Scrambling in Persian: An Optimality Theoretic Approach

Sajedeh yousefi1
Ahmad Reza Lotfi2
1, 2 Department of Foreign Languages, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran

Abstract: The present study sought to examine the optimality theoretic scrambled order of Persian syntactic structures. The article also gives a brief summary of Krimi’s (2005) work dealing with scrambling in Persian. This paper demonstrated how Optimality Theory is useful in identifying the order of structure which undergo scrambling and concluded with a discussion of what Optimality Theory can contribute to the understanding of scrambled forms of statements in Persian language. The analysis also found out the universal use on adjunction to more precise scrambling in this language.

Key words: Optimality Theory, scrambling, word order

1. Introduction

This paper aimed at highlighting on questions of optimal scrambled orders of constituents in Persian language. Diachronic aspects of Persian have recently aroused strong interest in theoretical linguistics. Among the most remarkable works is Karimi (2005), A Minimalist Approach to Scrambling, which claims that some Persian syntactic structures undergo scrambling. The present study frames Optimality-Theoretic approaches to word order change in Persian syntactic structures. A minimalist approach to scrambling and OT accounts are discussed. The main goal of this article is to develop an approach to free word order structures in Persian. Also in this paper we have laid out the foundations on which scrambling is occurred and pointed out the most important syntactic constraint to be satisfied in this respect.

Optimality Theory (OT) has originally been developed for dealing with phonological problems, abandoning the assumption that grammatical constraints are inviolable (Prince & Smolensky 1993/2004, McCarthy & Prince 1995). The fundamental new ideas have then soon been adopted in other grammatical domains, too, in syntax (Grimshaw 1997, Ackema & Neeleman 1998, Grimshaw & Samek-Lodovici 1998, Barbosa et al. 1998, Dekkers et al. 2000, Legendre et al. 2001) as well as in morphology (Legendre 2000, Wunderlich 2001).
The main innovations of OT can be summarized as follows: (i) Grammatical constraints can conflict with each other and are violable under certain conditions (not inviolable, as assumed before). (ii) Grammatical constraints are ordered according to their respective weight (not equally ranked). (iii) Different rankings of constraints are responsible for differences between individual grammars or languages. (What was formerly conceived as grammatical parameters is now reconstructed as constraints being differently ranked.) (iv) A construction is grammatical (not only by virtue of its own properties, or by virtue of its generative history, but rather) if it wins the competition in a candidate set, because it satisfies best the higher-ranked constraints.

2. Modern Persian: An overview

Persian belongs to the Iranian language family, which is a branch of Indo-Iranian, an Eastern branch of the Indo-European language family. It is spoken in Iran, Afghanistan, and Tajikistan where it is called Farsi, Dari, and Tajiki, respectively. “This group, Indo-Iranian, consists of two subgroups: Indic and Iranian. The Indic subgroup includes Sanskrit and many languages widely spoken in India and Pakistan today. Other Iranian languages still in use are Balochi spoken in West Pakistan; Pashtu or Afgan, the language of Afghanistan; Kurdish, spoken in western Iran and Turkey; and Ossetic spoken in the northern Caucasus.”

Persian is a pro-drop and scrambling language whose unmarked sentential constituent order is mainly SOV.

2.1. Word Order in Modern Persian

Some language typologists classify language syntax according to a head directionality parameter in word order, that is, whether a phrase is head-initial (= right-branching) or head-final (= left-branching), assuming that it has a fixed word order at all. Persian is more head-initial than head-final. As Greenberg (1963) classifies, Persian is a type III language. These languages are typically verb-final and postpositional. The genitive case and the modifying adjective precede the head noun, providing genitive-noun (GN) and adjective-noun (AN) orders, respectively. Although the unmarked Persian word order is SOV, all phrasal categories other than VP are head-initial. Karimi (2005) argues that Persian is not a typical type III language. She states that “written Persian exhibits a rigid SOV order, except sentential arguments of the verb systematically appear in post-verbal position. The colloquial language, however, allows a great degree of rearrangements.”

From language typology’s view point, Persian syntactic structures demonstrate a very considerable case. Ghomeshi (1996) observes that Persian flouts one of the most basic tendencies, with regard to Greenberg’s (1963) work on language universals, namely that the order of a verb and its complement and that of an ad-position and its complement are the same. Thus, Persian as a Type III (i.e. SOV) language is predicted to have postpositions. Greenberg notes that this is "overwhelmingly" true for SOV languages but mentions there exists four
exceptions, Persian is one (the others are Iraqw, a southern Cushitic language; Khamti, a Thai language; and Amharic).

Comrie (1981), also states that Persian is exceptional with respect to language typology and points out that Persian is an OV language that in almost all other respects has the properties of VO languages. Also Dryer (1988, 1992) suggests that Persian goes against strong tendencies in language typology. Dryer (1988), with reference to a broader sample of languages than that of Greenberg, points out that there is no evidence for any universal relationship between the order of an adjective and noun and the order of an object and verb. He also notes, however, that adjectives modifying nouns are often single words and suggests that the position of branching categories such as genitives and relative clauses may be more important.

Thus, the basic word order Karimi (2005) suggests, with respect to the verb and its phrasal arguments, looks like (1):

\[(1) a. (S) (O+râ) (PP) V \]
\[ b. (S) (PP) (O) V \]

“In (1), the subject precedes the internal arguments of the verb. (1a) shows a sentence with a direct object marked by \( râ \), the Case marker for specific objects. PP represents the position occupied by any subcategorized prepositional phrase. O in (1b) stands for the nonspecific object. Persian allows only one thematic object. Thus O+râ and the bare O cannot both appear in a given sentence.” [ibid] For instance,

### 2.2. Scrambled word order in Persian

#### 2.2.1. Some basic issues

The major claims Müller (1999) stipulates in his paper are: (i) scrambling is triggered by a subhierarchy of violable and ranked linearization constraints. (ii) Optimality under at least one linearization constraint results in grammaticality, optimality under the whole subhierarchy results in an unmarked structure (unmarked structures do not correspond to D-structures, as is often assumed). (iii) The distinction between subhierarchies and matrix hierarchies in optimality theory parallels the traditional distinction between weak and strong rules. It accounts for the difference between weak pronoun fronting to a Wackernagel position, which results in a fixed order, and scrambling to VP, which does not. (iv) Language-specific variation with respect to scrambling options is due to constraint reranking.

#### 2.2.2. Scrambling: Background

The term *scrambling* was coined in the 60s by John Robert Ross, who devised the first formulation of it, and stated that it was to be distinguished from the phenomenon called *free word order*. The distinction resided in the idea that scrambling involved movement. He further stated that scrambling was of *stylistic* nature (Ross 1967, Karimi 2005). Karimi (2005) states that there exist different types of elements in Persian language among which some undergo
scrambling, some are subject to limited scrambling, and elements that do not undergo scrambling at all. All phrasal arguments and adjuncts are subject to scrambling in Persian. The only exception is the nonspecific subject and object that may undergo scrambling in a limited fashion. Moreover, this language allows multiple scrambling in a clause. The following examples exhibit scrambling of arguments in the main and subordinate clauses. Depending on the stress of the scrambled element, it may receive a topic or a contrastive focus interpretation.

(1) Scrambling of the specific object over the subject
   a. ketāb-o Ali barā Sarah xānd
      book -rā A for S read
      ‘Ali read the book to Sarah.’ Or
      ‘It was the book that Ali read to Sarah.’

(2) Scrambling of the Indirect Object over the Subject
   a. be Sara hame pul mi-dah-and
to S everyone DUR-give-3pl
   ‘to Sara, everyone gives money.’

Persian allows long-distance scrambling as well. The embedded subject, specific object, and indirect object move into the matrix clause. Karimi (2005) points out that all scrambled elements can be interpreted as focus or topic, depending on the stress they carry.

(3) Long distance scrambling of the embedded subject
   Ali pro mi-dun-am ke in ketāb-ro xund-e
   A dur-know-1sg that this book -rā read -3sg
   I know that Ali has read this book.’ or “as for Ali, I know that (he) has read this book.”

(4) Long distance scrambling of the embedded specific direct object
   in ketāb-ro pro mi-dun-am ke Ali xund-e
   this book -rā dur-know-1sg that A read-3sg
   ‘As for this book, I know that Ali has read (it).’

(5) Long distance scrambling of the embedded indirect object
   be Kimea man fekr mi-kon-am ke Arezu un ketāb-ro dād-e
   to K I thought dur-do-1sg that A that book-rā gave-3sg
   ‘To Kimea I think that Arezu has given that book.’[Karimi,2005]
Furthermore, scrambling of multiple arguments is also possible, as shown in the following sentences.

(6) [be Sara] [ketāb-ā -ro] man dād – am
to S book -pl – rā I gave- 1sg
‘As for the books, TO SARA I gave (them).’ or
‘As for Sara, THE BOOKS I gave her.’

Adjuncts undergo scrambling as well:

(7) tu amuzešgāh pro fekr mi-kon-am pro Sara-ro did-an
in institute thought dur-do-1sg Sara -rā saw-2pl
‘in the institute, I think they saw Sara.’

Syntactic optionality has been named for a situation in which different ways of saying what seems to be the same thing show a clear correspondence in form. Such a situation may or may not be problematic for a given syntactic theory. Classic transformational grammar of the sixties acknowledges syntactic optionality by introducing a distinction between obligatory and optional transformations. Instances of syntactic optionality can be traced back to transformations that apply optionally. It is a well-known claim that scrambling is optional in free word order languages like Persian and German. But if we consider scrambling a universal phenomenon through which adjunction to the spec position of IP, CP is allowed in every language.

2.2.3. True Optionality

Given the definition of optimality, a candidate can be optimal without having a better constraint profile than all competitors; it suffices if there is no competitor that has a better constraint profile. Hence, if two candidates have an identical constraint profile, they can both be optimal; true optionality can arise within a single candidate set. Such an approach has been pursued by Grimshaw (1997) and Vikner (1999).

Müller (1999) counts several problems that arise with free word order structures which have not received a convincing solution so far. First, given economy constraints that block unforced movement (cf. Chomsky (1995)), scrambling cannot strictly speaking be an optional movement operation; rather, a trigger must be identified that forces scrambling. It is, however, not quite clear what this trigger might look like. Second, the issue of markedness arises: Clause-internal word order in scrambling languages often exhibits degrees of markedness, rather than complete wellformedness or illformedness, and this fact is still in need of an explanation. Third, it must be clarified why a free word order language like German does in fact exhibit a fixed order domain in the Mittelfeld, viz., the Wackernagel position, to which weak pronouns are fronted. Finally, the question arises of how language-specific variation with respect to scrambling options
Müller (1999) suggests that the constraint hierarchy must be split up into a matrix hierarchy and a subhierarchy. This modification will accommodate the scrambling evidence in German and leave previous reasonings in optimality theoretic syntax intact.

2.2.3.1. Illformedness vs. Markedness

Müller (1999) argues that scrambling structures pose a problem in economy and a markedness problem. A solution of the markedness problem that, at first sight, looks appealing is proposed by Keller (1996, 50). The basic idea is to give up the assumption that suboptimal candidates are invariably ungrammatical. More specifically, the notion of grammaticality is replaced by the notion of suboptimality.

Suboptimality:
A structure $\mathcal{S}_i$ is suboptimal with respect to a structure $\mathcal{S}_j$ if there are subsets $\mathcal{R}_i$ and $\mathcal{R}_j$ of the candidate set such that $\mathcal{S}_i$ is optimal for $\mathcal{R}_i$ and $\mathcal{S}_j$ is optimal for $\mathcal{R}_j$ and $\mathcal{R}_i \neq \mathcal{R}_j$ holds. A structure $\mathcal{S}_i$ is less grammatical than a structure $\mathcal{S}_j$ if $\mathcal{S}_i$ is suboptimal with respect to $\mathcal{S}_j$.

He reinterprets “less grammatical” as “more marked,” it follows that the optimal candidate is now viewed as the least marked; the second-best candidate is more marked than the optimal candidate, but less marked than the third-best candidate, and so forth. However, this graded approach to grammaticality faces a fundamental problem: It is a characteristic of most (if not all) syntactic analyses that have been developed within optimality theory that the “second-best” candidate is not less marked than the “third-best.” In fact, in most cases, all suboptimal candidates are decidedly ungrammatical, with no variability involved. Grimshaw (1997, 378) adopts constraints that force wh-movement (OP-SPEC), prohibit empty heads (OB-HD), and prohibit movement in general (STAY). From these assumptions it follows that (8a) is optimal, and (8b,c,d) are blocked as suboptimal (as matrix wh-questions), with (8b) emerging as the second-best candidate (OP-SPEC is satisfied, but OB-HD is not), (8c) as the third-best (OP-SPEC is violated, OB-HD and STAY are respected), and (13d) as the worst (OP-SPEC and STAY are violated).

(8 ) a. What will John read ?
   b. *What John will read ?
   c. *John will read what ?
   d. *Will John read what ?

By adopting Keller’s (1996) concept of suboptimality, we should expect that the sentences in (8b, c, d) are not completely ill formed, and that their degree of markedness should increase from top to bottom. Neither prediction seems correct, though. Thus, if the optimality theoretic
concept of grammaticality is replaced by the concept of suboptimality, this implies abandoning the main bulk of analyses that have been proposed in optimality theoretic syntax.

He suggests that this dilemma (the standard assumption that suboptimal candidates are ungrammatical, and at the same time permit markedness differences among grammatical candidates) can be solved by distinguishing two constraint levels, a matrix hierarchy and a subhierarchy. Fatal violations on the matrix hierarchy induce strict ungrammaticality. The constraints discussed up to now all belong to this hierarchy, and accordingly, suboptimal candidates in the competitions discussed above are correctly predicted to be ill formed. In contrast, fatal violations on the subhierarchy only induce markedness. The constraints that trigger scrambling belong to this latter hierarchy.

2.2.3.2. A Subhierarchy

Müller (ibid) supposes that scrambling is triggered by a Scrambling Criterion (Scr-Crit) that outranks Stay. He stipulates that Scr-Crit is in itself a subhierarchy of (potentially conflicting) linearization constraints. Among the linearization constraints are those listed below.

**Scr-Crit: in the VP domain,**

a. Nom (‘Nominative constraint’): [+nom] precedes [–nom] >
b. Def (‘Definiteness constraint’): [+def] precedes [–def] >
c. An (‘Animacy constraint’): [+animate] precedes [–animate] >
d. Foc (‘Focus constraint’): [–focus] precedes [+focus] >
e. Dat (‘Dative constraint’): [+dat] precedes [+acc] >
f. Adv (‘Adverb constraint’): [+NP] precedes [+adv] >
g. Per (‘Permutation constraint’, ‘Anti-Par-Move’): If ”c-commands # at level Ln, then “does not c-commands # at level Ln+1.

But we don’t need such classification of subhierarchy of scrambling in Persian as mentioned in previous section, some special elements undergo scrambling in this language like phrasal arguments and adjuncts, nonspecific subject and object that may undergo scrambling Persian also allows multiple scrambling in a clause. So we just consider the general case of scrambling (not the above subhierarchy of Müller) whenever we use Scr-Crit constraint.

As can be seen pronoun fronting is obviously shown in the following tableaux for the sake of focusing or topicalization.

**Tableau 1. Scrambling in Simple Interrogative**

Selected input:

(1) Tou kio didi ?
   You who(m) saw2S?
   “Who did you see?”
This tableau demonstrates how moving the object to the spec IP is optimal. If Scr-Crit out ranks STAY and OP-SPEC, s over the other candidate which incurs the violation (a) candidate Scr-Crit. This optimal candidate just violates the second constraint of this set, STAY, which grammar cares less about its violation here. As mentioned, specific object can precede the subject in Persian IP structures. Therefore the optimal structure maintains an OVS configuration. The verb stays in situ. Persian as a SOV language simply lacks verb movement. See the corresponding conflict in the following tableau.

Tableau 2. Scrambling of the specific object over the subject

Selected input: (2) a. Ali ketāb-o barā Sarah xānd
   A book-rā for S read
   “Ali read the book to Sarah.”

<table>
<thead>
<tr>
<th>Candidates</th>
<th>Scr-Crit</th>
<th>STAY</th>
<th>OP-SPEC</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. [IP Ali [VP ketāb-o barā Sarah xānd]]</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. [CP e [IP Ali [VP ketāb-o barā Sarah xānd]]]</td>
<td>*!</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>c. [IP ketāb₁-o Ali [VP ti₁ barā Sarah xānd]]</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. [IP barā Sarah₁ Ali [VP ketāb₁-o ti₁ xānd]]</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. [IP barā Sarah₁ ketāb₁-o Ali [VP ti₁ jxānd]]</td>
<td>* *!</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Two remarks on the competition in this tableau: first note that winner (c) simply adjoins specific objects to T. obviously rā-phrases seem to be generally freer to be moved out of their internal base position. Second: with respect to the unmarked basic word order, in a presence of this specific object, the indirect object is allowed to move into spec of IP. This form, candidate (d), is also
optimal but comparing to candidate (c) it seems somehow marked because, an indirect object must follow the direct object. Ghomeshi (1997) confirms that the râ-DP precedes the indirect object in the unmarked case.

**Tableau 3. Scrambling of the Indirect Object over the Subject**

Selected input: (3) a. be Ali hame mi-xand-and
to Ali everyone DUR-laugh-3pl
“At Ali, everyone laughs”

<table>
<thead>
<tr>
<th>Candidates</th>
<th>Scr-Crit</th>
<th>STAY</th>
<th>OB-HD</th>
<th>OP-SPEC</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. [IP hame [VP be Ali mi-xand-and]]</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. [CP e [IP hame [VP be Ali mi-xand-and]]]</td>
<td>*!</td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>c. [IP be Ali hame [VP t mi-xand-and]]</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Under the ranking of the above tableau a Persian declarative clause with scrambling is obtained. Since Persian doesn’t care very much about keeping the subject in the beginning of an IP structure, the indirect object scrambles over the subject. So here candidate (c) wins over the other two candidates here. They are out as they fail to satisfy the high ranked constraint of this set.

**Tableau 4. Matrix Declarative with Scrambling**

Selected input:
(4) Una lebâs xâhand xarid,
They cloth FUT-3P- buy
“They will buy cloth.”

<table>
<thead>
<tr>
<th>Candidates</th>
<th>Scr-Crit</th>
<th>STAY</th>
<th>OB-HD</th>
<th>OP-SPEC</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. [IP Una [VP lebâs xâhand xarid]]</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. [CP e [IP Una [VP lebâs xâhand xarid]]]</td>
<td>*!</td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>c. [IP lebâs, una [VP t, xâhand xarid]]</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. [CP xâhand, [IP Una [VP lebâs e, xarid]]]</td>
<td>*!</td>
<td>*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
In this tableau candidate (c) wins over the other three ones which incur the violation of Scr-Crit. Candidate (b) loses the competition as it violates the third constraints of this set. Despite the scrambled order of (d) which moves the auxiliary to the spec CP, it fails to satisfy Scr-Crit. It is certainly the case that whenever scrambling occur it makes a pragmatic contribution to the clause, either topicalization or focusing. So this not only the question of syntax but also pragmatics.

**Tableau 5. Matrix Interrogative with Scrambling**

Selected input:

(5)  Una chio mi-xâhand bexarand?

They what will-2P SUBJ-buy-3P

“What will they buy?

<table>
<thead>
<tr>
<th>Candidates</th>
<th>Scr-Crit</th>
<th>STAY</th>
<th>OB-HD</th>
<th>OP-SPEC</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. [IP una [VP chio mi-xâhand bexarand]]</td>
<td>∗!</td>
<td>∗</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. [CP e [IP una [VP chio mi-xâhand bexarand]]]</td>
<td>∗!</td>
<td>∗</td>
<td>∗</td>
<td></td>
</tr>
<tr>
<td>c. [CP Chio1 e [IP una [t1 mi-xâhand bexarand]]]</td>
<td>∗!</td>
<td>∗</td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. [CP Chio1 mi-xâhand [IP una [t1 e1 bexarand]]]</td>
<td></td>
<td></td>
<td>∗</td>
<td></td>
</tr>
<tr>
<td>e. [CP mi-xâhand1 [IP una [VP Chio e1 bexarand]]]</td>
<td>∗!</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The above tableau shows how a candidate with a scrambled word order is obtained. As can be seen candidate (a) and (b) violate the tow leftmost constraints more severely than any of their competitors. Candidate (c) is also ruled out by OB HD since the IP projection is left headless. candidate (d) here satisfies all existing constraints of this set therefore it wins the race by satisfying the three top-ranked constraints. However, it incurs the violation of the last constraint which Persian cares less about the satisfaction of that.

**Tableau 6. Multiple Interrogative with Scrambling**

Selected input:

(6)  una Chio koja mi-baran ?

they what where take3P

“What will they put where?”

<table>
<thead>
<tr>
<th>Candidates</th>
<th>Scr-Crit</th>
<th>STAY</th>
<th>OB-HD</th>
<th>OP-SPEC</th>
</tr>
</thead>
</table>
This shows how a multiple interrogative with scrambling is obtained in Persian. The evaluation in the above tableau shows that the satisfaction of the leftmost constrains SCR- Crit requires the violations of STAY. The movement of the wh-object results in a violation of STAY. This, therefore, explains the elimination of candidate (c) and (d) as they incur fatal violations. From a general standpoint, candidate (c) is not better than (d) since they both fail to satisfy Scr- Crit. Candidate(a) also is eliminated for it incurs double violations of the second leftmost constraint of this set, STAY. Candidate (b), on the other hand, comes out as optimal by satisfying the rightmost ranked constraint SCR- Crit.

3. Conclusion
This paper demonstrated how a range of syntactic structures can be triggered by the dominance of a single constraint namely Scr-Crit depending on its interaction with other constraints like STAY, OP-SPEC, OB-HD. All that the data decisively suggest is that Persian cares about the satisfaction of these constraints. This paper also indicated that there will be a tendency in an OV language like Persian for genitives and relative clauses to precede the noun. Of course, this is not true in Persian, where such constituents always follow the noun. It is tempting then, given the above discussion, to posit an underlying SVO order for Persian so that it will then fit nicely into the predicted patterns. Unfortunately, there is no evidence whatsoever for this underlying order. On the contrary, positive evidence can be found for SOV order (Karimi, 1989, 1994; Ghomeishi, 1996). I have tried to show that optimality theory makes it possible to develop a precise and testable account of free word order in Persian.

References:

Development: probing the evolutionary model of language change. Berlin: Mouton de Gruyter. 75-102


