



universität
wien

Emmerich Kelih

Institut für Slawistik

Quantitative Ansätze in der russischen Sprach- und Literaturwissenschaft: Die 20er und 30er Jahre

SCIENTIA QUANTITATIS

QUANTITATIVE LITERATURWISSENSCHAFT IN SYSTEMATISCHER UND HISTORISCHER PERSPEKTIVE

SCHLOSS HERRENHAUSEN, HANNOVER 30. SEPTEMBER - 02. OKTOBER 2014

Überblick – Einleitung

- Beginn der quantitativen Literaturwissenschaft im russischen Kontext
- Die Rolle von Andrej Belyj
- Statistische Studien im Nahbereich des Russischen Formalismus
- Von der Intuition zum System: B.I. Jarcho und seine exakte Literaturwissenschaft
- Zusammenfassung

Wann soll man die Geschichte beginnen lassen?

Prisma: Quantitative Methoden

- Umbruchzeiten Ende des 19. Jhd./ Anfang des 20. Jhd.
- Ausdifferenzierung von Philologien, Positionierung einer entstehenden Literaturwissenschaft, Sprachwissenschaft, Textwissenschaft?
- experimentelle Phonetik
- praktische Fragestellungen (Konstruktion von ökonomischen Alphabeten, Sprachdidaktik, Typographie, Stenografie)
- Häufigkeitwörterbücher – Brücke zur Untersuchung von Stilfragen
- quantitative Versanalyse
- Fortschritte der Statistik, Wahrscheinlichkeitstheorie, Mathematik hinsichtlich ihrer Methoden & zunehmendes Interesse an einer Erweiterung der Anwendungsgebiete

Fallbeispiel (1): Von den Anfängen der Stilometrie

Wissenschaftler, Kulturtheoretiker, Philosoph, Chemiker

N.A. Morozov (1854-1946)

Autor von:

Istorija čelovečeskoj kul'tury v estetstvenno-naučnom osveščanii // Die Geschichte der menschlichen Kultur im Lichte der Naturwissenschaften - verkürzt Christos // Christos

Einführung des Begriffs *stilometrija // Stilometrie* in Russland
(vgl. Morozov 1915 bzw. 1916)

(1915): „Lingvističeskie spektry: Spektry dlja otličeniija plagiatorov ot istinych proizvedenij togo ili inogo izvestnogo avtora“, in: *Izvestija otdelenija russkago jazyka i slovesnosti imperatorskoj akademii nauk* 20, 1/4; 95–127.

(1916): *Lingvističeskie spektry: Spektry dlja otličeniija plagiatorov ot istinych proizvedenij togo ili inogo izvestnogo avtora*. Petrograd: Tipografija Akademii Nauk.

(1998): *Bog i slovo*. Moskva: Kraft + Lean. [= *Istorija čelovečeskoj kul'tury v estetstvenno-naučnom osveščanii*, 3]



Konzeption & Ausrichtung:

- Verweis auf deutsche und polnische Vorläufer: (Th. Gomperz, W. Dittenberger, W. Lutosławski)
- linguistische Spektralverfahren in Analogie zur Chemie gesetzt
- Finden von stilometrischen Gesetzen, in Analogie zu Gesetzmäßigkeiten in der Natur und im gesellschaftlichen Leben
- **stabile quantitative Sprach- und Textmerkmale** auszumachen, die **aber gleichzeitig** Auskunft über die **Individualität eines Autors** geben
- Demonstration seines Verfahren anhand der Auszählung synsemantischer Wörter

inhaltliche Input:

- prozentuelle Anteil der Präpositionen *v* („in“), *na* („auf“) und *s* („mit“) in Texten von A.S. Puškin, L.N. Tolstoj und M.Ju. Lermontov
- Auszählung von Texten mit je 1000 Wörtern
- graphische Darstellung der Häufigkeiten in den Texten = „linguistische Spektren“

Ein prominenter Kritiker: A.A. Markov

Andrej A. Markov (1856-1922), seit 1886 Professor an der Universität Petersburg und 1896 Mitglied der Petersburger Akademie der Wissenschaften, Mathematiker & Wahrscheinlichkeitstheoretiker

(1913): „Primer statističeskogo izsledovanija nad tekstom ‘Evgenija Onegina’ illjustrirujuščij svjaz’ ispytanij v cep““, in: *Izvestija Imperatorskoj Akademii Nauk*, Serija VI, 7, 3; 153–162.

(1916): „Ob odnom primenenii statističeskogo metoda“, in: *Izvestija Imperatorskoj Akademii Nauk*, Serija VI, 10, 4; 239–242.

Kritik:

- ungenügende Stichprobenumfang der Analysen von N.A. Morozov
- Bedeutung des Gesetzes der Großen Zahlen

Und liefert gleichzeitig einen bedeutenden Beitrag:

Markov (1913)

- statistische Untersuchung an Puškins Versepos *Evgenij Onegin*
- analysiert die ersten 20.000 Buchstaben
- Aufeinanderfolgen von Vokalen und Konsonanten
- 200 Blöcke à 100 Buchstaben und ordnete diese zeilenweise fortlaufend in einer Matrix von 10 x 10 Buchstaben an.
- 200 Blöcke betrachtete er als unabhängige Variablen, wobei jede Variable als das Ergebnis von 100 abhängigen Versuchen anzusehen ist.

1	2	3	4	5	6	7	8	9	10
М	о	й	д	я	д	я	с	а	М
ы	х	ч	е	с	т	н	ы	х	п
р	а	в	и	л	к	о	г	д	а
н	е	в	ш	у	т	к	у	з	а
н	е	м	о	г	о	н	у	в	а
ж	а	т	с	е	б	я	з	а	с
т	а	в	и	л	и	л	у	ч	ш
е	в	ы	д	у	м	а	т	н	е
м	о	г	е	г	о	п	р	и	м
е	р	д	р	у	г	и	м	н	а
3	7	2	5	5	3	5	4	3	5

- Markov entwickelt einen theoretischen Erwartungswert für das Auftreten von Vokalen bzw. Konsonanten in bestimmter Position
- Nachweis gelungen, dass das Gesetz der großen Zahlen auch für abhängige Variablen gültig ist
- Wohlgeformtheit der Buchstabenketten

Zwischenresümee (1)

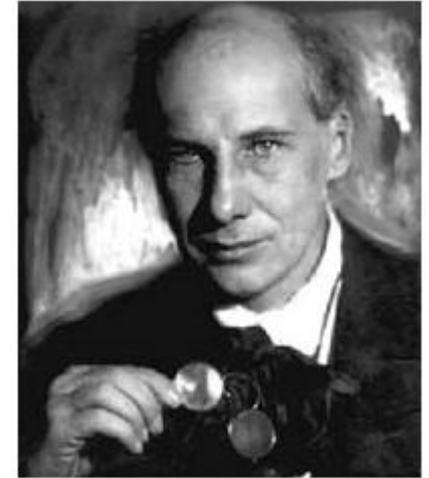
Mathematik/Statistik

Naturwissenschaft <----> Geisteswissenschaft

„Naturgeisteswissenschaft“

Die Rolle von Andrej Belyj (1880-1934)

- bedeutende russische Symbolist (Dichter, Theoretiker, Kulturphilosoph, Versspezialist u.v.m.)
- “positivistisches“ Elternhaus
- Einfluss auf den russischen Formalismus, insbesondere formale Versstudien
- Fokus auf **exakte Ästhetik**
- Mischung aus mystisch-philosophischer und naturwissenschaftlicher Anschauungen



Die Rolle von Andrej Belyj (1880-1934)

*Мой дядя самых честных правил,
Когда не в шутку занемог, [...]*
Mein Onkel, brav und hoch zu preisen,
scheint nun doch ernsthaft krank zu sein

- Pyrrhichien als „Beschleunigungsfaktor“
- wahrnehmungsästhetische Rolle
- Visualisierung
- statistische Auswertung
- rhythmische Kraft eines Textes (nicht Metrum)
- Rolle von Abweichungen

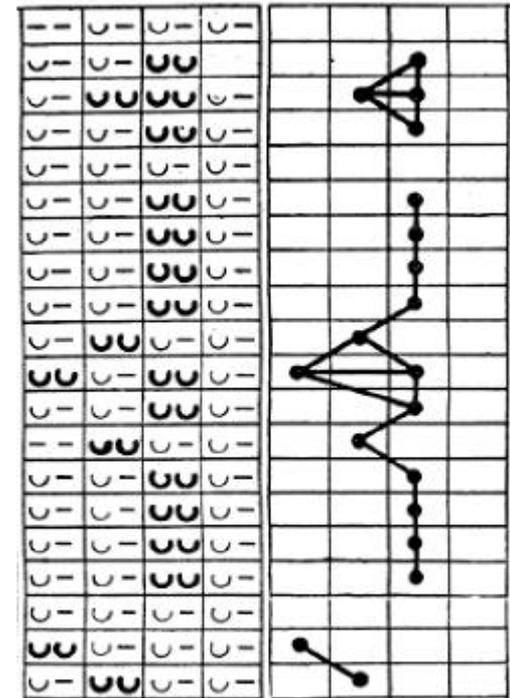


Abb. 3.1a: Visualisierung von Pyrrhichien nach Belyj (1910, 260)

Dach	umgekehrtes Dach	Quadrat	Leiter	Rhombus	Kreuz

Abb. 3.1b: „Figuren“ nach Belyj (1910)

kritische, aber breite Rezeption der Arbeiten von A. Belyj

- Breitenwirkung: „Rhythmischer Kreis“ wird eingerichtet
- viele Nachahmer unter Literaten gefunden (V.A. Čudovskij – logometrisches System, S.P. Bobrov)

Aber, eine durchaus treffende Einschätzung:

„[. . .] v oblasti stichovedenija rugat'sja rano - ešče ne sostavlena tablica umnoženija // [. . .] auf dem Gebiet der Verswissenschaft noch zu früh sei, zu polemisieren, da die Einmaleinstafel dafür noch nicht erstellt worden sei“ (vgl. Postoutenko 1990: 147).

durch **Boris V. Tomaševskij (1890-1957)**

Zwischenresümee (2)

Mathematik/Statistik <----> Literaten

Naturwissenschaft <----> Geisteswissenschaft

„Naturgeisteswissenschaft“

Statistische Studien im Nahbereich des Russischen Formalismus

- Nahverhältnis von formalistischer Verstheorie und statistischer Versanalyse
- nomothetische Charakter des Russischen Formalismus
- federführende Rolle hat **B.V. Tomaševskij**

Diese Methode, die in philologischen Kreisen auf Widerstand stößt [. . .] hat das Ziel, Fakten und Zusammenhänge aufzudecken, aber liefert, dies ist klar, niemals eine 'Erklärung'. Die Verwendung dieser Methode verlangt jedoch Vorsichtigkeit und eine kritische Herangehensweise, die jedoch nur Personen zugänglich ist, die mit der mathematischen Kultur vertraut sind. (Tomaševskij 1923c: 52).

Man soll die Philologie nicht der Statistik gegenüber stellen, da erstere durch das Objekt (Sprache) und zweitere durch die Methode der Konstatierung von Fakten bestimmt ist. Man darf in die Wissenschaft nicht das Verbot einer Methode zur Feststellung von Fakten einführen; dies würde nach Obskurantismus riechen. Die Zahlen, Formeln und Graphiken sind genauso Symbole des Denkens wie Wörter – sie sind nur jenen verständlich, die wissen, wie man mit diesem System von Symbolen umzugehen hat. Die Scheu der Philologen vor der Zahl ist hauptsächlich durch eine falsche Vorstellung der Rolle der Zahl in der wissenschaftlichen Arbeit bedingt. Es ist doch so, dass die Zahl nichts 'löst', d.h. sie interpretiert nicht, sondern dient ausschließlich als Mittel der Beobachtung und Beschreibung von Fakten. [. . .] Dass Zahlen und Graphiken missbraucht werden, ist keine Diffamierung der Methode selbst: Schuld ist derjenige, der diese Methode missbraucht hat und nicht das, was missbraucht wird. (Tomaševskij 1920/1929: 275f.).

Sonstige Verdienste von **B.V. Tomaševskij (gerafft)**

- Untersuchung des vier- und fünffüßigen Jambus
- Verssprache determiniert durch metrischen Reichtum einer Sprache und organisierende Kraft des Autors
- Berechnung der Wahrscheinlichkeit für das Auftreten einer jambischen Struktur
 - Einführung eines probabilistischen Modelldenkens in die Versanalyse
- Untersuchung des Prosarhythmus: Distanz der Betonungen (in Zusammenhang mit Mikrostruktur des poetischen Textes)
 - Einführung von Häufigkeitsverteilungen und deren theoretischer Modellierung (Spielarten der Binomialverteilung)
- Ausgewiesener Versexperte und gleichzeitig hohe Kompetenz in Statistik und Wahrscheinlichkeitstheorie

Zwischenresümee (3)

Mathematik/Statistik <----> Literaturwissenschaftler

[Naturwissenschaft] [Geisteswissenschaft]

~~„Naturgeisteswissenschaft“~~

B.I. Jarchos (1889-1942)
"exakte" Literaturwissenschaft: Kontext, Umfang und
heutige Relevanz



[...] mečtal o total'noj kvantivikacii filologičeskogo znanija // [...] träumte von der totalen Quantifikation des philologischen Wissens [...]
(Šapir 1990: 281)

"Yarkho's view is rather mechanistic and simplistic [...]"
(Margolin 1979: 7)

[...] raboty bolee pochoži na buchgalterskij otčet, čem na issledovanie o literature // seine Arbeiten ähneln mehr einer Buchhalterrechnung als einer Untersuchung über Literatur
(M.P. Štokmar in Jarcho (1997: 286))

- Biographische Einbettung
- Empirisch-statistische Literaturwissenschaft: Grundpositionen
- Wechselbeziehungen im literarischen Text
- Kompensationsmechanismen als textkonstitutives Merkmal?
- heutige Relevanz innerhalb der synergetischen Text- und Sprachanalyse

Biographische Einbettung

Arbeitsgebiete:

„Koe-komu ja izvesten, kak avtor toj ili inoj stat'i po ego spezial'nosti; a tak kak ja pisal po raznym „special'nostjam“ (slavistika, germanistika, lat(inskoe) srednevekov'e, drevnerusskaja l(iteratu)ra, fol'klor, romanistika), to do konca i ostalsja „avtorom stateek“: slavist ne znal menja, kak latinista. (Jarcho 2006: 4).

- Mitarbeit im M.L.K.
- ab 1922: RACHn, dann GACHN bis 1930
- Leiter der Sektionen für allgemeine Literaturwissenschaft, theoretische Poetik und künstlerische Übersetzung
- methodologisches Spezialgebiet: Anwendung statistischer Methoden in der Literaturwissenschaft

Synthese durch B.I. Jarcho: Quellen

- 1925: Granicy naučnogo literaturovedenija. Isskustvo. *Žurnal gosudarstvennoj akademii chudožestvennyh nauk*, Nr. 2, 45-60.
- 1927: Granicy naučnogo literaturovedenija. Isskustvo. *Žurnal gosudarstvennoj akademii chudožestvennyh nauk* 1927, tom III, Vyp. I; 16-38.
- 1969: Metodologija točnogo literaturovedenija (nabrosok plana). *Trudy po znakovym sistemam*, vyp. 4, 515-526.
- 1984: Metodologija točnogo literaturovedenija (nabrosok plana). *Kontekst*, 197-237.
- 2006: *Metodologija točnogo literaturovedenija. Izbrannye raboty po teorii literatury*. Moskva: Jazyki slavjanskich kul'tur. [1935]



Was ist exakte Literaturwissenschaft und was sind ihre Grenzen?

[...] "literaturovedenie est' izučenie slovesnych proizvedenij s točki zrenija chudožestvennosti, t.e. s točki zrenija éstetičeskikh élementov. // Literaturwissenschaft ist die Analyse künstlerischer Texte aus der Perspektive der Kunsthaftigkeit, d.h. aus der Perspektive der ästhetischen Elemente

Sovokupnost' élementov slovesnosti, sposobnych dejstvovat' na éstetičeskoe čuvstvo (v položitel'nuju ili otricateľnuju storonu) my nazываем chudožestvennoj formoj. Itak, literaturovedenie est' nauka, izučajuščaja literaturnuju formu. // Die Gesamtheit aller künstlerischen Elemente, die in der Lage sind auf das ästhetische Gefühl (in positiver oder negativer Weise) zu wirken, nennen wir künstlerische Form. Somit gilt, dass die Literaturwissenschaft eine Wissenschaft ist, die die literarische Form analysiert. // (Jarcho 1925: 47f.)

Aber welche Rolle nehmen statistische Methoden?

„[...] interes literaturoveda k proizvedeniju tem bol'she, čem mnogočislennee i raznoobraznee formal'nye élementy étogo proizvedenija [...]// das Interesse des Literaturwissenschaftlers an einem Werk ist umso größer, je mehr und vielfältiger die formalen Elemente des Werkes sind

[...] skol'ko" ili v kakom količestvennom sootnešenii? // [...] wieviel und in welchem quantitativem Verhältnis (ibid. 56)

→ je mehr formale Elemente, desto „literarischer“?

→ oder ist eine „bestimmte“ Häufigkeit gemeint?

Erwecken eines ästhetisches Gefühl: „Ungewöhnlichkeit“

- in Abhängigkeit von Betrachter
- im Vergleich zur gewöhnlichen Rede
- in Bezug auf die Proportion der formalen Elemente in einem literarischen Werk
- in Bezug auf die Kombination und Sukzessivität der Elemente
- und ob sie Teil von ungewöhnlichen **Wechselbeziehungen** sind !

Hilfswissenschaften der Literaturwissenschaft

1. Linguistik:

[...] ni odna statističeskaja veličina ne vvoditsja bez morfoložičeskogo analiza [...]

// [...] keine einzige statistische Größe wird ohne morphologische Analyse eingeführt [...] Jarcho (1984a: 198)

2. deskriptive Statistik

zuständig für eine exakte Analyse

„maksimal'noj ob'ektivnosti kategorii čisla // maximale Objektivität der Kategorie der Zahlen

2.1. Experiment

Welches Erkenntnisziel?

→ Entdecken von „Gesetzmäßigkeiten“ des literarischen Textes!

1. Ustanovlenie faktov // Bestimmung der Fakten (Deskription)
2. svjaz' priznakov // Wechselbeziehung der Merkmale

„tipičnie svjazi [...] nosjat nazvanie „zakonov prirody“. Zakon est' tipičeskaja, t.e. očen' často povtorjajuščajasja, svjaz' meždu javlenijami. // typische Wechselbeziehungen [...] tragen die Bezeichnung „Gesetze der Natur“. Ein Gesetz ist eine typische, d.h. sehr häufig sich wiederholende Beziehung zwischen den Phänomenen (Jarcho 1984: 211).

3. [...] „na toj stadii naučnosti, na kotoroj my sejčas nachodimsja, ... nam gorazdo važnee ustanovit' kak možno bol'se svjazej, čem objasnit' ich. [...] // zum gegenwärtigen Stand der Wissenschaft ... ist es weitaus wichtiger die Wechselbeziehungen festzustellen als diese zu erklären [...]

Literaturwissenschaft als nomothetische Wissenschaft:

- statistische „Gesetzmäßigkeit“ = Regularitäten
- induktive Finden von Zusammenhängen/Wechselbeziehungen!
- epistemologisch: Bewährung auf zeitlicher Achse + bestimmter Grad an Wahrscheinlichkeit !

Literarische Text:

- unterschiedliche Grade der Organisiertheit eines lit. Werkes
- kein einheitliches Prinzip
- **Wirken von Kompensationsmechanismen**

(1935): „Organische Struktur des russischen Schnaderhüpfels (Častuška). (Mit Ausblicken auf das deutsche Schnaderhüpfel)“, in: *Germanoslavica* 1/2, 31-64.

6000 Častuški

1000 spanische Cantes Populares

600 deutsche Schnaderhüpfel

2 Merkmale

Reim verbundene Zeilen

(Frequenz)

abcb

aabb

(aa)



Position der rhetorischen

Figuren + Frequenz

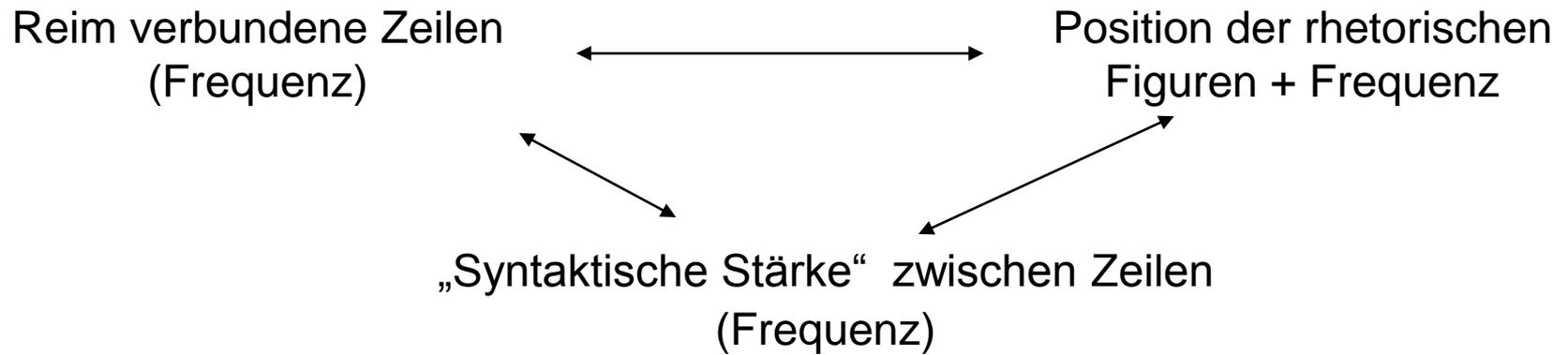
Anaphora

Epiphora

+

-

→ Wo die Hälften durch Reim verbunden sind, ist das „Figurenband“ schwächer (= weniger häufig vertreten).



→ Durch Reim verbundene Viertel (a-a, b-b) sind syntaktisch schwächer verbunden als die nicht reimenden.

Probleme:

- fehlende morphologische Analyse
- Bildung von (willkürlichen) Sub-Gruppen
- Verwendung einer Ordinalskala
- Zuordnung nicht nachvollziehbar
- Berechnung und Interpretation von Koeffizienten und Prozentzahlen

Ein tiefgreifendes organisches „Kompensationsgesetz“

Esli dva priznaka, nachodjaščiesja v otnošenii obratnoj koreljacii, vypolnjajut kakuju-nibud' odinakovuju funkciju, to, pri umen'shenii odnogo i odnovremennom roste drugogo, polučaetsja vpečatlenie, čto oni kak-by zamenjajut, kompensirujut drug druga pri vypolnenii obščej funkcii. // Wenn sich zwei Merkmale, die sich in einer gegenseitigen Wechselbeziehung befinden, die gleiche Funktion erfüllen, so, ergibt sich, dass bei der Verringerung der einen Funktion, sich die andere vergrößert. Damit ergibt sich der Eindruck, dass sich diese beiden gegenseitig ergänzen bei der Erfüllung ihrer allgemeinen Funktion ersetzen. (Jarcho 2006: 227)

- gegenseitige Regulation von Häufigkeiten!
- Wirken von Ausgleichsmechanismen!
- zeitgleiche Diskussion des Kompensationsprinzips in Biologie und Ökonomie
- Psycho-Biologie der Sprache von G.K. Zipf (1935)
- letztlich eine Brücke zu heutigen Ansätzen der synergetischen Linguistik

Zwischenresümee (4)

exakte Literaturwissenschaft

[Statistik, Linguistik als Hilfswissenschaften]

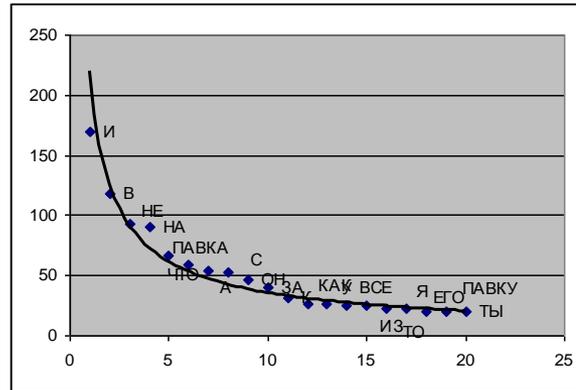
[[offenes Set an Wissenschaften]]

Zusammenfassung (mit Fragezeichen)

- kulturspezifische Spezifika
- Konstruktion – Rekonstruktion – ...
- Vielschichtigkeit
- Pluralität
- Verwischen der Grenzen, aber neues bewusstes Ziehen der Grenzen
- Kontinuität über die 20er und 30er Jahre hinaus
- 50er Jahre Verschiebung: Linguistik im Vordergrund

Exkurs: Gesetzmäßigkeit I: Zipf'sche Gesetz

N	Word	Freq.
1	И	169
2	В	118
3	НЕ	93
4	НА	90
5	ПАВКА	67
6	ЧТО	59
7	А	54
8	С	53
9	ОН	46
10	ЗА	40
11	К	32
12	КАК	26
13	У	26
14	ВСЕ	25
15	ИЗ	25



“Principle of least effort”

Sprecher



Hörer

Unifikation

Diversifikation

Dynamic equilibrium

Ausgewählte Literaturangaben:

Grzybek, Peter; Kelih, Emmerich (2005): Zur Vorgeschichte quantitativer Ansätze in der russischen Sprach- und Literaturwissenschaft. In: Köhler, Altmann, Piotrowski (Hg.): *Quantitative Linguistik. Quantitative Linguistics. Ein internationales Handbuch. An International Handbook*. Berlin, New York: de Gruyter (Handbücher zur Sprach- und Kommunikationswissenschaft, 27), S. 23–64.

Kelih, Emmerich; Grzybek, Peter (2005): Neuanfang und Etablierung quantitativer Verfahren in der sowjetischen Sprach- und Literaturwissenschaft (1956--1962). In: Köhler, Altmann, Piotrowski (Hg.), S. 65–82.

Kelih, Emmerich (2008): *Geschichte der Anwendung quantitativer Verfahren in der russischen Sprach- und Literaturwissenschaft*. Hamburg: Kovač (Studien zur Slavistik, 19).

Kelih, Emmerich (2013): B.I. Jarchos "exakte" Literaturwissenschaft: Kontext und Umfang. In: Aage A. Hansen-Löve, Brigitte Obermayr und Georg Witte (Hg.): *Form und Wirkung. Phänomenologische und empirische Kunstwissenschaft in der Sowjetunion der 1920er Jahre*. München: Fink, S. 411–426.

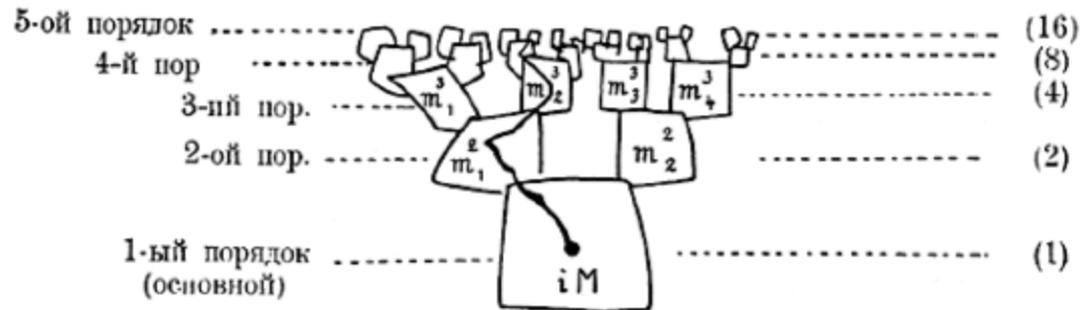


Abb. 3.3: Narrative Bifurkationen nach Bobrik (1921, 18)

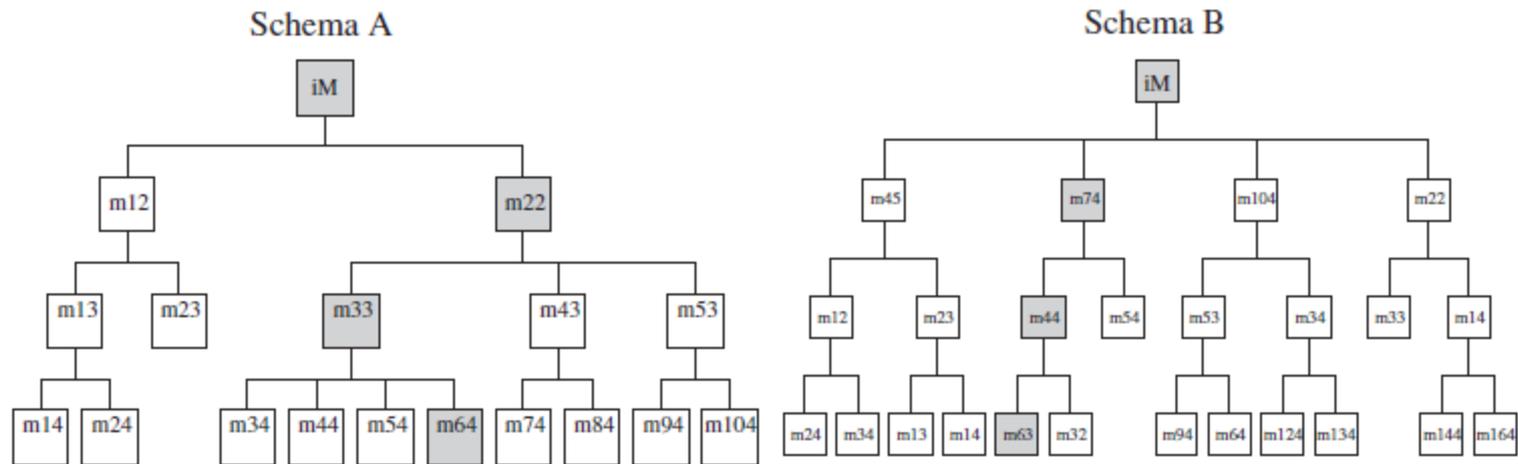


Abb. 3.4: Vergleichende Motivschemata nach Bobrik (1921, 24)

Zeitschrift *Glottology*

<http://www.degruyter.com/view/j/plot>

- erscheint 2 mal jährlich bei de Gruyter
- eds. Kelih/Köhler/Roelcke
- peer-reviewed
- Einreichungen in deutscher und englischer Sprache willkommen!
- Fokus: Sprach- und Literaturwissenschaft zwischen Quantität/Qualität
- Förderung des Dialoges zwischen quantitativen und qualitativen Ansätzen

