

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

Setayesh Dashti (Oxford) & Ash Asudeh (Rochester)

This paper offers two key contributions. The first contribution is a novel, LFG analysis of the syntax of Persian modals. We show that this analysis is not only economical but also accounts well for the main patterns. The second contribution is evidence and argumentation, based on the modal analysis, for the so-called *apocopated infinitive* (i.e., short infinitive) in modern Persian. Some have considered this an archaic form *qua* infinitive that is no longer active synchronically and instead have analyzed it as a PAST.3SG form (with which it is identical) in modern Persian. We argue that the identity of form is a morphological fact (e.g., captured by paradigm cell referral) but that the function of this ambiguous form is in fact infinitival in the relevant cases.

Background. Persian is an SOV language with *pro*-drop.¹ Verbal morphology follows a two-stem system, traditionally called present (e.g., *xor* ‘eat’) and past stems (e.g., *xord* ‘eat’). Modulo suppletive patterns, the past stem is regularly marked with *-d* and its allomorphs (Anoushe 2018).² Persian has no overt present tense marker. The present stem always occurs with either aspectual or mood markers; *mi-* for imperfective aspect (1a) and *be-* for subjunctive mood (1b). The unprefixed past stem with agreement suffixes is used to show the perfective aspect (1c). 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-* (1d).³

- (1) a. Nika be madrese mi-rav-ad.
Nika to school IMP-go.PRES-3SG
‘Nika goes to the school.’
- b. Nika šāyad be madrese be-rav-ad.
Nika maybe to school SBJV-go.PRES-3SG
‘Maybe, Nika goes 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 IMP-go-PAST-3PL
‘The kids had gone to school every day.’

While Persian contains several adverbial and complex predicate modals, there are two main modal auxiliaries in this language, *bāyestan* (necessity/□) and *šodan* (possibility/◇).⁴ These modals always appear in the default third person singular form: *bāyad* (□.PRES)/*bāyest* (□.PAST) and *mi-še* (IMP-◇.PRES)/*mi-šod* (IMP-◇.PAST). They can either occur with: 1. a finite complement (2), marked with subjunctive mood in present tense (2a) or imperfective aspect in past tense (2b); or 2. a non-finite complement (3) In the latter case, the non-finite complement has a simple past stem, which resembles the third person singular past inflection, and is historically an *apocopated infinitive*; it is interpreted as an impersonal (3).

- (2) a. *bāyad* be xune be-rav-am.
□.PRES to home SBJV-go.PRES-3SG
‘I have to go home.’
- b. *bāyad* bačehā be xune mi-raf-t-and.
□.PRES child-PL to home IMP-go-PAST-3PL
‘The children had to go home.’
- (3) *bāyad* zood be xune raf-t.
□.PRES early to home go-PAST.3SG
‘It’s necessary to go home early.’
- ‘One must go home early.’

The modal can also occur with a subjunctive past to express an epistemic necessity modal embedding a perfective:

- (4) *bāyad* bačehā be xune rafte bāš-and.
□.PRES child-PL to home go.PP AUX-3PL
‘The children must have gone home.’

When the modal occurs with a finite complement, it is possible to topicalize the embedded subject to the left:⁵

- (5) a. *bāyad* bačehā be xune mi-raf-t-and.
□.PRES child-PL to home SBJV-go-PAST-3PL
‘The children had to go home.’
- b. bačehā *bāyad* be xune mi-raf-t-and.
child-PL □.PRES to home SBJV-go-PAST-3PL
‘The children had to go home.’

Challenges. The main challenge in giving a formal characterization of Persian modals arises when dealing with modals in an impersonal construction with a non-finite complement, as in (3) or (6a). This non-finite complement is not, at least synchronically, the productive Persian infinitive: for example, it lacks the final *-an* marker (*raftan* ‘to go’ vs. *raft*). Moreover, it looks superficially like the simple third person singular past form; see (1c). However, there are reasons to believe that it is not an agreeing past form. For example, while the past finite complement of the modal should bear imperfective marking, adding this marking to this non-finite complement renders an impersonal reading unavailable and requires it to have a personal reading instead (6b).⁶

¹The data throughout this paper is from colloquial spoken Persian, not from the written standard.

²Glosses are abbreviated as follows: AUX–auxiliary, IMP–imperfect, INF–infinitive, PP–past participle, PRES–present tense, PAST–past tense, SBJV–subjunctive mood, SG–singular, PL–plural, DO–direct object.

³Past imperfective also functions as a fake past to convey counterfactuality, regardless of tense Bjorkman and Halpert (2017).

⁴While some literature such as Karimi (2005) and Taleghani (2008) also names *šāyestan* as another modal auxiliary in Persian, Labbafankhosh and Darzi (2015) have shown that this is rather a modal adverb.

⁵The subjunctive in the past tense has the same form as the imperative.

⁶Sentence (6b) can have another interpretation in which the subject of the verb is *pro*-dropped, which will translate to ‘he/she 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.

- (6) a. *bāyad šab-hā hašt sā'at xāb-id.*
 □.PRES night-PL eight hour sleep-PAST.3SG?
 '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 SBJV-sleep-PAST.3SG
 # 'It's necessary to sleep for eight hours a night./One must sleep for eight hours a night.'
 ✓ 'She/he/it 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., *raft* go.?).

But, as we have just seen, this does not get the correct interpretation for impersonal examples. We instead assume that this verbal form is infinitival and thus unmarked for TENSE/ASPECT/MOOD. The future construction, shown in (7), provides further evidence for non-finiteness of this verbal form, now glossed INF. This builds on Lowe's (2019) claim that non-finite forms generally appear in periphrastic constructions as the lexical content of the clausal predicate.

- (7) *Ali farda be madrese xāh-ad raft.*
 Ali tomorrow to school want-3SG go.INF
 'Ali will go to school tomorrow.'

We use a template below to generate this defective/infinitival verbal form. In sum, we argue that the correct gloss for the apparent past tense form in (3) and (6a), just as in (7), is *still* an infinitival form (synchronically), and that the apocopated infinitive is morphologically formed by referral to the past stem, which explains their identity of form; but they have different functions.⁷

Analysis. The following examples illustrate the c-structure position of the modal and the general structure of the CP and IP:

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

Example (8a) 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 (8b) 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 (8b), *bacheha* is in a non-agreeing topic position, reflected by the lack of plural agreement on the modal (which is generally not possible). Example (8c) 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.

The following rules license the left periphery in the c-structures in (8):

- | | |
|---|---|
| <p>(9) a. $CP \rightarrow XP \quad C'$
 $(\uparrow \text{ DIS}) = (\uparrow \text{ DISPATH}) \quad \uparrow = \downarrow$</p> <p>b. $C' \rightarrow C \quad IP$
 $\uparrow = \downarrow \quad \uparrow = \downarrow$</p> <p>c. $IP \rightarrow XP \quad I'$
 $(\uparrow \text{ SUBJ}) = \downarrow \quad \uparrow = \downarrow$</p> | <p>d. $I' \rightarrow I \quad VP$
 $\uparrow = \downarrow \quad \uparrow = \downarrow$</p> <p>e. $IP \rightarrow XP \quad IP$
 $(\uparrow \text{ DIS}) = (\uparrow \text{ DISPATH}) \quad \uparrow = \downarrow$</p> <p>f. $I' \rightarrow XP \quad I'$
 $(\uparrow \text{ DIS}) = (\uparrow \text{ DISPATH}) \quad \uparrow = \downarrow$
 $(\uparrow \text{ DIS})\sigma \in (\uparrow \sigma \iota \text{ TOPIC})$</p> |
|---|---|

The lexical entry for *bayad* is as follows:

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

⁷The 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 we are not aware of any theoretical or formal account of the apocopated infinitive. The theoretical literature has instead taken it to be the PAST.3SG form of the verb (for instance, Karimi 2008, Mirrazi 2022).

The meta-function CF is defined as a functional uncertainty: $CF = \{COMP|XCOMP\}$.

The lefthand case in (10) is for finite-complement cases, such as (2a–b) and (5a–b). The finite-complement case requires the subject to be an expletive, using the templates in (11) and (12) and requires its complement’s mood to have been specified as subjunctive.

$$(11) \text{ EXPL-SUBJ} := \neg(\uparrow \text{SUBJ PRED}) \\ @3SG$$

$$(12) \text{ 3SG} := (\uparrow \text{SUBJ PERS}) = 3 \\ (\uparrow \text{SUBJ NUM}) = \text{SG}$$

The righthand case in (10) is for the nonfinite-complements containing apocopated infinitives, such as (3). It is a standard subject-raising predicate which identifies its subject with that of its (infinitival) complement. We define the following templates for apocopated infinitives:

$$(13) \text{ APINF(Pred)} := (\uparrow \text{PRED}) = \text{Pred} \\ @\text{NO-TAM} \\ @\text{IMPERS-SUBJ}$$

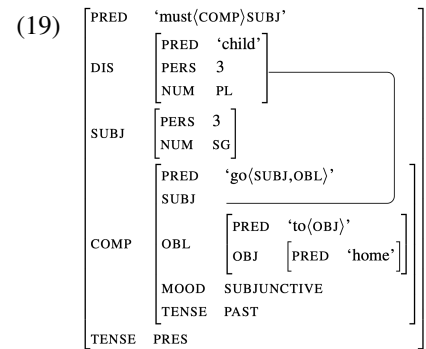
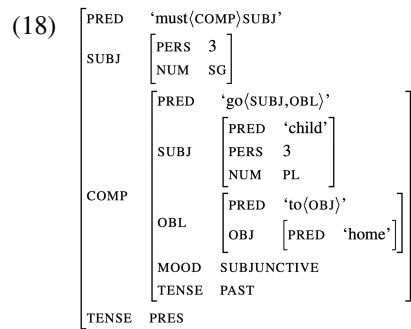
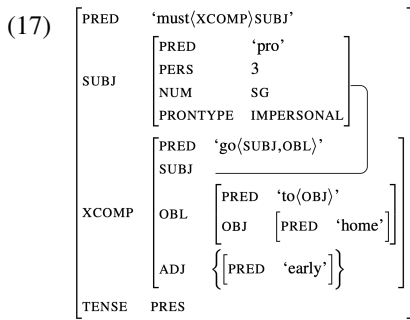
$$(14) \text{ NO-TAM} := \neg(\uparrow \text{TENSE}) \\ \neg(\uparrow \text{ASPECT}) \\ \neg(\uparrow \text{MOOD})$$

$$(15) \text{ IMPERS-SUBJ} := (\uparrow \text{SUBJ PRED}) = \text{'pro'} \\ (\uparrow \text{SUBJ PRONTYPE}) = \text{IMPERSONAL} \\ @3SG$$

The lexical entry for a sample apocopated infinitive, *raft*, is:

$$(16) \text{ raft} \quad V \quad @\text{APINF('go(SUBJ,OBL)')}$$

The f-structures for examples (3), (5a), and (5b) respectively are shown in (17)–(19); the corresponding examples are repeated in (20)–(22).



(20) *bāyad zood be xune raft.*
□.PRES early to home go.INF
'It's necessary to go home early.'
'One must go home early.'

(21) *bāyad bačehā be xune mi-raf-t-and.*
□.PRES child-PL to home
SBJV-go-PAST-3PL
'The children had to go home.'

(22) *bačehā bāyad be xune mi-raf-t-and.*
child-PL □.PRES to home
SBJV-go-PAST-3PL
'The children had to go home.'

Summary. We have shown that a fairly simple LFG analysis of Persian modal syntax is possible using standard tools of the framework. Furthermore, we have shown that 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.

References

- Anoushe, Mazdak. 2018. A Revision of Persian Past Tense Inflection: A Distributed Morphology Approach. *Language Related Research* 9(1): 57–80.
- 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.
- 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 Samiiian, 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: Šāyad ‘Perhaps’ and Bāyad ‘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.
- Mirrazi, Zahra. 2022. Tense in Conditionals: Ins and Outs. Ph.D. thesis, University of Massachusetts Amherst.
- 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.