

CHALLENGES IN COMPILING A DICTIONARY OF TOPONYMS: LINGUISTIC, HISTORICAL AND LEXICOGRAPHIC PROBLEMS

Ismatova Nodiraxon Maxmudovna

PhD researcher at Kokand State University

Email: nodiraismatova1981@gmail.com

<https://doi.org/10.5281/zenodo.20760293>

Abstract

A dictionary of toponyms is more than an alphabetical list of geographical names. It is a specialised lexicographic work that must describe the spelling, pronunciation, structure, origin, historical development and cultural meaning of place names. The compilation of such a dictionary is complicated by phonetic changes, dialectal variation, multilingual origins, semantic ambiguity, conflicting etymologies, repeated names and insufficient historical evidence. Based on the works of H. To'rayev, T. Nafasov, V. Nafasova and S. Qorayev, this paper analyses the principal difficulties encountered in preparing explanatory dictionaries of Uzbek toponyms, with particular reference to the Bukhara region. The study shows that a reliable toponymic dictionary requires the combined use of linguistic, historical, geographical, ethnographic and cartographic evidence. It also argues that uncertain interpretations should be clearly distinguished from established etymologies rather than presented as indisputable facts.

Keywords: toponymy, toponymic dictionary, lexicography, etymology, phonetic change, dialect, geographical names, Bukhara, standardisation.

Introduction

Toponyms are geographical proper names used to identify cities, villages, neighbourhoods, streets, rivers, lakes, mountains, deserts and other spatial objects. They function as linguistic units, but they also preserve information about the history, geography, ethnic composition, occupations and worldview of the communities that created them.

Nafasov and Nafasova describe toponyms as linguistic expressions of geographical space and as verbal records of a people's history. According to their approach, geographical names are valuable not only for linguistics but also for history, geography, ecology and cultural studies. Consequently, compiling a dictionary of toponyms requires more than explaining the modern lexical meaning of each name.

A toponymic dictionary should ideally answer several questions: What geographical object does the name identify? What is its standard written form? What earlier forms did it have? From which language does it originate? What are its structural components? Why was the place given this name? How certain is the proposed interpretation?

The purpose of this paper is to analyse the main linguistic, historical and methodological challenges involved in creating a dictionary of toponyms.

Research Materials and Methods

The study is based on the following sources:

- H. To'rayev's *Buxoro toponimikasi*;
- T. Nafasov and V. Nafasova's *O'zbek tili toponimlarining o'quv izohli lug'ati*;
- S. Qorayev's *Toponimika*.

The descriptive, comparative, etymological, componential and lexicographic methods were employed. Particular attention was given to cases involving variant forms, phonetic transformation, multilingual composition and competing etymological interpretations.

1. Determining the Original Form of a Toponym

One of the most difficult tasks is reconstructing the earliest or original form of a place name. Toponyms are often transmitted orally for centuries. During this process, unfamiliar sounds, morphemes and word combinations may be simplified or adapted to local pronunciation.

For example, To'rayev explains that **Parxon** may represent a shortened and phonetically altered form of *parixon*. Similarly, **Ofandar** is interpreted as a changed form of *bofanda*, meaning “weaver”.

Such transformations create a serious problem for lexicographers. The contemporary form may no longer clearly display its original lexical components. A compiler must therefore determine whether a name is:

- an original lexical formation;
- a contracted form;
- a dialectal pronunciation;
- a folk-etymological reinterpretation;
- or a corrupted written form.

The reconstruction cannot be based only on phonetic similarity. It should be supported by historical records, local pronunciation, geographical conditions and information about former occupations or ethnic groups.

2. Phonetic and Dialectal Variation

Toponyms are particularly sensitive to dialectal influence. The same name may be pronounced differently by local inhabitants, officials and speakers from other regions. In Bukhara, this issue is especially important because local speech reflects interaction among Uzbek, Tajik and historically Iranian linguistic elements.

The name **Soktari**, for instance, is associated with more than one possible reconstruction. One interpretation derives it from a dialectal form of the Tajik word *sokit*, connected with the meanings “quiet”, “calm” or “peaceful”. Another popular interpretation relates it to *sad piri*, understood as “one hundred spiritual teachers”.

This example illustrates that dialectal change may obscure a name's structure so extensively that several interpretations become possible. A dictionary compiler should therefore document:

1. the officially registered form;
2. the locally pronounced form;
3. historical written variants;
4. possible reconstructed forms.

Without this distinction, a dictionary may incorrectly present a secondary pronunciation as the original name.

3. Multilingual Origins

Uzbek toponymy was formed through prolonged contact among Turkic, Persian-Tajik, Arabic, Sogdian, Mongolian, Russian and other languages. Consequently, many place names are hybrid formations whose components belong to different linguistic systems.

Nafasov and Nafasova demonstrate that the interpretation of a toponym requires the identification of the language of each component. For example, **Novqat** is connected with Sogdian *nav* and *kat*, meaning “new fortress” or “new settlement”, while **Kitob** is analysed through ancient components associated with a settlement near water. Their dictionary also marks linguistic origins using labels such as Turkic, Mongolian, Persian-Tajik, Arabic, Sogdian and ancient Iranian.

The challenge is that similar components may occur in several related languages. A morpheme may also have changed in both form and meaning over time. Therefore, identifying the donor language solely from the modern appearance of a name may be misleading.

A reliable analysis requires comparison with:

- historical dictionaries;
- medieval manuscripts;
- neighbouring regional names;
- related Iranian and Turkic forms;
- local historical and geographical evidence.

4. Semantic Opacity

Many toponyms are semantically opaque to present-day speakers. Their components may contain obsolete words or meanings no longer used in ordinary speech.

For example, **Dalvarzin** is explained through ancient Iranian components associated with elevation, fortification and place. **Darxan** is connected with a Turkic-Mongolian social term denoting a privileged person, artisan or individual exempted from taxation.

Without knowledge of these historical meanings, modern speakers may reinterpret such names through familiar contemporary words. This produces folk etymology: an explanation based on phonetic resemblance rather than historical evidence.

The lexicographer must distinguish between:

- the present-day lexical association;
- the historical meaning of the base;
- the motivation for naming the geographical object.

These three levels do not always coincide.

5. Multiple Etymological Interpretations

A single place name may have several plausible explanations. This is one of the most serious methodological problems in toponymic lexicography.

To'rayev notes that, in some cases, all known interpretations are presented because the available evidence does not permit a final conclusion. Toponyms requiring additional research are intentionally left open to further investigation.

The name **Soktari** provides a clear example of this problem. Its derivation from a word meaning “quiet” is linguistically possible, but the explanation involving “one hundred spiritual teachers” reflects local historical memory. Neither version should automatically be treated as definitive unless supported by stronger documentary evidence.

A dictionary entry should classify interpretations according to their reliability:

- **documented etymology** — confirmed by early written forms;
- **probable etymology** — linguistically and historically well supported;
- **possible interpretation** — plausible but insufficiently documented;
- **folk explanation** — preserved in local oral tradition but linguistically uncertain.

Such labelling protects the dictionary from presenting hypotheses as established facts.

6. The Problem of Historical Written Forms

Historical documents may record the same name in Arabic, Persian, Cyrillic or Latin scripts. In addition, scribes often adapted names to the orthographic rules of the language in which they wrote.

For example, the modern name **Yorika** appears in documents of the Bukhara Emirate in the form *Yorigoh*, interpreted as “a place of assistance” or “a place that provides support”. This older form preserves a structure that is no longer transparent in the modern pronunciation.

7. Standardisation and Official Spelling

Local pronunciation and official spelling do not always correspond. Administrative documents may contain simplified, Russified, incorrectly separated or inconsistently transliterated forms.

For example, To‘rayev argues that **Ko‘lmiri** is more meaningfully understood through the form *Ko‘li amiri*, meaning “the emir’s lake” or “a lake belonging to the emirate”. However, changing an established official name solely on the basis of an etymological reconstruction may create practical problems.

A dictionary should therefore distinguish among:

- the official contemporary name;
- the historically reconstructed form;
- variant spellings;
- local pronunciation.

The headword should normally follow the officially accepted form, while the historical and linguistic variants should be provided in the explanation.

8. Repeated and Identical Toponyms

The same toponym may occur in several districts or villages. Names such as **Arablar**, **Arabxona**, **Qipchoq** and **Xalach** are found in different parts of the Bukhara region. To‘rayev stresses that repeated names should not automatically be merged because each may refer to a separate settlement with its own historical community.

This raises several lexicographic questions:

- Should identical names receive one general entry or several local entries?
- How should each settlement be geographically identified?
- Do the names share the same origin?
- Could one name have been transferred from another locality?

The best solution is to provide a separate sub-entry for each geographical object, including its district, type of object and local historical information.

9. Identifying the Type of Geographical Object

A geographical name may refer to more than one object. The same form can designate a river, district, village, street or neighbourhood. This process is known as transonymisation—the transfer of a proper name from one type of object to another.

For example, the name of a river may later become the name of a region, district or settlement. A dictionary that does not identify the object type may confuse the primary name with its later applications.

Each entry should therefore indicate whether the headword refers to a:

- city; village; neighbourhood; district; river; canal; lake; mountain; hill; street; historical monument.

The dictionary must also explain, where evidence permits, which object carried the name first.

10. Establishing the Motivation for Naming

Even when the lexical meaning of a toponym is understandable, determining why it was applied to a particular place can be difficult.

For instance, **Sho‘robod** is interpreted as a settlement created through the improvement of saline land, while **Miyonjo‘y** means “middle canal” and refers to a village situated along the middle part of a watercourse.

11. Separating Anthroponyms, Ethnonyms and Descriptive Names

Many toponyms originate from personal names, tribal names, professions or social groups. Their surface structure may resemble ordinary descriptive words.

For example:

- **Rahimobod** is interpreted as a settlement developed by a person named Rahim;
- **Mirzoyon** refers to a village inhabited by scribes or officials;
- **To‘prez** is associated with cannon-makers or metalworkers;
- **Arablar** and **Arabxona** preserve ethnic identity.

A lexicographer must determine whether the first component refers to:

- an individual;
- a family;
- an ethnic group;
- an occupation;
- a religious title;
- or a physical characteristic.

This classification requires historical population records and ethnographic evidence, not merely morphological analysis.

12. The Selection and Arrangement of Entries

A dictionary may be organised alphabetically, geographically, historically or thematically. Each model has advantages and disadvantages.

Nafasov and Nafasova use an alphabetic explanatory format, while To‘rayev organises many Bukhara names according to districts and local citizens’ assemblies. To‘rayev explains that this geographically based structure deliberately departs from a strictly alphabetical arrangement.

Alphabetical organisation facilitates rapid access to a name. Geographical organisation, however, better reveals relationships among neighbouring settlements, irrigation networks, ethnic groups and landscape features.

A comprehensive electronic dictionary could combine both systems by allowing users to search:

- alphabetically;
- by district;
- by object type;
- by source language;
- by semantic category;
- by historical period.

13. Incomplete and Uneven Evidence

Some famous toponyms are documented in numerous historical sources, while small villages may be known only through oral accounts. This creates an imbalance in dictionary entries.

A well-documented name may include:

- historical spellings;
- chronological changes;
- linguistic components;
- archaeological evidence;
- several scholarly interpretations.

A poorly documented name may have only a contemporary form and a local legend. The compiler must avoid filling such gaps with unsupported speculation.

Where evidence is insufficient, the dictionary should explicitly state:

“The etymology remains uncertain and requires additional historical and field research.”

Such transparency increases the scientific reliability of the work.

Recommendations for Compiling a Reliable Toponymic Dictionary

A scientifically grounded dictionary should follow several principles.

First, every entry should contain the standard form, object type, administrative location, pronunciation or dialectal variant, historical spellings, morphological structure, source language, semantic explanation and level of etymological certainty.

Second, linguistic evidence should be compared with archival, geographical, archaeological and ethnographic data.

Third, local oral traditions should be recorded but clearly distinguished from documented etymologies.

Fourth, all competing interpretations should be cited when the evidence is inconclusive.

Fifth, a digital database should be used to connect each name with maps, coordinates, historical documents, photographs and sound recordings of local pronunciation.

A recommended entry model is:

TOPONYM — object type; district or region. Historical variants. Structural components and source language. Literal meaning. Historical or geographical motivation. Alternative interpretations. Sources. Degree of certainty.

Conclusion

The compilation of a toponymic dictionary is a complex interdisciplinary task. Its principal difficulties arise from phonetic change, dialectal pronunciation, multilingual origins, obsolete vocabulary, semantic opacity, conflicting etymologies, historical spelling variation, repeated names and insufficient documentary evidence.

The works of To‘rayev, Nafasov, Nafasova and Qorayev demonstrate that geographical names cannot be interpreted through linguistic intuition alone. A toponym represents the interaction of language, history, geography, ethnic identity and cultural memory.

The main responsibility of the lexicographer is not merely to provide an attractive explanation but to distinguish evidence from assumption. A trustworthy toponymic dictionary must preserve historical variants, record local knowledge, indicate uncertainties and offer interpretations supported by several types of evidence. Such a dictionary becomes not only a reference work but also a linguistic archive of the historical and cultural landscape of a nation.

Adabiyotlar, References, Литературы:

1. Nafasov, T., & Nafasova, V. (2007). *O‘zbek tili toponimlarining o‘quv izohli lug‘ati*. Tashkent: Yangi asr avlodi.
2. Qorayev, S. (2006). *Toponimika*. Tashkent: O‘zbekiston faylasuflari milliy jamiyati nashriyoti.
3. To‘rayev, H. (2021). *Buxoro toponimikasi*. Bukhara: Durdona.