

Corpus-Pragmatic Analysis of Uzbek Paremias: Contextual Variation of Illocutionary Functions and Cross-Genre Distribution

YAKHYOYEVA UMIDA IBROHIM KIZI
Asia International University, Uzbekistan

ABSTRACT

This article examines Uzbek paremias at the intersection of paremiology, corpus linguistics, and pragmatics, with particular attention to the contextual variability of their illocutionary functions. The study proceeds from the hypothesis that the pragmatic status of a proverb cannot be adequately determined from its decontextualized dictionary form alone; rather, it emerges through the interaction of the paremic unit, discourse context, speaker intention, addressee, genre, and communicative situation. A pilot corpus is constructed from contextual occurrences of Uzbek proverbs in literary texts by Abdulla Qahhor, Said Ahmad, Abdulla Qodiriy, Hamza Hakimzoda Niyoziy, Omon Muxtor, and Xayriddin Sulton. The corpus-based procedure combines concordance analysis, contextual annotation, speech-act classification, cross-author and cross-genre comparison, and elementary quantitative analysis. Particular attention is paid to the paremias Tekinga mushuk oftobga chiqmaydi and Quruq qoshiq og'iz yirtar, which occur in different authorial and genre environments. The analysis demonstrates that formally identical or semantically related paremias retain a relatively stable propositional core while acquiring different pragmatic nuances, evaluative orientations, communicative aims, and perlocutionary effects across contexts. The article proposes a multi-layer annotation model for an Uzbek paremiological corpus including lemma, variant, author, work, genre, speaker, addressee, speech mode, illocutionary class, pragmatic subfunction, evaluative polarity, implicitness, and perlocutionary effect. The findings suggest

that corpus pragmatics provides a productive methodological framework for moving Uzbek paremiology from isolated interpretation toward reproducible and quantitatively supported contextual analysis.

Keywords: Paremia, proverb, corpus linguistics, corpus pragmatics, speech act, illocutionary function, concordance, pragmatic annotation, genre, contextual variation, Uzbek language.

1. INTRODUCTION

In twenty-first-century linguistics, the study of language has increasingly shifted from exclusively theoretical description toward empirical analysis based on naturally occurring data. Corpus linguistics has played a major role in this development by making it possible to collect, annotate, retrieve, compare, and statistically process large bodies of texts. The integration of corpus methodology with pragmatics has, in turn, established corpus pragmatics as a field concerned not only with the frequency of linguistic forms but also with the communicative functions that those forms perform in authentic discourse.

This perspective is particularly productive for paremiology. Proverbs and related paremic units are stable expressions that encode collective experience, cultural knowledge, and generalized judgments; however, their function in discourse is not exhausted by their dictionary meaning. In actual communication, a speaker may employ a proverb to assert, justify, advise, warn, criticize, ridicule, reject, persuade, or indirectly request something. Consequently, the same proverb may perform different speech acts depending on who uses it, to whom it is addressed, in what genre it occurs, and what communicative situation surrounds it.

The distinction between semantic content and pragmatic function is especially important in the analysis of Uzbek paremias. The proverb *Tekinga mushuk oftobga chiqmaydi* literally conveys the generalized idea that even a cat does not go out into the sun for nothing. Its propositional core can be

interpreted as the assumption that an action is normally motivated by some benefit or interest. Yet in literary discourse this stable proposition may support substantially different communicative intentions: it can justify expenditure, imply a demand for payment, characterize greed, or express distrust toward seemingly disinterested assistance.

For this reason, the present study treats the *paremia* not as an isolated folkloric formula but as a context-sensitive corpus unit. The proposed analytical chain is: *paremia* → textual variant → discourse context → speaker/addressee configuration → communicative intention → illocutionary function → evaluative orientation → perlocutionary effect. Such a model makes it possible to connect the traditional concerns of *paremiology* with the empirical procedures of corpus linguistics and the functional categories of pragmatics.

The main objective of the study is to develop and test a corpus-pragmatic framework for identifying contextual changes in the illocutionary functions of Uzbek *paremias* and for comparing their distribution across authors and genres. The study addresses four interrelated questions: how does the pragmatic function of a *paremia* differ from its stable propositional content; how does the same *paremia* vary across authorial contexts; how does genre influence its speech-act realization; and how can such information be represented through a reproducible corpus annotation model?

2. THEORETICAL FRAMEWORK

Speech act theory provides the principal pragmatic foundation of the analysis. Austin distinguishes locutionary, illocutionary, and perlocutionary dimensions of utterances, while Searle systematizes illocutionary acts into representative/assertive, directive, commissive, expressive, and declarative classes. Applied to *paremias*, this framework allows the researcher to move beyond thematic classification and ask what action a proverb performs in a particular communicative episode.

A central methodological difficulty is that illocutionary force is not always explicitly encoded in grammatical form. An

imperative proverb such as *Yetti o'lchab, bir kes* ('Measure seven times, cut once') has an evident directive potential. By contrast, *Tekinga mushuk oftobga chiqmaydi* has the grammatical form of a declarative sentence, but in context it may function as a justification, an indirect demand, a skeptical evaluation, or a characterization of self-interest. Therefore, a corpus-pragmatic analysis cannot equate sentence type with speech-act function.

Corpus pragmatics is useful precisely because it connects recurrent linguistic forms with their contextual functions. Instead of asking only what a proverb means in the abstract, the researcher can investigate who uses it, in which work, in which genre, in what speech mode, with what pragmatic aim, and with what observable response. Repeated contextual observations then make it possible to describe not a single predetermined function but a distribution of functions.

This study therefore proposes three interconnected annotation layers. The paremiological layer identifies the canonical proverb and its textual variants. The discourse layer records author, work, genre, speaker, addressee, and speech mode. The pragmatic layer records the illocutionary class, more specific pragmatic subfunction, evaluative polarity, degree of implicitness, and, where recoverable from the text, the perlocutionary outcome. The combination of these layers constitutes the basis of a corpus-pragmatic description of Uzbek paremias.

3. MATERIALS AND METHODS

The empirical basis of the study is a pilot dataset of contextual occurrences of Uzbek paremias in literary works by Abdulla Qahhor [10: 330], Said Ahmad [11: 280], Abdulla Qodiriy [12: 264], Hamza Hakimzoda Niyoziy [13], Omon Muxtor [14: 260], and Xayriddin Sulton [15: 230]. The present article concentrates on two semantically related paremic models: P1, *Tekinga mushuk oftobga chiqmaydi*, and P2, *Quruq qoshiq og'iz yirtar*. Their variants are examined in seven literary contexts. Because the dataset is deliberately small, the numerical results are treated as

pilot statistics rather than as estimates of the overall distribution of Uzbek proverbs.

The analysis combines five procedures. First, concordance analysis is used to examine the proverb together with its preceding and following textual environment rather than as an isolated sentence. Second, contextual-pragmatic analysis identifies the speaker, addressee, communicative situation, presuppositions, and immediate purpose of the paremia. Third, speech-act analysis distinguishes the general illocutionary class from a more specific contextual subfunction. Fourth, corpus annotation assigns structured metadata to every occurrence. Fifth, elementary quantitative analysis compares the distribution of the selected occurrences across authors and genres.

For quantitative description, the proportion of a category is calculated as $P = n/N \times 100$, where P is the percentage of the relevant category, n is the number of occurrences assigned to that category, and N is the total number of contextual occurrences in the pilot dataset. Since $N = 7$ in the present analysis, the percentages are descriptive and should not be generalized to the full Uzbek literary tradition.

A further methodological distinction is made between paremia-type and paremia-token. Parema-type refers to the normalized canonical proverb, whereas paremia-token refers to the actual textual realization found in a specific work. This distinction allows orthographic, lexical, dialectal, and authorial variants to be grouped under one searchable unit without erasing their individual textual forms.

4. RESULTS

4.1. *Paremiological type and contextual token*

The corpus perspective shows that the canonical form of a proverb is the starting point rather than the endpoint of analysis. The type *Tekinga mushuk oftobga chiqmaydi* is realized through textual variants such as *Mushuk tekinga oftobga chiqmaydi*, *Mishiq oftobga tekin chiqmaydir*, and *Mushuk bekorga oftobga chiqmaydi*. These variants differ lexically and stylistically but

preserve a recognizable semantic core: action is associated with an underlying benefit, interest, or motivation.

Normalizing the variants under one paremiological lemma makes cross-text retrieval possible, while preserving the token form allows the researcher to observe dialectal, historical, stylistic, and authorial variation. The type/token distinction is therefore essential for a corpus of paremias because proverb variation is itself a meaningful linguistic phenomenon.

4.2. *Contextual variation of* Tekinga mushuk oftobga chiqmaydi

In Abdulla Qahhor's short story 'O'g'ri,' [10: 330] the proverb occurs in the context of Qobil bobo's reasoning that something must be given to the ellikboshi. The paremia functions as an argument that naturalizes a social expectation: assistance or official intervention is not assumed to be free. The contextual subfunction is therefore justification or confirmation of a practical decision. Its perlocutionary relevance lies in strengthening the character's readiness to act according to this perceived rule.

In Said Ahmad's [11: 280] context, a closely related form is directed toward interlocutors and is associated more directly with material interest. The stable proposition about non-gratuitous action is pragmatically reconfigured as an implication that some benefit or payment is expected. The utterance consequently moves from general observation toward an indirect demand.

In Abdulla Qodiriy's [12:264] textual variant *Mishiq oftobga tekin chiqmaydir*, the paremia contributes to characterization. It is not merely an independent statement about human motivation; it becomes a discourse marker through which acquisitiveness and self-interest are associated with a character. The pragmatic force is therefore intertwined with ironic or critical evaluation.

In Omon Muxtor's [14: 260] *Egilgan bosh*, the paremia is linked with skepticism toward apparently disinterested assistance. The speaker interprets an offer through the presupposition that no action is entirely without motive. Here the proverb functions as an expression of distrust and supports rejection or resistance rather than a request for benefit.

These contexts demonstrate a central corpus-pragmatic pattern: the propositional core of a paremia may remain relatively stable while its illocutionary realization and pragmatic subfunction vary substantially. The same paremia-type can participate in justification, indirect requesting, character evaluation, and skeptical rejection.

Author	Textual variant	Contextual pragmatic function	Evaluation
A. Qahhor	<i>Tekinga mushuk oftobga chiqmaydi</i>	justification / confirmation	context-dependent
S. Ahmad	<i>Mushuk tekinga oftobga chiqmaydi</i>	indirect material demand	negative
A. Qodiriy	<i>Mishiq oftobga tekin chiqmaydir</i>	characterization of self-interest	negative
O. Muxtor	<i>Mushuk bekorga oftobga chiqmaydi</i>	distrust / skeptical evaluation	negative

4.3. Contextual variation of *Quruq qoshiq og'iz yirtar*

The second paremia belongs to the same broad conceptual field but reveals a different contextual range. In Qahhor's *O'g'ri*, *Quruq qoshiq og'iz yirtadi* is associated with Qobil bobo's understanding that one cannot approach an official empty-handed. The proverb thus functions as a pragmatic motivation for material preparation before action.

In Hamza Hakimzoda Niyoziy's [13] drama *Boy ila xizmatchi*, the paremia occurs in direct interaction. The dramatic form foregrounds the speaker-addressee relationship and makes the connection between utterance and expected response especially visible. The proverb implies the necessity of material inducement and contributes to the progression of the dialogue.

In Xayriddin Sulton's [15: 230] prose, the paremia is connected more broadly with material necessity and livelihood. In this context, its pragmatic scope is not limited to bribery or exchange; rather, it supports a general claim that practical needs cannot be satisfied by words alone. This difference shows that semantic relatedness between paremias does not entail identical pragmatic realization in every context.

4.4. *Distribution across authors and genres*

The seven pilot occurrences are distributed across six authors. Abdulla Qahhor contributes two tokens, or 28.6% of the pilot dataset, because both selected paremic models occur in *O'g'ri*. Each of the other five authors contributes one token, or 14.3%. These figures do not measure the authors' general paremiological density; they describe only the selected two-proverb sample.

Genre distribution is similarly descriptive. Three occurrences belong to short-story contexts (42.9%), two to novels (28.6%), one to drama (14.3%), and one to a novella or qissa (14.3%). The value of this comparison lies not in ranking genres but in demonstrating that genre should be included as a corpus variable. Narrative prose can embed paremias in authorial narration, internal speech, and dialogue, whereas drama normally makes the speaker-addressee configuration more immediately observable. Longer prose forms can additionally integrate paremias into sustained characterization and social evaluation.

5. A PRAGMATIC ANNOTATION MODEL FOR AN UZBEK PAREMIOLOGICAL CORPUS

The findings indicate that an electronic collection of proverb texts and bibliographic sources would not be sufficient for corpus-pragmatic research. A specialized paremiological corpus should encode both linguistic form and contextual function. The following annotation fields are proposed as a minimum working model.

Annotation field	Content
PAREMIA_ID	unique corpus identifier
LEMMA	normalized canonical form
VARIANT	actual textual form
AUTHOR / WORK	bibliographic source
GENRE	short story, novel, drama, qissa, etc.
SPEAKER / ADDRESSEE	participants in the communicative act
SPEECH_MODE	dialogue, monologue, internal speech, narration
ILLOCUTION	representative, directive, commissive, expressive, declarative
SUBFUNCTION	justification, warning, advice, irony, criticism,

	request, etc.
EVALUATION	positive, negative, neutral/context-dependent
IMPLICITNESS	explicit or implicit pragmatic force
PERLOCUTION	textually observable communicative outcome
CONTEXT	expanded concordance window

An important feature of the model is the separation of ILLOCUTION from SUBFUNCTION. Two occurrences may belong to the same broad Searlean class while performing different contextual actions. For example, representative uses may function as confirmation, justification, skeptical evaluation, or characterization. This distinction reduces the risk of forcing contextually diverse data into overly broad categories.

6. THE ILLOCUTIONARY PROFILE OF A PAREMIA

On the basis of corpus-pragmatic analysis, the concept of an illocutionary profile can be introduced. The illocutionary profile of a paremia is the qualitative and quantitative configuration of the speech-act classes and pragmatic subfunctions attested across all of its corpus occurrences. It can be represented schematically as $IP(P) = \{IA \square f \square, IA \square f \square, \dots, IA \square f \square\}$, where P is a paremia, IA is an attested illocutionary or pragmatic function, and f is its corpus frequency.

This model makes it possible to replace categorical statements such as “this proverb is directive” with empirically more precise descriptions. In a sufficiently large corpus, a proverb may display a dominant directive profile while also occurring in representative, expressive, or ironic-evaluative functions. The corpus can therefore reveal both dominant tendencies and secondary contextual realizations.

The notion of an illocutionary profile is particularly relevant to paremias because their formal stability often creates an impression of functional stability. Corpus evidence demonstrates, however, that stable wording can support variable communicative actions. Quantitative distribution and qualitative contextual interpretation must therefore be treated as complementary rather than competing methods.

7. DISCUSSION

Traditional paremiological studies often classify proverbs according to themes such as labour, friendship, homeland, morality, family, good and evil. Such classifications remain useful for semantic and cultural description, but they do not fully account for what a proverb does in discourse. Corpus pragmatics changes the central question from “What does this proverb mean?” to “Who uses this proverb, to whom, in what situation, for what purpose, in what genre, and with what communicative consequence?”

This shift has several methodological consequences. First, context must be preserved in the corpus rather than reduced to an isolated citation. Second, variants must be normalized without deleting their original forms. Third, author and genre metadata must be searchable. Fourth, speech-act categories should be supplemented with finer pragmatic subfunctions. Fifth, quantitative claims must be based on clearly defined denominators and should distinguish pilot observations from generalizable corpus results.

The pilot analysis also suggests that authorial comparison should not be limited to raw frequency. Authors may employ the same paremia in different discourse positions: internal reasoning, direct dialogue, ironic characterization, or evaluative narration. Genre comparison likewise requires more than counting occurrences, because genre affects the accessibility of speaker intention and perlocutionary consequences. Drama makes interaction explicit, whereas prose may require reconstruction of pragmatic force from narration and characterization.

A larger project can extend this methodology to hundreds or thousands of paremic occurrences. The workflow would include automatic or semi-automatic proverb detection, normalization of variants, concordance extraction, metadata assignment, expert pragmatic annotation, inter-annotator agreement testing, and statistical modelling. Once a sufficiently reliable manually annotated dataset is available, computational classification may be used to assist – but not replace – expert pragmatic analysis.

8. SCIENTIFIC NOVELTY AND RESEARCH IMPLICATIONS

The study contributes to Uzbek paremiology by integrating paremiological description, speech-act theory, and corpus methodology in a single analytical framework. Rather than assigning a proverb to one fixed pragmatic category, it proposes the reconstruction of its corpus-based illocutionary profile.

The distinction between *paremia-type* and *paremia-token* provides a practical solution to the problem of textual variation. The proposed multi-layer annotation model also makes author, genre, participant structure, pragmatic subfunction, evaluation, and perlocutionary outcome searchable corpus variables.

The approach creates a basis for statistically testing contextual variability. With a larger balanced corpus, researchers could compare the proportion of directive, representative, expressive, commissive, and other functions across genres, historical periods, authors, and discourse types; test whether particular *paremias* have statistically dominant functions; and identify contexts in which the same proverb systematically changes pragmatic force.

For Uzbek linguistics, this creates a bridge between traditional paremiological scholarship and contemporary digital humanities. A pragmatically annotated paremiological corpus could support research in pragmalinguistics, stylistics, literary linguistics, sociolinguistics, lexicography, language teaching, and computational processing of figurative and formulaic language.

9. CONCLUSION

The corpus-pragmatic analysis of Uzbek *paremias* demonstrates that the speech-act value of a proverb cannot be determined adequately from its canonical form alone. Its actual pragmatic function emerges from the interaction of semantic content, discourse context, speaker intention, addressee, genre, and the broader communicative situation.

The pilot analysis of “*Tekinga mushuk oftobga chiqmaydi*” shows that a relatively stable propositional core can support

different pragmatic subfunctions, including justification, indirect material demand, critical characterization, distrust, and skeptical evaluation. The contextual realizations of *Quruq qoshiq og'iz yirtar* similarly show that a related semantic principle may be activated in contexts of material inducement, practical necessity, and livelihood.

The quantitative distribution of the seven pilot occurrences – 42.9% in short stories, 28.6% in novels, 14.3% in drama, and 14.3% in *qissa* – should not be generalized beyond the sample. It nevertheless demonstrates the methodological value of genre as an independent corpus variable and illustrates how qualitative pragmatic interpretation can be combined with transparent quantitative description.

A specialized Uzbek paremiological corpus should therefore annotate not only the proverb and its source but also its canonical lemma, textual variant, author, work, genre, speaker, addressee, speech mode, illocutionary class, pragmatic subfunction, evaluative polarity, implicitness, perlocutionary effect, and expanded context.

The concept of an illocutionary profile offers a further step toward empirical paremiology. By summarizing all attested functions of a paremia and their frequencies, it becomes possible to identify dominant and secondary pragmatic tendencies without reducing a context-sensitive unit to one fixed category. The integration of paremiology, corpus linguistics, and pragmatics thus provides a promising framework for the development of a pragmatically annotated corpus of Uzbek paremias and for statistically grounded comparisons across authors, genres, and discourse types.

REFERENCES

1. Pirs. Ch. S. 1959. Collected papers of Charles Sanders Peirce. *Philosophy of Science*, 4/1, 109-121.
2. Safarov, Sh. 2008. Pragmalingvistika: monografiya. T.: O'zbekiston milliy ensiklopediyasi Davlat ilmiy nashriyoti.
3. Searle, J. R. 1969. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press.

4. Searle, J. R. 1979. *Expression and Meaning: Studies in the Theory of Speech Acts*. Cambridge: University Press.
5. Searle, J. & Daniel, V. 1985. *The Foundations of Illocutionary*. Cambridge University Press.
6. Searle, J. R. 1976. *A Classification of Illocutionary Acts*, 5/1, 1-23.
7. Серль Дж. 1986. *Классификация иллокутивных актов*, НЗЛ. Вып XVII, 170-187.
8. Серль Дж.Р. 1986. Что такое речевой акт?. Зарубежная лингвистика: Пер. с англ. Общ. ред. В.А.Звегинцева, Б.А.Успенского, Б.Ю.Городицкого. НЗЛ. Вып XVII, 151-169.
9. Yaxyoyeva U. I. 2026. O'zbek xalq maqollarida nutqiy akt (A.Qahhor asarlari misolida): filologiya fanlari bo'yicha falsafa doktori (PhD) diss. avtoref. Buxoro.
10. Қаҳҳор, А. 1987. Асарлар. 5 жилдлик. 1-жилд. (Ўғри). Т.: Адабиёт ва санъат нашриёти.
11. Аҳмад, С. 2000. Танланган асарлар: 3 жилдлик. 2-жилд. Ҳажвиялар. Пьесалар. Офтобга чиққан мушук. Т.: Шарк.
12. Qodiriy, A. 2019. *O'tkan kunlar*. Т.:Yangi nashr.
13. Ниёзий, Х. 1988. Тўла асарлар тўплами: 5 жилдлик. 3-жилд: Драмалар. Т.:Фан.
14. Muxtor, O. 1989. *Egilgan bosh*. Т.: G'afur G'ulom nomidagi nashriyot.
15. Sulton, X. 2005. *Ko'ngul ozodadur*. Т.: G'afur G'ulom nomidagi nashriyot- matbaa uyi.

YAKHYOYEVA UMIDA IBROHIM KIZI

ASIA INTERNATIONAL UNIVERSITY,
UZBEKISTAN.

ORCID: -0000-3548-578X

E-MAIL: <UMIDAYAHYOYEVA12@GMAIL.COM>