Some Unique Semantic Properties of Farsi

1   Introduction

Language acquisition can be characterized as the process of discovering forms of the target language, discovering the concepts that are communicated in that language, and mapping those forms to the correct concepts (Clark, 2009). Each step of acquisition has its own challenges to overcome. In this paper I discuss issues that are closely related to the last step: when learning a new language, how do learners discover the right mapping of forms and meanings? How do they converge on the correct characterization of a word's meaning among many plausible candidate meanings? Are some words or elements of a language more difficult to learn because of what they mean? How can we facilitate learning in cases where the semantics of words make their learning challenging?

Words in almost any language divide into two basic categories: content words and function words. Content words are the ones that are most word-like to native speakers! They are the flagbearers of the lexicon. They constitute nouns like *cat*, verbs like *run*, and adjectives like *red*; words that if you ask someone to name a word, they will most likely name those. Function words on the other hand are hardly recognized. They are small elements like *the*, *and*, *or*, and *every*, which are unlikely to be mentioned if someone is asked to name a word. Despite this lack of recognition, the role of function words in language is truly remarkable. They are the nuts and bolts that put content words together and create the sentences of a language. In short, function words form the backbone of a language structure and languages

cannot function without them.

Function words are small in number and they rarely add a new member. This is why they are called a "closed class". Content words on the other hand are much more numerous and easily add new members; hence the name "open class". Function words are present everywhere. If we compute the frequency of words in a document like this one, the most frequent words are going to be function words. Given their small number and high frequency, one might think that function words are easy to learn. However, this is the opposite of what we find in language acquisition. Children master function words of their first language much later than content words. But why are function words hard to learn?

Despite their frequency and ubiquitous presence, function words have meanings that are extremely abstract and hard to pin down. In introductory classes, I often ask students to provide their intuitions on the meaning of some content words and some function words. Students are often quick in responding to content words. The meanings seem very clear to them. However, with function words, they often do not really know what the meaning is, even though they know how to correctly use them in a sentence. Research in formal semantics and pragmatics has shown that the meanings of function words are indeed extremely subtle. We often need specific and precise mathematical tools to be able to capture their semantics. Therefore, it is not surprising that despite their high frequency in language, function words are learned later in the process of acquisition. To find the correct form-meaning mapping for function words, learners may need more data and time. More importantly, to distinguish subtle differences in meaning, learners may need to rely on crucial data points that differentiate semantic hypotheses from each other. But is it possible to facilitate the process of learning function words?

It is, in principle, possible to facilitate the process of form-meaning mapping if we provide the crucial data that learners need to converge on the right semantic hypotheses for words and expressions. For learners of a second language, it is also possible to provide grammatical generalizations that accurately capture the semantic contribution of functional elements.

Research in theoretical and formal semantics and pragmatics can help with both of these tasks. The goal of the present chapter is to provide recent findings on the semantics of some functional elements in Farsi, and provide examples that in my view count as "crucial examples" for learning, because they bring out the crucial semantic distinctions that those functional elements encode. The chapter also aims at demonstrating how theoretical research on semantics of function words can contribute to the literature and practices in second language acquisition.

I first start with a discussion of diglossia in Farsi (Persian). I argue that the informal and formal varieties of Farsi differ significantly in their phonology, syntax, (and most relevant to the discussion here) their semantics. Rules and generalizations that apply to one do not necessarily carry to the other. Therefore, in research and teaching of Farsi, it is important to keep the two systems apart and systematically highlight the similarities and differences. Section 3 discusses subtleties in the interpretation of bare nominals in Farsi. Section 4 explains the semantic contribution of some functional markers that appear on singular nominals while Section 5 focuses on plural marking. Finally, Section 6 discusses the object marker $r\bar{a}$ which is infamous for creating major difficulties for speakers of Farsi as a second language. In each case, I present the type of data that I consider crucial for bringing out the core semantic contribution of each marker and hence helpful in the process of language acquisition.

2   Diglossia

One of the most challenging aspects of learning Farsi is that what is spoken, namely Colloquial or Informal Farsi is considerably different from what is taught and written, i.e. Formal Farsi. The formal and informal varieties of Farsi are closely and systematically related but obey different rules and must be considered two separate systems. While Informal Farsi has been subject to rapid linguistic change, Formal Farsi has remained relatively closer to Literary Persian, which was spoken hundreds of years ago. Examples in (1) show similar

sentences in Literary Persian (1a), Formal Farsi (1b), and Informal Farsi (1c). The Literary example in (1a) which dates back to 700 years ago is quit similar to the Formal example in (1b). However, the Informal example in (1c) is very different from the other two. I have shown the differences between (1b) and (1c) with numbered boxes:

(1) a. Literary (1300-1371 CE): SOV

　　　chon $[(u)]_S$ [be xāne $]_{PP}$ [raft-ø$]_V$
　　　when (he)   to home   went-3.SG

　　　"When (he) went home."   (Ubayd Zākāni[1])

　　b. Modern Formal: $\boxed{\text{SOV}}^1$

　　　$[\text{Rezā}]_S$ $[\boxed{\text{be}}^3$ $\boxed{\text{xāne}}^2$ $]_{PP}$ [raft-$\boxed{ø}^4$ $]_V$
　　　Reza   to   home   went-3.SG

　　　"Reza went home"

　　c. Modern Colloquial: $\boxed{\text{SVO}}^1$

　　　$[\text{Rezā}]_S$ [raft-$\boxed{\text{esh}}^4$ $]_V$ $\square^3$ [$\boxed{\text{xune}}^2$ $]_{NP}$
　　　Reza   went-3.SG.CLC   home

　　　"Reza went home."

First, as (1c) shows, in Informal Farsi it is more natural to use the SVO word order for the sentence "Reza went home". However, in Formal Farsi it is more acceptable to use the SOV word order as in (1b). Second, the phonological form of "home" changes from /xune/ to /xane/ when we switch to Formal Farsi. Third, "home" can appear as an NP without a preposition, next to the verb *raft* "go" in Informal Farsi. This is ungrammatical in the formal variety. As (1b) shows, we need "home" to be preceded by the preposition *be* "to". Fourth, it is possible to use the third person singular clitic *esh* on the verb in Informal Farsi to show agreement with the subject of the sentence. This is totally ungrammatical in

---

[1]Resāle-ye Delgoshā: Ārmān Dozdi

4

Formal Farsi. Instead the verb should bear the third person subject-agreement suffix, which is zero-marking.

Most speakers of "Farsi" grow up learning an informal variety such as Tehrani or Shirāzi Farsi and only learn Formal Farsi through primary education and schooling where they learn to read and write. While Formal Farsi is useful in understanding the language of the news, literary texts, or formal communications, it is not as useful as Informal Farsi in day-to-day conversations. Many students of Farsi as a second language learn it because they would like to communicate with Farsi-speaking friends and family. This includes a growing number of heritage speakers who have had limited exposure to Farsi at home and would like to improve their conversational skills. However, most programs teaching Farsi as a second language focus on teaching Formal Farsi. In fact, many students are not aware of the relatively large gap between Formal and Informal Farsi, and are often surprised to find out that what they learned in class does not apply to day-to-day conversation. In the following sections, we see that the differences between Formal and Informal Farsi extend to semantics as well. Therefore, it is important to teach these differences and make students aware that Farsi has two interconnected varieties.

3   Bare Nominals

Nominals often appear bare in Farsi, which means they do not receive any morphological marking. A bare nominal can be interpreted in multiple ways. Consider the examples in (2). The bare nominal *mashin* "car" can be interpreted as a definite (2a), an indefinite (2b), a generic (2c), or a "numberless" nominal (2d). I call examples like (2d) numberless, because they are equally felicitous describing a singular or plural state of affairs. For example, (2d) can be used to warn someone about an incoming car in the street or several cars that are approaching. My own intuition is that numberlessness is something that is available in Farsi but not English. A similar example may be gender in the pronominal systems of Farsi and

English. Farsi does not make gender distinctions in its pronominal system but English does for the third person. As a result, it is hard to translate the pronominal clitic in (2b) into English. Since there is no exact lexical item available for translation, we have to resort to a disjunction such as "him or her".

(2) a. māshin xarāb-e
     car   broken-be.PRE.3.SG
     "The car is broken."

  b. māshin be-sh zad
     car   to-3.SG hit.PST.3.SG
     "A car hit him/her."

  c. māshin gerun-e
     car   expensive-be.PRE.3.SG
     "Cars are expensive."

  d. māshin mi-ād
     car   HAB-come.PRE.3.SG
     "Cars are coming (or a car is coming)."

The fact that bare nominals can be interpreted in multiple ways makes them particularly hard to learn for speakers of languages like English that often require nominals to be marked by determiners. Therefore it might be helpful for learners to focus on bare nominals and discuss examples like (2) which illustrate the wide range of interpretations bare nominals receive. Such examples can also be contrasted with examples from the learners' first language. The comparison helps learners see how distinctions made by functional elements in their first language collapse in Informal Farsi.

4   Singular Marking

In this section I discuss three functional markers: the indefinite determiner *ye(k)*, the indefinite clitic *-i*, and the nominal suffix *-e*. I focus on their semantics in Formal and Informal Farsi. For a more detailed semantic analysis of these elements please refer to (Jasbi, 2016, pearb). Let us start with the indefinite determiner *ye(k)*. The split between Formal and Informal Farsi shows itself immediately. The indefinite determiner in Formal Farsi is *yek* while in Informal Farsi it is *ye*. In Informal Farsi, the semantic behavior of *ye* is very close to the English indefinite determiner *a(n)*. The main difference is that *ye* is much more resilient

against being interpreted under negation than *a(n)*. For example in (3a), the indefinite determiner *ye* interacts with the universal *hame* "all" in a way that both wide and narrow scope readings are available: everyone watched the same movie (wide), everyone watched a different movie (narrow). This is very similar to the English translation of the sentence. However, when *ye* is forced under negation in (3b), the utterance becomes infelicitous (# is used to mark the infelicity). This is not the case for the English equivalent. The sentence "Nobody watched a movie" is a felicitous utterance conveying that no movie was watched (narrow scope).

(3) a. hame ye film tamāshā kard-an
     everyone ID film watch do.PST-PL
    "Everyone watched a movie."

   b. # hishki ye film tamāshā na-kard
     nobody ID film watch NEG-do.PST.3.SG
    "Nobody watched a movie."

This behavior of *ye* might be due to its division of labor with the indefinite clitic *-i* in informal Farsi. The semantic behavior of the clitic *-i* most resembles the determiner *any* in English. The clitic *-i* is unacceptable in simple positive episodic environments without modification as (4a) shows. Now compare (4a) to (4b) which is the formal version of the same sentence using the formal form of the verb "come" *āmad*. The indefinite clitic can be used to convey an indefinite meaning on its own in Formal Farsi but not informal Farsi. I should add that it is easy to find the indefinite clitic as the main marker of indefiniteness in older texts of Farsi. This suggests that the difference between Formal and Informal uses of this clitic may be due to historical shift in its meaning.

(4) a. * zan-i   umad
     woman-IC come.PST.3.SG
    "A woman came."

   b. zan-i   āmad
     woman-IC come.PST.3.SG
    "A woman came."

Similar to *any* in English, the indefinite clitic becomes acceptable in positive episodic environments if it is further modified; known as "subtrigging" in the linguistics literature (LeGrand, 1975). The example in (5) is identical to the one in (4a) except that it is modified by the relative clause *ke goft-i* "that you talked about". To most native speakers, (5) sounds a lot better than (4a) in informal Farsi. You may have noticed that the addition of the relative clause made something curious happen: the interpretation of the nominal "woman" is now definite rather than indefinite. This phenomenon has puzzled Iranian linguists for decades. Data like (5) make it hard to classify *-i* as a simple indefinite marker. Therefore, some Iranian linguists such as Moin (1958, 235) and Natel-Khanlari (1972, 255) proposed that the indefinite clitic is polysemous. In (Jasbi, 2016), I argued that the definite interpretation in examples like (5) is not due to the clitic *-i* but rather the result of the compositional structure of the sentence. In this analysis, *-i* conveys that the nominal *zan* "woman" is non-unique. If the nominal stays unmodified, then the sentence is most compatible with an indefinite reading given that it conveys the non-uniqueness of "woman". However, when the *i*-marked nominal is further modified by a restrictive relative clause such as *ke gofti* "that you talked about", then the whole NP "woman that you talked about" can pick out a unique woman. Therefore, the most likely interpretation in such cases is a definite one. (Jasbi, 2016) provides a more detailed and formal account of this analysis.

(5) zan-i   ke goft-i   umad
     woman-IC that tell.PST-2.SG come.PST.3.SG
     "The woman that you talked about came."

Again similar to *any* in English, the indefinite clitic in Farsi (especially Informal Farsi) is licensed in "downward entailing" or "non-veridical" environments (Ladusaw, 1980; Giannakidou, 1998). (6) shows examples of such environments (negative sentences, questions, conditionals) for the English word *any*. While *any* is not acceptable in positive episodic sentences (6a), it sounds very natural in downward entailing or non-veridical environments (6b-6d).

(6)   a.   # He bought any book.

      b.   He didn't buy any book.

      c.   Did he buy any book?

      d.   Tell us if he bought any book!

We see a parallel situation with the indefinite clitic -*i* in Farsi. Examples in (7) show this clearly. It is important to note that all the examples in (7), including (7a), are acceptable in Formal Farsi. Therefore, in Formal Farsi the clitic -*i* does not conform to the pattern of *any* in English. Students of Farsi that are mainly taught the Formal variety, therefore, may sound Formal or even poetic, because of using structures like (7a) in their colloquial speech. An important part of fluency in speaking Farsi, is to learn the separate syntactic or semantic rules that govern Formal and Informal Farsi, and be able to apply them depending on the context.

(7)   a.   # ketāb-i   xarid
          book-IC buy.PST.3.SG
       "S/he bought a book."

      b.   ketāb-i   na-xarid
        book-IC NEG-buy.PST.3.SG
       "S/he did not buy any book."

      c.   ketāb-i   xarid?
        book-IC buy.PST.3.SG
       "Did s/he buy any book?"

      d.   age ketāb-i   xarid,   be mā   be-gu!
        if   book-IC buy.PST.3.SG, to 1.PL SUB-say
       "If s/he bought any book, tell us!"

The third marker that I discuss here is the nominal suffix -*e*. This suffix is an innovation of Informal Farsi and does not seem to be used in the Formal variety at all. I bring example (8) below to illustrate this point. Notice the formal and informal forms of the verb "come".

(8) a. zan-e   umad
        woman-UM come.PST.3.SG

     "The woman came."

     b. * zan-e   āmad
          woman-UM come.PST.3.SG

     "The woman came."

This suffix is often described as the informal definiteness marker. Again, the story is not as simple as that. Examples like (9) show that the same suffix can appear with the indefinite determiner *ye* and convey an indefinite interpretation. So what does *-e* really do in informal Farsi?

(9) ye zan-e   umad
     ID woman-UM come.PST.3.SG

     "A (certain) woman came."

The clue is the word *certain* in the translation of (9). In short, the suffix *-e* acts similar to the adjective *certain* in English (see Jasbi (pearb) for a more detailed semantic analysis). In other words, it adds determinedness to the referent or value of a nominal such as *zan* "woman". Therefore, a noun modified by *-e* such as *ketāb-e* communicates that its referent is fixed. Now, this fixedness may be because the referent is known to the conversational participants, in which case no indefinite marker accompanies it and a definite reading is derived. Or alternatively, the referent might not be known the conversational participants, yet the speaker may want to convey that despite being unknown, the referent of the nominal is fixed.

This "determinedness" or "fixedness" of the nominals marked with *-e* results in interesting semantic patterns. For example, in (10a) below, the sentence without the nominal suffix is ambiguous between two readings: 1. everyone said "hello" to a different professor, and 2. everyone said "hello" to the same professor. Once *-e* appears on the nominal *ostād* "professor" in (10b), the referent of "professor" becomes fixed and can't vary with different individuals who said hello. Therefore, in (10b) the only available reading is the one where

everyone said hello to the same professor.

(10) a. emruz hame   be ye ostād   salām kard-im
         today everyone to ID professor hello   do-1.PL
         "Today we all said hello to (a different/ the same) professor."

      b. emruz hame   be ye ostād-e   salām kard-im
         today everyone to ID professor-UM hello   do-1.PL
         "Today we said hello to the same professor."

Examples below show that this phenomenon is systematic and not isolated to universal quantification with *hame*. In (11) the nominal suffix can disambiguate that Sara always gets into fights with the same boy and not different ones. In (12), it helps us know that the girl Amir is going to marry is determined and Amir is not just looking for some girl or other to marry.

(11) a. Sārā hamishe bā   ye pesar davā-sh   mi-sh-e
         Sara always   with ID boy   quarrel-3.SG IPFV-become-3.SG
         "Sara always gets into a fight with (a different/ the same) boy."

      b. Sārā hamishe bā   ye pesar-e davā-sh   mi-sh-e
         Sara always   with ID boy-UM quarrel-3.SG IPFV-become-3.SG
         "Sara always gets into a fight with the same boy."

(12) a. Amir mi-xā-d   bā   ye doxtar ezdevāj kon-e
         Amir IPFV-want-3.SG with ID girl   marry   do-3.SG
         "Amir wants to marry a girl." (1. ∃ > WANT 2. WANT > ∃)

      b. Amir mi-xā-d   bā   ye doxtar-e ezdevāj kon-e
         Amir IPFV-want-3.SG with ID girl-UM   marry   do-3.SG
         "There is a girl Amir wants to marry." (∃ > WANT)

In the next section, I explore the role of functional elements in creating plural nominals in Farsi, focusing more on Informal Farsi.

5 Plural Marking

Plurals in Farsi are formed using two different mechanisms. The first mechanism is the plural morpheme *hā* as in (13a) . The second mechanism is the combination of a plural numeral

such as *do* "two" or *chand* "many" as in (13b). These two mechanisms cannot be used at the same time. As sentence (13c) shows, the plural marker *hā* and the plural numeral *chand* cannot appear together. I should add that, if a sentence that contains *chand* is given a rising question intonation, then *chand* acts like a question word such as "how many".

(13)   a. zabān-shenās-(h)ā   injā neshast-an
             language-expert-PL.DEF here sit-3.PL
             "The linguists are sitting here."

        b. chand tā zabān-shenās   injā neshast-an
             many CL language-expert here sit-3.PL
             "Some linguists are sitting here."

        c. * chand tā zabān-shenās-hā   injā neshast-an
              many CL language-expert-PL.DEF here sit-3.PL

These two mechanisms interact with definiteness. In (13a) where "linguist" bears the plural suffix *hā*, the sentence receives a definite interpretation: "the linguists". In (13b) where "linguist" is only modified by the numeral modifier *chand*, the sentences receives an indefinite interpretation: something like "some linguists" or "several linguists". (13c) shows that numeral modifiers and the plural marker *hā* cannot appear together to mark plurality. Based on such examples, linguists such as Ghomeshi (2003) and Gebhardt (2009), have suggested that an NP marked by *hā* is both plural and definite.

It is also tempting to conclude that an NP with a numeral and a classifier such as *chand ta* is a plural indefinite. However, the data in (14) below show that such constructions can receive a definite interpretation.

(14)   a. in   chand tā aks-o   pāk kon!
             this/these some CL picture-OM clean do-3.SG
             "Delete these pictures!"

        b. in   aks-hā-ro   pāk kon!
             this/these picture-PL.DEF-OM clean do-3.SG
             "Delete these pictures!"

In (14a), *chand* is modifying *aks* meaning "picture" but the interpretation of the NP is

definite ("these pictures") due to the presence of the demonstrative *in* meaning "this/these". (14b) shows that the same meaning can be expressed by the plural definite suffix *hā*. The main difference between (14a) and (14b) is that the former has a partitive meaning; it is implied that there are more pictures and only some of them (the ones the speaker is referring to) should be deleted. We see the same pattern in singular nouns with *ye* meaning "one":

(15)   a.   ye   aks-o   pāk   kon!
             one picture-OM clean do.3.SG
             "Delete a picture!"

          b.   in   ye   aks-o   pāk kon!
             this one picture OM   clean do-3.SG
             "Delete this one picture!"

In (15a) where we have no demonstrative pronoun, the NP "one picture" is interpreted as indefinite. However, in (14b) the same NP is interpreted as definite due to the presence of the demonstrative *in* "this". This suggests that numerals such as *ye(k)* "one" or *chand* "many" in Farsi can act both as an indefinite determiner or a simple cardinal number. One possible account is that numerals in Farsi only provide number information and (in)definiteness is provided via covert semantic operations.

6   Object Marking

The Persian object marker, formally known as *rā*, is pronounced in colloquial Persian as *ro* or simply *o*. *ro* is used in the phonological environment where the preceding phoneme is a vowel and *o* is used if the preceding phoneme is a consonant:

Persian Object Marker	V__	C__
Formal Persian	*rā*	*rā*
Colloquial Persian	*ro*	*o*

The distribution of the object marker *rā* in Persian is determined by the interaction of syntactic and semantic factors. Syntax provides the environment where appearance of

$r\bar{a}$ is possible and semantics determines the conditions which make the occurrence of this marker necessary. I first explain where $r\bar{a}$ is allowed syntactically and then describe where the semantics of the nominal determines the occurrence of $r\bar{a}$.

The object marker $r\bar{a}$ appears only on nominals. It is ungrammatical on subjects (16a) and PP arguments of the verb (16b). It is grammatical on direct objects (16c) and certain nominal adverbials (16d). It can also participate in constructions such as (16e) which are called Clitic-Binder Constructions by Karimi (1990)[2].

(16) a. on subjects[3]:

[Maryam]$_S$ $\boxed{(\text{*o})}$ [keik]$_{DO}$ [xord-ø]$_V$
Maryam   OM   cake   ate-3.SG

"Maryam ate cake."

b. on PP arguments of the verb:

[Maryam]$_S$ be [barādar-esh]$_{IO}$ $\boxed{(\text{*o})}$ [keik]$_{DO}$ [dād-ø]$_V$
Maryam   to   brother-3.SG   OM   cake   gave-3.SG

"Maryam gave cake to his brother"

c. on direct objects:

[Maryam]$_S$ [keik]$_{DO}$ $\boxed{(\text{o})}$ [xord-ø]$_V$
Maryam   cake   OM   ate-3.SG

"Maryam ate (the) cake."

d. on nominal adverbs denoting duration or path:

---

[2]I use (*) to show that the sentence is ungrammatical with the object marker but grammatical without. I use *() to show that the sentence is ungrammatical without the object marker and grammatical with it. Finally, I use parentheses alone () to show that the sentence is grammatical with or without the object marker. In such cases the sentences may differ in meaning.

[3]I should note here that this sentence receives an interpretation but the one in which *Maryam* is the object: "The cake ate Maryam."

[Fardâ]$_{Adv}$ (ro) [Maryam]$_S$ [keik]$_{DO}$ [mi-xor-e]$_V$
Tomorrow $_{OM}$ Maryam   cake   IMP-eat-3.SG

"Tomorrow (all day), Maryam eats cake."

e. On extracted or left-dislocated objects:

Maryam$_i$ *(o) [[keik e **sh$_i$**]$_{DO}$ *(o) [xord-i]$_V$]$_{CP}$?
Maryam $_{OM}$ cake of him $_{OM}$ ate-2.SG

"As of Maryam, you ate his cake?"

As mentioned, semantics and pragmatics determine where the object marker is necessary. The occurrence of $r\bar{a}$ is obligatory on pronominal elements in Farsi: personal and demonstrative pronouns (17a), reflexive pronouns (17b), reciprocal pronouns (17c), and demonstrative nouns (17d). It is also required on superlatives (17e), question-words *kodum* "which" (17f) and *ki* "who" (17k), strong quantifiers such as *hame* "all" (17g), *bishtar* "most" (17h), *hardo* "both" (17i), and plurals with the plural marker $h\bar{a}$ (17j). I should add that $r\bar{a}$ also seems to be obligatory on *kas* which means "person".

(17)   a. on personal/demonstrative pronouns:

[Amir]$_S$ [un]$_{DO}$ *(o) [mi-shenās-e]$_V$
Amir   that $_{OM}$ IMP-know-3.SG

"Amir knows him."

b. on reflexive pronouns:

[Amir]$_S$ [xod-esh]$_{DO}$ *(o) [mi-shnās-e]$_V$
Amir   self-3.SG $_{OM}$ IMP-know-3.SG

"Amir knows himself."

c. on reciprocal pronouns:

15

[dāneshju-hā]$_S$ [hamdige]$_{DO}$ | *(ro) | [mi-shnās-an]$_V$
students-PL   eachother   OM   IMP-know-3.PL

"The students know each other"

d. on demonstrative nouns:

[Tāhā]$_S$ [un   keik]$_{DO}$ | *(o) | [did-ø]$_V$
Taha   that cake   OM   saw-3.SG

"Taha saw that cake."

e. on superlatives :

[Amir]$_S$ [behtarin ketāb]$_{DO}$ | *(o) | [xar-id]$_V$
Amin   best   book   OM   bought-3.SG

"Amin wants to read the best book."

f. on *kodum* "which":

[Sara]$_S$ [kodum keik]$_{DO}$ | *(o) | [xord-ø]$_V$?
Sara   which   cake   OM   ate-3.SG

"Which cake did Sara eat?"

g. on *hame* "all":

[Ali]$_S$ [hame ye   ketāb]$_{DO}$ | *(o) | [xund-ø]$_V$
Ali   all   IZAFE book   OM   read--3.SG

"Ali read all the book."

h. on *bishtar* "most":

[Ali]$_S$ [bisthar e   ketāb]$_{DO}$ | *(o) | [xund-ø]$_V$
Ali   more   IZAFE book   OM   read--3.SG

"Ali read most of the book."

i. on *har do* "both":

[Ali]$_S$ [har do ketāb]$_{DO}$ |*(ro)| [xund-ø]$_V$
Ali   each two book   OM   read--3.SG

"Ali read both books."

j. on plurals with the plural marker *hā*:

[Sara]$_S$ [keik-hā]$_{DO}$ |*(ro)| [xord-ø]$_V$
Sara   cake-PL   OM   ate-3.SG

"Sara ate the cakes."

k. on *who*:

[Ahmad]$_S$ [ki]$_{DO}$ |*(ro)| [did-ø]$_V$?
Ahmad   who   OM   saw-3.SG

"Who did Ahmad see?"

l. on *kas* "person":

    i. Mortezā ye kas-i |ro| dust   dār-e.
      Mortezā one person-*i*   OM friend have.PRES-3.SG

      Mortezā likes someone.

    ii. Mortezā hich kas-i |ro| dust   na-dār-e.
      Mortezā no   person-*i*   OM friend NEG-have.PRES-3.SG

      Mortezā doesn't like anyone.

Rā can also occur obligatorily or optionally on generics as the following examples show. In (18a) above, *rā* is obligatory while in (18b) it is optional. These sentences have a non-generic reading as well. As far as I can see, examples such as (18a) where *rā* is obligatory with a generic reading are rather rare. It is often the case that when the object NP has a generic reading, *rā* is optional.

(18)   a. [Serke]$_S$ [shir]$_{DO}$ |*(o)| [mi-bor-e]$_V$
         vinegar milk   OM   IMP-curdle-3.SG

"Vinegar curdles milk."

b. [oqāb]$_S$ [mush]$_{DO}$ (o) [shekār mi-kon-e]$_V$
   eagle   mouse   OM   hunt   IMP-do-3.SG

"Eagles hunt mice."

Those familiar with the literature on the object marker $r\bar{a}$ may have noticed that when I listed the environments where $r\bar{a}$ seems obligatory, I left out proper names. After all, examples like (19), suggest that $r\bar{a}$ is obligatory on proper names too.

(19)   [Amir]$_S$ [Barack Obama]$_{DO}$ ro [mi-shnās-e]$_V$
       Amir   Barack Obama   OM IMP-know-3.SG

"Amir knows Barack Obama."

However, while $r\bar{a}$ is obligatory on proper names in most contexts, there are context in which it does not appear on proper names. Such contexts shed light on the semantic contribution of the object marker and count as crucial data for learning its meaning. Consider the examples and contexts in (20) below. In (20a), the speaker does not presuppose that there is anyone with the name "Ali Saburi" when he asks the question. In other words, the question does not entail that there is anyone with that name in the discourse context. In fact, the point of asking the question is to inquire if such a person exists.

However, the opposite is true in (20b). The speaker knows for sure that there is someone called "Ali Saburi" and he does not consider that fact up for negotiation. In other words, he presupposes the existence of someone named "Ali Saburi", and simply asks if the addressee knows that person or not. Notice that the only difference in the form of the sentences in (20a) and (20b) is the presence or absence of the object marker $r\bar{a}$. Therefore, it is likely that $r\bar{a}$ is the culprit here in introducing the presupposition of existence into the discourse.

(20)   a. [Context: Hasan received a spam-like email from someone named Ali Saburi who claimed is an acquaintance of Reza. He is not sure if Reza knows anyone with this name. He asks Reza:]

18

Ali (e)   Saburi mi-shnās-i?
Ali (EZ)   Saburi MI-know-2.SG
"Do you know anyone named Ali Saburi?"

b. [Context: Ali Saburi is a famous Iranian singer. Hasan wants to know whether Reza knows

him. He asks Reza:]

Ali (e)   Saburi-[ro] mi-shnās-i?
Ali (EZ)   Saburi-OM MI-know-2.SG
"Do you know Ali Saburi?"

Note that a parallel distinction can be made in English using the indefinite determiner as
(21) shows[4]. In English, proper names such as Robert Moore in (21b) presuppose a unique
reference by default. However, the addition of the indefinite determiner in (21a) coerces the
proper name to drop its presuppositional status and instead, makes the existence of such an
individual the at-issue content of the question. Therefore, English takes the presuppositional
status of proper names as default and only marks them if this is not the case. On the other
hand, in the object position in Farsi, proper names are not presuppositional by default and
are only made so using the object marker.

(21)   a. Do you know a Robert Moore?

        b. Do you know Robert Moore?

We can test this presupposition of existence by explicitly denying it in a followup sentence
and see if we derive a contradiction. Consider the examples in (22). In (22a), the first
statement denies that there is any "work" to be done by Ali in the context of the utterance.
The following statement (after *pas* "so") explains that Ali did not do any work which is
consistent with him not having had any work to do! In example (22b) all we have done is
add *rā* to the nominal *kār* in the followup statement. This results in a contradiction. The
first statement of (22b) denies that there is any work, but it somehow seems like the second

---

[4]I would like to thank the anonymous reviewer for suggesting the addition of this point as well as providing
the example in (21).

statement does insist that there was work to do, but Ali did not do them. This is what we expected if *rā* contributed a presupposition of existence.

(22) a. Ali emruz kār-i na-dāsht, pas kār-i anjām na-dād.
     Ali today work-IC NEG-have.PST so work-IC finish NEG-give.3.SG
     "Today Ali didn't have anything to do so he didn't do anything."

    b. # Ali emruz kār-i na-dāsht, pas kār-i-[ro] anjām na-dād.
      Ali today work-IC NEG-have.PST so work-IC-OM finish NEG-give.3.SG

Another prediction is that if we change the first clause to assert that there is work to do for Ali, then using the object marker should not result in any contradiction. This is what (23) shows bellow. (22) and (23) together provide evidence that *rā* implies that the nominal it modifies is instantiated (exists) in the utterance context.

(23) Ali emruz xeyli kār dāsht vali kār-i-[ro] anjām na-dād
   Ali today very work have.PST but work-IC-OM finish NEG-give.3.SG
   "Ali had a lot of work to do but he didn't do any of them."

To provide further evidence for the semantic contribution of *rā* as an existential presupposition, consider the examples in (24) below. In (24a) the quantificational nominal *hichchiz-i* "nothing" appears without the object marker while in (24b), it appears with it. The sentences are do not convey the same meaning and have a subtle difference. The one in (24a) does not comment on whether there were things to buy for Ali or not. However, the one in (24b) does convey that there were things to buy. One way to translate this intuition into English is to use the partitive: "Ali bought none of them." I should add that partitives do not do justice to what *rā* contributes here. The meaning is a lot subtler and harder to translate. In Jasbi (peara), I provide a more formal and compositional account of *rā* in Informal Farsi.

(24) a. Ali hich chiz-i na-xar-id
     Ali no   thing-IC NEG-buy.PST-3.SG
     "Ali bought nothing."

    b. Ali hich chiz-i ro na-xar-id
     Ali no   thing-IC OM NEG-buy.PST-3.SG

"Ali bought none of them."

7   Discussion

I discussed several functional elements in Farsi and provided examples that shed light on their abstract and subtle meanings. Research on the meaning of function words in Farsi has only scratched the surface so far. There are numerous functional elements whose meanings are poorly understood, yet play a pivotal role in day-to-day conversations. Advances in theoretical and formal semantics can discover the meanings of function words and provide the crucial examples that illustrate their main functions in Farsi. These discoveries can in turn inform research and practice in language acquisition. Providing the crucial data for learning as well as communicating accurate generalizations on the meaning of functional elements can substantially boost and facilitate learning for learners of Farsi as a second language.

8   Glossing Abbreviations

1	First Person	2	Second Person	3	Third Person
EZ	Ezafe Marker	IC	Indefinite Clitic	ID	Indefinite Determiner
UM	Uniqueness Marker	NEG	Negation	OM	Object Marker
PL	Plural	PST	Past Tense	SG	Singular

References

Clark, E. V. (2009). *First language acquisition.* Cambridge University Press.

Gebhardt, L. (2009). *Numeral classifiers and the structure of DP.* PhD thesis, Northwestern University.

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

Giannakidou, A. (1998). *Polarity sensitivity as (non) veridical dependency*, volume 23. John Benjamins Publishing.

Jasbi, M. (2016). Three types of indefinites in Persian: Simple, complex, and antidefinite. In Mary Moroney, Carol-Rose Little, D. B. and Collard, J., editors, *Proceedings of Semantics and Linguistic Theory 26*, pages 244–263. Ithaca, NY: CLC Publications.

Jasbi, M. (to appeara). The meaning of the farsi object marker ra: What it is not, and what it (probably) is. that makes farsi nouns unique. *Advances in Iranian Linguistics*.

Jasbi, M. (to appearb). The suffix that makes farsi nouns unique. *Advances in Iranian Linguistics*.

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

Ladusaw, W. A. (1980). On the notion affective in the analysis of negative-polarity items. *Formal semantics: The essential readings*, pages 457–470.

LeGrand, J. E. (1975). *"or" and "any": The semantics and syntax of two logical operators*. PhD thesis, University of Chicago, Chicago.

Moin, M. (1958). *Mofrad va Jam, Ma'refeh va Nakareh; Tarh-e Dastur-e Zabān-e Fārsi [Singular and Plural, Definite, and Indefinite; The Design of Persian Grammar]*, volume 5. Tehran University Press.

Natel-Khanlari, P. (1972). *Dastur-e Zabān-e Fārsi [The Grammar of Persian]*. Bonyād-e Farhang-e Iran.