

Ambiguous Stripping Sentences in Persian: An Interpretation and Naturalness Rating Study¹

Vahideh Rasekhi & Jesse Harris

vrasekhi@humnet.ucla.edu, jharris@humnet.ucla.edu

ECBAE @ UC Davis, July 13

1. Introduction

- In stripping ellipsis such as (1), in the second clause there is a remnant (e.g. *the police*) and the negative marker *not*.

Depending on the accent placement, the remnant can contrast either with the subject or object in the antecedent clause. Therefore, the sentence can have two interpretations (i.e. the remnant (*police*) can contrast with both the subject (*smuggler*) and object (*gangster*) in the antecedent clause), as shown in (1a) and (1b).

- (1) The smuggler followed the gangster, but **not the police**.
- a. *Subject contrast interpretation*: The smuggler followed the gangster, but **the police** didn't follow the gangster.
- b. *Object contrast interpretation*: The smuggler followed the gangster, but the smuggler didn't follow **the police**.
- How do speakers choose the right correlate for the remnant?
 - Correlate-remnant pairing is important as it determines the interpretation and the structure of the ellipsis site.
 - The purpose of this study is to determine what strategies native speakers use to disambiguate ambiguous elliptical sentences such as (1).
 - Previous studies in English have found a preference for object contrast

¹ Many thanks to Maziar Toosarvandani, Katy Carlson, the participants at the UCLA Psycholinguistics Seminar, and ECBAE reviewers for their feedback! The first author is also thankful to the Roshan Cultural Heritage Institute and especially its Chair and President, Dr. Elahé Omidyar Mir-Djalali for making this research possible.

in sluicing² and stripping constructions (Carlson 2014; Harris 2015; Harris & Carlson 2017).

- However, studies on how the correlate-remnant pairing is resolved in English ellipsis have also found a strong effect of parallelism, in that subject reading can become more salient when the subject noun is marked for focus, contains a contrastive adjective, or is otherwise semantically parallel with the remnant (2).

(2) *Focus marking via particle only*

Only the smuggler followed the gangster, but not the police.

- In addition, in languages such as German that have rich morphology, case marking on DPs help disambiguate the contrast (3).

(3) Am Dienstag hat **der** Direktor **den** Schüler getadelt,
on Tuesday has the.NOM principal the.ACC student criticized
und nicht **der/den** Lehrer
and not the.NOM/ACC teacher

'On Tuesday, the principal criticized the students, and not the teacher.'

(Stolterfoht et al. 2007)

- Persian has only one case marking *-ra* which occurs on the definite/specific object (4a). When there is no *-ra* marking, the sentence becomes ambiguous (4b).³

(4) a. qāchāqchi gangster-ra donbāl=kard, vali **polis-ra** na
smuggler gangster-ra follow=did but police-ra not
'The smuggler followed the gangster but the smuggler did not follow **the police**.'

b. qāchāqchi gangster-ra donbāl=kard, vali **polis** na
smuggler gangster-ra follow=did but police not

² In sluicing such as (i), it is preferred to have the object (*someone*) as the correlate of the remnant *wh*-phrase (*who*) (see, for example, Frazier and Clifton 1998, Carlson et al. 2009).
(i) *Somebody* claimed that the president fired *someone*, but nobody knows *who*.

³ Refer to Rasekhi (2018) for the discussion on the syntax of stripping constructions in Persian.

- i. *Object contrast interpretation*: The smuggler followed the gangster but the smuggler did not follow **the police**.’
 - ii. *Subject contrast interpretation*: The smuggler followed the gangster but **the police** did not follow the gangster.
- In this study, the following questions are addressed:
 - Does the object bias that we see in English hold in other languages including Persian?
 - There are various strategies such as prosody, discourse, and overt focus marking that can help interpret ambiguous sentences. But how do speakers in the absence of such cues process and interpret ambiguous elliptical sentences (e.g. in silent reading)?
 - How do speakers identify the corresponding correlate of the remnant?
 - How do other factors such as structural and morphological parallelism help individuals locate a correlate for remnant in elliptical structures?
 - Ambiguous ellipsis sentences whose resolution depends on finding a contrastive correlate in the antecedent clause provides a testing ground for these questions.
 - In this study, we use Persian as a test ground for disambiguation since Persian has two characteristics that don’t exist in English:
 - The object is optionally marked for case.
 - We test how the existence/absence of case-marking affects interpretation of ambiguous sentences.
 - Persian allows scrambling, in which an element is focalized and moved to the sentence initial position.
 - We test whether word order (information structure) plays a role in interpreting ambiguous sentences.
 - Manipulating the degree of morphological and structural parallelism

between the matrix and the ellipsis clause helps us determine how speakers interpret ambiguous sentences.

2. Background on Persian

- The canonical word order in Persian is SOV.
- The direct object can be realized as a bare noun (5) or with an object marker *-ra* (6) (Karimi 1990).

(5) shir **babr** kosht
lion tiger killed
 Lit: ‘Lion killed tiger.’

(6) shir **babr-ra** kosht
lion tiger-ra killed
 Lit: ‘Lion killed **the** tiger.’

- Definite objects can be focalized and moved to the sentence initial position (7).

(7) **babr-ra** shir kosht
tiger-ra lion killed
 ‘It was the tiger that lion killed.’

3. Hypotheses for Disambiguation

- a. **Locality**: Previous experimental studies in English have found a strong preference for the remnant to contrast with the most local possible correlate (typically the object) in sluicing and stripping constructions (Carlson 2014; Harris 2015; Harris & Carlson 2017).
- Based on the locality hypothesis, in (1), it is preferred to have the object *gangster* as the correlate of the remnant *the police*. However, we want to test whether this bias for object contrast interpretation holds in other languages including Persian.

b. **Morphological Parallelism⁴**: An additional hypothesis is morphological parallelism, in which there is a strong preference for the correlate-remnant pair to share a similar morphological shape. This would potentially override any preference for the most local correlate.

- Based on morphological parallelism, if the remnant and the object in the antecedent clause have *-ra* marker, the sentence would have an object interpretation (8). On the other hand, if the remnant does not have *-ra* marker, subject interpretation would be preferred.

(8) qāchāqchi gangster-ra donbāl=kard, vali polis-**ra** na
smuggler gangster-ra follow=did but police-ra not
 ‘The smuggler followed the gangster but the smuggler did not follow the

police.’

- Persian allows us to compare the relative effects of morphological and structural similarity between the remnant and potential correlates in the antecedent clause via the *-ra* marker on the definite/specific object and object scrambling.

3. Method

- We manipulated the degree of morphological and structural parallelism between the matrix and the ellipsis clause by manipulating *-ra* mark and word order in 3x2 design (9). The test items were designed to be plausible on both subject and object readings.

(9) Possible variations of stripping construction with regard to *-ra* marking and scrambling

Antecedent clause	No <i>-ra</i> in remnant	<i>-ra</i> in remnant
Canonical: SOV	a. koose māhi gereft, vali panguan na <i>shark fish caught but penguin not</i>	d. koose māhi gereft, vali panguan- ra na <i>shark fish caught but penguin-ra not</i>
Canonical <i>-ra</i> marked: SO- <i>ra</i> V	b. koose māhi- ra gereft, vali panguan na <i>shark fish-ra caught but penguin not</i>	e. koose māhi- ra gereft, vali panguan- ra na <i>shark fish-ra caught but penguin-ra not</i>
Scrambled: O- <i>ra</i> SV	c. māhi- ra koose gereft, vali panguan na <i>fish-ra shark caught but penguin not</i>	f. māhi- ra koose gereft, vali panguan- ra na <i>fish-ra shark caught but penguin-ra not</i>

- An Internet questionnaire (conducted on Ibex Farm) was completed by 60 native speakers of Persian.
- 24 sextets were designed along the lines of those in (9).
- Test items were presented along with filler items and catch items in counterbalanced and randomized order (80 sentences total). Each participant saw only one condition per item.
- Naturalness of test items was pre-tested with six native speakers. This

was done to ensure that both subject and object readings are equally natural.

- The participants rated items like (9) on their naturalness on the scale of 1 (completely unnatural) to 7 (completely natural), and then answered a comprehension question, choosing between a Subject interpretation, Object interpretation, Both, or Neither.

⁴ It should be noted that Morphological Parallelism is not intended to override other factors like plausibility or Semantic Parallelism. The test items used in this study were designed so that both Subject and Object contrast interpretation were plausible.

In addition, it should be noted that the Locality bias and Morphological Parallelism are compatible. We are testing for the strength of Morphological Parallelism vs. other biases guiding correlate-remnant pairing.

- **Sample Questionnaire item:**

- Naturalness rating task

koose māhi gereft vali panguan na
shark fish caught but penguin not

(completely unnatural) 1 2 3 4 5 6 7 (completely natural)

- Comprehension question

mafhoome jomleye “koose māhi gereft vali panguan na”
meaning sentence shark fish caught but penguin not
 kodām yek az gozinehāye zir mibāshad?
which one of options below is

Which one of the following is the meaning of “shark caught fish, but penguin not”?

- ‘Shark caught fish, but penguin did not catch fish.’
- ‘Shark caught fish, but shark did not catch penguin.’
- Both (a) and (b)
- I am not sure

3. Predictions

- Locality:** There is a general locality preference (object contrast preference in 9b and subject contrast preference in 9c) independent from morphological parallelism.
- Morphological parallelism:**
 - Remnants with *-ra* marking (conditions 9d-9f) disambiguate to the Object, regardless of the morphological marking or order in the antecedent clause (provides a “sanity check” of intuitions).
 - Remnants without *-ra* marking:
 - Ambiguous with respect to the Subject/Object for antecedents with Canonical word orders and no morphological marking (condition 9a).
 - When there is *-ra* marking on the antecedent:
 - If ambiguous remnants are biased towards (morphologically

parallel) Subject interpretation, regardless of word order (i.e. in both 9b and 9c), this means that morphological parallelism is stronger than locality / information structure in pairing a correlate with a remnant.

- Regarding naturalness, our predictions were as follows:
 - Sentences that were morphologically matched would be rated as more natural.
 - Scrambled sentences would be rated less natural.

5. Results

Results were modeled as (G)LMERs with by-subject and by-item random slopes. Predictors were sum coded.

5.1. Interpretation Results

We analyzed answers for **comprehension questions** (removing Neither responses for convenience; < 4% data loss). Fig. 1 illustrates the results for each condition, summarized in (10).

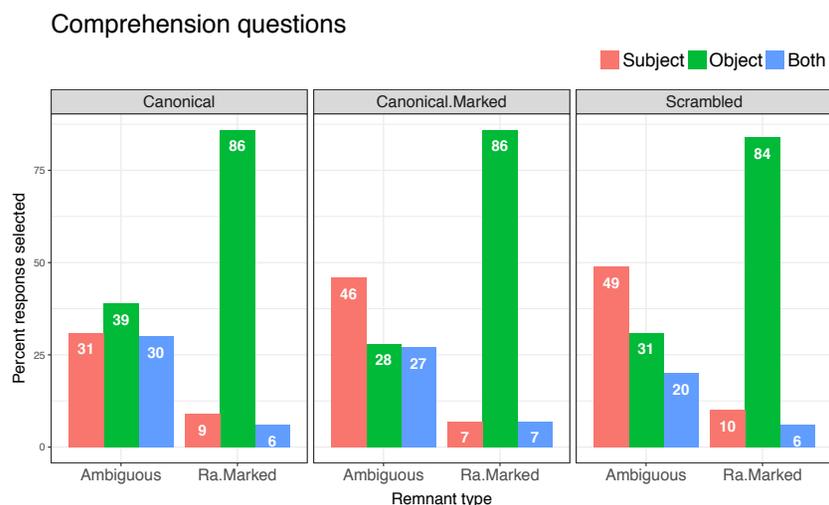


Figure 1. Comprehension questions

- (10) a. For all antecedent types: *-ra* marking on the Remnant strongly biases towards an Object contrast interpretation, regardless of whether the object is also *-ra* marked in the antecedent clause.
- b. Subject and Object contrast interpretations are both possible when there is no *-ra* marking on the remnant:
- i. When there was no *ra*-marking on the remnant (9a), there was a slight preference for object contrast interpretation (39 vs. 31).
 - ii. However, participants chose the Subject contrast interpretation more often when the object in the antecedent clause was *-ra* marked (46% in Canonical Marked and 49% in Scrambled word orders).
- c. The results suggest:
- i. Absence of Locality effect.
 - ii. Comprehenders used Morphological Parallelism to resolve ambiguous remnants, pairing two nouns that were (not) *-ra* marked as the remnant and correlate.
 - iii. When the object is *-ra* marked in the antecedent clause, the presence of a *-ra* in the remnant determines the interpretation of the remnant in stripping ellipsis:
 - a) If there is *-ra* marking in remnant, then we have Object contrast interpretation.
 - b) If there is no *-ra* marking in remnant, then we have Subject contrast interpretation.

5.1. Acceptability Ratings Results

- a. There was an interaction between Remnant type and Antecedent clause for acceptability ratings, as illustrated in Fig. 2.
- b. Antecedents with no *ra*-marking were equally acceptable with both non-marked and *ra*-marked remnants, however, *ra*-marking on the remnant improved acceptability ratings by .78 for Canonical

Marked conditions and by .81 for Scrambled Marked conditions, p 's < 0.001 (confirmed by planned by-subject and by-item paired t-test comparisons).

- Acceptability increases when there is morphological parallelism.
- In addition, Scrambled word order in the antecedent decreased acceptability ratings overall ($p < .001$).

Acceptability ratings

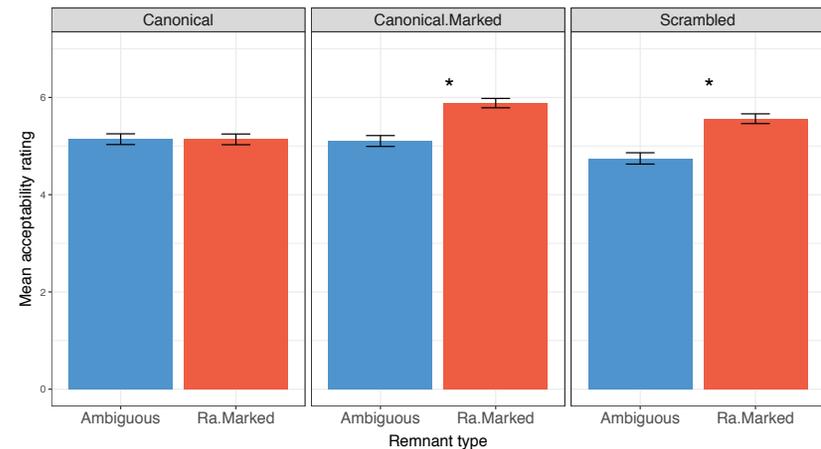


Figure 2. Acceptability ratings

5.2. Interaction of Acceptability Rating with Response Type in Ambiguous Items

- In ambiguous items⁵ (9a-9c), there was an interaction between the response type (Object contrast interpretation vs. Subject contrast interpretation) and acceptability rating, as illustrated in Fig. 3.
- Ambiguous sentences involving scrambling were rated lower than the other two sentence types.
- Sentences with scrambled word order showed greater increase in ratings when Subject interpretation was selected.
- Ambiguous items provide evidence against Locality effect.

6. Summary and Conclusion

- If there is *-ra* marking on the remnant and on the object on the antecedent clause, there is an Object contrast interpretation. However, if there is no *-ra* marking on remnant, there is a strong preference for Subject contrast interpretation.
- In ambiguous sentences, participants prefer Subjects contrast interpretation as the remnant is morphologically parallel to its antecedent (i.e. subject).
- When there is morphological parallelism, acceptability rating increases.
- This study contributes a novel experiment on Persian to the growing literature on the processing and interpretation of ellipsis. The findings are consistent with the Morphological Parallelism hypothesis, in which comprehenders prefer to contrast remnants with correlates that match on morphological shape, even overriding structural biases. This means that morphological parallelism is a stronger cue to remnant-correlate pairing than locality and word order in interpreting ambiguous sentences.

Acceptability ratings

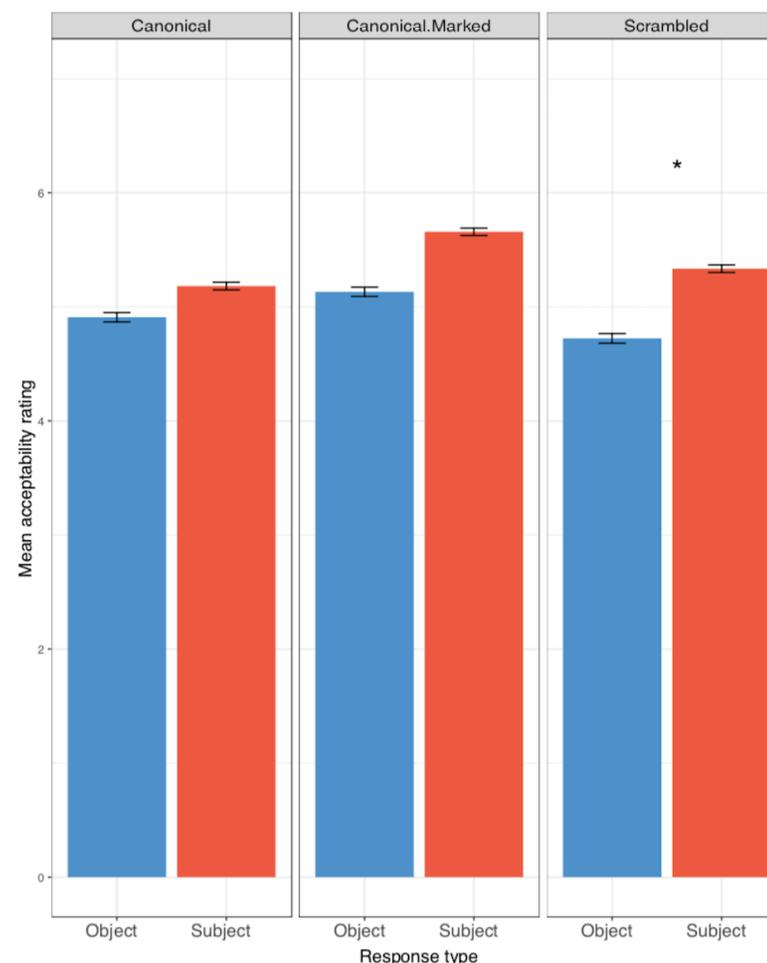


Figure 3: Interaction of Acceptability Rating with Response Type

⁵ Recall that non-*-ra*-marked sentences are ambiguous.

References

- Carlson, K. (2001). Predicting contrast in sentences with and without focus marking. *Lingua*, 150, 78-91.
- Carlson, K., Dickey, M.W., Frazier, L., & Clifton, C. (2009). Information structure expectations in sentence comprehension. *The Quarterly Journal of Experimental Psychology*, 62, 114-139.
- Carlson, K. (2014). Predicting contrast in sentences with and without focus marking. *Lingua* 1, 78-91.
- Frazier, L., & Clifton, C. (1998). Comprehension of sluiced sentences. *Language and Cognitive Processes*, 13, 499-520
- Harris, J. (2015). Structure modulates similarity-based interference in sluicing: An eye-tracking study. *Frontiers in Psychology*, 6. Encoding and navigating linguistic representations in memory.
- Harris, J. & Carlson, K. (2017). Information structure preferences in focus-sensitive ellipsis: How default persist. *Language and Speech*, 61, 480-512.
- Karimi, S. (1990). Obliqueness, specificity, and discourse functions: *Ra* in Persian. *Linguistic Analysis*, 20, 139-191.
- Rasekhi, V. (2018). *Ellipsis and information structure: Evidence from Persian*. Ph.D. Dissertation, Stony Brook University.
- Stolterfoht, B., Friederici, AD., Alter, K., Steube, A. (2007). Processing focus structure and implicit prosody during reading: Differential ERP effects. *Cognition*, 104, 565-590.

Appendix

Follow up studies:

Our current results suggest that morphological parallelism overrides locality in finding a corresponding correlate for the remnant. To further investigate this, we will have two follow up studies:

I. Sentence completion task

- In the above study, participants were forced to choose an interpretation for the remnant (*object reading, subject reading, both, neither*).
- In the follow up study, participants will be given sentences as in (11), in which they need to complete the sentence by writing the remnant to avoid (i) imposing potentially implausible relations into the sentence, and (ii) the “both” and “neither” response options.

(11) a. mohandes dar āvord, vali ____ na (Canonical: SOV)
engineer door brought but not

b. mohandes dar-ra āvord, vali ____ na (Canonical -ra marked:
engineer door-ra brought but not SO-raV)

c. dar-ra mohandes āvord, vali ____ na (Scrambled: O-raSV)
door-ra engineer brought but not

- In this study, we will use animate and inanimate nouns to help determine whether a Subject or Object contrast interpretation was selected.
- **Predictions:**
 - Bias towards animate NPs (indicating a subject remnant interpretation) in Canonical orders.
 - Increased rate of inanimate NPs completions (indicating an Object contrast interpretation) in Canonical Ra-Marked and Scrambled conditions.

II. Speeded grammaticality judgment task

- Participants will see sentences such as (12) word by word in RSVP.
 - As before, the subject and object in the antecedent clause differ in their animacy. However, the remnant will be specified with either animate (*workers*) or inanimate (*windows*) NPs.
- (12) a. mohandes dar-ra āvord vali panjare na
engineer door-ra brought but window not
'The engineer brought the door but he didn't bring windows.'
- b. mohandes dar-ra āvord vali kārgar na
engineer door-ra brought but worker not
i. 'The engineer brought the door but he didn't bring workers.'
ii. 'The engineer brought the door but the worker didn't bring doors.'
- After each sentence, participants will press buttons as fast as they can to indicate whether the sentence is grammatical or not. The speed at which they respond would indicate how fast/slow they resolve the

remnant contrast.

- In (12a), the remnant can only contrast with the object while the remnant in (12b) can contrast with either the object or the subject.
- We know that when the remnant is not *-ra* marked, there is a bias towards subject contrast (when NPs are equally plausible). However, in (12a), in which the remnant is inanimate and is not *-ra* marked, there is object bias.
- In this example, the remnant is morphologically similar to the subject but semantically it has to be object.

Hypothesis:

Since in elliptical structures, remnant needs to contrast with an element in the antecedent clause (either subject or object), we expect morphological parallelism to make processing faster. However, if this is constrained by animacy (i.e. inanimate remnant cannot be plausibly understood as the subject), then semantic information would override morphological parallelism.