

# UC Santa Cruz

## UC Santa Cruz Electronic Theses and Dissertations

### Title

Differential object marking in spoken Persian: towards an enriched typology

### Permalink

<https://escholarship.org/uc/item/1t98p7nd>

### Author

Clair, Nathaniel William

### Publication Date

2016

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

SANTA CRUZ

**DIFFERENTIAL OBJECT MARKING IN SPOKEN PERSIAN:  
TOWARDS AN ENRICHED TYPOLOGY**

A thesis submitted in partial satisfaction  
of the degree requirements for the degree of

MASTER OF ARTS

in

LINGUISTICS

by

**Nathaniel William Clair**

June 2016

The Thesis of Nathaniel Clair  
is approved:

---

Professor Donka Farkas, Chair

---

Associate Professor Adrian Brasoveanu

---

Assistant Professor Maziar Toosarvandani

---

Tyrus Miller  
Vice Provost and Dean of Graduate Studies

Copyright © by

Nathaniel Clair

2016

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Morphosyntax of Persian Nominals</b>	<b>2</b>
2.1	Overview of Persian . . . . .	2
2.2	Definite and bare objects . . . . .	2
2.3	Indefinites in Persian . . . . .	3
2.3.1	Simple indefinites . . . . .	3
2.3.2	Enclitic Indefinites . . . . .	3
2.3.3	Complex indefinites . . . . .	4
<b>3</b>	<b>Differential Object Marking</b>	<b>5</b>
3.1	Overview . . . . .	5
3.2	DOM in Persian . . . . .	6
<b>4</b>	<b>Aissen (2003): Prominence hierarchies</b>	<b>9</b>
<b>5</b>	<b>The semantics of indefinites</b>	<b>11</b>
5.1	Specificity . . . . .	12
5.1.1	Enç (1991) . . . . .	12
5.1.2	Karimi (2003) . . . . .	13
5.1.3	Farkas (2002a) . . . . .	15
5.2	Identifiability . . . . .	17
<b>6</b>	<b>DOM in Persian</b>	<b>20</b>
<b>7</b>	<b>A previous account</b>	<b>27</b>
7.1	Jasbi 2014 . . . . .	27
<b>8</b>	<b>An enriched hierarchy of definiteness</b>	<b>31</b>
8.1	Specific indefinites in Aissen (2003) . . . . .	32
8.2	Partitive specificity . . . . .	33

8.3 Strong identifiability . . . . .	35
<b>9 An updated OT analysis</b>	<b>37</b>
<b>10 Conclusion</b>	<b>40</b>
<b>References</b>	<b>40</b>

# **Differential Object Marking in Spoken Persian: Towards an Enriched Typology<sup>1</sup>**

Nathaniel Clair

## **Abstract**

In this thesis, I provide an account of differential object marking (DOM) in spoken Persian couched in the framework of Aissen (2003). In analyzing Persian indefinite patterns and their interaction with DOM in this way, I illustrate how DOM in Persian is sensitive to distinctions in definiteness that are not included in the traditional definiteness hierarchy. I argue that these distinctions concern the referential stability of the object in both current discourse contexts and in future discourse contexts. In this regard, I propose a structural distinction between partitive and epistemic specificity, which in turn has an important cross-linguistic prediction. I propose that the property of identifiability in principle also be represented in the definiteness hierarchy in which an additional distinction between strong and weak varieties of identifiability is included. Finally, I indicate how other analyses of Persian DOM that do not make reference to prominence hierarchies, but assign object marking a particular semantic denotation are problematic in the larger picture.

---

<sup>1</sup>Many thanks to my consultants Yasaman Deljouie, Hediyyeh Hosseini, Sahar Sanavi, Masoud Jahromi Shirazi, and Ehsan for providing their judgments in this project. Many thanks as well to my advisor Donka Farkas and committee members Adrian Brasoveanu and Maziar Toosarvandani, and the members of the Winter 2016 Research Seminar, participants of the 2016 Linguistics at Santa Cruz, and the UCSC Graduate Research Seminar for their insight and feedback on this project. Any mistakes herein are my own.

# 1 Introduction

In this thesis, I analyze the phenomenon of differential object marking (DOM) in spoken Persian by means of the markedness typology formalized in Aissen (2003). I argue that the definiteness hierarchy in this framework should be enriched in certain ways to account for Persian. First, I show that DOM is sensitive to both epistemic and partitive specificity, and argue on this basis that these properties represent unique rungs on the definiteness hierarchy. Second, I argue that DOM is also sensitive to the property of identifiability in principle, as defined in Farkas (2002b). More particularly, I show that DOM in Persian contextualizes this latter property in such a way that necessitates a distinction between strong and weak identifiability, and indicates that these properties should also be included into Aissen's definiteness hierarchy. Finally, I argue that this markedness typology is an ideal framework in which to discuss DOM generally because it avoids the problem of mapping differential object markers to any distinct semantic property, and provides a structured way of understanding cross-linguistic variation in DOM systems.

In section 2, I provide an overview of Persian nominal morphosyntax which focuses on the pragmatics of Persian indefinite patterns. In section 3, I provide a general overview of DOM cross-linguistically, and then illustrate the basic patterns of DOM in Persian. This is followed by a discussion of the markedness typology presented in Aissen (2003) with focus on the definiteness hierarchy in section 4. In section 5, I highlight some properties relevant to the referential restrictions placed upon indefinites, notably the varieties of specificity observed in Farkas (2002a), and the property of identifiability in principle from Farkas (2002b). In section 6, I illustrate how these properties are tied with indefinite patterns in Persian, and argue that it is necessary to distinguish between subtypes of identifiability. In section 7, I address a theory of DOM which does not adopt this typology, and indicate how it cannot account for some additional data. In section 8, I provide my own account in which the definiteness hierarchy of Aissen (2003) is enriched in ways that account for these recalcitrant data, followed by an updated Optimality Theoretic account in the spirit of Aissen (2003) in section 9 which mirrors the enrichments from section 8. In section 10, I conclude.

## 2 Morphosyntax of Persian Nominals

### 2.1 Overview of Persian

Persian, a Western Indo-Iranian language, generally has SOV word order. The dialect of Persian examined in this project is the spoken register of metropolitan Tehran, Iran. This register differs in some important ways from the more literary register, particularly with respect to the methods of marking indefinites. Thus I have taken care here to ensure that the register analyzed in this paper comes solely from the colloquial register, as past literature on Persian often conflate these registers, when in fact they differ in some of the areas discussed in this paper.

### 2.2 Definite and bare objects

Persian lacks an overt definite article. In general, definiteness may be detected on direct objects by the accompanying object marker *-ro*<sup>2</sup>. DP objects containing the overt demonstratives *in* ‘this’ and *un* ‘that’ similarly require *-ro*. Bare objects are also common in Persian and are distinguished morphologically from definite objects by the absence of *-ro*. The examples below illustrate the distinction between a definite object (1a), a bare object (1b), and an object with an overt demonstrative (1c).

- (1) a. *vis gol-ro chid.*  
Vis flower-ro picked  
‘Vis picked the flower.’ definite DP object
- b. *vis gol chid.*  
Vis flower picked  
‘Vis picked flowers.’ bare nominal object
- c. *vis in gol-ro chid*  
Vis this flower-ro picked  
‘Vis picked this flower.’ DP object with demonstrative

It is important to note that the marker *-ro* appears also on object pronouns, proper names, and some indefinites. This will be expanded upon in section 3.

---

<sup>2</sup>On a phonological note, the object marker is an unstressed enclitic and is generally pronounced *ro* (*-râ* in the formal register) when the preceding word ends in a vowel, and *o* when it ends in a consonant. However, this is not a strict phonological rule, and exceptions are common. For consistency, it will henceforth be represented as *-ro* in interlinear glosses.



## 2.3 Indefinites in Persian

Indefiniteness is expressed in Persian by three morphological patterns. Each pattern has a distinct set of semantic characteristics that constrain their distribution. I will refer to these patterns as simple indefinites, enclitic indefinites, and complex indefinites. In the following subsections, I discuss the semantic characteristics of each pattern in turn, independently of object marking. In section 6, I provide a follow-up discussion of the interaction of DOM with these patterns.

### 2.3.1 Simple indefinites

Simple indefinites are marked solely with the indefinite article *ye*. Similar to the English indefinite *a*, this article marks an indefinite in Persian that is compatible with both wide and narrow scope with respect to quantifiers or negation. The exact scopal outcome may be contextually defined (Jasbi (2014), Toosarvandani and Nasser (2015)), but generally it is ambiguous. In (2), a wide scope reading is preferred, whereas in (3), a narrow scope reading is preferred, which is emphasized by the presence of *hattâ* ‘even’ and *ham* ‘also’:

(2) *man emruz ye kâr anjâm na-dâd-am.*  
I today a task complete NEG-gave.-1SG  
‘There is a task I did not do today.’ Jasbi (2014), p38

(3) *man (hattâ) ye ketâb-(ham) na-xarid-am.*  
I (even) a book-(also) NEG-bought-1SG  
‘I didn’t by a book/any books.’ Toosarvandani and Nasser (2015, p7)

### 2.3.2 Enclitic Indefinites

Enclitic indefinites are marked solely with the indefinite suffix *-i*. Generally, this pattern contributes no existential force (Jasbi (2014), Toosarvandani and Nasser (2015)) and therefore cannot appear in the place of a simple existential indefinite, as shown in (4).

(4) *#pesar-i un ketâb-o xarid*  
boy-IND that book-ro bought.3SG  
Intended: ‘A boy bought that book.’

This indefinite pattern is licit only when embedded under a question (5a), under negation (5b), within a conditional (5c) or a possibility modal (5d) (Jasbi (2014)).

- (5) a. *râmin ketâb-i xarid-e?*  
 Ramin book-IND bought-3SG  
 ‘Has Ramin bought any books?’
- b. *pesar-i tâ hâlâ un ketâb-o na-xarid-e.*  
 boy-IND until now that book-ro NEG-bought-3SG  
 ‘No boy has bought that book yet.’ Toosarvandani and Nasser (2015, p8)
- c. *age mariam deraxt-i be-kâr-e, motma’en bâsh-id pish-e shâm*  
 if Mariam tree-IND SUBJ-plant-3SG certain be-2PL before dinner  
*dastâ-sh-o shostoshu kon-e.*  
 hands-3SGro washing do:SUBJ-3SG  
 ‘If Mariam plants a tree, make certain she washes her hands before dinner.’
- d. *momken-e ketâb-i xund-in.*  
 possible-is book-IND read-2PL  
 ‘Maybe you read a book.’

Toosarvandani and Nasser (2015) show that these indefinites always take low scope relative to negation, as in (6).

- (6) *man ketâb-i na-xarid-am*  
 I book-I NEG-bought-1SG  
 I didn’t buy any book/books.  $\neg > \exists, * \exists > \neg$

### 2.3.3 Complex indefinites

Complex indefinites are marked with both the article *ye* and the indefinite suffix *-i*. Unlike simple indefinites, these always take high scope with negation, shown in (7).

- (7) *ye shâgerd-i be daftar-am nay-umad bâ in ke sâ’at-hâ*  
 a student-IND to office-1SG NEG-came with this that hour-PL  
*montazer-esh bud-am.*  
 expecting-3SG was-1SG  
 ‘A student never came to my office, though I waited for him for hours.’  $\exists > \neg, * \neg > \exists$

The semantic distinctions between the simple and complex patterns are subtle, and will be explained further in section 6.

### 3 Differential Object Marking

#### 3.1 Overview

Differential object marking (DOM) has long been a topic of discussion in theoretical literature, from both generative and functional traditions (Enç (1991); Hopper and Thompson (1980); Croft (1988); Bossong (1991)). Broadly, it is a phenomenon in which a language uses overt structural case marking to distinguish direct objects that have particular semantic or pragmatic properties from direct objects that lack those properties. For example, Spanish has a DOM system which is sensitive to the animacy of the object. In (8a), the human object DP *la mujer* is obligatorily marked with the differential marker *a*, while in (8b) the non-human object *la casa* rejects this marker:

- (8) a. *Juan vió \*(a) la mujer.*  
Juan saw DOM the woman  
'Juan saw the woman.'
- b. *Juan vió (\*a) la casa.*  
Juan saw DOM the house.  
'Juan saw the house.'

In Hebrew, definiteness, rather than animacy, triggers DOM. Direct objects in Hebrew that are definite, as in (9a) obligatorily take the object marker *et*, accompanying the definite article *ha-*, while this marker is impossible with indefinite objects, as in (9b).

- (9) a. *dina ohevet \*(et)-ha-sefer*  
Dina loves DOM-the-book  
'Dina loves the book.'
- b. *dina ohevet (\*et)-sefer.*  
Dina loves DOM-book  
'Dina loves a book.'

A third type of DOM is characteristic of languages such as Romanian, in which both animacy and definiteness jointly play a role in distinguishing marked and unmarked nominals. Farkas and von Heusinger (2003) illustrate how this interaction plays out in Romanian. In (10a), DOM is obligatory when the human object is either a proper name or a personal pro-

noun. In (10b), DOM is optional when the object is a human definite description. Finally, in (10c), DOM is impossible when the object non-human.

- (10) a. *Maria \*(l)-a desenat \*(pe) Matei / el.*  
 Maria \*(CL)-has drawn \*(PE) Matei / him.  
 ‘Maria drew Matei / him.’
- b. *Maria (l)-a desenat (pe) băiat-ul vecinului.*  
 Maria (CL)-has drawn (PE) boy-DEF neighbor.GEN  
 ‘Maria drew the neighbor’s son.’
- c. *Maria \*(l)-a desenat \*(pe) tren-ul verde.*  
 Maria \*(CL)-has drawn \*(PE) train-DEF green  
 ‘Maria drew the green train.’ Farkas and von Heusinger (2003, p1)

### 3.2 DOM in Persian

Like Hebrew, Persian is strongly sensitive to definiteness, and to a lesser extent animacy. As in similar discussions with respect to other languages, DOM in Persian has often been observed in theoretical literature to accompany a specificity effect (Karimi (2005); Enç (1991); Aissen (2003)) or definiteness. Definite pronouns, personal names, and definite expressions all trigger the obligatory object marker *-ro*. This is exemplified in (11), (12) and (13).

*Object pronouns:*

- (11) a. *man to-\*(ro) dust dâr-am.*  
 I you-ro friend have-1SG.PRES  
 ‘I love you.’
- b. *moqe-i shekârchi-â rubâh-ro did-an, be un hamle kard-an.*  
 when-IND hunter-PL fox-ro saw-3PL to it attack did-3PL  
 ‘When the hunters saw the fox, they attacked it.’

*Proper names:*

- (12) a. *ruzbe maryam-\*(o) moteqâ’ed kard ke taklif-e*  
 Ruzbeh Maryam-ro convince do.3SG.PAST that homework-EZ  
*xod-esh-o tamum kon-e.*  
 self-3SG-ro complete do.SUBJ-3SG.  
 ‘Ruzbeh convinced Maryam to do her own homework.’
- b. *raisjomhur be mardom farmân dâd-e ke hargez irân-\*(o)*  
 president to people order gave-PERF.3SG that never Iran-ro  
*farâmush na-kon-an.*  
 forget NEG-do-3PL

‘The president ordered the people never to forget Iran.’

*Definite descriptions:*

- (13) a. *Maryam zamin-\*(o) ziroru kard ke deraxt-\*(o) be-kâr-e.*  
Maryam earth-ro overturn do.3SG.PAST that tree-ro SUBJ-plant-3SG  
‘Maryam tilled the earth so that she could plant the tree.’
- b. *maryam in maqâla-\*(ro) nevesht.*  
Maryam this letter-ro wrote.3SG.PAST  
‘Maryam wrote this letter.’

In addition, Aissen (2003) observes that Persian also marks indefinite objects which are specific. She defines specific indefinites as objects that are equivalent to *a certain* in English, indicating that the speaker has a particular entity in mind for the object, as well as indefinites that are partitive. This is very much in line with the definition of specificity present in Enç (1991) with respect to DOM in Turkish and to Karimi (2003) for Persian, although these accounts have some important differences which will be discussed later. These effects are shown for Persian in (14).

*Specific indefinites:*

- (14) a. *mariam ye ketâb-\*(ro) xarid.*  
Mariam a book-ro bought  
‘Mariam read a (certain) book.’
- b. *mariam yeki az un ketâb-â-\*(ro) xarid.*  
Mariam one of those book-PL-ro bought.  
‘Mariam read one of those books.’

A brief note is required here concerning the effect of animacy on DOM in Persian. Some analyses of Persian argue that animate objects are more likely to trigger DOM than inanimate ones, such as is shown in (15):

*Non-specific indefinites:*

- (15) a. *mariam ye xodkâr-\*(ro) mi-xâ-d.*  
Mariam a pen-ro PROG-wants.3SG  
‘Mariam wants a pen’
- b. *mariam ye shohar-(ro) mixâd*  
Mariam a husband-ro PROG-wants.3SG  
‘Mariam wants a husband.’

This project focuses on the sensitivity of DOM in Persian to definiteness, and leaves aside the effect of animacy. It has been observed in previous literature that effects of animacy are only observable when definiteness would not otherwise trigger DOM. For this reason, an account which included animacy would be compatible with the analysis sketched here.

*Bare nominal objects:*

In contrast to other objects, Persian also allows bare objects (Jasbi (2014); Mahootian and Gebhardt (1997)), which lack both articles and DOM. I suggest that this is because they are semantically incorporated, and as such are unavailable for object marking, as argued in Modarresi (2015). There are a number of reasons to believe that they are incorporated. First, bare objects can never take wide scope with respect to negation, as in (16).

- (16) *man sib na-xord-am*  
 I apple NEG-ate-1SG  
 ‘I didn’t eat apples. /\*There is an apple I didn’t eat.’

Bare objects also generally cannot be anaphoric as an atomic entity, but prefer kind-denotations, as shown in (17).

- (17) *diruz man ketâb xund-am.*  
 yesterday I book read-1SG  
 ‘I read a book/books yesterday.’

In (18), the bare object is raised to the left-periphery of the clause, in which the kind-denoting entity can receive contrastive focus.

- (18) *KETÂB<sub>i</sub> man diruz t<sub>i</sub> xarid-am*  
 book I yesterday *t* bought-1SG  
 ‘I bought *books* yesterday (as opposed to pens or paper).’

Finally, bare objects are generally non-referential and cannot serve as the antecedent for anaphora.

- (19) *man sib xarid-am. #xeyli xoshmaza-s.*  
 I apple bought-1SG. very delicious-is  
 Intended: ‘I bought an apple. It is very tasty.’

Modarresi 2014: 25

Based on these properties, it strongly appears as though bare objects contribute no discourse referents. DOM is attested on kind-denoting objects, but this is generally limited to generic contexts, and its distribution seems to follow a different set of restrictions than those illustrated here for DP objects. A well-noted example of this is portrayed in 20, below:

- (20) *sirke shir-o mi-bur-e.*  
vinegar milk-ro PROG-curdle-3SG  
Intended: ‘Vinegar curdles milk.’ Modarresi (2015, p93)

In this project, I focus on the distribution of DOM on entity-denoting DPs and leave DOM on kind-denoting terms aside.

#### 4 Aissen (2003): Prominence hierarchies

Aissen (2003) presents a formalized typology of DOM in which the properties of animacy and definiteness are defined in distinct, but interacting prominence hierarchies. This typology explains the structural uniformity observed in DOM systems while also providing a valuable tool for understanding and predicting cross-linguistic variation. The animacy hierarchy is presented in (21a), and the definiteness hierarchy in (21b), below:

- (21) a. Human > Animate > Inanimate  
b. Personal pronoun > Proper name > Definite NP > Indefinite Specific > Non-specific NP

The hierarchical nature of these properties is central; objects with greater prominence are more likely to trigger DOM, explained in Aissen (2003) as follows:

“The higher in prominence a direct object, the more likely it is to be overtly case-marked.” (Aissen (2003, p436))

This feature of DOM accounts for the uniformity of DOM systems, as each can be explained in reference to one of these hierarchies, as in the cases of Spanish and Hebrew, or an interaction of the hierarchies, as in Romanian. The underlying similarity among these languages is that even though languages draw the distinction between marked and unmarked objects at different

places on these hierarchies, DOM is only found on a particular rung *or higher*, and is constrained by the properties that these hierarchies portray.

In the animacy hierarchy, human objects are the most marked, followed by animate objects, and finally inanimate objects. In the definiteness hierarchy, the most prominent objects are personal pronouns, followed by proper names, then definite descriptions, and specific indefinites, and finally non-specific indefinites.

The spirit of this typology is to explain how DOM functions to distinguish prominent objects from subjects; the idea being that objects with high degrees of prominence on one or both of these hierarchies are more similar to canonical subjects, and must thus be differentiated from them by means of DOM. DOM is thus a morphosyntactic reflex of an object having a degree of semantic markedness. Aissen's account is couched in the framework of OT where the preference for DOM on objects that are ranked higher directly corresponds to constraint rankings which penalize lack of DOM on such objects. The details of this framework are discussed in section 9.

One prediction of this system is that a language like Hebrew which marks definites, but not indefinites, must also obligatorily mark all objects higher than definites on the definiteness hierarchy. Thus, proper names and anaphoric pronouns are predicted to take this marker as well, shown in (22).

- (22) c. *dina ohevet \*(et)-dani*  
Dina loves et-Dani  
'Dina loves Dani.'
- d. *dina ohevet \*(ot)-o*  
Dina loves et-3MSG  
'Dina loves him.'

Much cross-linguistic variation stems from two sources: (a.) which hierarchy the DOM system is sensitive to, and (b.) where the line between marked and unmarked objects is drawn. Spanish is sensitive to animacy, and the division between marked and unmarked objects is just below Human. Hebrew is sensitive to definiteness, and marks all objects that meet the level of definite or higher. Romanian is sensitive to both animacy and definiteness. Importantly for diachronic studies of DOM is the fact that the internal structure of these hierarchies is maintained,



while often the level at which the marked-unmarked distinction is drawn may vary.

This analysis focuses on a discussion of the properties underlying the definiteness hierarchy and the distinctions made within it. Farkas and von Stechow (2003) argue that the definiteness hierarchy reflects a gradience in the extent to which the variable introduced by an object's discourse referent is referentially stable. In other words, DOM systems that are constrained by definiteness employ structural case to reflect the extent to which the object has a stable discourse reference (Farkas (2002a)).

This builds upon the assumption that referential stability is constrained by a language's determiner system, which is responsible for producing different types of specificity or identifiability readings on indefinite DP's, in addition to greater contextual cues. This association explains why *a certain* indefinites and partitive indefinites should form a natural class with definite descriptions distinct from non-specific indefinites; all place similar referential restrictions on the DP. This discussion of what properties of indefinites are relevant to DOM such that some but not others may be marked requires special attention, as the distinction between marked and unmarked objects in Persian lies in this domain.

In the next section, I provide an overview of different properties that constrain the referential stability of indefinites. This will serve as the basis for a discussion of how these properties can be analyzed as having a hierarchical connection to one another in a way that explains our observations of DOM in Persian.

## **5 The semantics of indefinites**

We have seen that Persian draws the distinction between marked and unmarked objects within the domain of indefinites. It was argued in Aissen (2003) that the relevant distinguishing property is that of specificity. One issue this section covers is the fact that specificity cannot be viewed as a single property, rather it is a cover term that includes things such as partitives, wide-scoping indefinites and indefinites with *a certain* interpretation. Here, I do not assume that these properties can be placed at the same level of referential stability; in contrast, I argue that these distinctions can be represented structurally within the definiteness hierarchy, in concert with Farkas (2002b) with respect to partitive specificity more particularly.

A second problem discussed in this section is the fact that DOM can appear on indefinites in

Persian which are non-specific in the ways defined in previous analyses. This is shown in (23) below, where the indefinite can receive DOM even though the speaker overtly denies epistemic specificity:

- (23) *mârk ye kâghaz-o râje-be in mozu nevesht ke be-xâter-esh mashhur*  
 Marc a paper-ro about this topic wrote.3SG that about-3SG famous  
*shod, vali ne-mi-dun-am kodum-e.*  
 became.3SG but NEG-PROG-know-1SG which-is  
 ‘Marc has written a paper on this topic, for which he has become famous, but I don’t know which it is.’

For this reason, I provide a discussion of the property of identifiability in principle which I argue has distinct implications for, and makes more fine-grained distinctions within the realm of referential stability.

## 5.1 Specificity

Specificity is one property that is frequently discussed in connection to DOM systems cross-linguistically, following the observation that in some languages, such as Persian and Turkish, specific but not non-specific DPs bear DOM. I briefly address three ways in which specificity has been defined with respect to DOM, Enç (1991), Karimi (2003), and Farkas (2002a), and adopt the latter as an account best equipped to discuss Persian.

### 5.1.1 Enç (1991)

Enç defines specificity in terms of the types of prospective antecedents a DP has in the discourse. Definite DPs take *strong* antecedents. This means that the DP in question is able to serve as an anaphor for a DP in the previous discourse. This naturally includes definite pronouns, proper names and definite descriptions. Specific indefinites, by contrast, have *weak* antecedents. Having a weak antecedent means that the referent of the indefinite DP has the property of *inclusion*, i.e. either its referent is contained within a (usually definite) partitive set, or its referent is linked to prior discourse by an assignment function, in the case of indefinites with *a certain* interpretation. In the case of the latter, the DPs are specific because they are assigned values by assignment functions in a way which links the DP denotations to the previous discourse. Non-specific indefinites differ from specific indefinites in having no antecedents or links to other DPs.

Enç provides an example from Turkish of an indefinite object which is specific in the sense that it pairs with *a certain* in English, and is non-partitive. This is shown in (24a). A consequence of this reading is that the indefinite with DOM can only be interpreted as having wide scope with respect to the propositional attitude verb *want*, while its counterpart in (24b), which lacks DOM, has no such restriction:

- (24) a. *Ali bir piyano-yu kiralamak istiyor.*  $\exists > \text{want}/*\text{want} > \exists$   
 Ali a piano-DOM to rent wants  
 ‘Ali want’s to rent a (certain) piano.’
- b. *Ali bir piyano kiralamak istiyor.*  $\exists > \text{want}/\text{want} > \exists$   
 Ali a piano to rent wants.  
 ‘Ali wants to rent a piano.’ Enç (1991, p4-5)

The assignment function values the indefinite in (24a) with respect to a DP which is already in the discourse (*Ali*). Thus, we may informally understand *piyano* to denote something like  $(\exists x)[\text{piano}(x) \ \& \ \text{Ali wants to rent}(x)]$ , which additionally accounts for its preference for wide-scope. In contrast, the indefinite in (25) below is specific in Enç’s analysis by virtue of having an implicitly partitive interpretation, in that the presence of DOM signals that the value for *two girls* must come from the previously introduced partitive set.

- (25) *Several children entered my room...*
- İki kız-ı tanyordum*  
 two girl-DOM I knew  
 ‘I knew two (of the) girls.’ Enç (1991, p6)

### 5.1.2 Karimi (2003)

Karimi (2003) provides a similar account of Persian in which she argues that the object marker *-ro* is a marker of specificity. In order to do this, she argues that the notion of what is considered ‘specific’ in Persian should be expanded from the definition of Enç (1991). More explicitly, she argues that specificity should be expanded to include indefinite objects that have restrictive relative clauses, which is a crucial part of the pragmatic acceptability of such marked objects. This is exemplified in (26):

- (26) a. *man emruz se tâ bacha-ro did-am ke bâham da’vâ mi-kard-an.*  
 I today three CL child-ro saw-1SG that together argue PROG-did-3PL  
 ‘Today I saw three children that were arguing with each other.’ Karimi (2003, p9)

- b. *\*man emruz se tâ bacha-ro did-am.*  
 I today three CL child-ro saw-1SG Karimi (2003, p10)  
 Intended: ‘Today I saw three children.’

She also notes that in addition to appearing on partitive expressions, marked objects in Persian additionally take wide scope with respect to negation, which is independent of them being partitive or included having *a certain* interpretation. This is shown in (27):

- (27) *Kimeâ se tâ ketâb-ro na-xund-e.*  
 Kimeâ three CL book-ro NEG-read-3SG  
 ‘Kimea has not read three (specific) books.’

The divergence between these two analyses is strongly apparent with respect to non-specific DPs. Under Karimi’s definition of specificity, non-specific DPs have the ability to establish stable discourse reference via pronominal anaphors, while Enç predicts that all non-specific DPs should be unable to do this. Thus, the bare nominal in (28a), which can only have a kind-level denotation, cannot be coreferential with the definite pronoun *un* in the following sentence. By contrast, the non-specific indefinite object in (28b) is existential, and can be referred back to with such a pronoun:

- (28) a. *Kimeâ tunest mâhi be-gir-e. \*un xeyli châgh-e.*  
 Kimea managed fish SUBJ-get-3SG it very fat-be.3SG  
 ‘Kimea managed to catch fish. It is very fat.’
- b. *Kimeâ tunest apârtêmân peydâ kon-e. un xeyli ghashang-e.*  
 Kimea managed apartment find do-3SG it very pretty-be.3SG  
 ‘Kimea managed to find an apartment. It is very pretty.’ Karimi (2003, p10)

Thus, for Karimi the category of non-specific indefinites includes both existential and non-existential (kind-denoting) objects, while Enç predicts that only the latter can be considered non-specific.

Both accounts assume that DOM is a marker of specificity, although to reach this end, they are forced to different conclusions about what that entails. Despite their distinctiveness, they do converge on the idea that partitives, wide-scoping indefinites, and objects with a fairly restricted reference are considered specific (although they discuss possible sources of such a restriction: restrictive relative clauses vs. *a certain* interpretation). Importantly, the two accounts diverge

from one another with respect to the kinds of properties non-specific indefinites may have. Non-specific objects in Enç's analysis are always non-referential, while Karimi's analysis permits non-specific objects to be referential.

The properties identified with the term specificity in these analyses overlaps somewhat with that outlined in Aissen (2003), but a typology of DOM which acts as a structured way of understanding how DOM systematically varies ideally include multiple definitions of specificity. Recall that one way Aissen's typology predicts languages will differ from one another is in where they draw the line between marked and unmarked objects. We may thus reason that the difference between Turkish and Persian lies not in the definition of specificity, but where on the definiteness DOM is triggered in kind with Aissen (2003).

Considering the additional observation that DOM is triggered on indefinites in Persian which appear very non-specific, shown in (23) above, I suggest that a more robust typological understanding of specificity and potentially other properties is required. I turn to a third account of specificity which argues that some of the properties observed above should not be included under the same label, but constitute independent semantic properties, with independent consequences for referential stability.

### **5.1.3 Farkas (2002a)**

We have observed above that the definition of specificity has shifted to match language-particular descriptions, which is detrimental to a cross-linguistic approach to specificity in connection with DOM, Farkas (2002a) identifies three distinct semantic properties that have been grouped in this way, each of which has unique referential restrictions on objects' discourse referents. She defines these properties as scopal specificity, epistemic specificity, and partitive specificity.

Scopal specificity occurs when the discourse referent for a DP is interpreted independently of other quantifiers or operators. In other words, it is the effect of a DP taking wide scope relative to those operators. Narrow scope, on the other hand, occurs when the referent of a DP is interpreted dependently with respect to the quantifier or operator. In English, this distinction is observed in example (29), below:

(29) Every student read a paper about specificity. Farkas (2002a, p239)

This sentence is ambiguous between a scopally specific reading, in which every student read the same paper, and a scopally non-specific reading in which every student read a possibly different paper. It is not the case that the paper in question come from a partitive set in either interpretation. In this paper, I will be discussing scopal (non-)specificity relative solely to negation, and leave aside scopal specificity with respect to universal or other quantifiers for future analysis.

Epistemic specificity arises when the speaker has a particular value in mind for the object's discourse referent. Farkas (2002a) notes that this is the dominant reading in the example in (30), below:

- (30) a. A painting is missing from this room.  
b. A student cheated on the Syntax 1 exam. Farkas (2002a, p239)

The most natural interpretation for these examples is one in which the speaker has a particular referent in mind for the indefinite DPs. More formally, the value for the indefinites' reference is stable relative to the speaker's knowledge in the case of epistemically specific DPs, even though the discourse context does not uniquely stabilize the DP's reference, as is the case of definite DPs.

Finally, a DP is partitively specific if the value for its discourse referent comes from a familiar or unique partitive set. Since the referent is chosen from this more restricted set, the choice of value is more restricted than in other contexts. In (31) below, the value for the book is chosen from the domain contributed by the partitive phrase, rather than the domain consisting of the set of all books:

- (31) Marion bought one of the books.

Importantly for our discussion of specificity generally is that it is possible for a DP to have one variety of specificity without either of the others. The example in (31) above is perfectly licit as a partitive specific indefinite, even if epistemic specificity is overtly denied. This is the case in (32), below:

- (32) Marion bought one of the books, but I don't know which.

In the account of Aissen (2003), partitive specificity was one property which triggered DOM. The other occurrence was with indefinites with an *a certain* interpretation. It is possible that one property expressed in this description is that of epistemic specificity. Another possibility is that of identifiability in principle, which is discussed more in the next subsection. One benefit of including epistemic specificity among the properties that trigger DOM in Persian is that it also accounts for properties that rank above indefinites in the definiteness hierarchy, namely definite objects, proper names and definite pronouns. In this account, such objects are marked by virtue of the fact that definites are often also epistemically specific.<sup>3</sup>

The definiteness hierarchy in (21b) combines this latter property with partitive specificity into the single rung *specific indefinites*, following the intuition that both present a restriction on the valuation instructions of an object's discourse referent. What we have seen from Farkas' analysis is that we have a formal way of viewing the definiteness hierarchy such that these subdivisions are represented as unique hierarchical positions.

As these properties perhaps have different connotations with respect to the referential stability of the object, then this raises the question of whether we might encounter a DOM system indicating that these properties should not be conjoined in this way. Farkas (2002b) suggests that partitive specificity constitutes a rung of its own on this hierarchy. In the following section, I argue in the same vein for a bifurcation of this rung into two distinct rungs, one for partitive specificity and one for epistemic specificity.

## 5.2 Identifiability

Farkas (2002b) identifies an additional property relevant to indefinites, which she refers to as identifiability. This property concerns whether or not the value of the discourse referent of the DP is identified in the current context, and runs in tandem with the varieties of specificity. An object is said to be identified in the current context if a unique value exists for its discourse referent in the current context. This is the case for definite DPs, proper names and definite pronouns. The definition of identifiable in principle is presented below:

"There is no requirement that anybody possess identifying knowledge or that the

---

<sup>3</sup>Superlatives are an interesting exception to this; superlatives trigger obligatory DOM in Persian, even if they are not definite in the traditional sense.

variable be identified in the immediate future of the conversation, but only that it be in principle identifiable." Farkas (2002b, p16)

Next, an entity may be unidentified in the current context, either because the speaker cannot identify it, or chooses to leave it unidentified. The indefinite *some* is used to this effect in (33).

(33) Marc wrote some paper (or other) on indefinites and now he considers himself a specialist.

Farkas (2002b, p11)

This property of identifiability is distinct from epistemic specificity which is not connected to the mind of the speaker, but to the context set. This distinction allows epistemically non-specific indefinites to still be identifiable if there is at least one world in the context set in which the value is identified, as is the case of (33) above. Importantly, an indefinite may be identifiable in principle, even if no participant possesses any knowledge that uniquely identifies the referent.

Still other indefinites may be altogether unidentifiable, in which case the referent remains unidentified throughout the context. Farkas (2002b) notes that this property is also compatible with *some* DPs, for which she provides the example in (34) from Emily Dickinson:

(34) Our lives are Swiss,/ So still, so cool,/ Till some odd afternoon,/ The Alps neglect  
their curtains,/ And we look farther on.

Farkas (2002b, p12)

In Farkas (2002b)'s analysis, a speaker uses *some* to avoid identifying the referent of a DP, i.e. to leave the referent unidentified in the current context, though it may be identifiable in principle or unidentifiable. This contrasts with the indefinite *a certain*, which requires its NP complements to be identifiable in principle by the speaker. In (35), the indefinite *a certain* indicates that the speaker can identify the value of the indefinite *high official*.

(35) I spoke to a certain high official yesterday, who assured me that everything is all right.

Farkas (2002b, p15)

At this point, it is worth noting that we can discern an identifiability hierarchy, which portrays definiteness in a different perspective. This hierarchy is illustrated below in (36):



- (36) identified in current context > identified in speaker's mind > identifiable in principle  
> unidentifiable in principle

The first rung includes DPs that are identified in the current context. This property holds of definite pronouns, proper names, and definite descriptions; in each case, the referent for the individual in question is in the common ground. The second rung includes DPs that are identified in the speaker's mind, but are not in the common ground. This is true of DPs that are epistemically specific. This is followed by DPs that are identifiable in principle, in which case neither the speaker nor her interlocutor can identify the DP's referent in the current context, though given more information, the referent may be identified in some future context. Epistemic specificity further has the property of identifiability in principle, but is a more restricted version of it. In relation to the indefinites described in Farkas (2002b), *a certain* is compatible with the rungs corresponding to DPs that are identified in the speakers mind and identifiable in principle, while *some* corresponds to both identifiable in principle and unidentifiable.

This hierarchy captures some distinctions within the domain of referential stability that the definiteness hierarchy does not. First, we can distinguish indefinite DPs that are identifiable in principle from those that are unidentifiable. Both of the categories would be subsumed under the rung of non-specific indefinites. Secondly, this hierarchy further allows us to distinguish epistemic specificity from scopal specificity. In the latter, the wide-scoping object is identifiable in principle, but the speaker need not have a particular referent in mind. Thus, while not all identifiable objects need be scopally specific (if there are no negation operators to enter into scope relationships with), an object must be at least identifiable in principle to be scopally specific with respect to negation.

The fact that identifiability in principle similarly represents a restriction in the valuation instructions of objects' discourse referents brings into question whether or not the definiteness hierarchy should be enriched to include this property. The necessity of such an enrichment would be supported if there were a DOM system which was shown to be sensitive to the distinction between identifiable and unidentifiable indefinite DP objects, as the marked and unmarked cases, respectively. In the following section, I provide evidence that Persian is such a language.

## 6 DOM in Persian

In this section, I provide an outline of the general patterns of DOM in Persian. This discussion is informed by the understanding of the varieties of Persian indefinites presented in section 2, and the properties of indefinites discussed in section 5. Through including the properties of identifiability in principle defined in Farkas (2002b) and the varieties of specificity defined in Farkas (2002a), we can use the generalizations observed below to form an account which uses language-independent semantic properties, and avoids reshaping the definition of specificity to fit language-particular needs.

Previous accounts of DOM in Persian make a wide range of claims in connection with some of the semantic properties of objects it coincides with. Generative accounts of DOM in Persian argue that it signals definiteness and topichood Ghomeshi (1997, 2003). Other accounts Karimi (1996, 2003, 2005) provide an analysis of *-ro* as a marker of specificity in which DOM is triggered in connection with the structural position of marked objects in a position outside of the VP domain. In contrast, Dabir-Moghaddam (1992) provides an analysis combining both generative and functional assumptions in which syntax and pragmatics are closely linked, and in which *-ro* is simultaneously a syntactic reflex of definiteness and a pragmatic function indicating secondary topic. A more recent account (Jasbi (2014)) argues that *-ro* contributes an existential presupposition. This latter account will be discussed further in section 7.

I have illustrated how the simple (*ye* marked) indefinite pattern in Persian is compatible with both high and low scope readings with respect to negation. In case an simple indefinite object is scopally specific with respect to negation, it triggers DOM. This is illustrated in (37). Crucially, this wide scope interpretation disappears in the absence of DOM:

- (37) a. *hanuz ye dâstân-o na-xund-am.*  
 still a story-ro NEG-read-1SG  
 ‘There is a story I still haven’t read.’  $\exists > \neg$
- b. *hanuz ye dâstân na-xund-am.*  
 still a story-ro NEG-read-1SG  
 ‘I still haven’t read a story.’  $\neg > \exists$

We observe a similar situation with this pattern allowing, but not requiring, epistemic specificity as well. This is shown in the contrast between (38a), in which it is epistemically specific,

and (38b), in which it is not:

- (38) a. *man ye kârshenâs-\*(o) mi-sh(e)nâs-am, shâyad be-tun-e komak-et*  
 I an expert-ro PROG-know-1SG maybe SUBJ-can-3SG help-2SG  
*kon-e.*  
 do-2SG  
 ‘I know an expert, maybe he could help you.’
- b. *mi-sh-e ye xune tu in mahalle peydâ sh-e?*  
 PROG-become-3SG a house in this neighborhood find become:SUBJ-3SG  
 Is it possible to find a house in this neighborhood?

Example (38a) indicates that when the intended reading is epistemically specific, DOM is obligatory. This is not surprising if the *specific indefinite* layer in Aissen’s hierarchy is associated in part with epistemic specificity.

In some situations, the object marker can be used to indicate that the object is identified as belonging to a contextually salient subset of the set denoted by the nominal, i.e. it is an implicit partitive, thus allowing partitive specificity. This is observed in Jasbi (2014) in (39):

- (39) *Context: there were three cakes in the fridge.*  
*man ye keik-o tanhâyi xord-am.*  
 I a cake-ro alone ate-1SG  
 ‘I ate a cake all alone.’ (meaning: one of the cakes in the fridge) Jasbi (2014, p34)

**Generalization 1:** Simple (*ye* marked) indefinites are compatible with readings of scopal, epistemic, and partitive specificity. In each case, DOM is triggered.

Unlike the simple indefinites, the complex (*ye + -i* marked) indefinites prefer a wide scope reading, as in (40a). Jasbi (2014) notes that in such cases, DOM is strongly preferred, as in (40b).

- (40) a. *hanuz ye dâneshtju-i taklif-(o) be man na-dâd.*  
 still a student-IND homework-3SG-ro to me NEG-gave.3SG  
 ‘A student still hasn’t given me the homework.’  $\exists > \neg$
- b. *emruz ye kâr-i-#(ro) anjâm na-dâd-am.*  
 today a task-ind complete NEG-gave-1SG  
 ‘There is a task I did not do today.’  $\exists > \neg$  Jasbi (2014, p)

While the scopally specific reading with negation is prominent in these sentences, the effect of this pattern is to deny epistemic specificity. In other words, the speaker uses the complex indefinite pattern to introduce a discourse referent, but to leave its value unidentified either because they do not know the value, or choose to leave it unidentified. In this manner, the complex pattern is similar to the English indefinite *some*. It is possible, for instance, that the speaker of (40a and b) does not know the identity of the object's discourse referent, even though they are strongly indicating that a value exists and is identifiable. This is shown in example (41) below, where the dominant reading is epistemically specific:

- (41) a. *ye doktor-(#i) in qors-â-ro be man dâd.*  
 a doctor-IND this pill-PL-ro to me gave.3SG  
 'A doctor gave me these pills.'

The complex pattern is licit under an interpretation in which the speaker was less familiar to the speaker, or the fact that the person was a doctor is unimportant or happenstance.

The complex pattern is also compatible with a referent that is unidentifiable. An example of an unidentifiable object with the complex pattern is shown in example (42), from Toosarvandani and Nasser (2015):

- (42) *yani hamun-ham in tor-i nist ke shomâ boland*  
 I.mean that.same-also this way-IND NEG.be.3SG that you standing  
*sh-in be-r-in ye dâneshgâh-i be-g-in man*  
 become.SUBJ.2PL SUBJ-go-2PL a university-IND SUBJ-say-2PL I  
*mi-xâ-m in dars-ro dars be-d-am.*  
 PROG-want-1SG this class-ro class SUBJ-give-1sg  
 'I mean, it's not like you can just get up and go to **some university** and tell them that you want to teach a certain class...'

When placed in the direct object position, DOM is not allowed if its referent is unidentifiable:

- (43) *maryam bâyard ye hendevâne-i barâ-ye shâm be-xun-e, vali*  
 Maryam must a watermelon-IND for-EZ dinner SUBJ-buy-3SG but  
*pul na-dâr-e.*  
 money NEG-have-3SG  
 'Maryam has to buy a watermelon for dinner, but she doesn't have enough money.'

In this example, the indefinite *university* is not identifiable; the speaker does not have any referent in mind, nor does she indicate that there is one in particular to be identified. This shows that we cannot say that this pattern contributes either the property of identifiability in principle or epistemic specificity, though (40) shows us it is compatible with the former with a preference for DOM. Given the preference for the simple pattern in (41), I suggest that this pattern is used in contexts very similar to that of *some* in English, which permits its NP complement to be either identifiable in principle, as in (40) and (41), or unidentifiable, as in (42). I suggest that it is the nature of the complex pattern that to pair with identifiability in principle when it accompanies DOM, and prefer epistemic non-specificity. When DOM is not triggered, the complex indefinite object is unidentifiable in principle.

**Generalization 2:** DOM is not triggered on unidentifiable objects.

The main difference between the simple (*ye* marked) indefinites and complex (*ye + -i* marked) indefinites becomes apparent when DOM is present. Simple indefinites trigger DOM when they have a reading of epistemic specificity or identifiability in principle, but remain unmarked when neither of these conditions are present. However, indefinites of the complex type with DOM are only licit with identifiability in principle (explaining their persistent wide-scope), but reject epistemic specificity. I have suggested above that the reason for this lies in the semantics of the indefinites.

The sensitivity of DOM to identifiability in principle is strongly indicated by the fact that DOM is triggered on identifiable DP objects, but not on unidentifiable ones. However, this observation is problematic for an account in which DOM is constrained by the definiteness hierarchy. In other words, we have identified criteria to which DOM is sensitive which are not represented in the traditional definiteness hierarchy, i.e. the contrast between identifiable and unidentifiable objects. Without reference to these properties, the appearance of DOM on (40b) remains a mystery as both are predicted under Aissen's account to be left unmarked.

This puzzle also appears with respect to objects in the simple indefinite pattern. Similarly to (40b) in the complex pattern, such objects in the simple pattern may be epistemically non-specific, yet DOM is still present. This is exemplified in (44), below:

- (44) *mârk ye kâghaz-o râje-be in mozu neveshte ke be-xâter-esh mashhur*  
 Marc a paper-ro about this topic wrote.3SG that about-3SG famous  
*shod.*  
 became.3SG  
 ‘Marc has written a paper on this topic, for which he has become famous.’

These sentences challenge the analysis of Persian sketched in Aissen (2003). In (44), the speaker may have no value in mind for the referent of *paper*, nor is there a prominent partitive reading, yet the object marker is still possible. Similarly, the object is inanimate, and therefore is predicted to be left unmarked, contrary to fact. The object above is still identifiable in principle; neither the speaker nor the addressee may be able to identify the paper in question, but it can be identified given a more informed context state. I suggest that this reading is made more salient by the relative clause, which individuates the entity in question, without identifying it outright.

At this point, we may be tempted to link the appearance of DOM on objects to the property of identifiability in principle. A closer look indicates that the picture is yet more complex. The answer in (45) indicates that a finer distinction is in order:

- (45) Q: *diruz shomâ che kâr kard-in?*  
 yesterday you.PL what work did-2PL  
 ‘What did you do yesterday?’
- A: *man ye kâghaz nevesht-am va mariam kafsh xarid.*  
 I a paper wrote-1SG and Mariam shoes bought.3SG  
 ‘I wrote a paper and Mariam went shopping.’

The sentence in (45) contains an object which is in principle identifiable, but does not receive DOM. This presents a counterexample to the idea that identifiability in principle triggers DOM. The example in (45) differs crucially from (44) with respect to the identity of the discourse referent for the DP *paper* in the future of the discourse. In (44), there is a sense that the speaker uses DOM as a means of directing the conversation towards a context in which the object comes closer to being identified. More formally, we may say that the use of DOM on this kind of indefinite signals that the speaker is privileging a subset of the context set in which the referent of the indefinite DP is identified. This contrasts with (45) which is less felicitous if the speaker has a similar intention. Crucially, then, the distinction between an indefinite being

identifiable in principle as opposed to unidentifiable is not fine-grained enough of a distinction to fully capture speakers' intuitions about the use of DOM.

I claim that the division between marked and unmarked objects in Persian, as illustrated by the contrast in (44) and (45) can be explained with reference to a further distinction within the category of objects that are identifiable in principle. This distinction concerns the identity of the object DP's discourse referent in the future of the discourse. If an indefinite object is not defined in the current discourse, but is identifiable in principle, then the speaker may privilege a subset of the context set (a set of worlds corresponding to possible continuations of the discourse) in which the object is identifiable. In Persian, the speaker uses DOM to indicate this.

Alternatively, she may choose not to privilege the context set in this way, thus omitting DOM indicates a more stable lack of identification with respect to the object. I categorize the former property as *strong identifiability in principle*; an indefinite may be strongly identifiable in principle in case the speaker intends the object's referent to be stable in the future of the discourse. This contrasts with *weak identifiability* in which case the indefinite is still identifiable in principle, but the speaker has no additional intention of privileging the context set in the same manner.

**Generalization 3:** We can distinguish two flavors of identifiability, to which DOM is sensitive. In Persian, strong identifiability triggers DOM, weak identifiability does not.

Turning finally to the pattern of enclitic indefinites, if we consider the observation that this pattern behaves in some ways similarly to the English indefinite *any*, then we might expect it to block DOM. This expectation stems from the analysis of *any* in which the value for such nominals comes from the unrestricted domain of the NP (Kadmon and Landman (1993)); in other words, this article acts as a domain widener, rather than a domain restrictor. Recall that this pattern is compatible only with narrow scope readings with respect to negation, and is disallowed in contexts in which the speaker may have a particular referent in mind, as well as contexts which are simply existential.

This expectation, however, does not completely pan out. Jasbi (2014) observes that this pattern is compatible with an interpretation in which the indefinite DP is implicitly partitive. This reading was only found if there was DOM, as shown in (46):

- (46) *man ketâb-i-ro na-xarid-am.*  
 I book-IND NEG-bought-1SG  
 ‘I didn’t buy a book/any of the books.’

If no such reading is available, then DOM is strongly dispreferred. This returns us to the question of whether or not it remains suitable that epistemic and partitive specificity be jointly represented in the definiteness hierarchy. It provides evidence that partitive specificity is a salient property of definiteness to trigger DOM, and is independent of identifiability. As Farkas (2002a, 2002b) argue, these properties are really distinct semantic properties which effect distinct referential restrictions onto discourse referents. An additional situation in which the direct object is partitively, but not epistemically specific is shown in (46). Thus lacking other referential restrictions, partitives will always trigger DOM qua partitives. The example in (47) below supports this hypothesis with the appearance of an overt partitive whose referent identifiable but epistemically non-specific:

- (47) *keyhân yeki az aks-â-m-o forukht, vali ne-mi-dun-am*  
 keyhân one of photo-PL-1SG-ro sold-3SG, but NEG-PROG-know-1SG  
*kodum.*  
 which.  
 ‘Keyhân sold one of my photos, but I don’t know which.’

An important observation for the current analysis is that partitive specificity triggers DOM independently of identifiability in principle as well. This can be seen in the example (48) where the identity of the partitive object *one of Ali’s daughters* is overtly denied the property of identifiability in principle, yet DOM is still obligatory.

- (48) *man mi-xâ-m yeki az ketâb-â-sh-\*(o) be-xun-am, vali tasmim*  
 I PROG-want-1SG one of book-PL-3SG-ro SUBJ-read-1SG but decision  
*na-gereft-am kodum-o.*  
 NEG-get-1SG which-ro  
 ‘I want to read one of his books, but I haven’t decided which.’

**Generalization 4:** Partitive specificity triggers obligatory DOM independently of epistemic specificity or identifiability in principle.

To summarize this section, I have outlined the patterns of DOM in Persian taking into account the varieties of specificity indefinite objects may have, in addition to the property of



identifiability in principle. In doing so, I have presented two instances of DOM in Persian which are not accounted for in the definiteness hierarchy in (21b). First, we observed non-specific indefinites which seem to optionally trigger DOM on the basis of whether or not the speaker directed the future of the discourse towards a subset of the context set in which the referent of the indefinite DP was identified. This led to formalizing this distinction between strong and weak identifiability in principle. Second, we observed that partitive specificity was unilaterally able to trigger DOM on objects, both in cases where the partitive reading was implicit, and in cases where it was overt. Both of these situations present distinct puzzles for the definiteness hierarchy in its current form – it seems additional features should be included. In the next section, I outline another account of DOM in Persian which does not assume it is constrained by the definiteness hierarchy.

## **7 A previous account**

### **7.1 Jasbi 2014**

In this section, I outline the account of DOM in Persian in Jasbi (2014) which is not couched in a typology assuming that semantic markedness accounts for DOM. I argue that the data supporting this account can be reinterpreted in a way that is compatible with the account in this paper. Furthermore, I show that Jasbi's account and my account share some shortcomings, but suggest that an account in which DOM is couched in the markedness framework adapted from Aissen (2003) is more ideally equipped to resolve them.

To account for some of the more problematic cases of object marking in Persian, Jasbi argues that the Persian object marker *-ro* contributes an existential presupposition, thereby indicating that the set denoted by the predicative condition of the noun is non-empty. This theory accounts for the fact that DOM appears on indefinites which lack any of the types of specificity defined above, such as in (49):

- (49) Context: "suppose that my three-year-old cousin takes my phone and accidentally deletes a picture on my camera roll. When I open the camera app and look at the number of pictures written at the top, I notice it is one less than what it should be. In such a context I might tell my sister: "

*in sheytun ye aks-o pâk kard-e*  
 this Satan a photo-ro clean did-3SG  
 ‘This (little) rascal deleted a picture.’

Jasbi (2014, p32)

In this analysis, this example indicates the object marker in Persian cannot be said to contribute epistemic (or any other variety of) specificity. Rather, the addition of *-ro* signals that the set of objects denoted by an NP like *aks* ‘picture’ is non-empty. Being a presupposition, this has the advantage of explaining why DOM is possible on objects with low scope with respect to negation. This was seen in the example (40b) above, repeated below in (50):

(50) *man emruz kâr-i-(ro) anjâm na-dâd-am.*  
 I today task-IND-ro complete NEG-gave-1SG  
 ‘I didn’t do a task today.’

$\neg > \exists / * \exists > \neg$

According to Jasbi (2014), the presence of DOM on the object here shows that it cannot be an indication of scopal specificity but expresses the intuition that while the speaker did not do any work, there was a set of work for the speaker to complete. If DOM is absent from this sentence, then the only interpretation is *I didn’t do any task today*, in which case the referent of *task* is unidentifiable.

The examples above, however, do not rule out the possibility of either implicit partitive specificity or strong identifiability in principle. In fact, both readings are available for 49 and 50. In the first situation, it is understood that the objects in question is identified as belonging to the discursively salient subset. In (49), this is set of pictures on the camera, while in (50), this is the set of tasks the speaker has to do, not the set of all tasks in the domain. The fact that DOM is present in this situation is then unsurprising, as DOM is obligatory on all partitives.

The second possibility for these examples is that the objects have the property of strong identifiability in principle, such as in 50. Even though the speaker may currently be unable to identify which picture was deleted or which task they failed to perform, the speaker chooses to direct the discourse towards a subset of the context set in which their referents are identified given a more informed understanding, thus the existential presupposition goes hand in hand with being identifiable in principle.

Some support for the notion that *-ro* does not have a semantic contribution comes from a few cases where an existential presupposition is not present, yet DOM persists. This is the case

in (51) below where the speaker is directly questioning the existence of the object:

- (51) *ye doktor-(i)-(ro) mišnâse ke be-tun-e tu in vaz komak-ef*  
 a doctor-IND-ro know.3SG who SUBJ-can.3SG in this condition help-3SG  
*kon-e?*  
 do.SUBJ-3SG  
 ‘Does he know a doctor who can help with his condition?’

Here, the speaker is not presupposing the existence of such a doctor. This then is problematic for an account in which *-ro* contributes an existential presupposition. On the other hand, this is not surprising if the use of DOM correlated to the referential restrictions placed on the object. When we replace the relative clause with an adjective, or remove it altogether, the intuition of speakers is that both alternatives are fine, but *-ro* is dispreferred. Its optional nature in this situation shows that DOM does not always affect the truth-conditional output of the sentence, but suggests a pragmatic-based account in which contextual factors make one version more appropriate.

- (52) a. *ye doktor-(i)-(#ro) mišnâs-e?*  
 a doctor-IND-ro know-3SG  
 ‘Does he know a doctor?’ (any doctor will do)
- b. *ye doktor-e xub-(#ro) mišnâse?*  
 a doctor-ez good-ro know-3SG  
 ‘Does he know a good doctor?’

The correlation between the availability of DOM and the amount of restrictions placed on the object are compatible with an account in which strong identifiability and partitive specificity are analyzed as unique triggers of DOM in Persian. In the case of (51), the relative clause serves to individuate without actually identifying the object, and this in turn allows the speaker to direct the conversation towards a subset of the context set in which the object can become identified. This is compatible with the observations about DOM use in Persian by Karimi (2003), where DOM was strongly preferred on DP objects with restrictive relative clauses. This was exemplified in the contrast between (26a) and (26b), above, repeated below in (53).

- (53) a. *man emruz se tâ bacha-ro did-am ke bâham da’vâ mi-kard-an.*  
 I today three CL child-ro saw-1SG that together argue PROG-did-3PL  
 ‘Today I saw three children that were arguing with each other.’ Karimi (2003, p9)

- b. *\*man emruz se tâ bacha-ro did-am.*  
 I today three CL child-ro saw-1SG Karimi (2003, p10)  
 Intended: ‘Today I saw three children.’

In this situation, the speaker is a direct witness of the event depicted in the relative clause. This in effect shows that while the speaker may not have epistemic knowledge sufficient to identify the children in question, she is showing that they are identifiable. This is further supported by examples such as (54), below, in which the object rejects DOM despite having a relative clause. In this example, the articles are identifiable in principle, but the speaker is directing the context towards the content of the articles, not the articles themselves.

- (54) *man se tâ maqâle-ye diga-m khund-am ke hamin-o mi-goft-and.*  
 I three CL article-EZ other-also read-1SG that this-ro PROG-say-3PL  
 ‘I read three articles that said that very thing.’ source: [www.afkarnews.ir](http://www.afkarnews.ir)

In this sense, we cannot say that DOM is triggered on objects that head relative clauses generally. Rather, DOM is triggered on such indefinites when it is the speaker’s intention to highlight a subset of the context set in which the object is identified, thereby bringing the referent of the DP object into the common ground. Thus, the ungrammaticality of (53b) is a result of the fact that the object receives DOM without having been made identifiable by the relative clause.

Recall that one benefit of this markedness analysis is that it accounts for the distribution of DOM in Persian without needing to associate the marker to any particular semantic property. Indeed, if DOM is simply structural Case, then it has no meaning, rather its appearance is constrained by the pragmatic and semantic factors that contextualize it. More specifically, we have observed that DOM in Persian is obligatory on all objects which are epistemically specific, but as it is possible on non-specific objects it would be a faulty assumption that it contributes this property. Similarly, we cannot claim that *-ro* in Persian contributes the property of strong identifiability as it is possible for a partitive indefinite to lack this property, yet DOM is still triggered by virtue of the fact that it is partitive, and partitive ranks high on the definiteness hierarchy. An example of this is seen in (55), below:

- (55) *aslan ne-mi-tun-id yeki az un deraxt-â-\*(ro) bexar-id age balad*  
 at.all NEG-PROG-can-2PL one of that tree-PL-ro buy.2PL if know

*na-bâsh-id az-esh morâqebat kon-id.*  
NEG-be-2PL of-3SG care do-2PL

‘You can’t buy one of those trees if you don’t know how to take care of it.’

In this example, the speaker is not pointing out any particular entity, nor is she privileging any subset of the context set in which it is identified; any tree from the restricted domain will suffice. DOM is triggered by virtue of the fact that this unidentifiable object is partitive.

To summarize this section, I have outlined an alternative analysis of DOM provided in Jasbi (2014) which argues that the object marker *-ro* contributes an existential presupposition. I have argued to the contrary in providing examples in which DOM persists despite lacking such a presupposition. I have also shown that the cases in which this presupposition were apparent could be reinterpreted as either partitive specificity or strong identifiability in principle, and suggested that the cases of DOM in questions in which the existential presupposition was unavailable might still be accounted for under an analysis in which strong identifiability were included on the definiteness hierarchy, in anticipation of the following section. In sum, however, it is possible to analyze the cases of DOM in this other approach without claiming that the object marker *-ro* has a semantic contribution its own, but is merely a morphological reflex of the object meeting a certain condition or higher on the definiteness hierarchy. In the next section, I provide an account in which the facts above can be accounted for by means of an enriched hierarchy of definiteness.

## **8 An enriched hierarchy of definiteness**

In this section, I provide a formal analysis of DOM in Persian in which I show that it is possible to account for the recalcitrant cases of DOM observed above by adopting and modifying the typology of Aissen (2003) by enriching the definiteness hierarchy.

The first problem concerned the obligatory appearance of DOM on all indefinite partitives, independently of epistemic specificity or identifiability in principle. This was puzzling for the analysis of specificity in Karimi (2003) as well, who noted that DOM also appeared on non-partitive indefinite objects which headed relative clauses.

The second problem concerned the appearance of DOM on indefinite objects which lacked the properties of specificity defined in Aissen (2003), but which had the property of strong

identifiability in principle. The secondary problem with these cases was to account for the absence of DOM on indefinites which had the property of weak identifiability in principle or were altogether unidentifiable. To begin solving these problems, I unravel the properties that were included in the rung of specific indefinites in Aissen (2003).

### 8.1 Specific indefinites in Aissen (2003)

In her analysis, Aissen (2003) associates the rung of specific indefinites with a number of distinct properties which distinguish these indefinites from their unmarked counterparts. First, it is associated with indefinites with a partitive interpretation, all of which triggered DOM in Persian. Second, it was associated with indefinites with the interpretation of *a certain* in English. According to the analysis of these indefinites in Farkas (2002b), *a certain* is compatible with two interpretations. First, these indefinites must be at least identifiable in principle. One example of such a situation is presented in (56), below.

(56) The man had a certain look in his eye, which I couldn't quite put my finger on.

Identifiability in principle is the minimal qualification for such indefinites. We can further observe from Farkas's analysis that these indefinites are compatible with epistemic specificity, although they do not require it. This is because the referent for an indefinite may be identified in the speaker's mind, while remaining unidentified the shared context provided it is at least identifiable in principle. An example of an *a certain* indefinite which is epistemically specific is provided in (57a), and an indefinite which is epistemically non-specific in (57b).

- (57) a. I want to send an invitation to a certain biographer, but I don't know his address.  
b. Marc wants to plant a certain tree near the fence, but I don't know which.

In sum, the indefinites incorporated into this rung in the definiteness hierarchy in Aissen's account represent indefinites with three distinct semantic properties: partitive specificity, epistemic specificity, and identifiability in principle. In the following subsections, I indicate how these properties can be included into the hierarchy as distinct rungs, and point out some important consequences of this for DOM cross-linguistically.

## 8.2 Partitive specificity

The first puzzle for this analysis is accounting for the obligatory nature of DOM on indefinites that have one of the properties of partitive or epistemic specificity, but not the other. This was the case for example (47), repeated below in (58a). Similarly, (46), repeated below in (58b) presented a situation in which DOM was triggered on a partitive indefinite which was unidentifiable in principle.

- (58) a. *keyhân majbur-e yeki az un ketâb-â-\*(ro) be-xun-e, vali*  
 keyhân obliged-is one of those book-PL-ro SUBJ-read-3SG but  
*ne-mi-dun-am kodum.*  
 NEG-PROG-know-1SG which  
 ‘Keyhân has to read one of those books, but I don’t know which.’
- b. *keyhân mi-xâd yeki az doxtar-â-ye Ali-\*(ro) be mehmuni*  
 keyhân PROG-want.3SG one of daughter-PL-ez Ali-ro to party  
*davat bo-kon-e, va mohem nist kodum.*  
 invite SUBJ-do-3SG and important is.not which  
 ‘Keyhân wants to invite one of Ali’s daughters to the party, and doesn’t care which (daughter).’

These examples highlight a property of partitive expressions generally that they not entail either epistemic specificity or identifiability in principle. In this sense, it differs from epistemic specificity, as we observed with *a certain* indefinites above, which entails the latter. This is because the restrictions placed on the valuation instructions of the object stem from being part of a constrained domain (i.e a partitive set), rather than the epistemic state of the speaker. There is therefore no constraint placed on the referent of partitives that they be minimally identifiable in principle. Thus, the discourse referent for partitives is capable of being epistemically non-specific, as in (58a), as well as unidentifiable, as in (58b). Epistemically specific indefinites, however, are not identified in the current discourse, but are identified in the speaker’s mind, and therefore identifiable in principle with respect to the discourse as a whole.

The independent nature of these properties raised the question of whether partitives should be properly included in the same position as epistemically specific indefinites. To account for this observation, I argue that the properties of partitive and epistemic specificity be assigned to two distinct positions in the definiteness hierarchy, as argued for in Farkas (2002b). Our updated definiteness hierarchy is therefore as in (59):

- (59) Personal pronoun > Proper name > Definite > **Partitive spec.** > **Epistemic spec.** > Non-specific

This enrichment makes a prediction that there may exist a language in which the line between marked and unmarked objects was drawn between partitives and epistemically specific indefinites. Hungarian provides a good example of such a language (D. Farkas 2016 p.c.). In Hungarian, the choice of verbal conjugation is sensitive to the definiteness of the object. Definite objects (or higher) trigger definite conjugation agreement suffixes, while indefinite objects trigger indefinite conjugation suffixes. The sentences in (60)-(62) exemplify this interaction. In (60), the indefinite object triggers the indefinite conjugation on the verb, even if a reading is brought out by the adjective *bizonyos* ‘certain’ in which it is minimally identifiable in principle.

- (60) a. *Lát-**ok*** egy (*bizonyos*) *diák-ot*.  
 see-1SG<sub>indef</sub> a certain student-ACC  
 ‘I see a (certain) student.’
- b. \**Lát-**om*** egy (*bizonyos*) *diák-ot*.  
 see-1SG<sub>def</sub> a certain student-ACC  
 Intended: ‘I see a (certain) student.’<sup>4</sup>

The example (61) below illustrates an analogous example in which the context itself favors a reading of epistemic specificity. In this case, the indefinite conjugation is still obligatory, and the definite conjugation is ungrammatical.

- (61) a. *Meg fog-**ok*** hívni egy *szomszéd-ot*.  
 PRTL AUX-1SG<sub>indef</sub> invite a neighbor-ACC  
 ‘I will invite a neighbor.’
- b. \**Meg fog-**om*** hívni egy *szomszéd-ot*.  
 PRTL AUX-1SG<sub>def</sub> invite a neighbor-ACC  
 Intended: ‘I will invite a neighbor.’

We observe a different effect for partitive indefinites. Unlike epistemically specific indefinites, partitives obligatorily trigger the definite conjugation suffixes on the verb, shown in (62):

- (62) a. *Lát-**om*** egyik *barátnő-d-et*.  
 see-1SG<sub>def</sub> one.of friend-2SG-ACC  
 ‘I see one of your friends.’

<sup>4</sup>Thanks to Donka Farkas for these examples.



- b. \**Lát-ok*      *egyik barát(n)ő-d-et*.  
 see-1SG<sub>indef</sub> one.of friend-2SG-ACC  
 Intended: ‘I see one of your friends.’

Hungarian thus has a DOM system which is sensitive to the distinction between partitive indefinites and all other types of indefinites. If partitive specificity constitutes a distinct rung on the definiteness hierarchy which is ranked higher than other semantic properties of indefinites (e.g. epistemically specificity, identifiability, or non-specificity), then we can show why both definites and partitive behave similarly in Hungarian.

While Persian does not make the markedness distinction between epistemic and partitive specificity that Hungarian does, the fact that partitive specificity unilaterally triggered DOM in Persian was evidence that the distinction is important enough to be reflected in the hierarchy constraining DOM, as argued for in Farkas (2002b). I take the observation from Hungarian to be confirming evidence that this distinction is manifested in the definiteness hierarchy.

### 8.3 Strong identifiability

With the separation of partitive specificity from other properties relevant to the referential stability of indefinites, we are left to account for the other cases of *a certain* indefinites which were identifiable in principle, but were non-specific in both the partitive and epistemic sense. In the previous section, we furnished our typology with both strong and weak varieties of this property. In English, *a certain* was compatible with both types, but in Persian the distinction was signaled by the presence of DOM. This contrast was shown in (44) and (45), repeated below in (63a) and (63b).

- (63) *mârk ye kâghaz-o râje-be in mozu neveshte ke be-khâter-esh*  
 Marc a paper-ro about this topic wrote.3SG that about-3SG  
*mashhur shod.*  
 famous became.3SG  
 ‘Mark has written a paper on this topic, for which he has become famous.’

- b. *man ye kâghaz nevesht-am va mariam kafsh xarid.*  
 I a paper wrote-1SG and Mariam shoes bought.3SG  
 ‘I wrote a paper and Mariam went shopping.’

In other words, in Persian there is a requirement that DOM appear on identifiable objects whose identity is singled out in the future discourse. This differentiates them from objects that

are identifiable, but which are not singled out in this way. I explicate this requirement by embedding the distinction between strong and weak identifiability in principle within the definiteness hierarchy. This leaves nominals which are unidentifiable as the lowest rung of the hierarchy, distinct from weak identifiability. In sum, the modifications indicate that the category of specific indefinites in the hierarchy in our original hierarchy grouped together four distinct categories of referential stability: partitive specificity, epistemic specific, strong identifiability in principle and weak identifiability in principle.

With these modifications, our updated hierarchy is as in (64), below.

(64) Personal pronoun > Proper name > Definite > Partitive Spec. > Epistemic Spec. >  
**Strong Identifiability > Weak Identifiability > Unidentifiable in principle**

Our conclusion for Persian is that DOM is triggered on all objects that are strongly identifiable, or higher, while those that were weakly identifiable or were unidentifiable in principle were left unmarked. This provides an explanation for the apparent optionality of DOM on epistemically non-specific indefinites like those in (63) on the basis whether or not they were strongly identifiable. This also allows us to explain why DOM was still obligatory on partitives that were unidentifiable or epistemically specific.

The enrichments illustrated above make some additional predictions about other possible types of DOM systems cross-linguistically. For example, this predicts that there may exist a language with a DOM system in which DOM were triggered on all objects meeting the criteria of weak identifiability in principle, but not on unidentifiable objects. Similarly, another language may mark all objects with epistemic specificity or higher, leaving all non-specific objects with either strong or weak identifiability unmarked. Currently, I can find no such languages, but leave this as an open question for future study.

Before moving on, a final word is required concerning scopal specificity. It was observed in the examples above that DOM is obligatory in Persian on all wide-scoping indefinites with respect to negation. This includes *a certain* indefinites, which require the property of identifiability in principle. On this basis, we might suggest that we require an additional rung to accommodate the property of scopal specificity. However, we have seen that independently of this latter property that DOM is obligatory on objects with strong identifiability generally. Thus

while scopal specificity is functionally distinct from its epistemic and partitive counterparts, it entails that the DP be at least strongly identifiable in principle. This is a one-way entailment, however. Recall example (50) above from Jasbi (2014), repeated below in (65):

- (65) *man emruz kâr-i-(ro) anjâm na-dâd-am.*  
 I today task-IND-ro complete NEG-gave-1SG  
 'I didn't do a task today.' ¬ > ∃ /\*∃ > ¬

Here, Jasbi argues that the object is scopally non-specific, yet is able to receive DOM. If this is a correct analysis, then in the current approach, we could reanalyze this as an instance in which the task in question is selected from an implicit partitive subset (i.e. the tasks I had to do today). Alternatively, we could analyze this as being strongly identifiable in principle. In the latter situation, the fact that we have analyzed scopal specificity and strong identifiability in principle as distinct properties is greatly beneficial. It allows us to explain the appearance of DOM on scopally specific nouns by the fact that DOM is obligatory on DPs which have strong identifiability in principle, and to include a rung solely for scopal specificity would be to include a redundancy into our typology.

## 9 An updated OT analysis

In this section, I provide an updated analysis of the OT account of DOM as formalized in Aissen (2003) which reflects the enrichments made to the definiteness hierarchy in (64). Her analysis reflects the intuition that DOM is the outcome of a contention between *iconicity* and *economy*. More particularly, DOM is in part a reflection on the fact that properties high up on these hierarchies are more iconic to subjects, and less iconic to objects. Similarly, the lower parts of the hierarchy are more iconic of objects. The principle of economy constrains the distribution of overt structural Case marking on the object. With respect to definiteness, this pattern is captured by the hierarchies in (66):

- (66) a. Su/Pro > Su/PN > Su/Def > Su/Spec > Su/NSpec  
 b. Oj/Nspec > Oj/Spec > Oj/Def > Oj/PN > Oj/Pro Aissen (2003, p445)

The hierarchical nature of this phenomenon makes it a good candidate for an Optimality Theoretic analysis. In OT terms, the hierarchies above are closely analogous to ranking of

markedness constraints, as presented in (67), below. The highest constraint in (67b) penalizes objects that are pronominal, followed by a constraint which penalizes objects that are proper names, etc. This is the inverse of the ranking in (67a) which penalizes non-specific indefinite subjects over specific indefinite subjects.

- (67) a. \*Su/NSpec > \*Su/Spec > \*Su/Def > \*Su/PN > \*Su/Pro  
 b. \*Oj/Pro > \*Oj/PN > \*Oj/Def > \*Oj/Spec > \*Oj/NSpec Aissen (2003, p445)

This captures the intuition that having a great degree of definiteness, and consequently greater referential stability, makes an NP a more canonical subject, rather than an object. This captures the variation among DOM systems among the world's languages, in addition to the related phenomenon of differential subject marking, in a typology which also accounts for their structural similarities.

While this set of constraints determines how marked a certain association between objects and definiteness will be, other constraints are required to favor or penalize DOM as an instance of structural Case marking. To capture this, Aissen provides two additional constraints: \* $\emptyset_c$  and \*STRUC<sub>c</sub>.

The constraint \* $\emptyset_c$  penalizes the absence of structural Case marking, indirectly triggering overt Case morphology. This constraint conjoins with each of the constraints in (67b) via local constraint conjunction. This is to the effect that DOM is triggered more forcefully on more marked objects than on less marked objects. The result is the ranked set of locally conjoined constraints in (68) below, for objects:

- (68) \*Oj/Pro & \* $\emptyset_c$  > \*Oj/PN & \* $\emptyset_c$  > \*Oj/Def & \* $\emptyset_c$  > \*Oj/Spec & \* $\emptyset_c$  > \*Oj/NSpec & \* $\emptyset_c$   
 Aissen (2003, p448)

The constraint \*STRUC<sub>c</sub>, which penalizes instances of structural Case, regulates DOM and accounts for the cross-linguistic variation with respect to where on the hierarchies will a language draw the distinction between marked and unmarked objects by being placed at a certain position in the ranking in (68).

To account for a language like Hebrew, \*STRUC<sub>c</sub> must be ranked in between \*Oj/Def & \*ø<sub>c</sub> and \*Oj/Spec & \*ø<sub>c</sub>. As a result, DOM is ruled out on everything below definite descriptions, given the violations that would be incurred by \*STRUC<sub>c</sub>. DOM would be allowed on definite objects and higher, which is what we observe. For Persian, Aissen assumes \*STRUC<sub>c</sub> is ranked below \*Oj/Spec, thus accounting for DOM on specific indefinites.

In the analysis above, I have replaced the rung of specific indefinites with a more complex arrangement of semantic properties. I have distinguished partitive specificity from epistemic specificity, and I have distinguished strong identifiability from weak and unidentifiability. In our OT framework, these can be easily accounted for by mapping these properties onto unique constraints used in the ranking. To account for partitives, I propose a constraint \*Oj/Part, properly ranked between \*Oj/Def and \*Oj/Spec. For the sake of clarity, the constraint \*Oj/Spec will still be used to refer to objects that are epistemically specific. Similarly, we require an additional constraint \*Oj/StrongID which favors DOM on objects that are strongly identifiable that ranks directly below \*Oj/Spec. This latter constraint ranks above \*Oj/NSpec, which includes weak identifiability in principle and unidentifiability. Our updated ranking is thus as in (69):

$$(69) \quad *Oj/Pro \ \& \ *ø_c > *Oj/PN \ \& \ *ø_c > *Oj/Def \ \& \ *ø_c > *Oj/Part \ \& \ *ø_c > *Oj/Spec \ \& \ *ø_c > *Oj/StrongID \ \& \ *ø_c > *Oj/NSpec \ \& \ *ø_c$$

With this updated list of constraints and their proper ranking, we can now account for the fact that Persian marks all objects with strong identifiability or higher by placing \*STRUC<sub>c</sub> below \*Oj/StrongID, and above \*Oj/NSpec. This is shown in (70):

$$(70) \quad *Oj/Pro \ \& \ *ø_c > *Oj/PN \ \& \ *ø_c > *Oj/Def \ \& \ *ø_c > *Oj/Part \ \& \ *ø_c > *Oj/Spec \ \& \ *ø_c > *Oj/StrongID \ \& \ *ø_c > *STRUC_c > *Oj/NSpec \ \& \ *ø_c$$

As our enriched typology of definiteness permits a further distinction between weak identifiability in principle and unidentifiability, I contend for the sake of simplicity that we do not currently have need of such a subdivision to account for Persian. However, my analysis does predict the existence of a language with a DOM system that differentiated objects on this very

distinction. In such a situation, this could easily be accommodated in the OT framework presented here by the inclusion of a constraint \*Oj/WeakID, corresponding to weak identifiability in principle, with \*STRUC<sub>c</sub> appropriately ranked.

## 10 Conclusion

In this paper, I have shed light on the complex phenomenon of DOM in Persian in the typological framework of Aissen (2003). I provided evidence that the rung of specific indefinites on the definiteness hierarchy can be subdivided in four ways to account for instances of DOM in Persian that otherwise are unexplained. In line with Farkas (2002b) in which the definiteness hierarchy represents the referential stability of objects' discourse referents, I argue that partitive specificity should be separated from epistemic specificity, and both represent unique rungs on the definiteness hierarchy. Partitive specificity ranks above epistemic specificity as indefinites must receive DOM if they are partitive even if they are epistemically non-specific. While both types of specificity trigger DOM in Persian, this enrichment has the benefit of predicting a DOM system which marks partitive indefinites, but not indefinites that are epistemically specific but are not partitives. We saw that such a case is instantiated in Hungarian which differentiates objects in this manner by means of the choice of verbal conjugation.

I also argued that DOM is sensitive to the property of identifiability in principle. I indicated that the DOM system of Persian was sensitive to an additional distinction between strong and weak identifiability on the basis of the role of the indefinite in the future of the discourse. I argued the obligatory nature of DOM on indefinites with strong identifiability could be captured by including both strong and weak flavors into the definiteness hierarchy. With these distinction in referential stability incorporated into the definiteness hierarchy, I argued that Persian marks objects with the property of strong identifiability or higher, and leaves unmarked indefinites that are only weakly identifiable, unidentifiable, and non-specific.

## References

- Aissen, J. (2003). Differential object marking: Iconicity vs. economy. *Natural Language & Linguistic Theory*, 21(3), 435–483.
- Bossong, G. (1991). Differential object marking in romance and beyond. *New analyses in*

- Romance linguistics*, 143–170.
- Croft, W. (1988). Agreement vs. case marking and direct objects. *Agreement in natural language: Approaches, theories, descriptions*, 159–179.
- Dabir-Moghaddam, M. (1992). *On the (in)dependence of syntax and pragmatics: Evidence from the postposition "-râ" in persian*. Mouton de Gruyter.
- Enç, M. (1991). The semantics of specificity. *Linguistic inquiry*, 1–25.
- Farkas, D. (2002a). Specificity distinctions. *Journal of semantics*, 19(3), 213–243.
- Farkas, D. (2002b). Varieties of indefinites. In *Semantics and linguistic theory* (Vol. 12, pp. 59–83).
- Farkas, D., & von Stechow, K. (2003). Stability of reference and object marking in romanian. *Ms. Universität Stuttgart*.
- 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.
- Hopper, P. J., & Thompson, S. A. (1980). Transitivity in grammar and discourse. *Language*, 251–299.
- Jasbi, M. (2014). *The semantics of differential object marking in Persian* (Unpublished master's thesis). Stanford University.
- Kadmon, N., & Landman, F. (1993). Any. *Linguistics and philosophy*, 16(4), 353–422.
- Karimi, S. (1996). Case and specificity: Persian râ revisited. *Linguistic Analysis*.
- Karimi, S. (2003). On object positions, specificity, and scrambling in persian. *Word order and scrambling*, 91–124.
- Karimi, S. (2005). *A minimalist approach to scrambling: Evidence from persian* (Vol. 76). Walter de Gruyter.
- Mahootian, S., & Gebhardt, L. (1997). *Persian. descriptive grammars*. London: Routledge.
- Modarresi, F. (2015). *Bare nouns in persian* (Unpublished doctoral dissertation). Humboldt-Universität zu Berlin, Philosophische Fakultät II.
- Toosarvandani, M., & Nasser, H. (2015). Quantification in persian. *Handbook of quantifiers in natural language, 2nd edn*. Springer.