COPY RAISING IN PERSIAN*

Dennis Ryan Storoshenko, Elias Abdollahnejad, and Mahyar Nakhaei
University of Calgary

In this paper, we explore the possibility that so-called raising constructions in Persian are not a unitary phenomenon. First, we note that some speakers allow sentences with the predicate *be nazar âmadan* which cannot be accounted for under the A′-movement analysis of Karimi (2018). Rather, we claim that these sentences are examples of Copy Raising, in the sense of Landau (2011). Using psycholinguistic experimentation, we show that in sentences which can only be derived by A′-movement, the predictions of Karimi's analysis hold. However, participants do not demonstrate responses expected under an A′-movement analysis when another derivation is available.

1. Raising in Persian

The Persian complex predicate *be nazar âmadan* 'to come to view' is typically translated to English as 'seem' or 'appear'. In this section, we recount the most widely-held syntactic analysis of sentences using this predicate, as described in Karimi (2018). It is true that over the past 25 years there has been some change in how these examples are described in syntactic literature, but we hold discussion of this until section four. Our primary goal in this section is to present the current state of affairs, and to introduce new data which challenge the existing analysis. Throughout the paper, we refer to sentences with *be nazar âmadan* as examples of raising, mainly based on their semantic similarity to English sentences containing the raising verbs *seem* or *appear*. Descriptive use of the term *raising* in reference to sentences of Persian should not be taken to presume a particular syntactic analysis.

According to Karimi, Persian raising sentences are derived by A′ movement. The argument for this can be reconstructed from the following examples, beginning with (1):[1]

(1) be-nazar mi-yâ-d [(ke) bachche-hâ xaste bâsh-an]
to-view ASP-come-3SG that child-PL tired SUBJ.be-3PL
'It seems that the children are tired. (Adapted from Karimi 2018)'

From (1), we can observe the first of three key properties of A′-movement Karimi identifies. The embedded clause subject *bachche-hâ* remains in the embedded clause, yet the sentence is grammatical. This demonstrates the optionality of the movement. A parallel version of

*We would like to thank the participants in our lab studies, and the reviewers and attendees at the CLA for their comments. In particular, thanks to Morgan Mameni and Sahar Taghipour for their discussion of these data. All errors are our own.

[1]The following glossing abbreviations are used in this paper: ASP-(durative) aspect, CL-clitic, EZ-*ezafe*, OBJ-object marking, PL-plural, PPT-past participle, SG-singular, SUBJ-subjunctive

Actes du congrès annuel de l'Association canadienne de linguistique 2020.
Proceedings of the 2020 annual conference of the Canadian Linguistic Association.

the sentence, with *bachche-hâ* appearing before the inflected raising verb would also be grammatical; this is illustrated in (2):

(2) bachche-hâ$_i$ be-nazar mi-yâ-d [(ke) t$_i$ xaste bâsh-an]
child-PL to-view ASP-come-3SG that tired SUBJ.be-3PL
'The children seem tired.'

Note though that when the plural subject of the embedded clause is raised, there is no agreement with the matrix verb; the raising verb still bears 3SG agreement. Indeed, Karimi claims that the raising predicate always bears 3SG agreement. This lack of agreement with displaced arguments is the second key property of Persian raising sentences which supports an A′-movement analysis.

The last key property of Persian raising sentences which leads Karimi to propose an A′-movement analysis is illustrated in (3).

(3) [ketâb-â-ro]$_i$ be-nazar mi-yâ-d [(ke) bachche-hâ t$_i$ xunde bâsh-an]
book-PL-OBJ to-view ASP-come-3SG that child-PL read.PPT SUBJ.be-3PL
'As for the books, it seems that children have read (them).' (Karimi 2018)

In this sentence, it is the embedded clause object, *ketâb-â-ro* 'the books', which appears in the matrix clause. In typical subject-to-subject (STS) raising, a reflex of the matrix clause EPP, only the subject of the embedded clause should be able to move up, for locality reasons. Furthermore, if the presence of the definite object marker *-ro* is indicative of assignment of accusative case to the object in the embedded clause, as suggested in Abdollahnejad (2020), then the embedded object DP should not be a viable target for movement into what is typically an A-position. Thus, the movement into the matrix clause of a Persian raising sentence is symmetric, the third key property. In effect, this sentence is doubly-marked for A′-movement, as the plural non-subject has moved, and yet the matrix agreement is still singular. It should be noted that we do not dispute Karimi's analysis for examples of this type: the only viable analysis for (3) is, in our view, A′-movement.

However, some speakers accept Persian raising sentences such as the following:

(4) bachche-hâ$_i$ be-nazar mi-yâ-**n** [(ke) t$_i$ xaste bâsh-an]
child-PL to-view ASP-come-3PL that tired SUBJ.be-3PL
'The children seem like they are tired.'

This is a sentence parallel to (2), except now the matrix verb agrees with the displaced subject from the embedded clause. While it is already clear from (1) that this movement is optional, this is a clear counter-example to the claim that the raising verb must always have 3SG agreement. And, as this is a displaced subject from an intransitive embedded clause, there is no way to affirmatively claim that symmetry applies.

In the following sections, we develop the argument that (4) is in fact best analyzed as an example of Copy Raising (CR), as described by Landau (2011). In Section 2, we illustrate Landau's CR analysis using English, most crucially bringing to light a semantic requirement on CR subjects, which we claim also applies to (4), but not (2). In Section 3,

we discuss results from two psycholinguistic experiments where Persian raising sentences were included among trial items. For both, we show that the results are not in line with the prediction that all Persian raising sentences are derived from A′ movement. Our claim is that when a 3SG subject of an embedded clause appears in the subject position of a *be nazar âmadan* clause, two analyses are available. Such cases are compatible with either the A′-movement analysis of Karimi (2018), or they are compatible with our proposed CR analysis. We propose that this derivational ambiguity is the source of unexpected results in our studies. Lastly, in Section 4, we provide some suggested avenues for further research.

2. Copy raising

STS raising has a much less-discussed counterpart, known in the literature as Copy Raising (CR). An overview of early work on this construction appears in Potsdam and Runner (2001), where it is noted that while the construction appears in many languages, it is relatively understudied in English. One fact worth noting is that "Raising" is a misnomer in this case, as contemporary analyses of this construction do not involve movement. We continue to use this label despite the suggestion of a movement analysis.

We take the analysis presented in Landau (2011) as our jumping off point, though much of the same empirical ground is covered in Asudeh and Toivonen (2012), working in a different theoretical framework. Typical English CR is shown in (5):

(5) a. It seems like the kids are tired.

b. The kids$_i$ seem like they$_i$ are tired.

The clearest distinction between STS raising and CR sentences in English is the presence of the preposition *like* over the embedded clause, and the fact that the embedded clause is not an infinitive, regardless of the presence of an expletive in the matrix clause. As discussed by Landau, such sentences present a conundrum for syntactic theory, as the ability to take an expletive subject, as in (5a) suggests that, just as in STS raising, *seem* does not itself assign a theta role to the element that ultimately ends up in subject position. However, when there is a non-expletive subject, as in (5b), that subject determines the agreement on *seem*, while binding a pronoun in the embedded subject position. This appears to violate the theta-criterion, as the matrix subject *the kids* in (5b) does not have an evident theta-role. If it did, then presumably it should not be necessary to have a copy in the embedded clause. Such examples are marginal at best for English speakers:

(6) ? The house seems like the kids are tired.

For speakers who accept (6), *seem* would have a thematic subject, and accounting for (5b) is not a problem. However, there is a larger proportion of English speakers for whom (5b) is acceptable, while (6) is marginal if not ungrammatical. For these speakers, the raising verb does not assign a theta-role.

Landau's proposal is that the preposition *like* allows the merge of a null operator abstracting over the embedded clause, yielding an interim semantic structure for (5b) along the lines of (7):

(7) λx.[like x are tired]

In effect, the pronoun *they* becomes a variable in a newly-derived one-place property. This property is taken as the complement of the raising verb. The matrix subject merges directly in the matrix clause, but receives its semantic interpretation by saturating the predicate that has been formed from the abstraction over the embedded clause. The identity between arguments in the matrix and embedded clauses is a result of the fact that the semantics essentially unifies them as taking the same role in the embedded predicate. This replaces earlier accounts where the pronoun in the embedded clause is an overt trace of movement.

For Asudeh and Toivonen, and for Landau, there is a requirement that the subject of the matrix clause be a perceptual source. While they formalize this in slightly different ways, the intuition is that the subject of the raising verb should be interpretable as the speaker's source of information for the proposition in the embedded clause. If the raising verb is actually assigning a theta role, then the role would be one of perceptual source; this would be the role assigned to *the house* for speakers accepting (6). Otherwise, Landau proposes that while the requirement is not formally part of the semantics, a pragmatic constraint on the usage of such sentences restricts them to contexts where the embedded proposition can be plausibly inferred, and direct observation of the copied argument is the easiest route to this plausibility.

This requirement that the subject of a CR predicate be a perceptual source, arising either from pragmatics or from direct theta role assignment, is at the heart of our proposal that CR is also possible in Persian. The crucial minimal pair comes in examining two examples from the previous section:

(8) a. bachche-hâ$_i$ be-nazar mi-yâ-**d** [(ke) t$_i$ xaste bâsh-an]
child-PL to-view ASP-come-3SG that tired SUBJ-be-3PL
'The children seem tired.'

b. bachche-hâ$_i$ be-nazar mi-yâ-**n** [(ke) t$_i$ xaste bâsh-an]
child-PL to-view ASP-come-3PL that tired SUBJ-be-3PL
'The children seem like they are tired.'

In (8a), repeated from (2), there is no agreement with the plural subject, making this one of the examples supporting the A′-movement analysis of Persian raising. Conversely, (8b) shows 3PL agreement, suggesting that *bachche-hâ* is indeed the subject of the matrix clause. Furthermore, there is an additional interpretive distinction between the sentences: (8b) is only felicitous after direct observation of the children. This is reminiscent of the constraint on CR subjects, which are interpreted as perceptual sources. Landau also notes that in a *pro*-drop language, it would be entirely possible for the embedded clause copy of a CR subject to be covert. Thus, (8b) could be reanalyzed as in (9):

(9) bachche-hâ$_i$ be-nazar mi-yâ-**n** [(ke) *pro*$_i$ xaste bâsh-an]
child-PL to-view ASP-come-3PL that tired SUBJ-be-3PL
'The children seem like they are tired.'

The embedded clause subject position is occupied by a co-indexed *pro*, and there is no movement relationship between the two clauses. Despite the lack of movement, this fits the description of a CR sentence: there is an interpretive restriction on the subject of the raising verb, which also triggers agreement with the raising verb, while being co-referential with a pronoun in the embedded clause. The lack of such an interpretive restriction, along with the lack of agreement, in (8a), means that that example is not an example of either CR or STS raising.

If our analysis of (9) is on the right track, and CR is available in Persian, there is an interesting consequence for raising sentences with 3SG subjects. Given that a null *pro* is a possible realization of the embedded clause subject, derivational ambiguity arises:

(10) a. bachche$_i$ be-nazar mi-yâ-d [(ke) t$_i$ xaste bâsh-e]
child to-view ASP-come-3SG that tired SUBJ-be-3SG
'The child seems tired.'

b. bachche$_i$ be-nazar mi-yâ-d [(ke) *pro*$_i$ xaste bâsh-e]
child to-view ASP-come-3SG that tired SUBJ-be-3SG
'The child seems like they are tired.'

(10a) shows the analysis predicted by Karimi: *bachche* undergoes A′-movement from the embedded clause to the left periphery of the matrix clause, and 3SG agreement is co-incidental. (10b) is the CR derivation for the same string, where *bachche* merges directly in the subject position of the higher clause, and agreement is actually triggered by the ϕ-features of the subject. The only distinction would be at the level of interpretation: a sentence with the structure in (10b) would have a requirement of direct observation of the child, however this requirement would never be an absolute over the string in (10), as the alternative derivation in (10a) is always available. In the next section, we present evidence for this derivational ambiguity arising from two different psycholinguistic studies where strings such as those in (10) are used.

3. Psycholinguistic evidence

In this section, we report of the results of two studies which included sentences containing Persian raising structures. It should be noted that in both cases, these were not the main point of the study. Indeed, the trials were initially intended to function as control trials, as the existing literature makes clear categorical predictions regarding participant responses. In both cases, what we find is that when a trial contains a sentence with 3SG agreement on the raising verb, along with a displaced 3SG subject from the embedded clause, participants do not give the predicted responses. Our explanation is that the derivational ambiguity described above is the source of these unexpected responses.

3.1 Pairwise rating of sentences

The first study is a task involving the pairwise rating of sentences, conducted as a stimulus normalization task for the project described in Nakhaei (2019). Participants are presented with two sentences on screen, forming a minimal pair. The sentences are displayed as opposing ends of a 7-pt Lickert scale. Participants are instructed to rate the sentence pair as 4 (the centre of the scale) if there is no difference in acceptability between the members of the minimal pair. If there is a difference, participants should move along the scale toward the sentence they find most preferable. In the main portion of the study, participants were evaluating different subjects for a given verb, so that there would be no semantic interference in a later study on pronominal reference resolution. Raising sentences were included as distractor items in this study, along with sentences with very straightforward agreement mismatches, and sentences involving the placement of negation and the NPI *hichkas*.

The raising sentences used all had the predicate *be nazar âmadan*, inflected for 3SG agreement, as in (10). Five different combinations of embedded clause configurations were chosen, and minimal pairs were constructed on the basis of whether or not the subject or object of the embedded clause moved up. Thus, there are five schematic types of minimal pair, as in (11):

(11) a. 3SGSUB be-nazar mi-yâ-d [t_i 3SGOBJ] vs
3SGOBJ be-nazar mi-yâ-d [3SGSUB t_i]

b. 3SGSUB be-nazar mi-yâ-d [t_i 3PLOBJ] vs
3PLOBJ be-nazar mi-yâ-d [3SGSUB t_i]

c. 3PLSUB be-nazar mi-yâ-d [t_i 3SGOBJ] vs
3SGOBJ be-nazar mi-yâ-d [3PLSUB t_i]

d. 3SGSUB be-nazar mi-yâ-d [t_i 1SGOBJ] vs
1SGOBJ be-nazar mi-yâ-d [3SGSUB t_i]

e. 1SGSUB be-nazar mi-yâ-d [t_i 3SGOBJ] vs
3SGOBJ be-nazar mi-yâ-d [1SGSUB t_i]

To make the original positions of the arguments (SUBJ = embedded subject, OBJ = embedded object) clear, all objects were marked with the object marker *-ro*. A sample pair is given in (12):

(12) a. man$_i$ be-nazar mi-yâ-d [t_i muzik-ro mi-shnav-am]
I to-view ASP-come-3SG music-OBJ ASP-hear-1SG
'I seem to hear music.'

b. muzik-ro$_i$ be-nazar mi-yâ-d [man t_i mi-shnav-am]
music-OBJ to-view ASP-come-3SG I ASP-hear-1SG
'I seem to hear music.'

This represents the schema shown in (11e), where a first person singular subject is paired with a third person singular object. This pairing is notable in that it is one of the two (the

other being (11c)) where both options can only be derived by A′ movement. The rest, where there is a 3SG subject in the embedded clause, have one form which can be derived either by A′-movement, or by CR. However, given that A′-movement is always available for all sentences, we should expect no deviation from 4 on the ratings: all sentences are perfectly grammatical in each pair. Within each pair, displays were counterbalanced such that the sentence appearing with the embedded subject at the beginning of the sentence appeared on the left hand side of the scale once. However, results are normalized for analysis such that 1 always represents the option with the embedded subject appearing in the matrix clause, while 7 represents the version where the embedded object is displaced.

In this study, 20 native Persian speakers judged two trials in each of the five sentence configurations. Before analysis, all data points with a response time greater than 2.5 standard deviations from the mean response time across the whole study were removed; this removal process means that there is not an even number of responses for each of the conditions. The response patterns for each of the configurations are given in Figures 1 through 5. As is clear, the responses are quite different from the predicted results.

For the statistical analysis, we performed a one-sample t-test, checking to see whether the mean of all responses in each condition is significantly different from 4. Only one configuration, (11b) yields a result that is significantly different from 4 ($p < 0.001$). This would seem to suggest that by and large the predictions of the A′-movement analysis are confirmed.

However, examining Figures 1 through 5, we see that only in the configurations (11b), (11c), and (11e) is the mode ever 4. This is doubly surprising, as we find that 4 is not the mode in two conditions where the mean was not significantly different from 4 (those being (11a) and (11d)), and we find that 4 *is* the mode in the one condition where the mean is significantly different from 4. We discuss these different situations in turn below.

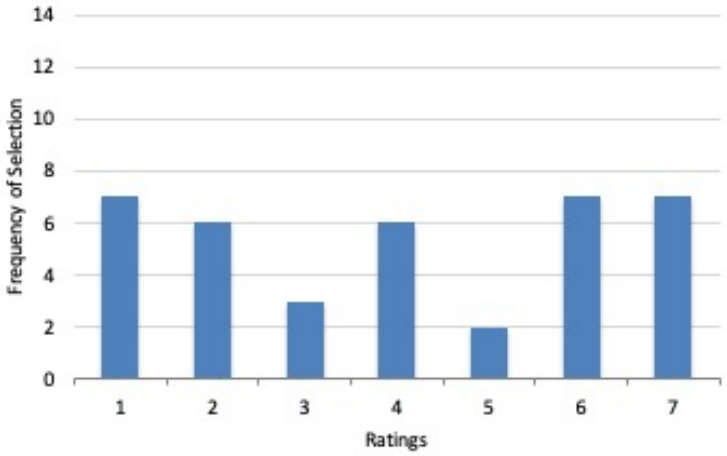

Figure 1. Distribution of responses for the configuration in (11a): 3SGSUB-3SGOBJ

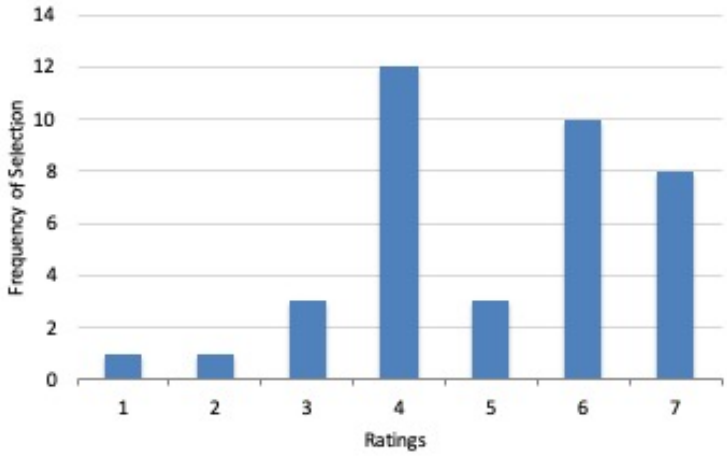

Figure 2. Distribution of responses for the configuration in (11b): 3SGSUB-3PLOBJ

For (11a), the configuration where both arguments are 3SG, there is a great deal of inconsistency between participants, with responses 1, 2, 4, 6, 7 all having either 6 or 7 selections. This symmetry explains how the mean is not significantly different from 4, though the true pattern is that the bulk of responses fall to one end or the other of the scale. A more extreme version of this pattern is seen in Figure 4 for (11d), where 13 of the responses are either 1 or 2, while 15 are either 6 or 7. Though the neutral response 4 is actually one of the most infrequently selected, the symmetry in the responses leads to a central mean. While this balanced response might be indicative of the A′-movement analysis, where there seems to be equal chance of selecting either configuration as acceptable, it would remain unclear as to why participants would not express this equal chance through the selection of the middle of the scale. Instead, what seems to be the case is that the two configurations are somehow different from each other, but participants do not respond uniformly.

Another clear case of a non-uniform response is in the one situation where there is a mean that is significantly different from 4, the configuration in (11b), shown in Figure 2. Here, the mode is actually 4, with 12 responses, though this is outweighed by a large preference toward the version of the sentence where the embedded clause object has been displaced to the matrix clause. To arrive at a potential explanation for this, it is necessary to first look at the three configurations discussed so far in more detail.

Firstly, these are the three where the embedded clause subject is 3SG. This means that the version of the sentence where the embedded clause subject appears in the matrix clause is derivationally ambiguous, arising either from A′-movement, with "accidental" agreement, or arising from CR with an additional semantic restriction on the subject. As such, there may be more considerations involved in the rating of the sentences than simple grammatical structure. Furthermore, in the case of (11d), there is a person asymmetry. Typologically, clauses with third person agents acting upon first person themes are somewhat

marked, and thus there is the added complexity of the version of (11d) where the embedded object appears in the matrix clause yielding a less-marked first person-initial structure. This competition of factors might explain the polarization of results versus the relative flatness of the results for (11a), where both arguments are 3SG. A different contrast again is manifested in (11b), where participants may be weighing the distinction between the semantic restrictions of CR versus the fronting of a plural object. The preference for displacing the embedded object, which we still maintain is derivable only by A′-movement, is unusual, as this is the costliest of all possible derivations.

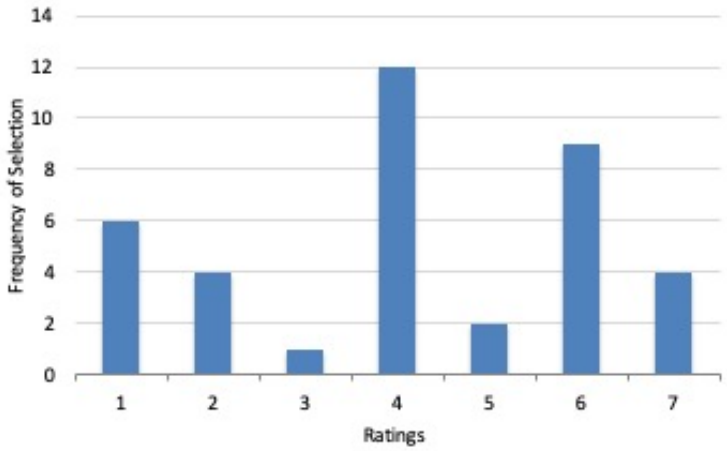

Figure 3. Distribution of responses for the configuration in (11c): 3PLSUB-3SGOBJ

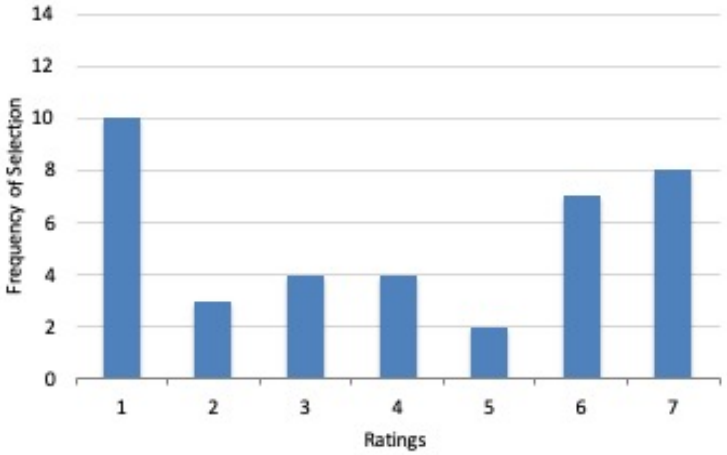

Figure 4. Distribution of responses for the configuration in (11d): 3SGSUB-1SGOBJ

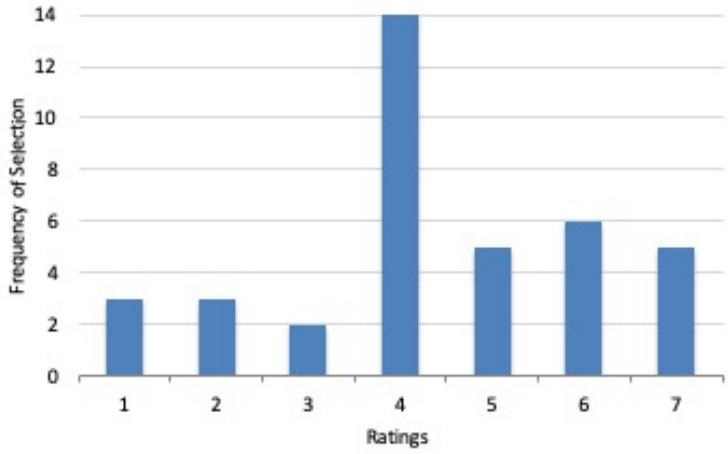

Figure 5. Distribution of responses for the configuration in (11e): 1SGSUB-3SGOBJ

Something similar seems to be at play in Figure 3 for (11c). Note that the derivational ambiguity is not at play here: the embedded clause subject is 3PL, and matrix agreement makes it clear that A′-movement is the only possible derivation. While the embedded object is 3SG, it still carries object marking in the displaced version of the sentence. Though there remains a slight preference toward displacing the object, removing the derivational ambiguity has allowed 4 to emerge as a clear mode. This is even clearer in Figure 5, showing the results for (11e): when the derivational ambiguity is removed, a pattern more consistent with the results of A′-movement emerges.

While it is clear that there are additional factors, such as markedness of different configurations of person features at play in these sentence pairs, there is one last piece of evidence supporting our contention that these results suggest that at least some speakers are considering CR as a possible derivation for the configurations in (11a), (11b), and (11d). This lies in the very mixed response to (11a), where the person and number features are held constant. Under an A′-movement analysis, there should be no interfering factors making one sentence more preferable over the other, and yet there is a clear split. This split, we believe, is the result of the addition of a CR derivation into the mix of factors to be considered when rating the relative acceptability of the sentences. We do not find direct evidence for the CR analysis of these sentences when the subject is displaced, but we do obtain results which are hard to explain if A′-movement is the only possible derivation. A similar situation arises in the next sub-section.

3.2 Self-paced reading

The results here arise from trials which are a part of the self-paced reading study reported in Abdollahnejad (2020). There, self-paced reading is used to determine whether or not there is evidence for movement in Persian ditransitive clauses where a definite direct object

appears to the left of a PP indirect object. As discussed in Featherston (2001), self-paced reading can be used as a tool to detect movement gaps. Put simply, participants should noticeably slow down in their reading when they process a movement gap.

Much as in the earlier study, the Abdollahnejad study uses *be nazar âmadan* sentences as distractor items which should have predictable results. To fit the overall study schema of sentences divided into seven regions, including PP indirect objects, the raising sentences are themselves embedded:

(13) Sinâ be Shirin | goft | ke | Bahrâm$_i$ | be-nazar mi-yâ-d | [t$_i$ az sharik-e
Sina to Shirin said that Bahram to-view ASP-come-3SG of partner-EZ
tejari-sh | jodâ be-sh-e]
commercial-3SG separate SUBJ-become-3SG
'Sina told Shirin that Bahram seems to separate from his business partner.'

(13) shows the moved condition of the sentence, where the subject of the most deeply embedded clause, *Bahrâm* appears to the left of *be-nazar mi-yâ-d*.[2] If this sentence is indeed derived by A′-movement, then readers should be aware of the movement upon encountering the region 5 of the sentence containing the raising verb. This should be manifested in a significant slowdown either in region 5 or 6, as participants might anticipate the movement gap to be immediately following the raising verb, but they should definitely be aware of it by the time the embedded clause arguments are being processed. As a basis of comparison, sentences with the structure of (14) are also included:

(14) Sinâ be Shirin | goft | ke | be-nazar mi-yâ-d | [Bahrâm$_i$ | az sharik-e
Sina to Shirin said that to-view ASP-come-3SG Bahram of partner-EZ
tejari-sh | jodâ be-sh-e]
commercial-3SG separate SUBJ-become3SG
'Sina told Shirin that Bahram seems to separate from his business partner.'

Here, the fourth and fifth regions are reversed, indicating that no movement has taken place. We should thus expect some combination of the following to occur if the only way that *Bahrâm* gets to the left of *be-nazar mi-yâ-d* in (13) is via A′-movement: either there will be a difference in the length of time taken to process the *be-nazar mi-yâ-d* regions of both sentences, or we should observe some significant difference in region 6, where the structure with displacement has a significantly longer reading time. To bring this result into sharper focus, we plot residual reading times in Figure 6. That is, instead of the mean reading time for each region, we first calculate an average reading time for each participant using their overall speed across all trials, and then plot mean residual reading times at each region.

As shown in Figure 6, the simplest description of the results is simply that the reading times for regions 4 and 5 appear to have been reversed. Statistical analysis using linear

[2]Descriptions 'left' and 'right' here are used in relation to the romanized examples. In the self-paced reading study, participants read right-to-left in Persian orthography.

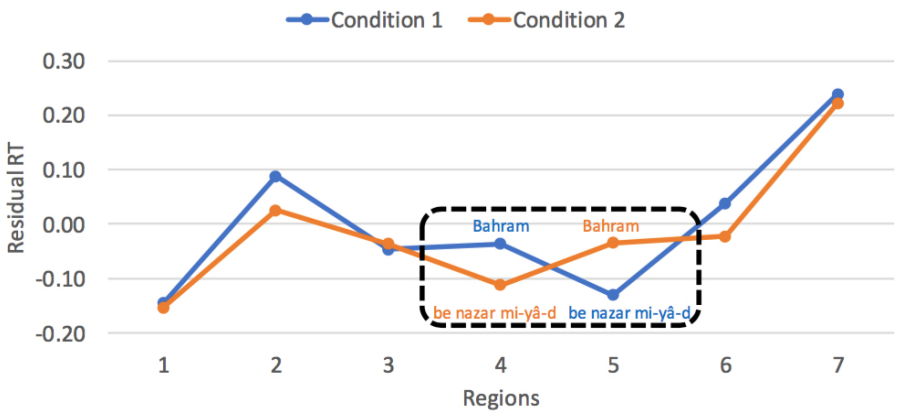

Figure 6. Residual reading times for *be-nazar mi-yâ-d* in the Abdollahnejad (2020) self-paced reading study. Condition 1 refers to the "moved" structure in (13), while Condition 2 refers to the (14) structure where no element of the most deeply-embedded clause appears to the left of the raising verb.

mixed effects modelling (Bates et al. 2015) shows no significant difference between the *be-nazar mi-yâ-d* regions between conditions (p=0.86), and also no significant difference between region 6 in either condition (p=0.311). In short, there is no evidence for a movement trace; both sentences are processed as though *Bahrâm* is merged directly into its surface position. It should also be noted that in the main portion of the study, this methodology did detect movement traces in ditransitives, corroborated by a separate online processing study of similar sentences. Lacking any evidence for movement, this is the clearest support for our claim that CR is available as an analysis for sentences such as (13).

4. Conclusion and future work

To recap, our claim is not that all sentences containing *be-nazar âmadan* in Persian are instances of CR. Rather, we claim that when this predicate is inflected for a subject that is not 3SG, the sentence is derived by CR, detectable from an additional semantic constraint on the subject, as described for (8b). This is similar to the observation that English contains both STS raising and CR, though there is a clear difference between the two in that English CR requires a preposition. This would make Persian CR somewhat unusual, in that there is no evidence for the preposition in the sentences we have tested. Below, we speculate on two potential avenues for exploration on how CR might arise in Persian.

First, we note that CR is present in Turkish (Moore 1998), and that the kind of double agreement seen in (8b) is also available in Turkish raising (Öztürk 2008). Given that sentences such as (8b) are not (to our knowledge) attested in the literature, more direct testing of their acceptability and possible demographic distribution is called for. It is possible that contact with languages that have CR is leading to the development of a Persian parallel.

Secondly, we note that there is inconsistency in the discussion of Persian raising in syntactic literature. While there is broad agreement on the empirical facts around agreement and optionality described in (1)-(2), the form of the raising predicate varies, as does the analysis. Darzi (1996) argues for an analysis closer to STS raising, using examples with the raising predicate *be-nazar residan* 'to arrive to view'. Ghomeshi (2001) argues for an A′ scrambling analysis of Persian raising, but lists five possible verbs, giving both *âmadan* and *residan* as possible light verbs with *be-nazar*. The other predicates are more modal in flavour: *lâzem budan* 'to be necessary', *momken budan* 'to be possible' and *ehtemâl dâshtan* 'to be possible'. More interestingly though, Ghomeshi provides another structure with an English gloss of 'seems' that uses none of these. This occurs with the phrase *mesl-e in-e ke* 'it is like this that', which is described as a frozen single word:

(15) dust-â-mun mesl-in-ke inǰâ-an
friend-PL-1PL.CL seems here-3PL
'It seems/looks as if our friends are here.' (Ghomeshi 2001)

This would appear to be a frozen CR construction, containing the prepositional element *mesl*. Interestingly, the English gloss here employs a CR structure with *as if*, while *be-nazar mi-yâ-d* is glossed as *seems*. Given that there seems to be some variation in the choice of light verb for *be nazar*, and evidence for CR in a similar construction that Ghomeshi describes as occupying the string position of a raising predicate, it is possible that a reanalysis underlies the acceptance of (8b) among our consultants. This in turn could make sentences such as (10), with displaced 3SG subjects before *be-nazar mi-yâ-d*, derivationally ambiguous, explaining the unexpected results in our studies.

If this is the case, then we would expect that careful construction of contexts in a self-paced reading study might resolve the ambiguity. Specifically, if it were clear in context that the speaker of a sentence such as (10) had not directly observed the child, then the perceptual source requirement on the subject would be violated, and the A′-movement analysis should once again be the only possible analysis. This, along with more direct testing for the acceptability of sentences such as (8b) is left for future work.

References

Abdollahnejad, Elias. 2020. *Online and offline examinations of constituent order in Persian ditransitives*. Ph.D. thesis, Univeristy of Calgary.

Asudeh, Ash and Ida Toivonen. 2012. Copy raising and perception. *Natural Language and Linguistic Theory* 30: 321–380.

Bates, Douglas, Martin Maechler, Ben Bolker, and Steve Walker. 2015. Fitting linear mixed-effects models using *lme4*. *Journal of Statistical Software* 67(1): 1–48.

Darzi, Ali. 1996. *Word order, NP movements, and opacity conditions in Persian*. Ph.D. thesis, University of Illinois at Urbana-Champaign.

Featherston, Sam. 2001. *Empty categories in sentence processing*. Amsterdam: John Benjamins.

Ghomeshi, Jila. 2001. Control and thematic agreement. *Canadian Journal of Linguistics* 46(1/2): 9–40.

Karimi, Simin. 2018. Generative approaches to syntax. In *The Oxford handbook of Persian linguistics*, ed. Anousha Sedighi and Pouneh Shabani-Jadidi. Oxford: Oxford University Press, 161–204.

Landau, Idan. 2011. Predication vs. aboutness in copy raising. *Natural Language and Linguistic Theory* 29(3): 779–813.

Moore, John. 1998. Turkish copy-raising and A-chain locality. *Natural Language and Linguistic Theory* 16(1): 149–189.

Nakhaei, Mahyar. 2019. *The interplay of memory and sentence structure on the resolution of persian pronouns*. Master's thesis, Univeristy of Calgary.

Öztürk, Balkiz. 2008. Non-configurationality: Free word order and argument drop in Turkish. In *The limits of syntactic variation*, ed. Theresa Biberauer, vol. 132 of *Linguistik Aktuell*. Amsterdam: John Benjamins, 411–440.

Potsdam, Eric and Jeffrey T. Runner. 2001. Richard returns: Copy raising and its implications. In *Papers from the 37th regional meeting of the Chicago Linguistic Society*, ed. Mary Andronis, Christopher Ball, Heidi Elston, and Sylvain Neuvel. Chicago: Chicago Linguistic Society, 453–486.