
**Tajik-Uzbek Contact
Morphophonology: Case-Marker
Variation and Persian-Tajik
Derivational Affixes in Uzbek**

Khasanov Maxmudjan Abdurahmanovich
makhmud.khasanov@gmail.com
Senior teacher,
Uzbekistan State World Languages University

Khayrullayeva Dilorom Sayfutdinovna
khayrullayevad@gmail.com
Senior teacher,
Uzbekistan State World Languages University

Annotation *This article analyzes morphophonological variation in Tajik-Uzbek contact settings. It describes (i) interference-driven competition between Uzbek case markers in bilingual zones and (ii) the integration of Persian-Tajik derivational affixes and affixoids into Uzbek word formation. Using a comparative descriptive approach, the study shows how contact weakens strict native conditioning and creates hybrid formations shaped by both donor patterns and Uzbek phonotactics.*

Keywords *Morphophonology, language contact, allomorphy, Uzbek, Tajik, derivational affixes, case marking, NLP*

**Таджикско-узбекская
контактная морфонология:
вариативность падежных
показателей и персидско-
таджикские
словообразовательные
аффиксы в узбекском языке**

Хасанов Махмуджан Абдурахманович
makhmud.khasanov@gmail.com
Старший преподаватель,
Узбекский государственный университет
мировых языков

Хайруллаева Дилором Сайфутдиновна
khayrullayevad@gmail.com
Старший преподаватель,
Узбекский государственный университет
мировых языков

Аннотация *В статье рассматривается морфонологическая вариативность в условиях таджикско-узбекского контакта. Описываются (1) интерференционные сдвиги в употреблении падежных показателей и (2) интеграция персидско-таджикских словообразовательных аффиксов/аффиксоидов в узбекскую систему. Показано, что контакт ослабляет строгость родных правил распределения и создаёт гибридные образования, подчиняющиеся как донорским моделям, так и узбекской фонотактике.*

Ключевые слова *Языковой контакт, интерференция, узбекско-таджикский билингвизм, заимствованные аффиксы, морфонологическая адаптация, гибридное словообразование, узбекские диалекты*

Tojik-o'zbek tillararo o'zaro ta'sir morfonologiyasi: kelishik ko'rsatkichlarining variantlanishi va o'zbek tilidagi fors-tojik so'z yasovchi affikslar

Xasanov Maxmudjan Abdurahmanovich
makhmud.khasanov@gmail.com

*Katta o'qituvchi,
O'zbekiston davlat jahon tillari universiteti*

Xayrullayeva Dilorom Sayfutdinovna
khayrullayevad@gmail.com

*Katta o'qituvchi,
O'zbekiston davlat jahon tillari universiteti*

Annotatsiya *Maqolada tojik-o'zbek bilingv hududlarida yuzaga keladigan morfonologik o'zgarishlar tahlil qilinadi. Asosiy e'tibor (1) kelishik ko'rsatkichlarida interferensiya va funksional raqobat, (2) fors-tojik affikslari hamda affiksoidlarning o'zbek tiliga kirib, fonotaktik moslashuv asosida integratsiyalashuviga qaratiladi. Tadqiqot kontakt sharoitida allomorf tanlovi va o'zak-qo'shimcha chegarasidagi moslashuvlar kengayishini ko'rsatadi.*

Kalit so'zlar *Til kontakti, interferensiya, o'zbek-tojik ikki tilliligi, o'zlashgan affikslar, morfonologik moslashuv, gibrid yasama, o'zbek shevalari*

Introduction

Morphophonology is the interface domain where morphological structure systematically conditions phonological alternations and yields predictable surface variation (Haspelmath & Sims, 2010). In language-contact settings, this interface becomes a sensitive indicator of convergence: speakers may replicate distributional patterns, extend existing alternations to new environments, or introduce new affixal material whose surface shape is reinterpreted under the receiving language's phonology (Matras, 2009; Thomason & Kaufman, 1988).

For Uzbek, long-term contact with Tajik (Persian) varieties is well documented, especially in historically bilingual regions such as Bukhara and Samarkand. A reference manual of standard Uzbek notes that Uzbek, while based on urban Turkic dialects, contains elements borrowed from local Persian dialects, with noticeable impact on the lexicon and the vowel system (Corps, 1992). Beyond lexical borrowing, contact may also affect

morphological behavior and morphophonological boundary choices in bilingual speech communities.

This paper focuses on Tajik-Uzbek contact morphophonology in two domains: (i) interference patterns in case marking and boundary alternations in bilingual dialect zones, and (ii) Persian-Tajik derivational affixes and affix-like elements that enter Uzbek and become morphophonologically integrated.

Literature review

Research on Uzbek morphophonology has traditionally emphasized regular alternations at morpheme boundaries and the systematic relationship between invariant morphemic meaning and variant surface shapes. In Uzbek phonology-and-morphophonology scholarship, boundary alternations are often treated as rule-governed and therefore suitable for abstract representation (Abduazizov, 1992; Nurmonov, 1990). Within Uzbek morphological theory, the description of form-building (*forma yasalishi*) highlights how suffixal attachment can trigger

phonetic adjustments and yield stable paradigmatic patterns rather than isolated exceptions (Hojiyev, 1979). Pedagogical descriptions and manuals further systematize alternations and affix variants for teaching purposes, reinforcing the idea that allomorphy is predictable once conditioning environments are made explicit (Rahmatullayev, 2006; Zokirova, 2016).

In general contact theory, morphological change is expected to be shaped both by sociolinguistic intensity and by structural compatibility between languages (Thomason & Kaufman, 1988). Modern contact linguistics underscores that morphology is not immune to borrowing and pattern replication: derivational morphology in particular may be transferred and subsequently adapted to recipient-language phonotactics (Gardani, Arkadiev, & Amiridze, 2015; Matras, 2009). For Turkic languages, contact studies note that bilingual regions can show functional overlap of grammatical markers and the gradual reanalysis of morphosyntactic distributions (Johanson, 2015).

Uzbek-specific work on Tajik-Persian material in Uzbek word formation documents a sizable inventory of borrowed affixes/affixoids (e.g., *ser-*, *-xon*, *-boz*, *-gar*, *-bon*) and discusses their productivity and stylistic distribution (Ermатов & Dehqonova, 2021). Other studies describe the borrowing of prefixes and suffixes from Persian-Tajik sources into Uzbek and note their role in expanding derivational patterns (Alimova, 2023). Dictionary-based analyses also highlight individual borrowed formatives (e.g., the diminutive *-ak*) and demonstrate their conventionalization in Uzbek lexicographic material (Abduvaliev, 2023).

Finally, recent Uzbek NLP research has increased the practical relevance of morphophonological description. Rule-based or hybrid approaches to stemming, lemmatization, and morphological analysis must either encode boundary alternations explicitly or normalize them during preprocessing (Salaev, 2024; Sharipov & Salaev,

2022). As a result, contact-driven variation and hybrid formations represent an additional challenge for robust analyzers designed for dialectal or noisy user-generated data.

Materials and Methods

The analysis is comparative and descriptive. We juxtapose (a) native Uzbek morphophonological tendencies (allomorph selection and boundary alternations) with (b) contact-related variants and borrowed derivational material described in published manuals and linguistic studies.

The theoretical framing follows standard definitions of allomorphy and morphophonology (Haspelmath & Sims, 2010) and Uzbek scholarship that treats boundary alternations as systematizable (Abduazizov, 1992; Nurmonov, 1990). For borrowed affixal material, the analysis draws on studies of Tajik-Persian affixes and affixoids in Uzbek word formation (Alimova, 2023; Ermатов & Dehqonova, 2021; Abduvaliev, 2023). Examples are presented as illustrative schemata typical for descriptive morphophonology; they are intended to show structural possibilities rather than provide frequency counts.

Result

Interference in case marking and boundary choices

In native Uzbek, case marking is normally realized through suffixation with well-known patterns of phonological accommodation at morpheme boundaries (Abduazizov, 1992; Rahmatullayev, 2006). In contact zones, however, bilingual practice may introduce functional overlap and competing choices between markers. A salient pattern is the extension of locative marking (*-da*) into contexts that are directional in Standard Uzbek, where the dative (*-ga*) would be expected. This results in competition that is not fully predicted by the native phonological conditioning of suffix shape.

Illustrative examples (contact-driven variation; schematic):

- 1) *uy + ga* → *uyga* 'to home' (standard) | *uy + da* → *uyda* (contact/variation);

- 2) *bozor + ga* → *bozorga* 'to the market' (standard) | *bozor + da* → *bozorda* (contact/variation);
- 3) *Toshkent + ga* → *Toshkentga* 'to Tashkent' (standard) | *Toshkent + da* → *Toshkentda* (contact/variation).

From a morphophonological viewpoint, the system gains additional surface variants whose distribution depends on bilingual norms and pragmatic/functional reanalysis, not only on segmental environment.

Persian-Tajik derivational affixes and their Uzbek integration

A central contact-related mechanism is the borrowing of derivational affixes (or affix-like elements) that attach to Uzbek stems and participate in Uzbek morphophonology. Studies of Uzbek word formation list numerous Tajik-Persian affixes/affixoids used in Uzbek (Ermatov & Dehqonova, 2021) and describe the borrowing of prefixes and suffixes from Persian-Tajik languages (Alimova, 2023). Such elements may become productive, especially in learned, bookish, or stylistically marked vocabulary.

Borrowed affixes often undergo phonological accommodation. Even when a borrowed suffix preserves its overall shape, the boundary between an Uzbek stem and a Persian affix may trigger familiar Uzbek adjustments (assimilation, epenthesis, voicing), consistent with the general principle that morphologically complex forms are optimized for phonotactic well-formedness (Haspelmath & Sims, 2010).

Examples (Uzbek stem + borrowed Persian-Tajik affix; illustrative):

- 1) *hunar + -mand* → *hunarmand* 'skilled; artisan';
- 2) *bog' + -bon* → *bog'bon* 'gardener';
- 3) *savdo + -gar* → *savdogar* 'trader';
- 4) *o'yin + -boz* → *o'yinboz* 'game-player; playful'.

In the Uzbek dialects of Bukhara, Samarkand and Khojand under the influence of the Tajik language, the substitution of the suffix -ga for -da is considered as an inter-dialect

morphological alliance. Not used in some dialects: *maktab bordim*, *shaxar bordim* that occur in such forms. This alternative to > \emptyset is also specific to dialects. It should be noted that the use of the suffix of the accusative in place of the suffix of the accusative occurs in the dialects.

The suffix -da is also used in dialects where the ending of a word ends in a consonant: *Toshkentda* (*Toshkentta*).

The affix of the dative has allomorphs -dan, -tan: *maktabdan* (*maktaptan*) in these examples, D shifts from -tan to T as a result of progressive assimilation in the allomorph.

At this point, the D - T exchange is historically phonetically conditioned.

The possessive affix -i, which indicates the third person singular, has its -sh, -isi allomorphs. However, -si, -isi allomorphs are added to morphemes ending in a vowel, and -i is added to morphemes ending in a consonant.

-isi allomorph sometimes joins after the possessive affix to form affixal pleonasm.

U. Kramsky says that in Uzbek the -si third person singular form suffix sometimes serves as a definite suffix in oral speech: *ota-otasi*.

In some cases, the affix -s can be used not only to denote the plural, but also in the sense of stylistic reinforcement: *Bozorlarni aylandik, ishlar besh*.

Sometimes the root of a compound word is Turkish and the next compound word is Tajik. Such words can be used synonymously. But a sound in the stem of the first Turkish word is eaten, because it is in the unstressed state: *mo'ri-mo'rkon*, *ariza-arzacha*. Here the morphoneme "l" - \emptyset differs. When the possessive affix is added in some words, the morpheme a- \emptyset differs: *shaxar-shaxri*. The same thing happens when a word-forming affix is added to the stem of some Turkish words: *bosqi-bosiqich*. In this case, the morphoneme i- \emptyset also differs.

Dictionary-oriented discussion of borrowed formatives also shows conventionalization of particular suffixes. For instance, the Persian diminutive suffix -ak is

reported as widely represented in Uzbek lexicographic material, supporting the idea that borrowed formatives can stabilize as part of Uzbek derivational resources (Abduvaliev, 2023).

Hybrid formations and morphophonological "double conditioning"

In contact-induced hybrid word formation, the output is shaped by two constraint sets:

- donor-pattern semantics and combinability;
- recipient-language phonotactics and boundary rules.

Consequently, borrowed suffixes may show "double conditioning": they retain donor-like meanings and distributional preferences, but their surface realization can shift under Uzbek morphophonology.

This perspective aligns with contact linguistics accounts where borrowing involves both matter (forms) and patterns (structural replication), and where morphological material may be integrated through gradual phonological and morphotactic accommodation (Johanson, 2015; Matras, 2009).

Discussion

The results suggest that Tajik-Uzbek contact morphophonology manifests in two complementary ways. First, in bilingual dialect zones, contact can introduce functional overlap between markers (e.g., locative vs. dative usage), expanding the inventory of competing variants beyond strictly phonological conditioning. Second, borrowed derivational material enters Uzbek and becomes integrated into its morphophonological system: borrowed affixes attach to Uzbek stems, and the resulting forms undergo accommodation and boundary adjustments.

Sometimes the root of a compound word is Turkish and the next compound word is Tajik. Such words can be used synonymously. But a sound in the stem of the first Turkish word is eaten, because it is in the unstressed state: mo'ri-mo'rkon, ariza-arzacha. Here the

morphoneme "l" - \emptyset differs. When the possessive affix is added in some words, the morpheme a- \emptyset differs: shaxar-shaxri. The same thing happens when a word-forming affix is added to the stem of some Turkish words: bosqi-bosiqich. In this case, the morphoneme i- \emptyset also differs.

The conjunctive used between compound words does not perform an affixal morphophonological function: gultojikhoro. In a number of synonymous word a sound is added at the border of the correct, stem and affix morphemes: in the paradigm "qochoq-qochqoq-qochqin" \emptyset -q, o-i alternate vowel and consonant phonemes. "Axoli-axl": o- \emptyset , i- \emptyset drops two vowels. The correctness of the noun and verb forms in some words indicates that they are in fact synonymous: bosvoldi - bosib oldi (used as a homophone in oral speech, especially in Tashkent dialect: bosvoldi).

In this case, the transition to ib-v occurs under the influence of "o" at the beginning of the next word, ie as a result of regressive assimilation. One of these words belongs to the synthetic form (bosvoldi), the other to the analytical form (bosib oldi). The synthetic form may be formed as a result of some contractions, simplifications, and other combinatorial-positional changes in the phonological structure of the word. For example: tog' + olucha – tog'olcha, sakkiz+o'n-sakson etc.

In these examples, the morphemes u- \emptyset , i- \emptyset differ. However, in the example of "sakson" the simplification (-kiz + u) simplifies the distinctive assimilation of the unvoiced k-q consonant to the "z" consonant in them, and due to the inability of the i-o vowels to combine "k" and "q" under the influence of the vowel the tongue became the back "o".

The above simplification phenomenon is also associated with the process of phonetic economics, because the adaptation of the movement of the members of the speech, etc., is the result of this. When a possessive suffix is added to some nouns, a vowel in the core morpheme is dropped: og'iz-og'zi, o'g'il – o'g'li

etc. In this case, the morphemes $i-\emptyset$, $u-\emptyset$ differ. Some words (ulug'-ulug'i, buyruq-buyrug'i) are an exception. In the paradigm *vaxma-vaxma(li)-vaxmaq-vaxshat* the morphoneme $i-\emptyset$ is defined. In the words *ko'klam-ko'kalam* (*ko'kat*), the affixal morpheme was simplified and removed a vowel (*ko'klam*). In the example *sadoqat-sodiq*, one syllable appears to be shortened, and the $a-$ $-o$ morphoneme performs a morphological function: *sadoqat* (noun) - *sodiq* (adjective). In some words, the syllable is complicated, and an abstract noun is distinguished from a noun that represents a profession: *raqs-raqqos*, *naqsh-naqqosh*, *dalolat-dallol*. The nouns, which are originally compound words, are simplified by the reduction of sounds and syllables: *sariq-yog'-saryog'*, *bu(l) kun-bugun*. In *sariq-yog'* "q" may be redistributed and shortened on subsequent exposure. In the next word *bugun-bu(l)-kun* it fell into an intervocal state and resounded.

Importantly, contact does not erase the native system; rather, it adds a layer of socially distributed variation. For description and teaching, this means that analysts must distinguish (a) native allomorphy patterns that are primarily phonological and regular (Abduazizov, 1992; Nurmonov, 1990) from (b) contact-driven variants whose distribution is governed by bilingual practice and functional reanalysis (Johanson, 2015; Matras, 2009).

For Uzbek NLP and standardization, recognizing contact-driven variation is crucial. Rule-based analyzers that rely on affix stripping or inflectional-ending inventories must handle alternative suffix choices, borrowed formatives, and dialectal forms to remain robust (Salaev, 2024; Sharipov & Salaev, 2022).

Conclusion

This study systematizes two major pathways by which Tajik-Uzbek contact expands Uzbek morphophonological variation: (1) interference in the functional distribution of case markers in bilingual zones and (2) the borrowing and integration of Persian-Tajik derivational affixes and affixoids. The findings support a layered view of Uzbek

morphophonology: a relatively stable native component with rule-governed boundary behavior, and an additional contact layer in which competing markers and hybrid formations are licensed by sociolinguistic norms.

The descriptive contribution is twofold. First, the paper highlights that contact-driven case-marker competition can weaken otherwise tight links between grammatical function and suffix choice, creating surface variants that are not predictable from phonological environment alone. Second, it shows that borrowed derivational material is not "external" to morphophonology once it becomes productive: Uzbek speakers apply familiar accommodation strategies at stem-suffix boundaries, and borrowed suffixes gradually become subject to Uzbek morphotactic expectations.

These conclusions have practical implications. In language teaching, especially in regions with widespread Uzbek-Tajik bilingualism, it is useful to teach learners both the standard paradigms and the typical contact variants as distributionally conditioned options rather than as random errors. In applied linguistics and lexicography, a clearer inventory of productive borrowed affixes supports more consistent dictionary treatment and more transparent morphological segmentation.

For NLP (Native language process), contact variation matters because many Uzbek tools adopt affix-stripping or ending-based morphological analysis. Incorporating a contact-aware layer such as optional variant mappings for case markers and curated lists of borrowed derivational formatives can improve normalization and downstream tasks (e.g., retrieval, lemmatization, tagging) on dialectal or user-generated data.

Limitations of the present study include its reliance on descriptive sources rather than corpus counts and fieldwork. Future research should:

- collect corpus evidence from bilingual regions;

- quantify the frequency and constraints of specific competing patterns by dialect zone;
- evaluate the impact of contact-aware rules in computational morphological analyzers.

References:

1. Abduazizov, A. (1992). *O'zbek tili fonologiyasi va morfonologiyasi* [Uzbek phonology and morphophonology]. O'qituvchi.
2. Abduvaliev, A. (2023). Some remarks on Persian variations in Uzbek language (the example of suffix -ak). *American Journal of Philological Sciences*, 3(02), 79-82.
3. Alimova, Z. (2023). On borrowing prefixes and suffixes from Persian-Tajik languages into Uzbek. *Scientific Journal of the Fergana State University*, 1, 124-127.
4. Ermatov, I., & Dehqonova, L. (2021). Ob affiksakh, zaimstvovannykh iz tadzhikskopersidskogo v uzbekskii yazyk [On affixes borrowed from Tajik–Persian into the Uzbek language]. *Obshchestvo i innovatsii*, 2(2/S), 462-469.
5. Gardani, F., Arkadiev, P., & Amiridze, N. (Eds.). (2015). *Borrowed morphology*. De Gruyter Mouton.
6. Haspelmath, M., & Sims, A. D. (2010). *Understanding morphology (2nd ed.)*. Hodder Education.
7. Hojiyev, A. (1979). *Hozirgi o'zbek tilida forma yasalishi* [Form-building in modern Uzbek]. (Publisher information unavailable in consulted sources).
8. Johanson, L. (2015). Language contacts. In M. Robbeets & M. Savelyev (Eds.), *The Oxford handbook of the Turkic languages*. Oxford University Press.
9. Matras, Y. (2009). *Language contact*. Cambridge University Press.
10. Nurmonov, A. N. (1990). *O'zbek tili fonologiyasi va morfonologiyasi* [Uzbek phonology and morphophonology]. O'qituvchi.
11. Peace Corps. (1992). *Uzbek language competencies for Peace Corps volunteers in Uzbekistan*. Peace Corps.
12. Rahmatullayev, S. (2006). *Hozirgi o'zbek adabiy tili* [Modern literary Uzbek]. O'qituvchi.
13. Salaev, U. (2024). *UzMorphAnalyser: A morphological analysis model for the Uzbek language using inflectional endings*. arXiv.
14. Sharipov, M., & Salaev, U. (2022). *Uzbek affix finite state machine for stemming*. arXiv.
15. Sharipov, M., Kuriyozov, E., Yuldashev, O., & Sobirov, O. (2023). *UzbekTagger: The rule-based POS tagger for Uzbek language*. arXiv.
16. Thomason, S. G., & Kaufman, T. (1988). *Language contact, creolization, and genetic linguistics*. University of California Press.
17. Zokirova, H. R. (2016). *O'zbek tili morfonologiyasi (o'quv-uslubiy qo'llanma)* [Uzbek morphophonology: A teaching manual]. Andijon.