

## Research article

Jamie Rinder\*

# Marking the specific direct object in restrictive relative clauses in spoken Farsi

<https://doi.org/10.33063/diva-454313>

**Abstract:** This study is an empirical investigation of both the distribution of the specific direct object marker *rā* in restrictive relative clauses in spoken Farsi and the degrees of acceptance of different patterns of this structure among 50 Farsi speakers. Examples of the target structure are taken from modern films and a spoken translation test and then used in an acceptability judgement test. The findings show that there is a considerable disparity between the written relative clauses studied in previous research and spoken versions of the same structure. More specifically, the findings in this study show that the omission of the *rā* marker is common, that the use of *rā*-marked referent pronouns no longer seems to be an acceptable feature of spoken Farsi, and that the use of the *rā* marker and the judgement of its acceptability change according to the syntactic function(s) of the head noun phrase.

**Keywords:** object marking, relative clause, Farsi, spoken language, acceptability judgement test

## 1 Introduction

In Farsi<sup>1</sup>, syntactic relations between noun phrases can be conveyed by means of several prepositions and one postposition, *rā*<sup>2</sup>. *Rā* marks the specific direct object: in theory, the more specific the direct object noun phrase, the more likely it is to be *rā* marked:

1 *man diruz ketāb xund-am*<sup>3</sup>  
 PN.1SG yesterday book read.PST-1SG<sup>4</sup>  
 I read yesterday / I did some (book) reading yesterday.

2 *man diruz ketāb=o<sup>5</sup> xund-am*  
 PN.1SG yesterday book=OM read.PST-1SG  
 I read the book yesterday.

Both examples 1 and 2 contain a subject (*man* ‘I’), a verb (*xundam* ‘read’), and a direct object (*ketāb* ‘book’), but only the *rā* marker (=o) in example 2 lends specificity to the direct object.

In sentences with relative clauses, the head noun phrase can have more than one syntactic function. For instance, the head noun phrase *ketāb* in example 3 functions both as the specific direct object of the verb in the relative clause (*xaridam* ‘I bought’) and as the subject of the complex predicate (*gerun bud* ‘was expensive’) in the main clause:

- 1 The variation of Modern New Persian spoken in Iran and the Iranian diaspora.
- 2 [ra] is the underlying phonemic representation for *rā*. However, it is conventionally pronounced [ro] after most vowel sounds and [o] after most consonant sounds (Jam 2017: 2). It is always unstressed.
- 3 Examples 1–6 were suggested by Farsi speakers and judged as correct or acceptable in a pilot study.
- 4 For glossing conventions, see the list of abbreviations at the end of this paper.
- 5 All examples in this paper represent spoken language and the transcriptions therefore include typical phonological transformations, for example ‘ro’ or ‘o’ for *rā* and ‘un’ for *ān*. The Farsi letters *te* and *tā* are represented by ‘t’; the letters *se*, *sin*, and *sād* are represented by ‘s’; the letters *zāl*, *ze*, *zād*, and *zā* are represented by ‘z’. Glottal stops are not marked.

\* Unit of Language and Communication, KTH Royal Institute of Technology, Stockholm, E-mail: jamier@kth.se

- 3 *ketāb=i*      *ke*      *dišab*      *xarid-am*      *gerun*      *bud*  
 book=IND      CLM      last.night      buy.PST-1SG      expensive      be.PST.3SG  
 The book that I bought last night was expensive.

The dual syntactic function of the head noun phrase potentially complicates *rā* marking (Elwell-Sutton 1963: 113–114), and this complication is approached in a number of different ways in grammars and textbooks. Lazard (1957: 222), for example, seems to advocate the use of referent pronouns<sup>6</sup> within the relative clause when the head noun phrase functions as a specific direct object in the relative clause:

- 4 *ketāb=i*      *ke*      *un=o*      *dišab*      *xarid-am*      *gerun*      *bud*  
 book=IND      CLM      DEM.DIST=OM      last.night      buy.PST-1SG      expensive      be.PST.3SG  
 The book that I bought last night was expensive.

In Thackston (1993: 86–87), *rā* markers are commonly applied immediately after the head noun phrase:

- 5 *ketāb=i=ro*      *ke*      *dišab*      *xarid-am*      *gerun*      *bud*  
 book=IND=OM      CLM      last.night      buy.PST-1SG      expensive      be.PST.3SG  
 The book that I bought last night was expensive.

According to Lambton (1974: 76), the *rā* marker in such a sentence is optional; according to Dahlén (2010: 343) it is incorrect. In Rafiee (1988: 344), no *rā* markers are used in such relative clauses.

In Dahlén (2012: 200), the *rā* marker is sometimes applied after the relative clause:

- 6 *ketāb=i*      *ke*      *xarid-am=o*      *xund-am*  
 book=IND      CLM      buy.PST-1SG=OM      read.PST-1SG  
 I read the book that I bought.

This pattern is increasingly common among Farsi speakers, but it has been considered “much less acceptable” and “a recent innovation employed in mass media and by younger speakers” in a study on syntactic theory (Karimi 2001: 63–64).

In previous descriptions of relative clauses in Farsi, the focus has been on the written language (Abdollahnejad and Marefat 2017) and on the *i* enclitic as it is applied to the head noun phrase (Jahani 2000). No study has dealt with the use and position of the *rā* marker in spoken Farsi and no empirical data for its distribution has been provided. The purpose of this paper is therefore to investigate the distribution of the *rā* marker in restrictive relative clauses in spoken Farsi and the degrees of acceptance of the *rā* marker’s varied use and position among Farsi speakers.

The target structure for this study is sentences with a restrictive relative clause and a head noun phrase that functions as the specific direct object in at least one clause, thereby legitimising a *rā* marker. Examples of the target structure are taken from ten modern Iranian films and a spoken translation test taken by 50 fluent Farsi speakers. Examples from both the films and the translation tests are then used in an acceptability judgement test by the same Farsi speakers to gauge the perceived correctness of different uses and positions of the *rā* marker.

<sup>6</sup> Also called resumptive pronouns, these are personal pronouns or demonstrative pronouns within the relative clause.

## 2 Method

### 2.1 Gathering data from modern Iranian films

Ten modern Iranian films<sup>7</sup> were chosen that contained realistic plots and natural dialogue. They were viewed repeatedly with the aim of noting all the examples of the target structure. The examples were checked by several Farsi-speaking friends and colleagues, added to a spreadsheet, and classified according to the following criteria:

- i. the head noun phrase's syntactic function in both the relative clause and the main clause
- ii. the presence and position of the *rā* marker
- iii. the context of the speech (gender and social positions of the speaker and listener)<sup>8</sup>

### 2.2 Gathering more data from an English-to-Farsi spoken translation test

In order to gain more examples of the target structure for this study, an English-to-Farsi spoken translation test was devised. The results provided a step towards ensuring the validity of the examples gathered in the film corpus.

The translation test consisted of 17 sentences inspired by examples in Jasbi (2014) and discussions with Farsi-speaking friends and colleagues. The sentences were presented to the test participants in a jumbled order. The target structure was required in six sentences. Sentences i and ii were designed to elicit restrictive relative clauses in Farsi whose head noun phrase functions as the specific direct object in both the relative clause and the main clause:

- i. I don't know the student you chose.
- ii. I didn't like the apple I ate yesterday.

Sentences iii and iv were designed to yield head noun phrases that function as the specific direct object in the relative clause and as the subject in the main clause:

- iii. Where is the photograph you sold?
- iv. The coffee he drank was cold.

The head noun phrase of sentence v should function as the specific direct object in the main clause and as the subject in the relative clause:

- v. I saw the children who had just arrived.

The head noun phrase in sentence vi should function as the specific direct object of the main clause and as the prepositional object of the relative clause:

- vi. She saw the teacher she was afraid of.

The other sentences were distractors. In translations of sentences vii, ix, and xv, the participants' use of colloquial language was gauged. In translations of the other sentences, the use of the *rā* marker was tested in sentences without relative clauses.

- vii. Reza went home.
- viii. Amir knows her.

<sup>7</sup> See the list of films at the end of this paper.

<sup>8</sup> Upon later analysis, no connection between context and the distribution of the *rā* marker was found.

- ix. If you go to Tehran, call me.
- x. Taha saw that film.
- xi. Where is the vinegar?
- xii. Which cake did Sara eat?
- xiii. Ali read both newspapers.
- xiv. Farzad wants a computer.
- xv. What's your telephone number?
- xvi. Eagles hunt mice.
- xvii. Majid wants a wife.

### 2.3 Testing Farsi speakers' judgements

In order to gauge the degrees of acceptance among the Farsi speakers of the *rā* marker's varied use and position, an acceptability judgement test was devised. The test comprised two parts:

- 1) Eight questions to determine each participant's approximate age and lingua-cultural background. Most of the questions were inspired by Skutnabb-Kangas (2008) and required a short answer:

- Which age category are you in? 20–30, 31–40, 41–50, 51–60, 61–70, 71–80.
- Which language did you learn first?
- Which language do you identify most with today?
- Which language do you know best?
- Which language do you use most today?
- Do you listen to Farsi-speaking media at least once per week?
- Do you live in Iran? If not, when did you stop living there?
- Other than Farsi, which other languages can you speak with ease?

2) An acceptability judgement exercise with 22 sentences (listed below) inspired by examples from the film corpus and/or devised in cooperation with Farsi-speaking friends and colleagues. The sentences were pre-recorded by a native speaker, spoken in a natural manner, and presented to the test participants in a jumbled order.

Sentences i to v are variations of a line taken from Sabbaghzade (2017) meaning *Did you sell the jumper I bought you for your birthday?* The head noun phrase, *puliver*, is the direct object in both the main clause and the relative clause. In sentence i, there is no *rā* marking; in sentence ii, the *rā* marker is pre-posed<sup>9</sup>; in sentence iii, the *rā* marker is post-posed<sup>10</sup>; in sentence iv, the enclitic *eš* is used to refer back to the head noun phrase; in sentence v, the relative clause is extraposed.

i	<i>puliver=i</i> jumper=IND	<i>ke</i> CLM	<i>man</i> PN.1SG	<i>kādo=ye</i> gift=EZ	<i>tavalod</i> birthday	<i>barā=t</i> for=PC.2SG	<i>xarid-am</i> buy.PST-1SG	<i>foruxt-i?</i> sell.PST-2SG
ii	<i>puliver=i=ro</i> jumper=IND=OM	<i>ke</i> CLM	<i>man</i> PN.1SG	<i>kādo=ye</i> gift=EZ	<i>tavalod</i> birthday	<i>barā=t</i> for=PC.2SG	<i>xarid-am</i> buy.PST-1SG	<i>foruxt-i?</i> sell.PST-2SG
iii	<i>puliver=i</i> jumper=IND	<i>ke</i> CLM	<i>man</i> PN.1SG	<i>kādo=ye</i> gift=EZ	<i>tavalod</i> birthday	<i>barā=t</i> for=PC.2SG	<i>xarid-am=o</i> buy.PST-1SG=OM	<i>foruxt-i?</i> sell.PST-2SG
iv	<i>puliver=i</i> jumper=IND	<i>ke</i> CLM	<i>man</i> PN.1SG	<i>kādo=ye</i> gift=EZ	<i>tavalod</i> birthday	<i>barā=t</i> for=PC.2SG	<i>xarid-am=eš</i> buy.PST-1SG=PC.3SG	<i>foruxt-i?</i> sell.PST-2SG

<sup>9</sup> The *rā* marker is placed before the relative clause.

<sup>10</sup> The *rā* marker is placed after the relative clause.

v	<i>puliver=i-ro</i> jumper=IND=OM	<i>foruxt-i</i> sell.PST-2SG	<i>ke</i> CLM	<i>man</i> PN.1SG	<i>kādo=ye</i> gift=EZ	<i>tavalod</i> birthday	<i>barā=t</i> for=PC.2SG	<i>xarid-am?</i> buy.PST-1SG
---	--------------------------------------	---------------------------------	------------------	----------------------	---------------------------	----------------------------	-----------------------------	---------------------------------

Sentences vi to x mean *I don't know the student you chose* and they are variations of a sentence I originally used with colleagues to illustrate my research idea. Again, the head noun phrase *šāgerd* is the specific direct object in both the main clause and the relative clause. In sentence vi, there is no *rā* marking; in sentence vii, the *rā* marker is pre-posed; in sentence viii, the *rā* marker is post-posed; in sentence ix, the relative clause is extraposed; in sentence x, a referent pronoun is used.

vi	<i>šāgerd=i</i> student=IND	<i>ke</i> CLM	<i>entexāb</i> choice	<i>kard-i</i> do.PST-2SG	<i>ne-mi-šenās-am</i> NEG-IMP-know.PRS-1SG	
vii	<i>šāgerd=i-ro</i> student=IND=OM	<i>ke</i> CLM	<i>entexāb</i> choice	<i>kard-i</i> do.PST-2SG	<i>ne-mi-šenās-am</i> NEG-IMP-know.PRS-1SG	
viii	<i>šāgerd=i</i> student=IND	<i>ke</i> CLM	<i>entexāb</i> choice	<i>kard-i-ro</i> do.PST-2SG=OM	<i>ne-mi-šenās-am</i> NEG-IMP-know.PRS-1SG	
ix	<i>šāgerd=i</i> student=IND	<i>ne-mi-šenās-am</i> NEG-IMP-know.PRS-1SG	<i>ke</i> CLM	<i>entexāb</i> choice	<i>kard-i</i> do.PST-2SG	
x	<i>šāgerd=i</i> student=IND	<i>ke</i> CLM	<i>un=o</i> DEM.DIST=OM	<i>entexāb</i> choice	<i>kard-i</i> do.PST-2SG	<i>ne-mi-šenās-am</i> NEG-IMP-know.PRS-1SG

Sentences xi to xiii are variations of a line taken from Kamyar (2017) meaning *I didn't eat the cold food left over from yesterday*. The head noun phrase *gāzāye sard* functions as the direct object in the main clause and as the subject in the relative clause. In sentence xi, the *rā* marker is post-posed and in sentence xii it is pre-posed. Sentence xiii is unmarked.

xi	<i>gāza=ye</i> food=EZ	<i>sard=i</i> cold=IND	<i>ke</i> CLM	<i>az</i> from	<i>diruz</i> yesterday	<i>mund-e-ro</i> remain.PST-PP=OM	<i>na-xord-am</i> NEG-eat.PST-1SG
xii	<i>gāza=ye</i> food=EZ	<i>sard=i-ro</i> cold=IND=OM	<i>ke</i> CLM	<i>az</i> from	<i>diruz</i> yesterday	<i>mund-e</i> remain.PST-PP	<i>na-xord-am</i> NEG-eat.PST-1SG
xiii	<i>gāza=ye</i> food=EZ	<i>sard=i</i> cold=IND	<i>ke</i> CLM	<i>az</i> from	<i>diruz</i> yesterday	<i>mund-e</i> remain.PST-PP	<i>na-xord-am</i> NEG-eat.PST-1SG

Sentences xiv and xv are variations of a line taken from Moghadam (2012) meaning *You added so much opiate to the bread that it killed a girl*. The head noun phrase *moxader* functions as the direct object in the relative clause and as the subject in the main clause. Sentence xiv has a pre-posed *rā* marker and sentence xv is unmarked.

xiv	<i>moxader=i-ro</i> opiate=IND=OM	<i>ke</i> CLM	<i>be</i> to	<i>nun=eš</i> bread=PC.3SG	<i>zad-i</i> put.PST-2SG	<i>enqadr</i> so	<i>ziād</i> much	<i>bud</i> be.PST.3SG
xv	<i>moxader=i</i> opiate=IND	<i>ke</i> CLM	<i>be</i> to	<i>nun=eš</i> bread=PC.3SG	<i>zad-i</i> put.PST-2SG	<i>enqadr</i> so	<i>ziād</i> much	<i>bud</i> be.PST.3SG
	<i>ke</i> CLM	<i>ye</i> one	<i>doxtar=o</i> girl=OM	<i>košt</i> kill.PST.3SG				

The test included a number of distractors, which were also used in the translation test. These distractors either presented common features of spoken Farsi (sentences xvi and xvii) or tested the par-

participants' sensitivity to the *rā* marker in sentences without relative clauses (sentences xviii, xix, xx, and xxii). Sentence xxi includes an error.

- xvi *age* *tehrān* *mi-r-i* *zang=am* *be-zan*  
 if Tehran IMP-go.PRS-2SG bell=PC.1SG SBJV-hit.2SG  
 If you go to Tehran, call me.
- xvii *rezā* *raft=eš* *xune*  
 Reza go.PST.3SG=PC.3SG home  
 Reza went home.
- xviii *amir* *un=o* *mi-šenās-e*  
 Amir DEM.DIST=OM IMP-know.PRS-3SG  
 Amir knows her.
- xix *tāhā* *un* *film=o* *did*  
 Taha DEM.DIST film=OM see.PST.3SG  
 Taha saw that film.
- xx *serke* *kojā=st?*  
 vinegar where=COP.PRS.3SG  
 Where's the vinegar?
- xxi *sārā* *kodum* *keik=i* *xord?*  
 Sara which cake=IND eat.PST.3SG  
 (Error) Lit. Which a cake did Sara eat?
- xxii *ali* *har* *do* *ruznāme=ro* *xund*  
 Ali each two newspaper=OM read.PST.3SG  
 Ali read both newspapers.

The test participants were asked to judge each sentence after the second hearing using a three-point Likert scale: *Correct*, *Acceptable*, or *Incorrect*.

## 2.4 Conducting the tests

After a short round of pilot tests with friends and colleagues, potential test participants were invited by e-mail to take part in the study. In the invitation, it was made clear that the study was about spoken Farsi, that the test would take a maximum of one hour, and that it would take place online using video conferencing software. The participants' rights were also clarified<sup>11</sup>. All tests were conducted in April and May 2020.

At the beginning of the online meeting and after greetings and formalities, the participants were reminded of their rights and of the fact that their fluency in Farsi was not being tested. If the participants were happy, the recording was started and I began with the questions regarding each participant's lingua-cultural background. If participants struggled with these questions, I rephrased them. I spoke in English during the formal parts of the test.

Thereafter followed the translation test. After reading out each sentence in English, I asked the participants to translate orally using their everyday, naturally spoken Farsi. A few participants asked if they could write the sentences down, and I made clear that writing was unnecessary.

<sup>11</sup> The participants could refuse to take part or refuse to continue to take part in the study at any time. No personal details such as name or exact age would be noted. All responses would be anonymised in the written report. The test would be recorded but not played to anyone outside my seminar group. Participants could request a copy of the recording.

In the final part of the test, the participants were informed that they would hear a native Farsi speaker read out some sentences twice. They were asked to judge the sentences after the second hearing using a three-point Likert scale (*Correct*, *Acceptable*, or *Incorrect*). They were reminded that the test was not about pronunciation, intonation, or the likelihood of them using the same vocabulary as the speaker in the recording, but rather about the correctness of the utterances. The test was concluded with some formalities and thanks.

After each recording, the participants' spoken translations were transcribed and their correct / acceptable / incorrect judgements and any additional comments were added to a spreadsheet. Later, the judgements were given a numerical value (10, 5, and 0 respectively) to facilitate the calculation of mean average judgement scores.

## 2.5 Description of the fifty participants

Half the participants were male and half were female. Their ages ranged from 20s to 70s (19 participants aged 20–29; 23 aged 30–39; 2 aged 40–49; 2 aged 50–59; 3 aged 60–69; 1 aged 70–79).

All participants were confident Farsi speakers. First languages were Farsi (43), Azeri (6), and Arabic (1). All participants had completed at least their secondary education in Iran.

Thirty-three of the participants lived in Iran or had left the country in the last five years. Of the 17 participants who had lived outside Iran for more than five years, 15 lived in Sweden, 1 in Italy, and 1 in Germany.

## 2.6 Limitations of the data-gathering methods

An ideal source for this investigation would be naturally occurring, non-elicited, everyday speech by monolingual Farsi speakers. In the absence of this kind of material, films and translation sentences were chosen as sources for the target structure. Speech in film is to a greater degree scripted; however, care was taken to choose films with realistic plots and natural dialogue. The translation test was a way of generating more examples and validating the results of the film corpus. Ultimately, however, these sources are secondary in quality to naturally occurring speech.

A second important limitation of this study was the narrow range of the participants' backgrounds. For instance, only eight of the participants were over the age of 40. All of the 50 participants were well educated and at least bilingual, and many of them had an interest in language. The participants' high level of education should be taken into consideration when evaluating the results.

Finally, a comment about the Likert scale and the scope of this study must be made. The three points (*Correct*, *Acceptable*, or *Incorrect*) forced the participants to make rapid judgements about complex linguistic questions, such as the possibility for non-standard language to be correct and the precise boundary between acceptable and incorrect language. It is most likely that the participants had different views about language correctness and acceptability, but an analysis of such issues goes beyond the scope of this study.

## 3 Results

The aims of this study are, first, to investigate the distribution of the *rā* marker in restrictive relative clauses in spoken Farsi and, second, to explore the degrees of acceptance among Farsi speakers of different patterns of restrictive relative clauses where the use of the *rā* marker is legitimate. As described in the method section, examples of the target structure were taken from ten modern Iranian films and an English-to-Farsi spoken translation test. These examples were then used in an acceptability judgement test, where they were judged as correct, acceptable, or incorrect.

The examples of the target structure are classified as follows:

- Sentences where the head noun phrase functions as the specific direct object in both the main clause and the relative clause (*DO in both clauses*).
- Sentences where the head noun phrase functions as the specific direct object in the main clause (*DO in the MC*). The head noun phrase has a different syntactic function in the relative clause.
- Sentences where the head noun phrase functions as the specific direct object in the relative clause (*DO in the RC*). The head noun phrase has a different syntactic function in the main clause.

As explained in the introduction to this paper, the *rā* marker can be applied in the target structure in the following four ways:

i) The *rā* marker can be applied before the relative clause. This position of the *rā* marker is possible for all examples of the target structure.

7	<i>puliver=i-ro</i>	<i>ke</i>	<i>man</i>	<i>barā=t</i>	<i>xarid-am</i>	<i>foruxt-i?</i>
	jumper=IND=OM	CLM	PN.1SG	for=PC.2SG	buy.PST-1SG	sell.PST-2SG
	Did you sell the jumper I bought you for your birthday?					

ii) If the head noun functions as the specific direct object in the relative clause, the *rā* marker can be applied to a referent pronoun in the relative clause:

8	<i>puliver=i</i>	<i>ke</i>	<i>man</i>	<i>un=o</i>	<i>barā=t</i>	<i>xarid-am</i>	<i>foruxt-i?</i>
	jumper=IND	CLM	PN.1SG	DEM.DIST=OM	for=PC.2SG	buy.PST-1SG	sell.PST-2SG
	Did you sell the jumper I bought you for your birthday?						

iii) The *rā* marker can be applied after the relative clause if the head noun phrase functions as the specific direct object in the main clause:

9	<i>puliver=i</i>	<i>ke</i>	<i>man</i>	<i>barā=t</i>	<i>xarid-am=o</i>	<i>foruxt-i?</i>
	jumper=IND	CLM	PN.1SG	for=PC.2SG	buy.PST-1SG=OM	sell.PST-2SG
	Did you sell the jumper I bought you for your birthday?					

iv) The *rā* marker can be omitted and the head noun phrase is said to be unmarked. The omission of the *rā* marker is possible for all examples of the target structure.

10	<i>puliver=i</i>	<i>ke</i>	<i>man</i>	<i>barā=t</i>	<i>xarid-am</i>	<i>foruxt-i?</i>
	jumper=IND	CLM	PN.1SG	for=PC.2SG	buy.PST-1SG	sell.PST-2SG
	Did you sell the jumper I bought you for your birthday?					

### 3.1 Distribution of the *rā* marker in the target structure in the film and spoken translation corpora

#### DO in both clauses

When the head noun phrase functions as the specific direct object in both the main clause and the relative clause, the *rā* marker can be omitted:

11	<i>avalin</i>	<i>kādo=ye</i>	<i>tavalod=i</i>	<i>ke</i>	<i>barfā</i>	<i>barā=m</i>	<i>gereft-e</i>	<i>bud</i>
	first	gift=EZ	birthday=IND	CLM	Barfa	for=PC.1SG	get.PST-PP	be.PST.3SG
	<i>foruxt-e</i>	<i>bud-am</i>						
	sell.PST-PP	be.PST-1SG						
	I had sold the first birthday gift that Barfa had got me. (Sabbaghzade 2017: 01:36:39)							



12 *tavaqo* *dār-in* *man* *bāvar* *kon-am* *in* *čiz=i*  
 expectation have.PRS-2PL PN.1SG belief SBJV.do.PRS-1SG DEM.PROX thing=IND  
*ke* *dār-in* *mi-g-in?*  
 CLM have.PRS-2PL IMP-say.PRS-2PL  
 Do you expect me to believe what you're saying? (Khajooei 2008: 00:39:55)

13 *dānešju=i* *ke* *entexāb* *kard-i* *ne-mi-šenās-am*  
 student=IND CLM choice do.PST-2SG NEG-IMP-know.PRS-1SG  
 I don't know the student you chose. (Participant 35)

The *rā* marker can also be applied. In both corpora, post-posed *rā* marking (examples 14, 15, and 16) is more common than pre-posed (examples 17 and 18):

14 *mā* *eddeā=i* *ke* *dār-in=o* *ne-mi-tun-im* *be-fahm-im*  
 PN.1PL claim=IND CLM have.PRS-2PL=OM NEG-IMP-can.PRS-1PL SBJV-understand.PRS-1PL  
 We can't understand the claims that you make. (Seyyedi 2016: 01:20:00)

15 *dānešāmuz=i* *ke* *entexāb* *kard-i=o* *ne-mi-šenās-am*  
 student=IND CLM choice do.PST-2SG=OM NEG-IMP-know.PRS-1SG  
 I don't know the student you chose. (Participant 32)

16 *engār* *dār-an* *tamum=e* *kārt-ā=ye* *sadāfarin* *va* *bist-ā=i*  
 imagination have.PRS-3PL all=EZ card-PL=EZ hundred.congratulation and twenty-PL=IND  
*ke* *tu* *bačcegi* *beh=em* *dād-e* *bud-an=o* *yek=i* *yek=i*  
 CLM in childhood to=PC.1SG give.PST-PP be.PST-3PL=OM one=IND one=IND  
*az=am* *pas* *mi-gir-an*  
 from=PC.1SG back IMP-take.PRS-3PL  
 It's as if they're taking back all the congratulations cards and top grades they had given me when I was a child. (Sadrameli 2010: 00:10:31)

17 *man* *dānešāmuz=i=ro* *ke* *entexāb* *kard-i* *ne-mi-šenās-am*  
 PN.1SG student=IND=OM CLM choice do.PST-2SG NEG-IMP-know.PRS-1SG  
 I don't know the student you chose. (Participant 3)

18 *nosrat* *bānu* *umad* *dastmozd=e* *lebās=i=ro* *ke* *duxt-e*  
 Nosrat lady came.PST.3SG wages=EZ clothes=IND=OM CLM sew.PST-PP  
*bud-am* *dād<sup>2</sup>*  
 be.PST-1SG give.PST.3SG  
 Miss Nosrat paid for the dress I made her. (Moghadam 2012: 00:26:57)

Table 1 shows the distribution of *rā* marking in 'DO in both clauses' sentences in the two corpora. As the results show, *rā* marking is more common in the translation corpus than in the film corpus, and both corpora show that post-posed *rā* marking is more common than pre-posed. There are no examples of referent pronouns in the data.

**Table 1:** Distribution of the *rā* marker in 'DO in both clauses' sentences from the film and spoken translation corpora

	Position of the <i>rā</i> marker			
	Before the RC	In the RC	After the RC	No <i>rā</i> marking
Film corpus (n=12)	2 (17%)	0	4 (33%)	6 (50%)
Translation corpus (n=89 <sup>13</sup> )	28 (31%)	0	47 (53%)	14 (16%)

<sup>12</sup> *dastmozd* is the direct object of the verb in the main clause (*dād*); *lebās* functions both as the genitive attribute in the main clause and the direct object of the verb in the relative clause (*duxt-e budam*).

<sup>13</sup> Eleven translations were considered spoilt due to misunderstandings of the original English, and they were therefore not counted.

**DO in the MC**

When the head noun phrase functions as the specific direct object in only the main clause, a *rā* marker can be applied. It can appear before the relative clause (examples 19, 20, and 21) or after the relative clause (examples 22 and 23):

- 19 *če gonāh=i kard-am ke bāy-ad kas=i=ro ke hame=ye omr*  
 what sin=IND do.PST-1SG CLM must.PRS-3SG person=IND=OM CLM all=EZ life  
*beh=eš del bast-am az dast be-d-am?*  
 to=PC.3SG heart attach.PST-1SG from hand SBJV-give.PRS-1SG  
 What sin did I commit to have to lose someone I've loved all my life? (Moghadam 2012: 01:09:44)
- 20 *hič surat o sedā=i=ro ne-mi-tun-i peydā kon-i*  
 no face and sound=IND=OM NEG-IMP-can.PRS-2SG evident SBJV.do.PRS-2SG  
*ke āšenā bāš-e*  
 CLM familiar SBJV.be-3SG  
 No familiar faces or sounds can be found. (Sadrameli 2010: 00:08:30)
- 21 *man bačče-hā=i=ro ke tāze resid-e bud-an did-am*  
 PN.1SG child-PL=IND=OM CLM just arrive.PST-PP be.PST-3PL see.PST-1SG  
 I saw the children who had just arrived. (Participant 36)
- 22 *bačče-hā=i ke tāze resid-e bud-an=o did-am*  
 child-PL=IND CLM just arrive.PST-PP be.PST-3PL=OM see.PST-1SG  
 I saw the children who had just arrived. (Participant 34)
- 23 *man ne-mi-tun-am do tā ādam=i ke be ye andāze*  
 PN.1SG NEG-IMP-can.PRS-1SG two CL person=IND CLM to one measure  
*maznun=an=o be-šun-am kenār=e ham*  
 suspected=COP.PRS.3PL=OM SBJV-seat.PRS-1SG side=EZ each.other  
 I can't seat two equally suspected people next to one another. (Seyyedi 2016: 01:06:58)

The *rā* marker can also be omitted:

- 24 *či šod ke sar az injā dar āvord-am o dār-am*  
 what happen.PST.3SG CLM head from here PREV bring.PST-1SG and have.PRS-1SG  
*kār=i mi-kon-am ke aslan tasavor=eš=o ne-mi-kard-am?*  
 job=IND IMP-do.PRS-1SG CLM at.all imagination=PC.3SG=OM NEG-IMP-do.PST-1SG  
 How did I end up here, doing a job I never at all imagined I'd do? (Khajooei 2008: 01:12:36)
- 25 *be man qarārdād=i nešun dād-an ke bast-e šod-e o*  
 to PN.1SG contract=IND show give.PST-3PL CLM close.PST-PP become.PST-PP and  
*sar tā pā eškāl=e*  
 head to foot problem=COP.PRS.3SG  
 They've shown me a contract that has already been signed and is full of mistakes. (Raisian 2019: 00:24:29)
- 26 *bačče-hā=i ke tāze umad-an did-am*  
 child-PL=IND CLM just come.PST-3PL see.PST-1SG  
 I saw the children who had just arrived. (Participant 48)

Table 2 shows that most of the 'DO in the MC' examples in the corpora are *rā* marked. Post-posed *rā* markers are more common in the translation corpus.

**Table 2:** Distribution of the *rā* marker in ‘DO in the MC’ sentences from the film and spoken translation corpora

	Position of the <i>rā</i> marker			
	Before the RC	In the RC	After the RC	No <i>rā</i> marking
Film corpus (n=7)	3 (43%)	-	3 (43%)	1 (14%)
Translation corpus (n=93)	26 (28%)	-	54 (58%)	13 (14%)

### DO in the RC

When the head noun phrase functions as the specific direct object in only the relative clause, the *rā* marker is almost always omitted:

27 *un* *āmpul=e* *tārix* *masraf* *gozašte=i* *ke* *beh=em* *dād-i*  
 DEM.DIST ampoule=EZ date consumption past=IND CLM to=PC.1SG give.PST-2SG  
*yek=i=ro* *košt*  
 one=IND=OM kill.PST.3SG  
 That expired ampoule you gave me killed someone. (Kamyar 2017: 00:58:44)

28 *čiz=e* *jozi* *mund-e* *ke* *išun* *peigiri* *mi-kon-e*  
 thing=EZ trivial remain.PST-PP CLM PN.3SG investigation IMP-do.PRS-3SG  
 There’s still a trivial matter that he’s following up on. (Raisian 2019: 00:49:00)

29 *aks=i* *ke* *foruxt-i* *kojā=st?*  
 photograph=IND CLM sell.PST-2SG where=COP.PRS.3SG  
 Where is the photograph you sold? (Participant 5)

In the spoken translation corpus, only one translation has a *rā*-marked head noun phrase:

30 *un* *aks=i=ro* *ke* *to* *foruxt-i* *kojā=st?*  
 DEM.DIST photograph=IND=OM CLM PN.2SG sell.PST-2SG where=COP.PRS.3SG  
 Where is the photograph you sold? (Participant 41)

Table 3 shows a clear pattern for the ‘DO in the RC’ sentences in both corpora. Note the continued absence of referent pronouns:

**Table 3:** Distribution of the *rā* marker in ‘DO in the RC’ sentences from the film and spoken translation corpora

	Position of the <i>rā</i> marker			
	Before the RC	In the RC	After the RC	No <i>rā</i> marking
Film corpus (n=13)	0	0	-	13 (100%)
Translation corpus (n=95)	1 (1%)	0	-	94 (99%)

## 3.2 Results of the acceptability judgement tests

### DO in both clauses

The results in tables 4 and 5 show that examples of the target structure whose head noun phrase functions as the specific direct object in both the main clause and the relative clause are considered correct with either a pre-posed *rā* marker or a post-posed *rā* marker. Unmarked head noun phrases are also considered correct, albeit less so. According to the mean judgement scores, the extraposed relative clauses are clearly judged as incorrect, as are anaphoric strategies such as the use of the referent pronoun *un* and the enclitic referent pronoun *eš*.

**Table 4:** Results of the acceptability judgement test of the sentence *I don't know the student you chose* (n=50). Mean judgement scores were calculated by giving the following numerical scores to the test participants' acceptability judgements: Correct = 10, Acceptable = 5, Incorrect = 0.

	Correct	Acceptable	Incorrect	Mean judgement / 10
<i>šāgerdiro ke entexāb kardi nemišenāsam</i> (pre-posed <i>rā</i> marker)	46	4	0	9.6
<i>šāgerdi ke entexāb kardi nemišenāsam</i> (post-posed <i>rā</i> marker)	44	3	3	9.1
<i>šāgerdi ke entexāb kardi nemišenāsam</i> (unmarked)	36	10	4	8.2
<i>šāgerdi ke uno entexāb kardi nemišenāsam</i> ( <i>rā</i> -marked referent pronoun)	5	10	35	2.0
<i>šāgerdi nemišenāsam ke entexāb kardi</i> (unmarked and extraposed)	0	5	45	0.5

**Table 5:** Results of the acceptability judgement test of the sentence *Did you sell the jumper I bought you for your birthday?* (n=50)

	Correct	Acceptable	Incorrect	Mean judgement / 10
<i>puliveri ke man kādoye tavalod barāt xaridamo foruxti?</i> (post-posed <i>rā</i> marker)	35	13	2	8.3
<i>puliveriro ke man kādoye tavalod barāt xaridam foruxti?</i> (pre-posed <i>rā</i> marker)	34	14	2	8.2
<i>puliveri ke man kādoye tavalod barāt xaridam foruxti?</i> (unmarked)	34	13	3	8.1
<i>puliveri ke man kādoye tavalod barāt xaridameš foruxti?</i> (marked with an enclitic referent pronoun)	7	13	30	2.7
<i>puliveriro foruxti ke man kādoye tavalod barāt xaridam?</i> (unmarked and extraposed)	4	6	40	1.4

### DO in the MC

The results in table 6 show that examples of the target structure whose head noun phrase functions as the specific direct object in only the main clause are considered correct with either a pre-posed *rā* marker or a post-posed *rā* marker. Unmarked head nouns are also considered acceptable. The mean judgement scores show the same order of acceptability as for the *šāgerd* sentence in Table 4.

**Table 6:** Results of the acceptability judgement test of the sentence *I didn't eat the cold food left over from yesterday* (n=50).

	Correct	Acceptable	Incorrect	Mean judgement / 10
<i>ġazāye sardi ke az diruz munde naxordam</i> (pre-posed <i>rā</i> marker)	44	6	0	9.4
<i>ġazāye sardi ke az diruz mundero naxordam</i> (post-posed <i>rā</i> marker)	38	9	3	8.5
<i>ġazāye sardi ke az diruz munde naxordam</i> (unmarked)	20	19	11	5.9

### DO in the RC

The results in table 7 show that sentences with marked and unmarked head noun phrases that function as the specific direct object in only the relative clause are considered similarly acceptable, with a slight preference for unmarked head noun phrases.

**Table 7:** Results of the acceptability judgement test of the sentence *You added so much opiate to the bread that it killed a girl* (n=50)

	Correct	Acceptable	Incorrect	Mean judgement / 10
<i>moxaderi ke be nuneš zadi enqadr ziād bud ke ye doxtaro košt</i> (post-posed <i>rā</i> marker)	38	8	4	8.4
<i>moxaderiro ke be nuneš zadi enqadr ziād bud ke ye doxtaro košt</i> (unmarked)	34	10	6	7.8

## 4 Discussion

A mixed picture has emerged for examples of the target structure where the head noun phrase functions as the specific direct object in both the main clause and the relative clause. Fifty percent of the 'DO in both clauses' examples from the film corpus are unmarked, and the mean acceptability judgement score for unmarked sentences with canonical word order is 8.15 out of 10. These results clearly suggest that *rā* marking is unnecessary for acceptability in such relative clauses. However, the translations seem to reveal that the participants know that specific direct object marking is grammatically more correct, because 84% of their translations of 'DO in both clauses' sentences are *rā* marked<sup>14</sup>.

A greater degree of sensitivity to the need for specific direct object marking can be seen in examples where the head noun phrase functions as the specific direct object in only the main clause. Eighty-six percent of 'DO in the MC' examples from the film corpus are *rā* marked, and the *rā* marker is post-posed in half of those examples. However, the acceptability judgement scores show that pre-posed *rā* markers are considered marginally more correct. Unmarked head noun phrases in 'DO in the MC' sentences represent only 14% of the examples in both corpora, and such sentences receive a mean acceptability judgement score of only 5.9 out of 10.

<sup>14</sup> It is not surprising that translation, an explicit linguistic exercise, yields more correct language than film scripts that aim to imitate naturally occurring speech.

A clear picture has emerged for examples where the head noun phrase functions as the specific direct object in only the relative clause. Only one example in the whole corpus for this study has a *rā*-marked head noun phrase, meaning 99.5% of examples are unmarked. This result reveals how the participants speak; however, it does not demonstrate that a *rā* marker is considered incorrect. The *rā*-marked example in the acceptability judgement test (*moxaderiro ke be nuneš zadi enqadr ziād bud ke ye doxtaro košt*) has a mean score of 7.8 out of 10, only 0.6 lower than the unmarked version of the sentence.

Post-posed *rā* marking is used in almost half of the examples of ‘DO in both clauses’ and ‘DO in the MC’ sentences in the two corpora, suggesting that the post-posed *rā* marker has become well established in relative clauses in spoken Farsi. Very few participants in this study judged post-posed *rā* marking as incorrect and, overall, the scores of the acceptability judgement tests show that post-posed *rā* marking is considered only marginally less correct than pre-posed *rā* marking (Tables 4–6). Taken together, these results suggest that a great deal has changed since Karimi’s (2001) study, where the use of post-posed *rā* marking was dismissed as a “much less acceptable” innovation. This study’s results for post-posed *rā* marking also contrast with Abdollahnejad and Marefat’s (2017) study of newspaper editorials, where only 18.5% of the *rā* markings in relative clauses were post-posed. The combination of our results confirms a notable difference between written and spoken Farsi.

A significant result of this empirical investigation must be the omission of the *rā* marker in about half of the examples presented in the results section, despite the fact that *rā* marking is legitimate in every case. It is noteworthy that this omission occurs not only in examples where the head noun phrase has a dual syntactic function (as noted in Elwell-Sutton (1963: 114) and Lambton (1974: 76), but also in examples where the head noun phrase functions as the specific direct object in both clauses. It is also noteworthy that unmarked sentences are considered more than acceptable by the participants in this study, with a mean acceptability judgement score of 7.5 out of 10.

A final significant result of this study is the absence of *rā*-marked referent pronouns. Indeed, no example from the film and translation corpora includes this strategy. The one example provided in the acceptability judgement test (*šāgerdi ke uno entexāb kardi nemišenāsam*) has a mean acceptability judgement score of only 2 out of 10, and it is the sentence that aroused most reactions from the participants<sup>15</sup>. This result suggests that the once optional use of *rā*-marked referent pronouns in restrictive relative clauses (as noted in Lazard 1957: 222; Elwell-Sutton 1963: 112; Thackston 1993: 86) is no longer considered acceptable in spoken Farsi.

## 5 Conclusion

The combined results of this study reveal patterns that motivate the following conclusions about *rā* marking in restrictive relative clauses in spoken Farsi:

1. Correct *rā* marking of the head noun phrase in restrictive relative clauses is dependent on the syntactic function of the head noun phrase:
  - a. When the head noun phrase functions as the specific direct object in both the main clause and the relative clause, it can either be left unmarked, or marked with either a pre-posed or post-posed *rā* marker. A *rā*-marked head noun phrase is considered slightly more correct.
  - b. When the head noun phrase functions as the specific direct object in only the main clause, it can either be left unmarked, or marked with either a pre-posed or post-posed *rā* marker. A *rā*-marked head noun phrase is considered more correct.

<sup>15</sup> “uno is redundant” (Participant 14); “I should say it’s correct” (Participant 41); “Impossible!” (Participant 30); “Unsure... Maybe it’s wrong” (Participant 8).

- c. When the head noun phrase functions as the specific direct object in only the relative clause, the *rā* marker is omitted. However, a pre-posed *rā* marker in this context is not considered incorrect.
2. The use of the *rā*-marked referent pronoun in the relative clause is considered incorrect.

Together, these conclusions provide an up-to-date picture of *rā* marking in restrictive relative clauses in spoken Farsi. This description is necessary, because previous studies have either presented a misleadingly uniform pattern in the target structure (Rafiee, 1988) or presented only examples from written language (Thackston, 1993). Hopefully, the conclusions in this study provide not only some clarity for teachers and students of Farsi language as they deal with relative clauses, but also a contribution to research on the use of more complex grammatical structures in speech.

## Acknowledgement

This paper is based on a degree project written under the supervision of Professor Carina Jahani, whom I would like to thank for her generosity, dedication, and support not just during the project but over the last three years. I am also indebted to two anonymous reviewers of *Orientalia Suecana* for their constructive comments on an earlier version of this paper.

## Abbreviations

CL	classifier
CLM	clause linkage marker
COP	copula
DEM.DIST	demonstrative pronoun or determiner to show distal deixis
DEM.PROX	demonstrative pronoun or determiner to show proximal deixis
EŽ	ezāfe
IMP	imperfective (for past and non-past verb forms)
IND	individuation clitic
INT	interrogative pronoun
NEG	negative
OM	object marker
PC	person-marking clitic
PL	plural
PN	pronoun
PP	past participle
PREV	preverb
PRS	present tense
PST	past tense
REF	referential
SBJV	subjunctive
SG	singular

## References

- Abdollahnejad, E. and Marefat, H. 2017. Relative Clauses in Persian. A Small-Scale Corpus Study, *Linguisticae Investigationes*, 40(2), pp. 135–149.
- Dahlén, A. 2010. *Modern Persisk Grammatik*. Stockholm: Ferdosi Publication.
- . 2012. *Persiska för nybörjare*. Stockholm: Ferdosi Publication.

- Elwell-Sutton, L.P. 1963. *Elementary Persian Grammar*. Cambridge: Cambridge University Press.
- Jahani, C. 2000. Restrictive Relative Clauses in Classical and Modern New Persian and the Marking of the Antecedent, *Orientalia Suecana*, XLIX, pp. 33–53.
- Jam, B. 2017. On the Pronunciation of Postposition “rā” in Persian, *Journal of Linguistics and Khorasan Dialects*, 1 (Serial Number 16).
- Jasbi, M. 2014. *Semantics of Differential Object Marking in Persian*. Available at: [https://jasbi.github.io/research/persian\\_dom\\_semantics\\_draft.pdf](https://jasbi.github.io/research/persian_dom_semantics_draft.pdf) (Last accessed 19th April 2020).
- Karimi, S. 2001. Persian Complex DPs: How mysterious are they? *Canadian Journal of Linguistics/ Revue canadienne de linguistique*, 46(1/2), pp. 63–96.
- Lambton, A.K.S. 1974. *Persian Grammar*. Cambridge: Cambridge University Press.
- Lazard, G. 1957 *Grammaire du persan contemporain*. Paris: Librairie C. Klincksieck.
- Rafiee, A. 1988. *Colloquial Persian. The Complete Course for Beginners*. London: Routledge.
- Skutnabb-Kangas, T. 2008. Bilingual Education and Sign Language as the Mother Tongue of Deaf children. in Bidoli, K., Cynthia, J., and Ochse, E. (eds.) *English in International Deaf Communication*. Bern: Peter Lang, pp. 75–94.
- Thackston, W.M. 1993. *An Introduction to Persian*. 3rd ed. Bethesda, Maryland: Ilex Publishers.

## Films

- Jeyrani, F. 2002. *Šāme Āxar*. Available at: <https://www.imvbox.com/watch-persian-movie-iranian-movies/the-last-supper-shaame-akhar> (Last accessed 17th May 2020).
- Kamyar, B. 2017. *Mordād*. Available at: <https://www.imvbox.com/watch-persian-movie-iranian-movies/august-mordad> (Last accessed 17th May 2020).
- Khajooei, A. 2008. *Mahiā*. Available at: <https://www.imvbox.com/watch-persian-movie-iranian-movies/mahya> (Last accessed 17th May 2020).
- Moghadam, S. 2012. *Yek 'Āšegāneye Sāde*. Available at: <https://www.imvbox.com/watch-persian-movie-iranian-movies/a-simple-love-story-yek-asheghaneye-sadeh> (Last accessed 17th May 2020).
- Nematollah, H. 2017. *Rage Xvāb*. Available at: <https://www.imvbox.com/watch-persian-movie-iranian-movies/subdued-rage-khab> (Last accessed 17th May 2020).
- Raisian, A. 2019. *Mardi bedune Sāye*. Available at: <https://www.imvbox.com/watch-persian-movie-iranian-movies/a-man-without-a-shadow-mardi-bedoone-sayeh> (Last accessed 17th May 2020).
- Sabbaghzade, K. 2017. *Itāliā Itāliā*. Available at: <https://www.imvbox.com/watch-persian-movie-iranian-movies/italy-italy-italia-italia> (Last accessed 17th May 2020).
- Sadrameli, R. 2010. *Zendegi bā Ćešmāne Baste*. Available at: <https://www.imvbox.com/watch-persian-movie-iranian-movies/living-with-closed-eyes-zandegi-ba-cheshmane-baste> (Last accessed 17th May 2020).
- Saghafi, A. A. 2017. *Helen*. Available at: <https://www.imvbox.com/watch-persian-movie-iranian-movies/helen> (Last accessed 17th May 2020).
- Seyyedi, H. 2016. *Xāšm o Hayāhu*. Available at: <https://www.imvbox.com/watch-persian-movie-iranian-movies/sound-and-fury-khashm-va-hayahoo> (Last accessed 17th May 2020).