Syntactic and Semantic Constraints on Pronoun and Anaphor Resolution in Persian

Elias Abdollahnejad & Dennis Ryan Storoshenko

University of Calgary, Alberta

The First North American Conference in Iranian Linguistics

Stony Brook University

28-30 April, 2017

The Issue

What is the mechanism of reference resolution for the colloquial pronoun *un* "(s)he" and the anaphor (reflexive) *xod-eš* "self-3sg"?

Pronoun *un* shows a clear Condition B effect, while the behaviour of anaphor *xod-eš* is more unexpected.

1. **sohrab**$_i$ be **araš**$_j$ goft [ke   **mina**$_k$ un$_{i/i/*k}$ /xod-eš$_{i/i/k}$ =ro dust dare].

  S   to A   said that mina (s)he /self-PC.3sg =OM like have

  'Sohrab said to Arash that Mina likes her-him/self.'

Main Claims

- While both forms (pronoun & reflexive) can appear in overlapping environments, and are subject to some of the same constraints, the reference resolution mechanisms for *un* and *xod-eš* are different.

  - Pronoun *un* functions as a "standard" co-referential pronoun, drawing its reference from context alone.

$$[\![...un_1...]\!]^{g[1\to x]}$$

  - Reflexive *xod-eš* shows some hallmarks of a bound variable.

$$[\![...\lambda x_1...xod\text{-}eš_1...]\!]$$

Outline

1) Constraints on Reference Resolution

2) Behaviour and Features of Persian Reflexives

3) Semantics (source vs. perceiver) and Reference Resolution

4) Conclusion and Future Work

Outline

1) **Constraints on Reference Resolution**

2) Behaviour and Features of Persian Reflexives

3) Semantics (source vs. perceiver) and Reference Resolution

4) Conclusion and Future Work

Syntax vs. Semantics

- **Syntactic** and **Semantic** information have been argued to have different determining values in reference resolution. (Kuno, 1987; Tenny, 2003; Kaiser et. al. 2009; among others)

  - Within clause: Syntax > Semantics
  - Across clauses: Overlap
  - Between sentences: Syntax < Semantics

Roles and Hierarchies

- Other than the pure syntactic theories, **binding relations** have also been argued for in terms of argument **hierarchies** (e.g. HPSG) or specific maps to **theta roles** (e.g. Arnold 2001).

- Specific relations have been discussed to be more relevant to specific types of anaphors:

  - Preference for **source** of information as antecedent of **reflexives** (Kuno, 1987),
  - Preference for **perceiver** of information as the antecedent of **pronouns** (Tenny, 2003),

Form-Specific Multiple-Constraints Framework (Kaiser, 2003; Kaiser & Trueswell, 2008)

- The *either-or* classification of reference resolution based on **structural** or **non-structural** constraints is an "oversimplification" (Kaiser et. al. 2009).

- **Form-specific multiple-constraints framework:** "anaphor resolution is the result of the interaction of multiple constraints" guiding "reference resolution to be **weighted differently** for different referential forms" (Kaiser et. al. 2009, p. 56).

Form-Specific Multiple-Constraints Framework (Kaiser et. al., 2009)

- Kaiser et. al. (2009) used the verb to manipulate the source/perceiver status of the subject and object in English sentences with PNPs.

source
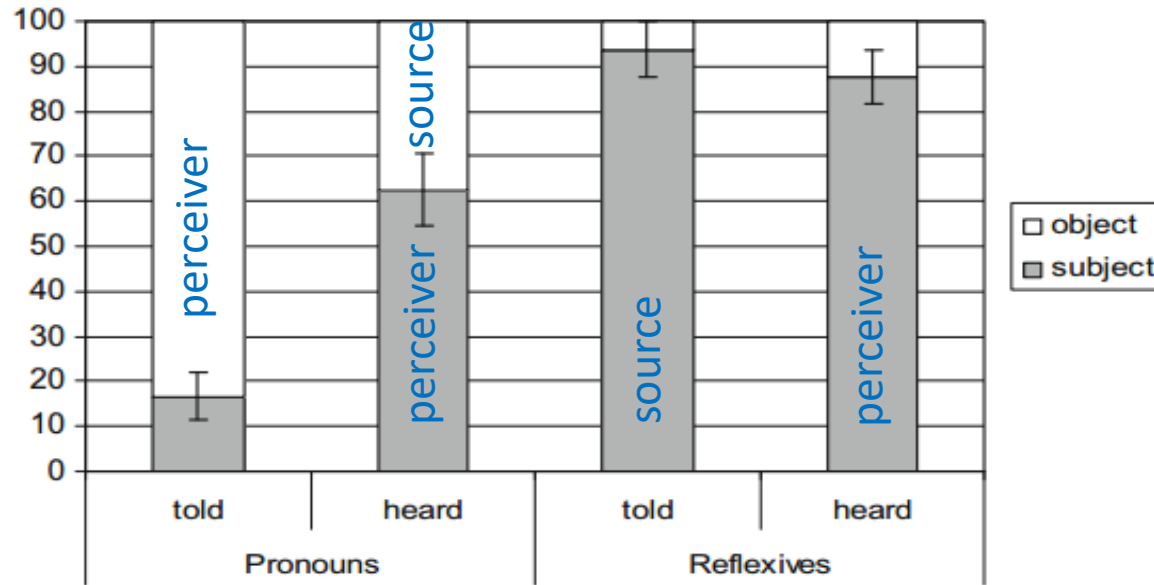4. Peter **told** Andrew about the picture of {**him**/**himself**} on the wall.
source
5. Peter **heard from** Andrew about the picture of {**him**/**himself**} on the wall.

- They used this contrast to measure the effects of structure vs. semantic roles,

Implications of This Framework (Kaiser et. al., 2009)

Percentage of subject-picture choices and object-picture choices

- For reflexives, structure is equally important regardless of the semantics,

- For pronouns neither bias fully determines the result,

- The weight of biases is different for each form,

Outline

Reflexives in Persian

- Moyne (1971) uses Persian to discuss distinct syntactic and semantic features of reflexive and emphatic elements.

- Mahootian & Gebhardt (1997):
  - Reflexive pronouns are **anaphoric** with two main functions:
    - ①   indicating coreference of object and subject,
    - ②   serving as intensifiers,
  - In Persian, "the scope of reflexivity is **not restricted to the clause**", i.e. antecedent and reflexive can occur in separate clauses (p. 96).

Persian Reflexive Forms

- Reflexivity appears in two forms in Persian:

i. The simplex expression *xod* 'self' with all persons and numbers (used more in formal and written context),

ii. *xod* plus a Pronominal Clitic (*xod-PC*). The clitic determines the number and person of the reflexive element (used more in colloquial and informal language),

Pronouns and Anaphors in Persian

Binding

6. [**minā**ᵢ un*ᵢ /ₖ ⁄ xod-ešᵢ/?*ₖ ⁄ xodᵢ/*ₖ =ro   moarefi   kard].

   mina (s)he / self-PC.3sg/ self   =OM introduce did

   'Mina introduced her-him/self.'

7. **sohrāb**ⱼ goft   [ke   **minā**ᵢ un*ᵢ /ⱼ /xod-ešᵢ/ⱼ   /xodᵢ/*ⱼ =ro   dust dāre].

   sohrab say.3sg.past that mina (s)he /self-PC.3sg/ self   =OM like   have

   'Sohrab said that Mina likes her-him/self.'

14

Diagnostics for Anaphors

Bound variables tend to have sloppy reading with <u>quantifier</u> antecedents.

8. hær-kæsi$_i$ xod-eš$_i$ / xod$_i$ =ro   dust dâre.

   Every-body self-PC.3sg / self =OM like   have

   'Everybody likes self.'

Sloppy reading: ∀ x [x likes x] = John likes John, Bill likes Bill, …

Strict reading: Everybody likes the very same person.

- Both forms of reflexive are bound variables.

Diagnostics for Anaphors

VP Ellipsis

- If after VP ellipsis only the sloppy reading is possible (not the strict reading) the reflexive element is a bound variable (not a free variable).

9. sohrab$_i$ xod-eš$_i$ / xod$_i$ -ro   dust dâre, væli sârâ$_j$ ~~xod-eš~~$_j$ ~~/xod~~$_j$ ~~=o   dust~~ næ-dâre.

   S.   self-PC.3sg / self -OM like   have but Sara self-PC.3sg/self=OM like neg-have

   'Sohrab likes self, but Sara doesn't ~~like self~~.'

Sohrab likes Sohrab,

=but Sara doesn't like Sara. (sloppy reading) Bound Variable (**preferred**)

=but Sara doesn't like Sohrab. (strict reading)   Free variable

Diagnostics for Anaphors

- One additional piece of evidence that *xod-eš* is subject to binding requirements is that when two instances of *xod-eš* occur in the same sentence, they must co-refer.

10. sohrāb$_i$ be mina$_j$ goft ke mi-dune [ke faqat mādar-e xod-eš$_i$ hičvaqt xod-eš$_i$ =o tanhā ne-mi-zāre].

xod-eš$_j$ xod-eš$_j$

S.   to M.   said that DUR-know that only mother-EZ self-3sg never self-3sg =OM alone neg-DUR-put.

'Sohrab$_i$ said to Mina$_j$ that he knows that only self$_i$'s mother does not leave self$_i$ alone.'

self$_j$'s mother   self$_j$

- They are bound by the same (lambda) binder.

Outline

Constraints in Persian

11. source   perceiver

sohrāb$_i$ **be** arash$_j$ **goft** [ke   minā$_k$ hatman   bā   un$_{i/j/*k}$ / xod-eš$_{i/j/\#k}$ _tamās mi-gire_].

S.   to A.   said that M.   certainly with (s)he / self-3sg   contact DUR-get

'Sohrab **said to** Arash that Mina will certainly contact him/self.'

12. perceiver   source

sohrāb$_i$ **az**   arash$_j$ **šenid** [ke   minā$_k$ hatman   bā   un$_{i/j/*k}$ / xod-eš$_{i/j/\#k}$ _tamās mi-gire_].

S.   from A.   heard that M.   certainly with (s)he / self-3sg   contact DUR-get

'Sohrab **heard from** Arash that Mina will certainly contact him/self.'

- Does the manipulation of source/perceiver status of subject and object cause any difference in the likelihood of potential antecedents?

Constraints in Persian

11. sohrāb$_i$ **be** arash$_j$ **goft** [ke   minā$_k$ hatman   bā   xod-eš$_{i/j/\#k}$ _tamās   mi-gire_].

    S.   to A.   said that M.   certainly with   self-3sg   contact DUR-get

    'Sohrab **said to** Arash that Mina will certainly contact him/self.'

12. sohrāb$_i$ **az**   arash$_j$ **šenid** [ke   minā$_k$ hatman   bā   xod-eš$_{i/j/\#k}$ _tamās   mi-gire_].

    S.   from A.   heard that M.   certainly with   self-3sg   contact DUR-get

    'Sohrab **heard from** Arash that Mina will certainly contact him/self.'

Reflexive: more likely to be bound by matrix subject.

- Strong influence of syntactic information on reflexives in general,

- Weak Subject Orientation as a violable preference for subject antecedents (Sohng 2004),

Constraints in Persian

perceiver

11. sohrāb$_i$ **be** arash$_j$ **goft** [ke   minā$_k$ hatman   bā   un$_{i/j/*k}$ _tamās   mi-gire_].

   S.   to A.   said that M.   certainly with (s)he   contact DUR-get

   'Sohrab **said to** Arash that Mina will certainly contact him/self.'

perceiver

12. sohrāb$_i$ **az**   arash$_j$ **šenid** [ke   minā$_k$ hatman   bā   un$_{i/j/*k}$ _tamās   mi-gire_].

   S.   from A.   heard that M.   certainly with (s)he   contact DUR-get

   'Sohrab **heard from** Arash that Mina will certainly contact him/self.'

Pronoun: It seems that semantics (i.e. bias for perceiver) also plays a major role.

Picture Noun Phrases (PNPs)

- The same effect seems to exist for *un* and *xod-eš* inside PNPs acting as the direct object of mono-clausal ditransitives.

source   perceiver

13. sohrāb$_i$ [$_{DO}$ šaye-ye   jadid darmored-e un$_{i/j}$ / xod-eš$_{i/*j}$ =ro] [$_{IO}$ be arash$_j$] **goft**.

      S.   rumor-EZ new   about-EZ   (s)he / self-3sg   =OM   to A.   said

      'Sohrab **said to** Arash the new rumor about him/self.'

Picture Noun Phrases (PNPs)

- The same effect seems to exist for *un* and *xod-eš* inside PNPs acting as the direct object of mono-clausal ditransitives.

source

13. sohrāb$_i$ [$_{DO}$ šaye-ye   jadid darmored-e   xod-eš$_{i/*j}$ =ro] [$_{IO}$ be arash$_j$] **goft**.

    S.   rumor-EZ new   about-EZ   self-3sg =OM   to A.   said

    'Sohrab **said to** Arash the new rumor about him/self.'

Reflexive:

- Subject (source of info) is the possible antecedent.

- Indirect object cannot be coindexed with *xod-eš* (it does not c-command the reflexive).

Picture Noun Phrases (PNPs)

- The same effect seems to exist for *un* and *xod-eš* inside PNPs acting as the direct object of mono-clausal ditransitives.

perceiver

13. sohrāb$_i$ [$_{DO}$ šaye-ye   jadid darmored-e un$_{i/j}$ =ro] [$_{IO}$ be arash$_j$] **goft**.

   S.   rumor-EZ new   about-EZ   (s)he   =OM   to A.   said

   'Sohrab **said to** Arash the new rumor about him/self.'

Pronoun: subject and indirect object are both possible references (perceiver (object) more likely).

Picture Noun Phrases (PNPs)

perceiver   source

14. sohrāb$_i$ [$_{DO}$ šaye-ye   jadid darmored-e un$_{i/j}$ / xod-eš$_{i/?*j}$ =ro] [$_{IO}$ az   arash$_j$] **šenid**.

     S.   rumor-EZ new   about-EZ   (s)he / self-3sg   =OM   from A.   heard

     'Sohrab **heard from** Arash the new rumor about him/self.'

Picture Noun Phrases (PNPs)

perceiver

14. [sohrāb$_i$] [$_{DO}$ šaye-ye jadid darmored-e   xod-eš$_{i/?*j}$ =ro] [$_{IO}$ az   arash$_j$] **šenid**.

    S.   rumor-EZ new   about-EZ   self-3sg   =OM   from A.   heard

   'Sohrab **heard from** Arash the new rumor about him/self.'

Reflexive:

- Subject (perceiver) is the possible antecedent,
- Indirect object cannot (?) be coindexed with *xod-eš* (it does not c-command the reflexive).

26

Picture Noun Phrases (PNPs)

perceiver

14. sohrāb_i [_{DO} šaye-ye   jadid darmored-e un_{i/j}   =ro] [_{IO} az   arash_j] **šenid**.

      S.   rumor-EZ new   about-EZ   (s)he   =OM   from A.   heard

     'Sohrab **heard from** Arash the new rumor about him/self.'

Pronoun: subject and indirect object are both possible references (perceiver (subject) more likely).

Outline

1) Constraints on Reference Resolution

2) Behaviour and Features of Persian Reflexives

3) Semantics (source vs. perceiver) and Reference Resolution

4) **Conclusion and Future Work**

General Conclusion

- Overall, the picture is that *un* and *xod-eš* are subject to different but possibly overlapping sets of constraints with different weights for each.

- To account for the inconsistent behaviour of *xod-eš*, we leave for future work the possibility that there may be semantically different but homophonous forms of *xod-eš* (c.f. Anand (2006) for Mandarin *ziji*),

Future Work

- Having concluded that *xod-eš* is indeed a bound anaphor and not merely a pronoun, we will conduct experiments to determine the relative weights of the binding constraints.

- Visual world paradigm eye tracking,

Future Work

We will pursue further tests for logophoricity, based on the distinction in 15 and 16 (c.f. Anand 2006).

15. sohrāb$_i$ fekr kard [ke arash$_j$ be **un**$_i$ gofte [ke māšin-e xod-eš$_{i/j}$ =o dozdid-an]].

S. thought did that A. to (s)he said that car-e self-3sg =OM stole-3PL

'Sohrab$_i$ thought that Arash$_j$ has said to **him**$_i$ that they have stolen self$_{i/j}$'s car.'

16 sohrāb$_i$ fekr kard [ke arash$_j$ be **pedar-e** **un**$_i$ gofte [ke māšin-e xod-eš$_{i/j}$ =o dozdid-an]].

S. thought did that A. to father-EZ (s)he said that car-EZ self-3sg =OM stole-3PL

'Sohrab$_i$ thought that Arash$_j$ has said to **his**$_i$ **father** that they have stolen self$_{i/j}$'s car.'

Take Home Message

- Literature on binding in Persian is still relatively scant – Good news: lots of work to do!

- The data are subtle, and the possible analyses quite complex (as multiple factor analyses seem likely).

- All of this must be kept in mind before binding is used as a diagnostic for syntactic structure (locality and maybe even c-command).

Thank You

سپاس

Bibliography

- **Abdollahnejad, E**. (2016). Reflexivity in Persian. In L. Hracs (Ed.), *The proceedings of the Canadian Linguistic Association Annual Conference*, Calgary, Canada. Retrived from: http://cla-acl.ca/actes-2016-proceedings.

- **Anand, P.** (2006). *De de se*. Doctoral dissertation, MIT.

- **Arnold, J. E.** (2001). The effect of thematic roles on pronoun use and frequency of reference continuation. Discourse Processes, 31(2), 137-162.

- **Kaiser, E**. (2003). *The quest for a referent: A crosslinguistic look at reference resolution.* Doctoral dissertation, Philadelphia, PA: University of Pennsylvania.

- **Kaiser, E., & Trueswell, J. C.** (2008). Interpreting pronouns and demonstratives in Finnish: Evidence for a form-specific approach to reference resolution. Language and Cognitive Processes, 23(5), 709-748.

- **Kaiser, E., Runner, J. T., Sussman, R. S., & Tanenhaus, M. K.** (2009). Structural and semantic constraints on the resolution of pronouns and reflexives. *Cognition, 112*(1), 55-80.

- **Kuno, S.** (1987). *Functional syntax: Anaphora, discourse and empathy.* Chicago: University of Chicago Press.

- **Mahootian, Sh, & Gebhardt. L**. (1997). Persian: Descriptive grammars, London: Routledge.

- **Moyne, J. A.** (1971). Reflexive and emphatic. Language, 141-163.

- **Sohng, Hong-Ki,** (2004). A minimalist analysis of X0 reflexivization in Chinese and Korean. *Studies in Generative Grammar*, 14(3): 375–96.

- **Tenny, C.** (2003). Short distance pronouns, argument structure, and the grammar of sentience. *Manuscript, Carnegie Mellon University*.