

MORPHOLOGICAL FEATURES OF THE PHENOMENON OF PREFIXATION IN LOANWORDS OF THE UZBEK LANGUAGE

Narzullayeva Muborak Sherzod qizi

Independent Researcher, Gulistan State University, Uzbekistan

E-mail: narzullayevamuboraksherzodovna@gmail.com

Abstract

The article describes the morphological features of prefixation in the loanword stratum of modern Uzbek. All prefixal elements of the language are of foreign origin, and they entered together with the Arabic, Persian and Russian lexical layers. The study establishes five structural criteria of prefixal status for a suffixing language, applies them to the whole inventory, and shows that the elements of that inventory occupy different degrees of morphological integration. Three types of evidence are used. The first is the behaviour of the prefixed complex under inflection and further derivation, where forms such as *noaniqlik*, *hamkorlashmoq* and *g'ayriqonuniylik* prove that the complex functions as a single stem. The second is combinability with native Turkic stems, which distinguishes a prefix of Uzbek from a foreign element carried in by a ready-made loanword. The third is morphonological behaviour at the boundary, where the absence of harmony and of assimilation and the tolerance of vowel sequences of the type *noaniq* mark the prefix boundary as stronger than any suffix boundary of the language. The article also examines the competition between prefixation and the native suffixes *-siz*, *-li* and *-dosh*, which produces doublets such as *besuv* and *suvsiz* or *hamkasb* and *kasbdosh*, and it proposes a scale of analysability for loan stems that begin with a prefix-like syllable.

Keywords: Prefixation, loanwords, Uzbek language, morphology, affix criteria, hybrid formation, morphological integration, morphonology, word formation, doublets.

Introduction

Uzbek forms words by suffixation. Derivational and inflectional morphemes follow the stem in a fixed order, and this order is one of the defining properties of the Turkic morphological type [1]. Prefixation nevertheless exists in the language, and its whole inventory consists of elements taken from other languages: from Persian and Tajik, from Arabic, and from Russian together with the international vocabulary transmitted through it. The phenomenon is therefore located entirely within the loanword stratum, and its description belongs simultaneously to word formation theory and to contact morphology.

The theoretical difficulty is that the notion of a prefix is defined for the general case by position and by boundness [2]. A prefix is a bound morpheme that precedes the stem and contributes a regular meaning. In a language whose native morphology contains no prefixes, this definition does not by itself decide the concrete cases. A word borrowed as a whole from Persian also has a bound element in initial position, but for the borrowing language this element may be no more than a sequence of sounds. The task of a morphological description is to distinguish the two situations, and this requires criteria that go beyond position and boundness.

The aim of the article is to establish the morphological features of prefixation in Uzbek loanwords and to determine the degree of integration of each part of the inventory. The tasks are the following: to formulate structural criteria of prefixal status applicable to a suffixing language; to classify the inventory by origin, productivity and semantic function; to examine the behaviour of prefixed complexes under Uzbek inflection and derivation; to describe the morphonological properties of the prefix boundary; to analyse the competition between prefixal and suffixal expression of the same meanings; and to build a scale of analysability for stems that only appear to contain a prefix. The material consists of prefixed items of the modern literary language recorded in explanatory dictionaries, in terminological vocabularies and in the language of the press.

2. Criteria of prefixal status in a suffixing language

Five criteria are used in this study, and they are applied cumulatively rather than alternatively. The first is non-autonomy. The element must not occur as an independent word with the same meaning. The elements *be-*, *no-*, *ser-*, *g'ayri-* and

bad- satisfy this criterion without reservation, since none of them is a word of Uzbek. The elements ham and kam do not, because both exist as free items: ham is a particle meaning also, as in men ham bordim for I also went, and kam is an adjective meaning few or insufficient. These two are therefore transitional cases. The second criterion is series formation. A prefix must occur in an open set of formations with a constant meaning, not in two or three isolated items. By this test be-, no-, ser-, ham-, g'ayri-, anti- and mikro- are prefixes, while lo- in loqayd for indifferent and in loaqal for at least is not, since it appears only in a closed and archaic group. The third criterion is combinability with native stems. If an element attaches to a Turkic stem, the resulting word cannot have been borrowed as a whole and must have been made inside Uzbek. The fourth criterion is the behaviour of the complex under further morphology: Uzbek suffixes must attach to the prefixed word as to any other stem. The fifth is semantic regularity, that is, the predictability of the meaning of the whole from the meaning of its parts [3]. These criteria produce a graded rather than a binary result, which corresponds to what is known about affix status in other languages, where the boundary between an affix and a compound member is likewise a matter of degree [4]. The following sections apply the criteria to the material.

3. The inventory by origin, productivity and semantic function

By origin the inventory falls into three groups. The Persian and Tajik group contains be-, no-, ser-, ba-, bo-, ham-, kam-, bad- and xush-. The Arabic group contains g'ayri- and the unproductive lo- and bilo-. The Russian and international group contains anti-, kontr-, re-, de-, dis-, sub-, super-, ultra-, ekstra-, trans-, post-, pro-, mikro-, makro- and, in the vocabulary of the last three decades, kiber-, veb- and elektron-.

By productivity three degrees are distinguished. Highly productive elements form new words freely in the present-day language: be-, no-, ser-, ham- and, within terminology, anti-, mikro-, makro- and post-. Elements of limited productivity form a defined but non-expanding set: ba-, bo-, kam-, bad-, xush- and g'ayri-. Unproductive elements appear only in inherited items: lo- and bilo-. Productivity is measured here by the presence of formations created after the middle of the twentieth century, which is a conservative criterion but a verifiable one.

By semantic function five classes emerge. The privative and negative class contains be-, no- and g'ayri-, together with the international anti-, de- and dis-. The class of abundance contains ser-. The comitative and collective class contains ham-. The evaluative class contains bad- and xush-, and it differs from the others in that the prefix carries an assessment rather than a logical operation. The quantitative and dimensional class contains kam-, mikro-, makro-, super- and ultra-. A separate spatial and temporal group is formed by trans-, post-, sub- and pro-, which belongs entirely to the international layer and to terminology [5]. The classes are not equally represented in general usage: privative formations account for the greater part of the prefixed vocabulary of the everyday language, whereas the spatial and temporal group is almost confined to written specialist registers. A remark on the boundaries of the inventory is necessary before the morphological analysis. Not every element that appears in initial position in the international vocabulary is a prefix, and several are combining forms, that is, bound stems of Greek or Latin origin that function as the first member of a compound. The elements avia-, tele-, foto-, agro- and bio- belong to this group. They differ from prefixes in that they carry lexical rather than operational meaning, and they differ from ordinary compound members in that they have no independent use. Uzbek treats them exactly as it treats the international prefixes, which is one reason why descriptions of the language often place the two groups together. The present study keeps them apart, since the criterion of semantic regularity applies to them in a different way: a prefix modifies the meaning of the stem, whereas a combining form contributes a second lexical component.

4. Morphological integration: inflection and further derivation

The strongest morphological evidence that a prefixed loanword is a single stem for Uzbek is the behaviour of the whole complex under suffixation. If the prefix were a separate element, suffixes would be expected to attach inside the complex or to be blocked by it. Neither happens. The abstract noun suffix -lik attaches to the outer edge: noaniq yields noaniqlik for imprecision, beg'am yields beg'amlik for carefreeness, serhosil yields serhosillik for fruitfulness, hamkor yields hamkorlik for cooperation and g'ayriqonuniy yields g'ayriqonuniylik for unlawfulness. The verbalising suffixes behave in the same way: hamkor yields

hamkorlashmoq for to cooperate, and hamfikr yields hamfikr bo'lmoq in the analytical construction.

Case and number affixes give the same result. Hamkor forms hamkorlar, hamkorlarimiz and hamkorlarimizga without any interference from the prefix. Adverbial derivatives are formed on the whole complex as well: bekor gives bekorga for in vain and bekorchi for an idler, behuda gives behudaga. Comparison operates over the complex too, since noaniqroq for more imprecise is regularly formed and eng noaniq for the most imprecise is the normal superlative construction.

Two conclusions follow. First, for Uzbek morphology the prefixed loanword is a stem and not a syntactic group, which distinguishes prefixation from the analytical constructions with which it competes, such as aniq emas for not precise. Second, the prefix is outside the domain of Uzbek suffixation but inside the domain of the word, which is exactly the position occupied by prefixes in languages that possess them natively [2]. The Uzbek system has thus accommodated a foreign morphological element without altering the architecture of its own suffix chain.

5. Word class of the output

Prefixation in Uzbek is predominantly adjective-forming, which distinguishes it sharply from suffixation, where nominal, verbal and adjectival derivation are all well represented. The privative be- attaches to nouns and produces adjectives: pul for money gives bepul for free of charge, suv for water gives besuv for waterless, g'am for sorrow gives beg'am for carefree, tob for strength gives betob for unwell. The negative no- attaches to adjectives and to adjectival nouns and preserves the class: to'g'ri for correct gives noto'g'ri for incorrect, aniq for precise gives noaniq for imprecise, haq for just gives nohaq for unjust, qonuniy for lawful gives noqonuniy for unlawful. The element ser- attaches to nouns and produces adjectives of abundance: hosil for harvest gives serhosil for fruitful, gap for talk gives sergap for talkative, quyosh for sun gives serquyosh for sunny.

The comitative ham- is the principal exception, since it produces nouns denoting persons: kor for work gives hamkor for partner, suhbat for conversation gives hamsuhbat for interlocutor, kasb for profession gives hamkasb for colleague, yurt for homeland gives hamyurt for compatriot, fikr for thought gives hamfikr for a

like-minded person. The evaluative elements are mixed: xushbo'y for fragrant and badxulq for ill-tempered are adjectives, whereas xushxabar for good news is a noun. The international elements are mostly noun-forming when attached to nouns, as in antivirus, mikroiqlim and postmodernizm, and adjective-forming when attached to adjectives, as in transmilliy for transnational and antiinsoniy for inhuman.

The distribution has a systemic explanation. The meanings expressed by the productive prefixes, that is, absence, negation and abundance, are prototypically property meanings, and property meanings are expressed by adjectives. The native competitors of these prefixes, the suffixes -siz, -li and -dor, are likewise adjective-forming, which places prefixation and suffixation in direct competition within one and the same word class [6].

6. Combinability and hybrid formation

Combinability with native Turkic stems is the decisive test of integration, because a hybrid formation cannot be a borrowing. It does not exist in the donor language and must have been produced by the rules of the recipient language. Uzbek has a substantial set of such hybrids. With no-: noto'g'ri, where to'g'ri for straight and correct is Turkic; noo'rin for inappropriate, where o'rin for place is Turkic; notinch for restless, where tinch for calm is Turkic. With be-: beto'xtov for uninterrupted, on the Turkic verb to'xta for to stop; bepoyon for boundless, on the Turkic poyon for end in its Uzbek use. With ser-: sersuv for well watered, seryomg'ir for rainy, serqatnov for busy with traffic, all on Turkic stems. With ham-: hamyurt for compatriot, on the Turkic yurt for homeland.

The international elements are at an earlier stage of the same process. The majority still attach only to stems of the same origin, as in antivirus, mikro kredit and postmodernizm, where both parts are international. A growing set, however, has crossed the boundary: infratuzilma for infrastructure contains the Turkic tuzilma for structure, avtoyo'l for motorway contains the Turkic yo'l for road, teleko'rsatuv for television programme contains the Turkic ko'rsatuv for showing, and kiberxavfsizlik for cybersecurity contains a stem with the Uzbek suffix -lik. These formations show that the international elements are beginning to function as word-forming means of Uzbek and not merely as parts of imported terms.

Combinability is also restricted from the side of the stem. The prefix *be-* normally requires a noun denoting a possessible entity, so *bepul* and *besuv* are well formed while an attachment to a verbal noun is not. The prefix *no-* prefers stems with an evaluative or a qualitative component, which explains the acceptability of *noto'g'ri* and the oddity of an attachment to a concrete object noun. These restrictions are semantic rather than formal, and they are one of the reasons why prefix choice is a frequent source of error in unedited texts.

7. Morphological features of the prefix boundary

The prefix boundary in Uzbek behaves differently from the suffix boundary, and the difference is systematic. The first feature is invariance. Uzbek suffixes historically alternated according to the vowel and the final consonant of the stem, and traces of this alternation remain in the literary language. Prefixes show no alternation at all: *be-*, *no-* and *ser-* have one shape before any stem, whatever its vowels [7].

The second feature is the tolerance of vowel sequences. Native Uzbek phonotactics avoids two adjacent vowels within a word, and at suffix boundaries the sequence is normally resolved. At the prefix boundary it is retained: *no* plus *aniq* gives *noaniq*, *be* plus *adab* gives *beadab*, *ser* plus *unum* gives *serunum*, *ham* plus *ohang* gives *hamohang* for harmonious. The retention of these sequences shows that the boundary is phonologically strong and that the two parts are processed as distinct units even though they form one grammatical word.

The third feature is the absence of the consonant adjustments found elsewhere in the morphology. There is no voicing assimilation and no degemination across the prefix boundary, and the final consonant of the prefix and the initial consonant of the stem remain unchanged in all combinations. The fourth feature concerns stress. Uzbek stress falls on the final syllable of the word, and prefixed words follow the general rule, so that the prefix is unstressed in neutral pronunciation. In contrastive use the prefix can attract emphatic stress, as when a speaker opposes *noqonuniy* to *qonuniy*, but this is a phrasal phenomenon and not a lexical property. This last observation is presented as a probable generalisation, since instrumental data on Uzbek prefix stress are limited [8].

A further morphological property distinguishes the prefix from the suffix in Uzbek, and it concerns coordination. Turkic languages permit suffix suspension,

that is, the omission of a shared suffix on all but the last member of a coordinated series, so that a case or plural affix marked once at the end of the series is understood over the whole of it. Prefixes do not permit the corresponding operation. A shared prefix must be repeated on each member of a coordinated pair, and an attempt to mark it once produces either an ungrammatical string or a reading in which only the marked member is affected. The asymmetry follows from the position of the elements relative to the stem and from the fact that suspension operates at the right edge of the phrase, but it is worth recording as an independent piece of evidence that Uzbek treats its prefixes and its suffixes as different kinds of morphological object. This generalisation is presented as probable, since it rests on constructed examples and native speaker judgement rather than on corpus counts.

8. Competition between prefixation and native suffixation

Because the productive prefixes express privation, negation and abundance, they compete with native suffixes that express the same meanings. The suffix *-siz* forms privative adjectives, the suffixes *-li* and *-dor* form adjectives of possession and abundance, and the suffix *-dosh* forms nouns of joint participation. The competition produces doublets, and the doublets are informative because they show what prefixation adds to the system.

Three types of relation are observed. In the first type both forms exist and are close in meaning: *besuv* and *suvsiz* for waterless, *beg'am* and *g'amsiz* for carefree, *hamkasb* and *kasbdosh* for colleague, *hamyurt* and *yurtdosh* for compatriot, *hamfikir* and *fikrdosh* for a like-minded person, *hamsuhbat* and *suhbatdosh* for interlocutor. In such pairs the prefixal member usually belongs to the bookish and the literary register and the suffixal member to the neutral or colloquial one. This is a stylistic and not a denotational difference, and it is a probable rather than an established generalisation, since no systematic register study of these pairs exists. In the second type only the prefixal form exists, because the stem has no independent use in Uzbek or because the suffixal formation is blocked: *behuda* for futile has no suffixal counterpart, and the same holds for *bemalol* for unhindered. In the third type the two forms have diverged in meaning, which is the most interesting case for lexicography and the most dangerous one for editing. The pair *noqonuniy* and *g'ayriqonuniy* is the clearest example. Both are translated

into English as unlawful, but in the language of Uzbek legislation they are not interchangeable in every context, and the difference is one of degree of formality and of established terminological usage rather than of logical content. A translator who substitutes one for the other in a statutory text changes a term, not a stylistic register [9].

The general conclusion of this section is that prefixation in Uzbek is not a redundant duplication of native means. It supplies a second series of formations with its own register value, it fills gaps where the suffixal pattern is blocked, and it supports terminological differentiation where two closely related notions must be kept apart.

Productivity in this domain is difficult to measure with the usual instruments, and the difficulty deserves a remark. Type counts in a dictionary reflect the accumulated history of the language rather than its present activity, and they will always favour be- and no-, which have been accumulating items since the Chagatai period. Frequency counts in a corpus reflect the register composition of the corpus, and since prefixed formations are concentrated in written and specialist registers, a corpus built on the press will overstate the international elements and understate the older ones. The criterion adopted here, the presence of formations created after the middle of the twentieth century, avoids both distortions but requires dated lexicographic material, which exists for Uzbek only in part. A reliable productivity ranking for the whole inventory therefore remains a task for future work with a diachronically stratified corpus.

9. Prefixoids and native semi-prefixes

Beside the borrowed prefixes proper, modern Uzbek contains a set of elements that occupy the same position and perform the same function but are native or fully assimilated words. They are conventionally called prefixoids, and their treatment is necessary here because they show what the borrowed model produced inside the recipient language.

The element yarim for half forms an open series in terminology: yarimo'tkazgich for semiconductor, yarimavtomat for semiautomatic device, yarimorol for peninsula, yarimfabrikat for semifinished product. The element umum for general forms umumxalq for nationwide, umummilliy for national and umumdavlat for state-wide. The element ko'p for many forms ko'pburchak for polygon, ko'pqirrali

for many-sided and ko'pfunksiyali for multifunctional. The element o'z for self forms o'zboshimcha for arbitrary and o'zaro for mutual. Each of these elements is also a free word of Uzbek with the same meaning, which by the first criterion of section two excludes them from the class of prefixes. Each of them nevertheless forms an open series with a constant and partly bleached meaning, attaches to stems of any origin and yields a single stem for further suffixation, which by the remaining four criteria places them close to prefixes.

The origin of this set is the decisive point. The series listed above are for the most part calques of Russian and international models: the yarim series corresponds to formations with polu- and semi-, the umum series to formations with obshche-, the ko'p series to formations with mnogo- and multi-, and the o'z series to formations with samo- and auto-. Uzbek therefore borrowed from the international layer not only the prefixal elements themselves but also the structural pattern of left-hand modification, and it then filled that pattern with native material. This is a case of pattern borrowing as distinct from matter borrowing, and it is the strongest available evidence that prefixation has become a productive resource of Uzbek word formation and not merely a property of imported vocabulary [11].

The practical consequence concerns classification. A description that recognises only borrowed elements as prefixes understates the extent of the phenomenon, because the terminological vocabulary created since the middle of the twentieth century relies heavily on the prefixoid series. A description that treats every left-hand element as a prefix overstates it, because it erases the difference between a bound morpheme and the first member of a compound. The graded treatment adopted here, in which the five criteria are applied cumulatively and the result is a position on a scale, avoids both errors and reflects the actual state of the material.

10. Degrees of analysability of loan stems

A large group of Uzbek loanwords begins with a syllable identical to a prefix but is not prefixed in the modern language. The description of prefixation is incomplete without a scale that separates these cases, since the same surface string requires different morphological treatment.

The first degree is full analysability. The prefix is productive, the stem is a free word of Uzbek, and the meaning of the whole is predictable: bepul, besuv, beg'am, noaniq, noto'g'ri, serhosil, hamkor. The second degree is partial analysability. The prefix is recognisable and the stem exists in Uzbek, but the meaning of the whole is not fully derivable from the parts: bechora for a poor unfortunate person contains chora for remedy, yet the modern meaning is lexicalised; bemalol for unhindered contains a stem of restricted use. The third degree is etymological analysability only. The whole word was borrowed with its prefix, the stem does not exist in Uzbek, and no synchronic segmentation is possible: bemor for an ill person is a Persian item in which the second part is not a stem of Uzbek [10].

The fourth degree is false analysability, which is the practically important one. Here the initial syllable coincides with a prefix by accident and any segmentation is an error. Nozik for delicate, notiq for orator and nomzod for candidate begin with a sequence identical to the prefix no-, but none of them contains it: the first two continue Persian and Arabic stems and the third is a Persian compound of a different structure. Nonushta for breakfast begins with the noun non for bread. Antiqua for antique and quaint has nothing to do with anti-. Sertifikat for certificate has nothing to do with ser-. Bezak for ornament and beshik for cradle are Turkic words that begin with the sequence be [11].

The distinction has direct applications. In lexicography it determines whether an item is entered under the stem or as an independent headword. In language teaching it determines what may be presented as a productive model. In automatic morphological analysis it determines the content of the exception list, since a stemmer that strips be- from bemor produces a non-existent stem and corrupts every operation built on it, from search to machine translation.

A final observation concerns the direction of change. The three groups of the inventory entered at different periods and are at different stages of integration, but the movement is uniform in direction: from restricted combinability towards free combinability, and from the vocabulary of one register towards general use. Nothing in the material indicates a movement in the opposite direction, apart from the loss of individual archaic elements such as lo- and bilo-, which is a lexical rather than a systemic change. If this direction continues, the international elements will occupy in the twenty-first century the position that the Persian

elements reached in the Chagatai period, and the number of prefixes of Uzbek in the full structural sense will increase.

11. Conclusion

Prefixation in Uzbek is a phenomenon of the loanword stratum, but it is no longer a foreign phenomenon in morphological terms. The five criteria applied in this study give a graded result. The elements *be-*, *no-* and *ser-* satisfy all five: they are non-autonomous, they form open series, they attach to native Turkic stems, they allow the whole complex to take Uzbek suffixes, and they contribute a predictable meaning. The elements *ham-* and *kam-* fail the criterion of non-autonomy and occupy a transitional position between prefixation and compounding. The element *g'ayri-* and most of the international elements satisfy all criteria except free combinability with native stems, and several of the latter are now crossing that boundary, as *infratuzilma*, *avtoyo'l* and *teleko'rsatuv* show.

The morphological behaviour of prefixed loanwords is uniform where it matters. Inflection and further derivation always apply to the outer edge of the complex, which proves that the complex is a single stem for Uzbek. The prefix boundary is nevertheless phonologically stronger than any suffix boundary of the language: it shows no alternation, it tolerates vowel sequences that native phonotactics avoids elsewhere, and it triggers no consonant adjustment. Uzbek has therefore integrated prefixation at the level of the word while keeping it outside the phonological processes of its own affixation.

Two results have practical value beyond the theoretical description. The first is the account of the competition between prefixation and the suffixes *-siz*, *-li*, *-dor* and *-dosh*, which shows that the doublets are distributed by register and by terminological convention rather than by logical content, and that a substitution inside a pair such as *noqonuniy* and *g'ayriqonuniy* is a terminological decision. The second is the four-degree scale of analysability, which separates genuine prefixed formations from lexicalised and from falsely segmentable items and thereby defines the exception list that lexicography, teaching and automatic analysis all require.

References

1. Johanson L., Csato E. A. (eds.). The Turkic Languages. London: Routledge, 1998. 474 p.
2. Haspelmath M., Sims A. D. Understanding Morphology. 2nd ed. London: Hodder Education, 2010. 366 p.
3. Booij G. The Grammar of Words: An Introduction to Linguistic Morphology. 3rd ed. Oxford: Oxford University Press, 2012. 353 p.
4. Bauer L. English Word-Formation. Cambridge: Cambridge University Press, 1983. 311 p.
5. Plag I. Word-Formation in English. Cambridge: Cambridge University Press, 2003. 240 p.
6. Hojiyev A. O'zbek tilida so'z yasalishi. Toshkent: O'qituvchi, 1989.
7. Sjoberg A. F. Uzbek Structural Grammar. Bloomington: Indiana University Press, 1963. 158 p.
8. Bodrogligeti A. J. E. A Grammar of Modern Literary Uzbek. Munich: LINCOM Europa, 2003. 480 p.
9. Kononov A. N. Grammatika sovremennogo uzbekskogo literaturnogo yazyka. Moscow, Leningrad: Izdatelstvo AN SSSR, 1960. 446 p.
10. Rahmatullayev Sh. O'zbek tilining etimologik lug'ati. Toshkent: Universitet, 2000.
11. Haspelmath M., Tadmor U. (eds.). Loanwords in the World's Languages: A Comparative Handbook. Berlin: De Gruyter Mouton, 2009. 1081 p.