

MORPHOTACTIC RULES IN THE MORPHOLOGICAL ANALYZER OF THE UZBEK LANGUAGE

Khamroeva Shahlo¹

¹Doctoral student of Tashkent State University of the Uzbek Language and Literature,
Doctor of Philosophy (PhD) in Philology

Abstract – The article analyzes some issues of morphotactic rules of the Uzbek language morphoanalyzer. The article provides information about morphotactics, grammatical valence, its internal and external factors, describes the specifics of morphotactic rules between morphemes. The features of word-selective and non-word-selective morphemes are also analyzed in the development of morphotactic rules.

Key words: *uzbek language, morphoanalyzer, morphotactics, valence, word formation, word-selective / non-word-selecting morpheme.*

I. Introduction

It is known that there are several stages in the formation of a morphoanalyzer. One of these is the creation of a system of *morphotactic rules*. Also, the rules of morphotactics have a special place in the components of the linguistic support of the morphoanalyzer: without these rules, qualitative analysis cannot be achieved. Morphotactic rules are developed mainly on the basis of the morphotactic capacity of the verb and the noun base: it serves to determine the relationship between the base and affixal groups.

II. Main part

Morphotactics and valence. Morphotactics is the study of the valence (the possibility of accumulation) of morphemes in the process of word formation or syntactic communication, the study of their morpheme relations [10]. In other words, morphotactics is the branch of morphemes that studies the laws of conjugation of so-called morphemes. Morphotactics is also an integral part of word formation, which studies the cohesiveness, valence properties of morphemes and their morpheme relations (Ye.A. Zemskeya).[11]

The morpheme relationship of affixes. Each affixal morpheme has its own circle and is added only to a certain range of words or bases, forming this or that word form. This special process of word formation and word formation is called a *morpheme relationship* (N.M. Shansky, A.N. Tikhonov) [11].

Valence (lat. *valentia* - force) - the basis of fabrication in the fabrication and the ability of the connecting part to combine; the ability to combine a modifier with a base or other type of attachment. In syntax, the term valence refers to the ability of a word to coalesce: such a possibility is called *external valence*. The combination of morphemes in a word is *internal valence* and is called the morpheme connection or *morphotactics* of suffixes [10].

Linguist M. Abuzalova in her research on substantial morphology, valence and syntactic structure in the Uzbek language studied the issue of valence in a monographic plan [1]. The term valence of grammatical form is also raised in the researches of S.H. Muhamedova [4], the problem of grammatical form valence Sh. Shahobiddinova [9], B. Mengliyev [5].

Defining the concept of valence, M. Abuzalova writes: “The concept of lexical / semantic, semantic / syntactic valence is initially applied to a lexical phenomenon, and to this day valence is mainly understood as lexical / semantic valence, but in the substantive approach lexical valence. At the same time, the concept of grammatical valence - the valence of grammatical / morphological forms, as well as syntactic categories (positions) differs” [1; 64].

The linguist describes grammatical valence as follows: “The ability of a particular word to normally combine with words of a particular spiritual group is word valence. This possibility in the linguistic valence of words is called semantic valence. Like lexemes, grammatical forms have valences, which can change the valence of a lexeme and affect the valence of a lexeme. In other words, the ability of affixal morphemes to coalesce is the valence of the grammatical form” [1; 67-69].

The grammatical valence or morphotactics mentioned by M. Abuzalova plays a special role in the development of the rules of morphotactics of the automatic morphoanalyzer.

In recent years, the linguist emphasizes the difference in valence in the concept of valence in Uzbek linguistics to three aspects: word valence; syntactic valence; the grammatical form. According to him, there are two problems with word valence: linguistic valence and speech valence.

Valence ability in grammatical forms, in particular, categories that are classified as syntactic forms of Turkic languages (H. Nematov, G. Zikrillayev, B. Bahridinova) or forms of communication [6; 3; 2;] in grammatical forms the valence ability is clearly visible. For example, in the future, ownership will require additional consideration. If the combination of possessive suffixes with the possessive suffixes is a necessary (obligatory) valence, then the possessive

suffixes can be combined with the possessive suffixes is an optional (optional) valence. Only this phenomenon is no longer specific to the lexical unit, but also to the grammatical formant [1].

According to M. Abuzalova, in the process of speech, a linguistic lexeme occurs as a certain word, in a certain syntactic position. The linguistic (potential) lexical-semantic valence of a lexeme is combined in this process with the valence of the grammatical form [1; 77-78]. The influence of grammatical valence on the lexical-semantic (potential) valence of a word must be taken into account in terms of morphological form and syntactic position, morphology and syntax in two interrelated but relatively independent areas. Therefore, the grammatical valence itself is divided into two: a) the valence of the morphological form (category); b) should be divided into groups such as syntactic position valence [1; 80].

Based on the above, it can be concluded that there are two types of morphotactics: linguistic morphotactics and speech morphotactics. Linguistic morphotactics is the ability of morphemes to combine (valence). Speech morphotactics is the occurrence of linguistic valence and the factors that determine it (lexical, syntactic, morphological and phonetic factors). For example, in a speech situation, *manim* ("my" word form) option is influenced by external factors.

The interaction of consecutive morphemes and the effect of a morpheme on a dominant word on a morpheme on a subordinate word are also external factors. For example: *I read a book // the book was read*. In the second embodiment, the dominant proportion in the ruling clause requires that the subordinate clause be in the general agreement.

In the development of morphotactic rules of morphoanalyst, we consider the optimal way to develop models of word-forming affixes and their combinations to core morphemes and to include them in the linguistic base. Because word-forming affixes make up a significant amount. The rules of morphotactics include word-forming affixes and their arrangements. Of course, word modeling localization modeling takes place after morphotactics. In our opinion, if special attention is not paid to the phonetic phenomena that occur in the nucleus when a word-forming affix is added, the scope of work in morphological analysis will expand, and it will not be as expected as in the analysis.

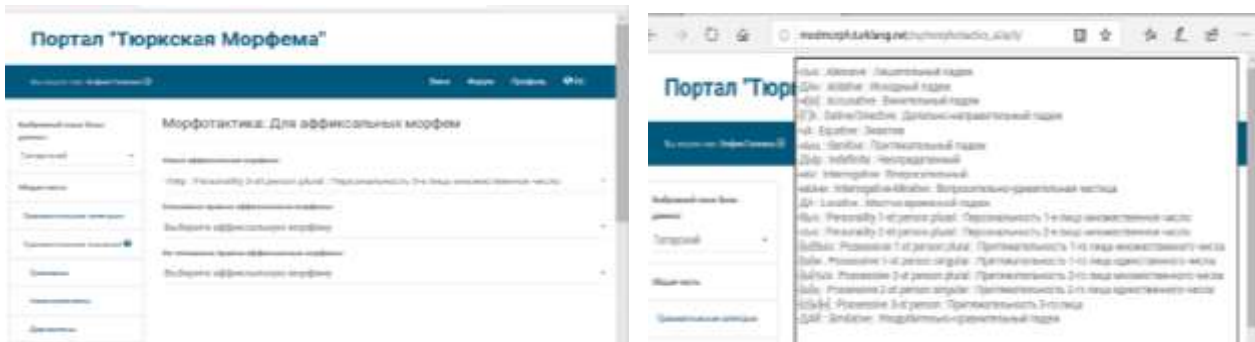
The following is an example of the morphotactics of some morpheme-forming, and the question of the arrangement of word-forming, lexical, and syntactic-forming morphemes will be the subject of the next chapter.

It is known that in Uzbek affixal morphemes are added after the stem. The rules of morphotactics determine the sequence of morphemes and allomorphs in a word. There are instructions in the reference list [7; 34-39] on this. Based on this, the following table shows the possibility of linking some morphemes in the Uzbek language.

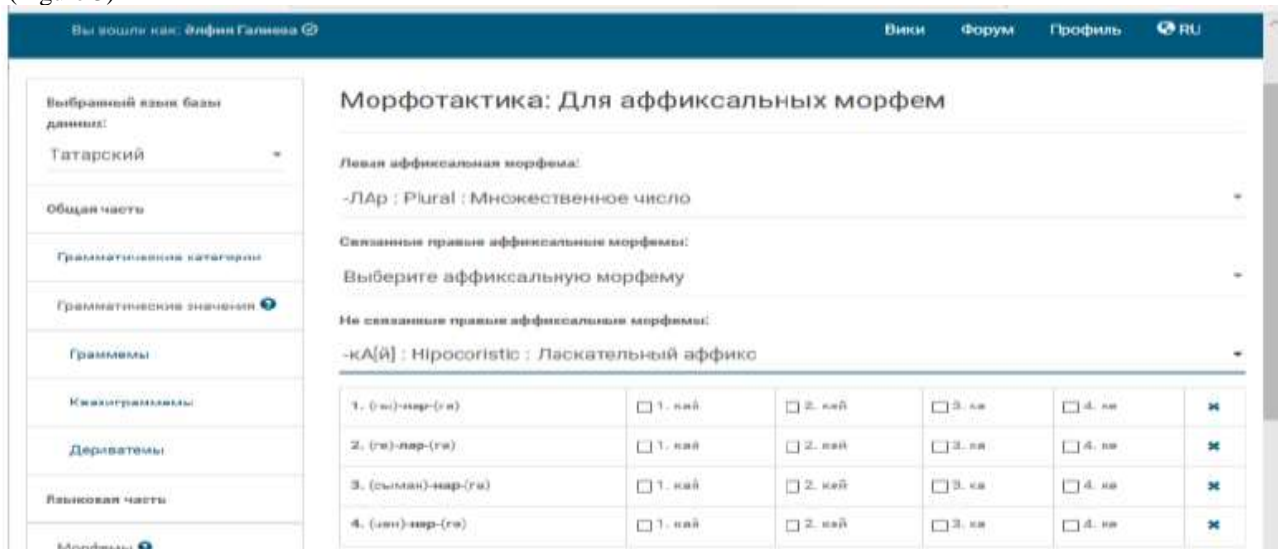
Left side layout morphemes			
№	List of morphemes	Symbol: features of left side layout (+) / (-)	Example
1	-лар	+	бола+лар+дан
2	-ман	-	бор+а+ман
3	-сан	-	ўқи+й+сан
4	-миз	-	ёз+а+миз
5	-сиз	-	айт+иб+сиз
6	-инг	+	кел+инг
7	-им	+	китоб+им
8	-имиз	+	дафтар+имиз
9	-си	+	ука+си
10	-ни	-	ашула+ни

As can be seen from the table, suffixes such as *-лар*, *-им*, *-инг*, *-имиз*, *-си*, *-ни* can be added directly to the core, they are added to the left side of another affixal morpheme. is added. The suffixes *-ман*, *-сан*, *-миз*, *-сиз* cannot be added directly to the stem. Because before them, another morpheme (for example, a tense suffix) is added to the stem, followed by this group of suffixes, which are located to the right side of affixal morphemes.

It can be seen that in some morpheme-based analysis programs [12] a very convenient way of defining morphotactic rules has been chosen. To activate this function, once placing all the affixal morphemes existing in the language of the program database, the function of setting the rules of morphotactics is activated A specific suffix is defined from the list of affixal morphemes. This morpheme is marked and saved after this suffix (see Figure 1-2). In this order, additional morphemes are attached to each of the additional morphemes on the right side (see Figure 3). Our observations show that this method is the most optimal and convenient way to establish the rules of morphotactics. (Figure 1, 2)



(Figure 3)



It can be said that in order to compile the rules of morphotactics of the Uzbek language, a list of all suffixes in the language is formed and the usual order of their placement is determined.

In the next chapter, we will discuss in detail the arrangement of morphemes, that is the surrounding of affix morphemes in a word, the specialization of meaning in the surrounding, and the modeling (arrangement) of the surrounding of morphemes.

3.3.2. About non-word-selective and word-selective morphemes

Computer linguistics, in particular, morphoanalysis experts., distinguish between *word-selective* and *non-word-selective* morphemes [8; 81-83]. In Turkic languages, including Uzbek, there are a number of suffixes that are added to all stems belonging to the same morphological group. For example, the suffix *-ra* can be added to all stems in the morphological group N. Such suffixes are called *non-word-selective* morphemes. However, there is a group of suffixes that do not apply to all stems belonging to the same morphological group. This group of suffixes can include personal and passive suffixes (-ин, -н, -ил, -л). They are potentially added to the verb-stem morpheme, but in practice (in context) a certain group of stems cannot accept this suffix. For example, the stem *безамоқ* takes both suffixes: *то безамоқ*, *безалмоқ*. There is no law on whether or not these additions are to the core. Determining the possibility of addition / subtraction of ratio additives simplifies the process of morphoanalysis, simplifies computer processing and speeds up synthesis. There is no any rule as to whether these additions are included in the suffix or not. Determining the possibility of the relative type of suffixes included / not included in the stem simplifies the process of morphoanalysis, facilitates the work of a computer, and accelerates synthesis.

Let's analyze the relative clauses of the word form *жойламоқ*. In order to accurately convey the meaning, we use the time and the personal clause.

- 1) active voice: *жойлади* – *Сабзавотларни қутиларга тартиб билан жойлади*;
- 2) reflexive voice *жойлаиди* – *Ўқтам янги квартирасига жойлаиди*;
- 3) passive voice: *жойланди* – *Машиий техника буюмлари картон қутиларга жойланди*.
- 4) joint voice: *жойлаиди* – *Сабзавотларни қутига тартиб билан жойлаиди*;

5) causative voice: *жойлаштирди* – *Машиини техника буюмларини картон қутиларга жойлаштирди*.

Apparently, the meaning of reflexive voice is expressed with the suffix *-и*. This means that the Uzbek language has grammemes that are added only to certain word forms, which must be sorted among the *word-selecting suffixes*. Otherwise, the morphological analysis process will not give the expected result.

There is another unusual use of suffixes. In the classical grammar of the Turkic languages, word-formings belong to the group of non-wordselective suffixes. *-дай, -ли, -сиз; -дай* makes adverb, *-ли, -сиз* makes adjective. For example: *қуш+дай, бир+дай* (adverb), *туз+ли, туз+сиз* (adjective).

The constructive part of these derivative words must be included in the database along with the artificial words. Obviously, the addition of such constructions to the lexicon leads to an extraordinary increase in the size of the base. In bilingual lexicographic resources, only lexeme-equivalent units are included in the translated language in order to prevent the lexicon from being artificial enlargement.

Compare:

- 1) *тузли (salty) – соленый; ақлли (smart) – умный; истеъдодли (talented) – талантливый;*
- 2) *Kitobli (with book) – “*” – kitobga ega; burunli (with nose) – “*” – burni bor.*

Apparently, the words of the second line do not take place in the dictionary, therefore, the reflection of such derivatives in dictionaries is a lexicographic problem to study. This issue remains unresolved both in lexicography and in computer linguistics.

Let's analyze the examples in more detail: *гулли* – floral (цветочный), *давомли* – continuous (продолжительный), *сувли* – watery (сочный), *тишли* – *toothed* (имеющий зубья, зубчатый), *ташвишли* – *anxious* (озабоченный, обеспокоенный, встревоженный, тревожный).

Here are some examples of their use in context: *"The rain slowly turned to snow. Continuous applause that turned into an ovation. On the eve of Navruz, the elderly make white clothes, and young people sew colorful floral dresses. The "virtues" of Alimkhan are innumerable. When an owner of grounds and waters, rich man dies, his heirs are being torn apart with each other. Universal's big toothed wheels made the yard cut down like a tesha. Transmission via toothed wheels or gears are widely used in the technology. The fire in the hearth turned Gulamjon's anxious heart into hope. The chief's condition was really worrying."*

There are such words that in translation dictionaries do not have their own equivalent. For example, *ватанли, уйли-жойли* – *settled somewhere* (имеющий дом, кров; обосновавшийся где-либо; обзаведшийся домом, хозяйством), *ишли* – *having job* (занятый делом; имеющий дело). It is clear from the translation in parentheses that the above words are not given by a single lexeme in the dictionary.

A similar example can be found in words with the *-сиз* morpheme. For example: *сифатсиз* (poor quality) (некачественный; низкого, плохого качества; недоброкачественный; ~ **махсулот (product)** недоброкачественная продукция), *шафқатсиз* (ruthless) (безжалостный; бездушный, беспощадный, жестокий; бесчеловечный;), *штатсиз* (part-time) (внештатный; // вне штата; ~ **лектор (lecturer)** (внештатный лектор).

Here are some examples of their use in context: *Since the opportunity is missed, there is a rush and as a result, the work is performed in poor quality (639). Part-time departments operate on a community basis (644). "They are very ruthless people, these inspectors," the young man muttered as he reached a safe place, - if you get caught, you will not get away with paying a fine (644).*

But the word kumobсuz (without a book) has no equivalent in the dictionary. We think that it is not permissible to include words of this type into the analyzer dictionary.

Conclusion

In conclusion, it should be said that in order to formulate the rules of morphotactics of the Uzbek language, a complete list of all suffixes of the Uzbek language is required. At the next stage, the usual order of the suffixes in the word is determined. There are such morphemes in the Uzbek language that not all words belonging to one group can be united equally. Although they theoretically belong to the same group, in practice they are able to connect only to a specific word. In developing the rules of morphotactics, it is necessary to take into account such word-selecting morphemes, to develop a complete list of them for linguistic support.

References:

1. Abuzalova M.K. Substantial morphology, valence and syntactic device: Doctor of Philological Sciences (DSc) dissertation. - Samarkand, 2018. - 224 p.
2. Bakhriddinova B. The system of lexical forms of verbs. The category of Verb. Dissertation of the candidate of philological sciences. - Tashkent, 2002.
3. Zikrillayev G. Independence and literary language. - Tashkent: Fan, 2004.

4. Mukhamedova S.Kh. Semantic and valence features of action verbs in Uzbek language. Doctoral dissertation of philological sciences. - Tashkent, 2007.
5. Mengliyev B. Spiritual features and semantic possibilities of morphological means: Abstract of dissertation of the candidate of philological sciences. - Tashkent, 1996;
6. Nigmatov Kh. G. Functional morphology of Turkish-speaking monuments of the XI-XII centuries. Monograph. – Tashkent: Fan, 1989.
7. Suleymanov D. sh., Gatiatullin A. R. Structural and functional computer model of Tatar morphemes. - Kazan: Feng, 2003. - 220 p.
8. Formal models and software tools for computer processing of the Tatar language. / R. R. Gataullin, A. R. Gatiatullin, O. A. Nevrozova, D. R. Mukhamedshin, D. Sh. Suleimanov, B. E. Khakimov, A. F. Khusainov. – Academy of Sciences of RT, Institute of applied semiotics AN RT. – Kazan: publishing house of Academy of Sciences RT, 2019. – 260 p.
9. Shakhobiddinova Sh. On the interpretation of grammatical meaning: abstract of dissertation of the candidate of philological sciences, - Samarkand, 1993.
10. <http://www.slovorod.ru/russian-wordform.html>
11. http://window.edu.ru/catalog/pdf2txt/516/25516/8411?p_page=10
12. http://modmorph.turklang.net/ru/morphotactics_a2a
13. МАТИБАЙЕВА, R. (2019). THE DEVELOPMENT OF KHANAPHISM IN MAVARANNAKHR. *The Light of Islam*, 2019(3), 16.
14. Мамасалиев, М. М., & Бўриев, И. И. (2019). ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ УЧАЩИХСЯ-ОДНО ИЗ ОСНОВНЫХ НАПРАВЛЕНИЙ ДЕЯТЕЛЬНОСТИ ПЕДАГОГА. In *ФУНДАМЕНТАЛЬНЫЕ ОСНОВЫ ИННОВАЦИОННОГО РАЗВИТИЯ НАУКИ И ОБРАЗОВАНИЯ* (pp. 188-190).
15. Очилова, Н. Р., Мамасалиев, М. М., & Астанов, А. Р. (2020). ВНИМАНИЕ ВОСПИТАНИЮ ГАРМОНИЧНО РАЗВИТОГО МОЛОДОГО ПОКОЛЕНИЯ. *International Independent Scientific Journal*, (14-2), 18-20.