DOI: https://doi.org/10.37547/supsci-ojp-06-05-18 Pages: 167-174

Oriental Journal of Philology

ORIENTAL JOURNAL OF PHILOLOGY

journal homepage:
http://www.supportscience.uz/index.php/ojp/about

STATISTICAL ANALYSIS OF ROBOTICS TERMINOLOGY IN ENGLISH AND UZBEK

Gulmira Rasulova
PhD Researcher of Namangan State Institute of technical university
E-mail: rgulmira0509@gmail.com
Science ID: FNM-0926-0072
Namangan, Uzbekistan

ABOUT ARTICLE

Key words: robotics terminology, statistical analysis, terminological density, corpus linguistics, term, term component, lexical diversity, TTR, standardization, English, Uzbek.

Received: 10.09.26
Accepted: 11.09.26
Published: 12.09.26

Abstract: The article investigates the statistical characteristics of robotics terminology in English and Uzbek through a comparative corpus-based approach. The study aims to determine the quantitative distribution, terminological density, structural and component features, and degree of standardization of robotics terms in scientific and educational sources. The empirical material was extracted from works by J.J.Craig, D.Staple, N.B.Alimova, and M.W.Spong, S.Hutchinson and M.Vidyasagar. A total of 1,228 terminological units were identified in the four principal sources, of which 753 were classified as robotics-specific terms. In addition, a specialized research corpus containing 146 terms, a list of 170 ISO-related standardized terms, and a robotics glossary comprising 341 units were employed as control datasets.

INGLIZ VA O‘ZBEK TILLARIDAGI ROBOTOTEXNIKA TERMINOLOGIYASINING STATISTIK TAHLILI

Gulmira Rasulova
Namangan davlat texnika universiteti tayanch doktoranti
E-mail: rgulmira0509@gmail.com
Ilmiy identifikatsiya raqami: FNM-0926-0072
Namangan, O‘zbekiston

MAQOLA HAQIDA

Kalit so‘zlar: robototexnika terminologiyasi, statistik tahlil, terminologik

Annotatsiya: Mazkur maqolada ingliz va o‘zbek tillaridagi robototexnika sohasiga oid

zichlik, korpus lingvistikasi, termin, termin-komponent, leksik xilma-xillik, TTR, standartlashtirish, ingliz tili, oʻzbek tili.

terminlarning statistik xususiyatlari qiyosiy-korpus yondashuvi asosida tadqiq etiladi. Tadqiqotning asosiy maqsadi robototexnika terminologiyasining turli ilmiy va oʻquv manbalaridagi miqdoriy taqsimoti, terminologik zichligi, struktur-komponent xususiyatlari hamda standartlashuv darajasini aniqlashdan iborat. Empirik material sifatida J.J.Craig, D.Staple, N.B.Alimova hamda M.W.Spong, S.Hutchinson va M.Vidyasagar asarlaridan ajratib olingan terminologik birliklardan foydalanildi. Toʻrtta asosiy manbada jami 1228 termin aniqlanib, ulardan 753 tasi bevosita robototexnika sohasiga oid termin sifatida tasniflandi. Bundan tashqari, 146 termindan iborat maxsus ilmiy korpus, 170 termindan iborat ISO standartlari roʻyxati va 341 birlikdan iborat robototexnika glossariysi nazorat manbalari sifatida qoʻllanildi.

СТАТИСТИЧЕСКИЙ АНАЛИЗ РОБОТОТЕХНИЧЕСКОЙ ТЕРМИНОЛОГИИ В АНГЛИЙСКОМ И УЗБЕКСКОМ ЯЗЫКАХ

Гульмира Расулова
Докторант Наманганского государственного технического университета
E-mail: rgulmira0509@gmail.com
Научный идентификационный номер: FNM-0926-0072
Наманган, Узбекистан

О СТАТЬЕ

Ключевые слова: терминология робототехники, статистический анализ, терминологическая плотность, корпусная лингвистика, термин, термин-компонент, лексическое разнообразие, TTR, стандартизация, английский язык, узбекский язык.

Аннотация: В статье исследуются статистические особенности терминов робототехники в английском и узбекском языках на основе сравнительно-корпусного подхода. Цель исследования заключается в выявлении количественного распределения, терминологической плотности, структурно-компонентных характеристик и степени стандартизации робототехнической терминологии. Материалом послужили терминологические единицы из трудов J.J.Craig, D.Staple, Н.Б.Алимовой, M.W.Spong, S.Hutchinson и M.Vidyasagar. В четырех основных источниках выявлено 1228 терминологических единиц, из которых 753 классифицированы как непосредственно относящиеся к робототехнике. Дополнительно использованы специальный корпус из 146 терминов, список из 170 терминов стандартов ISO и глоссарий из 341 единицы.

Introduction. The rapid development of robotics has resulted in a considerable expansion of specialized terminology across engineering, computer science, artificial intelligence, mechatronics, automation, control theory and human-machine interaction. Contemporary robotic systems integrate mechanical structures, electronic devices, sensing technologies, control algorithms, programming environments and intelligent decision-making mechanisms. Consequently, robotics terminology represents a dynamically developing interdisciplinary lexical system rather than a closed collection of technical terms.

In contemporary terminology studies, the description of individual lexical units alone is insufficient for revealing the systemic properties of specialized vocabulary. Quantitative methods make it possible to establish the frequency, distribution, structural organization, semantic concentration and cross-linguistic correspondence of terms. Corpus-based statistical analysis therefore provides an objective basis for determining which lexical units constitute the terminological core of a discipline and how terminological systems vary across sources and languages.

Corpus linguistics has established an important methodological basis for such analysis. Sinclair's work on corpus, concordance and collocation demonstrates the value of recurrent contextual evidence in lexical research, while Biber, Conrad and Reppen show how quantitative corpus methods can reveal structural and functional variation across texts. These principles are especially relevant to terminology because specialized meaning is often realized through recurring multiword combinations rather than isolated lexical items.

The relevance of statistical approaches becomes particularly apparent in robotics. A term may be central in a theoretical textbook but relatively uncommon in programming-oriented literature. Similarly, a general scientific lexical item may acquire a highly specialized function when incorporated into a robotics-related multiword term. The lexical item force, for example, belongs to general physics; however, combinations such as force control, force sensor and contact force function as specialized robotics terms. The same principle applies to system, which forms terminological combinations such as control system, robot system, vision system and sensor system.

The purpose of this study is to statistically investigate English and Uzbek robotics terminology and identify quantitative tendencies in its distribution across sources of different functional orientations. The study establishes the number and proportion of robotics-specific terms, compares terminological density, describes methodological criteria for identifying specialized terms, and interprets the quantitative differences from a comparative linguistic perspective.

Literature review. Quantitative investigation of specialized vocabulary is closely associated with corpus linguistics. Sinclair argues for the systematic analysis of lexical units through concordance and collocational evidence. In terminology research, this approach allows a term to be evaluated in its recurrent professional context rather than as an isolated dictionary unit. Biber, Conrad and Reppen further developed corpus methodology by demonstrating how frequency distributions reveal functional and structural variation across texts. Their approach is relevant to robotics terminology because term frequency and density may vary significantly according to genre, communicative purpose and disciplinary specialization.

Terminological theory also emphasizes the relationship between specialized concepts and their linguistic designations. Felber's terminology manual provides a systematic framework for terminology work and standardization. From this perspective, terminological density can be treated as a useful quantitative indicator of the degree to which a text is concentrated around a specialized field.

Robotics provides a productive field for this type of analysis because its terminology reflects the interaction of several scientific domains. Craig's Introduction to Robotics: Mechanics and Control represents a major source of theoretical and applied robotics vocabulary, particularly terms associated with manipulators, kinematics, transformations, trajectories, control and mechanical movement.

Spong, Hutchinson and Vidyasagar's Robot Modeling and Control has a strong mathematical and theoretical orientation. Its vocabulary includes terminology related to manipulator dynamics, kinematic modelling, control theory and mathematical representation. Staple's Learn Robotics Programming, in contrast, combines robotics-specific vocabulary with programming, electronics, computing and practical implementation terminology. Alimova's Uzbek-language textbook represents an educational and technical source in which robotic and mechatronic devices, information systems, sensors and actuators are central.

Methodology. The study employs a comparative, corpus-based and quantitative research design. Quantitative analysis is combined with terminological classification and contextual interpretation to identify both numerical and functional properties of robotics terminology. The empirical database consists of four principal sources representing different functional orientations. Craig provides 329 identified technical terms, including 194 robotics-specific units; Staple contains 438 identified terms, of which 191 are robotics-specific; Alimova contains 201 terms, of which 186 are robotics-specific; and Spong et al. contains 260 terms, including 182 robotics-specific units. In total, 1,228 terminological units were examined, of which 753 were classified as robotics-specific.

In addition to these four sources, three supplementary datasets were used for methodological control and comparison: a specialized research corpus containing 146 robotics terms, a standardized list containing 170 ISO-related robotics terms, and a robotics terminology glossary containing 341 units.

The principal selection criterion was the concept of a robotics-specific term. A lexical unit was included when its meaning was directly associated with a robot device, robot system, manipulator, actuator, sensor, control mechanism, motion-planning process, kinematic or dynamic model, robot programming, human–robot interaction or robotic intelligence. General scientific terms from physics, mathematics, electronics, informatics and mechanics were included only when they performed a specialized terminological function within robotics discourse.

Terminological density was calculated according to the formula TD = (R/T) × 100, where TD is terminological density, R is the number of robotics-specific terms, and T is the total number of identified terminological units in the source. Component analysis treated independent lexical units as components of terms. Hyphenated forms denoting a single concept, such as end-effector, were treated consistently as terminological units. Abbreviations such as SLAM, HRI, PID, IMU and TCP were accepted as independent terminological units or components according to their function in the corpus.

Source	**Total terms**	**Robotics-specific**	**Density**	**Language**	**Orientation**
Craig (2004)	329	194	59.0%	English	Theoretical-applied
Staple (2021)	438	191	43.6%	English	Applied-programming
Alimova (2023)	201	186	92.5%	Uzbek	Educational-technical
Spong et al. (2006)	260	182	70.0%	English	Mathematical-theoretical
TOTAL	1,228	753	61.3%	2 languages	4 orientations

Table 1. Quantitative characteristics of the terminological datasets

Results. The analysis identified 1,228 terminological units across the four principal sources. Of these, 753 units met the established criteria for robotics-specific terminology. The overall average density of specialized robotics terms was therefore 61.3%. The distribution, however, differed substantially among the sources.

The highest terminological density was observed in N.B. Alimova's Uzbek-language textbook. Of 201 identified terms, 186 were classified as robotics-specific, producing a density of 92.5%. This result indicates that the source is highly concentrated on specialized technical concepts.

The second-highest value was recorded in Spong, Hutchinson and Vidyasagar's Robot Modeling and Control. Of 260 terminological units, 182 were classified as robotics-specific. The resulting density of 70.0% reflects the book's strong mathematical and theoretical specialization.

Craig's Introduction to Robotics: Mechanics and Control contained 329 identified terminological units, 194 of which were classified as robotics-specific. Its terminological density was therefore 59.0%. The result is consistent with the theoretical-applied nature of the source, which integrates specialized robotics concepts with terminology from mechanics, mathematics and control engineering.

The lowest density among the four sources was identified in Staple's Learn Robotics Programming. Although the source contained the largest total terminological inventory - 438 units - only 191 were classified as robotics-specific. Consequently, its terminological density was 43.6%. This result demonstrates that absolute term quantity and terminological specialization represent different statistical dimensions.

A comparison of the raw quantities provides further evidence. Craig contributes 194 robotics-specific units, Staple 191, Alimova 186 and Spong et al. 182. The difference between the largest and smallest specialized inventories is only 12 terms. Thus, despite considerable variation in the total number of identified technical terms, the number of specialized robotics units remains relatively stable across the four datasets.

The combined distribution of 753 robotics-specific terms is relatively balanced: Craig accounts for approximately 25.8%, Staple 25.4%, Alimova 24.7%, and Spong et al. 24.2%. This balance suggests that the selected sources provide comparable quantities of core terminological material despite differences in their overall lexical composition.

The supplementary datasets extend the comparison. The specialized research corpus contains 146 selected terms, the ISO-related list contains 170 units, and the robotics glossary contains 341 units. Their inclusion makes it possible to check terminology extracted from individual textbooks against broader specialized and standardized inventories.

Discussion. The quantitative differences identified in the study should primarily be interpreted in relation to source specialization rather than as a direct contrast between English and Uzbek as linguistic systems. The exceptionally high density of 92.5% in Alimova's Uzbek-language textbook reflects the highly specialized educational and technical character of the selected source.

Conversely, Staple's density of 43.6% should not be interpreted as terminological insufficiency. The work has a programming-oriented and practical character. Robot programming requires extensive terminology from software engineering, electronics, computer systems and information technology. Many of these units are essential to the subject matter but do not satisfy the narrower criterion of a robotics-specific term adopted in the present research.

The difference between Craig's 59.0% and Spong et al.'s 70.0% similarly reflects functional orientation. Craig combines theoretical foundations with practical mechanics and control applications, whereas Spong et al. emphasize mathematical modelling and robot control. Thus, terminological density is better understood as a measure of specialization than as a measure of the lexical capacity of a language.

From a linguistic perspective, multi-component structures are especially significant in English robotics terminology. A general lexical component can acquire specialized status through combination with another element. Examples such as force control, force sensor, control system, robot system and vision system demonstrate that terminological meaning is frequently distributed across the entire lexical construction rather than contained in one component.

The use of ISO-related terminology as a control dataset is important for standardization. Standardized terminological resources can serve as reference points for distinguishing stable internationally recognized terms from source-specific or variable forms. For Uzbek, this procedure can support the identification of cases in which several competing equivalents exist for one English concept.

From the perspective of corpus methodology, the findings confirm that absolute frequency should be combined with relative measures. If only the total number of terms were considered, Staple's source with 438 units would appear to be the most terminologically saturated. When robotics-specific density is calculated, however, Alimova's source becomes the most specialized at 92.5%. This illustrates why normalization and proportional analysis are essential in comparative terminology studies.

Conclusion. The statistical analysis of English and Uzbek robotics terminology demonstrates that quantitative methods provide an effective instrument for identifying structural and functional characteristics of specialized vocabulary. The four principal sources contained 1,228 terminological units, of which 753, or 61.3%, were classified as robotics-specific terms. Despite comparable numbers of specialized units in the individual sources - 194 in Craig, 191 in Staple, 186 in Alimova and 182 in Spong et al.- their terminological densities differed considerably. Alimova demonstrated the highest density at 92.5%, followed by Spong et al. at

70.0%, Craig at 59.0% and Staple at 43.6%. These differences primarily reflect functional orientation and degree of specialization rather than the quantitative capacity of the languages themselves.

The research also established that the identification of robotics-specific terminology requires contextual analysis. General scientific lexical items may acquire specialized terminological functions in multi-component structures. Consequently, component analysis represents an essential stage in the statistical description of robotics vocabulary. The use of a 146-term specialized research corpus, a 170-term ISO-related dataset and a 341-unit robotics glossary provides additional opportunities for verifying terminological coverage and examining standardization. The results can be applied to the standardization of Uzbek robotics terminology, technical translation, terminology teaching, corpus compilation, and the development of English–Uzbek specialized dictionaries and electronic terminological databases.

References:

1. Sinclair J. Corpus, Concordance, Collocation. Oxford: Oxford University Press, 1991. 179 p.

2. Biber D., Conrad S., Reppen R. Corpus Linguistics: Investigating Language Structure and Use. Cambridge: Cambridge University Press, 1998. 300 p.

3. Craig J.J. Introduction to Robotics: Mechanics and Control. 3rd ed. Pearson Education, 2004. 408 p.

4. Staple D. Learn Robotics Programming. 2nd ed. Birmingham: Packt Publishing, 2021. 602 p.

5. Alimova N.B. Mexatron modullar va robotlarning informatsion qurilmalari. Toshkent: Ziyo nashr-matbaa, 2023. 215 b.

6. Spong M.W., Hutchinson S., Vidyasagar M. Robot Modeling and Control. John Wiley & Sons, 2006. 496 p.

7. Felber H. Terminology Manual. Paris: UNESCO & Infoterm, 1984. 426 p.