

# Modal syntax cuts short the claim that modern Persian lacks apocopated infinitives

Setayesh Dashti      Ash Asudeh  
University of Oxford    University of Rochester

LFG 2023  
University of Rochester

July 23, 2023

## 1 Background<sup>1</sup>

- Persian is an SOV Indo-European language with ‘*pro*-drop’.
  - Verbal morphology follows a two-stem system, traditionally called
    1. *Present stem*: no overt present tense marker; and
      - (1) *xor* ‘eat’
    2. *Past stem*: modulo suppletive patterns, the past tense is regularly marked with *-d* and its allomorphs (Anoushe 2018).
      - (2) *xord* ‘eat’
  - The present stem always occurs with either aspectual or mood markers; *mi-* for imperfective aspect (3a) and *be-* for subjunctive mood (3b).<sup>2,3</sup>
  - The unprefixd past stem with agreement suffixes is used to show the perfective aspect (3c).
  - Past imperfective, progressive and perfect are also derived from the past stem with agreement suffixes; for example, past imperfective is formed with the same prefix as present imperfective, *mi-* (3d).<sup>4</sup>
- (3) a. Nika be madrese mi-rav-ad.  
Nika to school IPFV-go.PRES-3SG  
‘Nika goes to school.’
- b. Nika šāyad be madrese be-rav-ad.  
Nika may to school SBJV-go.PRES-3SG  
‘Nika might go to school.’
- c. Nika be madrese raf-t.  
Nika to school go-PAST.3SG  
‘Nika went to school.’
- d. bače-hā har ruz be madrese mi-raf-t-and.  
child-PL every day to school IPFV-go-PAST-3PL  
‘The kids used to go to school every day.’

---

<sup>1</sup>The dialect reported on here is colloquial spoken Persian, not the written standard.

<sup>2</sup>Glosses are abbreviated as follows: AUX–auxiliary, COP–copula, EZ–*ezafe* (nominal linker), IPFV–imperfect, INF–infinitive, NEG–negation, PP–past participle, PRES–present tense, PAST–past tense, SBJV–subjunctive mood, SG–singular, PL–plural, DO–direct object. We use the hyphen (-) to indicate an affix boundary and an equal sign (=) to indicate clitic attachment.

<sup>3</sup>What we have glossed as IPFV–imperfective is sometimes glossed as DUR–durative. We prefer to gloss it based on its morphological form rather than its typical morphosyntactic function.

<sup>4</sup>Past imperfective also functions as a fake past to convey counterfactuality, regardless of tense (Bjorkman and Halpert 2017).

- Persian contains several adverbial and complex predicate modals, but there are two main simplex verbal modal auxiliaries, *bāyestan* (necessity/□) and *šodan* (possibility/◇).<sup>5</sup>
  - These modals always appear in the default third person singular form: *bāyad* (□.PRES)/*bāyest* (□.PAST) and *mi-še* (IPFV-◇.PRES)/*mi-šod* (IPFV-◇.PAST).
  - They can either occur with:
    1. a finite complement (4), marked with subjunctive mood in present tense (4a) or imperfective aspect in past tense (4b); or
    2. a nonfinite complement (5)
  - In the second case, the verb in the complement has a simple past stem, which *resembles* the third person singular past inflection, but is historically an *apocopated infinitive* (short infinitive); importantly, it is interpreted as an impersonal (5).
- (4) a. *bāyad be xune be-rav-am.* (5) *bāyad zood be xune raf-t.*  
 □.PRES to home SBJV-go.PRES-1SG □.PRES early to home go-??  
 ‘I have to go home.’ ‘It’s necessary to go home early.’/
- b. *bāyad bačehā be xune mi-raf-t-and.* (5) *bačehā zood be xune mi-raf-t-and.*  
 □.PRES child-PL to home IPFV-go-PAST-3PL □.PRES child-PL to home IPFV-go-PAST-3PL  
 ‘The children had to go home.’ ‘One must go home early.’
- When the modal occurs with a finite complement, it is possible to topicalize the embedded subject to the left:<sup>6</sup>
- (6) a. *bāyad bačehā be xune mi-raf-t-and.* b. *bačehā bāyad be xune mi-raf-t-and.*  
 □.PRES child-PL to home IPFV-go-PAST-3PL child-PL □.PRES to home IPFV-go-PAST-3PL  
 ‘The children had to go home.’ ~ ‘As for the children, they had to go home.’
- A verb that works very similarly to the simplex modal verbs is *be nazar āmad-an* (lit. ‘to opinion come-INF’/‘to seem’), which is the equivalent of *seem*.
  - Unlike *bāyad* (□/necessity), which never inflects for agreement, some speakers allow both the non-agreeing/default form (7b) and the agreeing form (7c); all speakers allow the first, non-agreeing form:
- (7) a. *be nazar mi-ā-d ke bače-hā xaste šo-d-an.*  
 to opinion IPFV-come.PRES-3SG that child-PL tired become-PAST-3PL  
 ‘It seems that the children have gotten tired.’
- b. *bače-hā be nazar mi-ā-d ke xaste šo-d-an.*  
 child-PL to opinion IPFV-come.PRES-3SG that tired become-PAST-3PL  
 ‘As for the children, it seems that they have gotten tired.’
- c. % *bače-hā be nazar mi-ā-n ke xaste šo-d-an.*  
 child-PL to opinion IPFV-come.PRES-3PL that tired become-PAST-3PL  
 ‘The children seem to have gotten tired.’
- Note that in (7) we are emphasizing the colloquialness of the reported example by using certain spoken-only forms, such as the contracted 3PL in (7c): *-(a)n* rather than *-(a)nd*.

<sup>5</sup>There is some debate over the status of the modal *šāyestan*. Some literature, such as Karimi (2005) and Taleghani (2008), treats it as another modal auxiliary, while other literature, such as Labbafankhosh and Darzi (2015), treats it as a modal adverb.

<sup>6</sup>The subjunctive in the past tense has the same form as the imperfective.

## 2 Puzzles/questions

1. How should we account for the complement in (5)?

- (5) bāyad zood be xune raf-t.  
 □.PRES early to home go-??  
 ‘It’s necessary to go home early.’/ ‘One must go home early.’

Is it a past tense form or a short infinitive (synchronously as well as diachronically)?

2. How can we capture the impersonal and personal readings of modals like (5) vs. (4)?

- (4) a. bāyad be xune be-rav-am.                      b. bāyad bačehā be xune mi-raf-t-and.  
 □.PRES to home SBJV-go.PRES-1SG                      □.PRES child-PL to home IPFV-go-PAST-3PL  
 ‘I have to go home.’    ‘The children had to go home.’

3. What is the syntactic structure of simplex modal constructions?

4. How should the variable agreement displayed in (7) be explained?

- (7) b. bače-hā be nazar mi-ā-d                      ke xaste šo-d-an.  
 child-PL to opinion IPFV-come.PRES-3SG that tired become-PAST-3PL  
 ‘As for the children, it seems that they have gotten tired.’  
 c. % bače-hā be nazar mi-ā-n                      ke xaste šo-d-an.  
 child-PL to opinion IPFV-come.PRES-3PL that tired become-PAST-3PL  
 ‘The children seem to have gotten tired.’

## 3 The syntax of Persian modals

- The main challenge is the construction with an impersonal interpretation, as in (5), shown here with the complement structure made explicit, or the similar example (8):

- (5′) bāyad [zood be xune raf-t].  
 □.PRES [early to home go-??]  
 ‘It’s necessary to go home early.’/ ‘One must go home early.’

- (8) bāyad [šab-hā hašt sâ’at xāb-id].  
 □.PRES night-PL eight hour sleep-??  
 ‘It’s necessary to sleep for eight hours a night.’/ ‘One must sleep for eight hours a night.’

- This is often called the *impersonal construction* (IC; e.g. Karimi 2008).
- Let us assume that the complement is a clause, because there is apparent inflection on the embedded predicate.
  - The question is: what kind of clause?
    - Is it finite, as indicated by its shared form with the past tense (see, e.g., Karimi 2008)?
    - Is it infinitival, despite the shared form (Karimi 2005, Samvelian 2018)?
    - Perhaps it is not a clause at all,<sup>7</sup> but presumably some kind of nominal?

<sup>7</sup>For example Darzi and Kwak (2015: 2) write, “This language, like Greek and unlike English, lacks nonfinite clauses.”

### 3.1 Infinitives in Persian

- The morphosyntactically distinguished Persian infinitive is formed from the past stem and the suffix *-an*.

- (9) a. *dav-id-an barāye salāmati mofid-e*  
 run-PAST-INF for health beneficial-COP.3SG  
 ‘Running has health benefits.’
- b. *Nika dav-id-an=ro dust dār-e*  
 Nika run-PAST-INF=DO friend have-PRES.3SG  
 ‘Nike likes running.’

- Contrast these with similar impersonal complements:

- (10) a. *bāyad barāye salāmati dav-id*  
 □.PRES for health run-??  
 ‘One should run for health reasons.’
- b. *bāyad dav-id-an=ro dust dāsh-t*  
 □.PRES run-PAST-INF=DO friend have-??  
 ‘One must like running.’

#### 3.1.1 The syntax of Persian infinitives

- The forms in (9) and (10) do not have the same distribution.
- There is, indeed, a case for arguing that the long infinitives are nominals, whereas the other forms are short/apocopated infinitives forming nonfinite verbal clauses.

1. This immediately suggests a function for the suffix *-an*: it is a nominalizer.
2. Long infinitives can take modifiers with the nominal linker *-e* (*ezafe*). The infinitive in (11a) has a genitive modifier, the one in (11b) has an adjectival one.

- (11) a. *xand-id-an-e nowzād ārāmešbaxš-e*  
 laugh-PAST-INF-EZ baby calming-COP.3SG  
 ‘Babies’ laughing is calming.’
- b. *Ali dav-id-an-e ārum=ro be šenā kar-d-an-e sari’ tarjih mi-d-e*  
 Ali run-PAST-INF-EZ slow=DO to swim do-PAST-INF-EZ fast preference IMPF-give.PRES-3SG  
 ‘Ali prefers running slowly to swimming fast.’

3. It explains the fact that a) unbounded dependencies can be formed into the impersonal complement marked by the short infinitive, as shown in (6b) repeated here as (12a), since it is a clause, b) whereas unbounded dependencies cannot be formed into long infinitives as putative complex nominals (a):

- (12) a. *bačehā bāyad be xune mi-raf-t-and.*  
 child-PL □.PRES to home IPFV-go-PAST-3PL  
 ~‘As for the children, they had to go home.’

- (13) a. Sarina tanhā be bače šir dā-d-an=ro dust na-dāre  
Sarina alone to baby milk give-PAST-INF=DO friend NEG-have.PRES.3SG  
'Sarina doesn't like to breast-feed the child alone.'
- b. \* tanhā Sarina be bače šir dā-d-an=ro dust na-dāre  
alone Sarina to baby milk give-PAST-INF=DO friend NEG-have.PRES.3SG  
'Sarina doesn't like to breast-feed the child alone.'
- c. \* be bače Sarina tanhā šir dā-d-an=ro dust na-dāre  
to baby Sarina alone milk give-PAST-INF=DO friend NEG-have.PRES.3SG  
'Sarina doesn't like to breast-feed the child alone.'

- In sum, we have presented syntactic arguments for why the long infinitive is a nominal and for why the form in the impersonal construction is not.
- This already suggests that the impersonal complement is a nonfinite clause, as this is the obvious alternative to its being a nominal.
- In the next section, we will show that in fact the form in the impersonal is nonfinite, since it cannot take tense/aspect/mood-marking without losing its impersonal interpretation.
- The assumption that the form in the impersonal is a short infinitive that heads a clause explains why it does not take the *-an* suffix: this suffix is a nominalizer and the short infinitive is not a nominal.
- The fact that this looks superficially like the past form is a fact of morphosyntactic syncretism and nothing more, because the short infinitive does not *function* as a past form or have past meaning, as we will show next.

### 3.1.2 The function and interpretation of Persian infinitives

- The past finite complement of the modal should bear imperfective marking, but adding this marking to the sort of complement under discussion renders an impersonal reading unavailable and requires it to have a personal reading. Contrast (14a), repeated from (8) above, with (14b):<sup>8</sup>

- (14) a. bāyad šab-hā hašt sā'at xāb-id.  
□.PRES night-PL eight hour sleep-??  
'It's necessary to sleep for eight hours a night.'/'One must sleep for eight hours a night.'
- b. bāyad šab-hā hašt sā'at mi-xāb-id.  
□.PRES night-PL eight hour IPFV-sleep-PAST.3SG  
# 'It's necessary to sleep for eight hours a night.'/'One must sleep for eight hours a night.'  
✓ 'pro.3SG had to sleep for eight hours a night.'

- Persian is sometimes assumed to lack a nonfinite clause (Darzi and Kwak 2015), exactly because of the similarity in morphological form between the third singular past form, which is unmarked for agreement morphology (e.g., *raf-t* go-PAST.3SG) and the simple stem form in question (e.g., *raf-t* go-??).
- But, as we have just seen, this does not account correctly for the impersonal readings.
- We instead assume that this verbal form is infinitival and thus unmarked for TENSE/ASPECT/MOOD.

<sup>8</sup>Sentence (14b) can have another interpretation in which the subject of the verb is *pro*-dropped, which will translate to 'pro.3SG should have slept eight hours a night'. This is a different construction than the one in question; most importantly, the alternative construction is never impersonal.

- The future construction, shown in (15), provides further evidence for nonfiniteness of this verbal form, now glossed INF.
- This builds on Lowe's (2019) claim that nonfinite forms generally appear in periphrastic constructions as the lexical content of the clausal predicate.

(15) Ali farda be madrese xāh-ad raf-t.  
 Ali tomorrow to school want-3SG go-INF  
 'Ali will go to school tomorrow.'

- In the analysis section §4, a template (Dalrymple et al. 2004, Asudeh et al. 2013) is used to generate this defective/infinitival verbal form.

## Summary

- The first question posed in §2 above was *whether the morphologically ambiguous form in the impersonal modal construction is a past tense form or a short infinitive (apocopated infinitive)*.
- We argued that the form in question, just as in (15), is an infinitival form *synchronically*, and that the apocopated infinitive is morphologically formed by referral to the past stem, which explains their identity.
- However, the agreeing past tense form and the non-agreeing short infinitive have different functions and interpretations.<sup>9</sup>

## 4 An LFG analysis of Persian modal syntax

- LFG assumes a separation of syntax into two levels, c(onstituent)-structure and f(unctional)-structure.
  - C-structure represents syntactic distribution, via categories, constituency, hierarchy, and linear order.
  - F-structure represents relational aspects of syntax, such as grammatical functions, agreement, case-marking, as well as local (control/raising) and non-local (unbounded dependencies) relations.
- The following illustrate the c-structure position of the modal and the general structure of the CP and IP:

- (16) a. Mariam goft  
 Mariam said  
 $[_{CP} [_{C'} [_{C} \text{ke}] [_{IP} \text{kodoom ketab-ha-ro} [_{IP} [_{I'} [_{I} \text{bayad}] [_{VP} \text{bače-ha be-xun-and} ]]]]]]$   
 that which book-PL=DO must child-PL SBJV-read-3PL  
 'Mariam said that the children must read WHICH BOOKS?'
- b. Mariam goft  
 $[_{CP} [_{C'} [_{C} \text{ke}] [_{IP} \text{kodoom ketab-ha-ro} [_{IP} [_{I'} \text{bače-ha} [_{I'} [_{I} \text{bayad}] [_{VP} \text{be-xun-and} ]]]]]]]]$   
 that which book-PL=DO child-PL must SBJV-read-3PL  
 'Mariam said that the children must read WHICH BOOKS?'
- c. Mariam goft  $[_{CP} [_{C'} [_{C} \text{ke}] [_{IP} \text{kodoom ketab-ha-ro} [_{IP} \text{bače-ha} [_{I'} [_{VP} \text{xun-d-and} ]]]]]]$   
 that which book-PL=DO child-PL read-PAST-3PL  
 'Mariam said that the children read WHICH BOOKS?'
- d. Mariam goft  $[_{CP} [_{C'} [_{C} \text{ke}] [_{IP} \text{bače-ha} [_{IP} \text{kodoom ketab-ha-ro} [_{IP} [_{VP} \text{xun-d-and} ]]]]]]$   
 that child-PL which book-PL=DO read-PAST-3PL  
 'Mariam said that, as for the children, they read WHICH BOOKS?'

<sup>9</sup>The insight that the so-called past stem in these constructions is the apocopated infinitive is not novel (especially in the context of the future construction; Windfuhr 1979), but the theoretical literature seems largely to have taken it to be the PAST.3SG form of the verb (for instance, Karimi 2008, Mirrazi 2022).

- Example (16a) shows that there is a position for the top of an unbounded dependency below C, since the C position is occupied by an overt complementizer.
  - We assume that this position is an IP-adjunct, since otherwise the *wh*-phrase would be in regular subject position in SpecIP.
- Example (16b) shows that there is a position for an internal topic below this IP-adjunct position.
  - We postulate that this is an I'-adjunct.
  - Thus, in (16b), *bačeha* is in a non-agreeing topic position, reflected by the lack of plural agreement on the modal (which is not possible).
- Example (16c) shows that when an agreeing subject is present, in a simple case without a modal, it can be assumed to occur in the standard SpecIP subject position.
- Example (16d) shows the IP-adjunction in (16c) can be reversed freely, with a topicalized *bačeha* occurring adjoined to an IP that itself contains an IP-adjoined *wh*-phrase.
- The following rules license the left periphery in the c-structures in (16):<sup>10</sup>

$$\begin{array}{ll}
 (17) \text{ a. } CP \rightarrow & \begin{array}{c} \text{XP} \quad \text{C}' \\ (\uparrow \text{DIS}) = (\uparrow \text{DISPATH}) \quad \uparrow = \downarrow \end{array} & \text{d. } I' \rightarrow & \begin{array}{c} \text{I} \quad \text{VP} \\ \uparrow = \downarrow \quad \uparrow = \downarrow \end{array} \\
 \text{b. } C' \rightarrow & \begin{array}{c} \text{C} \quad \text{IP} \\ \uparrow = \downarrow \quad \uparrow = \downarrow \end{array} & \text{e. } IP \rightarrow & \begin{array}{c} \text{XP} \quad \text{IP} \\ (\uparrow \text{DIS}) = (\uparrow \text{DISPATH}) \quad \uparrow = \downarrow \end{array} \\
 \text{c. } IP \rightarrow & \begin{array}{c} \text{XP} \quad I' \\ (\uparrow \text{SUBJ}) = \downarrow \quad \uparrow = \downarrow \end{array} & \text{f. } I' \rightarrow & \begin{array}{c} \text{XP} \quad I' \\ (\uparrow \text{DIS}) = (\uparrow \text{DISPATH}) \quad \uparrow = \downarrow \\ (\uparrow \text{DIS})\sigma \in (\uparrow \sigma \iota \text{ TOPIC}) \end{array}
 \end{array}$$

- We assume the following lexical entry for *bāyad* ( $\square$ .PRES):

$$(18) \text{ bayad } \quad \text{I} \quad \left. \begin{array}{l} (\uparrow \text{PRED}) = \text{'must}\langle \text{CF} \rangle \text{SUBJ}' \\ (\uparrow \text{TENSE}) = \text{PRES} \\ \left\{ \begin{array}{l} @\text{EXPL-SUBJ} \\ (\uparrow \text{COMP MOOD}) =_c \text{SUBJUNCTIVE} \end{array} \right\} \left| \begin{array}{l} (\uparrow \text{SUBJ}) = (\uparrow \text{XCOMP SUBJ}) \end{array} \right\}
 \end{array}$$

- This lexical entry is for both the personal and impersonal present modal construction, so some information is shared, but the information needs to diverge at some point.
- The modal in both constructions occupies an identical position, hence the category I is shared.
- The modal in both constructions is present tense, hence the specification of [TENSE PRES].
- The two modals are also forms of the same basic predicate, so have the same PRED value.
  - The distinction is that the personal construction takes a closed sentential complement, COMP, which can realize its own subject, whereas the impersonal is a kind of subject raising construction, taking an open sentential complement, XCOMP, which cannot realize its own subject.
  - The personal construction also requires that its complement independently have subjunctive mood.<sup>11</sup>

<sup>10</sup>The equation regarding DIS in (17e) connects the top and bottom of the unbounded dependency in the corresponding f-structure (Dalrymple et al. 2019: 39ff). The grammatical function DIS is a way of unifying the previously distinguished f-structural functions of TOPIC and FOCUS, which are properly elements of i-structure rather than f-structure, as an *overlay function* that captures the abstract f-structural role of the top of unbounded dependencies; this was originally proposed by Asudeh (2004), where the function was named UDF, for *unbounded dependency function*. The set statement regarding DIS and TOPIC in (17f) states that the top of the unbounded dependency encodes a TOPIC at i(nformation)-structure (Dalrymple and Nikolaeva 2011).

<sup>11</sup>The fact that the subjunctive mood requirement is information that is checked by the modal, rather than information that is actually contributed by it, is modelled by the *constraining* equation, marked =<sub>c</sub> rather than simply =.

- The lefthand side of (18) calls a template, EXPL-SUBJ.
- A template call is marked by @.
- The semantics of template invocation is very simple (Dalrymple et al. 2004): the template just defines a bundle of lexical information and gives it a name; when the template is invoked, the corresponding information it encodes is substituted in.
- Note that a template may call other templates, so there may be multiple such substitutions; this is also exemplified by EXPL-SUBJ.

$$(19) \text{ EXPL-SUBJ} := \neg(\uparrow \text{SUBJ PRED}) \quad (20) \text{ SUBJ-3SG} := (\uparrow \text{SUBJ PERS}) = 3$$

$$\quad \quad \quad @\text{SUBJ-3SG} \quad \quad \quad (\uparrow \text{SUBJ NUM}) = \text{SG}$$

- The righthand case in (18) is for the nonfinite-complements containing apocopated infinitives, e.g. (5).
- We define the following templates for apocopated infinitives:

$$(21) \text{ APINF}(P) := (\uparrow \text{PRED}) = P \quad (22) \text{ NO-TAM} := \neg(\uparrow \text{TENSE})$$

$$\quad \quad \quad @\text{NO-TAM} \quad \quad \quad \neg(\uparrow \text{ASPECT})$$

$$\quad \quad \quad @\text{IMPERS-SUBJ} \quad \quad \quad \neg(\uparrow \text{MOOD})$$

$$(23) \text{ IMPERS-SUBJ} := (\uparrow \text{SUBJ PRED}) = \text{'pro'}$$

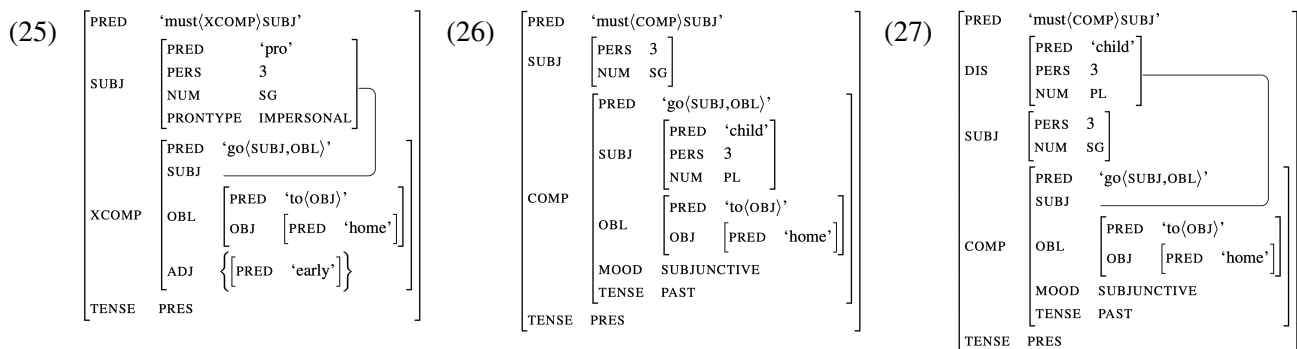
$$\quad \quad \quad (\uparrow \text{SUBJ PRONTYPE}) = \text{IMPERSONAL}$$

$$\quad \quad \quad @\text{SUBJ-3SG}$$

- Note that the APINF template is one that takes an argument: whatever is passed in as the argument becomes the value of PRED.
- The lexical entry for a sample apocopated infinitive, *raft* ('go'), is:

$$(24) \text{ raft} \quad \text{V} \quad @\text{APINF}(\text{'go(SUBJ,OBL)'})$$

- The f-structures for examples (5), (6a), and (6b) respectively are shown in (25)–(27); the corresponding examples are repeated in (28)–(30).<sup>12,13</sup>



- (28) *bāyad zood be xune raf-t.*  
 □.PRES early to home  
 go-INF  
 'It's necessary to go home early.'  
 'One must go home early.'
- (29) *bāyad bačehā be xune mi-raf-t-and.*  
 □.PRES child-PL to home  
 IMPFV-go-PAST-3PL  
 'The children had to go home.'
- (30) *bačehā bāyad be xune mi-raf-t-and.*  
 child-PL □.PRES to home  
 IMPFV-go-PAST-3PL  
 'The children had to go home.'

<sup>12</sup>Note that the imperfective and the subjunctive are syncretic in the past tense. Therefore, although for consistency we have always glossed *mi-* as imperfective (IPFV), we assume that it can convey subjunctive mood and hence satisfy the constraining equation in (18). We do not attempt to account for this syncretism here.

<sup>13</sup>See footnote 10 regarding the function DIS in (27).



## Interim summary

- We are now in a position to answer the second and third questions in §2.
- The second question asked, *What is the syntactic structure of simplex modal constructions in Persian?*
  - Persian modals occupy the category I; this is unsurprising from an LFG-theoretic perspective, since modals in general are base-generated in this category or C (depending on distribution).
  - This interacts with the general structure of the left periphery that we have provided—see (16) and the c-structure rules in (17)—such that all and only the valid orderings are captured.
- The third question asked, *How can we capture the personal and impersonal readings of modals like (4a) vs. (5)?*

- |     |    |   |       |   |
|-----|----|---|-------|---|
| (4) | a. | bāyad be xune be-rav-am.<br>□.PRES to home SBJV-go.PRES-1SG<br>'I have to go home.' | (5'') | bāyad zood be xune raf-t.<br>□.PRES early to home go-INF<br>'It's necessary to go home early.'/ |
|     |    | 'I have to go home.'  |       | 'One must go home early.'   |

- The lexical entry for the modal *bāyad* (□.PRES) in (18) explains the differences by treating the personal as a subcategorized subjunctive COMP and treating the impersonal as a raising predicate which allows the requirements of the apocopated infinitive, as captured in template (21), to control the reading, with the modal simply wrapping necessity around this.

## 4.1 Capturing the variation

- The fourth question in §2 asked, *How should the variable agreement displayed in (7) be explained?*

- |     |      |  |
|-----|------|--|
| (7) | b.   | bačē-hā be nazar mi-ā-d ke xaste šo-d-an.<br>child-PL to opinion IPFV-come.PRES-3SG that tired become-PAST-3PL<br>'As for the children, it seems that they have gotten tired.' |
|     | c. % | bačē-hā be nazar mi-ā-n ke xaste šo-d-an.<br>child-PL to opinion IPFV-come.PRES-3PL that tired become-PAST-3PL<br>'The children seem to have gotten tired.'                    |

- Our proposal may have been anticipated by now:
  1. Speakers who only allow the non-agreeing form (7b) maintain an analysis of the preposed nominal, *bačēhā* ('children'), as a TOPIC.
    - It is a general fact about Persian (and perhaps universally), that topichood is not sufficient to directly trigger agreement.
  2. Speakers who do allow the agreeing form have instead analyzed the preposed nominal as a SUBJ, which robustly triggers agreement in Persian.
    - The light verb, *āmadan* ('to come'), in this construction, unlike the modals, is a fully agreeing form (*miān*), which allows for (7c).
    - For these speakers, *be nazar āmadan* 'seems', when it shows agreement with a preposed element, is akin to English copy raising (Rogers 1973, Postal 1974):
 

(31) Harry seems like he is tired.

- However, since Persian is *pro*-drop, the embedded pronominal does not surface.<sup>14</sup>
- When it does not show agreement, as in (7b), it is akin to English *seems that* with topicalization; i.e., there is an (in Persian, unrealized) expletive subject with the bare-topicalized nominal occurring in only apparent subject position:  
(32) As for Harry, it seems that he is tired.

## Summary

- We hope to have shown that a fairly simple LFG analysis of Persian modal syntax is possible using standard tools of the framework.
- This analysis lends further support to the view that synchronic Persian grammar indeed does contain an apocopated infinitive, and that this short infinitive's formal resemblance to the past stem/zero-marked PAST.3SG form is misleading.

## 5 Conclusion

- We have now answered all of the questions that we posed in §2:

1. *How should we account for the complement in the impersonal modal construction?*

⇒ It is a (short/apocopated) infinitival which is formally but not functionally identical to the past stem.

The formal identity can be captured by standard means, such as rules of referral (Stump 2016) or their alternatives in other frameworks.

For example, in L<sub>R</sub>FG (Melchin et al. 2020, Asudeh and Siddiqi 2022, among others; lrfg.online), the required statement could look like this:

$$(33) \langle [V], @AP-INF( \_ ) \rangle \xrightarrow{\nu} \nu \langle [V], @PAST \rangle$$

The statement in (33) is intended as a *rule schema* or *meta-rule*. Its lefthand side is underspecified for the parameter of the AP-INF template and can thus match any instance of it in the Vocabulary. The schema states that the exponent of this vocabulary item is the exponent of the PAST template, which controls contribution of [TENSE PAST] to the f-structure. As mentioned above, the past tense exponent is *-d* and its allomorphs (Anoushe 2018). Thus, the schema in (33) elegantly captures the fact that the apocopated infinitive form of 'eat' is *xord*, as in *Bayad xord* ('One must eat.'), that the one of 'go' is *raft*, as in (28) above, and that the one of 'sleep' is *xābid*, as in (14a) above. But it does so without ever referring to any particular form.

2. *How can we capture the personal vs. impersonal readings of the modals?*

⇒ The distinction is governed by the lexical entry for the modal and the templates that it uses.

3. *What is the syntactic structure of the simplex modal constructions?*

⇒ The modal is in I. There is a topic position above this, but below C.

4. *How should the variable speaker agreement displayed for the subject of the raising/perception verb be nazar āmadan (lit. 'to opinion come' / ~ 'to seem like/that') be captured?*

⇒ The light verb that anchors this predicate, *āmadan*, is a fully agreeing predicate, unlike the modals.

Some speakers have reanalyzed the preposed topic as a subject, since the position it occupies is in many cases string-identical to subject position. On this analysis, the verb must agree with the subject, as is the case overall in Persian grammar. However, the other analysis, in which the preposed nominal is actually a topic, is also available, but does not trigger agreement. Therefore, these speakers display variation in their utterances.

<sup>14</sup>In fact, one could possibly get it to surface given enough discourse support, but it is difficult because of opposing discourse forces.

## References

- Anoushe, Mazdak. 2018. A Revision of Persian Past Tense Inflection: A Distributed Morphology Approach. *Language Related Research* 9(1): 57–80.
- Asudeh, Ash. 2004. Resumption as Resource Management. Ph.D. thesis, Stanford University.
- Asudeh, Ash, Mary Dalrymple, and Ida Toivonen. 2013. Constructions with Lexical Integrity. *Journal of Language Modelling* 1(1): 1–54.
- Asudeh, Ash, and Daniel Siddiqi. 2022. Morphology in LFG. To appear in Dalrymple (2022).
- Bjorkman, Bronwyn M., and Claire Halpert. 2017. In an imperfect world: Deriving the typology of counterfactual marking. In Ana Arregui, María Luisa Rivero, and Andrés Salanova, eds., *Modality Across Syntactic Categories*, 157–178. Oxford University Press.
- Dalrymple, Mary, ed. 2022. *The Handbook of Lexical Functional Grammar*. Language Science Press. Forthcoming.
- Dalrymple, Mary, Ronald M. Kaplan, and Tracy Holloway King. 2004. Linguistic Generalizations over Descriptions. In Miriam Butt and Tracy Holloway King, eds., *Proceedings of the LFG04 Conference*, 199–208. Stanford, CA: CSLI Publications.
- Dalrymple, Mary, John J. Lowe, and Louise Mycock. 2019. *The Oxford Reference Guide to Lexical Functional Grammar*. Oxford: Oxford University Press.
- Dalrymple, Mary, and Irina Nikolaeva. 2011. *Objects and Information Structure*. Cambridge: Cambridge University Press.
- Darzi, Ali, and Saera Kwak. 2015. Syntax and semantics of subjunctive clauses in Persian. *Lingua* 153: 1–13.
- Karimi, Simin. 2005. *A Minimalist Approach to Scrambling: Evidence from Persian*. Berlin: De Gruyter Mouton.
- . 2008. Raising and Control in Persian. In Simin Karimi, Vida Samiian, and Don Stilo, eds., *Aspects of Iranian Linguistics*, 177–208. Cambridge Scholars Publishers Ltd.
- Labafankhosh, Zahra, and Ali Darzi. 2015. On the Syntactic Analysis of Persian Modals: Šâyad ‘Perhaps’ and Bâyard ‘Must’ [In Persian]. *Journal of Linguistic Research* 5(2): 97–112.
- Lowe, John J. 2019. The Syntax and Semantics of Nonfinite Forms. *Annual Review of Linguistics* 5(1): 309–328.
- Melchin, Paul B., Ash Asudeh, and Dan Siddiqi. 2020. Ojibwe Agreement in Lexical-Realizational Functional Grammar. In Miriam Butt and Ida Toivonen, eds., *Proceedings of the LFG20 Conference*, 268–288. Stanford, CA: CSLI Publications.
- Mirrazi, Zahra. 2022. Tense in Conditionals: Ins and Outs. Ph.D. thesis, University of Massachusetts Amherst.
- Postal, Paul M. 1974. *On Raising: One Rule of English Grammar and its Theoretical Implications*. Cambridge, MA: MIT Press.
- Rogers, Andy. 1973. Physical Perception Verbs in English: A Study in Lexical Relatedness. Ph.D. thesis, UCLA.

- Samvelian, Pollet. 2018. Specific Features of Persian Syntax: The Ezâfe Construction, Differential Object Marking and Complex Predictaes. In Anousha Sedighi and Pouneh Shabani-Jadidi, eds., *The Oxford Handbook of Persian Linguistics*, 226–269. Oxford University Press.
- Stump, Gregory T. 2016. *Inflectional Paradigms: Content and Form at the Syntax-Morphology Interface*. Cambridge: Cambridge University Press.
- Taleghani, Azita H. 2008. *Modality, Aspect and Negation in Persian*. Amsterdam: John Benjamins.
- Windfuhr, Gernot L. 1979. *Persian Grammar: History and State of Its Study*. The Hague: Mouton.