

**POSSIBILITIES OF REPRESENTING THE STYLISTIC USE OF NOUN
GRAMMATICAL FORMS IN THE CORPUS****Elova Dilrabo Qudratullayevna**

Doctor of Philosophy (PhD) in Philological Sciences,
Associate Professor, Tashkent State University of Uzbek Language and Literature
named after Alisher Navoi
dilraboelova@gmail.com

Abstract. This article analyzes the functional-stylistic characteristics of subjective evaluative forms in the Uzbek language (diminutive, endearing, intensifying, and derogatory nuances) and explores the possibilities of tagging them in language corpora. It identifies the semantic-pragmatic features, connotative meanings, and frequency of use of lexical noun-forming suffixes (-cha, -gina/-kina/-qina, -jon, -xon, -oy, -boy, -bek, -toy, etc.) across various speech styles. The article proposes a two-stage algorithm for linguistic modeling of subjective evaluative forms and develops a system of tags that determine the stylistic markedness of linguistic units. The analysis demonstrates the active use of these forms in artistic and conversational speech, substantiates the stylistic function of simile forms (-dek/-day) and diminutive-endearing forms, and examines their synonymous and polysemantic nature. The results of the study contribute to improving the process of creating functional-stylistic annotations in the National Corpus of the Uzbek Language.

Keywords: subjective evaluative forms, stylistic markedness, linguistic modeling, tagging system, functional styles, affix, affixoids, connotative meaning, Uzbek language corpus, diminutive, endearing forms, simile form.

Introduction. Corpus linguistics is one of the relatively young fields in Uzbek linguistics, and in recent years the number of researchers conducting studies in this direction has significantly increased. During this period, various aspects of the field have been explored, including the principles of constructing an Uzbek authorial corpus [Khamroyeva Sh. *Linguistic foundations of constructing the Uzbek authorial corpus*: PhD dissertation in Philological Sciences. – Karshi, 2018. – 250 p.], the principles of building the linguistic database of a language corpus [Eshmuminov A.A. *The synonym database of the Uzbek National Corpus*: PhD dissertation in Philological Sciences. – Karshi, 2019. – 140 p.], the linguistic foundations of semantic tagging of Uzbek appellative units in corpora [Akhmedova D.B. *Linguistic foundations and models of lexical-semantic tagging of appellative units for Uzbek corpora*: PhD dissertation in Philological Sciences. – Bukhara, 2020. – 247 p.], analyses of English language corpora [Atabayev N. *Functional characteristics of the English corpus (based on COCA)*: PhD author's abstract. – Tashkent, 2020. – 58 p.], and the theoretical foundations of constructing an English-Uzbek parallel corpus [Karimov R., Mengliev B. *Theoretical fundamentals of the Uzbek–English parallel*

corpus. Journal of Critical Reviews, Vol. 7, Issue 17, 2020, pp. 73–76; Karimov R.A., Mengliev B.R. *The role of the parallel corpus in linguistics, its importance and interpretation possibilities*. International Journal of Engineering and Advanced Technology (IJEAT). Vol. 8, Issue 5S3, July 2019, pp. 388–391]. Significant work has also been carried out on developing the stylistic principles and linguistic infrastructure of the Uzbek language corpus [Elova D.Q. *Developing stylistic principles and linguistic support for the units of the Uzbek language corpus*: PhD dissertation in Philological Sciences. – Tashkent, 2022. – 155 p.], creating a parsing system for the Uzbek National Corpus [Khurov O. *Linguistic foundations of developing a parsing system for the National Corpus*: PhD author's abstract. – Jizzakh, 2021. – 55 p.], and building an idiom database for the Uzbek National Corpus [Begmatova G. *Creating an idiom database in the Uzbek National Corpus*: PhD author's abstract. – Termiz, 2021. – 48 p.]. In addition to corpus-related tools, a number of monographic works have emerged in the field of developing software resources indirectly related to corpus construction. These include works on designing linguistic support for computer programs based on motion verbs [Muhammedova S. *Creating linguistic support for computer programs based on motion verbs*. Methodological guide. – Tashkent, 2006], the linguistic infrastructure of the Uzbek morphological analyzer [Khamroyeva Sh. *Linguistic support for the Uzbek morphological analyzer*: Doctoral dissertation in Philological Sciences. – Fergana, 2021. – 76 p.], and the linguistic foundations of the Uzbek semantic analyzer [Gulyamova Sh. *Linguistic foundations of the Uzbek semantic analyzer*: Doctoral dissertation in Philological Sciences. – Fergana, 2022. – 76 p.]. These works collectively contribute significantly to the development of Uzbek corpus linguistics. In addition to software tools, studies dedicated to the linguistic-statistical analysis of the language have also formed a separate direction within Uzbek computational linguistics. The aforementioned research covers issues related to corpus construction, development of its software tools, and preparation of its linguistic support.

Main Part. The presence of annotations in language corpora indicating which style linguistic units belong to and which functional style of the language they represent enhances opportunities for studying the Uzbek language and analyzing its stylistic characteristics. The style of a text is influenced by several factors, including its type (academic, popular-scientific, official-administrative, neutral, vulgar, individual-authorial style, dialectal, etc.), as well as the age and educational level of its intended audience. Therefore, without identifying the theoretical foundations and issues of stylistics and speech culture, it is impossible to reflect the stylistic distinctiveness, expressiveness, and culture of speech within language corpora.

Based on existing classifications of speech styles in Uzbek linguistics, it is advisable to develop stylistic annotations and stylistic search markers for the Uzbek National Corpus. Style-marking tags are closely related to the semantic features of linguistic units, and developing such a tagging system as well as forming its

linguistic database is one of the important tasks in natural language processing. In the future, these will serve as the foundation for implementing automatic text analysis in Uzbek language corpora. For this purpose, the process of linguistically modeling the stylistic features of subjective evaluative forms in Uzbek is analyzed below.

In modern Uzbek, affixes and affixoids such as *-gina* (*-kina*, *-qina*), *-cha*, *-choq*, *-chak*, *-loq*, *-jon*, *-xon*, *-oy*, *-voy* serve as subjective evaluative forms that express the speaker's modal attitude toward objective reality and their own speech [Qilichev E. 1992:21]. These forms convey connotative meanings such as affection, endearment, intensification, and disparagement. Additional meanings carried by subjective evaluative forms play an important stylistic role in literary texts by revealing specific characteristics of characters – such as their tenderness, softness, and likeability. Diminutive and endearing forms also appear in spoken language, and their meanings depend on the overall context of speech.

Synonymy, polysemy, and stylistic belonging of subjective evaluative forms determine their stylistic function [Qilichev E. 1992:21]. The morpheme *-cha* is used uniformly across all functional speech styles. In addition to its objective meaning of “smallness,” it also expresses subjective evaluation. For example, in *qozoncha* (small pot), *taqsimcha* (small plate), *daftarcha* (small notebook), *bolacha* (little child), *yigitcha* (young lad), *xotincha* (little woman), the suffix *-cha* conveys both diminutive and emotional meaning: *kechga qolding*, *tulkicham*, *kech keldimi echkicham?* (from a fairy tale).

Based on the information above, the stylistic tagging of the suffix *-cha* may be represented as follows:

-cha	Stylistically marked	S ₁	B ₁	P ₁	P ₁	I ₁
------	----------------------	----------------	----------------	----------------	----------------	----------------

When the *-cha* morpheme is added to proper nouns, as well as to common nouns denoting positions, professions, or occupations, it emphasizes the speaker's negative emotional attitude:

- *Shunday emasmi, Fayzicha? – deb so‘radi boy. (S.A.)*

- *“Isn’t that so, Fayzicha?” asked the rich man. (S.A.)*

Based on this information, this phenomenon can be modeled as follows:
[proper noun + -cha] = word with negative connotation;

[personal nouns + -cha] = word with negative connotation.

Additionally, when the form **-cha** occurs alongside possessive affixes, it may express irony, sarcasm, or humor, for example: *oyimchamiz*, *xotinchangiz*. This phenomenon can be modeled as follows:

[-cha + possessive form] = irony, sarcasm, humor → conversational style

The representation of this in the database would appear as follows:

-cha +	Possessive forms (-im, -ing, -i, imi, -ingiz, -lari)	Stylistically marked	Sarcasm, mockery, humor	S ₁
-----------	--	----------------------	-------------------------	----------------

The **-cha** morpheme can sometimes be synonymous with other subjective evaluative forms: *qizcha – qizaloq*, *toycha – toychoq*, *kelincha – kelinchak*, and so on.

The forms **-gina** (**-kina**, **-qina**) function in speech both as particles expressing limitation or restriction and as subjective evaluative forms expressing endearment. When **-gina** serves as a modal formative, it is added to nouns before possessive and case suffixes and receives stress. This feature strengthens its modal meaning. The intensifying function of this suffix depends both on the part of speech it attaches to and on the context. When the affixes **-gina** (**-kina**, **-qina**) are attached to nouns, they intensify the speaker's attitude toward the listener: *bo 'yginangdan o 'rgilay*; *Oyijon, oyi-ya*, *boshginam og 'riydi-ya!* (from a song)

Since this phenomenon is mainly characteristic of artistic and conversational styles, it is annotated in the database with the following tag:

-gina (-kina , -qina)	Stylistically marked	S ₁	B ₁
--	----------------------	----------------	----------------

In our language, some words lose their independent meanings and turn into affixoids that express diminutive or endearing meanings. Such affixoids as **-jon**, **-xon**, **-oy**, **-niso**, **-sho**, **-boy** (**-voy**), **-bek**, **-toy**, **-bonu**, **-gul** are widely used in spoken language, especially in literary style. These affixoids express modal meanings such as *affection*, *endearment*, *respect*, and *humor*. Since the additional emotional-expressive meanings in these types of affixoids are close to one another, they form synonymous series: *Dilbarjon – Dilbarxon – Dilbaroy*, *Salimboy – Salimbek – Salimsho*, etc. [Qilichev E. 1992:22]. Such synonymous usage ensures the diversity and richness of speech. Based on the information above, the stylistic characterization of these suffixes can be tagged as follows:

-jon -xon -oy -niso -sho -boy -(voy) -bek -toy -bonu -gul	Stylistically marked	Conversational style	S ₁	Literary style	B ₁
--	----------------------	----------------------	----------------	----------------	----------------

Some forms of subjective evaluation with endearing nuances are added only to personal names or kinship terms, and these affixes mainly express affection or fondness. Suffixes such as *-boy*, *-bek*, *-sho*, *-mirza*, *-qul* are attached to male names, while *-oy*, *-gul*, *-bonu*, *-beka*, *-bibi*, *-niso* are added to female names. Some affixoids of both types form internal synonymous series: *Matlabxon / jon / oy / niso*; *Ergashjon / boy / bek / qul*. We have discussed their stylistic features above. By adding further explanation regarding the expression of affection and fondness, we expand the annotation.

Some endearing affixoids are used less frequently in speech (*-qul*, *-niso*, *-bekach*, *-bonu*). At the same time, in Uzbek, especially in the speech of young people, Russian diminutive forms can sometimes be observed: *Nazarchik* (*Nazar*), *Layloshka* (*Laylo*), *Sayish* (*Saida*), etc. Some Russian forms are also used synonymously with Uzbek endearing forms.

Dedilar bir yig‘inda, kerakmas hech xon demoq,
Xayrixonning o‘rniga Xayrichkasi yaxshiroq.
(They said at a gathering, “No need to say *xon*,
Xayrichkasi is better instead of *Xayrixon*.” (V. Abdullaev)

E. Qilichev emphasizes that the connotative meanings expressed by subjective-evaluation forms must be compatible with the lexical meaning of the base they attach to [Qilichev E. 1985:36]. Otherwise, a stylistic flaw emerges. For example, the word *kichik* (“small”) already contains a semantic indication of small quantity. Therefore, it cannot take the diminutive suffix *-cha* (*kichikcha* is not used). Instead, the affectionate suffix *-kina* is added (*kichkina*). The word *katta* (“big”) also does not take the diminutive suffix *-cha* (*kattacha* is not used), because its semantic degree of largeness does not allow it. This phenomenon applies to all other subjective-evaluation forms. Although the language does not favor unnecessary repetition of the same element, speakers or writers often use repeated forms intentionally for stylistic purposes. Repetition can appear at the lexical level (*ulug‘ ayyom kunlari*; *tuz-namak bo‘ldik*; *bosh-oyoq sarpo qildi*) or at the morphological

level (*ko'rpacha, qizaloqcha, bo'taloqcha*). By layering subjective-evaluation forms, speech becomes more expressive, and the speaker's positive or negative attitude toward the described event is emphasized [Qilichev E. 1992:23].

Thus, repeated subjective-evaluation forms serve stylistic functions. If modeled as *-cha + -cha* or *-loq + -cha*, there is a possibility of becoming similar to other patterns. Therefore, listing all such forms and adding them to an exception base is advisable.

Pleonasm in subjective-evaluation forms arises from the close interconnection between diminutive and affectionate meanings. In such repeated forms, the first affix usually expresses smallness, while the second conveys an affectionate nuance. The connotative meanings inherent to subjective-evaluation forms ensure expressiveness in artistic language, which is why writers effectively use such affixed words in their works. Among the functional forms of nouns, simile forms and paired or reduplicated forms stand out with distinctive stylistic features [Qo'ng'urov R. 1989].

The simile form is created with the suffix *-dek (-day)*, which likens an object or person to the one described by the *-dek* form: *sherdek kuchli* ("strong as a lion"), *asalday shirin* ("sweet as honey"). These simile forms belong mainly to artistic style and are widely used in descriptive expression. Based on these facts, simile forms can be tagged as follows:

<i>-dek / -day</i>	Stylistically marked	Literary style	B ₁
--------------------	----------------------	----------------	----------------

According to their stylistic use, *synthetic forms of adjective degrees, distinguished by their compactness, are primarily used in poetry, whereas analytic forms of adjectives are employed in journalistic and scientific styles*. Modal forms in adjectives also possess certain stylistic features. Modal suffixes do not fundamentally alter the meaning of the adjective they attach to; rather, they add additional emotional-expressive nuances such as *affection, pity, endearment, or irony*. For example:

"Oh Gulsanamim, O'rtoqjonim, mehribonginam." (O.) Men o'z qo'lim bilan shiringina palov qilib beraman. (O.)

(Oh, my dear Gulsanam, my comrade, my sweet one. (O.) I will cook delicious pilaf for you with my own hands. (O.))

The affixes *-gina (-kina, -qina)*, when attached to adjective stems, can convey not only pleasantness or smallness but also negative evaluative nuances: *u qiz xunukkina-ku; Xayr, o'g'rigina bolam, kelib tur* (*That girl is rather ugly; Goodbye, my little thief, come by often*). (G'. G'.)

These data indicate that it is not always possible to represent the stylistic nuances of adjectives directly in a database. In such cases, a modeling approach can be applied:

[*adjective + -gina (-kina, -qina)*] = *journalistic, literary style*]

This demonstrates that certain stylistic features can be identified through modeling and a two-stage algorithm:

1. In the first stage, the stem of the word and the position of its grammatical form are modeled.
2. In the second stage, the stylistic classification is determined based on the algorithm derived from this model.

Conclusion. The stylistic annotation of subjective-evaluation forms was carried out based on theoretical perspectives on speech stylistics in Uzbek linguistics. Although morphological forms are generally used uniformly across different styles, some of them exhibit variants in various functional styles. Certain forms are predominantly used in a specific style and are considered characteristic of that style, while a single form may serve multiple stylistic functions. In such cases, semantic shifts in the form can also occur. This allows for the classification of morphological forms and categories from the perspective of functional styles, which are a central concern in stylistics, and their representation in Uzbek language corpora. In Uzbek, subjective-evaluation forms are important morphological tools that enhance the emotional-expressive potential of speech. During this study, mechanisms for annotating these forms in corpus linguistics were developed, based on their semantic, pragmatic, and stylistic properties. The scope of usage of affixes and affixoids such as *-cha, -gina (-kina, -qina), -jon, -xon, -oy, -boy, -bek, -toy* in different speech styles was determined, and models for their stylistic annotation were proposed. Furthermore, the stylistic function of simile forms (*-dek/-day*), endearing, and intensifying forms in literary language was substantiated. Cases of synonymy, polysemy, and pleonasm in subjective-evaluation forms highlighted the necessity of including certain forms in an exceptions list. The results of this study provide a foundation for creating a system of functional-stylistic tags in the Uzbek National Corpus, improving algorithms for automatic analysis, and offering a robust theoretical basis for stylistic research.

References

1. Khamroeva Sh. *Linguistic Foundations for Creating the Authorship Corpus of the Uzbek Language*. PhD dissertation in Philology – Qarshi, 2018. – 250 pp.
2. Eshmomminov A.A. *Synonym Word Base of the National Corpus of the Uzbek Language*. PhD dissertation in Philology – Qarshi, 2019. – 140 pp.

3. Akhmedova D.B. *Linguistic Foundations and Models for Lexical-Semantic Tagging of Denotative Units in Uzbek Language Corpora*. PhD dissertation in Philology – Bukhara, 2020. – 247 pp.
4. Atabayev N. *Functional Features of the English Language Corpus (Using the SOSA Example)*. PhD dissertation abstract in Philology – Tashkent, 2020. – 58 pp.
5. Karimov R., Mengliev B. *Theoretical Fundamentals of the Uzbek-English Parallel Corpus*. Journal of Critical Reviews, ISSN 2394-5125, Vol. 7, Issue 17, 2020, pp. 73-76.
Karimov R.A., Mengliev B.R. *The Role of the Parallel Corpus in Linguistics: Importance and Possibilities of Interpretation*. International Journal of Engineering and Advanced Technology (IJEAT), ISSN 2249–8958, Vol. 8, Issue 5S3, July 2019, pp. 388-391.
6. Elova D.Q. *Principles for Creating Stylistic Features of Units in the Uzbek Language Corpus and Their Linguistic Support*. PhD dissertation in Philology – Tashkent, 2022. – 155 pp.
7. Khidirov O. *Linguistic Foundations for Creating a Parsing Program for the National Corpus*. PhD dissertation abstract in Philology – Jizzakh, 2021. – 55 pp.
8. Begmatova G. *Creating an Idiom Database in the Uzbek National Corpus*. PhD dissertation abstract in Philology – Termez, 2021. – 48 pp.
9. Muhammadova S. *Creating Linguistic Support for Computer Programs Based on Action Verbs: Methodological Guide*. – Tashkent, 2006.
10. Khamroeva Sh. *Linguistic Support for the Morphological Analyzer of the Uzbek Language*. PhD dissertation in Philology – Fergana, 2021. – 76 pp.
11. Gulyamova Sh. *Linguistic Foundations of the Semantic Analyzer of the Uzbek Language*. PhD dissertation abstract in Philology – Fergana, 2022. – 76 pp.
12. Shomaksudov A., Rasulov I., Qongurov R., Rustamov H. *Stylistics of the Uzbek Language: A Guide for University and Pedagogical Institutes, Faculty of Philology*. Tashkent: O‘qituvchi, 1983.
13. Mamajonov A., Mahmudov U. *Repetition as a Stylistic Device*. “Uzbek Language and Literature” Journal, 1997.
14. Vinogradov V.V. *Stylistics: Theory of Poetic Speech. Poetics*. Moscow, 1963. – p. 593
15. Tozhiev Y., Khasanova N., Tozhimatov Kh. *Culture of Uzbek Speech and Foundations of Stylistics*. Tashkent, 1994.
16. Lomonosov M.V. *A Brief Guide to Rhetoric*. 1743.
17. Qilichev E. *Practical Stylistics of the Uzbek Language (Grammar)*. Textbook – Tashkent: O‘qituvchi, 1992. – 160 pp.
18. Qilichev E. *Practical Stylistics of the Uzbek Language*. Textbook – Tashkent: O‘qituvchi, 1985. – 104 pp.
19. Qongurov R. *Semantic and Stylistic Features of Subjective-Evaluation Forms*. Monograph – Tashkent: Fan, 1989.

20. Garside R., Leech G. *Corpus Annotation: Linguistic Information from Computer Text Corpora*. Longman, 1997.