The semantics of the Persian object marker *-râ*[*]

Maryam Hosseini Fatemi
Carleton University

SUMMARY

This paper presents an analysis of the semantics of the Persian object marker *-râ*. The semantics of this morpheme has been identified with definiteness marking (Sadeghi, 1970; Vazinpour, 1977), specificity marking (Browne, 1970; Karimi, 1989, 1990, 1996, 2003) and presupposition marking (Ghomeshi, 1996). In this study, I will challenge the assumptions presented in previous works and will argue that while definiteness, specificity and presupposition capture important aspects of the meaning of *-râ*, none of them adequately characterize its semantics. Specifically, I argue that previous proposals do not account for occurrences of *-râ* in question words, contrastive topics, donkey sentences, plurals and interactions with indefinites. I argue that a unified analysis can be given if we assume that *-râ* has the semantic property of picking out the maximal member of the denotation of its argument.

RÉSUMÉ

Ce document présente une analyse de la sémantique d'un marqueur d'objet direct Perse *-râ*. Ce morphéme a été identifié à partir d'un marqueur définitif (Sadeghi, 1970; Vazinpour, 1977), un marqueur de spécificité (Browne, 1970; Karimi, 1989, 1990, 1996, 2003) et un marqueur de présupposition (Ghomeshi, 1996). Dans le cadre de cette étude, je contesterai les hypothéses pésentées dans les travaux précédents et ferai valoir que tandis que des marqueurs définitifs, de spécificité et de présupposition capturent les aspects importants du sens de *-râ*, aucun d'entre eux caractérisent adéquatement sa sémantique. Plus précisément, je soutiens que les propositions antérieures ignorent les occurrences de *-râ* pour les pronoms interrogatifs, thémes contrastes, « donkey sentences », les formes plurielles et les interactions indéfinis. Je soutiens qu'une analyse unifiée peut être offerte avec la supposition que *-râ* a la propriété sémantique de choisir le membre maximal de la désignation de son argument.

---

1   INTRODUCTION

Object noun phrases are marked with *-râ* if some semantic/pragmatic conditions are satisfied. Primarily, *-râ* seems to mark noun phrases for definiteness as shown in (1a),[2] it obligatorily appears with proper names, pronouns, nouns modified by a demonstrative, etc. But it also co-occurs with indefinite determiners (i.e. *-i*, *-ye(k)*) as shown in (1b) which should be impossible if it marks definiteness.

(1)   (Ghomeshi, 1997, 138.6)

    a.   Sara ketâb-o xarid.
       Sara book-râ bought.3SGs
       'Sara bought the book.'

    b.   man ketâb-i-ro   xarid-am.
       I   book-IND-râ bought-1SGs
       'I bought a (certain/particular) book.'

The co-occurrence of *-râ* and indefinites present a problem for the semantics of this morpheme. Therefore, researchers have appealed to a variety of notions to account for the meaning of *-râ* marked indefinites. The most prominent semantic proposal is that *-râ* marks noun phrases for specificity (Browne, 1970; Karimi, 1990).

2   -RÂ AS A SPECIFICITY MARKER

The semantics of the direct object marker *-râ* has been commonly equated with specificity marking (Karimi, 1990). According to Karimi, a noun phrase is specific if the speaker has a particular referent in mind. Specificity is defined as the selection of a particular individual from a set of individuals. She divides direct object NPs into specific (definite/indefinite) and non-specific (indefinite/generic), where only the specific ones can be marked with *-râ*. Karimi (1990) applies the following test to show the difference between indefinite specific and indefinite nonspecific use of a noun phrase (cf. Karttunen, 1976).

(2)   Mary was looking for a pen, and

    a.   she found one. (Nonspecific)
    b.   she found it. (Specific)

Under Karimi's definition, a specific definite has a referent that is known to the speaker and presupposed to be known to the hearer. A specific indefinite on the other hand, denotes a referent that is known to the speaker but presupposed not to be known to the hearer. The referent of a non-specific NP is unfamiliar to both speaker and hearer. For example, (2b) implies that the speaker has a specific book in mind or is referring to. 'N-râ' can be used when there is no specific N that is referred to or that the speaker has in mind (similar issues arise with other previous analyses of the semantics of *-râ*).

---

[2] The morpheme *-râ* is pronounced as [râ] in formal contexts, while it appears as [ro] or [o] in colloquial language.

3   DIFFICULTIES WITH THE SPECIFICITY ANALYSIS

As mentioned earlier, a commonly accepted proposal on the semantics of *-râ* is Karimi's specificity analyses. While Karimi's proposal offers valuable insights, it seems to face some difficulties arising from the observations that some specific noun phrases do not appear with *-râ* and that 'N-râ' does not refer to a specific N. Dabir-Moghaddam (1992) and Ghomeshi (1997) argue that specific noun phrases are not always followed with *-râ* as shown in (3).

(3)   Ali mi-xast   ye ketâb-i   peyda kone   va peydâ-sh kard
      Ali CONT-wanted.3SG one book-IND find do.3SG and find-it   did.3SG
      'Ali wanted to find a book and he found it.' (Dabir-Moghaddam, 1992, 557)

In (3) the clitic (*-sh* 'it') in the second part refers back to *ye ketâb-i* 'a book' ; however, *ye ketL'b-i* while being specific is not followed with *-râ*.

Another problem with the specificity analysis is Dabir-Moghaddam (1992) and Ghomeshi (2003) observe that generics in direct object position can appear with *-râ* as indicated in (4) and (5).

(4)   Sirka   shir-râ mi-burrad.
      vinegar milk-râ DUR-curdle.3SG.S
      'Vinegar curdles milk.' (Dabir-Moghaddam, 1992, 557)

(5)   xod-at   ke mard-hâ-ro mi-shenas-i.
      self-2SG that man-PL-râ DUE-know-2SG.S
      'You yourself know men.' (Like Water for Chocolate, p. 130, by Laura Esquivel, translated by Maryam Bayat, from Ghomeshi, 2003 :51)

While the direct objects in (4) and (5) are both followed with *-râ* , they do not receive a specific reading. In fact, there is no particular entity that is being referred to or that the speaker has in mind.

Another problem for the specificity analysis, previously unnoticed in the literature, is that *-râ* can anaphorically pick up indefinite antecedents in 'donkey' sentences. For instance, in (6) below, the *-râ* marked *-ketâb* 'book' doesn't refer to a specific book ; in fact, there is no book in the world that is the referent of the *-râ* marked *-ketâb*.

(6)   Agar Sârâ ketâb dâsht,   ketâb-a-ro   (un-o) mi-xund.
      If   Sara book had.3SG, book-DEF-râ (it-râ) DUE-read.3SG
      'If Sara had a book, she would read the book (it).'

Finally, *-râ* obligatorily appears on some *wh*-phrases, such as *which* and *who* :

(7)   Târâ kodum ketâb-hâ-*(ro) xund ?
      Tara which book-PL-râ   read.3SG.S ?
      'Which books did Tara read ?'

(8)   Ki-*(ro) diruz   tu mehmuni did-i ?
      Who-râ yesterday in party   saw-2SG.S ?
      'Who did you see in the party yesterday ?'

Browne (1970) and Karimi (1990) have proposed that *wh*-phrases *ki* ('who') and *kodum* ('which') are specific in nature. Support for this conclusion comes from distributional evidence specially co-occurrence of these *wh*-phrases with *-râ*. However, in (8), again, there is no particular book that the speaker has in mind here.

In sum, if by specificity we mean that there is a particular referent that the speaker has in mind, then the data in (3)-(8) remain unexplained. One approach to resolving these difficulties might be to formulate an alternative analysis of "specificity" that accounts for its non-referential use. I will take a different perspective, one that divorces *-râ* from reference. Specifically, I will argue that *-râ* is a maximality operator. The starting point of my analysis is Link's proposal for the meaning of 'the'.

4   PROPOSAL

4.1   *-râ* AS A MAXIMALITY OPERATOR

Link (1983) argues for a unified semantics of the definite determiner "the" which accounts for both aspects of the meaning of the determiner; namely, *uniqueness* and *universality*. When "the" appears with singular count nouns, it expresses the uniqueness of the referent of the NP and when used with plural count nouns it functions as a universal quantifier.

Link's semantics is based on an ontologically rich model structure. This model includes not only *atomic objects* but also plural *non-atomic objects*, which are the *sums* of the atomic objects. Atomic and non-atomic objects in Link's model are ordered by the *part-of-relation*. To take a concrete example, suppose there are three books in the domain of discourse : *Blue book*, *Green book* and *Red book*, b, g and r say. Then, the denotation of the singular noun *book* will be the set $\{b, g, r\}$, and the plural noun *books* will denote $\{b+g, g+r, b+r, b+g+r\}$. In this example, the denotation of the books in Link's terminology is the maximal element of the set [[books]] that is the unique element of the set which includes all the other elements as its parts. In the case of singular definite NPs, consider the singular NP the book in "I bought the book." The noun book denotes the set of atoms which are books and contains no non-atomic objects. Since all the elements of the set are atomic and there is no part-of-relation, then max([[book]]) is defined iff (if and only if) there is exactly one book in the domain of discourse. For example, in the above context where [[book]] = $\{b, g, r\}$, max([[book]]) will be undefined, since there is no maximal element of the set (the objects b, g, and r are not ordered by the part-of relation). The only way to get max to be defined with atomic objects is if there is one object, and only one object, say [[book]] = $\{g\}$, in which case max([[book]]) = g. In the case of plurals, where [[books]] = $\{b+g, g+r, b+r, b+g+r\}$, max([[books]]) = b+g+r (the maximal element in the ordering). Link's proposal identifies *the* with max. Returning to Persian, in the case of singular NPs as in (1a), repeated below as (9), the noun *ketab* 'book' denotes the set of atoms which are books and contain no non-atomic objects. In such a context, [[ketab]] = $\{k1\}$, and the maximality operator *-râ* picks out k_1. Since all the elements of the set are atomic and there is no part-of-relation, max ([[*ketab*]]) is defined iff there is exactly one book in the domain of discourse.

(9)   Sara ketâb-o xarid.
        Sara book-râ bought-3SG.S
        'Sara bought the book.'

(10) Sara ketâb-hâ-ro xarid.
      Sara book-PL-râ bought-3SG.S
      'Sara bought the books.'

Unlike previous analyses, the maximality approach correctly predicts that when *ketab* is plural-marked, *ketâb-hâ*, as in (12),[3] appending *-râ* to this will pick out all the books in the context. Thus, the morpheme *-râ* on the plural noun *ketâb-hâ* 'books', collects the unique non-atomic element of the set which is the sum of all the books. Following Link, I assume that *-râ* has the same semantic property of picking out the maximal member of the denotation of its argument.

(11) ***-râ* as a Max Operator (revised in (14)**, (17)) : [[*N-râ*]] = Max([[N]])

By identifying the semantics of "N-râ" in Persian with the semantics of "the-N" in English, the account readily extends to donkey-sentences, as in (6), and to generics, as in (4). Under the maximality approach the meaning of the donkey sentence in (6) would be analysed like the corresponding English paraphrase : *If Sara had a book, she would read the book*. The donkey anaphora usage in (6) can be accounted for following from a situation semantics approach where the antecedent introduces a minimal situation with a book in it, e.g., Heim (1990) ; Elbourne (2005) ; *ketâb-râ* then denotes this unique book. However, under the specificity definition this meaning cannot be derived. The maximality approach to *-râ*, while not directly predicting that *-râ* should appear with generics, becomes unified with the observation that generics cross-linguistically appear with the definite article (Krifka et al. (1995) ; e.g., *The potato was first cultivated in South America*).

The definition in (12) does not however extend in any obvious way to questions. In the next section I propose that if we allow *-râ* some flexibility in the semantic types that it can assume, while constraining it to be a max operator, we can appeal to maximality approaches to the semantics of questions (most directly, Rullmann, 1995) to capture its appearance on *wh*-phrases. That is, I suggest that *-râ* has the following property :

(12) ***-râ* is a type-flexible max-operator** (revised in (17)) : The Persian object marker *-râ* has the semantics of maximality across semantic types and syntactic categories.

4.2 MAXIMALITY AND EXHAUSTIVENESS IN *Wh*-WORDS

In Karttunen (1977), a question denotes the set of all its true propositions as the answer. Rullmann (1995) adds maximality to Karttunen's theory of questions, and I propose that *-râ* is an overt reflection of this maximality. In Rullmann's analysis, all questions denote singleton sets which contain exactly one proposition, the strongest true answer to the question. For example, suppose that (12), repeated below as (13), is asked in a situation in which *Tara* read *Moby Dick* and *Anna Karenina*.

(13) Târâ kodum ketâb-hâ-*(ro) xund ?
      Tara which book-PL-râ   read.3SG.S ?

---

[3] It should be noted that *-hâ* is the default plural marker on Persian nouns. However, it is not the only way to mark plurality in Persian.

'Which books did Tara read ?'

Then, the Karttunen semantics for (13) selects all the true answers of the form 'Târâ read x,' and maximality selects from these true answers the maximally informative (logically strongest) one. So if Tara read *Moby Dick* and *Anna Karenina*, the Karttunen answers will be {Tara read Moby Dick, Tara read Anna Karenina, Tara read Moby Dick and Anna Karenina}, and the maximally informative member of this set is the proposition that Tara read *Moby Dick* and *Anna Karenina*. Semantics for (13), following Rullmann, is given in (14), (see also Groenendijk & Stokhof, 1982 ; 1984) :

(14) $\iota p \exists x[[p(w) \wedge p = \lambda w[[x = \max (\lambda y[[read(w)(t,y)]])]]]]]$

Note that the maximality requirement here differs from the nominal domain in that it selects maximally informative propositions (from a set of alternatives) instead of maximal individuals ordered by a part-of relation. This is what gives rise to exhaustivity effects in question-answer pairs.

In the next section, it will be shown that an entry for *-râ* that selects maximally informative true propositions, instead of merely maximal individuals, also captures exhaustivity effects in contrastive topics in Persian, which have remained unexplained in previous approaches to the semantics of *-râ*.

4.3   EXHAUSTIVENESS IN CONTRASTIVE TOPICS

It has been observed that when *-râ* -marked phrases are contrastively focused, they have an exhaustified meaning (Karimi, 1990). However, to my knowledge, a formal derivation of exhaustivity has not been derived for these cases. The maximality analysis directly predicts the exhaustified meaning of contrastive topics. In (15), for example, I assume that focus-marking generates a set of alternatives of the form 'I danced with x' (Rooth, 1992), and that *-râ* marking generates the meaning that Ali is the maximal individual x that the speaker danced with. Another way to say this is that 'I danced with Ali' is the maximally informative true proposition in the set of focus-alternatives {I danced with x : x a salient individual} ; in (15) it is the presence of *-râ* that gives rise to the exhaustivity effect.

(15)   ALI ro bâhâsh   raghsid-am.
         ALI -râ with.him danced-1SG.S
         'ALI, I danced with him' (It was ALI I danced with (not John or others))

5   A COMPOSITIONALITY PROBLEM

The semantic entry in (12) has a clear compositional interpretation : *-râ* appears adjacent to the head noun, and picks the maximal element in the set denoted by the noun. However, when *-râ* is extended to a max operator over propositions, as in *questions* and *contrastive* topics, *-râ* appears adjacent to an embedded noun but applies maximality over the entire sentence :

(16)   **Morphosyntax** : $[_S ... [_{DP} \text{N-râ}] ...]$
         Semantics : S is the maximally informative true proposition in some set of alternatives to

S.

In other words, *-râ* appears within a sentence, but its semantics seems to require access to the meaning of the entire sentence containing it (as well as a set of alternatives to the sentence). Compositionality thus appears to break down, making it difficult to see how the meanings in questions and contrastive topics discussed above can be derived. In response to this difficulty I suggest that *-râ* itself has no meaning, but instead merely signals that it is within the scope of (is c-commanded by) a covert maximality operator, either the one over individuals or the one over propositions :

(17)   ***-râ* signals the presence of a higher max** : The presence of *-râ* attached to a noun, N-râ, signals that there is a maximality operator that c-commands this constituent.

We have identified two max operators whose presence *-râ* seems to signal : one over individuals, which I will here call MAX_{IND}, and one over propositions, which (following Fox & Hackl, 2006) I will call MAX_{INF}. Given the focus-sensitivity of both questions and contrastive topics in Persian, I will also assume that the alternative propositions needed in the entry for MAX_{INF} are the focus-alternatives of the sentence (Rooth, 1992). Here I give the lexical entries for MAX_{IND} and MAX_{INF}, and discuss below how they account for the meanings in both noun phrases and in propositions.

(18)   *Maximality over individuals* : $MAX_{IND}([[X]])$ = the maximal element in the set denoted by X, where [[X]] is of type <e,t> and is ordered by the part-of-relation (Link, 1983).

(19)   *Maximality over propositions* :
   a. $MAX_{INF}(ALT(S))(S)$ asserts that S is the maximally informative true sentence in the alternatives to S
   b.   X is the maximally informative sentence in a set Z if for all Y in Z, X entails Y
   c.   ALT(S) is the set of focus-alternatives to S, derived by replacing focus-marked constituents in S with other elements of the same semantic type (Rooth, 1992).

When *-râ* functions as MAX_{IND}, as in (1a), I propose that there is a null definite determiner $\emptyset$ with the meaning of MAX_{IND} occupying D :

(20)   a.   Morphosyntax : $[_{DP} [_D \emptyset] [_{NP}...[_N$ *ketab-râ*$]]]$
   b.   Semantics of the DP : $MAX_{IND}([[$*ketab*$]])$

When *-râ* functions as MAX_{INF}, as in (15), I argue that the presence of *-râ* signals that there is a null operator higher in the structure with the meaning of MAX_{INF} :

(21)   a.   Morphosyntax of (15) : $MAX_{INF}(ALT(15))([_S [_{DP}$ *Ali-r*$] [_{VP}$ *bâhâsh raghsid-am*$]])$
   b.   Semantics of (22a) : $MAX_{INF}(ALT(15))(15)$ = the maximally informative true proposition in ALT(15) is that I (the speaker) danced with Ali
   c.   ALT(15) = {I danced with Ali, I danced with Sara, I danced with Ehsan,É} (some contextually restricted subset of all sentences of the form 'I danced with x')

6   A NEW PREDICTION : *-râ* AND INDEFINITES

I have argued that the compositionality problem can be overcome by assuming that *-râ* signals the presence of either MAX_{IND} or MAX_{INF} higher in the structure. Here I show that this approach can solve a puzzle concerning the co-occurrence of *-râ* with the indefinite enclitic *-i*. Consider (1b) repeated here as (22) :

(22)   man ketâb-i-ro   xarid-am.
         I   book-IND-râ bought-1SG.S
         'I bought a (certain/particular) book.'

The co-occurrence of the indefinite enclitic *-i* with *-râ* has been argued to often require a relative clause to be felicitous (Karimi, 1990; Dabir-Moghaddam, 1992). For example, there is a strong intuition that (23) is infelicitous unless it is modified by a relative clause :

(23)   man ketâb-i-ro   xarid-am   ke ghermez bud.
         I   book-IND-*râ* bought-1SG.S that red   was
         'I bought a book that was red'

However, this requirement is not strict ; it has been noted that (23) is acceptable in some contexts (e.g., Ghomeshi, 2003).

I propose to explain (23) and (23) in terms of the semantics of *-râ*. Specifically, I suggest that the complementary distribution might mirror the two Max operators whose presence *-râ* signals. To see how, note that my proposal predicts that when *-râ* is in the scope of MAX_{IND}, there is a uniqueness requirement imposed on the head noun (recall our discussion in section 4.1). I suggest that when is felt to be odd unless modified by a relative clause, this happens only in those contexts in which the uniqueness requirement would not be satisfied without the relative clause, but does get satisfied with it. For example, if there are multiple books in the context, (23) cannot be used because it would fail to satisfy the uniqueness requirement imposed by MAX_{IND} ; to satisfy this, a relative clause is needed to pick out a unique book, say the one that is red, as in (23). The predicted meaning of sentences like (23) then is thus paraphrasable as "I bought book x out of the set of books {x,y,z...}," and the relative clause (e.g., the book that is red) helps us uniquely identify x. To my knowledge, neither a statement of this generalization nor or a formal account of it have previously been offered.

When a relative clause is not required to make (23) felicitous, my proposal suggests that it must be MAX_{INF} that *-râ*. is signalling (for otherwise the uniqueness requirement of MAX_{IND} would fail). Therefore, my proposal predicts that sentence (23) without a relative clause must be interpreted exhaustively, as "I only bought a book," or "a book is the only thing I bought". This is because the parse of the sentence is actually as in (24) :

(24)   a.   Predicted parse of (23) when felicitous : $MAX_{INF}(ALT(23))(23)$
         b.   Predicted meaning of (23) : (23) is the strongest true proposition in ALT(23), where
              ALT(23) is a set of propositions of the form "I bought x"

Evidence supporting this analysis comes from the observation that (25) is odd :

(25)   a.   # man ketab-i-ro xarid-am va hamchenin jakat-i ro xarid-am.
        b.   # I bought book-IND-*râ*, and I also bought jacket-IND-*râ*

Thus, the possibility of (23) and (23) both being felicitous, and the corresponding meanings under each (partitive in the case of (23), exhaustive in the case of ((23) follow from my analysis. To my knowledge, these predictions are not made by any previous account of -*râ*.

7   DISCUSSION

In this paper, it has been argued that -*râ* itself has no meaning, but signals the presence of maximality operators higher in the structure. Nevertheless, more research is needed to work out more of the formal details and also to relate this to the general typology of differential object markers.

It has recently been argued that the semantic entry for the given by Link (1983) should be modified in favour of a MAX_{INF} type entry (Kai Von Fintel and Iatridou, 2012; Schlenker, 2012). If this is true, the entry for -*râ* in (17) might be again revised to say that it signals the presence of a higher covert MAX_{INF} operator (see Fox & Hackl, 2006 for a typology of MAX_{INF} operators).

RÉFÉRENCES

Browne, W. (1970). More on definiteness markers : interrogatives in Persian. *Linguistic Inquiry*, 1(3) :359–363.

Dabir-Moghaddam, M. (1992). On the (in) dependence of syntax and pragmatics : Evidence from the postposition-ra in Persian. *Cooperating with written texts*, pages 549–573.

Elbourne, P. D. (2005). *Situations and individuals*, volume 90. MIT Press Cambridge, MA.

Ghomeshi, J. (1996). *Projection and inflection : A study of Persian phrase structure*. PhD thesis, University of Toronto.

Ghomeshi, J. (1997). Topics in Persian VPs. *Lingua*, 102(2) :133–167.

Ghomeshi, J. (2003). Plural marking, indefiniteness, and the noun phrase. *Studia Linguistica*, 57(2) :47–74.

Heim, I. (1990). E-type pronouns and donkey anaphora. *Linguistics and philosophy*, 13(2) :137–177.

Kai Von Fintel, D. F. and Iatridou, S. (2005/2012). Definiteness and maximal informativeness. To be presented at GLOW 2012.

Karimi, S. (1989). *Aspects of Persian syntax, specificity, and the theory of grammar*. PhD thesis, University of Washington.

Karimi, S. (1990). Obliqueness, specificity, and discourse functions : Ra in Persian. *Linguistic Analysis*, 20 :139–191.

Karimi, S. (1996). Case and specificity : Persian ra revisited. *Linguistic Analysis*, 26(3/4) :173–194.

Karimi, S. (2003). On object positions, specificity and scrambling in Persian. *Word order and scrambling*, pages 91–124.

Karttunen, L. (1977). Syntax and semantics of questions. *Linguistics and philosophy*, 1(1) :3–44.

Krifka, M., Pelletier, F. J., Carlson, G., Ter Meulen, A., Chierchia, G., and Link, G. (1995). Genericity : an introduction. *The generic book*, pages 1–124.

Link, G. (1983). *In Meaning, use and the interpretation of language*, chapter The logical analysis of plurals and mass terms : A lattice-theoretical approach, pages 303–323. Walter de Gruyter, Berlin, New York.

Rooth, M. (1992). A theory of focus interpretation. *Natural language semantics*, 1(1) :75–116.

Rullmann, H. (1995). *Maximality in the semantics of wh-constructions*. PhD thesis, University of Massachusetts, Amherst.

Sadeghi, A. A. (1970). Ra dar Zaban-e-Farsi-e Emruz ('Ra in the Persian language of today'). Tabriz : Tabriz University, Faculty of Literature and Humanities.

Schlenker, P. (2012). Informativity-based Maximality Conditions.

Vazinpour, N. (1977). Dastur-e Zaban-e Farsi ('the grammar of the Persian language'). Tehran : Sepahian-e Enghelab University.