

## Disambiguating stripping ellipsis in Persian: How parallelism and locality interact

Vahideh Rasekhi & Jesse Harris (University of California, Los Angeles)

[vrasekhi@humnet.ucla.edu](mailto:vrasekhi@humnet.ucla.edu)

**Introduction.** Experimental studies of clausal ellipsis structures like stripping ellipsis (1) in English have found a general preference for the remnant noun (*the police*) to contrast with the structurally most local/recent correlate noun (typically the object, *the gangster*) (Locality Bias: Carlson 2014; Harris 2016). However, factors such as information structure modulate this preference: a non-local (e.g., the subject, *the smuggler*) noun is more salient as a correlate when it is marked for focus or is semantically parallel with the remnant (Carlson 2001; Harris & Carlson 2017). As such cues are relatively weak in English, we turned to Persian, which (i) marks information structure overtly via word order scrambling and (ii) uses the *-ra* suffix to mark definiteness/specificity. In two offline experiments, we investigated how speakers disambiguate stripping ellipsis in Persian (Toosarvandani 2015, Rasekhi 2018) to determine how potentially conflicting cues are used to find a correlate for the remnant. We find that while overt indicators of information structure and parallelism most strongly influence correlate resolution, a Locality Bias nevertheless emerges in cases of Scrambling where the *-ra* suffix is not optional.

**Experiment 1.** We manipulated the degree of morphological/semantic parallelism between nouns with/without the *-ra* marker and word order (canonical: SOV, canonical marked: SO-*ra*V, scrambled: O-*ra*SV) in a 2x3 design. Whereas *-ra* marking the Object is optional in Canonical order, it is required in Scrambling. We hypothesized a preference for Semantic/Morphological Parallelism, in which there is a strong preference for the correlate-remnant pair to share a similar semantic and/or morphological shape, potentially overriding a global preference for Locality. An Internet questionnaire was completed by 60 native speakers of Persian, who rated items like (2) for naturalness, and then answered an interpretation question choosing an option disambiguating between a Subject interpretation, Object interpretation, Both or Neither. Results were modeled as (G)LMERs. We found that Canonical SOV order was equally acceptable with both Ambiguous (without *-ra* marking) and *ra*-marked remnants (Fig. 1, left panel). However, *-ra* marking improved acceptability ratings for Canonical Marked and Scrambled conditions ( $p$ 's < 0.001; center and right panels in Fig 1). The analysis of answers for interpretation questions showed that unmarked remnants were compatible with either Subject (Non-Local) or Object (Local) contrasts, but that an Object contrast was strongly preferred when the remnant was *-ra* marked (Fig. 2). A Subject contrast was selected more often when the remnant did not have *-ra* marking while the object in the antecedent clause was *-ra* marked. Thus, Semantic/Morphological Parallelism resolved ambiguous ellipsis structures, pairing two nouns as the remnant and correlate that matched with regard to the presence/absence of *-ra* marking.

**Experiment 2.** 21 native speakers of Persian completed sentence fragments with remnants (3) to allow us to gauge production preferences with a less constrained method. Noun animacy in the antecedent clause was controlled to allow us to determine whether the completion contrasted with a Subject (animate) or an Object (inanimate) remnant. The results (analyzed as  $\chi^2$  tests due to categorical skew) showed that both Locality Bias and Semantic/Morphological Parallelism affect remnant choice. Object remnants were only *-ra* marked when the correlate was also *-ra* marked;  $\chi^2(2) = 48.59$ ,  $p < 0.001$ . Both types of Canonical orders (SOV and SO-*ra*V) elicited more Object contrasts (56-59% Object contrasts) compared to the Scrambled order (O-*ra*SV), which was biased towards Subject contrast completions (64% Subject contrasts);  $\chi^2(2) = 8.15$ ,  $p < 0.001$  (Locality Bias – Fig. 3 and Table 1). Taken together, the completion patterns suggest that comprehenders were more tempted by a Subject contrast when it was Local and the Object *-ra* marking in the antecedent clause was required independently by Scrambling.

**Conclusion.** Results indicate that language-specific indicators of information structure and a preference for parallelism in semantic/morphological forms are dominant factors in interpreting contrastive ellipsis structures. Nevertheless, a preference for Local correlates manifests in Scrambling, in which the morphological shape on the Object is constrained by the grammar.

- (1) The smuggler followed the gangster, but not **the police**. (Ambiguous Stripping Ellipsis)  
 a. The smuggler followed the gangster, but **the police** didn't follow the gangster. (Subj. interp.)  
 b. The smuggler followed the gangster, but the smuggler didn't follow **the police**. (Obj. interp.)

(2) **Sample sentence item from Experiment 1: Acceptability Judgment Task**

modir monshi estekhdām kard vali **moāven** na (Ambiguous Stripping Ellipsis)  
 manager secretary hire did but assistant not

- a. 'The manager hired a secretary but **the assistant** did not hire a secretary.' (Subj. interp.)  
 b. 'The manager hired a secretary but the manager did not hire **an assistant**.' (Obj. interp.)

(3) **Sample fragment item from Experiment 2: Sentence Completion Task**

**Antecedent clause**

a. Canonical: SOV	mard-e javān māshin kharid vali _____ na
	<i>man-EZ young car bought but not</i>
b. Canonical –ra marked: SO-raV	mard-e javān māshin- <b>ra</b> kharid vali _____ na
	<i>man-EZ young car bought but not</i>
c. Scrambled: O-raSV	māshin- <b>ra</b> mard-e javān kha <sup>id</sup> vali _____ na
	<i>car man-EZ young bought but not</i>

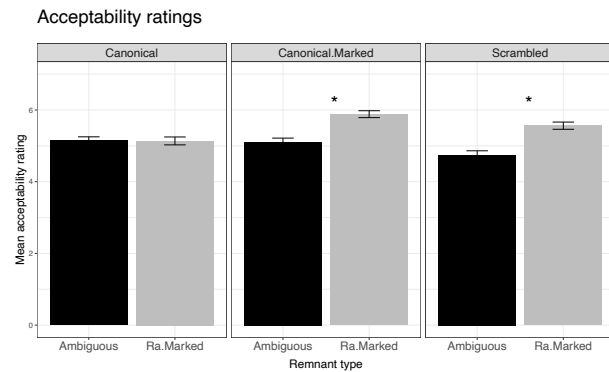


Fig 1: Acceptability ratings for Exp 1

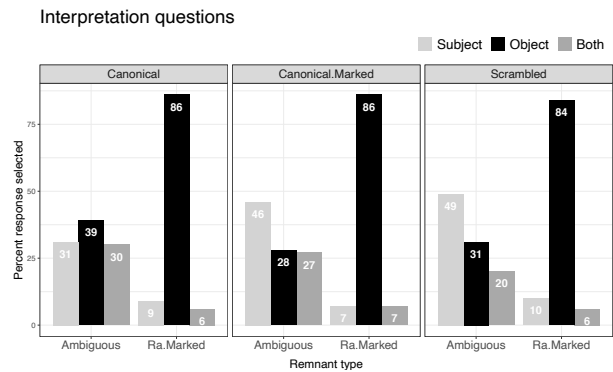


Fig 2: Interpretation questions for Exp 1

Completion study

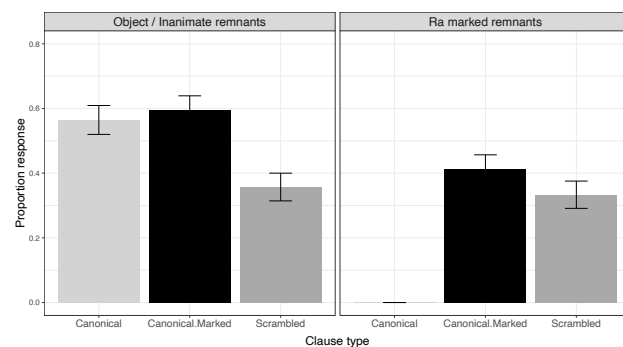


Fig 3: Completion proportions for Exp 2

Clause	Response	Ra. marked	Not - Ra. marked
a. Canonical	Object	0%	56%
	Subject	0%	44%
b. Canonical. Marked	Object	41%	18%
	Subject	0%	40%
c. Scrambled	Object	33%	2%
	Subject	0%	64%

Table 1: Summary of completions in Exp 2

**Selected references: Carlson (2014).** Predicting contrast in sentences with and without focus marking. *Lingua* 1, 78-91. **Harris & Carlson (2017).** Information structure preferences in focus-sensitive ellipsis: How default persist. *Language and Speech* 61, 480-512. *Frontiers in Psychology*, 6. **Rasekhi (2018).** Ellipsis and Information structure: Evidence from Persian. PhD Dissertation. Stony Brook University. **Toosarvandani (2015).** Persian. In: *The Oxford Handbook of Ellipsis*.