegija: R.G. Piotrovskij et al. Minsk: Izdatel'stvo BGU.

Statistika teksta II (StT–II)

1970 Statistika teksta. Sb. statej. T. II: Avtomatičeskaja pererabotka teksta. Red.: A.I. Kiselevskij et al. Minsk: Izdatel'stvo BGU.

Štejnfel'dt, Ė.A.:

1961 Slovar'-minimum russkogo jazyka dlja II–IV klassov éstonskich škol. Tallin.

- 1962 Častotnyj slovar' sovremennogo ruskogo jazyka. *Russkij jazyk v nacional'noj škole* 4, 32–39 und 6, 13–16.
- 1963 Častotnyj slovar' sovremennogo ruskogo jazyka. 2500 naibolee upotrebitel'nyh slov. Posobie dlja prepodavatelej ruskogo jazyka. Tallin.
- 1965a Opyt sostavlenija častotno slovarja sovremennogo ruskogo literaturnogo jazyka. In: *Materialy V-ogo meždunarodnogo metodičeskogo seminara prepodavatelej ruskogo jazyka stran socializma, Ijun' 1962*, Moskva, 84–114.
- 1965b Častotnyj slovar' sovremennogo ruskogo jazyka. Spravočnik dlja prepodavatelej ruskogo jazyka. Moskva.
- [o.J.] Häufigkeitswörterbuch der russischen Sprache. 2500 meistgebrauchte Wörter der modernen russischen Schriftsprache. Handbuch für Russischlehrer. Moskau: Progress.
- [o.J.] Dictionnaire des frequences de mots dans la langue russe moderne. Moscou.
- 1966 Russian Word Count of spoken Russian, The Soviet Usage (Častotnyj slovar' sovremennoj ruskoj reči). Ohio.
- Stepanova, E.M. (red.):
- 1970 Častotnyj slovar' obščenučnoj leksiki. Pod obščej red. E.M. Stepanovoj. Moskva: Izd. Mosk. Uni-ta.
- Stroevea, T.V.:
- 1968 Sopostavitel'naja statistika padežnyh slovoform imeni suščestvitel'nogo v nemeckom i ruskom jazykach. *Inostrannye jazyki v škole* 5.
- Struve, P.G.:
- 1918 Kto pervyj ukazal na primenenie statistiki k filologičeskim issledovanijam? In: *Izvestija Rossijskoj Akademii Nauk (Bulletin)*, VI ser., No. 13, 1317–1318.
- Suchotin, B.V.:
- 1962 Ėksperimental'noe vydelenie klassov bukv s pomošč'ju elektronnoj vyčislitel'noj mašiny. In: *Problemy strukturnoj lingvistiki 1962*, Moskva, 198–206.
- 1963 Algoritmy lingvističeskoj dešivrovki. In: *Problemy strukturnoj lingvistiki 1963*, Moskva, 75–100.
- 1976 Optimizacionnye metody issledovanija jazyka. Moskva: Nauka.
- 1984 Vydelenie morfem v tekstach bez probelov meždu slovami. Moskva: Nauka.
- Sudakov, G.V.:
- 1985 Leksičeskie dialektizmy i dialektnye ob'edinenija jazyka Moskovskoj Rusi. *Voprosy jazykoznanija* 5, 83–93.
- 1986 Predmetno-bytovaja leksika v onomasiologičeskom aspekte. *Voprosy jazykoznanija* 6, 105–112.
- Šuljak, N.T.:
- 1976 O nekotoryh rezul'tatach issledovanija metodiki postroenija častotnyh slovarej. In: *Strukturnaja i matematičeskaja lingvistika* 4, Kiev, 51–54.
- Šunevič, B.I.:
- 1985 Častotnyj anglo-ruskij slovar' po robo-technike. In: *Kvantitativnaja lingvistika i avtomatičeskij analiz tekstov. Trudy po lingvostatistike*, Tartu, 117–126.
- Suprun, A.E.:
- 1969 Slavjanskije čislitel'nye. Minsk: BGU.
- 1979 K količestvennoj ocenke leksičeskogo bogatstva teksta. *Filologičeskie nauki* 1, 44–48.

Švec, A.V.:

- 1979 Publicističeskij stil' sovremennogo russkogo literaturnogo jazyka (O nekotorych osobennostjach sintaksičeskogo stroja sovremennoj gazetnoj reči i tendencijach ee razvitija). Kiev.

Taranovski(j), K.:

- 1953 Ruski dvodelni ritmovi. I–II. Beograd.
1955-56 Ruski četverostopni jamb u prvim dvemadecenijama XX veka. *Južno-slovenski filolog* XVI.
1966a Četyrechstopnyj jamb Andreja Belogo. *International Journal of Slavic Linguistics and Poetics* X, 127–147.
1966b Osnovnye zadači statističeskogo izučenija slavjanskogo sticha. In: *Poetics* II, Warszawa—The Hague, 173–195.

Terechova, T.G.:

- 1966 Proiznošenie sočetańj trex soglasnyx v sovremennom russkom literaturnom jazyke. In: *Razvoitje fonetiki sovremennogo russkogo jazyka*, Moskva, 72–84.

Těšitelová, M.:

- 1980 Využití statistických metod v gramatice (*Studie a práce lingvistické* 15). Praha: Academia.
1985 Kvantitativní charakteristiky současné češtiny (*Studie a práce lingvistické* 19). Praha: Academia.
1987 O češtině v číslech. Praha: Academia.
1992 Quantitative Linguistics (*Linguistic and Literary Studies in Eastern Europe*, vol. 37) Amsterdam–Philadelphia: John Benjamins.

Tezisy Konferencii

- 1959 Tezisy Konferencii po matematičeskoj lingvistike. Leningrad.

Tichonov, A.N.:

- 1983 Sistema russkogo slovoobrazovanija s točki zrenija količestvennych dannych. In: Bartkov, B.I. (red.), *Issledovanie derivacionnoj podsistemy količestvennyjm metodom*, Vladivostok, 61–73.

Tolstaja, S.V.:

- 1968a Načal'nye i konečnye sočetańija soglasnyx v slavjanskich jazykach. Moskva: Avtoref. kand. diss.
1968b Sočetaemost' soglasnyx v svjazi s fonologičeskoj strukturoj slova v slavjanskich jazykach. *Sovetskoe slavjanovedenie* 1/4, 41–54.
1968c Fonologičeskoe rasstojanie i sočetaemost' soglasnyx v slavjanskich jazykach. *Voprosy jazykoznanija* 3, 66–81.
1974 K charakteristike konsonantnyx sočetańj v slavjanskich jazykach (načalo i konec slova). *Slavia* 43, 113–133.

Tomaševskij, B.:

- 1923 Pjatistopnyj jamb Puškina. Berlin.
1929 O stiche. Leningrad.

Toporov, V.N.:

- 1959 O vvedenii verojatnosti v jazykoznanie. *Voprosy jazykoznanija* 6, 28–35.
1966 Materialy dlja distribucii grafem v pis'mennoj forme russkogo jazyka. In: *Strukturnaja tipologija jazykov*, Moskva, 65–143.

Tornow, S.:

- 1980 (Rez.) Zazorina, L.N., Častotnyj slovar' russkogo jazyka. Moskva 1977. *Zeitschrift für Slavische Philologie* 41, 226–229.

- 1984 Die häufigsten Akzenttypen in der russischen Flexion (*Veröffentl. d. Abtlg. f. Slavische Sprachen und Literaturen des Osteuropa-Instituts an der FU Berlin*, Bd. 57). Berlin–Wiesbaden.
- Tuldava, Ju.A.:
- 1983 (ed.) *Kvantitativnaja lingvistika i stilistika*. Tartu.
- 1987 *Problemy i metody kvantitativno-sistemnogo issledovanija leksiki*. Tallin: Valgus.
- Turko, L.A.:
- 1968 *Častotnyj slovar' russkoj razgovornoj reči*. In: *Statistika reči*, Leningrad, 191–199.
- Trubeckoj, N.S.
- 1939 *Zur phonologischen Statistik*. *TCLP 7*, Prague, 230–241.
- Tvorogov, O.V.:
- 1967 *O primenenii častotnych slovarej i istoričeskoj leksikologii russkogo jazyka*. *Voprosy jazykoznanija 2*, 109–117.
- 1984 *Leksičeskij sostav "Povesti vremennyh let"* (Slovoukazatel' i častotnyj slovník). Kiev.
- Uluchanov, I.S.:
- 1964 *Predlogi pred" – pered" v russkom jazyke XI–XVII vv.* In: R.I. Avanesov (red.), *Issledovanija po istoričeskoj leksikologii drevnerusskogo jazyka*, Moskva, 125–160.
- Uspenskij, V.A.:
- 1958 *Soveščanie po statistike reči*. *Voprosy jazykoznanija 1*, 170–173.
- Vakar, N.P.:
- 1966 *A Word Count of Spoken Russian: The Soviet Usage*. Columbus: Ohio State University Press.
- 1967 *Statistical Methods in the Analysis of Russian*. *Slavic and East European Journal XI*, 59–65.
- Valnickij, A.:
- 1968 *Leksičeskij minimum dlja nerusskich, izučajuščich russkij jazyk*. *Russkij jazyk za rubežom 3*, 55–56.
- Vannikov, Ju.V.:
- 1976 *Lingvostatističeskij spravočnik russkoj razgovornoj reči i polučenie lingvističeskoj informacii vysšich urovnej*. In: *Voprosy lingvostatističeskogo analiza russkoj razgovornoj reči*. *Sb. statej*, Moskva, 13–66.
- Vannikov, Ju.V., Ichmal'jan, S.:
- 1969 *Statističeskij analiz distantno-pozicionnyh charakteristik sintaksičeskich struktur*. In: *Voprosy sopostavitel'nogo i obščego jazykoznanija*, Moskva.
- 1970 *Statističeskij analiz distantnyh, pozicionnyh, formal'nych i funkcional'nych svojstv élementarnyh sintaksičeskich struktur (v razgovornom i knižno-literaturnom stiljach russkogo jazyka)*. In: *Russkaja razgovornaja reč'*, Saratov, 192–205.
- Vannikov, Ju.V., Maslov, L.V.:
- 1969a *Častota i rasprostranennost'*. In: *V pomošč prepodavateljam russkogo jazyka kak inostrannogo*, Moskva: izd. MGU, 103–132.
- 1969b *Problema statističeskoj korrelnacii pokazatelej častoty i rasprostranennosti slova*. In: *Problemy lingvističeskogo analiza*, *Sb. statej*, Moskva, 33–63.
- Vasil'evič, A.P., Skokan, Ju.N.:
- 1986 *K metodike sopostavitel'nogo issledovanija (Na primere leksiki cvetooboznačenij)*. *Voprosy jazykoznanija 3*, 103–110.

Vejlert, A.A.:

- 1967 Sistema glagolov v verchnenemeckom govore s. Najdorf (Statističeskoe issledovanie). Kand. diss., Alma-Ata.
- 1973 Ob ispol'zovanii količestvennyh dannyh v dialektologii. *Voprosy jazykoznanija* 4, 119–123.
- 1974 Statističeskoe obosnovanie naibolee tipičnyh form vremeni v verchnenemeckom govore sela Najdorf. *Trudy Krasnodarskogo politechn. in-ta*, vyp. 50, 91–97.
- 1979 Russkoe slovo v nemeckoj dialektnoj reči (K metodike ocenki materiala). *Voprosy jazykoznanija* 3, 82–94.
- 1980 Izmerenie stepeni sootvetstvija sistem glasnyh nemeckogo literaturnogo standarta i ostrovgogo dialekta. *Voprosy jazykoznanija* 1, 102–113.

Vesler, I.Ju.:

- 1980 Algoritm sžatija russkich slov, učityvajuščij dlinu slov. In: *Strukturnaja i matematičeskaja lingvistika* 8, 14–22.

Vinarskaja, E.N.:

- 1969 K izučeniju verojatnostnyh otnošenij v rečevoj dejatel'nosti. *Voprosy jazykoznanija* 1, 93–97.

Vinogradova, V.N.:

- 1958 Izbytočnosť kategorii roda u russkogo glagola. In: *Tezisy Konferencii po mašinnomu perevodu*, Moskva, 84–86.

Vinokur, G.O.:

- 1959 Nasledstvo XVIII v. stichotvornom jazyke Puškina. In: ders., *Izbrannye raboty po russkomu jazyku*, Moskva, 328–387.

Vitov, M.V.:

- 1967 Severnorusskaja toponimija XV–XVIII vv. (K postanovke toponimičeskogo istočnikovvedenija). *Voprosy jazykoznanija* 4, 75–90.

Vjalkina, L.V.:

- 1964 Složnye slova v drevnerusskom jazyke i ich otnošenii k jazyku grečeskogo originala (Na materiale Efremovskoj kormčej). In: R.I. Avanesov (red.), *Issledovanija po istoričeskoj leksikologii drevnerusskogo jazyka*, Moskva, 94–118.

Vjalkina, L.V., Lukina, G.N.:

- 1964 Opyt primenenija nekotoryh metodov matematičeskoj statistiki k izučeniju leksiki drevnerusskich tekstov. In: R.I. Avanesov (red.), *Issledovanija po istoričeskoj leksikologii drevnerusskogo jazyka*, Moskva, 298–307.
- 1966 Materialy k častotnomu slovarju drevnerusskich tekstov. In: *Leksikologija i slovoobrazovanie drevnerusskogo jazyka*, Moskva, 263–292.

Volockaja, Z.M. et al.:

- 1958 O russkom slovare častotnosti na materiale matematičeskich tekstov. In: *Voprosy statistiki reči (materialy soveščanija)*, Leningrad, 93–99.
- 1963 Ob odnom podchode k tipologii slavjanskich jazykov (na materiale sistemy sklonenija suščestvitel'nogo). In: *Slavjanskoe jazykoznanie. Doklady sovetskoj delegacii. V Meždunarodnyj s'ezd slavistov (Sofija, 1963)*, Moskva, 510–551.

Voprosy lingvostatističeskogo analiza [VLARR]

- 1976 Voprosy lingvostatističeskogo analiza russkoj razgovornoj reči. Sbornik statej. Moskva.

Voprosy lingvostatistiki i avtomatizacii [VLiA]

- 1967 Voprosy lingvostatistiki i avtomatizacii lingvističeskich rabot. Vyp. 1 (*Trudy CNIPI*, serija 3), Moskva.

Voprosy statističeskoj stilistiki [VStS]

- 1974 Voprosy statističeskoj stilistiki. Red. kolegija: B.N. Golovin, V.I. Perebejnos et al. Kiev.

Voprosy statistiki reči [VStR]

- 1958 Voprosy statistiki reči. Materialy soveščanija. Leningrad (1957).

Voprosy stilistiki [VSt-I, VSt-II]

- 1962-65 Voprosy stilistiki. Vyp. 1-2. Saratov.

Voroncova, V.L.:

- 1976 Varianty fleksij *-ov* i \emptyset v roditel'nom padeže množestvennogo čisla suščestvitel'nyh mužskogo roda. In: L.P. Krysin, D.N. Šmelev (red.), *Social'no-lingvističeskie issledovanija*, Moskva, 129–144.
- 1977 Normy udarenija i statistika. In: *Jazykovaja norma i statistika*, Moskva, 219–244.

Vyčislitel'naja lingvistika

- 1976 Vyčislitel'naja lingvistika. Moskva: Nauka.

Wall, R.E.:

- 1958 *Statistical Aspects of the Scientific Russian Word Count*. Seattle.

Wedel, E.:

- 1993 Zur Poetik des Romantitels in der russischen Literatur des 19. Jahrhunderts. *Die Welt der Slaven XXXVIII*, 360–380.

Wenzel, Fr.:

- 1979 [Rez.] Častotnyj slovar' russkogo jazyka, pod red. L.N. Zazorinoj. Izd. Russkij jazyk, Moskva 1977. *Zeitschrift für Slawistik* 24, 297–301.

Wierzbicka, A.:

- 1966 K voprosu o porjadke slov v pol'skom i russkom stiche. In: *Poetics II*, Warszawa—The Hague, 345–369.

Winter, W.:

- 1964 Styles as Dialects. In: H.G. Lunt (ed.), *Proc. of the Ninth Int. Congr. of Ling.*, The Hague, 324–330 [repr. in Doležel/Bailey 1968, 3–9].

Yokoyama, O.T.:

- 1986 Lexical Frequency and its Implications: The Case of Contemporary edited Russian. *The Slavic and East European Journal* 30, 2, 147–166.
- 1987 High Frequency Vocabulary in Russian and in American English: A Sociolinguistic Comparison. In: A.L. Crone, C.V. Chvany (eds.), *New Studies in Russian Language and Literature*, Columbus/Ohio (1986/1987), 291–302.

Zacharova, V.P.:

- 1980 Verojatnostno-statističeskij analiz refleksov praslavjanskich DJ i TJ v Lavrent'evkoj letopisi. In: *Ėvoljucija i predystorija russkogo jazykovogo stroja*, Gor'kij, 105–109.
- 1981a Verojatnostno-statističeskij analiz refleksov praslavjanskich dj i tj v povestvovatel'nyh pis'mennyh tekstach nižegorodcev XIX-XX vv. In: *Problemy istorii kul'tury Volgo-Vjatsko-go regiona*, Gor'kij, 70–88.
- 1981b Verojatnostno-statističeskaja charakteristika refleksov praslavjanskogo v "Žitii" protopopa Avvakuma. In: *Ėvoljucija i predystorija russkogo jazykovogo stroja*, Gor'kij, 66–71.

Zaplatkina, N.I.:

- 1977 Obščee i različnoe v sistemach odnosložnych slov slavjanskich jazykov. In: *Strukturnaja i matematičeskaja lingvistika* 5, Moskva, 40–46.
- 1980 Tipologičeskoe issledovanie struktury odnosložnych slov v slavjanskich jazykach. Kiev.

Zasorina, L.N.:

- 1963 O statistike i avtomatizacii v leksikografii. *Naučnye doklady Vysšej školy, Filologičeskie nauki*, 4, 95–104.
- 1966 Avtomatizacija i statistika v leksikografii (Rabota nad častotnym slovarom russkogo jazyka). Leningrad: Izd. LGU.
- 1969 Častotnye i učebnye slovari. *Russkij jazyk v nacional'noj škole* 2, 20–25.
- 1977 Častotnyj slovar' sovremennogo russkogo jazyka. Moskva.

Zasorina, L.N., Zubkova, T.N., Kaplinskaja, K.P.:

- 1966 Častotnyj slovar' po slesarnomu delu. Leningrad.

Zil'bert, B.A.:

- 1970 Složnopodčinennye struktury v russkoj razgovornoj reči. In: *Russkaja razgovornaja reč'*, Saratov, 170–175.

Zinder, L.R.:

- 1957 O lingvističeskoj verojatnosti. In: *Voprosy statistiki reči*, Leningrad (1958), 58–61.
- 1958 O lingvističeskoj verojatnosti. *Voprosy jazykoznanija* 2, 121–125.

Zinin, S.I.:

- 1969 Struktura russkich antroponimov XVIII veka (na materiale aktovyh knig g. Moskvy). In: *Onomastika*, Moskva, 79–83.

Zipf, G.K.:

- 1949 Human behaviour and the principle of least effort. Cambridge.

Žuravlev, A.F.:

- 1967 Opyt verojatnostno-statističeskogo izučenija stilevyh različij. (Na materiale odnočlen-nyh složnopodčinennyh predloženij). In: *Jazyk i obščestvo*, Saratov, 87–111.
- 1970 O nekotoryh otičijach živoy razgovornoj reči ot stilizovanoj. In: *Russkaja razgovornaja reč'*, Saratov, 176–184.
- 1988 Leksikostatističeskaja ocenka genetičeskoj blizosti slavjanskich jazykov. *Voprosy jazykoznanija* 4, 37–51.
- 1992 Iz kvantitativnotipologičeskich nabljudenij nad leksikoj slavjanskich jazykov (Praslavjanskoe nasledie). *Voprosy jazykoznanija* 3, 106–118.
- 1993 Praslavjanskij slovník drevnenovgorodskogo dialekta s točki zrenija leksikostatistiki. *Voprosy jazykoznanija* 4, 87–98.
- 1994 Leksikostatističeskoe modelirovanie sistemy slavjanskogo jazykovogo rodstva. Moskva.

Žuravlev, A.P., Orlov, Ju.M., Brjakin, A.V.:

- 1968 Ėntropija elementov predloženija i problema zakončennosti. In: *Voprosy slavjanskogo jazykoznanija*, Saratov, 101–108.