Information Structure and Direct Object Indexation in Persian

Fatemeh Bahrami
Department of Linguistics, University of Isfahan, Isfahan, Iran

Vali Rezai
Department of Linguistics, University of Isfahan, Isfahan, Iran

Abstract—Object indexation is a grammatical phenomenon in which a clitic pronoun, co-referential with the free nominal object, appends to the verb. Optionality of indexation in Persian leads us to consider the role of pragmatic factors in its occurring. The present study aims at investigating the influence of information structure on direct object indexation in Persian. To do this, the data of standard spoken Persian including 540 cases were extracted from various resources and then analyzed within the framework of Role and Reference Grammar. The high frequency of topical direct object indexation confirms the role of information structure in direct object indexation. Analyzing the few cases of focal direct object indexation indicated that all cases of both topical and focal object indexation necessarily involve identifiable referents. Hence, the basic requirement in direct object indexation in Persian is identifiability of its referent. Considering the syntactic position of direct object in Persian clauses shows strong overlap of topical and focal objects. The post-core slot only belongs to the topical direct object and focal object cannot be placed there. Therefore, in order to identify the pragmatic relations of direct object in a clause, considering the context of the discourse is highly important.

Index Terms—object indexation, information structure, topic, focus, identifiability

I. INTRODUCTION

Object indexation occurs when a co-indexed clitic is placed on the verb beside the overt accusative noun phrase in the clause (Belloro, 2007, p. 72). This grammatical phenomenon is also available in Persian. In the following transitive clauses, in addition to the obligatory subject agreement marker, a pronominal clitic which is co-referential with the object is also joined to the verb.

(1) man   **dɑ̃niɑ̃l**   o[1]   tanhɑ̃   bozorg.**eʃ**[2]   kɑrdɑm
     I   Daniel   OM   lonely   large.obj index3SG   do.PAST.1SG
     I raised Danial by myself.

(2) **behʃɑ̃d**   pesɑr.e   ʔɑ̃qɑ̃.ye   moxtɑ̃ri   ro   miʃnɑ̃s.i.ʃ
     Behshad   son.EZ   Mr.EZ   Mokhtari   OM   know.PRES.2SG.obj index3SG
     Do you now Behshad, Mr. Mokhtari's son?

The point to be mentioned in these structures is that joining these pronominal clitic is optional; but this does not mean that all accusative noun phrases can be freely indexed. Rather, it seems that discoursal factors such as definiteness, identifiability and information structure of object noun phrases are also effective in indexation. In the present study, we are to investigate the indexation of direct object in Persian using the functional approach of Role and Reference Grammar; the issue not having been adequately considered in structuralist approaches due to ignoring pragmatic aspects of language.

Considering relatively free word order of clause elements in Persian on the one hand, and optionality of direct object indexation on the other hand, the information structure theory in Role and Reference Grammar seems to provide an appropriate tool for analyzing this grammatical phenomenon in Persian. In this study, we aim at answering the question of whether the pragmatic relationship of a direct object (its focus structure) and the object's referent status in terms of identifiability are effective in indexation. The second question is whether the topical and focal objects occupy different positions in the clause. To do this, after a brief review of some studies done on this issue, the theoretical framework is introduced in section 3. Afterwards, Persian examples are investigated and analyzed according to information structure theory.

---

1 . "râ", "ro" and "o" are allomorphs of the same morpheme i.e "object marker"
2 . Object Indexation Marker can join to the verbal or preverbal component of a compound verb

II. REVIEW OF LITERATURE

The linguistic evidence illustrates that object indexation is available in Roman, Balkan–Slavic, Albanian, Iranian, Celtic and Greek among the Indo-European family, Semitic languages among the Hemic-Semitic family and Bantu branches of East Africa (Borer, 1984; Anagnostopoulou, 2003; Torrego, 1998; Belloro, 2007 & etc.). Object indexation depends on various factors in different languages including identifiability, animacy, person, number and information structure (Comrie, 1981&1989; Croft, 1988 & 2003). Some of the most relevant studies are refered to in this section.

Kallulli (2000) states that direct object indexation in Albanian and Greek is used not only for animate-referring noun phrases, but also for definite and inanimate ones. The only constraint for indexation is on placing object in focal position; in this situation, in spite of the definiteness of the noun phrase, indexation is not allowed. In Larike (one of the Indonesian languages) object indexation depends on its discoursal prominence and in case of not being prominent, only one of the elements can be used, nominal or indexed one (Laidig & Laidig, 1990). In Macedonian, definite noun phrases have obligatory indexation in direct and oblique object positions. Petkova Schick (2000) believes that in this language, the indexed structures do not interact with information structure of clauses. This phenomenon works independently from topicality of the object noun phrases and only works as a grammatical tool to focus on definiteness. Also, in Swahili, there is a close correlation between object indexation marking and its topicality (Seidl & Dimitriadis, 1997). In Chicheŵa, object indexation is dependent on topicality of the accusative noun phrase. Also, the presence of object indexation may lead to deletion of the co-referential noun phrase which is dependent on contextual information (Bresnan & Mchombo, 1987).

In studying Ruwund, from the Bantu languages branch, Woolford (2001) points out that in this language no feature affects object indexation in isolation and various combinations of features are involved, including: [+animate, +identifiable]; [+identifiable, +focal]; [+animate, +target] and [+animate, +benefactive / malefactive]. Taking into account the universal aspects, Woolford (1999) adopts a typological approach to investigate the evidence from four African languages and presents various hierarchies in object indexation that these languages follow. With regard to information structure, she considers the influence of focal > non-focal hierarchy in these languages. Following optimality theory, to justify hierarchies involved, she relates the features to some constraints that are violable when in lower ranks.

Weissenrieder (1995) regards the nature of object indexation as agreement marker; however, she points out that object indexation is optional, contrary to subject agreement, and involves factors such as topicality, animacy, definiteness and specificity. Suñer (2000) mentions that indexation of independent pronouns is obligatory in all Spanish dialects because they always refer to the presupposed elements. Gutiérrez Rexach (2000) claims that indexed presupposed elements are necessarily non-focal and focal noun phrases cannot be indexed. Nevertheless, Suñer (1988, p.420) mentions some cases in which indexation of accusative noun phrases occurs in the focal position.

In order to completely describe the effective conditions on the occurrence of object indexation in Spanish, Belloro (2007) applies a functional approach and examines discursal factors. According to her, whenever the referent of a direct object is cognitively available, indexation is available. Reversely, in the case of oblique object indexes principally appear when the referents of their objects are inactive in the discourse. Accordingly, Belloro presents a precise explanation for indexation occurence in Spanish, without facing counter examples. Rarity of the indexed indefinite object is related to lack of identifiability.

Recently, some aspects of clitic elements have been studied in Persian; however, the pragmatic aspects of indexation have not been adequately considered. Megerdoomian (2006, p.16) states that although the function of accusative pronouns is limited to spoken language; they often appear in journalistic texts too. They have the possibility of being attached to prepositions and transitive verbs. Furthermore, they can join to preverbal element or the verbal part of compound verbs. Samvelian & Tseng (2010) present a similar analysis.

Rasekh Mahand (2005 & 2010) gives the most detailed discussion with respect to the grammatical nature of indexed accusative elements. Rasekh Mahand (2009) affirms that clitic elements only co-occur with definite direct objects and does not go with non-definite ones. However, witnessing some cases of indefinite direct object indexation, we present evidence which shows that definiteness may not be considered as a necessary condition in Persian direct object indexation. An example is presented here.

(3)	xodã	**harki**	ro	biʃtɑr	dus.eʃ	dãre	biʃtɑr
	God	everybody	OM	more	friend.obj index3SG	have.PRES.3SG	more
	gereftãr.eʃ		mikone				
	pushed.obj index3SG		do.PRES.3SG				
	The more God loves one, the more he troubles him.						

We will return to this issue, in section 4-1.

III. THEORETICAL FRAMEWORK

In Role and Reference Grammar theory which focuses on the interface between syntax and semantic/pragmatic factors, information structure, as its pragmatic part, forms one of the basic components of grammar. This approach

assumes a correlation between formal structure of sentences and the linguistic and extra-linguistic contexts in which sentences are used as units of propositional information. In RRG, the approach taken to information structure builds upon Lambrecht (1994, 2000) and is based on the idea that the structure of a sentence reflects in systematic and theoretically interesting ways a speaker's assumptions about the addressee's state of mind at the time of an utterance. Lambrecht (1994, p. 340) asserts that this approach answers the basic but often-neglected question of why grammars provide so many different ways of expressing the same proposition. In the following lines the most important parts of this theory will be introduced.

A. Focus Structure of the Clause

The most important concepts related to information structure include: presupposition, assertion3, topic, and focus. Lambrecht (1994) takes topic and focus as two information states related to referential expressions which have the pragmatic-discursal role and are not defined in terms of their syntactic position in the clause. In defining the topical and focal relations, he refers to the concepts of presupposition and assertion.

Lambrecht (1994, p. 52) defines presupposition as "the set of propositions lexico-grammatically evoked in a sentence which the speaker assumes the hearer already knows or is ready to take for granted at the time the sentence is uttered". He, in fact, takes the old information of a sentence as pragmatic presupposition and the presupposed entity constitutes topical elements in a syntactic structure. More precisely, the topic of a sentence is a referential expression whose referent is in the center of interest in the conversation and communicative interaction and about which more additional information is available to the addressee. Lambrecht insists on the fact that only by being presupposition or old information an element cannot be taken as topical, rather in addition to the mentioned condition, topical elements must enjoy centrality in discourse and communicative interaction (ibid).

According to him (1994, p. 52) the pragmatic assertion is "the proposition expressed by a sentence which the hearer is expected to know or take for granted as a result of hearing the sentence uttered". In general, it can be said that a referent is considered as the topic of proposition if in a specific communicative act, the proposition is basically ABOUT that referent i.e. it contains the information related to that referent (Lambrecht, 1994: 118). In Erteschick-Shir's opinion (2007, p. 20), previous mention is an indication of topichood; however topics are not necessarily associated with previous mention. Lambrecht (1994, p. 120) points out that the discourse context has a significant role in determining the topic in languages such as English in which neither grammatical relations nor linear configuration of structures act as a reliable tool in representing topic structure. Considering this issue in analyzing Persian is highly important.

Focus is the semantic component of the pragmatically structured proposition and whereby the presupposition and the assertion DIFFER from each other. Focus is basically the unpredictable or pragmatically non-recoverable element in an utterance (ibid, pp. 206-207). Lambrecht necessarily does not take the focal element as equivalent to new information (1994, p. 211). He defends his position by presenting the following example.

    (4)   Q: Where did you go last night, to the movies or to the restaurant?
            A: We went to the RESTAURANT.

In the response to this question the denotatum of the noun phrase the restaurant is discourse-active because it was mentioned in the immediately preceding question; therefore, information is not considered as new, but it is considered as non-recoverable or unpredictable at the time of utterance.

Lambrecht (1994, 2000) takes three foci which are as follows: argument focus, predicate focus and sentence focus. In argument focus, only a single constituent like subject, object or verb can be placed in focus relation. In Role and Reference Grammar, this focus is called narrow focus. The predicate focus is related to the unmarked topic-comment sentences in which the predicate is the focus and the subject (plus any other topical elements) is in the presupposition. Lambrecht (2000, p. 615) believes that this structure is the most unmarked information structure. The topic or the subject constituent is pragmatically presupposed and the comment ABOUT it contains new information. Finally, the sentence focus refers to a situation in which both topic and comment are new information; therefore, focus extends over both the subject and the predicate (minus any topical non-subject elements).

Since the topic-comment structures are assumed to be unmarked, subject is generally the sentence topic; however, this is not always the case and sometimes other structures may play the role of topic, too. This situation is presented in the following example provided by Lambrecht (1994, p. 146) in which the object is the topic and stress falls on the last constituent of the sentence.

    (5)   Pat said they called her TWICE.

According to Lambrecht (ibid, p. 147) the fact that in topicalization, a non-subject noun phrase can be topic does not entail that the subject must lose its topic status, but sometimes a clause may have two topics; the subject as the primary topic and the pre-posed structure as the secondary one. These two points have great importance in investigating the focus structure of indexed direct objects.

---

3. Van Valin (2005: 69) points out that Lambrecht's pragmatic assertion is equivalent to the concept of "comment" in Gundel (1998).

Lambrecht distinguishes marked narrow focus from unmarked one; the distinction depends on the position of focus constituent which is various in different languages. In Persian, the unmarked position of the focus is preverbal (Rezai & Tayyeb, 2006, p. 19).

Another distinction refers to the difference between the potential focus domain and the actual focus domain proposed by Van Valin (1993). The actual focus domain is the part of the sentence in which the focal element occurs and the potential domain refers to that part of the sentence in which the focus may potentially occur. In English and Persian, the potential focus domain is the whole sentence and therefore, the narrow focus can occur in every position of the sentence; however, based on its position it can be considered marked or unmarked.

In different languages, various strategies are used to encode the information status of structures, which mainly include morpho-syntactic tools, intonation and prosodic properties. In Persian, the narrow focus structure is indicated by putting stress on the focal element; in addition, the topicalization is also available as a syntactic tool. In the predicate focus structures, the focal stress falls on the predicate and the subject is in the topic position. Therefore, its presence is not obligatory and has the potentiality of being absent. But in the sentence focus, because of the presence of subject in the focal position, it cannot be deleted. As Van Valin (2005, p. 71) states, the most common function of sentence focus is in presentational structures and in Lambrecht's words (1994, p. 180), the most common presentational clauses among languages are those types whose predicates often have strictly non-agentive argument; like have, arrive, live, be at, be and their equivalents in other languages.

Rezai & Tayyeb (2006, p. 19) present the following discussion regarding the focus structure in Persian. The configuration of sentence structure in Persian is not always free and information structure imposes some constraints on the configuration of structures. As an example, the focus costituent cannot occur in the post-verbal position4. Furthermore, contrary to the expectations, the presence of the subject noun phrase is not always optional in Persian and if placed in the focal position, its presence is obligatory. As Van Valin (1998) points out, in SOV languages the preverbal position is the unmarked focal position and Rezai & Tayyeb (2006) confirm this claim by investigating Persian evidence. Also Yar Mohammadi (2002, p. 15) states that in Persian unmarked focus structures the stress falls on the element placed in preverbal position, but the marked focal position occurs in the sentence initial position. Therefore, narrow focus elements can be placed in the initial position as well. Rezai & Tayyeb (2006, p. 19) mention that in addition to focus elements, other non-focal elements can be influenced by topicalization and occur in the sentence initial position; however in this situation they lack prosodic prominence while focal elements are necessarily stress-bearing. As Lambrecht (1994, p. 201) remarks the possibility of topicalization is available for both focal and topical constituents in many languages and the only difference lies in their being stressed or not. If for some reasons, like prominence or emphasis, non-focal non-subject noun phrase is placed in the initial position of the clause, it will not be stress-bearing. As a summary, the configuration of the transitive clause constituents, especially direct object position in different kinds of focus, is presented in table 3-1. The focus elements are in bold.

TABLE I
THE CONFIGURATION OF STRUCTURES IN DIFFERENT KINDS OF FOCUS EXTRACTED FROM REZAI & TAYYEB (2006: 17)

Focus Type		Word order of the clause
predicate	unmarked (in situ)	(subject) + **object** + **verb**
	marked (initial position)	**Object** + **verb** +(subject)
Sentence	unmarked (in situ)	**Subject** + **object** + **verb**
	marked (initial position)	**Object** + **subject** + **verb**
narrow	unmarked (in situ)	**Subject** + object + verb Subject + **object** + verb Subject + object + **verb**
	marked (initial position)	**Object** + subject + verb Object + **subject** + verb

Roberts (2005, p. 22) examines the scrambling process in Persian based on focus structure and different syntactic positions. He also asserts that since Persian, in spite of its unmarked configuration of SVO, has a free word order, in addition to the stress position, the focus structure of the clause are encoded through the syntactic strategy, too. Exploring the Persian evidence and using the ideas of Mahoutian (1997), Rezai (2003) and Ganjavi (2003), he presents the syntactic position of the clause elements, and the direct object among them, regarding the information structure.

---

4. It should be noted that this situation is mainly true for transitive sentences but in ditransitive clauses, a focal oblique object can be placed in post-verbal position.

TABLE II
DIRECT OBJECT NOUN PHRASE PLACEMENT POSSIBILITIES IN PERSIAN EXTRACTED FROM ROBERTS (2005: 28)

	in situ (preverbal)	pre oblique	PrCS	PoCS	LDP[5]
Definite topical DO	-	-	+	-	+
indefinite topical DO	-	-	+	-	-
definite focal DO	-	+	+	-	-
indefinite focal DO	+	+	-	-	-

Although Roberts presents a detailed account of the syntactic structure of the clause in the framework of Role and Reference Grammar, his analysis of the clause focus structure and the position of topical and focal elements is not sufficiently precise. This analysis does not allow the possibility of the presence of the focus and topic in some positions of the clause which is not consistent with Persian evidence. This issue, especially with respect to the direct object noun phrase, has a great significance, concerning our subject matter and we will refer to it in section 4-2.

Having introduced the clause information structure, now another part of this theory is given, which refers to the cognitive state of the referents in discourse.

B. The Activation States of Referents

Role and Reference Grammar presents the cognitive state of referents based on the classifications by Prince (1981) and Chafe (1987), and believes that the speaker, according to the hearer's familiarity with the referents, chooses the proper linguistic forms to refer to them. Chafe (1987) affirms that the term "identifiability" is applied to differentiate between the present elements and the elements which have not yet come into existence in the universe of discourse. The unidentifiable referents refer to the elements which have not yet been shaped in the audience's mind and are not referable deictically. Using linguistic expressions, the speaker creates a new mental representation that can be recoverable in the next discourse. Lambrecht (1994, p. 77) metaphorically calls this primary mental representation as a new "referential file" in the discourse register, to which further elements of information may be added in the course of the conversation and which can be reopened in the future discourse; hence, the referent becomes identifiable for the audience. In Role and Reference Grammar, there are five levels of activation which are now presented and defined based on Van Valin (2005, p. 79):

Active is a referent which is actively under consideration in the discourse by means of direct mention. Zero and pronominal forms or unstressed referential expressions are used to represent these referents.

Accessible is a referent which is not actively under consideration but readily recognized by the addressee due either to knowledge of the world or to occurrence in the immediate environment of the speech situation. These referents are in the addressee's peripheral consciousness and the addressee has background knowledge regarding them, but at the time of utterance they are not in the center of attention. Chafe (1987, p. 22) uses two expressions of textual accessibility and inferential accessibility in order to refer to different kinds of accessible referents. Lambrecht (1994, pp. 99-100) adds the third kind of accessibility called situational accessibility. A textually accessible referent is deactivated from an earlier state, but because of being placed in the participants' peripheral consciousness, it still enjoys accessibility. In inferential accessibility, by referring to one referent, all referents associated to the related cognitive schema or framework will be accessible and finally sometimes referents of a noun phrase become accessible because of its presence in the text-external world.

Inactive refers to previously mentioned referents which are not actively under consideration at the time of utterance and not assumed by the speaker to be recognized by the addressee. These referents are in the participants' long-term memory and are not paid attention to. They may also be inactive until unknown time; however, as soon as they are referred to in discourse, they will become active. These referents are usually represented in the form of a stressed noun phrase.

Brand new-anchored is a referent which was not previously mentioned, but is related to something already introduced or accessible. In other words, this referent is anchored to another noun phrase or constituent in the text and is bound to it.

Brand new-unanchored is a referent not previously mentioned and not related to anything previously mentioned. Both kinds of brand new noun phrases are represented in the form of indefinite noun phrases (e.g. a guy). But in case of a brand new anchored referent, the indefinite noun phrase accompanies a definite noun phrase (e.g. a guy I work with).

---

5. Roberts (ibid) believes the possibility of the presence of a definite topical direct object in the left detached position and to confirm his claim he refers to an example presented by Mahoutian (1997:124). Regarding this issue, it should be noted that this example is a two-clause compound sentence; while in the present study, only one-clause sentences are considered. In addition, according to Lambrecht (1994:182), the most important criterion for identifying a sentence having the left detached position and its difference from clauses having object indexation is obligation of resumptive pronoun co-referential with the noun phrase, while in the example provided by Roberts and every other sample with direct object indexation in Persian, the indexed pronoun co-referential with the object is optional. Therefore, there is no evidence indicating the presence of the direct object in the left detached position in clauses with object indexation in Persian.

All active, accessible and inactive referents are identifiable, but brand new - unanchored and anchored referents are unidentifiable. More precisely, the identifiability decreases passing from active referents towards new unanchored referents.

Lambrecht (1994, pp. 110-111) by presenting an example, considers the activation states of the referents of noun phrases in discourse, to which we will refer to clarify the discussion. The relevant referential expressions are underlined. Small capitals indicate main points of pitch prominence.

(6)   I heard something TERRIBLE last night. (Ø) remember MARK, the guy we went HIKING with (Ø), who's GAY? His LOVER just die d of AIDS.

The deictic pronouns I, we and the first Ø are active and receive their salience from the text-external world. The referents of the anaphoric pronominal expressions who, his and the second Ø enjoy active status because of the text-internal world and they refer to the noun Mark which was previously mentioned. The active state of these referents is expressed via pronominal coding and absence of prosodic prominence. The referent of the noun phrase something terrible is brand new-unanchored, while Mark and AIDS are inactive. All three mentioned noun phrases are prosodically prominent. The referent of the temporal phrase last night is situationally accessible being deictically anchored with reference to the time of utterance. Due to its deictic status, it may go unstressed. Finally, the referent of the noun phrase his lover is inferentially accessible; both because of its relationship with the cognitive framework of the word gay and the anaphoric relationship with the now active referent Mark instantiated in the possessive determiner his. However, because of being stress-bearing, it is not active in discourse. In order to supplement the discussion, Lambrecht (ibid) expands his example:

(7)   Mark is terribly UPSET.

Following Lambrecht, if the speaker after some dialogue about the person died of AIDS shifts back to the person called Mark using the utterance (7), the referent of this noun phrase will be textually accessible. This referent cannot be coded by pronouns because due to the intervening discourse it has already been deactivated; however, it is still present in the current discourse and accessible. Therefore, it is coded via the lexical noun phrase and based on the speaker's assumption about the addressee's mental ability in recovering the referent it can be stressed or unstressed

One of the most significant issues with respect to information structure is the natural interaction between focus structure, the form of coding of the noun phrases and the activation states of the referents in the discourse. Role and Reference Grammar presents a hierarchy which shows the interaction between the pragmatic role (topic and focus) and their mode of coding. According to this continuum, by distancing away from the right, the identifiability of the referent decreases.
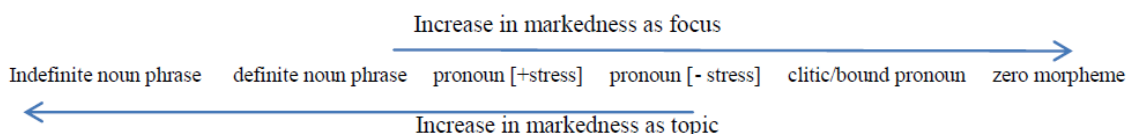
Increase in markedness as focus
→

Indefinite noun phrase   definite noun phrase   pronoun [+stress]   pronoun [- stress]   clitic/bound pronoun   zero morpheme

←
Increase in markedness as topic

Figure1. the relationship between the pragmatic role and the mode of coding extracted from Van Valin and LaPolla (1997**: 205)**

According to the figure, zero coding is the most unmarked form for topic and the indefinite noun phrase is the most unmarked choice for focus. Interestingly, the inactive referent cannot be used as a topic, but an active referent can be placed in the focus position. From another perspective, the relationship between two sides of the continuum is asymmetric, i.e. indefinite noun phrases (e.g. common nouns) can take the topic role according to the property of identifiability (see Lambrecht, 1994, pp. 167-170); however, no focus element can be coded as zero form or clitic.

Lambrecht (1994, p. 160-167) affirms that although the concepts of topic and identifiability are related to each other, it should be noted that they are totally different. The topical relation is not responsible for coding the referent activation state, but it establishes the referent of a noun phrase as the discussed issue in a clause, about which more additional information is available for the addressee. In other words, there is no obligation for all identifiable referents to be construed as topical. However, identifiability is the necessary (but not sufficient) condition for construing a referent as a topical constituent (ibid, p. 163) because some degree of identifiability of the referent is needed to consider it as a topic. It cannot be expected that more information should be available about a referent which has not been activated in discourse and has not been identifiable for the addressee yet. Accordingly, Lambrecht (1994, p. 165) presents the Topic Acceptability Scale, which indicates the relationship between the pragmatic relation of the referents and their activation level.

Active >accessible >inactive > brand new anchored > brand new unanchored
Figure2. the topic acceptability scale extracted from Lambrecht (1994, p. 165)

Having presented the theoretical framework, we will examine and analyze the Persian evidence in the following.

IV. DISCUSSION AND RESULTS

The data are analyzed in two separate sections. At first, the role of information structure in occurring indexation is examined. Afterwards, the positions of the focal and topical direct objects in the clause will be discussed.

A. *The Role of Information Structure*

As mentioned, object indexation in Persian is optional and the objective of the present study is to consider the role of the clause information structure in its occurrence. To do this, linguistic evidence including 540 samples of object indexation extracted from more than 90 TV shows, miniseries, lectures and some dialogues from standard spoken Persian were collected and scrutinized in terms of information structure. To be natural, the data were extracted from various TV programs so that a specific genre's or speaker's influence might not be imposed on the data. Furthermore, the daily live programs with diverse topics and several administrators in different days of the week were considered in data collection. In this line, it should be noted that only those conditions which make object indexation possible are of special importance in the present study and hence, even a specific genre's or individual's influence do not distort the data. However, it has been attempted to select randomly from diverse data sources. The results of investigating the data in terms of direct object indexation are presented in the following table.

TABLE III.
FOCUS STRUCTURE STATUS OF INDEXED DIRECT OBJECTS IN PERSIAN EVIDENCE

	topical direct object	focal direct object	total samples
sample	465	75	540
percentage	86.1	13.9	100

The Persian evidence illustrates that despite the high frequency of topical direct object indexation which allocates to itself about 86% of the sample, some cases of focal direct object indexation are also observed. Therefore, in this regard, Persian is similar to Chicheŵa, Swahili and Spanish, and different from Ruwund and some other African languages of Bantu branch (see Bresnan, J. and S. Mchombo, 1987; Woolford, 1999; and Suñer, 2000). Consequently, Persian confirms with the following hierarchy in direct object indexation.

Hierarchy of information structure: topical > focal

In this regard, the answer to this basic question is of great significance: how can one justify the high frequency of topical object indexation and a few cases of focus objects? Is it basically possible to explain all cases of object indexation in Persian based on a single criterion? To this end, the other part of information structure theory i.e. the identifiability of the indexed object referent, including both topic and focus, should be taken into account.

Lambrecht (1994, p. 262) points out that while the referent of a topical constituent should be identifiable and have some degree of pragmatic prominence in discourse; the focus structure is free from identifiability and activation. He adds that although new referents are generally placed in the focus position, the presence of the active referent in the focus position is also possible. Analyzing the Persian evidence in terms of the identifiability of the indexed direct object referents illustrates that not only do all topical and focal objects have identifiable referents, but they also are located in the higher levels of the continuum. It means that in the case of the presence of a single referent, deictic reference, possessive reference, anaphoric reference via pronouns and relation with a cognitive framework, the referent of object noun phrase in discourse is generally active and in some cases accessible, while not even a single case of direct object indexation with a new referent was found in the present data. In order to clarify the discussion, some instances of focal direct object indexation are presented here:

(8) ʔemruz   **pesɑr.e**   **mɑrhum**   **soleiman**   o   sɑrɑfrɑ̃z.**eʃ**   kɑrdin
     today   son.EZ   deceased   Soleiman   OM   unbowed.obj indx3SG   do.PRES.2PL
     Today, you have honored the late Soleiman's son.

(9) **xɑ̃nom**   o   d͡ʒelo.ye   bimɑ̃restɑ̃n   piɑ̃dɑ.**ʃun**   kon
     lady   OM   front.EZ   hospital   walk.obj indx3PL   do.IMP.2SG
     Would you please drop the lady in front of the hospital?

(10) ʔin   d͡ʒɑʃnvɑre   **kodum**   **bɑxʃ**   ɑz   ʔin   ʔertebɑ̃t   ro   mitune
     this   festival   which   part   from   this   relation   OM   can.PRES.3SG
     taqviɑt.**eʃ**   kone
     revival. obj indx3SG   do.PRES.3SG
     Which part of this relationship can this festival reinforce?

(11) mɑn   ʔemruz   ʔenʃɑlɑ̃   mixɑ̃m   **ʔin**   **bɑhs**   o
     I   today   if God wishes want. PRES.1SG this   discussion   OM
     bɑ̃z.**eʃ**   konɑm   ke...
     open. obj indx3SG   do.PRES.1SG   that…
     Today, if God wishes, I want to raise this discussion that…

In the above examples, the referents of indexed focal elements in discourse are totally identifiable (8-9 situationally and 10-11 textually). Therefore, all instances of direct object indexation in Persian, including topical and focal, are accounted for in terms of a single criterion. In addition, the high frequency of the indexation of topical objects can easily be justified because all topical elements are necessarily identifiable. This fact is in line with the Lambrecht's hierarchy of topical acceptability. Conversely, focal elements generally enjoy new referents and are unidentifiable in discourse. They just have high identifiability in a few cases. Therefore, it is not surprising that focal direct object indexation allocates to itself only 14% of the total data of the present study.

In section 2, it was mentioned that indefinite direct object indexation has been observed in some cases, which is unexpected according to Rasekh Mahand (2009). In this regard, considering the example (3) in terms of information structure will be helpful. To better introduce the context in which the utterance has been used, the speaker's previous sentences are presented as well.

(12) | ʔensɑ̃ | bɑ̃yad | dɑr | sɑxti.hɑ̃ | moqɑ̃vemat | kone | va | ʔetemɑ̃d.eʃ |
|---|---|---|---|---|---|---|---|
| man | should | in | difficulty.PL | resistance | do.SBJN.3SG | and | trust.POSS |
| be | xodɑ̃ | ro | ʔaz | dɑst | nɑde. | xodɑ̃ | **harki** | ro | biʃtar |
| to | God | OM | from | hand | NEG.give.SBJN.3SG | God | everybody | OM | more |
| dus.**eʃ** | | dɑ̃re | biʃtar | gereftɑ̃r.eʃ | | mikone. |
| friend.obj index3SG | | have.PRES.3SG | more | pushed.obj index3SG | | do.PRES.3SG |

A <u>man</u> should persist against calamities and shouldn't lose <u>his</u> trust in *God*. The more God loves <u>one</u>, the more *He* troubles <u>him</u>.

Lambrecht (1994, p. 95) claims that generic pronouns like English you and they and German man or words like people which have common referents, are undoubtedly active in discourse because their referents are so general that they may always be taken for granted and need not be activated. Accordingly, the referent of the word "ʔensɑ̃" (human being) at the beginning of the utterance has the property of identifiability and is active in discourse due to being general in Persian. Later on in the dialogue, one can see that the indefinite pronoun "harki" (everybody) replaces the word "ʔensɑ̃" (human being) and is still considered as a common referent for the mentioned humankind. In addition, due to its anaphoric reference, it is considered active in discourse; therefore, the necessary condition is provided for its indexation. On the other hand, in terms of information structure, it enjoys the topical state since it forms the presupposed part of proposition and is already at the core of concentration and communicative interaction. In other words, it is the element which has previously been mentioned in the context and about which some more additional information is available for the addressee. Consequently, in this sentence, there are two topical participants: the subject "xodɑ̃" (God) as primary topic and the direct object "harki" (everybody) as the secondary topic.

B. Possible Positions for Direct object

As a complementary discussion we are to review the syntactic position of both kinds of topical and focal direct objects in Persian clause. As mentioned in the introduction, due to the relatively free word order of Persian clauses, the issue which appeals to the attention is the possible syntactic positions for topical and focal direct objects in a clause.

Roberts (2005, p. 28) limits the possibility of placing topical definite and indefinite direct objects to the pre-core slot position (clause initial position) and clearly affirms that topical direct objects cannot be placed in situ (preverbal), i.e. before oblique object and in post-core slot positions (clause final position). However, our data illustrates the presence of a topical object in each of the mentioned places. Interestingly, in more than half of the cases, the topical direct object is found in the preverbal position. It should be noted that, in his classification, Roberts does not make it clear whether the topical direct object has been used in transitive or ditransitive clauses, but the examples he refers to include both kinds of transitive predicates. In the present study, it was identified that in 195 cases of transitive clauses, direct objects are placed in immediately preverbal position, which is the unmarked object position in Persian, a SOV language. In the ditransitive clauses, there are 8 cases in which direct object is placed in the preverbal position and the oblique object is in the post-verbal site. Here, just two examples of the topical direct object in the preverbal position in transitive and ditransitive clauses are referred to:

(13) | **piʧ.e** | baqal | dasti | ro | **bɑ̃z.eʃ** | kardam |
|---|---|---|---|---|---|
| screw.EZ | beside | hand | OM | open. obj indx3SG | do.PAST.1SG |

I unscrewed the adjacent screw.

(14) | ʔedʒɑ̃ze | bede | **to** | ro | befrestam.**et** | piʃ.e | ʔunɑ̃ |
|---|---|---|---|---|---|---|
| allow | give.IMP.2SG | you | OM | send.PRES.1SG. obj indx2SG | beside.EZ | them |

Let me sent you to them.

Furthermore, the presence of a topical direct object before oblique object and in post-core slot is confirmed in Persian; the positions which are not allowed according to Roberts. Here, an example is provided for each of them.

(15) | ye | deiqe | **ʔin** | o | be | man | midi.ʃ? |
|---|---|---|---|---|---|---|
| one | minute | this | OM | to | I | give.INT. 1SG. obj indx3SG |

Can you give it to me for just a minute?

(16) tʃeqɑd   mɑn   dus.**eʃ**   dɑ̃rɑm   **ʔin**   **ʔɑ̃dɑm**   o
      how much   I   friend.obj index3SG   have.PRES.1SG   this   man   OM
      How much I love this guy!

In addition, topical elements can be placed after the oblique object and also before and after the adjunct, about which Roberts has nothing to say. Now, one example is referred to for each of the mentioned issues respectively:

(17) bɑ̃   ye   tʃɑkoʃ   **nowdʒɑvun**   o   tɑhqir.**eʃ**   konim
      with   one   hammer   teenager   OM   disdain. obj index3SG   do.IMP.1PL
      Do not disdain a teenager with a hammer!

(18) ʔɑgr   **ʔun**   **nowdʒɑvun**   o   dɑr   hozur.e   dustɑ̃ʃ   baqɑl.**eʃ**
      if   that   teenager   OM   in   presence.EZ   friends.POSS   embrace.obj index3SG
      konin   vɑ̃koneʃ   neʃun   mide
      do.SBJN.2PL   reaction   show   give.PRES.3SG
      If you embrace that teenager in the presence of his friends, he will react.

(19) mɑ̃   bɑ̃yad   tu   hamin   hafte   **ʔaxtar**   o   ʃowhar.**eʃ**   bedim
      we   must   in   this   week   Akhtar   OM   husband.obj index3SG   give.SBJN.1PL
      We should marry Akhtar this very week.

It can be concluded that there is no limitation regarding the position of a topical direct object in Persian. As Roberts indicates, only in case that the topical direct object is indefinite, it cannot be placed before oblique object (dative). In this regard, it should be noted that absence of an indefinite direct object in this position is due to the semantic considerations of the verb itself not its topicality. In fact, semantically, the collocation of indefinite direct object and oblique object (dative) in a clause is impossible because basically an indefinite element lacking referential property cannot be related to an oblique object and be in some relation with it. For example, in case of the ditransitive verb "gereftan" (get) someone (subject) necessarily gets something (direct object) from someone (oblique object) whose identity is known to both interlocutors or only at least to the speaker. However, in case that by oblique object we mean every kind of objects except direct ones, it should be noted that there is no limitation for the presence of an indefinite topical direct object after or even before some oblique objects e.g instrument. This is the issue which should have been considered in Robert's analysis. The following examples confirm this claim:

(20) mɑn   gɑ̃hi   **ʃokolɑ̃t**   ro   bɑ̃   tʃɑ̃qu   teketek.**aʃ**   mikonɑm
      I   sometimes   chocolate   OM   with   knife   slice.obj index3SG   do.PRES.1SG
      I sometimes slice the chocolate with a knife.

(21) mɑn   gɑ̃hi   bɑ̃   tʃɑ̃qu   **ʃokolɑ̃t**   ro   teketek.**aʃ**   mikonɑm
      I   sometimes   with   knife   chocolate   OM   slice.obj index3SG   do.PRES.1SG
      I sometimes slice the chocolate with a knife.

Regarding the focal direct object, Roberts takes the admissible position as preverbal, before oblique object and pre-core slot positions. As Rezai & Tayyeb (2006) and Rezai (2003) assert, a focal direct object cannot be placed in the post-verbal position in Persian and because the post-core slot is necessarily placed in the post-verbal position, no example of the presence of a focal direct object in the post-core slot is observed. Below, there are some examples of the presence of the focal object in the mentioned positions respectively:

(22) **batʃe.hɑ̃**   ro   hamɑ̃me.**ʃun**   mikardɑm
      children   OM   bathroom.obj index3PL   do.PAST.1SG
      I gave the kids a bath.

(23) mɑn   **ʔin**   hadiye   ro   barɑ̃   to   xaridɑm.**eʃ**
      I   this   present   OM   for   you   buy.PAST.1SG. obj index3SG
      I have bought this present for you.

(24) **hamey.e**   **barge.hɑ̃.ye**   **ʔamal**   o   ʔun   ʔemzɑ̃.**ʃun**   karde
      all.EZ   sheet.PL.EZ   operation   OM   that   sign. obj index3PL   do.PAST.3SG
      He has signed all the operation sheets.

In the present linguistic evidence, the focal direct object is also observed in the positions after and before the adjunct.

(25) ʔenqɑd   **mɑn**   o   ʔazɑ̃b.**am**   nade
      so much   I   OM   torment.obj index1SG   NEG.give.IMP.2SG
      Don't annoy me please.

(26) ʔinke   **mɑhrox**   o   nɑ̃xɑ̃ste   ʔaziat.**eʃ**   konam
      that   Mahrokh   OM   inadvertently bother.obj index3SG   do.SBJN.1SG
      …that I inadvertently bother Mahrokh.

Therefore, the only difference in the positions of focal and topical objects is the impossibility of focal direct objects occurring in the post-core slot.

The last point to be regarded is that 64 cases of focal objects are positioned in predicate focus which is not unexpected due to the unmarkedness of topic- comment information structure. Only 2 cases are allocated to the narrow focus in which the direct object is in the form of question word and in the remaining 9 cases, the focus domain extends over the whole clause. The following examples indicate the narrow, predicate and sentence focuses respectively.

(27)   **kodum**   **ketɑ̃b.ɑm**   o   ʔandɑ̃xti.ʃ
      which   book.POSS   OM   throw out.PAST.2SG. obj index3SG
      Which book of mine did you throw out?

(28)   **dɑ̃dɑ̃ʃ.e**   **doqolu.m**   o   motɑ̃d.eʃ   kardɑm
      brother.EZ   Twin.POSS   OM   addicted.obj index3SG   do.PAST.1SG
      I addicted my twin brother.

(29)   ʔun   do   nafar   **mɑ̃**   ro   loxt.e.**mun**   kardan
      that   two   people   we   OM   naked.EZ.obj index1PL   do.PAST.3PL
      Nothing happened; those two guys (bandits) stripped us.

V. Conclusion

The indexation of topical direct objects occurs in such a high percentage in Persian that there is no doubt regarding the possibility of their indexation and it seems unlikely that there is any limitation for indexation of topical objects. But, in case of focal direct objects, there is no sufficient evidence about the possibility of indexation in all cases. The common property of both kinds of focal and topical indexed direct objects is their referent identifiability. Therefore, according to the second part of information structure theory, it can be concluded that what firstly plays a role as the basic condition in direct object indexation is the pragmatic status of both definite and indefinite object referent which should necessarily have some degree of identifiability and be preferably active in the discourse. Based on the topic acceptability scale which refers to the correlation between the topical element and identifiable referent, it seems unarguably logical that topical direct object indexation overtakes that of focal object because topical elements mostly enjoy high identifiability and are always active in the discourse in our data.

The comparison of the position of the topical and focal objects reveals that there is no limitation with respect to the syntactic position of focal and topical direct objects in Persian and it is possible to place them in any position of the clause. The only limitation is the impossibility of placing focal direct objects in the post-verbal and post-core positions. The syntactic tool of topicalization (positioning a constituent in pre-core slot) which creates a marked focus structure is used for both kinds of focal and topical objects and only the presence of stress on the pre-posed structure indicates its being focal. Therefore, as Lambrecht (1994, p. 120) claims, when neither grammatical relations nor the linear order of constituents do act as a reliable tool in recognizing a topical structure, the discourse context has a basic role in determining topics.

Abbreviations

DO	direct object
EZ	ezâfe
IMP	imperative mood
INT	interrogative mood
LDP	left detached position
NEG	negation
obj.index	object indexation marker
OM	object marker
PAST	past tense
PL	plural
PoCS	post core slot
POSS	possessive pronoun
PrCS	pre core slot
PRES	present tense
SG	singular
SBJN	subjunctive mood

References

[1]   Anagnostopoulou, E. (2003). The Syntax of Ditransitives: Evidence from clitics Berlin/ New York: Mouton de Gruyter.
[2]   Belloro, V. 2007. Spanish Clitic Doubling: A Stydy of the Syntax-Pragmatics Interface. PhD Dissertation, University of New York at Buffalo.
[3]   Borer, H. (1984). Parametric Syntax: Case Studies in Semitic and Romance Languages. Foris, Dordrecht.

[4] Bresnan, J. & S. Mchombo. (1987). Topic, pronoun, and agreement in Chichewa. Language 63: 741-782.
[5] Chafe, W. (1987). Cognitive Constraints on Information Flow. In R. Tomlin (eds.), *Coherence and Grounding in Discourse.* Amsterdam: John Benjamins, 21-51.
[6] Comrie, B. (1981). Language Universals and Linguistic Typology. Oxford: Basil Blackwell.
[7] Comrie, B. (1989).Language Typology and Linguistic Universals. Syntax and morphology, Oxford: Basil Blackwell.
[8] Croft,W. (1988). Agreement vs Case Marking and Direct Objects. In M. Barlow & C. Ferguson (eds.), *Agreement in Natural Language: Approaches, Theories, Descriptions*, CSLI: Stanford, 159-179.
[9] Croft, W. (2003). Typology and Universal (2nd edn.). Cambridge: University of Manchester.
[10] Erteschick-Shir, N. (2007). Information Structure: The Syntax-Discourse Intrface, New York: Oxford University Press.
[11] Ganjavi, S. (2007). Direct Objects in Persian, PhD dissertation, University of Southern California.
[12] Gutiérrez Rexach, J. (2000). The Formal Semantics of Clitic Doubling. *Journal of Semantics* 16: 315-380.
[13] Kallulli, D. (2000). Direct Object Clitic Doubling in Albanian and Greek. In F. Beukema & M. den Dikken (eds.). *Clitic Phenomena in European Languages*, Amsterdam: John Benjamins, 209-248.
[14] Laidig, W. & C. Laidig. (1990). Larike pronouns: duals and trials in a Central Moluccan languages. *Oceanic Linguistics* 29, 87-109.
[15] Lambrecht, K. (1994). Information Structure and Sentence Form. Cambridge: Cambridge University press.
[16] Lambrecht, K. (2000). When subjects behave like objects: a markedness analysis of sentence focus constructions across languages. *Studies in Language* 24, 611-82.
[17] Mahootian, S. (1997). *Persian,* New York: Routledge.
[18] Megerdoomian, K. (2000). Persian Computational Morphology: A Unification-Based Approach, Memoranda in Computer and Cognitive Science, Las Cruces, New Mexico: New Mexico State University.
[19] Petkova Schick, I. (2000). Clitic Doubling Constructions in Balkan-Slavic Languages. In F. Beukema & M. den Dikken (eds.), *Clitic Phenomena in European Languages*, Amsterdam: John Benjamins, 259-292.
[20] Prince, E. (1981). Toward a Taxonomy of Given-New Information. In P. Cole (eds.), *Radical Pragmatics*. New York: Academic Press, 223-254.
[21] Rasekh Mahand, M. (2005). Object Agreement Marker in Persian, Symposium of First conference of Linguistic Society of Iran, Tehran: Liguistic Society of Iran Publication, 275-285.
[22] Rasekh Mahand, M. (2009). Definite and Indefinite in Persian, *Grammar* 5, 81-103.
[23] Rasekh Mahand, M. (2010). Clitics on Verbs in Persian*, Linguistic Research* 2, 75-85.
[24] Rezai, V. (2003).A Role and Reference Grammar Analysis of Simple Sentences in Modern Persian (Farsi). PhD dissertation. University of Isfahan.
[25] Rezai, V. & T. MohammadTaghi. (2006). Information Structure and word order of the constituents in the sentence, *Grammar* 3, 1-19.
[26] Roberts, J. (2005). Scrambling in Persian: an RRG Approach, *SIL International*, 1-42.
[27] Samvelian, P. & J. Tseng. (2010). Persian object clitics and the syntax-morphology interface, Sorbonne, Proceeding of the HPSG10 Conference.
[28] Seidl, A. & A. Dimitriadis. (1997). The Discourse Function of Object Marking in Swahili. Chicago Linguistic Society. The Main Session.
[29] Suñer, M. (2000). El traslado de objetos: reflejos semánticos en la sintaxs del español y el germánico con algún detalle sobre el griego. V Encuentro Internacional de Lingüística en el Noroeste. Z. B. A. Estrada Fernández, Isabel. Hermosillo, *Editorial Unison.* 1: 17-38.
[30] Torrego, E. (1998). The Dependencies of Objects, Cambridge & Mass: MIT Press.
[31] Van Valin, R. (1993). Advances in Role and Reference Grammar. Amsterdam & Philadelphia: John Benjamin.
[32] Van Valin, R. (1998). The acquisition of WH-questions and the mechanisms of language aquisition. In M. Tomasello (eds.), *The new psychology of language: cognitive and functional approaches to language structure*. Mahwah, NJ, LEA, 221-249.
[33] Van Valin, R. (2005). Exploring the Syntax-Semantics Interface. Cambridge: Cambridge University Press.
[34] Van Valin, R. & R. LaPolla. (1997). Syntax, Structure, meaning and function. Cambridge: Cambridge University Press.
[35] Weissenrieder, M. (1995). Indirect Object Doubling: Saying Things Twice in Spanish. *Hispania* 78 (1), 169-177.
[36] Woolford, E. (1999).Animacy Hierarchy Effects on Object Agreement. In P. Kotey (eds.), *New Dimensions in African Linguistics and Languages* (Trends in African Linguistics 3), 203-216.
[37] Woolford, E. (2001).Conditions on Object Agreement in Ruwund (Bantu). In E. Benedicto (eds.), *The Umass Volume on Indigenous Languages*, University of Massachussets Occasional Papers in Linguistics 20. Amherst, MA: GLSA.
[38] Yar Mohammadi, L. (2002). A Contrastive Analysis of Persian & English, Tehran: University of Payam e Noor.

Fatemeh Bahrami was born in Isfahan, Islamic Republic of Iran. She is a PhD student in General Linguistics at the University of Isfahan, Iran. She completed a Master's degree in General Linguistics at the University of Isfahan in 2009 and got a BA in German Translation.

She has been teaching at University of Isfahan for four years. She also published several papers on Linguistics among them "Silbenstruktur des Deutschen und des Persischen" in ISI "Zeitschrift Für Germanische Linguistik" Journal. Her main Research Interests are Syntax, Typology and Pragmatics.

Vali Rezai is assistant professor of linguistics at the University of Isfahan, Isfahan, Iran. He received his Ph.D degree in General linguistics from the University of Isfahan in 2003. His primary research interests concern syntactic theories, typology and functional grammar. He has published many articles in Iranian and international journals.