

# **The information structure of phraseology: Mazanderani in correlation with Persian and Tajik**

Vladimir Borisovich Ivanov

*Institute of Asian and African Countries, Lomonosov Moscow State University  
Moscow, Russia  
iranorus@mail.ru*

Leyli Rahimovna Dodykhudoeva

*Institute of Linguistics of the Russian Academy of Sciences  
Moscow, Russia  
leiladod@yahoo.com*

The present paper examines proverbs in several Iranian languages that feature a unique type of verbless sentence. Methods of analysis developed for Persian verbless predicative structures are applied to data from the Mazanderani, Gilaki and Tajik languages.

The structures of verbless sentences that are contained in the proverbs in these Iranian languages are characterized by the absence of a copula and ellipsis. The absence of a verbal part enhances the style of the utterances, lending them additional expressiveness.

The unique style of the proverbs focuses the listener's attention on the message, lending it greater weight and making it more persuasive. Such proverbs are built on a familiar linguacultural image and, therefore, their language can be more concise.

In the construction of such verbless paroemias in different languages, we detect a number of features, such as presence of nominal parts of speech, which have the so-called "built-in predicativity" and cannot be combined with other predicative indicators, such as personal forms of the verb; within complex phrases, the components may be connected by a coordinating enclitic conjunction, or the connection can simply be implied. A significant number of two-member phraseological units contrast the first clause with the second.

**Keywords:** phraseology, phraseological units, information structure, predicativity, ellipsis, paroemiological data, verb

**For citation:** Ivanov V. B., Dodykhudoeva L. R. The information structure of phraseology: Mazanderani in correlation with Persian and Tajik. *Rodnoy yazyk. Linguistic Journal*, 2025, 2: 87–114.

**DOI:** 10.37892/2313-5816-2025-2-87-114

## **Информационная структура мазандеранской фразеологии в соотношении с персидским и таджикским материалом**

Владимир Борисович Иванов

*Институт стран Азии и Африки, Московский  
государственный университет имени М. В. Ломоносова  
Москва, Россия  
iranorus@mail.ru*

Лейли Рахимовна Додыхудоева

*Институт языкоznания РАН  
Москва, Россия  
leiladod@yahoo.com*

В статье рассматриваются паремии ряда иранских языков, в которых представлен особый тип безглагольных предложений. Методика анализа, разработанная для персидских безглагольных предикативных структур, в данной статье применяется к материалу мазандеранского, гилянского и таджикского языков.

Структуры безглагольных предложений, образующих паремии в указанных иранских языках, характеризуются: а) отсутствием связки и б) эллипсисом. Отсутствие глагольной составляющей усиливает стиль высказываний, придавая им дополнительную выразительность. Особый стиль паремий фокусирует внимание слушателя на сообщении, придавая высказыванию больший вес и делая его более убедительным. Такие паремии построены на известном лингвокультурном образе, и потому их язык более лаконичный.

В построении таких безглагольных паремий в разных языках обнаруживается ряд закономерных черт: 1) в некоторых представ-

лены именные части речи, которые имеют так называемую «встроенную предикативность» и которые не сочетаются с другими предикативными показателями, такими как личные формы глагола; 2) в двучленных фразеологических единицах используется союзная связь (сочинительный энклитический союз) и импликация, где первая часть противопоставляется второй.

**Ключевые слова:** фразеология, фразеологические единицы, информационная структура, предикативность, эллипсис, паремиологический материал, глагол

**Для цитирования:** Иванов В. Б., Додыхудоева Л. Р. Информационная структура мазандеранской фразеологии в соотношении с персидским и таджикским материалом. *Родной язык. Лингвистический журнал*, 2025, 2: 87–114.

**DOI:** 10.37892/2313-5816-2025-2-87-114

## 1. Introduction

### Mazanderan and the Mazanderani language

According to the most recent classification, the Mazanderani language and its dialects belong to the Central (formerly, North-western) subgroup of Iranian languages, and are spoken by about 2.5 million people living on the south-eastern coast of the Caspian Sea. Although Mazanderani is not an official, state-sanctioned written language, many Mazanderani texts have been published. They represent various genres: descriptive grammars of various dialects, works of Mazanderani writers and poets, dictionaries and collections of local folklore and proverbs. It is noteworthy that historical and literary texts in Mazanderani date back to the medieval period (for more details see: [Ivanov, Dodykhudoeva 2025; 2021; 2017]). Mazanderani speakers are for the most part bilingual, i.e., they speak Mazanderani and Persian as their native languages, and sometimes also local Turkic varieties.

This article examines Mazanderani along with some Persian and Tajik phraseology, particularly units that are verbless and thus based on ellipsis.

## State of research in the field

In this article we focus on phraseology featuring “verb phrase ellipsis” (VP ellipsis or VPE), a type of grammatical omission whereby a verb phrase is left out but its meaning can still be reconstructed from the context [Skovorodnikov 1973].

This type of ellipsis is typically found in colloquial speech, and is characterized by constructions lacking a component in a syntactic position. Ellipsis is determined by the circumstances of the utterance and the presence of non-verbal means of communication (gestures with specific semantics). It is sometimes also caused by the structural organization of the text and is widely used in fiction as a stylistic figure that lends dynamism and enriches expressiveness. Ellipsis has a wide range of interpretations in the areas of speech comprehension and text linguistics [Bel'chikov 1990], which require further research in Iranian studies.

We give special attention to phraseological units that constitute utterances, defined as “any linear segment of speech that, in a given speech situation, performs a communicative function and, in this situation, is sufficient to report something” [Shvedova 1980 II: 84]. We regard (syntactically meaningful) phraseological units as constructions that act either as idioms (noun phrases) or as paroemias, such as proverbs and other proverbial sayings.

It is widely acknowledged that the concept of an utterance is closely linked to that of a sentence and its associated predicativity. According to Yakov Testelets, “the grammatical property of predicativity, which distinguishes sentences from other types of groups, is associated with the typical use of sentences as utterances, i. e., speech segments appropriate in a specific speech situation in which there is a speaker, addressee, subject, time, place, and purpose of the message” [Testelets 2001: 233].

Furthermore, it should be noted that there are classes of words that cannot form utterances. These are function words such as prepositions, conjunctions, clitics, preverbs, articles, etc. However, some of these, including particles, interjections, ono-

matopoeic words (ideophones), and words representing a whole sentence, can form utterances classified as undivided sentences.

### **Iranian phrase structure:**

#### **Persian, Tajik and Mazanderani**

The verb plays an important role in all Iranian languages [Edel'man 2001a; 2001b; Veretennikov 1993]. Persian has basic SOV word order; however, this language permits scrambling. Most of the predicates in the language are complex predicates comprised of two parts, a verb and a nonverbal element. Their formation is productive and they comprise an ever-expanding segment of the verbal system. The class of simple verbs is mostly closed; there are about one hundred of them. Most of these verbs do not contribute to the core semantics of complex predicates, although they play a decisive role in determining the argument structure of the sentence.

There exists wide scholarship devoted to Persian phraseology, and recently attention has been paid to the description and classification of verbless sentences (cf. research into the phenomenon of the head and constituent ellipsis, widely discussed in Persian grammars [Ivanov 2019]). However, the case of phraseological ellipsis, considered one of the types of ellipsis (either lexical or syntactical), has not yet been fully researched. As we have encountered this phenomenon in our fieldwork, we examine here ellipsis in phraseological units, understood as the omission of a structurally necessary and semantically significant component of a word, phrase, or sentence. This is especially relevant to all Iranian languages and cultures, as at their core lies classical Persian literature, with its didactic sententious style and admonitions, including moralising idioms, maxims, proverbs and other figurative statements.

Although syntactic ellipsis has today been adequately considered in scholarly literature, an extensive volume of empirical material remains under-researched with regard to simplification processes in both written and spoken, syntactically significant, phraseological units having the status of a phrase (idiom) or especially a sentence (paroemia).

With this in mind, we study the ellipticity of meaning-relevant elements in the structure of phraseological units in modern Mazanderani. It is known that, through the worldview of the speaker, phraseological units serve to figuratively and emotionally evaluate objects and phenomena which already have designations in the language.

These paroemiological units are not created by the speaker, but are used as set models, reproduced with constant components and conventionalized meaning. In this sense, such units exist as a special database, accessible to representatives of the relevant linguacultural community. These units are significant in analyzing the linguacultural specifics of a particular community, reflect the value orientations of its speakers, their worldview and traditions. Proverbs of this kind are typically multifaceted. Moreover, the secondary meaning generally turns out to be an expansion of the primary meaning and tends to be more abstract in nature.

These units reinforce the multidisciplinary nature of our research into the specificity of Iranian linguacultural traditions (as represented by Mazanderanis and Gilakis, Persians and Tajiks) and the vocabulary and syntax of their languages. The resonance of phraseological units lies in their formal invariance within the dominant cultural context. Their cognitive and emotional impact on the speaker is supported by their structural and semantic components.

Among stable phraseological units, the most common are idioms with an initial (prepositional) ellipsis. There exist also frequent cases of medial ellipsis, implying the elimination of the article, as well as shortened forms of multi-component idioms. Thus, by transforming predicative constructions into semi-predicative or nominative ones, ellipsis can have the effect of altering the syntactic status of phraseological units. Consequently, it has a derivational significance in idiomatic phrases. This is evidenced, in particular, by the recent trend toward an occasional reduction in the number of words used in phraseological units, and the development of elliptical “clipped proverbs”, transformed into idioms which also have the potential to change the sense [Aleksandrova 2006: 120]. This trend can be illus-

ted by a variety of Tajik examples, such as *xonai bačador* – *bozor*, *xonai be bača* – *mazor* ‘Home with children is a market-place, home without children is a cemetery’, which is usually in speech truncated to *xona bo bača...* or even *xona* – *bozor*.

The structures of proverbial verbless sentences in Tajik, even more clearly than in Persian, retain the stylization of folk expressions in combination with archaic vocabulary, imagery and admonitions, as in the expression *rešai duo* – *sabz* (lit. the root of prayer is green) meaning ‘an answered prayer leads to perfection’, which goes back to Sufi imagery, or in the widespread proverb *Avval pursiš, ba'd kušiš* ‘Ask first, act later’.

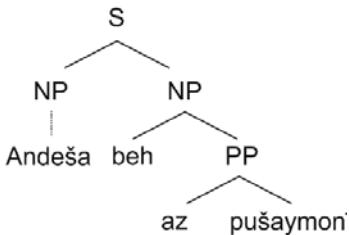
Some nominal parts of speech possess inherent predictability and therefore do not combine with other predicative units, such as personal forms of the verb. The majority of bipartite phraseological units oppose the first judgment to the second. Within complex judgments, the components are most often connected asyndetically or with the enclitical conjunction *-o* ‘and’. In many polyadic paroemias, the first clause contains a finite verb, and the subsequent clauses are elliptical sentences with an omitted verb. The copula is usually omitted in all clauses. A number of verbless sentences in paroemias are distinguished by an emphatic word order, with the attribute preceding its head. Most sayings, from the point of view of information structure, refer to situations as a whole and therefore are thematic sentences with a non-inherent topic and a given which is expressed in a preliminary context.

Such a situation can be seen in the Tajik proverb:

(Example 1)

Andeša	beh	az	pušaymoni
thought	better	than	remorse

‘Thought better than remorse.’  
(Cf. English ‘Look before you leap’.)



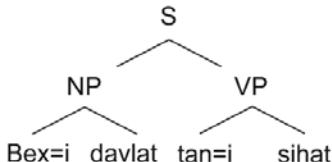
Ellipsis is distinguished from other types of null anaphora by a number of well-established criteria that define it as a surface anaphora (a term used by [Hankamer, Sag 1976]). Surface anaphoras have fully articulated syntactic structures, transformed at a later point through derivation [Ross 1967]. However, Merchant [2001] has argued that verb phrase ellipsis, along with sluicing, is derived just by non-pronunciation of the syntactic structure in phonetic form.

An example of this type of ellipsis can be seen in the following Tajik proverb. The copula after the last word is omitted:

(Example 2)

Bex=i	davlat	tan=i	sihat
root=IZ	happiness	body=IZ	health

‘A sound mind in a sound body.’



We base our hypothesis on Lazar Peysikov’s definition that phrases are grammatically formed units of language with a cohesive meaning that serve as the basis for constructing sentences [Peysikov 1959: 5]. In this perspective, a phrase is an extended form of designation, i. e., the nomination of some object or phenomenon. In some cases, a nominal group (noun phrase, NP) can be reduced to a single word.

We describe the most common types of verbless Mazanderani paroemias (proverbs, aphorisms and sayings), their syntactic and semantic structures and patterns of their functioning, with a view to demonstrating a brief figurative verbal expression of traditional values and views based on the life of Mazanderani people.

A sentence differs from a phrase in two essential features [Testelets 2001: 230]: by the presence of *predicativity*; and by the presence of information structure (division into theme and rheme, linking new information with already known information). The Mazanderani structural schema of a sentence is usually simplified to “noun – finite form of the verb”. In general, a Mazanderani sentence with the structure SOV (subject–object–verb) ends with a verb or a copula. However, there are sentences where no verb is used, and where *predicativity* may not be explicit even in writing. As Shvedova points out [Shvedova 1990: 395], the presence of the verb in the structural basis of the sentence is not necessary, and there are many types of sentences that are built just by means of nominal components (without a verb).

Structural analysis of phraseological units implies analysis of the syntactic units correlating with them, i. e., phrases and phraseological-unit sentences. These phraseological sentences differ from phraseological units of other types by the actual presence of *predicativity*.

In oral speech, *predicativity* is usually expressed just by intonation. To our knowledge, Mazanderani intonation has not yet been the subject of any specific research. However, in this paper, in connection with the aforementioned phraseology, we also draw on the results of studies of Mazanderani and Persian intonation currently in progress [Ivanov 2018a; 2018b]. Provisionally (judging by human ear) bilingual Mazanderani speakers use the same intonational constructions both for Mazanderani expressions and when interpreting these in Persian. Therefore, as a first approximation, we will apply the results of studies of Persian intonation.

If we treat intonation as a means of achieving *predicativity*, then *predicativity* remains the sole but sufficient feature that

distinguishes a sentence from both a phrase and a single word [Ladd 1996; Peysikov 1959: 208; Ivanov 2023a]. Generally, the division into *subject* and *predicate* components coincides with the division into **theme** and **rheme**. But there are cases where they diverge: the intonational expression of predicativity is based on the opposition of the intonational constructions of *incompleteness* and *completeness*. Incompleteness in the Mazanderani language is signaled by a significant rise in tone of the final syllable of the subject group and a noticeable pause before the predicate group.

## 2. Structures of Mazanderani paroemias

We examine verbless Mazanderani phraseology, with the emphasis on:

- a) omission of zero link-verbs and
- b) ellipsis structures [Ivanov 2022].

In this way the language of proverbs tends to be more economical: a thought is expressed in fewer words than in regular discourse, and omission of the verb heightens the style of statements, distancing them from everyday colloquial speech.

Elliptical and zero copula sentences are often found in paroemias. To source Mazanderani paroemias we have drawn on a number of works [Yazdānpanāh-e Lamuki 1997; Rahimyān 2004; Ansari 2011]. The most informative of these was [Ansari 2011], because in addition to giving a literary translation of Mazanderani paroemias into Persian, the author also provides a literal translation. Much of the material from this source was also voiced by our language consultant, which greatly facilitated the syntactic analysis of examples, making it possible, namely:

- a) to distinguish the unstressed ezafe *-e* (-ə) from the same stressed vowel in the absolute final of a word;
- b) to distinguish the unstressed copula *-e* (-ə) ‘is’ from the same stressed vowel in the absolute final of a word;
- c) to distinguish the unstressed conjunction *-o* (-ə) ‘and’ from the same stressed vowel in the absolute final of a word;

d) by pauses, to detect the differences between phrases and simple sentences in the composition of compound sentences, as well as to identify the boundaries between the components of the subject and predicate, theme and rheme.

In general, the scripts used by the source authors [Yazdān-panāh-e Lamuki 1997; Rahimyān 2004; Ansari 2011], whether Latin or Arabic-Persian, did not allow us in writing to make the kind of distinctions listed in (a) to (d).

### The intonational expression of predication in Mazanderani

The intonational expression of predication in Mazanderani can be observed in both verb and verbless phrases. The Mazanderani phrase differs from the Russian one by a more marked variation in tone. The stress in nominal parts of speech falls on the final syllable, which is emphasized by a higher  $F_0$  (fundamental frequency) [Ivanov 2014: 105; 2018a: 54].

Consider the example of a Mazanderani two-part sentence with a verb:

(Example 3)

Me	bār	<u>sabok</u>	bay-ye
My	burden	light	become.ind.PST=3sg

‘I felt better’



Me bār sabok bayye

*Fig. 1. Phrasal intonation in the simple sentence Me bār sabok bayye ‘I felt better’*

In example (3), the division into subject and predicate components coincides with the division into theme and rheme (underlined). Syntactic structure and information structure do not always coincide in this way. The cases where they diverge will be highlighted in further examples.

The intonational expression of predication is based on the opposition of the intonational constructions of incompleteness and completeness (for more on this topic, see [Bryzgunova 1969]). The same intonational constructions of incompleteness in the Mazanderani language consist of a significant rise in tone (major third – perfect fourth) on the final syllable of the subject group and in a noticeable pause before the predicate group (Fig. 1). In Arabic-Persian script this pause may be indicated by a comma, and in some cases by a colon, although these punctuation marks are often omitted. In cases where no punctuation was visible in writing, our native-speaker/consultant pronounced the relevant phrase, after having intuitively performed a syntactic analysis. In some cases, his first attempt was unsuccessful, and then after a pause, he pronounced the corrected version.

### Phraseological units

A significant number of verbless sentences exist in the form of interrogative phrases, as in the Tajik example:

(Example 4):

Az	po=i	lang	čī	sayr	va
from	foot=IZ	lame	what	stride	and

az	dast=i	gurusna	čī	xayr?
from	hand=IZ	hungry	what	good?

‘What stride can you get from a lame foot and what use can you get from a hungry man?’

(Cf. English ‘A hungry belly has no ears’.)

Without further details (see, for instance, [Ivanov 1995]), we note that intonational predication is also expressed in such cases, and that in two-part sentences the intonational construction of incompleteness is retained.

We have seen that an affirmative declarative sentence with the SOV structure is formed by an intonational construction of completion, characterized by a smooth drop in tone towards the end of the sentence (i. e., its predicate; Fig. 1). So, in the same way,

verbless sentences are formed through intonation, except that in those cases the tone falls not on the verb, but, as a rule, on the nominal parts of speech, which are in the final position.

The concept of “predication” is not usually applied in any of the grammars of Iranian languages or dialects known to us, including the Mazanderani language. In these grammars, a sentence is defined in two aspects: semantic and structural. See for instance: “In the Mazanderani dialect of the city of Sari, a sentence consists of a theme and a rheme (subject and predicate), e. g., *Hosayn šune<sup>1</sup>* ‘Hussain is coming’” [Shokri 1995: 134].

There exist a number of lexemes in various Iranian languages which, for various reasons, cannot be combined with a verb, including a copula.

Such words include the one-word sentences *ba:le* ‘yes’, *no* ‘no’; interjections such as *ay!*, *āy!* ‘hey!', *oh!*, *in-am* ‘here’ [BPRS]<sup>2</sup>, and some others, notably the predicative word *kū?* ‘where?’. The presence of predication in the interrogative word *kū?* means that it can be combined only with a so-called zero copula (i. e., used without a verb), thus excluding ordinary verbs and copulas.

(Example 5)

<u>Ande</u>	<u>čerā</u>	<u>hā</u> <u>kerd-i</u>		<u>pas</u>	<u>kū</u>	Te	dembe?
so	much	grazing[you did.IND.PST-2SG]	so	where	where	You	fat_tail?

‘(You) grazed so much, so where is your fat tail’

(Cf. English “However hard you try, you’re flogging a dead horse”)

In the compound sentence (example 5), the first clause contains a verb, but in the second, because of the predicative word *kū?* ‘where?’, there is no verb.

<sup>1</sup> Since in Mazanderani the stress generally falls on the last syllable of a word, the stressed vowel is highlighted in bold in cases where this is not the case, as in the example above and in those given below.

<sup>2</sup> Cf. also Tajik *ana*, *mana* ‘here it is, that’s it’.

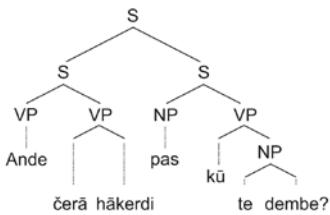


Fig. 2. Syntactic tree diagram of Example 5

(Example 6)

<u>In-am</u>	naz=e	šass=e	essā	<u>in-am</u>
here=also	caress=IZ	finger=IZ	teacher	here=also
mezz=e	dass=e	essā		
payment=IZ	hand=IZ	teacher		

‘Here the caress of the teacher’s (big) finger, here the payment of the teacher’s hand’.

This phrase is used in the sense that “one does harm, and another bears the blame”, i. e. “to be wrongfully accused”, cf. English “One law for the rich and another for the poor”.

There are no verbs in the compound sentence (example 6) because they cannot be combined with the predicative subordinating conjunctions *in-am* ‘here’, which form the rheme. The data providing information about the situation lies outside the sentence.

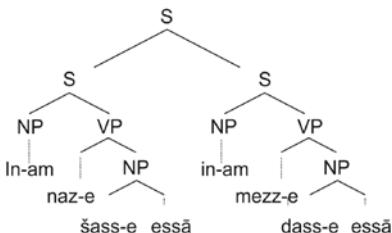


Fig. 3. Syntactic tree diagram of Example 6

This example corresponds to Persian

<u>In-am</u>	nāz=e	šast=e	ostād,
here=also	caress=IZ	finger=IZ	teacher

<u>in-ham</u>	<u>mozd=e</u>	dast=e	ostād
here=also	payment=IZ	hand=IZ	teacher

‘Here the caress of the teacher’s big finger, here the payment of the teacher’s hand.’

Cf. this sentence in Persian, close in sense but with a full sentence structure (including verbs), in Firdowsi’s “Šahnāme”:

*Bedānsān-ke šāhān navāzeš konand, bedān bandegān-niz navāzeš konand.*

‘Know that the rulers make kindness, know (they) make kindness (that) subjects (are yet to feel)’.

What is important for Iranian languages is that sentences with zero copula should be distinguished from so-called elliptical (incomplete, truncated) sentences, in which a verb could be included, but is actually omitted. “There may be various reasons for such a contraction, called an ellipsis. For example, part of the message can be omitted when it is clear to the speaker and hearer, due to their existing knowledge of the situation” [Testelets 2001: 253–254].

Between these two types of verbless sentences, it is hard to draw a clear boundary; the two sets overlap. In cases where the verbless sentence consists of several words, the question arises which word is the main one, which is the dependent one, and what is the hierarchy of these words.

Figure 4 shows the syntactic tree diagram corresponding to example (7).

(Example 7)

Ārd	<u>hame</u>	jā	nun	jā	<u>be</u>	jā
flour	all	place	bread	place	to	place

‘Flour everywhere, (but) bread here and there, i. e. to bake bread, you need to work.’

(Cf. English “No pain, no gain”)

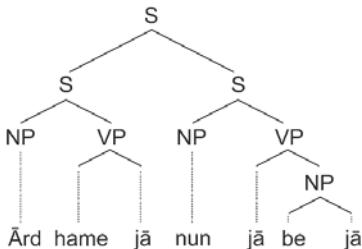
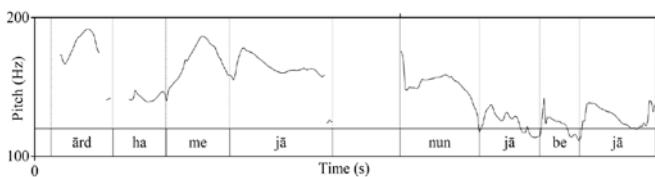


Fig. 4. Syntactic tree diagram of Example 7

Fig. 4 shows that the entire statement consists of two simple sentences (S). Each sentence consists of a noun phrase (NP) and a verb phrase (VP), in other words, the composition of the subject and predicate. The verbal groups in the example are rhemes. Here and in further examples, rhemes are underlined. Although copulas are omitted in these verbal groups, they could, in principle, be included, i. e., here we are dealing with an elliptical sentence.

We also give a Mazanderani example of intonation in example (7) ‘Flour everywhere, bread here and there’ – with a significant rise in tone of the final syllable of the subject group <ārd> and <nun> and a noticeable pause before the predicate group <hame jā> and <jā be jā>:



In written Persian this pause may be indicated by a comma or a colon. In cases where no punctuation was visible in writing, our language consultant pronounced the relevant phrase, based on his intuitive syntactic analysis. In some other cases, he made

several attempts or hesitated before he pronounced the correct version.

Thus, ellipsis helps to enhance the effect of imagery when encoding cognitive and expressive information and serves to eliminate redundancy in the verbal representation of meaning, specifically by elimination of the verb.

Example (7) is based on a very common model, whereby the first part of a statement is contrasted with the second. There are several logical variants within this model. Examples (7) and (8) demonstrate logical multiplication or conjunction, represented as  $a \wedge b$ , i. e., they are complex logical expressions, which are considered true if, and only if, both parts of the expression are true; in all other cases, the complex expression would be false. Hereinafter, the notation of formal logic, and its interpretation, are based on the work of [Pleskunov 2014: 41].

(Example 8)

Ādem	jenn=o	jeme	pari
man	jinn=and	clothes	peri

‘The man is a jinn, but by clothes is (like) a peri’

(Cf. English “A wolf in sheep’s clothing”, “Don’t judge a book by its cover”)

This model may have become established in the Persian and Mazanderani languages under the influence of Arabic, where the same model is present. In compound sentences, the components can be linked either with the help of the conjunction *-o* (-ā) ‘and’ (as in example (8)) or without a conjunction (as in example (7)). As a rule, in these cases the coordinating conjunction *-o* (-ā) ‘and’ is translated into English by the adversative conjunction ‘but’, less often by ‘while’.

As regards the information structure in example (8), the division into theme and rheme does not coincide with the division into given information and new information. Since example (8) is a compound sentence, it has two themes and two rhemes (underlined). It is a statement about someone who was mentioned

earlier, i. e., it is completely new, but the given information is outside the statement, previously stated in a preliminary context.

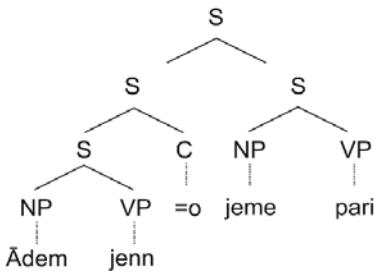


Fig. 5. Syntactic tree diagram of the verbless sentence in Example 8

Although, as an enclitic, the conjunction *-o* (-ā) ‘and’ adjoins the last word of the first proposition, it connects not only that word to those that follow, but the propositions as a whole to each other (as shown in Figure 5).

(Example 9)

Āš	kele=sari	zan	nūmzebāzi
soup	stove=on	woman	marriage_proposal

‘Soup on the stove, a woman (reaches) the marriageable age’ (Cf. English ‘Time and tide wait for no man’ or ‘Procrastination is the thief of time’).

Example (9) is a *thetic* sentence (communicatively undivided). So, in colloquial speech, only its first part can be used: *Āš kele-sari...* ‘Soup on the stove...’

It is a judgement about a situation which has been defined in a prior context. In terms of their information structure, sentences like these usually consist entirely of a *rheme*, i. e., they are *thetic* sentences or sentences with a non-inherent theme [Testelets 2001: 447–448].

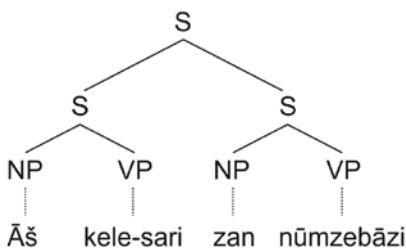


Fig. 6. Syntactic tree diagram of the verbless sentence in Example 9

New information coincides with the rheme; although given information is omitted, it can be clearly deduced from the context or situation. Although paroemias often consist in a fully polypredicative model of opposition containing a judgement expressed in the components of a sentence, a truncated model is also common, whereby the contrasted components are reduced to a phrase or even to a single word.

(Example 10)

Aš=ā	Assyubūni
bear=and	flour_milling

‘A bear and flour milling’ (meaning “one thing has nothing to do with another”, cf. English “Like chalk and cheese”)

In the simple sentence (10) the opposing judgements are expressed by nouns linked by the enclitic conjunction *-ā* ‘and’. This phrase constitutes a thetic sentence. It answers the question: What is happening? What is going on? Here the given information and the theme are not expressed; they are clear from the prior context.

Example (10) is a minimal adversative construction where each of the antithetical components does not exceed one word. In formal logic, it corresponds to the XOR gate (strict disjunction, excluding the possibility of “or”). This is represented as  $a \wedge b$ , i. e.,  $a \wedge b$  is true when either  $a$  or  $b$  is true, but both cannot be true at the same time.

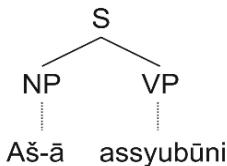


Fig. 7. Syntactic tree diagram of the verbless sentence in Example 10

### 3. Gilaki, Persian and Tajik phraseology with a similar structure

Paroemiological data (Persian *zarbolmasal-o goftehā*, Tajik *zarbolmasal-u maqol* ‘proverbs and sayings’) with a similar logical structure are widely known and used in both Persian and Tajik. Thus, in Persian we find this proverb:

(Example 11)

xāna=ye	xers	(=o	bādiye=ye	mes)
house=IZ	bear	(=and	cup=IZ	copper)

‘In the bear’s house, but a copper cup’ (mes ‘large copper cup for wine’)

This is used to denote an impossible situation, the manifestation of something contrary to expectations, or of things that are not related to each other. In colloquial speech a truncated part of the expression is widely used metaphorically: *xāna-yi xirs* ... ‘a bear-house’; it is even documented in the early vocabulary of Steingass [1892] with the metaphorical meaning: ‘appearance of a thing where it is not expected’. A similar proverb is known in Gilaki:

(Example 12)

xars	xane=yu	āb	angur?
bear	house=and	water	grape

‘In the bear’s house and grape juice?’

Here also the expression is understood as the juxtaposing of two incompatible things [Marashi 2003: 556]. Cf. Kurdish *hirç û govend*, *şivan û dîwan*, *file û pez*, *kurmanc û gameş* ‘bear and round dance, shepherd and power...’. In Tajik, we also find a closely related saying usually applied when denoting the worthlessness of something or somebody:

(Example 13a)

az	xirs	mū=e...
from	bear	hair=ART

‘From a bear – (at least) a strand of fur.’

In its full form it reads (Example 13b):

az	xirs	mū=e	az	gul	bū=e
from	bear	hair=ART	from	bear	hair=ART

‘From a bear – a strand of fur, from a flower – a fragrance.’

There is a variant with a similar sense that even has three simple phrases in a single unit (Example 13c):

az	gurgxurda	pūst	az	gul	bū=e,
from	wolf_eaten	skin	from	flower	smell=ART
az	xirs	mū=e			
from	bear	hair=ART			

‘From a (sheep) eaten by a wolf – a skin, from a flower – a fragrance, from a bear – fur/hair.’

In a colloquial Tajik situation only the last part of the saying is regularly used [Dodykhudoeva, Vinogradova, forthcoming]. Another notable example comes from the Ramsar region of Iran [Rahimiyān 2004: 51]:

(Example 14a):

on	te	sag	kuta
he	your	dog	puppy

‘He is your dog’s puppy’, i. e., he looks like your dog.

Cf. the Tajik expression with the similar sense *bača či gu-na – oča namuna* ‘what the child (looks like) – (alike) his mother’ [Fozylov 1963: 84] or (*čun*) *yak sebi dukafon* ‘two halves of one apple’ [Fozylov 1964: 417], Rus. *kak dve poloviny odnogo yabloka* ‘like two halves of one apple’, *kak dve kapli vody* ‘like two drops of water’.

The following Persian equivalent contains a copula omitted in Ramsari:

(Example 14b):

u	tule=ye	sag=e	to=st
he	puppy=IZ	dog=IZ	you=COP.3SG

‘He is your dog’s puppy’, i. e., he looks like your dog.’

This is equivalent to the English “The apple never falls far from the tree.”

This is a further original example from Ramsar:

(Example 15a):

“anār”	baxt	“oyoz”	tāla
pomegranate	chance	nut	fate

‘A pomegranate (is a sign of good luck) chance, (and/but) a nut (is the sign) of (predestined) fate.’

However, in [Rahimiyyān 2004: 50] we find Ramsari:

(Example 15b):

“anār”	baxt=ə	“oyoz”	tāla
pomegranate	chance=?	nut	fate

‘A pomegranate (is a sign of good luck) chance, (and/but) a nut (is the sign) of (predestined) fate.’

The =ə in the written text should be pronounced and treated as either an article or a copula (3Sg.), or can even be understood as a coordinating conjunction.

The Persian explanation is given with a copula in the first clause:

(Example 15c):

“anār”	nešāne=ye	šāns=ast,	gerdu	nešāne=ye	sarnevešt
pomegranate	sign=IZ	chance=3SG	nut	sign=IZ	fate

‘A pomegranate (is a sign of good luck) chance, (and/but) a nut (is the sign) of (predestined) fate.’

These examples suggest a type of fortune-telling, predicting good or bad luck, with interpretation (*ta'bir*) of what is in one's mind or seen in a dream.

### Specific features of Tajik *ana* and *mana*

A further noteworthy, under-researched paroemiological case can be found in Tajik data with the pronoun/particle *ana* ‘here it is, here!’, and also *mana* ‘here it is, that's it!’ which have their own “built-in” predicativity. On this basis there exist several proverbial sayings, such as:

(Example 16a):

Ana	gap=u	mana	gap!
here	talk=and	here	Talk

‘Would you believe it!?’

(Example 16b):

Ana	xalos!
here	Release

‘What a surprise!’

Both expressions can be used metaphorically in the sense: ‘That's all! That's all there is to it!’

## 4. Conclusion

We have observed that ellipsis is the process of simplifying the surface structure of a sentence without affecting its underlying essence (as in the Chomskyan perspective). Since the basis of ellipsis is the tendency towards optimisation of linguistic resources, it is widely used in idioms and proverbs.

In phraseological units, ellipsis becomes significant when omitting function words bearing limited semantic weight – auxiliaries, articles, and communicatively unloaded pronouns – and when replacing nouns with their adjectival or numerative substitutes. Such transformations often lead to changes in the syntactic status of phraseological units, converting predicative units (paroemias) into nominative ones (idioms). In proverbs and sayings, ellipsis acquires a systemic character, since its scope encompasses most fundamental, structurally significant components of the sentence – the subject and predicate, which can be regularly omitted within all positionally defined ellipsis patterns.

Thus, the process of simplifying the surface structure of a sentence without affecting its essence explains why we have no problem in decoding omitted elements of phraseological units; meaning-relevant words remain outside the scope of elimination, as they are not subject to abbreviation, and the fixed word order keeps the phraseological unit within the systemically defined parameters of the sentence. The proposed structural-semantic analysis of Mazanderani and the phraseological units of other Iranian languages allow us to include the presence of such a syntactic phenomenon as ellipsis in the scope of the linguacultural data.

The structures of verbless sentences forming Mazanderani paroemias are characterized by distinctive features: a) zero copula and b) ellipsis, as well as by the presence of predicative words having predicative status (*kü?* ‘where?’, *in-am* ‘here’).

In this way the language of such sentences becomes more economical and more concise. The idea is expressed in fewer words than in everyday speech. The elimination of the verb enhances the style of statements, giving them extra weight. The

deviation of the style of paroemias leads the interlocutor to listen more attentively to the speaker, thus making the speaker's statement more persuasive.

In polypredicative paroemias, the first simple sentence is complete, with a predicate expressed in the personal form of the verb, while the second and subsequent sentences (or clauses) are elliptical with a truncated verb.

In terms of their information structure, most sayings refer to situations as a whole and are therefore *thetica* sentences with a non-inherent theme and data, which were previously expressed in a preliminary context.

## References

Aleksandrova 2006 — Aleksandrova V. G. Ellipsis v strukture frazeologicheskikh edinits sovremennoogo angliyskogo yazyka [Ellipsis in the structure of phraseological units of the modern English language]. *Kul'tura narodov Prichernomor'ya*. 2006, 76: 119–122. (In Russ.)

Ansari 2011 — Ansari M. *Farhang-e zarb-ol-masalhā-ye māzandarāni* [Dictionary of Mazanderani phraseology]. Sāri, Entešārāt-e Šalfin, 2011 (1390). (In Pers.)

BPRS — *Bol'shoj persidsko-russkiy slovar'* [Complete Persian-Russian Dictionary]. Comp. V. B. Ivanov. Vol. I–III. Moskva, 2020–2024. (In Russ.)

Bryzgunova 1969 — Bryzgunova Ye. A. *Zvuki i intonatsiya russkoy rechi. Lingafonnyy kurs dlya inostrantsev* [Russian sounds and intonation. A linguaphonic course for foreigners]. Moskva, 1969. (In Russ.)

Bel'chikov 1990 — Bel'chikov Yu. A. Ellipsis [Ellipsis]. *Lingvisticheskiy entsiklopedicheskiy slovar'*. Yartseva V. N. (gl. red.). Moskva, 1990.

Dodykhudoeva, Vinogradova (forthcoming) — Dodykhudoeva L. R., Vinogradova S. P. K kharakteristike obraza medvedya v fol'klornoy i literaturnoy traditsiyakh iranskogo lingvokul'turnogo areala [On the characterization of the image of the bear in the folklore and literary traditions of the Iranian

linguocultural area. *Traditsionnye rechevyе praktiki* (po materialam Lingvisticheskogo Forum-a-2022). (In Russ.)

Edel'man 2001a — Edel'man D. I. Voprosy sravnitel'no-istoricheskogo sintaksisa iranskikh yazykov [Issues of Comparative-Historical Syntax of Iranian Languages]. *Issledovaniya po iranskoy filologii*. Vyp. 3. Moskva, 2001, 150–170. (In Russ.)

Edel'man 2001b — Edel'man D. I. K rekonstruktsii prairanskogo predlozheniya [On the Reconstruction of a Proto-Iranian Sentence]. *Yazyk i kul'tura. K 70-letiyu YU. S. Stepanova*. Moskva, 2001, 137–145. (In Russ.)

Fozylov 1963, 1964 — Fozylov M. *Farhangi iborahoi rextai hozirai tojik (farhangi frazologī)*. 1. Dushanbe, 1963. 2. Dushanbe, 1964. (In Taj.)

Hankamer, Sag 1976 — Hankamer J., Sag I. A. Deep and surface anaphora. *Linguistic Inquiry* 1976, 7: 391–426.

Ivanov 1995 — Ivanov V. B. Granitsy slova i inkapsulyatsiya v persidskom, tadzhikskom i dari [Word boundaries and encapsulation in Persian, Tajik and Dari]. *Voprosy yazykoznaniya*, 1995, 3:107–117, Moskva. (In Russ.)

Ivanov 2014 — Ivanov V. B. Mazanderanskaya prosodiya [Mazanderani prosody]. *Lomonosovskie chteniya. Vostokovedenie: tezisy dokladov nauchnoy konferentsii* (Moskva, 14 aprelya, 2014 g.). Red. I. I. Abylgaziev, M. S. Meyer. Moskva, 2014. (In Russ.)

Ivanov 2018a — Ivanov V. Navā-ye goftār dar zabān-e māzanderāni [Types of speech in Mazanderani]. *Pažuhešhā-yi dar bāre-ye kerānehā-ye janubi-ye daryā-ye Kāspiyen*. Tehrān, Mirmāh, 2018 (1397). (In Pers.)

Ivanov 2018b — Ivanov V. Navā-ye goftār dar zabān-e gilaki [Types of speech in Gilaki]. *Pažuhešhā-yi dar bāre-ye kerānehā-ye janubi-ye daryā-ye Kāspiyen*. Tehrān, Mirmāh, 2018 (1397), 59–62. (In Pers.)

Ivanov 2019 — Ivanov V. B. *Ocherk teorii persidskogo yazyka* [A sketch of a theory of the Persian language]. Vol. 1. Moskva, 2019. (In Russ.)

Ivanov 2022 — Ivanov V. B. Bezglagol'nye predlozheniya v mazanderanskikh paremiyakh [Verbless sentences in Mazan-

derani paroemias]. *Armyanskiy gumanitarnyy vestnik*, 2022, 8: 100–112 (In Russ.)

Ivanov 2023a — Ivanov V. B. Bezglagol'nye predikativnye struktury v persidskikh paremiyakh [Verbless predicative structures in Persian paroemias]. *Voprosy yazykoznaniya*, 2023, 5: 117–132. (In Russ.)

Ivanov 2023b — Ivanov V. B. Intonatsiya nezavershonnosti v iranskikh yazykakh [Intonation of incompleteness in Iranian languages]. *Lomonosovskie chteniya. Vostokovedenie i afrikanistika* (Moskva, 4–7 aprelya 2023 g.), Moskva, 2023, 179–181. (In Russ.)

Ivanov, Dodykhudoeva 2017 — Ivanov V. B., Dodykhudoeva L. R. Sintaksicheskie otnosheniya imon v severo-zapadnykh iranskikh yazykakh (na mazanderanskom i gilyakskom materiale) [Syntactic Relations of Names in Northwestern Iranian Languages (Based on Mazanderani and Gilaki data)]. *Voprosy yazykoznaniya*, 2017, 2: 77–95. (In Russ.).

Ivanov, Dodykhudoeva 2021 — Ivanov V. B., Dodykhudoeva L. R. Vižegihā-ye zabānšenāsi-o farhangi-ye gilaki o māzandarāni [Linguistic and cultural features of the Gilaki and Mazanderani languages]. *Pažuhešhā-ye dar bāre-ye kerānehāye ja-nubi-ye daryā-ye Kāspiyen*. (Studies of the coastal area of the Caspian Sea). Vol. 4. Tehran, 2021. (In Pers.)

Ivanov, Dodykhudoeva 2023 — Ivanov V. B., Dodykhudoeva L. R. Information Structure of Phraseology in Mazanderani. *ECIS 10. Conference Abstracts*. Individual Presenters & Panels. Leiden University, Leiden 21–25 August 2023. Leiden, 2023, 79.

Ivanov, Dodykhudoeva 2025 — Ivanov V. B., Dodykhudoeva L. R. Sotsiolingvisticheskiye osobennosti Prikaspinskogo Regiona: na materiale gilyakskogo i mazanderanskogo yazykov I [Sociolinguistic features of the Caspian Region: a study based on the Gilaki and Mazanderani languages]. *Sociolinguistica*, 2025, 2 (22): 24–61. (In Russ.)

Ladd 1996 — Ladd R. *Intonational phonology*. Cambridge, 1996.

Marashi 2003 — Marashi Ah. *Vāzhenāme-ye Guyeshe Gilaki be Enzemame Estelāhāt va zarbolmasalhāye Gilaki* [Dictionary of the Gilaki Dialect with Addition of Gilaki Proverbs and Sayings]. Rasht, 2003. (In Pers./Gilaki)

Merchant 2001 — Merchant J. *The syntax of silence: Sluicing, islands, and the theory of ellipsis*. Oxford, 2001.

Peysikov 1959 — Peysikov L. S. *Voprosy sintaksisa persidskogo yazyka* [Issues in Persian syntax]. Moskva, 1959. (In Russ.)

Pleskunov 2014 — Pleskunov M. A. *Osnovy formal'noy logiki* [Fundamentals of formal logic]. Yekaterinburg, 2014. (In Russ.)

Rahimiyān 2004 — Rahimiyān H. *Farhang-e zabānzadhā-ye Rāmsar (saxtsar)* [Dictionary of vernaculars of Ramsar]. Tehrān, Entešārāt-e Mo'in-Parvin, 2004 (1383). (In Pers.)

Ross 1967 — Ross J. R. *Constraints on variables in syntax*. Ph. D. Dissertation, Massachusetts Institute of Technology, 1967.

Shokri 1995 — Shokri G. *Guyeš-e Sāri (māzandarāni)* [Sari dialect (Mazanderani)]. Tehrān, Pažuhešgāh-e 'olum-e ensāni va motāle'āt-efarhangi, 1995 (1374). (In Pers.)

Skovorodnikov 1973 — Skovorodnikov A. P. O kriterii ellipsis tichnosti v russkom sintaksise [On the Criterion of Ellipticity in Russian Syntax]. *Voprosy yazykoznaniya*, 1973, 3. (In Russ.)

Shvedova 1980 — Shvedova N. YU. (ed.) *Russkaya grammatika. T. II. Sintaksis* [Russian grammar. II. Syntax]. Moskva, 1980. (In Russ.)

Shvedova 1990 — Shvedova N. YU. *Predlozhenie* [Sentence]. Yartseva V. N. (ed.). *Lingvisticheskiy entsiklopedicheskiy slovar'*. Moskva, 1990. (In Russ.)

Steingass 1892 — Steingass F. J. *A Comprehensive Persian-English dictionary, including the Arabic words and phrases to be met with in Persian literature*. London, 1892.

Testelets 2001 — Testelets Ya. G. *Vvedenie v obshchiy sintaksis* [Introduction to general syntax]. Moskva, 2001. (In Russ.)

Veretennikov 1993 — Veretennikov A. A. *Ocherki glagol'noi frazeologii persidskogo yazyka* [Outlines of Persian verbal phraseology]. Moskva, 1993. (In Russ.)

Yazdānpanāh-e Lamuki 1997 — Yazdānpanāh-e Lamuki T. *Farhang-e masalhā-ye māzandarāni* [Dictionary of Mazanderani phraseology]. Tehrān, Entešārāt-e Farzin, 1997 (1376). (In Pers.)