

TERMS OF LENGTH MEASUREMENT IN THE RUSSIAN AND UZBEK LANGUAGES: A COMPARATIVE ANALYSIS ON THE MATERIAL OF LINGUISTIC CORPORA

Очилдиева Дилфуза Шодиевна

Преподаватель Джизакского государственного
педагогического университета

Abstract

The article discusses the terms of length measurement in the Russian and Uzbek languages in a comparative aspect based on the data of linguistic corpora, explanatory dictionaries and dictionary and corpus interfaces. The relevance of the study is due to the fact that the units of length measurement are not only a special terminological subsystem that serves scientific, technical and educational discourse, but also an important fragment of the linguistic picture of the world, in which cognitive, historical, cultural, translation and pragmatic parameters. The paper shows that at least two layers of vocabulary coexist in both languages: the modern metric, focused on standardized international scientific communication, and the traditional one, associated with anthropometric, everyday and historically fixed ways of measuring space. It has been established that in the Russian language, archaic units of length, such as arshin, vershok, span, sazhen and versta, in modern use much more often realize the phraseological, expressive and culturally marked potential, while in the Uzbek language the units qarich, quloch, qadam, gaz, chaqirim demonstrate a more lively connection with the figurative nomination, approximate assessment of distance and national-cultural semantics. The comparative analysis made it possible to identify both zones of direct equivalence, primarily in the field of international metric terminology, and zones of incomplete, functional or contextually conditioned correlation, especially in artistic, historical and colloquial discourses. It is concluded that the corpus approach provides a more accurate description of the lexical and semantic structure of terms of length, their combined models, genre distribution and degree of terminologization, as well as creates a reliable basis for bilingual lexicography, translation and teaching of the Russian and Uzbek languages.

Keywords: Terminology; units of length measurement; Russian language; Uzbek language; comparative analysis; corpus linguistics; linguistic corpus; metric system; archaic vocabulary; equivalence; translation; national-cultural semantics.

Introduction

The study of the terms of length measurement in the comparative aspect requires going beyond a narrow nominative description, since this group of words is located at the intersection of terminology, historical lexicology, ethnolinguistics, cognitive semantics and translation theory. Formally, units of length can be perceived as a purely applied set of designations intended for the quantitative characterization of space, but the real functioning of these units in the language is significant more complicated: they participate in the formation of naïve and scientific pictures of the world, preserve traces of historical methods of measurement, are included in figurative and phraseological constructions, serve technical description, and also act as markers of cultural memory. That is why the topic does not tolerate a superficial approach: if you limit yourself to the school list "meter - centimeter - kilometer" or its Uzbek equivalents, then the main thing disappears - the way language stores different models of correlating a person with space. In the Russian language, modern metric nomenclature coexists with historical measures such as arshin, span, vershok, sazhen, versta; In the Uzbek language, next to the international designations METR, santimetr, millimetr, kilometr, the words qarich, quloch, qadam, gaz, chaqirim and other units that arose on an anthropometric or household basis continue to be significant. A comparison of these layers shows that the semantic organization of the terms of length in the two languages is not mirrored: even where an approximate correspondence is possible, there is no complete coincidence of cultural volume, stylistic coloring and frequency distribution. For example, the Russian arshin and the Uzbek gaz can be close in the historical and everyday dimension, but their connotative field, the degree of phraseologization and modern recognizability differ; span and qarich are similar in anthropometric base, but are not identical in the degree of idiomatization and the set of stable contexts; Versta and Chaqirim have the function of indicating a long distance, but historical values and associative series diverge. The theoretical significance of the topic is determined by the fact that it is precisely on such seemingly "simple" lexemes that the



dynamics of the transition from the pre-scientific dimension associated with the human body and direct experience to scientific standardization and internationalization of terminology are especially clearly manifested. The practical significance is obvious for bilingual lexicography, translation of literary and historical texts, teaching Russian as a foreign language and Uzbek as a native or non-native language, as well as for the preparation of special educational materials that require a precise distinction between terminological, everyday and figurative use. The purpose of the article is to identify the general and specific features of the functioning of length measurement terms in the Russian and Uzbek languages on the basis of corpus data. To achieve this goal, the following tasks are solved: to determine the composition of the main thematic subgroups of length terms in both languages; to establish the ratio of the modern metric and historical-traditional layers; to describe the features of their lexical and semantic structure, compatibility and genre distribution; to identify the types of interlingual equivalence and non-equivalence; to show how corpus data refine the idea of the real use of these units. The scientific novelty of the work lies in the fact that the terms of length are considered not as a static dictionary list, but as a mobile system, where each sign manifests itself in different modes of functioning: terminological, common language, historical, metaphorical and phraseological. In addition, it is fundamentally important to use the corpus approach, which allows you to rely not on the researcher's intuition, but on the real contexts of written and, in some cases, oral speech. For the Russian language, this opportunity is provided primarily by the Russian National Corpus, which is a representative collection of texts with a volume of over 2 billion word usages and includes main, media, oral, parallel, educational and other corpora. For the Uzbek language, an important empirical support is the Uzbek Electronic Corpus, where search by token, lemma, concordance, morph analyzer, parallel and educational corpus, as well as thesaurus and dictionary modules are available. As a result, the comparison acquires not a speculative, but a verifiable character: one can see where the unit really lives in the modern language, and where it is preserved mainly as a cultural trace, idiom or historicism.

Materials and methods of research

The material of the study consists of the data of the Russian National Corpus, interfaces and dictionary sections Uzbek Electronic Corpus and O‘zbek tilining



ta'limiy korpusi, as well as explanatory dictionary sources that allow correlating the terminological meaning of a unit with its usual and phraseological potential. It is fundamentally important to emphasize that the corpus method in this work is not replaced by a mechanical calculation of individual word forms: such an approach would be methodologically weak, because it would not take into account the differences between direct terminological use and secondary, figurative, phraseological or quotation implementation of the lexeme. Therefore, the study is based on a combination of several procedures. First, the method of continuous selection of key units from two thematic series is used: a) units of the international metric system, for example, meter, millimeter, centimeter, decimeter, kilometer and their Uzbek equivalents metr, millimetr, santimetr, detsimetr, kilometr; b) traditional and historical units related to the anthropometric and everyday dimension of space, for example, arshin, span, vershok, sazhen, verst, cubit, as well as qarich, quloch, qadam, gaz, chaqirim. Secondly, a component analysis of dictionary definitions is used, which makes it possible to distinguish obligatory and optional semes: 'unit of length', 'accuracy/approximation', 'anthropometric base', 'historical fixation', 'belonging to the metric system', 'archaism', 'figurative meaning', 'phraseological activity'. For example, the dictionary data of the Russian language record arshin as an ancient measure of length of about 0.71 m, vershok as 1/16 of an arshin or about 4.4 cm, sazhen as about 2.134 m, span as the distance between the outstretched fingers of the hand; Uzbek dictionary and corpus sources define gaz as an obsolete unit of length equal to about 0.71 m, and chaqirim as an ancient unit of length of about 1.06 km. For metric terms, quantitative constructions, numerical parameters, abbreviations, technical descriptions, designations of dimensions, drawings, construction, educational and scientific texts are significant; Historical narratives, literary texts, proverbs, fixed expressions and evaluative metaphors are important for traditional units. Fourth, a comparative method is carried out that makes it possible to distinguish between full equivalence, partial equivalence, functional analogy and non-equivalence. At the same time, full equivalence is understood as cases of international metric terminology, where the differences relate mainly to graphics and phonetics, and not to conceptual content; partial equivalence refers to cases when units are close in denotative function, but differ in historical magnitude, cultural fixation, or compatibility; functional analogy refers to cases of translation by role in the text, and not by

exact quantitative coincidence; Non-equivalence refers to cases when a unit requires commentary, tracing, descriptive translation, or preservation of national and cultural color. In addition, the cognitive-semantic aspect is taken into account: the models underlying the naming of length are investigated, primarily the model "body as a measure" and the model "standardized abstract unit". The corpus base for the Russian language is especially reliable in this regard, since the RNC combines the main corpus, media subcorpora, oral, parallel and educational segments; For the Uzbek language, the architecture of the electronic corpus itself is essential, including search by token and lemma, concordance, morph analyzer, parallel and educational modules. Such a set of tools makes it possible not to replace analysis with guesses and not to mechanically transfer the meanings of one language unit to another.

Results of the study

The analysis shows that the system of terms for measuring length in the Russian and Uzbek languages is organized according to a similar general principle, but differs in internal hierarchy, cultural load and the degree of preservation of the archaic layer. The first, most obvious result is that in both language systems, the core of the modern terminological field is formed by international metric units. Millimetr, santimetr, detsimetr, kilometr coincide in conceptual basis, serve educational, engineering, scientific and official business discourse and have a high degree of terminological stability. Their corpus behavior is characterized by regular combination with numerals, graphic abbreviations, parametric designations of width, length, height, depth, radius, diameter, distance, thickness, etc. In both languages, such units demonstrate minimal imagery and maximum denotative transparency: here the term works as a tool for the accurate transmission of a parameter, and not as a carrier of a cultural image. However, already at the next level, a discrepancy is revealed: the Russian system of traditional measures of length in the modern language has largely moved to the sphere of cultural memory, art-historical discourse and phraseology, while in the Uzbek language a number of traditional units retain greater semantic transparency and are more naturally involved in the colloquial or figurative-everyday assessment of space. For the Russian language, arshin, span, vershok, sazhen and verst are not just old units of measurement, but culturally loaded signs. An arshin, according to dictionary data, denotes an ancient measure of length of about 0.71



m, equal to 16 vershkas, but in modern linguistic practice it actively lives in expressions such as measure everyone by one arshin or as if a yard has swallowed, where the quantitative meaning is either completely erased or recedes into the background. A span retains a direct meaning as an anthropometric measure, but it is especially stable in the idioms ni span of earth, seven spans in the forehead, span by span, where physical magnitude is reinterpreted as a sign of intellectual strength, perseverance, consistent advancement or spatial minimum. A vershok equal to about 4.4 cm is more often recognized by a native speaker through a historical and everyday context or phraseological reminiscence than as a actually used measure. A sazhen, according to dictionary data, about 2.134 m, has a particularly strong figurative productivity in the formula oblique sazhen in the shoulders, where the quantitative measurement of the body turns into a stereotype of a massive figure and physical strength. Thus, the Russian historical series today gravitates towards two zones of functioning: a) historical description and b) phraseological symbolization. In the Uzbek language, the traditional series demonstrates a different configuration. Dictionary and corpus data show that gaz is defined as an obsolete unit of length, equal to about 0.71 m, that is, in terms of denotative value close to the Russian arshin; Chaqirim is described as an ancient unit of length of about 1.06 km and at the same time is close in the usual sense to the designation "kilometer"; qarich and quloch are included in a number of long measures that preserve bodily visual motivation; qadam combines the value of a step unit and the basis for figurative expressions. The second result of the study is especially important here: for the Uzbek language, the relationship between the body image and the unit of length in modern linguistic intuition often remains more transparent. Qarich is read as a "swing" of the fingers, quloch as a "swing of the hands", qadam as a step, and therefore the transfer to evaluative, approximate, spatial and everyday use turns out to be less sharp and less "museum-like" than in Russian. This can be seen from the dictionary and corpus data, where not only direct designations are recorded, but also such phraseologized contexts as burnini bir qarich ko'tarmoq, qadam o'lchab bosmoq, yuz chaqirim nariga qochmoq, as well as uses such as bir quloch, bir yarim quloch in spatial description. The third result concerns the structure of interlingual correspondences. Complete equivalence is characteristic almost exclusively of metric terminology: metre is metr, centimeter is santimetr, millimetre is millimetr, kilometr is kilometr. Here the translator does not work with a cultural

problem, but with graphic and orthographic adaptation. Partial equivalence is observed in pairs of arshins — gaz, span — qarich, verst/large traditional distance measure — chaqirim, where convergence is possible, but should not replace the difference in historical volume and stylistic marking. For example, arshin and gaz are close as pre-metric units and in size, but the Russian sign is deeply embedded in the phraseology of the Russian classics, while the Uzbek gaz in modern use is more often recognized as historicism or a vocabulary unit with a local cultural background. Span and qarich are similar in anthropometric model, but the Russian span is more "frozen" in well-known stable formulas, while qarich has a more natural connection with the everyday assessment of a small space. A functional analogy arises when direct translation destroys the cultural fabric of the text: in a historical novel, the preservation of arshins, versts, gaz, chaqirim may be preferable to a formal conversion into meters, if the task is to convey the epoch rather than technical accuracy. The fourth result is related to the genre distribution. Metric units in both languages dominate scientific, educational, technical, administrative, and informational discourse; Traditional units are concentrated in a literary text, historical narration, phraseology, ethnographic description, and colloquial-evaluative speech. In other words, the modern terminological center of both systems is equally standardized, but their cultural periphery is different. The fifth result is that the corpus approach makes it possible to see not just the presence of a word, but the mode of its life in the language. Without a corpus, a researcher easily makes a typical mistake: he believes that if a word is recorded in the dictionary, it means that it is equally accessible and equally active for a native speaker. In fact, this is not the case. Some units live as full-fledged instruments of measurement, others as cultural shadows, others as elements of stable imagery, and still others as historical marks of the era. That is why the comparison of Russian and Uzbek material cannot be reduced to a simple selection of "dictionary analogues": real equivalence is revealed only in the context, genre, pragmatics and frequency profile of use.

Discussion

The results obtained allow us to move from the description of individual units to a more rigid theoretical conclusion: the terms of length measurement in the Russian and Uzbek languages cannot be considered as a homogeneous class of words, because within this class there are different mechanisms of nomination,



standardization, semantic erosion and cultural reconstruction. The weakest strategy in such studies is to mechanically mix modern terminology and historical vocabulary in one row. without distinguishing their communicative status. If this is done, then the Russian centimeter and the Russian sazhen will turn out to be "equally units of length", although in practice the former functions as a full-fledged topical term, and the latter is mainly a historicism and a phraseological sign; Similarly, the Uzbek kilometr and qarich cannot be described in the same language without reservation, since the former serves precise standardized communication, while the latter is associated with bodily motivated approximate evaluation and cultural imagery. From this point of view, a comparison of Russian and Uzbek material shows not only coincidences, but also deep differences in the type of linguistic memory. The Russian system of old measures of length is intensively literary: its modern recognizability is largely ensured by school reading, classical texts, historical narration and stable phraseology. Hence the high symbolic density of the words arshin, span, sazhen, vershok: the bearer may never measure space with them, but he still recognizes them as signs of the cultural code. In the Uzbek system, traditional units of length are also culturally marked, but the anthropometric basis of the qarich-quloch-qadam series remains more intuitively accessible, making it easier to reuse them for approximate spatial estimation. This distinction is important for translation. A translator who strives for a rough quantitative correspondence runs the risk of destroying two types of meanings: either to erase the cultural flavor by replacing the historical measure with a neutral metric equivalent, or, on the contrary, to create a false exoticism if he begins to substitute a nationally colored unit where dry technical accuracy is required in the original. Therefore, the choice of a translation strategy should be determined not by "what one is equal to in meters", but by why it is used in the text. If we are talking about a drawing, engineering instructions, geometry textbook, construction specification, medical description of dimensions or other special text, strict metric normalization and terminological unambiguity have priority. If a unit of length is part of the fabric of a literary text, historical reconstruction, folklore description or idiom, its literal recalculation often turns out to be stylistically deadly: seven spans in the forehead cannot be translated as "a lot of centimeters between the fingers", just as burnini bir qarich ko'tarmoq cannot be neutrally smoothed down to "show arrogance" if the task is to preserve the national speech texture. Here the question arises about the degree of



permissible adaptation. A four-way approach seems to be optimal: with complete terminological equivalence, use a direct correspondence; in case of partial analysis, choose the nearest functional analogue, taking into account the genre; in culturally significant historicisms, keep the original unit with a commentary or contextual prompt; in phraseological use – to give priority to the meaning and expression, and not to the literal measure. The comparative analysis also shows that in both languages, units of length are involved in the conceptualization of space through two basic cognitive models. The first model, "the body as the measure of the world", is especially clearly represented in the span, qarich, quloch, qadam, and historically in a number of other units; The second model, the "standard as an abstract scale," is realized in the meter and derived units. These models do not just coexist, but compete in the mind of the wearer. Where accurate knowledge is needed, the standard dominates; Where a quick everyday grasp of scale, emotional evaluation or cultural memory is needed, language again resorts to corporeal, figurative, historically fixed measure. This explains the stability of archaic units in language even after they have completely lost their measuring function. Consequently, from the point of view of terminology, we are not dealing with "dead words", but with units that have changed their functional regime. For Russian and Uzbek lexicography, a practical conclusion follows from this: a dictionary entry should record not only the meaning of a unit in the system of measures, but also its frequency profile, genre status, phraseological connections, translation risks and cultural marking. There is also a direct consequence for the teaching methodology: it is pointless to teach these words as a simple set of historicisms; It is more productive to show them through contexts, comparisons, and corpus examples, otherwise the student gets a dead list instead of a working language model. Finally, for corpus linguistics, this study confirms a simple but fundamental thing: even a limited set of lexemes reveals the complex architecture of a language if analyzed not in the abstract, but in a real distribution by genres, meanings, and pragmatic tasks.

Conclusion

A comparative analysis of the terms of length measurement in the Russian and Uzbek languages based on the material of linguistic corpora showed that both language systems combine a universal metric layer and a nationally marked traditional layer, but the ratio of these layers is different. In the field of

international metric terminology, there is a high degree of interlingual equivalence, which ensures accurate translation and standardized use in scientific, technical and educational discourse. There is no complete coincidence of historical and anthropometric units: the Russian arshin, span, vershok, sazhen, verst and the Uzbek gaz, qarich, quloch, qadam, chaqirim are only partially related to each other, since they differ not only in size, but also in the degree of archaism, cultural recognition, phraseological activity and genre profile. It has been established that in the Russian language the old measures of length are especially actively preserved in fixed expressions and literary-historical code, while in the Uzbek language a number of traditional units remain more directly related to the bodily motivated and approximate assessment of space. The corpus approach made it possible to clarify the real functioning of the units under study and to avoid a false equality between dictionary fixation and live usage. The practical value of the results lies in the possibility of their application in bilingual lexicography, literary and scientific translation, teaching Russian and Uzbek languages, as well as in the development of educational materials on terminology, lexicology and corpus linguistics. Prospects for further research are associated with the expansion of the corpus of material through parallel corpora, quantitative analysis of frequency, the study of translation correspondences in specific genres, and the use of data from other Turkic and Slavic languages for typological comparison.

References

1. The National Corpus of the Russian Language. Official website.
2. The National Corpus of the Russian Language. Corpus statistics.
3. The National Corpus of the Russian Language. News archive.
4. Gramota.ru. Dictionary entry: **arshin**.
5. Gramota.ru. Dictionary entry: **span**.
6. Gramota.ru. Dictionary entry: **vershok**.
7. Gramota.ru. Dictionary entry: **sazhen**.
8. Uzbek Electronic Corpus. Official interface: search by token, lemma, concordance, morpho analyzer, parallel corpus, educational corpus.
9. O‘zbek tilining ta’limiy korpusi. Vocabulary: **gaz**, **chaqirim**, series of length units and corpus examples.



10. Abjalova M., Adali E., Iskandarov O. *Educational Corpus of the Uzbek Language and its Opportunities.*
11. Karimov S. *A national corpus of the Uzbek language will be created...*
12. Lyashevskaya O. N., Sharov S. A. *Frequency Dictionary of the Modern Russian Language (Based on the Materials of the National Corpus of the Russian Language).* Moscow, Azbukovnik Publ., 2009.