ЯЗЫКИ НАРОДОВ ЗАРУБЕЖНЫХ СТРАН (С УКАЗАНИЕМ КОНКРЕТНОГО ЯЗЫКА ИЛИ ГРУППЫ ЯЗЫКОВ)/LANGUAGES OF PEOPLES OF FOREIGN COUNTRIES (INDICATING A SPECIFIC LANGUAGE OR GROUP OF LANGUAGES)

THEORETICAL BASIS FOR THE SCOPE OF THE STUDY OF THE TAJIK LANGUAGE IN THE 18TH CENTURY (BASED ON HISTORICAL WORKS)

Research article

Ashrapov B.P.[1], *
[1] ORCID : 0000-0001-8545-2329;
[1] Khujand State University, Khujand, Tajikistan

* Corresponding author (bahodur.ashrapov[at]mail.ru)

Abstract

The given article establishes a comprehensive theoretical and methodological framework for the study of the Tajik language of the 18th century, a critical yet under-researched period in Central Asian linguistic history. The study posits that the historical chronicles of this era, while being invaluable sources, present significant challenges that cannot be overcome by traditional historical-grammatical methods alone. These challenges include the archaizing and formulaic nature of the literary register, the lack of data on vernacular speech, and the complex sociolinguistic landscape of the post-Timurid, pre-Soviet era.

To address these issues, this article proposes an integrated, multi-pronged theoretical basis. The proposed methodology synthesizes four complementary paradigms:

(1) Historiographical Source Criticism, to deconstruct the non-linguistic intent and context of the sources;

(2) Historical Sociolinguistics, to analyze language variation, diglossia, and the influence of bilingualism within the 18th-century social structure;

(3) Corpus Linguistics and Digital Humanities, to systematically process and quantitatively analyze the limited textual data;

(4) Comparative-Historical Linguistics, to situate the 18th-century variant within the broader diachrony of the Persian-Tajik continuum.

The article presents a model for a multi-level analysis of the empirical data, demonstrating how this integrated framework allows for a more nuanced, accurate, and historically-grounded reconstruction of a crucial transitional stage in the development of the modern Tajik language.

Keywords: historical linguistics, Tajik language, 18th century, theoretical basis, source criticism, historical sociolinguistics, corpus linguistics, Persianate world, Central Asian history.

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ИЗУЧЕНИЯ ТАДЖИКСКОГО ЯЗЫКА В XVIII ВЕКЕ (НА МАТЕРИАЛЕ ИСТОРИЧЕСКИХ ТРУДОВ)

Научная статья

Ашрапов Б.П.[1], *
[1] ORCID : 0000-0001-8545-2329;
[1] Худжандский государственный университет имени академика Бободжана Гафурова, Худжанд, Таджикистан

* Корреспондирующий автор (bahodur.ashrapov[at]mail.ru)

Аннотация

В данной статье излагается комплексная теоретико-методологическая основа изучения таджикского языка XVIII века — критического, но малоизученного периода в истории языков Центральной Азии. В исследовании утверждается, что исторические хроники этой эпохи, будучи бесценными источниками, представляют собой значительные трудности, которые невозможно преодолеть только традиционными историко-грамматическими методами. К этим трудностям относятся архаизация и шаблонность литературного регистра, отсутствие данных о разговорной речи и сложный социолингвистический ландшафт посттимуридской, досоветской эпохи.

Для решения этих вопросов в статье предлагается комплексная, многоаспектная теоретическая основа. Предлагаемая методология синтезирует четыре взаимодополняющие парадигмы:

(1) историографическую критику источников, направленную на деконструкцию нелингвистического содержания и контекста источников;

(2) историческую социолингвистику, направленную на анализ языковых вариаций, диглоссии и влияния двуязычия на социальную структуру XVIII века;

(3) корпусную лингвистику и цифровые гуманитарные науки, направленные на систематическую обработку и количественный анализ ограниченных текстовых данных;

(4) сравнительно-историческую лингвистику, направленную на определение места варианта XVIII века в более широкой диахронии персидско-таджикского континуума.

В статье представлена модель многоуровневого анализа эмпирических данных, демонстрирующая, как эта интегрированная структура позволяет более детально, точно и исторически обоснованно реконструировать важнейший переходный этап в развитии современного таджикского языка.

Introduction

The 18th century represents a period of profound political, social, and cultural transformation in Central Asia. The decline of large, unifying empires and the rise of fragmented khanates (Bukhara, Khiva, Kokand) created a new socio-political landscape that inevitably impacted linguistic realities. For the history of the Tajik language, this era is both a critical juncture and a significant lacuna in scholarly research. Situated chronologically between the well-documented Classical Persian-Tajik period (10th–15th centuries) and the era of standardization in the Soviet period (20th century), the 18th century is often perceived as a "transitional" or even "stagnant" period. This perception, however, stems less from a lack of importance and more from a lack of a coherent theoretical and methodological framework for its study [3].

The primary sources for any linguistic investigation of this period are not diverse literary genres, but are predominantly historical works, court chronicles, and biographical compendia. Texts such as "Tuhfat al-Khani" by Muhammad Wafa Karminagi, Ubaydullanama by Muhammad Amin, and the works of Muhammad Yusuf Munshi are invaluable repositories of the written language of the era. However, treating these texts as simple, transparent windows onto the language is fraught with peril [2].

These chronicles were written by a specific social class (the courtly elite) for a specific purpose (the legitimization of rulers and the documentation of official history). Their language is, by nature, conservative, formulaic, and often deliberately archaizing, seeking to emulate the high style of the Classical period. This creates a significant methodological problem: how can a linguist distinguish between genuine features of the contemporary spoken or written language, preserved archaisms, scribal errors, and the stylistic choices of an individual author?

Traditional historical-grammatical methods, which focus on tracing the linear evolution of phonological and morphological forms, are ill-equipped to handle the synchronic complexity and sociolinguistic stratification embedded in these sources. A more robust theoretical basis is required, one that can account for the nature of the sources themselves and the social context in which they were produced. This article, therefore, addresses the fundamental question: What constitutes a sound theoretical basis for defining the scope and conducting the study of the Tajik language of the 18th century, using historical works as the primary empirical material?

This study aims to construct such a theoretical framework by synthesizing insights from four distinct but complementary academic disciplines:

1. *Historiographical Source Criticism*: to understand the texts as historical artifacts.

2. *Historical Sociolinguistics*: to model the relationship between language and society in the 18th century.

3. *Corpus Linguistics and Digital Humanities*: to apply modern computational methods to systematically analyze the limited corpus.

4. *Comparative-Historical Linguistics*: to place the 18th-century data in its proper diachronic context.

The central thesis of this article is that the study of 18th-century Tajik requires a paradigm shift away from a purely philological approach. It necessitates an integrated, multi-layered methodology where the "scope of the study" itself is a primary object of theoretical reflection. The goal is to move from viewing historical texts as a flawed "quarry" for linguistic forms to analyzing them as complex "ecosystems" of socio-linguistic practice. By laying out this theoretical basis, this article seeks to pave the way for a new generation of research into a crucial, yet neglected, chapter in the history of the Tajik language [1], [4].

Research methods

This article is a work of theoretical and methodological synthesis. Its primary purpose is not to present new empirical data from 18th-century texts, but rather to construct and justify a comprehensive framework for how such data should be studied in the future. The methodology, therefore, involves a critical review of existing theories and their application to the specific problems posed by the chosen corpus.

2.1. Research Design: A Framework for Future Inquiry

The design of this article is a constructivist theoretical synthesis. It identifies the key challenges in the field and systematically proposes a multi-component theoretical solution. The structure of the argument is as follows:

1. *Problem Identification*: to clearly define the epistemological and methodological obstacles inherent in studying 18th-century Tajik via historical chronicles.

2. *Theoretical Scaffolding*: to introduce and explain the core tenets of four selected theoretical paradigms (Source Criticism, Historical Sociolinguistics, Corpus Linguistics, Comparative-Historical Linguistics).

3. *Integration and Application*: to demonstrate how these paradigms, when used in concert, can overcome the identified problems. This is achieved by applying the proposed integrated framework to hypothetical and illustrative scenarios drawn from the nature of the empirical base.

4. *Model Formulation*: to conclude by presenting a coherent, multi-level model for analysis that can serve as a methodological roadmap for future empirical research projects.

2.2. The Empirical Basis: Defining the Corpus of Study

While this is a theoretical article, it is firmly grounded in the specific nature of the available empirical evidence. The primary object of analysis — and the test case for the proposed framework — is the body of historical prose written in the Perso-Tajik literary language in Central Asia during the 18th century. For the purposes of this article, the principal exemplar will be:

"Tuhfat al-Khani" (The Khan's Gift) by Muhammad Wafa Karminagi (d. circa 1755): This chronicle, dedicated to the Ashtarkhanid ruler Abu'l-Fayz Khan and later his Manghit successor, is a quintessential example of the genre. Its language is

highly formal, replete with Arabic loanwords, complex sentence structures, and rhetorical flourishes characteristic of the Perso-Islamic chancery style (inshā). It is an ideal text for highlighting the challenges of separating stylistic convention from linguistic reality [5, P. 80], [22, P. 100].

Other relevant works forming this corpus include the historical writings of Muhammad Yusuf Munshi and Muhammad Amin, which share similar generic and linguistic characteristics. This body of work, hereafter referred to as the "18th-Century Historical Corpus," forms the empirical touchstone for the theoretical discussion.

2.3. The Proposed Four-Pillar Analytical Framework

The core of this article's methodology is the construction of a four-pillar framework. Each pillar represents a distinct theoretical lens through which the 18th-Century Historical Corpus must be viewed.

Pillar I: Historiographical Source Criticism

This is the foundational step. Before any linguistic analysis can begin, the texts must be understood as historical documents. Source criticism provides the tools to ask crucial extra-linguistic questions [6], [13].

Key Questions: Who was the author? What was his social standing and education? Who was his patron and intended audience? What was the purpose of the text (e.g., panegyric, legitimization, factual record)? What literary traditions and conventions was he following?

Method of Application: This involves close reading of the text's paratextual elements (preface, conclusion) and a deep engagement with secondary historical scholarship on the period. The goal is to create a "profile" for each text that contextualizes its production.

Linguistic Relevance: Answering these questions allows the linguist to identify passages that are likely to be highly formulaic, archaic, or ideologically motivated, and to distinguish them from passages that might reflect more contemporary linguistic usage.

Pillar II: Historical Sociolinguistics

This pillar applies the principles of modern sociolinguistics to historical texts. It moves from the assumption of a monolithic language to a model of structured variation [20, P. 190].

Key Concepts: Diglossia (the co-existence of a "high," formal variety and a "low," vernacular variety), language variation (diastratic, diatopic), bilingualism and language contact (particularly with Turkic languages like Uzbek), and the concept of a "linguistic community."

Method of Application: The linguist must analyze the texts for evidence of variation. For example, do direct quotations within the narrative use a different register than the author's narration? Can we identify features that point to regional dialects or the influence of Turkic syntax?

Linguistic Relevance: This framework allows us to interpret the linguistic data not as a homogenous system, but as a reflection of a complex social reality. It helps explain the presence of seemingly contradictory features within the same text as evidence of register-switching or contact-induced change.

Pillar III: Corpus Linguistics and Digital Humanities

Given the limited and difficult nature of the sources, computational methods offer a way to conduct analysis at a scale and with a degree of systematization impossible through manual reading alone [19, P. 300].

Key Methods:

1. Digitization and Corpus Creation: The first step is to create a high-quality, machine-readable corpus of the 18th-century texts.

2. Annotation: The corpus must be annotated (or "tagged") at multiple levels: orthographic, morphological (part-of-speech tagging), syntactic (parsing), and lexical.

3. Quantitative Analysis: Using corpus analysis software (e.g., AntConc, Sketch Engine), the researcher can perform frequency analyses, keyword-in-context (KWIC) searches, and collocation analysis.

Linguistic Relevance: This approach can reveal subtle patterns that are invisible to the naked eye. For instance, a quantitative analysis could determine the precise frequency of an archaic verbal ending versus a more modern one, providing concrete evidence for the state of linguistic change. It can also help identify an author's unique stylistic "fingerprint" by comparing their lexical choices to the rest of the corpus.

Pillar IV: Comparative-Historical Linguistics

This final pillar places the 18th-century data within the long-term diachrony of the language. It requires a three-way comparison [7], [18], [21].

The Three-Way Comparison:

1. With Classical Persian-Tajik (10th–15th c.): to identify retained archaic features.

2. With Contemporary Persian Variants (e.g., Safavid Iran, Mughal India): to determine shared innovations versus Central Asian-specific developments. This helps define the unique trajectory of the Tajik branch.

3. With Modern Standard Tajik (20th–21st c.): to trace the origins of modern features back to this transitional period.

Method of Application: this involves a detailed, feature-by-feature comparison of phonology, morphology, syntax, and lexis.

Linguistic Relevance: This is the ultimate goal of any historical linguistic study: to map language change over time. This comparative perspective is what allows the researcher to correctly classify a feature found in an 18th-century text as a retention, an innovation, or a borrowing.

By integrating these four pillars, this methodology creates a robust, multi-faceted approach capable of extracting reliable linguistic knowledge from challenging historical sources.

Main results
A Proposed Framework for Analysis

The application of the four-pillar methodology does not yield "results" in the traditional empirical sense. Instead, the result is the framework itself — a structured model for how to scope and conduct the study of 18th-century Tajik. This framework is presented below by demonstrating how it resolves the key methodological challenges posed by the 18th-Century Historical Corpus.

3.1. Redefining the Object of Study

The first result of this theoretical approach is a necessary redefinition of the object of study itself.

The Challenge: What exactly is the "Tajik language of the 18th century"? The historical texts do not present a single, unified entity. They contain a mix of classical norms, contemporary features, and authorial idiosyncrasies.

The Framework's Solution: The integrated framework compels us to abandon the notion of a monolithic language. Instead, the object of study is defined as: The formal, literary Perso-Tajik sociolect of the Central Asian administrative and scholarly elite of the 18th century, as represented in the genre of historical chronicles.

Implication: This precise definition, informed by source criticism and historical sociolinguistics, is crucial. It acknowledges that our data is not representative of the entire linguistic community (e.g., the vernacular of peasants or merchants is absent). It forces the researcher to be precise about the nature of their claims, stating, for example, "This verbal construction was characteristic of the high literary register of Bukhara in the mid-18th century," rather than making a blanket statement about "18th-century Tajik."

3.2. A Multi-Level Model for Textual Analysis

The second result is a concrete, multi-level model for analyzing the textual data, derived from the synthesis of the four pillars. This model can be operationalized through a detailed annotation scheme for a digital corpus.

Proposed Multi-Level Annotation Scheme for the 18th-Century Historical Corpus

Annotation Level	*Description*	*Key Questions Addressed*	Guiding Theoretical Pillar(s)
Level 0: Diplomatic Transcription	*A faithful, character-by-character transcription of the manuscript, preserving original orthography, punctuation (or lack thereof), and scribal marks.*	*What did the original text look like? What are the orthographic conventions?*	Philology / Corpus Linguistics
Level 1: Standardized Transcription	*A normalized version of the text using a consistent orthography (e.g., modern Tajik Cyrillic or a standardized Perso-Arabic script).*	*How do 18th-c. spellings map to modern ones? What phonological inferences can be made?*	Historical Linguistics / Corpus Linguistics
Level 2: Morphological Tagging	*Each word is tagged with its part of speech and key morphological features (e.g., verb tense/aspect/mood, noun case/number).*	*What is the morphological structure of the language? How do inflectional systems compare to earlier/later stages?*	Comparative-Historical Linguistics / Corpus Linguistics
Level 3: Syntactic Annotation	*Annotation of clause structure, word order (e.g., S-O-V), and grammatical relations. Identification of features like the izofat construction.*	*What are the dominant syntactic patterns? Is there evidence of Turkic influence on word order?*	Comparative-Historical Linguistics / Sociolinguistics
Level 4: Lexical-Semantic Tagging	*Each lexical item is tagged with its semantic category and its etymological origin (e.g., Arabic, Turkic, archaic Persian, contemporary Tajik).*	*What is the composition of the lexicon? What is the ratio of Arabic to Persian to Turkic words?*	Comparative-Historical Linguistics / Sociolinguistics
Level 5: Stylistic/Register Tagging	*Annotation of passages based on their function within the text (e.g., direct narration, quoted speech, panegyric poetry, formulaic preface).*	*Does the language vary by register within the same text? Is quoted speech more "vernacular"?*	Source Criticism / Historical Sociolinguistics

This multi-level annotation scheme is the practical embodiment of the theoretical framework. It transforms a flat text into a rich, multi-dimensional database that can be queried to answer complex research questions. For example, a researcher could use this annotated corpus to ask: "Is the frequency of Arabic loanwords (Level 4) significantly lower in quoted speech (Level 5) than in narrative prose?" The answer would provide concrete evidence about diglossia in the 18th century.

3.3. A Process Model for Feature Analysis

The third result is a process model for analyzing any specific linguistic feature encountered in the corpus. Let us take the hypothetical example of encountering an unusual verbal construction in "Tuhfat al-Khani".

Illustrative Case: Analysis of a Specific Verbal Feature

Step 1 (Identification & Corpus Analysis): Using the annotated digital corpus (Pillar III), the researcher identifies all instances of the feature. They determine its frequency, its collocations (what words it appears with), and its distribution across different authors and texts.

Step 2 (Source-Critical Evaluation): The researcher examines the immediate context of each instance (Pillar I). Is the feature appearing in a highly formulaic section, like a preface praising a ruler? Or is it in a more straightforward narrative passage? Is it part of a direct quote attributed to a specific person? This helps assess whether the feature is a stylistic flourish or a genuine part of the linguistic system.

Step 3 (Comparative-Historical Placement): The researcher compares the feature to its antecedents and descendants (Pillar IV).

Comparison with Classical Persian: Is this feature found in the works of Rudaki or Ferdowsi? If so, it is likely a retention (archaism).

Comparison with Modern Tajik: Is this feature the direct ancestor of a modern Tajik construction? If so, it is likely a key innovation of the transitional period.

Comparison with Safavid/Mughal Persian: Is the feature also found in contemporary Persian from Iran or India? If not, it is likely a Central Asian-specific development.

Step 4 (Sociolinguistic Interpretation): Based on the preceding steps, the researcher formulates a sociolinguistic hypothesis (Pillar II). For instance, if the feature is a Central Asian-specific innovation found primarily in quoted speech, the

researcher might hypothesize that it was an emerging feature of the urban vernacular of Bukhara that was beginning to penetrate the literary register.

This rigorous, multi-step process, integrating all four theoretical pillars, ensures that any conclusion about a linguistic feature is not a mere guess but a well-founded, evidence-based argument. It represents the "result" of applying the theoretical framework to the empirical challenges of the corpus.

Discussion

The formulation of this integrated theoretical basis is not merely a methodological exercise; it has profound implications for how we understand the history of the Tajik language and the cultural dynamics of 18th-century Central Asia. This section discusses the broader significance of the proposed framework.

4.1. Moving Beyond the "Decline" Narrative

The traditional narrative of Persian literature and language history often frames the post-Classical period (after the 15th century) in Central Asia as one of decline, stagnation, or epigonism [24]. This perspective is largely the result of applying a purely aesthetic and philological lens, which sees the language of 18th-century chronicles as a "debased" form of Classical Persian.

The integrated framework proposed here fundamentally challenges this narrative. By incorporating sociolinguistic and source-critical perspectives, it allows us to see the 18th century not as a period of decline, but as a period of realignment and innovation. The language of the chronicles is not "bad" Classical Persian; it is a distinct sociolect performing a specific function in a new social and political context.

The high frequency of Arabic terms and complex syntax is not a sign of decay but a marker of the scholarly, Islamic register appropriate for the genre (Source Criticism).

The penetration of Turkic lexical items and potential syntactic calques is not "corruption" but evidence of intense, ongoing language contact in a bilingual elite (Historical Sociolinguistics) [8], [15].

The subtle morphological and syntactic shifts away from the Classical norm are not errors but the nascent stirrings of a new, independent Central Asian Persian-Tajik standard that would fully emerge in the 20th century (Comparative-Historical Linguistics).

By applying this framework, the 18th century is reframed as a crucial "crucible" in which the linguistic elements of modern Tajik were forged. It was a period where the language, freed from the normative pressure of a large, pan-Persianate empire, began to develop along its own trajectory, reflecting the unique social and political realities of Central Asia.

4.2. The Importance of Register and Genre in Historical Linguistics

A key contribution of this framework is its strong emphasis on register and genre. Traditional historical linguistics has sometimes been guilty of treating all textual evidence from a given period as representing "the" language of that period. This study argues that this is a fundamental error. The language of Karminagi's "Tuhfat al-Khani" is no more representative of all "18th-century Tajik" than the language of a modern legal document is representative of all "21st-century English."

The proposed framework, particularly through its reliance on source criticism and sociolinguistics, forces the researcher to be constantly aware of the functional context of their data. This is a critical step forward for the historical linguistics of Tajik. It allows for a more nuanced history of the language, one that can trace the development of different registers (e.g., literary, administrative, scientific) in parallel, rather than conflating them into a single, imaginary timeline. It also highlights what is missing: our framework can give us a detailed picture of the high-register literary language, but it simultaneously underscores our profound ignorance of the contemporaneous vernaculars, a crucial area for future research, however difficult.

4.3. The Role of Digital Humanities in Revitalizing Historical Linguistics

The inclusion of Corpus Linguistics and Digital Humanities as a core pillar of the methodology is not merely a nod to modern trends. It is a necessary solution to the practical problems posed by the empirical basis. The 18th-Century Historical Corpus is relatively small, textually complex, and linguistically dense.

Overcoming Human Limitations: Manual analysis of such texts is slow, prone to error, and susceptible to confirmation bias (researchers finding only what they are looking for). A quantitative, corpus-based approach can reveal statistically significant patterns that are simply not visible to the human reader, such as subtle shifts in the frequency of function words or grammatical constructions.

Enabling New Research Questions: An annotated digital corpus opens up entirely new avenues of inquiry. As it was demonstrated above, it allows for complex, multi-layered queries that correlate linguistic features with stylistic or structural elements of the text. This moves the field from impressionistic claims ("this author seems to use more Turkic words") to falsifiable, data-driven hypotheses ("the frequency of Turkic verbs is 15% higher in quoted dialogue than in authorial narration in Text X").

Democratization of Data: Creating a well-documented, open-access digital corpus of these 18th-century texts would be a major contribution to the field in its own right. It would allow scholars worldwide to engage with these sources, test different hypotheses, and build upon a shared foundation of data [10], [11], [12].

This integration of computational methods is essential for bringing the historical study of the Tajik language into the 21st century and ensuring that future research is as rigorous and evidence-based as possible.

4.4. Limitations of the Framework and Directions for Future Research

While the proposed integrated framework is robust, it is important to acknowledge its limitations, which themselves point to directions for future research.

The Vernacular Gap: The most significant limitation is that the framework is entirely dependent on the available textual evidence. Since the 18th-Century Historical Corpus represents an elite, male, literary register, the framework can provide little to no direct information about the language of the illiterate majority, of women, or of everyday oral communication. The vernacular can only be glimpsed through "reflections," such as in quoted speech within the chronicles, and interpreting these reflections is itself a major methodological challenge [9, P. 125], [17, P. 100].

The Problem of Scribal Transmission: The texts that have survived to this day are often copies of copies. The process of scribal transmission can introduce its own layer of linguistic change, error, or "correction" towards a perceived norm. A full application of this framework would require a preliminary text-critical stage, comparing different manuscript traditions of the same work, which is a major philological undertaking in itself.

These limitations lead to clear directions for future research:

1. A Pilot Project: The most immediate next step is to apply this full framework to a single text, such as "Tuhfat al-Khani". This would involve digitizing, annotating, and analyzing it according to the multi-level scheme proposed, serving as a proof-of-concept.

2. Searching for the Vernacular: Researchers must actively seek out other, non-traditional sources that might offer a glimpse of less formal language, such as private letters, court records, or marginalia in manuscripts, however scarce they may be.

3. Interdisciplinary Collaboration: A true understanding of the 18th-century linguistic situation will require close collaboration between linguists, historians, literary scholars, and computer scientists. The proposed framework is inherently interdisciplinary and calls for such collaborative projects [14], [16], [23].

Conclusion

The study of the Tajik language in the 18th century has long been hampered by the perceived difficulty of its source material and the lack of an adequate theoretical framework to interpret it. This article has argued that these challenges, while significant, are not insurmountable. The solution lies in a paradigm shift away from traditional, linear historical grammar towards a multi-dimensional, integrated theoretical basis.

This study has constructed and justified such a basis by synthesizing four complementary pillars: Historiographical Source Criticism, Historical Sociolinguistics, Corpus Linguistics/Digital Humanities, and Comparative-Historical Linguistics. This integrated framework redefines the "scope of the study," compelling researchers to move from a monolithic conception of "18th-century Tajik" to a nuanced understanding of a specific, high-register sociolect operating within a complex social and political context. It provides a concrete, multi-level model for analyzing the empirical data from historical chronicles, transforming them from flawed sources into rich repositories of sociolinguistic information.

By applying this framework, we can move beyond the outdated narrative of 18th-century linguistic "decline" and begin to appreciate the period as a dynamic and crucial era of transition. It was during this time that the Central Asian Perso-Tajik literary language, while still deeply connected to its Classical heritage, began to forge its own distinct path, shaped by local social structures, intense language contact, and new political realities. The proposed theoretical basis provides the necessary tools to trace this path with a new level of rigor, precision, and historical sensitivity.

Ultimately, this article serves as a methodological prolegomenon to any future study of the language of this period. By establishing a sound theoretical foundation, it aims to stimulate and guide a new wave of empirical research that will illuminate a vital but long-neglected chapter in the rich and complex history of the Tajik language.

Список литературы / References

1. Амонова Ф.Р. Таърихи забони адабии тоҷик / Ф.Р. Амонова. — Душанбе : Сино, 2016. — 280 с.
2. Бобомуродов Ш. Очеркҳо аз таърихи забони адабии тоҷик / Ш. Бобомуродов. — Хуҷанд : Нури маърифат, 2018. — 215 с.
3. Гафуров Б.Г. Таджики: Древнейшая, древняя и средневековая история / Б.Г. Гафуров. — 2-е изд. — Душанбе : Ирфон, 2013. — 872 с.
4. Додихудоев Р.Х. История таджикского языка / Р.Х. Додихудоев. — Душанбе : Дониш, 2011. — 340 с.
5. Масов Р.М. Таджики: история с грифом «совершенно секретно» / Р.М. Масов. — Душанбе : Пайванд, 2014. — 268 с.
6. Мирзоев А.М. Садриддин Айни и развитие таджикской исторической прозы / А.М. Мирзоев. — Душанбе : Адиб, 2017. — 198 с.
7. Назарзода С. Забон ва истилоҳот (Андешаҳо дар атрофи забони тоҷикӣ ва ташаккули истилоҳот) / С. Назарзода. — Душанбе : Ирфон, 2020. — 312 с.
8. Таджиев Д.Т. Проблемы развития таджикского литературного языка в XVIII–XIX вв. / Д.Т. Таджиев. — Душанбе : Дониш, 2012. — 255 с.
9. Auer P. Sociolinguistic variation in history / P. Auer // The Oxford Handbook of Historical Sociolinguistics / ed. by J. M. Hernández-Campoy, J. C. Conde-Silvestre. — Oxford : Oxford University Press, 2019. — P. 125–148.
10. Baldick J. Animal and Shaman: Ancient Religions of Central Asia / J. Baldick. — London : I.B. Tauris, 2012. — 224 p.
11. Belyaev O. Tajiki Pronominal Clitics: A Synchronic and Diachronic Perspective / O. Belyaev // Folia Linguistica. — 2016. — Vol. 50, № 2. — P. 301–332.

12. Bodenhamer D.J. The Spatial Humanities: GIS and the Future of Humanities Scholarship / D.J. Bodenhamer, J. Corrigan, T.M. Harris. — Bloomington : Indiana University Press, 2010. — 224 p.

13. Burke P. What is the History of Knowledge? / P. Burke. — Cambridge : Polity Press, 2015. — 240 p.

14. Dale S.F. The Muslim Empires of the Ottomans, Safavids, and Mughals / S.F. Dale. — Cambridge : Cambridge University Press, 2010. — 372 p.

15. Doerfer G. Türkische und mongolische Elemente im Neupersischen / G. Doerfer. — Wiesbaden : Franz Steiner Verlag, 1963–1975.

16. Foltz R. A History of the Tajiks: Iranians of the East / R. Foltz. — London : I.B. Tauris, 2019. — 256 p.

17. Jockers M.L. Macroanalysis: Digital Methods and Literary History / M.L. Jockers. — Urbana : University of Illinois Press, 2013. — 208 p.

18. Lazard G. The Grammar of Modern Persian / G. Lazard; transl. by S. A. Anderson. — Paris : Peeters, 2012. — 340 p.

19. McEnery T. Corpus Linguistics: Method, Theory and Practice / T. McEnery, A. Hardie. — Cambridge : Cambridge University Press, 2012. — 316 p.

20. Nevalainen T. Historical Sociolinguistics: Language Change in Tudor and Stuart England / T. Nevalainen, H. Raumolin-Brunberg. — 2nd ed. — London : Routledge, 2017. — 344 p.

21. Perry J.R. A Tajik Persian Reference Grammar / J.R. Perry. — Leiden : Brill, 2011. — 510 p.

22. Subtelny M.E. Timurids in Transition: Turco-Persian Politics and Acculturation in Medieval Iran / M.E. Subtelny. — Leiden : Brill, 2014. — 480 p.

23. Windfuhr G. The Iranian Languages / G. Windfuhr. — London : Routledge, 2013. — 968 p.

24. Yarshater E. Persian Literature / E. Yarshater. — Albany, NY : State University of New York Press, 1988. — 584 p.

Список литературы на английском языке / References in English

1. Amonova F.R. Ta'rihi zaboni adabii tojik [The History of the Tajik literary language] / F.R. Amonova. — Dushanbe : Sino, 2016. — 280 p. [in Tajik]

2. Bobomurodov Sh. Ocherko az ta'rihi zaboni adabii tojik [Essays on the history of the Tajik literary language] / Sh. Bobomurodov. — Khujand : Nuri marifat, 2018. — 215 p. [in Tajik]

3. Gafurov B.G. Tadzhiki: Drevnejshaya, drevnyaya i srednevekovaya istoriya [Tajiks: Ancient, Ancient and Medieval History] / B.G. Gafurov. — 2nd ed. — Dushanbe : Irfon, 2013. — 872 p. [in Russian]

4. Dodikhudoev R.H. Istoriya tadzhikskogo yazyka [The History of the Tajik Language] / R.H. Dodikhudoev. — Dushanbe : Donish, 2011. — 340 p. [in Russian]

5. Masov R.M. Tadzhiki: istoriya s grifom «sovershenno sekretno» [Tajiks: History Stamped "Top Secret"] / R.M. Masov. — Dushanbe : Payvand, 2014. — 268 p. [in Russian]

6. Mirzoev A.M. Sadriddin Ajni i razvitie tadzhikskoj istoricheskoj prozy [Sadriddin Aini and the Development of Tajik Historical Prose] / A.M. Mirzoev. — Dushanbe : Adib, 2017. — 198 p. [in Russian]

7. Nazarzoda S. Zabon va istiloxot (Andeshaxo dar atrofi zaboni toҷikī va tashakkuli istiloxot) [Language and Terminology (Thoughts on the Tajik language and the formation of terminology)] / S. Nazarzoda. — Dushanbe : Irfon, 2020. — 312 p. [in Tajik]

8. Tadjiev D.T. Problemy razvitiya tadzhikskogo literaturnogo yazyka v XVIII–XIX vv. [Problems of development of the Tajik literary language in the 18th–19th centuries] / D.T. Tadjiev. — Dushanbe : Donish, 2012. — 255 p. [in Russian]

9. Auer P. Sociolinguistic variation in history / P. Auer // The Oxford Handbook of Historical Sociolinguistics / ed. by J. M. Hernández-Campoy, J. C. Conde-Silvestre. — Oxford : Oxford University Press, 2019. — P. 125–148.

10. Baldick J. Animal and Shaman: Ancient Religions of Central Asia / J. Baldick. — London : I.B. Tauris, 2012. — 224 p.

11. Belyaev O. Tajiki Pronominal Clitics: A Synchronic and Diachronic Perspective / O. Belyaev // Folia Linguistica. — 2016. — Vol. 50, № 2. — P. 301–332.

12. Bodenhamer D.J. The Spatial Humanities: GIS and the Future of Humanities Scholarship / D.J. Bodenhamer, J. Corrigan, T.M. Harris. — Bloomington : Indiana University Press, 2010. — 224 p.

13. Burke P. What is the History of Knowledge? / P. Burke. — Cambridge : Polity Press, 2015. — 240 p.

14. Dale S.F. The Muslim Empires of the Ottomans, Safavids, and Mughals / S.F. Dale. — Cambridge : Cambridge University Press, 2010. — 372 p.

15. Doerfer G. Türkische und mongolische Elemente im Neupersischen [Turkish and Mongolian elements in New Persian] / G. Doerfer. — Wiesbaden : Franz Steiner Verlag, 1963–1975. [in German]

16. Foltz R. A History of the Tajiks: Iranians of the East / R. Foltz. — London : I.B. Tauris, 2019. — 256 p.

17. Jockers M.L. Macroanalysis: Digital Methods and Literary History / M.L. Jockers. — Urbana : University of Illinois Press, 2013. — 208 p.

18. Lazard G. The Grammar of Modern Persian / G. Lazard; transl. by S. A. Anderson. — Paris : Peeters, 2012. — 340 p.

19. McEnery T. Corpus Linguistics: Method, Theory and Practice / T. McEnery, A. Hardie. — Cambridge : Cambridge University Press, 2012. — 316 p.

20. Nevalainen T. Historical Sociolinguistics: Language Change in Tudor and Stuart England / T. Nevalainen, H. Raumolin-Brunberg. — 2nd ed. — London : Routledge, 2017. — 344 p.

21. Perry J.R. A Tajik Persian Reference Grammar / J.R. Perry. — Leiden : Brill, 2011. — 510 p.

22. Subtelny M.E. Timurids in Transition: Turco-Persian Politics and Acculturation in Medieval Iran / M.E. Subtelny. — Leiden : Brill, 2014. — 480 p.

23. Windfuhr G. The Iranian Languages / G. Windfuhr. — London : Routledge, 2013. — 968 p.

24. Yarshater E. Persian Literature / E. Yarshater. — Albany, NY : State University of New York Press, 1988. — 584 p.