Revisiting the rā-marked vs. non-rā-marked dichotomy in the analysis of the Persian VP

Pegah Faghiri (Université Sorbonne Nouvelle)

Building on extensive quantitative studies on word order preferences in ditransitive constructions, this paper proposes to revisit the prevailing analysis of syntactic structure of the Persian VP which assumes two distinction syntactic positions for the DO depending on whether it carries the enclitic =*rā* or not.

The prevailing view of the Persian VP initially suggested in generative studies assumes a hierarchical structure, mainly motivated by the existence of differential object marking (DOM), which requires definite/specific DOs to be marked by =*rā*. In this view, rā-marked DOs are assumed to occupy a higher syntactic position than their non-*rā*-marked counterparts do (e.g. Browning & E. Karimi 1994; Ghomeshi 1997; Karimi 2003; Ganjavi 2007; Modarressi 2014). More precisely, in line with studies such as Diesing (1992), these studies claim that unmarked DOs are VP internal while rā-marked DOs are VP external. Note that most of these studies do not formulate their claim in terms of DOM but in terms of a binary feature such as specificity assumed to trigger *rā*-marking (cf. e.g. Karimi 2003:91). The phrase structures in (1), adopted from Karimi (2003:105), illustrate these two syntactic positions.

(1) a. $[_{VP} \, DP_{[+rā]} \, [_{V'} \, PP \, V]]$

      b. $[_{VP} \, [_{V'} \, PP \, [_{V'} \, DP_{[-rā]} \, V]]]$

(2) a. (S) $OD_{[+rā]}$ OI V

      b. (S) OI $OD_{[-rā]}$ V

The main argument put forward in support of this view is a broadly admitted hypothesis on the canonical/neutral/unmarked word order between the direct and the indirect object (IO). As illustrated in (2), it is assumed that in a neutral word order a *rā*-marked DO precedes while an unmark DO follows the IO (e.g. Browning & Karimi 1994; Mahootian 1997; Karimi 2003; Rasekhmahand 2004), as illustrated in (2). However, this hypothesis lacks solid empirical underpinning and has been put to question by recent quantitative studies (e.g. Faghiri & Samvelian 2014; Faghiri *et al.* 2014).

In order to have a better understanding of the issue at stake, let us consider the definition of the canonical word order. One of the most consensual definitions stipulates that the dominant word order is the order in which constituents appear in least pragmatically and stylistically marked or neutral sentences (cf. Siewierska 1988). Meanwhile, identifying the dominant word order among semantically, or truth-conditionally, equivalent word orders is not self-evident, especially based on pragmatic unmarkedness, which is not a straightforward notion (cf. e.g. Lambrecht 1996). Frequency is one single criterion upon which most authors rely to identify the dominant word order (if any) among competing orders in a given language. Naturally, in some cases intuitions are strong enough to be regarded as robust and reliable. Nevertheless, in other cases, where the difference between available alternatives are more subtle, intuitions are less stable. Regarding the relative order between the DO and the IO in Persian, while identifying the canonical order is straightforward for *rā*-marked and bare (single-word) DOs, as in ex. 3-a and 3-b, it is not for other types of unmarked DOs. For instance, some reference grammars accept both word orders for unmarked DOs carrying the indefinite enclitic =*i*, as illustrated in ex. 3-c (Givi Ahmadi & Anvari, 1995:305). In such cases, the quantitative approach, provided methodological standards are respected, remains the most reliable way to identify the canonical order.

(3) a. Yusef **ketāb=rā** be ketābxane dād
      Y   book=RA to library   gave
      'Yusef gave the book to the library.'
    b. Yusef az   ketābxāne **ketāb** gereft
      Y   from library   book took
      'Yusef took a/some book(s) from the library.'
    c. (i). Yusef **ketāb=i**   az   ketābxāne gereft
        Y   book=INDF from library   took
      (ii). Yusef az   ketābxāne **ketāb=i**   gereft
        Y   from library   book=INDF took
      'Yusef took a book from the library.

Figure 1 Word order by DO type

Indeed, recent quantitative studies on word order preferences between the DO and the IO, including a series of sentence production experiments, have allowed for a better understanding of word order preferences in the preverbal domain. Figure 1 summarizes the distribution of word order (*i.e.* DO-IO-V vs. IO-DO-V) for the above-mentioned three DO types in the sentence productions of 60 Persian native speakers, see Faghiri (2016:182-196) for details. As it is clearly illustrated by the data, the predictions of (1) and (2), are only verified for bare (single-word) and *rā*-marked DOs. Unmarked DOs carrying an indefinite determination, contrary to these predictions, group with *rā*-marked DOs, rather than bare DOs, and prefer the DO-IO-V order. At first sight, this implies that bareness should replace *rā*-marking in (2) in order to obtain an empirically valid generalization: non-bare DOs precede while bare DOs follow the IO. Meanwhile, the degree of determination (*i.e.* zero, indefinite, =*rā*) obtained by crossing the two features provides even better predictions. Furthermore, functional factors such as relative length and animacy are also shown to be involved. Interestingly, bare modified DOs (ex. *ketāb=e amuzeš=e akkassi*) are shown to display a significantly less strong preference to appear adjacent to the verb than their single-word counterparts. In sum these studies show that: 1) any generalization based solely on *rā*-marking is too strong and yields wrong predictions and 2) the linear position of the DO can be accounted for in terms of soft/functional constrains (*i.e.* the more the DO is salient and/or (semantically) independent from the verb, the more it is likely to be separate from the verb and thus precede the IO).

These findings strongly undermine the widespread dichotomous view of the Persian VP, while they are compatible with a flat structure view (cf. Samvelian 2001; Bonami & Samvelian 2015). In addition, in line with these findings, a close examination of the other existing arguments put forward in favor of this view (Faghiri & Samvelian 2016; Faghiri 2016:220-256) shows that there exists no conclusive evidence justifying a dichotomous view of the Persian VP.

References

Bonami, O. & P. Samvelian. 2015. The diversity of inflectional periphrasis in Persian. *Journal of Linguistics* 51(2). 327–382.

Browning, M. & E. Karimi. 1994. Scrambling to object position in Persian. In N. Cover, H. van Riemsdijk (eds), *Studies on scrambling*, 61-100.

Diesing, M. 1992. *Indefinites*. Linguistic Inquiry Monographs. Cambridge, Mass: MIT Press.

Faghiri, P. 2016. *La variation de l'ordre des constituants dans le domaine préverbal en persan : approche empirique,* PhD Dissertation, Université Sorbonne Nouvelle Paris 3.

Faghiri, P. & P. Samvelian. 2014. Constituent ordering in Persian and the weight factor. In Christopher Pinon (ed.), *Empirical issues in syntax and semantics 10*, 215 – 232, CNRS.

Faghiri P. & P. Samvelian. 2016, How much structure is needed? The case of the Persian VP, In D. Arnold, M. Butt, B. Crysmann, T. Holloway King & S. Müller (eds.) *Proceedings of the Joint 2016 Conference on Head-driven Phrase Structure Grammar and Lexical Functional Grammar,* 236–254, Stanford, CA: CSLI Publications.

Faghiri, P. *et al*. 2014. Accessibility and Word Order: The Case of Ditransitive Constructions in Persian. In S. Müller (ed.): Proceedings of the 21st HPSG Conference, 217–237. CSLI Publications.

Ganjavi, S. 2007. *Direct objects in Persian*, Doctoral dissertation, University of Southern California.

Givi Ahmadi H. & H. Anvari. 1995. Dastur zabāne fārsi [The grammar of Persian]. Mo'assese farhangi Fātemi.

Lambrecht, K. 1996. *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*. Cambridge University Press.

Karimi, S. 2003. Object positions, specificity and scrambling. In Karimi, S. (ed.) *Word Order and Scrambling*, 91-125.

Mahootian, S. 1997. *Persian*, New York: Routledge.

Modarresi, F. 2014. Bare nouns in Persian: Interpretation, grammar and prosody, PhD dissertation, Humboldt-Universität zu Berlin.

Rasekhmahand, M. 2004. Jāygāh maf'ul mostaqim dar fārsi [The position of the direct object in Persian]. *nāme farhangestān* 6:56–66.

Samvelian, P. 2001. Le statut syntaxique des objets nus en persan. *Bulletin de la Société de Linguistique de Paris* 96(1). 349–388.

Siewierska, A. 1988. *Word order rules*. Routledge.