Diachronic study of information structure in Persian

Marjan Ansari
Ph.D. Student, Department of Linguistics, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran.
ansarymarjan@yahoo.com

Bahram Hadian* (iD)
Department of Linguistics, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran.
bah.hadian@khuis.ac.ir

Vali Rezaei
Associate Professor, Department of Linguistics, University of Isfahan, Isfahan, Iran
vali.rezai@fgn.ui.ac.ir

Abstract

The present study aimed to investigate diachronic changes in the frequency and function of marked syntactic structures, namely passive constructions, preposing, and cleft sentences that manifest information structural elements in Persian. The study examined these structures across three periods of Persian, i.e., Middle Persian, Dari, and Modern Persian. The data for Middle Persian sourced from The Mēnōg-ī Khrad and Oshnar-I Dana, while for Dari, Tārīkh-i Bayhaqī and Safarnāma by Nasir Khusraw are used. For Modern Persian, written resources from various genres, such as short stories and scientific articles, are analyzed. The samples from each period are analyzed to determine the function and frequency of marked syntactic structures, which are then compared to identify any potential changes in their usage over time. The findings indicated that the frequency and function of these syntactic structures have changed over time. The findings suggested that passive constructions were commonly used across all three periods to present the patient as the pragmatic topic and maintain the topic-focus order. However, in modern Persian, passive construction was also used to emphasize the patient as the focus. Cleft constructions were not found in Middle Persian, but the increase in frequency of cleft constructions in Dari can be attributed to the fact that in this period, cleft sentences were used similarly to preposing, in order to observe the principle of separation of role and reference, rather than focusing on specific elements. Nonetheless, in modern Persian, cleft constructions were used to exclude other possibilities and emphasize the selected element as the focus. Preposing structures were used to present an element as the topic and provide new information about it. This study contributes to our understanding of language change and provides insights into the evolution of Persian syntax over time.

Keywords: Diachronic Linguistics, Information Structure, Middle Persia, Dari.

1. Introduction:

The study of language change is an essential aspect of linguistics. It provides insights into how languages evolve over time and the factors that drive these changes. One area of language change that has received considerable attention in recent years is the study of information structure and its manifestation in syntax. The study aims to investigate diachronic changes in the frequency and function of syntactic structures, namely passive constructions, preposing, and cleft sentences that manifest information structural elements in Persian. Each of these structures serves a specific pragmatic function in this regard. The study examines these structures across three periods of Persian: Middle Persian, Dari, and Modern Persian. The central question is whether and how the frequency and function of

*Corresponding author

these syntactic devices, encoding information structure have changed over time in Persian. The hypothesis is that there may have been shifts in preferences for particular structures or in their information structural functions from one period to another. By tracking these diachronic changes, the study seeks to gain insights into how Persian grammar has evolved in the domain of information structure.

Persian, an Indo-European language, has undergone several stages of evolution over the centuries. According to Abolghasemi (1993), Old Persian, which was spoken between the sixth and fourth centuries BC, is the earliest form of Persian. Old Persian was used mainly for administrative purposes and was the language of the court. Middle Persian, also known as Pahlavi, is the stage of Persian that was spoken from the 3rd century BC to the 7th century AD. Modern Persian, also known as Farsi, is the stage of Persian that has been spoken since the 9th century AD. Fischer (1968) explains that the modern Persian language has evolved from Middle Persian and has borrowed many words from Arabic, Turkish, and other languages. According to Monshizadeh (2020), Dari is an evolved variety of Middle Persian.

Languages undergo changes over time, and one aspect of these changes is reflected in their syntactic constructions. Diachronic studies typically focus on the written form of a language due to the availability of written documents from past stages of a language. In the written form of a language, marked syntactic structures, such as passive and cleft structures, are used to convey different aspects of information structure, while in spoken language, it is the phonetic devices, such as contrastive stress or intonation, that modify the information structure of the sentence.Consequently, a diachronic study of these syntactic devices can have significant implications for understanding historical changes in languages and their subsequent pragmatic facets.

2. Background

Information structure refers to the way in which speakers organize the information contained in a sentence to convey their intended meaning. It involves the identification and marking of the main elements of a sentence, such as topic, focus, and presupposition. The study of information structure in linguistics traces its roots back to the Prague School of Linguistics, which was founded in the 1920s by a group of linguists in Czechoslovakia. The Prague School was interested in studying how meaning is conveyed through different instances of language use. To this end, Prague School linguists developed a number of concepts and methods that are still influential in the study of information structure.

One of the key concepts developed by Prague School was the notion of "theme" and "rheme." Accordingly, theme refers to a part of a sentences that is already known or presupposed by the listener or reader. This is while rheme is associated with a part of a sentence that contains new information. For example, in the following sentence, "the cat" functions as the theme of the sentence, and the predicate, "chased a mouse", is the rheme.

The cat chased a mouse.

Here, "the cat" is assumed to be the known or presupposed information introduced earlier in the context, therefore it serves as the theme of the sentence. "Chased the mouse" functions as the rheme since it provides new information about what the cat did. Later, this concept was further developed by other linguists, such as Halliday (1967), who used the terms "given" and "new" instead of theme and rheme.

Another important concept developed by the Prague School was the notion of "focus." It is defined as "the part of the sentence which presents the most important or relevant information" (Hajicova et al., 2000, p. 6). Other linguists, such as Lambrecht (2001), further developed the notion of 'focus' and made distinctions between "topic" and "focus". He argued that the way in which information is structured in a sentence can have a significant impact on how it is understood by a listener or a reader.

In the mid-twentieth century, the study of information structure was further developed by linguists working in the field of discourse analysis. One of the key figures in this field was Michael Halliday, who developed the systemic functional linguistics approach to language analysis. According to Halliday (1967), the organization of information in a sentence depends on three factors: the grammatical structure of the sentence, the context in which the sentence is used, and the speaker's or writer's communicative intent. Halliday's work on information structure was influential in the development of the 'given/new' framework, which is still widely used in linguistic analysis. According to this framework, information in a sentence can be categorized as either given (already known to the listener or reader) or new (previously unknown or unexpected). Halliday states:

> "The speaker is obliged to chunk his speech into information units. He has to present his message in a series of packages. He is, however, free to decide how he wishes to package the information. He is free to decide where each information unit begins and ends, and how it is organised internally" (1967, p. 200, as cited in Brown, Yule, 1983, p. 155).

In the same vein, Lambrecht (2001) also argues that, 'information structure' refers to the way in which information is divided and organized in discourse and how it is conveyed to the listener or reader.

In the 1990s and 2000s, the study of information structure was extended to the realm of cognitive linguistics. One of the key figures in this field was Ronald Langacker, who developed the theory of cognitive grammar. According to Langacker (1987), the organization of information in a sentence depends on the way in which the speaker or writer conceptualizes the situation being described.

Langacker's work on cognitive grammar was influential in the development of the concept of "activation," which refers to the degree to which a particular element in a sentence is prominent or salient in the speaker's or writer's conceptualization of the situation. Elements that are more salient or prominent are more likely to be focused on or emphasized in the sentence.

In recent years, the study of information structure has continued to be a fertile active field of research in linguistics. Researchers have developed a variety of different frameworks and methods to analyze the ways in which information is structured and conveyed through discourse, including the given/new framework, the topic/focus framework, and the cognitive grammar framework. One of the recent theories about information structure is Lambrecht (2001).

One of the key concepts in Lambrecht's framework is the notion of "topic" and "focus." Lambrecht (2001) defines 'topic' as "that part of the sentence which refers to the entity or situation that the sentence is about and which is held constant or presupposed throughout the sentence" (p. 141). In other words, topic is a part of a sentence that provides particular background . r context for the rest of the sentence. On the other hand, 'focus' refers to "that part of the sentence which presents new information and which is not presupposed or held constant throughout the sentence" (Lambrecht, 2001, p. 141). In other words, the focus is the part of the sentence that contains the new or unexpected information that the speaker or writer wants to convey.

Lambrecht (2001) argues that the way in which information is structured in a sentence can have a significant impact on how it is understood by a listener or a reader. By highlighting certain parts of the sentence as the topic or focus, the speaker or writer can guide the listener's or reader's attention and emphasize certain aspects of the message. Languages are equipped with different morphological, syntactic, and phonetic devices to represent information structure. In the present study, the focus is on syntactic devices, which manifest the information structure of a sentence in written language. The justification for adopting Lambrecht's (2001) information structure framework in this study lies in its comprehensive and well-established nature within the field of linguistics. Lambrecht's framework has been widely used and accepted by researchers studying information structure across various languages. It offers a systematic and detailed analysis of the organization and packaging of information within sentences, focusing on the distribution and prominence of discourse referents. One of the advantages of Lambrecht's framework is its ability to capture both syntactic and pragmatic aspects of information structure, providing a holistic perspective on how information is encoded and conveyed in language. In comparison to a purely cognitive account, Lambrecht's framework offers a more formal and linguistically grounded approach. While cognitive accounts emphasize the cognitive processes involved in information structure and its interaction with human cognition, Lambrecht's framework provides a more explicit and structured framework for analyzing and describing the grammatical and pragmatic devices used to encode information.

In written Persian, like in many other languages, information structure is realized through a variety of syntactic structures, including passive, preposing, and cleft structures. These structures are marked and serve specific functions in the manifestation of information structure. Specifically, in written language where phonetic devices cannot be used, syntactically marked structures are exploited to manipulate the information structure in the sentences which make up that text.

One of the syntactic devices of information structure that is discussed in the present research in Middle Persian, Dari, and Modern Persian is passive construction. Passive construction plays an important role in the information structure by changing the unmarked order of the information units in the sentence. According to Keenan and Dryer (2007:325), passive construction serves to emphasize or topicalize a particular element, such as the semantic role of patient, by moving it to the beginning of the sentence.

Persian's passive structure has long been a source of debate among linguists. A group of linguists support the existence of this structure in Persian, including Marashi (1970), Palmer (1971), Soheyli-Esfahani (1976), and Hajati (1977). In contrast, Moyne (1974) and Vahedi-Langrudi (1998) argue that Persian does not employ passive construction to communicate events and states. Taking a formal approach, Moyne (1974) maintains that there is no structure known as passive in Persian grammar. Therefore, what is considered to be passive in Persian is in fact an "inchoative structure." Unlike this argument, Rezai (2003), Jabbari (2003), and Dabir-Moghadam (1985) distinguish passive constructions from inchoative structures.

The other marked syntactic structures analyzed in this research are cleft and pseudo-cleft structures. These constructions are used to emphasize a particular element in a sentence. According to Huddleston and Pullum (2002, p. 1416), a cleft sentence "a cleft clause is formed by dividing a more elementary clause into two parts." In these structures, the first part of the sentence typically begins with the expletive pronoun 'it'. For example, "It was the dog that ate my homework" (Huddleston and Pullum, 2002, p. 1414).

Similarly, Quirk et al. (1985) explain that a pseudo-cleft construction is composed of two clauses in which the focus resides in the second clause. The first part of the sentence typically begins with a wh-word (such as what, who, or which), followed by a form of be verbs . For example, "What I saw was a beautiful sunset." Halliday (1967) states that in pseudo-cleft sentences, all of the constituents in a sentence can be moved to the focus position.

Another marked syntactic structure discussed in this study is preposing. Preposing constructions are used to place a constituent, usually a noun phrase, at the beginning of a sentence for emphasis or contrast. According to Huddleston and Pullum (2002), preposing is the technique of moving a constituent from its normal position to a position at or near the

beginning of the clause. For example, in the sentence "But Mary, I didn't see," the nominal item "Mary" has been preposed to the beginning of the sentence for emphasis.

The significance of this study lies in its ability to shed light on the long-standing debate among linguists regarding the existence and nature of passive structures in Persian. As it was previously noted, the literature embraces various conflicting views on this issue. While some argue for the presence of the passive constructions in Persian, others refute its existence, proposing alternative interpretations and considering them an instance of inchoative structure. By examining this linguistic phenomenon over time, a diachronic study can provide valuable insights into the evolution and function of passive and inchoative structures in Persian and is liable to resolve the extant controversy. Furthermore, the research also encompasses the analysis of cleft and pseudo-cleft structures, which are syntactic constructions used for emphasizing specific elements in a sentence. By investigating these constructions, the study contributes to an understanding of Persian syntax and the ways through which information is structured and highlighted.

3. Methodology

The data for the present research come from authentic written resources remaining the three periods of Persian language: Middle Persian, Dari, and Modern Persian. These data are employed to investigate the frequency and function of marked syntactic structures that manifest the information structure of sentences in Persian. . we opted for the passages from Mēnōg-ī Khrad and Oshnar-I Dana for Middle Persian. And the samples of Dari are collected from resources like Tārīkh-i'Bayhaqīi and Safarnāma by Nasir Khusraw. For Modern Persian, the data are collected from written resources in different genres, including short stories[1] and scientific articles[2]. The sample size for each respective period consists of 1700 sentences. It is worth noting that the interpretation of the samples from Middle Persian was subject to certain limitations, which consequently constrained the researchers from selecting larger sample sizes. The texts and passages belonging to Old Persian were excluded from the data set due to the fact that the analysis and interpretation of the texts of this era solicits a particular expertise that falls well beyond the borders of linguistics.

The frequencies of these marked syntactic structures are counted in the samples of different periods. The frequencies are then contrasted to determine whether there has been an increase or reduction in the use of these syntactic devices of the information structure. In order to reach explanatory adequacy in this study, the functions of these syntactic structures are then analyzed from an information structure point of view. In each example analyzed here, the elements of information structure, including focus and topic are determined, and the pragmatic function of the syntactic structures under investigation in the text is discussed.

This study is based on Lambrecht's (2001) theoretical framework on information structure. Lambrecht's (2001) framework is a widely recognized and influential model in the field of linguistics for analyzing the way information is structured and conveyed in discourse. It is worth noting that in this study cognitive approaches could not be employed to account for information structure since for the analysis, the researchers had access only to written texts of Middle and Dari Persian.

4. Data analysis

In this part, some examples of each syntactic structure from the three periods of Persian are analyzed in order to investigate the pragmatic function of these structures. The frequency of the syntactic structures in the corresponding eras is then provided and contrasted.

The procedure followed in discussing each example is as follows: Firstly, the Persian sentence is presented, followed by its English translation. Next, the focus and type of focus are identified and stated, and then the pragmatic topic of the sentence is discussed. Finally, the sentence is analyzed. Due to the limitations of the article, only 11 analyzed sentences are provided here, while the remaining examples are taken into account in the study's conclusion.

4.1 Passive construction

The first structure presented here is a case of passive in Middle Persian text. In this period, the passive structure is made from the past participle with the verb `isted`, the third person singular of the verb `istadɑn`.. Subsequently, the person and number of the verb are considered to be its logical subject. Therefore, in the absence of the subject in the sentence, it can be argued that the sentence has a passive meaning.

The first example presented here is belongs to Mēnōg-ī Khrad, one of the middle Persian corpora of this study.

(1) ʃaʃom   kaʃ   nasa   azer   nigɑn   kard   isted
     Sixth   that-he   Nasa   from   view   make PAST-1SG   become

Literal Translation: The sixth is that under which Nasa is hidden from view.

Focus: `nasa azer nigɑn kard isted` (Literal Translation: Nasa is hidden from view.)

Type of focus: Predicate focus

Pragmatic Topic: ʃ (third person singular pronoun)

In the text, from which this sentence is extracted, Mēnōg-ī Khrad is naming all types of lands which are not fertile, and in this sentence, the pronoun »š« refers to the land. Therefore, the pragmatic topic is the referent of this pronoun, i.e., land, which is mentioned in the beginning of the sentence, and the rest of the sentence provides new information about it. The writer has used a passive sentence in order to present the semantic role of patient as topic and then present new information about it. So, the predicate of the sentence is in the focus position. This sentence is pragmatically unmarked since the topic is at the beginning of the sentence and the focus follow it.

(2) ud xwastɑg an weh d xwoʃtar   ī az   fraronih handuxt isted.

      and wealth   that better and pleasant-comparative that from righteousness collected is

Literal Translation: O wealth is better and more pleasant, which is collected (by righteousness, multiplied and preserved by (good) deeds and benevolence)

Focus: an weh ud xwoʃtar ī az fraronih handuxt isted.. (Literal Translation: is better and more pleasant, which is collected)

Type of focus: predicate focus

Pragmatic Topic: (xwāstag) (Meaning: wealth)

In this passive sentence, pragmatic topic is the reference of the word (xwāstag), about which the entire sentence provides new information. In this text, Mēnōg-ī Khrad describes certain features that are the characteristics of a strong fortress or shelter. Therefore, (xwāstag) is considered the pragmatic topic, and the rest of the sentence plays the role of focus, which gives new information about the topic. The writer has used a passive sentence in order to present the semantic role of patient at the beginning of the sentence and reserve the pragmatically unmarked order of the elements of information structure.

The following two examples are from Dari Persian.

(3) ʃtand ʃoɢl-i   mohem darɑm ke farize ast   ta   an ra bɑrgozar ayad.

      Several Job-GEN important have-1SG that necessary is-3SG that-COMP that OM executed become

Literal Translation: I have some important jobs that should be fulfilled.

Focus: bɑrgozar ayad (Literal Translation: be fulfilled)

Type of focus: predicate focus

Pragmatic Topic: "an" (Meaning: that)

In this sentence, the passive verb has promoted the object to the subject position of the sentence. As a result, there should not be any need for "ra", which is the object marker in the sentence. Although in Modern Persian, passive construction is always used without an object marker, in some Dari Persian texts, a passive verb like the above sentence is often used with the object marker "ra". Therefore, we can postulate that in Dari Persian, passive structure is in a transition state between active and what is called passive in its modern sense.

In the first clause of this compound sentence, the reference of the noun phrase "some important jobs" is introduced, and, thus, it is activated in the mind of the addressee. In the second clause, it is referred to by the pronoun "an" since it is in the position of the pragmatic topic (being active and known in the mind of the addressee). That is why a pronoun is used to refer to it in the second clause, and the rest of the sentence is the focus and provides new information about it. The writer has used a passive construction in order to present the semantic role of patient at the beginning of the sentence as the pragmatic topic and keep the pragmatically unmarked order of topic-focus in the text. In an active sentence, it is the subject that is in the position of pragmatic topic.

(4) iʃan ra moɢrar   gardid   ke afaridegar-e alam-e asrār ast

      They   OM clear   become-PAST that Creator-GEN World-GEN Secrets is

Literal Translation: For them it will be clear (They will realize) that the creator is aware of secrets of the world.

Focus: "afaridegar alam-e asrar ast" (Literal Translation: the creator is aware of secrets of the world.)

Type of focus: predicate focus

Pragmatic Topic: "ishan" (Meaning: they)

In the above passive construction, the pronoun "iʃan" is moved from object position to the initial subject position at the beginning of the sentence. In this period of Persian, passive sentences still use the case marker "*ra*" for the subject of the passive sentences. The referent of this pronoun is "xeradmandan"(wise men) which is mentioned in the previous context. Therefore, it is referred to here by a pronoun. The addressee is aware of this information. From an information structure point of view, this is the pragmatic topic of the sentence, and the rest of the sentence provides new information about it. The writer has kept the unmarked arrangement of topic-focus or old-new information by using a passive construction in order to present the object (or patient) at the beginning of the sentence.

The following two instances of passive structure are taken from the written date set of modern Persian.

(5) in   ʃarayet tɑvɑsoᴛe ideoloʒi tain   miʃavad

      These conditions   by   Ideology detenined-pp becom-PROG-1SG

Literal Translation: "these conditions are determined by ideology"

Focus: "tɑvɑsoṭe ideoloʒi tain miʃɑvɑd"

Type of focus: predicate focus

Pragmatic Topic: "in ʃarɑyet"

In this sentence, the semantic role of patient, "in ʃarɑyet", is promoted to the subject position of the passive sentence as the pragmatic topic. As its referent was previously established within the discourse context, the subject receives definite article marking. The remaining predicate of the sentence conveys new information about this topicalized patient. In most of the passive sentences analyzed in this study, the predominant motivation for promotion of the patient via passivization similarly aimed to represent the patient, or other low-agentive thematic roles such as theme or experiencer, as the pragmatic topic. Based on this pervasive pattern, it can be argued that passive constructions in modern Persian typically function to present a semantically patient-like, low-agency constituent as the point of departure for the ensuing comment.

(6) (bɑrɑye xolasenevisi motuṇe hoɢuɢi) hiʃt olgui erae naʃode ɑst

　　For　　　summarizing　　　texts-GEN legal　no　pattern　present-PP　　is

Literal Translation: For summarizing legal texts, no pattern is provided.

Focus: "hiʃt olgui erae naʃode ɑst" (Literal Translation: No pattern is provided)

Type of focus: sentence focus

Pragmatic Topic: "(bɑrɑye xolasenevisi motuṇe hoɢuɢi" (Meaning: is not provided)

In the above example, there are both preposing and passive structures at the same time. The pre-posed structure, "(bɑrɑye xolasenevisi motuṇe hoɢuɢi, is located to the beginning of the sentence as the topic and the rest of the sentence, "hiʃt olgui erae naʃode ɑst" presents new information about it. The phrase "hiʃt olgui" is the promoted element to the subject position and has the semantic role of the theme. Since this structure carries the infinitive markers, the words "hiʃt" and indefinite "i", it is presented as new information[1]. As a result, it is a part of the focus, along with the predicate. In this passive construction, the type of focus is sentence focus, since the entire clause, subject+predicagte, carries new information.

In contrast to the most of passive sentences analyzed in this study, here the promoted object of the passive sentence is part of the focus. Therefore, it can be concluded that in modern Persian, sometimes passive construction is used to present the semantic role of the patient as the focus and not the topic. This phenomenon can be explained by what Lambrecht (2001) calls the "non-existence of a one-to-one relationship between syntactic structure and pragmatic role".

4.2 Cleft constructions

In this section, some samples of cleft sentences from the three periods of Persian are discussed. There are no cleft sentences in our data set pertaining to the middle Persian texts . The reason for the absence of complex constructions like cleft and pseudo-cleft sentences in Middle Persian could be attributed to its rich inflection and the resulting flexibility in word order. This faculty has allowed for effective sentence focus. The first two examples of cleft sentences analyzed here is from Dari Persian, Tārīkh-i Bayhaqī.

7- "mɑrd an mɑrd ɑst ke goftɑnd　　　ɑlavf　　endɑlɢadre　bekar avɑrɑd"

　　Man that man is　　that say-PAST-3PL forgiveness　when-possible　into-use　bring-3SG

Literal Translation: That man is a man, who is said to be able to use forgiveness.

Focus: "ke goftɑnd ɑlavf endɑlɢadre bekar avɑrɑd" (literal translation: who is said to be able to use forgiveness)

Type of focus: predicate focus

Pragmatic Topic: "mard" (Meaning: man)

Cleft and pseudo-cleft constructions are used when there are several alternatives and the speaker/writer chooses one of the available options using this construction. In this sentence, there are two alternatives for "being a man". One is whoever can forgive, and the other is whoever can take revenge. Therefore, the writer used a cleft sentence to emphasize the first option.

From a pragmatic point of view, the presupposition is that "*someone is a man*", and the addressee doesn't know who he is. In other words, the context already contains the referring phrase "*mard*". However, as searched in the context of this sentence, there was no reference to this word in the previous sentences, and this word is not supposed to be definite. This paradox can be accounted for by a principle called "pragmatic accommodation" by Lambrecht (2001). According to this principle, the speaker/writer invites the addressee to cooperate and supposes that the referent of that expression is already known and available in the context.

---

[1] This means that no referent for this phrase is active in the audience's mind; in other words, its referent is as an unrecoverable or unrecognizable element in the audience's mind.

8- "xorjinaki bud ke ketab dar an minahadam".

    Saddlebag   was that book in   that put-PAST-PROG-1SG

Literal Translation: there was a saddlebag in which I put books.

Focus: "ketab dar an minahadam" (Literal Translation: in which I put books)

Type of focus: predicate focus

Pragmatic Topic: "xorjinak" (Meaning: saddlebag)

    In this sentence, the referent of the phrase "xorjinak", which is the topic of the sentence and the whole sentence explains it, is initially inactive in the mental state of the audience. Therefore, in order for the author to place it as the topic of the sentence and give further information about it, he must first introduce it in the first clause of the sentence for the first time and then in the second clause, put it in the topic position and give new information about it. This technique is referred to by Lambrecht (2001) as the "principle of the separation of Reference and Role" according to which introducing a referent and its pragmatic topic in two separate clauses can reduce the cognitive load on the audience's mind. In other words, the referent should be introduced in the first clause, and new information about it should be provided in a subsequent clause.

    The analysis of other examples from this period shows that the cleft constructions in this period are viable to lend themselves to be studied under this principle (Principle of Separation of Reference and Role), introducing the referent of a noun in the first clause and then providing new information about it in the next clause. This means that, in most cases, the occurrence of such constructions was not intended to emphasize a particular element in the sentence. The reason for this function of cleft constructions can be explained through accounting for the flexible word order of Persian at this time period, which does away with the requirement for intricate structures like cleft and pseudo-cleft sentences. The following sentence exemplifies a cleft construction in Modern Persian.

9- "in adabiyat ast ke zamineye fahme tarix ast"

    This literature is that basis-GEN   understanding literature is

Literal Translation: It is literature that is the basis for understanding history.

Focus: "adabiyat" (Literal Translation: literature)

Type of focus: argument focus

Pragmatic Topic: zamine fahme Tarikh (Meaning: the basis for understanding history)

    This is a two-clause structure with a main clause containing a relative verb and a subordinate clause. According to information structure, the pragmatic topic is in the second clause since the pragmatic topic is that information about which the rest of the sentence provides further information. In this framework, the information that provides new information about the topic serves as the focus. Therefore, the focus in the above sentence is the noun phrase "adabiyat", which the writer puts emphasis upon through this cleft sentence. In fact, this type of marked structure is used where there are several alternatives for the element in the focus position, thereby the writer rejects the other alternatives by using this focal syntactic construction. In the example above, the alternative to the element in the focus position is "tarikh", which might be rejected through the application of a cleft sentence.

4.3 Preposing

In this section, certain examples of prepositional construction in Middle Persian, Dari, and Modern Persian are discussed. It is worth noting that in most of the examples, the object is preposed to the beginning of the sentence. Nonetheless, in Middle Persian, it is also possible for verbs to be placed at the beginning of a sentence, as shown in the following example.

10- pursid   danag o minug i xrad   radih   weh ayab rastih ayab spasdarih.

    Ask-PAST-3SG Wise from Minog-i Kherad forgiveness is or honesty or thankfulness

Literal Translation: Dana asked Minog-i Kherad whether forgiveness is better or honesti or thankfulness.

Focus: "pursīd" (Literal Translation: asked)

Type of focus: predicate focus

Pragmatic Topic: danag o minug i xrad (Meaning: Dana from Minooye Kherad)

    This sentence can be analyzed on the two levels of clause and compound sentence in terms of pragmatic function. At the clause level, the verb is preposed to the beginning of the sentence. In so doing, the focus is placed on the preposed verb, which is. However, considering the entire compound sentence, the clause "pursid danag o minug i xrad" is an element that is used from the beginning of the Mēnōg-ī Khrad text as active and available information in the mind of the addressee and carries old information. This is while the rest of the compound sentence (the subordinate clause) contains new information. The next preposing construction is derived from Dari Persian.

11- -ketabi bozorg ke harʃte andar an baʃɑd dɑr fehrest ɑz an asari padide kardebaʃɑd.

    book   big   that every in   that be-3SG in list   from that trace created   become

Literal Translation: A great book, whatever is inside that book, has created an effect in the list.

Focus: "harʃte andar an baʃɑd dɑr fehrest ɑz an asari padide kardebaʃɑd." (Literal Translation: whatever is inside that book, has created an effect in the list.)

Type of focus: sentence focus

Pragmatic Topic: ketabi (Meaning:book)

The pragmatic topic in this sentence is the referent of the noun phrase "ketabi," about which the entire sentence provides new information. Since the writer primarily intends to introduce the pragmatic topic and then give further information about it, s/he uses two clauses. Through this technique, s/he follows the principle of separation of role and referent in order to decrease the cognitive load of the sentence. In the first clause, s/he introduces the topic, and in the second one, s/he provides new information about it. In this sentence, the pragmatic topic is definite. In Lambrecht's (2001) theoretical framework, definiteness or accessibility in the mind of the addressee is not a necessary condition for the pragmatic topic. An indefinite noun phrase can, under certain conditions, be the pragmatic topic of the sentence.

The next example of preposing structure is from Modern Persian text.

12- "dar dargiri-haye suez 9 nafar koʃte ʃodand"

      In   clashes-GEN   Suez 9 people killed   become-PAST

Literal translation: In the clashes in Suez, nine people were killed.

Focus: "9 nafar koʃte ʃodand" (Literal Translation: nine people were killed)

Type of focus: predicate focus

Pragmatic Topic: "dar dargiri-haye Suez" (Meaning: In the clashes in Suez)

In this sentence, the expression "dar dargiri-haye Suez" is preposed to the beginning of the sentence before the subject "9 people". This movement is triggered by pragmatic motivation. The writer has already referred to "dar dargiri-haye Suez" in the preceding context, and its referent is now active in the mind of the addressee. In a pragmatically unmarked arrangement of information, the old information is mentioned firstly as the point of departure for the new information. Hence, one may conclude that the writer has employed this preposing construction to adhere to the default organization of the information structure and ensure a seamless flow of information.

4.4 Statistical Analysis

In this section, the statistical data of the research is presented. In the following tables and graphs, the percentage of the occurrence of passive, cleft, and preposing constructions in three periods of Middle Persian, Dari, and Modern Persian is presented and compared. As mentioned in the methodology section, each period is represented by a corpus consisting of 1700 sentences, thus reflecting a substantial sample size. Consequently, the following analysis is based on the simple frequencies calculated from these 1700-sentence samples. In the first graph, the frequency of passive structures in Middle Persian is illustrated.
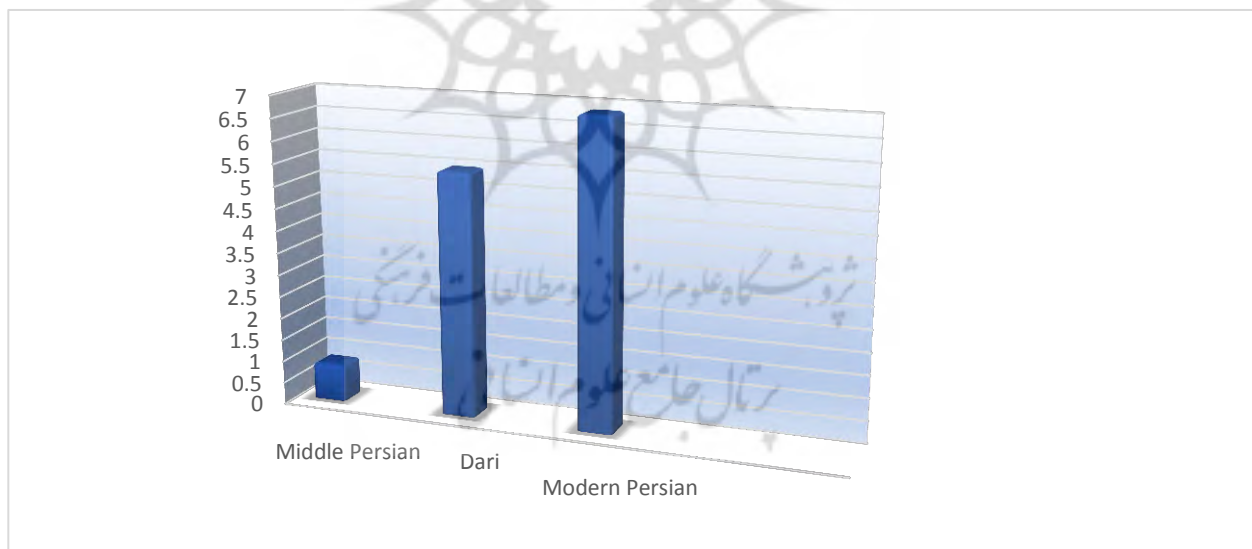
Graph 4-1- Frequency percentage of passive construction in different periods of Persian

Graph 4-1 indicates that the use of the passive construction as a syntactic device for manifesting information structure in Middle Persian is significantly less frequent compared to the two subsequent periods. This difference between Middle Persian and Dari Persian is significant. Furthermore, the transition from Dari Farsi to modern Farsi has increased the frequency of this construction across different texts.
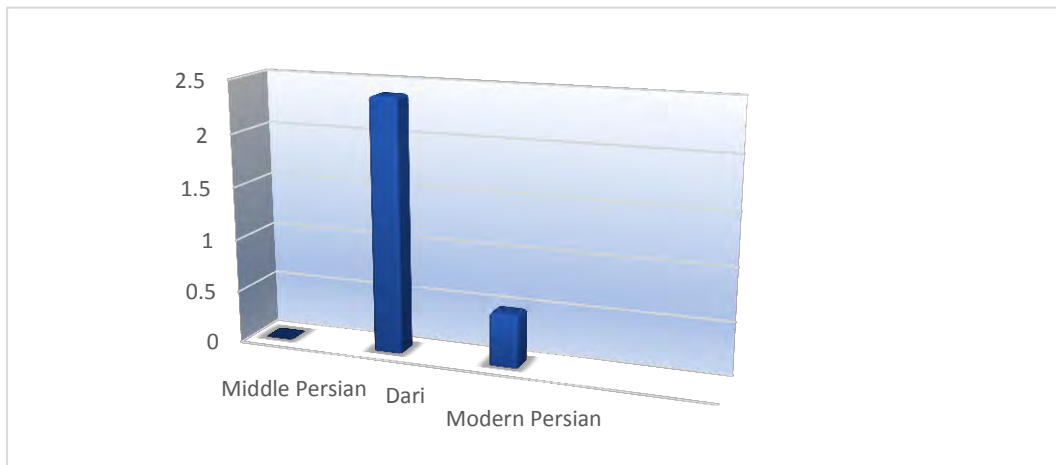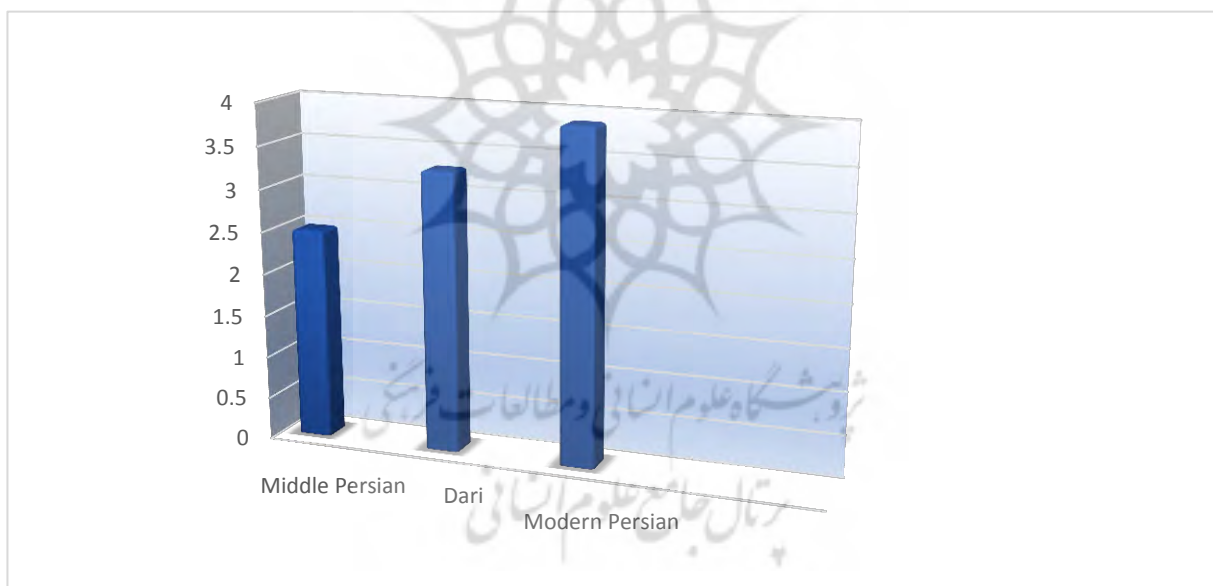
Graph 4-2- Frequency percentage of cleft construction in different periods of Persian

Figure 4-2 displays the percentage frequency of cleft constructions in the three historical periods. Evidently, a structure similar to what is currently referred to as "cleft constructions" was not found in the selected texts from the Middle Persian period. However, the highest frequency of this construction is observed in the selected texts in Dari Persian. In contemporary Persian texts, including narrative, scientific, and news genres, this structure is relatively less common compared to the Dari Persian period.

In what follows, the bar graph 4-3 illustrates the difference in the preposing construction, in which an element other than the subject is moved to the beginning of the sentence, across the three periods of the Persian language.

Graph 4-3- Frequency percentage of preposing construction in different periods of Persian

As it is seen in Figure 4-3, this structure has been increasingly used in Persian Dari to Modern Persian. Additionally, compared to the two former marked structures, namely passive and cleft constructions, this structure has the highest usage in the selected texts of this research.

5. Concluding remarks

The aim of this study was to examine how marked syntactic structures are used to convey information in Persian sentences across three distinct periods: Middle Persian, Dari, and Modern Persian. Data analysis showed that passive structure is usually used in these three periods to present the semantic role of patient at the beginning of the sentence as the pragmatic topic and keep the pragmatically unmarked order of topic-focus in the text. However, there were some cases in which the semantic role of patient was presented as the focus at the beginning of the passive sentence. Thus, it

can be inferred that in modern Persian, passive construction is occasionally employed to highlight the semantic role of the patient as the focus rather than the topic. This phenomenon can be attributed to the absence of a one-to-one correlation between syntactic structure and pragmatic role, as described by Lambrecht (2001). The present findings diverge from the description provided by Vahidian-Kamyar (2009) and Lambton (1983) which posited that Persian passivization is restricted to contexts where the agent is lacking, unnamed by the speaker, or representationally backgrounded information. However, the current research identified instances wherein the passive subject position is occupied by the agent constituent assigned focal prominence, while the patient role serves as the topical element introducing the proposition. This observation contradicts findings of studies such as those promulgated by Vahidian-Kamyar (2009) and Lambton (1983).

The Middle Persian corpus examined in this study did not contain any instances of cleft construction. In Dari, however, cleft construction is employed to present an element of the sentence as the topic and provide new information about it. The writer introduces the element in the first clause and then places it in the topic position in the second clause, according to the principle of the separation of Reference and Role as described by Lambrecht (2001). This technique helps to reduce the cognitive load on the audience's mind by introducing the referent and its pragmatic topic in separate clauses. In other words, the referent is first introduced in the first clause, and new information about it is then provided in a subsequent clause. The analysis of other examples from this period, i.e., Dari, shows that cleft construction in this period of Persian language was used only to comply with this principle and was not used frequently to emphasize an element in the sentence. In Modern Persian, however, this type of marked structure is used when there are multiple alternatives for the element in the focus position and the writer wishes to reject the other options. In other words, this focal syntactic construction is used to exclude other possibilities and emphasize the selected element as the focus of the sentence.

Examples of prepositional construction in Middle Persian, Dari, and Modern Persian were analyzed in this study. Data analysis showed that in most of the examples identified in these three periods, the object is preposed to the beginning of the sentence. Nonetheless, in Middle Persian, it is also possible for verbs to be placed at the beginning of a sentence. In most of the examples analyzed in this study, the writer has employed the preposing construction in order to present an element at the beginning of the sentence as the topic and then provide new information about it. The writers use this marked structure to adhere to the default organization of the information structure, i.e., old-new or topic-focus, and ensure a seamless flow of information. This is in line with Taheri and Gohari (2021) whose findings, based on an analyzed corpus, suggested that in approximately 80% of sentences exhibiting fronting, the proposed constituents functioned to represent already known information and play the role of pragamtic topic in the sentences.

The findings also show that the frequency of the syntactic devices of information structure has changed over time from Middle Persian to Modern Persian. Statistical analysis revealed that the use of passive construction as a syntactic device for conveying information structure in Middle Persian is significantly less frequent than in the two subsequent periods. The difference between Middle Persian and Dari Persian in this regard is significant. Moreover, the transition from Dari Farsi to Modern Farsi has led to an increase in the frequency of this construction across various texts.

No structure similar to what is now known as "cleft constructions" was found in the selected texts from the Middle Persian period. However, the highest frequency of this construction was observed in the selected texts in Dari Persian. The increase in frequency of cleft constructions in Dari can be attributed to the fact that in this period, cleft sentences were used similarly to preposing, in order to observe the principle of separation of role and reference rather than focusing on specific elements. In contemporary Persian texts, including narrative, scientific, and news genres, this structure is relatively less common compared to the Dari Persian period. The observations in the current study align with the account provided by Faghiri and Samvelian (2019) regarding the relative infrequency of cleft construction usage in Persian. Specifically, Faghiri and Samvelian (2019) noted the general paucity of clefting as a syntactico-pragmatic operation in the language based on their analysis.

Similar to passive sentences, the preposing structure has been increasingly utilized in Persian, from the Middle Persian to modern Persian. A possible explanation for the consistent increase in the use of marked syntactic structures could be the decrease in the flexibility of word order from Middle Persian to Modern Persian. Modern persian is discourse configurational language and displays free word order (Faghiri, 2016).

In conclusion, the findings indicated that the frequency and function of these syntactic structures have changed over time, reflecting changes in the flexibility of word order in Persian and changes in the functions of syntactic devices. This study contributes to our understanding of language change and provides insights into the evolution of Persian syntax over time.

Overall, the study offers valuable new insights into how Persian syntax has changed functionally and statistically to adapt to evolving discourse-semantic needs, thus enriching our understanding of both Persian linguistic development and the adaptive relationship between structural and pragmatic systems in a language.

References

Abolghasemi, M. (1993). *The History of Persian Language Tehran*: Samt. [In Persian]

Brown, G., & Yule, G. (1983). *Discourse analysis*. Cambridge: Cambridge University Press.

Dabir Moghaddam, M. (1985). The Persian Passive. *Journal of Linguistics*, 3, 31-46.

Faghiri, P., & Samvelian, P. (2019). A corpus-based description of cleft constructions in Persian. *Faits de Langues*, *52*(1), 183-206.

Faghiri, P. (2016). *La variation de l'ordre des constituants dans le domaine preverbal en persan : approche empirique´*, PhD dissertation, Universit´e Sorbonne Nouvelle .

Fischer, W. B. (1968). *The Canbricige History of Iran*. Cambridge: Cambridge University Press.

Hajat, A. K. (1977). Ke-constructions in Persian: Descriptive and Theoretical Aspects. PhD dissertation, University of Ilinois a Urbana-Champaign.

Hajičová, J., & Panevová, P. Sgall. (2000) *A manual for tectogrammatical tagging of the prague dependency treebank*. Praha: MFF UK

Halliday, M. A. K. (1967). Notes on transitivity and theme in English: Part 2. *Journal of Linguistics*, 3(2) 199-244.

Halliday, M. A. K. (1985). *An introduction to functional grammar*. London: Arnold.

Huddleston, R., and Pullum, G. K. (2002). *The cambridge grammar of english language*. Cambridge: Cambridge University Press

Jabbari, M.J. (2003). *The difference between passive in persian and english*. Linguistics *69*, 78-94.

Lambton, A. K. S. (1983). *Persian grammar*. Cambridge: Cambridge University Press.

Keenan E. L. and Dryer, M.S. (2007). Passive in the World's Languages. In *Clause Structure, Language Typology and Syntactic Description*. (1), 325-361.

Lambrecht, K. (2001). *Information structure and sentence form*. Cambridge: Cambridge University Press.

Langacker, R.W. (1987). *Foundations of cognitive grammar*: Theoretical prerequisites. Michigan:Stanford University Press.

Marashi, M. (1970). *The Persian verb: A partial description for pedagogical purposes*. University Microfilms. Mashhad: Ferdowsi University Press.

Modarresi, B. (2007). Information Structure and its Representation in Persian Unpublished PhD. dissertation. University of Tarbiyat Modares.

Monshizadeh, M. (2020). The Evolution of Verbal Prefix "mi" In Persian Language. *Language Research*. (112). 271-285

Moyne, I. (1974). The so-called Passive in Persian. *Foundation of Language*, 12, 249-267

Palmer, A_ (1971), The EZAFE Construction in Modern Standard Persian, PhD dissertation, University of Michigan Quirk, R. Greenbaum. S. Leech, G., Svantvik, I. (1985). *Comprehensive Grammar of the English Language*. London: Longman

Rezai, V. (2003). A Role and Reference Grammar Analysis of Simple Sentences in Farsi, Ph.D. dissertation. University of Isfahan.

Soheili-Isfahani, A. (1976). Noun phrase complementation in Persian, Unpublished Ph.D. dissertation. University of lion's, Urbana.

Taheri, K. & Gowhari, H. (2021). Information Structure and Thematic Structure of Persian Language A Corpus-based Analysis. *ZABANPAZHUHI (Journal of Language Research)*. 13(38), 213–242. [In Persian]

Vahedi Langrudi, M. (1998). Verbal structures of passive by Shodan in Persian. *Moddares Journal*. Tehran: Tarbiat Modares University press. [In Persian]

Vahidian-Kamyar, T. (2009). *Persian grammar*. Tehran: Samt. [In Persian]