



Section 3. General questions of philology and linguistics

DOI:10.29013/EJLL-25-1.2-20-26



ABOUT THE MOST RECENT PROBLEMS OF LEXICOGRAPHY

*Parvana Pashayeva Ismayil*¹, *Abbas Abbasov Gurban*¹

¹ Baku State University

Cite: Pashayeva P. I., Abbasov A. G. (2025). *About the Most Recent Problems of Lexicography*. *European Journal of Literature and Linguistics* 2025, No 1–2. <https://doi.org/10.29013/EJLL-25-1.2-20-26>

Abstract

This article investigates the theoretical dimensions of digital technologies in reshaping modern lexicography, focusing on their capacity to create adaptive, virtual lexicographic frameworks. It proposes a novel conceptual model for standardizing linguistic resources across cognate languages, with a specific emphasis on the Turkic languages. By analyzing the interplay of normalization, terminological unification, and multilingual lexicography, the study advances a theoretical foundation for linguistic convergence, contributing to the broader discourse on language standardization and digital lexicographic innovation.

Keywords: *lexicographic means, information technologies, terminology, word-term of course, orthography, orthoepy, rhetoric, language teaching methodology, Multiparametric city*

Introduction

Scientific importance. Dictionaries and reference works must evolve to meet diverse modern needs, prompting the question of what lexicographic tools are best suited for today's challenges. Electronic dictionaries, unlike their static predecessors, are dynamic and virtual, existing only when accessed, as Schryver, (2003). notes: they “needn't really be there anymore.” Atkins (1996) further describes this as their “virtual” quality, enabling real-time adaptability. This transformation allows lexicographers to rethink data storage, processing, and accessibility, foster-

ing innovative applications such as shared lexicographic systems for related languages like Turkic languages. By leveraging these capabilities, electronic lexicography not only addresses current demands but also paves the way for standardized, interconnected linguistic resources.

1. Normalization, unification and standardization in lexicography

Nowadays, research on four interdependent issues assumes particular relevance. These issues are as follows (Tarp, S., & Gouws, R. H., 2008).

1. Lexicography and terminography;
2. Averaging of terms;
3. Multilingual dictionaries;
4. Establishing the role of dictionaries within the national language corpus (language groups).

Undoubtedly, both evident and subtle differences exist between these areas in terms of their scope, complexity, and development. The comparison between the concepts of lexicography and terminography arises from the contrast between “word” and “term.” Previously, the development of theoretical and practical aspects of different types of dictionaries was referred to as *lexicography* (Gorodetsky B. Y., 1983).

Modern lexicography is not limited to theoretical studies on dictionary compilation and principles of dictionary development. It also includes the practical publication of various types of dictionaries to meet the demand for linguistic resources. Modern lexicographic tools must adhere to normalization requirements, at least at the orthographic level. That is, in a modern dictionary, a word must be presented according to the orthographic norms of its respective language. Thus, every lexicographic resource fulfills an additional function in maintaining language standardization (Swanepoel, P., 2006).

The results of lexicography are widely applied in areas such as orthography, orthoepy, rhetoric, and language teaching methodology. Lexicography is an integral part of linguistics. Scientific-technical lexicography focuses on the theory and practice of creating specialized (terminological) dictionaries. These dictionaries define and classify the nomenclature of various scientific disciplines and their associated terms. A. S. Gerd refers to scientific-technical lexicography as the initial branch of applied terminology (Gerd A. S., 1986, p. 24). At the same time, V. M. Leichic (1996), using the term “applied terminology,” pointed out that “terminography” is derived from the phrase “terminological lexicography.”

Lexicography is closely linked with lexicology, benefiting from its research findings while addressing scientific and practical issues. Conversely, lexicography supplies extensive material for lexicological study. Similar relationships exist between terminology and

terminography. In terminology, the selection, organization, and analysis of terms ultimately lead to their inclusion in dictionaries. This means that modern terminological dictionaries can contribute to the organization, unification, and standardization of terms. Terminography thus has the potential to evolve terminological dictionaries into standardized term dictionaries (Hjørland, B., 2023).

In fact, terminography emerged as a subfield of lexicology (Alberts, 2001). From the same perspective, terminology originated within lexicology before becoming an independent field. In solving terminographical problems, issues such as synonymy, homonymy, polysemy, conceptual equivalence, and definition point back to terminological studies. Therefore, there are no obstacles to regarding terminography as a subsection of terminology. However, in the modern era, the trend toward differentiation and integration in every scientific field is a recognized reality.

One of the increasingly urgent problems in modern linguistics is the issue of *averaging*, especially in the context of convergence among cognate languages. Moreover, averaging plays a significant role in the development of formal languages. At present, efforts and studies are ongoing in both directions. The process of averaging among cognate languages initially appears straightforward due to the existence of common units and components at various levels. These common elements serve as reference points in selecting and defining other tools for averaging.

It is essential to identify and document these shared units, study them comparatively, and determine their scope. In other words, it is necessary to clarify which language element is common or similar to the element of another language. This involves analyzing the identity and similarity between elements A in one language and B in another, then extracting the shared elements. If $A = B$ reflects mathematical identity, in linguistics, identity and equivalence are viewed from multiple angles – graphical, phonetic, semantic, and distributive. Hence, comparison and identification in linguistics are characterized by their multiparametric nature.

Currently, the problem of a *common language* receives particular attention in linguistics. While the concept itself is not new,

several languages are used as common or lingua franca in different contexts. Globally, English serves as the primary common language, and its domains of use continue to expand. Geographically, other languages may also function as regional lingua francas, depending on the language's prevalence and influence in that area. The integration of speakers of different native languages within a state often leads to the emergence of a regional common language. Historically, the language of a dominant state has often functioned as the regional lingua franca in political systems – for example, French, Spanish, Portuguese, and Russian (May, S., 2014).

2. Lexicographical issues of common Turkic language

Some developments in the modern world necessitate a fundamentally different approach to the issue of a common language. In particular, the growing convergence among Turkic peoples has intensified the need for a shared linguistic medium. In the current context, with political and ideological barriers removed, it has become more feasible for one of the existing Turkic languages to assume the role of a common language. Nevertheless, the issue remains a subject of diverse perspectives. Occasionally, discussions arise around the creation of an entirely new common Turkic language.

Addressing this debate, A. Akhundov noted:

“When we speak of a common Turkic language, what is meant is not the creation of a new language, but rather the establishment of all Turkic languages on the basis of unified principles – such as a common alphabet, orthographic norms, and shared terminology. The concept of a common Turkic language encompasses two critical aspects (İskender, H., 2023):

1. a common Turkic language as a universal tool for communication, and
2. a common Turkic language as one underpinned by a shared alphabet, orthography, and terminology” (Akhundov, 2003, p. 245).

The author emphasized that there is no necessity to artificially construct a new language by merging elements from various Turkic languages, as such an initiative would likely result in an unnatural and unproductive outcome.

Indeed, historical experience demonstrates that artificially created languages lacking national foundations and practical usage – such as Esperanto – have largely failed to integrate into actual communication systems.

In the context of intergovernmental relations and communication among Turkic-speaking communities across different states, selecting one of the existing Turkic languages as the common medium is a realistic alternative. This selection should be based on a set of objective criteria, including (Ouhssousou, S., & Tonkal, Ö., 2024):

- the language's scope of usage,
- its geographical distribution,
- accessibility for learning,
- number of speakers,
- phonetic and grammatical structure,
- and overall simplicity.

These features can serve as indicators of a language's potential for broader communicative adoption.

The solution to the problem of linguistic averaging among Turkic languages must begin with a comprehensive comparative analysis of their lexical foundations. According to A. Akhundov, two principles are central to addressing the common Turkic language issue through lexical alignment:

1. identifying the historical trajectory of borrowed lexicons in Turkic languages and selecting the most advanced variants in terms of modernity, and
2. adhering to historical convention, particularly in orthographic practices (Akhundov, 2003).

Naturally, the terminological system stands out as a key area of focus. The current linguistic status of each Turkic language allows for targeted interventions in its terminological system. In fact, many fields already display overlapping terminological units across Turkic languages, and further harmonization of these elements is entirely feasible.

A practical avenue for such harmonization lies in lexicography. It is therefore essential to conduct comparative studies of modern terminological dictionaries of various Turkic languages. Electronic lexicography, in particular, offers the advantage of compiling all field-specific terminological dictionaries into digital formats, thereby enabling wider accessibility and integration.

This task is also significant for the development of a common *corpus* of Turkic languages. In both world linguistics and Russian linguistic tradition, a *corpus* is understood as a systematically organized collection of texts in a natural language, stored electronically, and designed to support research on various linguistic features. As Mahmudov (2013) describes, corpora are created and used based on specific rules and offer valuable empirical data for linguistic studies.

Within the framework of a national corpus, two important sub-corpora can be defined: lexicology and terminography. These components assume particular relevance in addressing the problem of language averaging and establishing a common Turkic language. Naturally, each national corpus is an open and dynamic system, subject to refinement and expansion through the inclusion of new sub-corpora.

A sub-corpus on terminography should encompass existing terminological dictionaries and also project the development of new ones. The process of lexical averaging at the dictionary level requires, above all, the creation of comprehensive term inventories. Since averaging concerns multiple languages, the development of multilingual dictionaries becomes a key necessity. Ensuring concept-term equivalence, managing synonymy, and reducing polysemy are central tasks in this context, as these challenges can impede the early phases of linguistic convergence.

Constructing multilingual terminological dictionaries based on shared elements allows for partial resolution of these challenges. In practice, bilingual terminological dictionaries can serve as reference models. For instance, dictionaries such as English–Turkish, English–Azerbaijani, English–Turkmen, or English–Uzbek may be examined for comparative purposes.

Typically, in such dictionaries, the left column contains stable vocabulary items in the base language (e.g., English), while the right column lists their equivalents in the target language. However, selecting equivalents for terms often results in increased variability, optionality, or polysemy on the target-language side. Thus, several potential translations may correspond to a single base-language unit. Addressing this asym-

metry is an essential step in achieving meaningful lexical convergence.

The idea of a common Turkic language encompasses all structural levels of language, including phonology, morphology, syntax, and most critically, lexicon. Among these, the formulation of a unified lexical system stands out as a foundational component. In every language, new terms are continuously being created, and the number of borrowed or adapted terms in specialized terminologies is increasing. This dynamic facilitates the advancement of terminological commonality, which is essential for linguistic convergence (Kornfilt, Jaklin. 2018).

In contemporary linguistics, two principal tools for convergence among cognate languages are widely recognized: (1) the unification of orthographic systems through the adoption of a common alphabet, and (2) the harmonization of terminological systems. Both approaches support the broader goal of linguistic standardization and mutual intelligibility among related languages.

Currently, favorable conditions have emerged for the convergence and lexical averaging of specialized terminologies across different Turkic languages. There is a generally positive attitude toward this process, particularly in the domain of technical and scientific vocabulary. The presence of shared linguistic structures and historical interconnections among these languages further facilitates such efforts. In genetically related languages, opportunities for convergence tend to be more pronounced due to overlapping grammatical, phonetic, and lexical features.

To illustrate these possibilities, a comparative review of terminological usage in Azerbaijani, Kyrgyz, Turkmen, and Uzbek – specifically in the field of automated control systems – proves insightful. The analysis reveals that numerous technical terms, such as *vector*, *ventil*, *anode*, *cathode*, *harmonica*, *balance*, *camera*, *generator*, *aggregate*, *histogram*, *phase*, *scheme*, *controller*, *coder*, *graph*, *code*, *diapason*, and *diagram*, are employed in virtually identical forms across these four languages. This phenomenon is primarily attributable to the transfer of these terms via the Russian language during the Soviet era (Pavlenko, A., 2008).

Many of these terms exhibit no graphical or phonological variation. For instance, terms such as *diapason*, *anode*, *cathode*, *phase*, *vector*, *contactor*, and *assembly* are entirely consistent in both spelling and pronunciation across the examined languages. Such uniformity significantly strengthens the prospects for terminological unification and supports the broader process of linguistic averaging. These terms, by virtue of their identical or near-identical forms, can be classified as *shared terminology* within the Turkic linguistic domain.

However, the comparative analysis also indicates the presence of minor phonetic discrepancies in some terms. A representative example is the term *system*. While Azerbaijani employs the form *sistem*, Kyrgyz, Turkmen, and Uzbek continue to use *systema*, retaining the final “a” as in the Russian original. This suggests that phonetic divergence, even when minimal, persists among certain terms due to differential degrees of Russification or internal language reforms. Nevertheless, these minor variations can be systematically addressed and reduced through standardization initiatives (Grenoble, Lenore A., and Nadezhda Ja Bulatova. 2017).

Implementing a lexicographic filtering process for technical vocabulary across disciplines will allow scholars and experts to construct a unified terminological system. Such a system could then be incorporated into the national language corpora of the respective Turkic languages. The next logical phase involves analyzing the functional usage of these terms in various linguistic environments to determine their adaptability, frequency, and semantic consistency (Fantechi, A., Gnesi, S., Lami, G., & Maccari, A., 2003).

Through systematic standardization efforts, supported by lexicographic analysis and corpus-based studies, it becomes increasingly viable to lay the foundation for a unified and functional terminological framework for the Turkic languages.

3. Establishment of relations among the sections of lexicography

Lexicography, as a multifaceted domain within linguistics, encompasses a variety of subfields distinguished by their purpose, content, and application. Despite numerous at-

tempts to define its exact structural divisions, there remains an absence of universally accepted boundaries and clear-cut criteria to differentiate its sections comprehensively. Nevertheless, based on historical development, functional orientation, and typological features of dictionaries, several core sections of lexicography have been identified. According to H. Jackson (2013), these sections include:

1. Philological lexicography
2. Teaching lexicography
3. Terminological lexicography
4. Encyclopedic lexicography
5. Nomenclatural lexicography
6. Etymological lexicography
7. Historical lexicography

However, this classification, while functional, is not devoid of limitations. For instance, empirical research and practical analysis suggest that *etymological* and *historical* lexicography may be better understood not as distinct categories, but as overlapping or even integrated components. Given that historical lexicography often addresses the origins and diachronic evolution of lexical units, it inherently incorporates etymological insights. In this light, etymological lexicography may be more appropriately positioned as a subdomain within historical lexicography, rather than as an independent section (Pearsall, J., 2015).

Similarly, elements related to pedagogy and history can be embedded within various other branches – particularly in philological, terminological, encyclopaedic, and nomenclatural lexicography. For example, the compilation of etymological or historical dictionaries can also be regarded as a function of philological lexicography, depending on the linguistic materials used and the intended purpose of the resource.

One notable subset of lexicographic practice is *teaching lexicography*, which is primarily concerned with language acquisition and instructional materials. This direction heavily relies on *lexico-statistical* methodologies. Numerous studies have demonstrated the application of *linguo-statistical* analyses in shaping teaching-oriented dictionaries. These include frequency dictionaries, alphabet-frequency lists, and core vocabulary compilations, which are critical in determining lexical minima for educational purposes. Although this area ini-

tially appears as an independent field, current research indicates that teaching lexicography functions more appropriately as a *subsection*, rather than a separate major branch, of philological or applied lexicography.

Taking these factors into account, contemporary scholarship increasingly supports the consolidation of lexicographic classifications into four principal categories (Fuertes Olivera, Pedro A., 2008):

1. **Philological lexicography** – encompassing historical, etymological, and teaching subdomains.
2. **Terminological lexicography** – focused on subject-specific and scientific vocabulary.
3. **Encyclopaedic lexicography** – dealing with concept-based knowledge rather than purely linguistic content.
4. **Nomenclatural lexicography** – emphasizing standardization and classification of object names and scientific designations.

Additional subfields that emerge from these categories – based on thematic, functional, or linguistic materials – should be regarded as subdivisions rather than standalone branches.

Modern technological advancements, particularly in the realm of *electronic lexicography*, have significantly transformed the structural and functional relationships among these lexicographic sections. The use of digital platforms and electronic databases allows for multi-purpose storage, processing, and dissemination of lexicographic content. For instance, it is now feasible to derive specialized, small-scale educational or explana-

tory dictionaries from extensive general encyclopedic databases through digital filtering and thematic segmentation.

Moreover, the integration of *illustrative material* into dictionaries has become more efficient due to searchable digital corpora and online linguistic resources. Researchers and compilers no longer face the same level of difficulty in sourcing examples and usage data; instead, they can access vast textual repositories through computational tools and search engines, thus enhancing both the quality and reliability of illustrative entries in modern dictionaries.

In conclusion, while traditional lexicographic classifications retain relevance, the dynamic nature of lexicographic practice – fueled by technological innovation – necessitates a more integrated and functionally adaptive understanding of its subdivisions and interrelations.

Conclusion

Digital technologies redefine modern lexicography by enabling theoretical convergence between lexicography and terminography, particularly in standardizing terminology. This study proposes a conceptual model for a unified lexicographic framework among Turkic languages, grounded in normalization and comparative terminological analysis. Core lexicographic domains – philological, terminological, encyclopedic, and nomenclatural – inform this framework, enriched by pedagogical insights. Digital lexicography thus emerges as a pivotal field, advancing theoretical paradigms for linguistic standardization and multilingual communication across cognate languages.

References

- Akhundov A. (2003). Language and literature. Baku, Ganjlik, – 660 p.
- Atkins Jr, L. C. R. L., Lamb, L. C. D. A., Barger, M. M. L., Adair, M. L. N., & Tiernan, M. M. J. (1996). 2025 In-Time Information Integration System (I 3 S).
- Fantechi, A., Gnesi, S., Lami, G., & Maccari, A. (2003). Applications of linguistic techniques for use case analysis. Requirements Engineering, – 8. – P. 161–170.
- Fuertes Olivera, Pedro A., and Ascensión Arribas-Baño. “Pedagogical specialised lexicography”. 2008. – P. 1–176.
- Gerd A. S. (1986). Fundamentals of scientific and technical lexicography. – Leningrad, LSU, – 72 p.
- Gorodetsky B. Y. (1983). Problems and methods of modern lexicography // New in foreign linguistics. Publication 14. – Moscow, Progress, P. 5–22.

- Grenoble, Lenore A., and Nadezhda Ja Bulatova. "Language standardization in the aftermath of the Soviet language empire". *Standardizing minority languages*. Routledge, 2017. – P. 118–134.
- Hjørland, B. (2023). Terminology. *Ko Knowledge Organization*, – 50(2). – P. 111–127.
- İskender, H. (2023). A Case of Unique Lexical Engineering: Deliberate Etymological Misrepresentation in Turkish Language Reform. *Küllüye*, – 4(2). – P. 331–349.
- Jackson, Howard. *Lexicography: an introduction*. Routledge, 2013.
- Kornfilt, Jaklin. "Turkish and the Turkic languages". *The world's major languages*. – Routledge, 2018. – P. 536–561.
- Leichic V. M. (1996). Applied science of terminology and its directions // *Applied philology – SPSU: SPU*, – P. 276–286.
- Mahmudov M. (2013). Computer linguistics. Baku, "Science and Education", – 356 p.
- May, S. (2014). Contesting public monolingualism and diglossia: Rethinking political theory and language policy for a multilingual world. *Language policy*, – 13. – P. 371–393.
- Ouhssousou, S., & Tonkal, Ö. (2024, September). Analysis of Global Language Dynamics: A Cross-Cultural Examination of the Most Spoken Languages and Perceived Learning Ease. In *2024 8th International Artificial Intelligence and Data Processing Symposium (IDAP)* (pp. 1–6). IEEE.
- Pavlenko, A. (2008). Multilingualism in post-Soviet countries: Language revival, language removal, and sociolinguistic theory. *International journal of bilingual education and bilingualism*, – 11(3–4). – P. 275–314.
- Pearsall, J. (2015). Historical principles sv. synchronic approaches. U *International Handbook of Modern Lexis and Lexicography*, ur. Patrick Hanks i Gilles-Maurice de Schryver, – P. 1–14.
- Schryver, G. M. (2003). Lexicographers' dreams in the electronic dictionary age. *International journal of lexicography*, – 16(2). – P. 143–199.
- Swanepoel, P. (2006). 5 Dictionary typologies: A pragmatic approach. In *A practical guide to lexicography* (pp. 44–69). John Benjamins Publishing Company.
- Tarp, S., & Gouws, R. H. (2008). A lexicographic approach to language policy and recommendations for future dictionaries. *Lexikos*, – 18(1). – P. 232–255.

submitted 01.05.2025;

accepted for publication 15.05.2025;

published 27.06.2025

© Pashayeva P. I., Abbasov A. G.

Contact: parvanapashayeva@bsu.edu.az; abbasabbasov@bsu.edu.az