

LEXICAL NUCLEOLOGY IN THE DEVELOPMENT OF QUANTITATIVE STUDIES OF VOCABULARY

Asal Tagaeva Rustamovna

Lecturer, Department of Practical English

Navoi State University

Navoi, Uzbekistan

Tog'ayeva@nspi.uz

Abstract. *The article examines lexical nucleology as a historically grounded branch of quantitative vocabulary studies rather than an isolated recent proposal. Attention is given to the line that leads from early statistical observation of lexical behavior to parametric description of lexical systems and then to the study of lexical nuclei. The paper shows that nucleology grew out of measurable procedures for comparing dictionaries, lexical relations and central meanings across languages.*

Keywords. *quantitative linguistics, lexical nucleology, history of linguistics, vocabulary studies, lexical nucleus, parametric analysis, frequency, lexical systems, comparative lexicology, Slavic studies, language measurement.*

Introduction. The history of quantitative linguistics did not begin with digital methods. In the XX century, G.U.Yule linked literary vocabulary to statistical regularity, G.K.Zipf turned lexical distribution into a general problem of human economy, and G.Herdan attempted to organize scattered statistical observations about language into a unified theoretical system. Later surveys of the field stress that this line belongs to a much older scholarly tradition and that its continuity matters for present research. Lexical nucleology belongs to this historical line, but its object is narrower and more precise. It is concerned not

with vocabulary in general, but with the stable lexical center that can be identified through measurable parameters and then compared across languages. In the Slavic material, this approach was prepared by quantitative lexicology and then formulated as a distinct field in works by A.A.Kretov, I.A.Merkulova and V.T.Titov.

Methods and literature review. The study combines historical reconstruction, comparative analysis, bibliographic verification and close reading of programmatic texts. The historical part follows the movement from lexical counts and frequency regularities to mathematical modeling and then to parametric description of lexical systems. The comparative part focuses on how central meanings are selected, grouped, and interpreted in nucleological research. The analytical procedure is qualitative in exposition but rests on works that define quantitative linguistics as a discipline of measurable language relations.

The source base includes classical works by G.U.Yule, G.K.Zipf and G.Herdan, later generalizations by R.Köhler, G.Altmann, P.Meyer, O.Rottmann and K.H.Best and nucleological studies by A.A.Kretov, I.A.Merkulova and V.T.Titov. In this sequence, one can see three successive tasks. The first was to prove that lexical behavior is measurable. The second was to build explanatory models for recurrent regularities. The third was to move from general lexical measurement to the identification of lexical nuclei and key meanings within related languages.

Results. The first stage of the line under discussion centered on lexical quantity itself. Yule treated literary vocabulary as a domain where repetition and dispersion can be counted. Zipf moved further and connected lexical rank with a general principle of least effort. Herdan then expanded the field by integrating vocabulary statistics, style and broader questions of language structure. These authors did not yet speak about lexical nuclei in the later sense, but they created the methodological habit without which nucleology would

have been impossible. A second stage appeared when the field began to describe itself historically and theoretically. O.Rottmann writes that quantitative linguistics has a very old tradition beginning about 2500 years ago. K.H.Best adds that quantification and testing of hypotheses are documented in linguistics since the 17th century. The handbook edited by R.Köhler, G.Altmann and R.G.Piotrowski defines the field through scientific and technical approaches using quantitative methods for language and texts. G.Herdan, in turn, sought an organic whole for dispersed statistical work on language, while P.Meyer calls Zipf one of the founders of the contemporary methodology of QL. These brief formulations show that lexical nucleology entered a field that already possessed history, method and self reflection.

Period	Scholar or group	Main lexical object	Main result for later nucleology
1940s	G.U.Yule	literary vocabulary	proof that lexical behavior can be statistically described
1940s	G.K.Zipf	rank and frequency	regular relation between lexical frequency and ordering
1950s and 1960s	G.Herdan	vocabulary, style, language structure	attempt to unify lexical statistics into one theoretical system
2000s	R.Köhler, G.Altmann, R.G.Piotrowski and others	general quantitative linguistics	broad disciplinary framework for measurable language relations
2010s	A.A.Kretov, I.A.Merkulova, V.T.Titov	lexical systems of related languages	transfer from general lexical measurement to

			parametric lexical centers
late 2010s and 2020s	I.A.Merkulova	key meanings and lexical nuclei	explicit formulation of lexical nucleology

When the discussion reaches quantitative lexicology of Slavic languages, the object changes. Kretoy, Merkulova, and Titov describe the lexical system through functional, syntagmatic, paradigmatic and epidigmatic stratification and they compare dictionaries by dominant lexical relations. This step is decisive because it replaces isolated counts with a structured view of the lexicon. The nucleus is not simply the most frequent word set. It is the central zone where usage, combinability, semantic relations and derivational productivity converge. In this sense, lexical nucleology may be treated as the next historical phase of quantitative lexicology.

Merkulova formulates this phase directly. In one concise definition, lexical nucleology is a new section of lexicology. In a related study, she identifies the key concepts “hit”, “go”, “row”, “measure”, “have”, “circle” and relates them to large semantic domains shared within Slavic languages. These domains can be illustrated by stable lexical clusters such as strike, beat, blow, push for physical impact, go, come, move, walk for motion, have, keep, hold, own for possession, measure, count, size, number for ordering, circle, ring, round, wheel for closed shape and row, line, series, chain for linear arrangement.

Central meaning	English examples	Russian examples	Uzbek examples	Analytical comment
physical impact	hit, strike, beat, blow	ударить, бить, толкать	urmoq, tepmoq, turtmoq	dense action field with clear verb center

movement	go, come, walk, move	идти, ходить, прийти	bormoq, kelmoq, yurmoq	high lexical stability and broad combinability
possession	have, keep, hold, own	иметь, держать, владеть	ega bo‘lmoq, ushlamoq, saqlamoq	central field for social and material relations
ordering	measure, count, number, size	мерить, считать, число	o‘lchamoq, sanamoq, son	strong link between cognition and lexical classification
circularity	circle, ring, round, wheel	круг, кольцо, круглый	aylana, halqa, dumaloq	spatial pattern with rich nominal and adjectival extension
linear sequence	row, line, series, chain	ряд, линия, цепь	qator, chiziq, zanjir	useful for comparing nominal cores across languages

The historical value of nucleology lies in its change of scale. Earlier quantitative work often asked how many times a word form appears or how lexical richness varies between texts. Nucleology asks which meanings hold lexical systems together and by what measurable signs that centrality can be demonstrated. For comparative lexicology this is a stronger question. It allows the researcher to move from isolated lexical correspondences such as go and

идти, have and иметь, circle and круг, row and ряд, count and считать toward a structured account of lexical proximity between languages.

Discussion. The reviewed material shows that lexical nucleology did not emerge outside the main history of quantitative linguistics. It grew from the same insistence on measurable regularity that shaped lexical statistics, rank studies and mathematical modeling. What changed was the target of description. The focus shifted from counts and distributions alone to the lexical center of a language system and to the comparison of such centers across related languages.

This shift is productive for present research because it joins numerical procedure with semantic interpretation. In the cited nucleological works, the lexical nucleus is not reduced to a raw list. It becomes a structured set of meanings whose centrality can be argued through several parameters at once. That model is especially promising for comparative vocabulary studies, bilingual lexicography and typological work on closely related languages. It can also be adapted for English related research when the aim is to identify stable lexical centers rather than merely accumulate lexical counts.

Conclusion. Lexical nucleology should be read as a historically mature phase in quantitative studies of vocabulary. The route to it runs from Yule's statistical treatment of literary vocabulary through Zipf's lexical ordering and Herdan's system building to the later parametric studies of lexical structure. Within that route, the work of Kretov, Merkulova, and Titov marks the point where quantitative lexicology becomes the study of lexical nuclei and key meanings.

REFERENCES

1. Yule G. U. The Statistical Study of Literary Vocabulary / G. U. Yule. Cambridge, Cambridge University Press, 1944.
2. Zipf G. K. Human Behavior and the Principle of Least Effort / G. K. Zipf. Cambridge, MA, Addison Wesley Press, 1949.
3. Herdan G. Language as Choice and Chance / G. Herdan. Groningen, Noordhoff, 1956. XVI + 356 p.
4. Herdan G. The Advanced Theory of Language as Choice and Chance / G. Herdan. Berlin, Heidelberg, Springer, 1966. XV + 459 p.
5. Sobirovich A. S. COMPARATIVE FINE TUNING STRATEGIES FOR UZBEK AUTOMATIC SPEECH RECOGNITION ACROSS MULTILINGUAL FOUNDATION MODELS //CONFERENCE ON GLOBAL RESEARCH PERSPECTIVES. – 2026. – T. 4. – №. 1. – C. 145-154.
6. Sobirovich S. A. Robust Uzbek Asr AND Tts for Dialectal AND Noisy Settings //American Journal of Research in Humanities and Social Sciences. – 2025. – T. 40. – C. 1-4.