Topicilazation and Focalization in Persian

Vali Rezai
Manijeh Youhanaee
Hasan Ghasemi

University of Isfahan, Iran
Email: Youhanaee_m@hotmail.com

Abstract *This article provides a syntactic analysis of Topicalization and Focalization processes in Persian within the framework of minimalist syntax. According to Haegeman and Gueron's Split CP Hypothesis (1999), Topicalization is a recursive process and can apply to more than one constituent in a sentence, while Focalization is unique and therefore appears in a single projection. In this paper we argue that in Persian more than one constituent might undergo focalization. This implies that Focalization cannot be unique. Thus, we first propose two Focal Phrases in the periphery position of a clause. The specifier positions of the two Focal phrases are the landing sites of the focused constituents. Then we suggest the relative position of focal phrases with topic phrases in Persian. It will be concluded that the first focus phrase appears immediately after the complementizer phrase before the first recursive topic phrase followed by the second focus phrase and recursive topic phrase respectively .*

Keywords: *Focus phrase, given information, new information, Split CP Hypothesis, Topic phrase*

1. Introduction

Topicalization and focalization are two processes in Persian whose constituents do not appear in canonical word order. In these constructions Topics and/or Focus phrases move from their base positions to Periphery positions sentence initially. These discourse motivated movements yield marked orders with specific readings. A number of analyses have been proposed to account for topicalization and/or focalization in Persian e.g. Karimi 2005, Anushe 2010, Darzi & Beyraghdar 2010. These analyses are based on Haegeman and Gueron's split CP hypothesis (1999). However, these analyses suffer from the fact that they are not comprehensive in terms of considering all kinds of clauses (simple versus complex, independent versus dependent) and/or they do not account for sentences which instantiate more than one topicalized and focalized constituents simultaneously. In this study we reevaluate the split CP hypothesis to find the extent to which it can appropriately account for Persian data. So, we present a brief review this hypothesis as suggested by Haegeman & Gueron (1999) and then we discuss Topicalized and Focused phrases in Persian and shortcomings of presented analyses. Finally, we present arguments in favor of a second Focus phrase in the clausal periphery and the sequential orders for CP projections.

2. Topicalization

A variety of positions have been suggested to be the topic position in different languages. As for long-distance topicalization, Chomsky (1977) argues that this movement obeys the same island conditions as wh-movement. This means that topicalization instantiates an A-bar movement to specifier of CP. However, the fact that topicalized constituents can appear after a complementizer indicates that there must be more than one phrasal projection on the top of the TP. Furthermore, evidence from Persian shows that more than one topic might undergo such A-bar movement.

1.   [CP [TopP baĉehâ$_i$ [TopP maŝino$_j$ [TP benazarmiad t$_i$ t$_j$ pasandidand]
                   Children   car   seems   liked
            (It seems that the children liked the car.)

Another position, taken by Haegeman and Gueron (1999), claim that the CP layer like TP includes more than one projection. The first periphery projection they suggest is the topic phrase the head of which bears an uninterpretive topic feature. The specifier position of this projection is the landing site for the topicalized phrase which undergoes movement from its base position and deletes the uninterpretive feature of the head topic. The topic phrase represents old information and is recursive; i.e. there could be more than one topic in a single sentence. The second projection is the Focus Phrase. The head of this phrase bears an uninterpretive focus feature which triggers movement of a constituent bearing an interpretive focus feature to its specifier position. The higher projection in the CP layer is the usual complementizer phrase. Given these assumptions, they propose the following representation for the relative positions of these phrases:

2.   [CP [TopP* [FocP [TopP* [TP …]]]]]

Topicalization in Persian occurs in simple clauses and complex clauses. Movement in complex clauses is licensed both to the periphery of the dependent clause or the matrix clause. Examples (3) and (4) show elements *un mag*â*zaro* 'that shop' and *Lebasâŝo* 'his clothes' are topicalized and appear in the sepecifier of the topic projection:

3.   [TopP un   magâzaro$_i$ [TP ali   t$_i$ xarid]].
              that   shop   Ali   bought

(Ali bought that shop.)

4.   [TopP lebâsâ-ŝ-o$_i$ [TP peyman   t$_i$ ŝost]].
               clothes-his-Acc   Peyman   washed

(Peyman washed his clothes.)

As the examples (3) and (4) show, the two elements *un mag*â*zaro* 'that shop' and *Lebasâŝo* 'his clothes' refer to something that is familiar for the speaker and hearer, so they have given information. Since these elements move to initial position of sentence, they are topicalized constituents. Furthermore, in a simple sentence, two elements can simultaneously move to initial position of sentence as topic. In (5), two constituents *un ketâbro* 'that book' and *be ali* 'to Ali' move to initial position of sentence. These constituents contain given information and therefore are topicalized constituents:

5.   [TopP un ketâbro$_i$ [TopP be ali$_j$ [TP moallem-eŝ   t$_i$ t$_j$ dâde- bud]]]
                that book   to Ali   teacher- his   given- was

(His teacher had given that book to Ali.)

As it is illustrated (3-5), we can explain the simple marked sentences in Persian through Hageman and Gueron's split CP hypothesis (1999). According to this hypothesis, when one element is topicalized such as (3) and (4), the landing site of that element is one of the two specifiers of topic projections in the CP layer.

Also, when two elements move as topic their landing site are two specifiers of topic projections in CP layer such as in (5). In what follows, we will see the given hypothesis can account for compound and complex sentences including one or two topicalized elements.

In compound sentence (6) and complex sentence (7), two elements *barâdare*ŝo 'his brother' and *xunaro* 'the house' are topicalized. Like simple sentences, the landing site of the two topicalized constituents in (6) and (7) is specifiers of two topic projections.

6.   [TopP barâdar-eŝ-oᵢ   [TP ali tᵢ ovord   va   pro   raft   ]]
            brother-his-râ   Ali   brought   and   he   left

(Ali brought his brother and he left).

7.   u   goft   ke   [TopP xunaroᵢ [TP mâ tᵢ pârsal   kharidim]]
            he said   that   house   we   last year   bought

(He said that we bought the house last year.)

In a dependent clause, the topicalized constituents can move to front positions of either the dependent clause or the main clause. In (8) two elements are topicalized. The first topicalized constituent is subject of dependent clause *Soheyl* and the second topicalized constituent (*mâŝino* 'his car') is the direct object of this clause. These elements are topicalized and moved to the front of the main clause. In (9), two topicalized constituents in the dependent clause have moved to the front position of the same clause. So, two topicalized constituents *un ketâbaro* 'that book' *and az anbâr* 'from cellar' have undergone movement to periphery of their own clause.

8.   [TopP soheylᵢ [TopP mâŝinoⱼ [CP benazarmiyâd ke [TP tᵢ tⱼ be khâhar-eŝ hediye dâde]]]]
            Soheyl   car   seems   that   to sister-his present given

(It seems that Soheyl has given the car to his sister as present.)

9.   ali motmaen bud [CP ke [Top un ketabâroᵢ [Top az anbârⱼ [TP mâ tᵢ tⱼ xârej kardim]]]]
            Ali certain   was   that   that books   from cellar   we   out did

(Ali was certain that we took those books from the cellar.)

In examples (8) and (9), two elements are topicalized. As said by Hageman and Gueron's split CP hypothesis (1999), the landing sites of these elements are the two specifiers of topic phrases while the complementizer (ke) is in the head of CP. As we saw in the above examples, we could explain the order of complementizer and two topicalized constituents based on Hageman and Gueron's split CP hypothesis (1999). So, this hypothesis can explicate topicalization process in Persian. In what follows, we will test the given hypothesis when one sentence has simultaneously multiple topicalized and focused constituents.

3. Focalization

Focus is a part of a sentence that bears new information and it has contrast with the ground of sentence which has given information. Focus can be related to phonological structure of language (Vallduvi and Engdahl 1996 and Kiss 1998) Yet, it can also be represented syntactically. Focus is classified in two types: information focus and contrastive focus. Information focus only contains new information. While contrastive

focus not only has new information but also has a kind of contrast with given information of hearer and it corrects his or her earlier presupposition (Vallduvi and Engdahl 1996 and Kiss 1998).

Kiss (1998) and Rizzi (1997, 2003) believe that the position of focus is unique. It is implied that there is one focus phrase per sentence. Also in their hypothesis, Hageman and Gueron (1999) suggest that in the CP layer there is one focus projection. Kiss (1998) proposes that certain elements carry an inherent focus: *only*-phrases and *wh*-phrases that belong to this group. Under this view, Karimi (2005) believes that the position of focus is not unique. It is implied that Persian allows two elements bearing contrastive focus in the same sentence only if at least one of them has an inherent focus feature (Karimi 2005:13):

10.   Kimya   mâhe   gozaŝte   faqat se-tâ   film   dide-Ø
         Kimya   month previous   only   three-classifier movie   seen-is

(It was Kimya who has seen only three movies last month.)

In the above sentence, Karimi (2005) believes that there are two focuses. Subject of sentence *Kimya* is the contrastive focus because it has a contrast to other person who speaker has in her or his mind. Another focus is the object *faqat se tâ film* 'only three-classifier movie' since it has an inherent focus feature presented by *faqat* 'only'. As it is obvious, Karimi (2005) shows that this is plausible that two foci exist in a sentence. Now if these focuses move to the CP layer, where is their landing site? Before we respond to this question it is necessary to introduce classifying of adverbs. Cinque (1999) divides adverbs into two major groups of higher and lower adverbs. He suggests that higher (sentence) adverbs precede lower adverbs. That is, the sequence of two groups cannot change and the higher adverb always precedes the lower one. Under this view, Karimi (2005) divides adverbs into two groups in Persian High (sentence) Adverbs: *amdan* 'intentionally', *xoŝbaxtâne* 'fortunately', zâheran 'apparently' etc.; and Low (VP) Adverbs: *hamiŝe* 'always', *hanuz* 'still', *hargez* 'never' etc. In fact, the high adverbs precede TP and the low adverbs precede vP.

In (11), three elements are moved to the CP layer. One of these elements is the focused constituent and two other are topicalized constituents. The adverb *bâ ziraki* 'cleverly' that appears before VP (its original position), has moved to the CP layer through focalization process and occupies the specifier position of the focus phrase. Additionally, the direct object *penâlti* 'penalty' and the indirect object *az dâvar* 'from referee' have moved to the CP layer by topicalization process and occupy specifier positions of two topic projections.

11.   injâ moŝaxas-e [$_{cp}$ ke [$_{TopP}$ penâltiro$_i$ [$_{FocP}$ bâ ziraki$_j$ [$_{TopP}$ az   dâvar$_k$ [$_{TP}$ pro t$_j$ t$_i$ t$_k$ gerfte-Ø]]]]]
         here clear-is   that   penalty   cleverly   from referee   he   taken-is

(Here, it is clear that he cleverly took the penalty from the referee.)

We have so far been able to show that Persian sentences which have topicalized and focused constituents could be accounted for based on the split CP hypothesis. Moreover, the given hypothesis could explain that kind of sentence which had two topicalized constituents and one focused constituent. We continue with a discussion of the possibility of movement of two focused constituents to the CP layer in a single sentence. If this is the case (as we find below), Hageman and Gueron's split CP hypothesis (1999) cannot account for these kind of sentences.

Focus movement has been considered as an instance of displacement triggered by a discourse-functional feature (Karimi 1999, 2003b and Kiss 2005). Karimi (2005) also believes that Persian does not exhibit obligatory wh-movement, although these phrases are subject to scrambling. If scrambled, they force an answer that indicates a contrastive focus. In fact, wh-movement in Persian is considered as A-bar movement. Any constituent in position of a wh-phrase, such as wh-arguments and wh-adjuncts, receives a focus feature. Similarly any constituent which is a reply to a wh-phrase receives a focus feature too. Wh-

phrases and their replies contain new information; therefore they are focused (Robert & Van valin 2005, Vallduvi and Engdahl 1996). "A very common example of a narrow focus sentence is a WH-question like *What did you buy?* and the answer *I bought …* ; the WH-word and the NP filling the slot in the reply are both unmarked narrow foci. In a yes–no question like *Did* John *leave?* and the response *No,* Fred *did*, *John* and *Fred* are marked narrow foci" (Van valin 2005: 72). Under this view, in (12) and (14), two wh-phrases *koja* 'where' and *key* 'when' are focused constituents. Also, in (13) and (15) the replies *dâneŝgah* 'university' and *do-ruz piŝ* 'two days ago' are focused constituents too. These sentences can be analyzed based on the split CP hypothesis. the wh-pharses in (12 & 14) as well as their replies in (13 & 15) occupy the specifier of the focus projection.

12.   [$_{FocP}$ kojâ$_i$ [$_{TP}$ pro   t$_i$ miri]]
          Where   you   go

(Where do you go?)

13.   [$_{FocP}$ dâneŝgâh [$_{TP}$ pro t$_i$ miram]]
         university   I   go

(I go to university.)

14.   [$_{FocP}$ key$_i$ [$_{TP}$ pro   t$_i$ az   ŝirâz   bargaŝti]]
         When   you   from   Shiraz   returned

(When did you return from Shiraz?)

15.   [$_{FocP}$ do ruz piŝ$_i$ [$_{TP}$ pro   t$_i$ az   ŝirâz   barghaŝt-am]]
         two days ago   I   from Shiraz   returned

(I returned from Shiraz two days ago.)

Rizzi (1997) and Haegeman (2000) have argued that just focused constituents occupy the specifier position within a focus phrase and that topicalized constituents occupy the specifier position within a topic phrase (Radford 2004: 329). The data from Persian show that there could be two focused constituents in a sentence.

16.   [$_{FocP}$ barâ kesiyam$_i$ [$_{FocP}$ ĉizi$_j$ [$_{TP}$ pro   t$_j$ t$_i$ soqât   ovorde budan]]]?
        for   somebody   something   they   souvenir   brought   were

(Had they brought any souvenir for anybody?)

17.   [$_{FocP}$ faqat ye kif$_i$ [$_{FocP}$ barâ narges$_j$ [$_{TP}$ amu t$_i$ t$_j$ gerfte- bud]]]
        only   a   purse   for   Narges   uncle   taken-was

(Only a purse had Uncle bought Narges.)

18.   be-hem   migi [$_{CP}$ ke   [$_{FocP}$ kio$_i$ [$_{FocP}$ bâ   ki$_j$ [$_{TP}$ pro   t$_i$ t$_j$ didan-Ø]]]]?
         to-me   tell   that   who   with whom   they   seen-are

(Lit: Will you tell me who with whom they have seen?)

19.   [FocP saido_i [FocP bâ   maryam_j [TP diruz   baĉehâ   t_i   t_j   didan]]]
            Saeed   with Maryam   yesterday guys   saw

(The guys saw Saeed and Maryam together yesterday.)

In the above sentences, there are two focused constituents: wh-phrases and its replies. For example in (16), two wh-phrases *barâ kesiyam'* To somebody' and *ĉizi* 'something' are focused constituents because they have inherent focus feature. Also in (17), two constituents which are a reply for wh-phrases are focused constituents because they have new information.
Similarly, in (20), focused constituents are wh-phrase *ĉerâ* 'why' and object *faqat abbâs* 'only Abbâs

20.   fahmid-i [CP ke   [FocP ĉerâ_j [FocP faqat_i abbâso [TP dâneŝgâh t_i burs   karde-Ø t_j]]]]
            learned-you that   why   only Abbâs   university scholarship done-has

(Did you learn why, only to Abbas, has the university given a scholarship?)

As it was noted in the previous section, Hageman and Gueron (1999) suggest that the CP layer has only one focus projection. So, this hypothesis is not an appropriate one to account for Persian sentences that have simultaneously two focused constituents. Karimi (2005:137) argues that the two focused constituents cannot be separated; therefore, she suggests multiple specifiers for the sole focus projection. That is, she claims that the focused phrases occupy two specifier positions of the focus projection.

21.   [TopP ki_i bâ   ki_j [TP to   [vP fekr   mikoni [CP (ke)   t_i t_j beraqse]]]?
            who with who   you   think do   that   dance

(Li: Who with whom is it that you think will dance?)

22.   ?? ki_i   emruz bâ   ki_j   pro   fekr mikoni [CP (ke) t_i t_j beraghse]?
            who today with who   think do   that   dance

(Li: Who with whom is it that you think will dance today?)

However, this position cannot be borne out for the following reasons below based on which we conclude that, at least in Persian, two focused phrases can be maintained.
    First, it should be noted that while the sentence in (22) is highly marked , it is acceptable. Second, the verb *raqsidan* 'dance' is a predicate which has one argument *ki* 'who'. So, we must consider the focused constituent *bâ ki* 'with whom' as a wh-adjunct. This is contrast to Karimi's analysis in which wh-adjuncts cannot move to specifier of the focus projection.
    Third, according to Hornstein et.al. (2005), a movement operation is licensed only if it leads to elimination of an uninterpretive formal features. Under this view, if a constituent moves to another position, either the constituent itself or the head of the projection it moves to must have an uninterpretive feature. Under Karimi's multiple specifier view, we should assume that the two focused phrases bearing an uninterpretive feature would check the uninterpretive feature of the focus head twice. But, this is not a plausible analysis based on Chomsky (1995) whereby movement of one constituent deletes the uninterpretive feature of the functional category. In addition, data show that a constituent can appear between two focused elements in Persian.

23.   [FocP kio$_i$ [FocP bâ   ki$_j$ [TP doktor t$_i$ t$_j$ un ruz   eŝtebâh- gerefte- bud]]]?
          Who   with whom   doctor   that day   mistake-taken-was

(Lit: Who with whom is it that the doctor had confused that day?)

This sentence (23) can be replied in three different ways, with slightly different interpretation:

24.   [TopP un ruz$_i$ [FocP man-o$_j$ [FocP bâ   dâdâŝ-am$_k$ [TP doktor t$_i$ t$_j$ t$_k$ eŝtebâh -gerfte -bud]]]]
          that day   I-Acc   with brother-my   doctor   mistake-taken-was

25.   [FocP man-o$_j$ [FocP bâ dâdâŝ-am$_k$ [TopP un ruz$_i$ [TP doktor   t$_i$ t$_j$ t$_k$ eŝtebâh -gerfte -bud]]]]
          I-Acc   with brother-my   that day   doctor   mistake-taken-was

26.   [FocP man-o$_j$. [TopP un ruz$_i$ [FocP bâ   dâdâŝ-am$_k$ [TP doktor t$_i$ t$_j$ t$_k$ eŝtebâh -gerfte -bud]]]
          I-Acc   that day   with brother-my   doctor   mistake-taken-was

(The doctor had confused me with my brother that day.)

In the above sentence there are three fronted constituents. The constituents *man* 'me' and *bâ dâdâŝam* 'with my brother' are focused constituents because they are replies to wh-phrases and have new information. The constituents *un ruz* 'that day' is mentioned in previous discourse and because it bears given information is a topicalized constituent. Sentence (26) illustrates a topicalized constituent (i.e. *un ruz* 'that day') is placed between two focused constituents, contrary to Karimi's analysis (2005).

And finally, considering more examples, a variety of sequential orders can be observed, specifically in sentences which instantiate two Focused as well as two topicalized fronted constituents. Under each example below, the relevant configurationally representation indicates the order of the moved constituents.

27a.   [CP [TopP ketâbo$_i$ [TopP be minâ$_j$ [FocP ĉerâ [TP said   t$_i$ t$_j$ dâde-Ø   t$_k$]]]]?
            book   to Mina   why   Saeed   given-is

[CP [TopP [TopP [FocP [TP]]]]]

27b.   [CP [TopP ketâbo$_i$ [FopP ĉerâ [TopP be minâ$_j$ [TP said   t$_i$ t$_j$ dâde-Ø   t$_k$]]]]?
            book   why   to Mina   Saeed   given-is

[CP [TopP [FocP [TopP [TP]]]]]

(Why has Saeed given the book to Mina?)

28a. [CP [FocP ĉerâ$_i$ [FocP hanuz$_k$ [TopP in   duste-ŝ-o$_j$ [TP sohrâb t$_k$ t$_j$ be-mâ moarrefi   na-karde-Ø t$_i$]]]]?
            why   yet   this friend-his-Acc   Sohrâb   to us   introduce no-done-is

[CP [FocP [FocP [TopP [TP]]]]]

28b. [CP [FocP ĉerâ$_i$ [TopP in-dust-eŝ-o$_j$ [TopP   be-mâ$_g$ [FocP hanuz$_k$ [TP sohrâb t$_k$ t$_j$ t$_g$ moarrefi na-kard-e t$_i$]]]]]?
            why   this-friend-his-Acc to-us   yet   Sohrâb   introduce-no-done-is

[CP [FocP [TopP [TopP [FocP [TP]]]]]]

28c. [CP [FocP ĉerâi [FocP hanuzk [TopP in-dust-eŝ-oj   [TopP be-mâg [TP sohrâb tk tj tg   moarrefi na-kard-e ti]]]]]]?
            Why   yet   this-friend-his-Acc   to us   Sohrâb   introduce-no-done-is

[CP [FocP [FocP [TopP [TopP [TP]]]]]]

28d. [CP [FopP ĉerâi [TopP in-dust-eŝ-oj [FocP   hanuzk [TopP be-mâg [TP sohrâb tk tj tg moarrefi na-kard-e ti]]]]]]?
            Why   this-friend-his-Acc   yet   to us   Sohrâb   introduce-no-done-is

[CP [FocP [TopP [FocP [TopP [TP]]]]]]

(Why has not Sohrâb introduced his friend to us yet?

Given the arguments above and considering all the given examples, it seems that we must postulate a second focus projection in the CP layer in Persian. In order to account for the discussed data so far the following configuration is proposed for periphery positions in Persian (The * indicative recursive property)..

 [CP [FocP [TopP* [FocP [TopP* [TP]]]]]]

4. Conclusion

In this paper we analyzed Persian CP layer based on the Split CP Hypothesis proposed by Haegeman and Gueron (1999). We discussed discourse motivated movements to CP projections both in simple and complex sentences and to main and dependent clauses respectively. Our data, however, indicated that the Split CP Hypothesis cannot account for all the data from Persian. We also found that Karimi (2005)'s analysis assuming multiple specifiers for the focus phrase was not compatible with the data. Finally, we proposed a second focused phrase in Persian and a representation which successfully accounts for all the observed data.

References

Anushe, M. (2010). Topicalization and Focus movement in Persian: A Feature-Based Approach. Language research 1,1: 1-28
Chomsky, N. (1995). The Minimalist Program, Cambridge, Mass.: MIT Press.
Den Dikken, M. (2006). Relators and linkers: The Syntax of Perdication, Predicate Inversion, Copulas. Cambridge: MIT Press.
Darzi, A & Beyraghdar, R. M. (2011). A minimalist approach to the landing site of Persian topics. Research in Linguistics 2, 1: 1-18
Haegeman, L, and J. Gueron (1999). English Grammar: A Generative Perspective. Oxford: Blackwell Publishers.
Hornstein, N; Nunes, J. and Grohmann, K. (2005). Understanding Minimalism. Cambridge: Cambridge University Press
Jackendoff, R. (1972). Semantics Interpretation in Generative Grammar, Cambridge, Mass.: MIT Press.
Karimi, S. (1997). Persian complex verbs: idiomatic or compositional. Lexicology 3 (2): 273-318.
Karimi, S. (2003). Focus Movement and the Nature of Uninterpretable Features, In Carnie, Andrew, Heidi Karimi, S. (2005). A Minimalist
        Approach to Scrambling, Evidence from Persian, New York: Mouton de Gruyter.
Harley and Mary Willie (eds) Formal approach to Functional Focus, John Benjamins.
Kiss, K.E. (1998). Identificational Focus and Information Focus, Language 74:245-273.
Radford, A (2004). Minimalist Syntax: Exploring the Structure of English, Cambridge: Cambridge University Press.
Robert, D, and Van Valin, JR. (2005). Exploring the syntax-Semantics Interface, Cambridge: Cambridge University Press
Rizzi, L. (1997). The Fine Structure of the left Periphery, in L. Haegeman (ed.), Element of Grammar, 281-337, kluwer Academic
        Publishers
Rizzi, L. (2003). Locality and Left Periphery, to appear in A. Belletti (ed), Structures and Beyond: the Cartography of Syntactic
        Structures, vol. 2, Oxford: Oxford University Press.
Valluvi, E. & Engdahl, (1996). The linguistic realization of information packaging. Linguistics 34.