

# A SYNTACTIC ANALYSIS & GRAMMATICIZATION OF SIMPLE FUTURE, WITH XĀST-AN, IN PERSIAN

Mohammad Mehdi Vahedi Langrudi

Tarbiat Modarres University, Tehran, Iran

## Abstract

In this paper, I present a unified syntactic analysis of the main verb *xāst-an* "want" in its three distinct uses in Persian. This unified analysis observes economy in derivation, computation as well as in the lexicon by proposing that *xāst-an* takes three different types of complements in syntax, i.e., NP, CP, VP. The three types of complements indicate three stages of grammaticization of this verb in a unidirectional universal path from more concrete to more abstract meaning. I argue that *xāst-an* in all its three uses is a modal type main verb and not an auxiliary in Persian.

## I- Introduction

This article addresses the various uses, syntactic functions, and grammaticization of the (main) verb *xāst-an* "want", its modal-like meaning in Modern Persian. This verb is used in three different syntactic contexts in Persian. I argue that all these three uses may be unified under a single syntactic analysis. This then indicates the economy which prevails in language and shows that human language and grammar, despite their surface complexities, follow certain underlying properties and principles. The three uses of *xāst-an* are presented below:

**A- The 1<sup>st</sup> use of *xāst-an* "want", *xāst-an-1*:** Here the verb is a simple transitive lexical verb taking an NP-DO. Semantically, it indicates the modal notions desire, wish, and intention of the

subject/agent to possess the NP complement.

1a- Ali ketāb mi-xāh-ad "Ali wants a book."

1b- Ali ketāb rā mi-xāh-ad "Ali wants the book."

1c- mardom az dowlat edālat mi-xāh-and  
"People want justice from the government."

Similar to many simple lexical verbs in Persian, *xāst-an-1* may form a compound verb with a non-referential, predicative DO, (1c), as in *edālat-xāh-an* "justice-seekers".

**B- The 2<sup>nd</sup> use of *xāst-an*, *xāst-an-2*:** here the verb takes a full clausal, CP, complement. Semantically, it specifies the volition and desire of the main clause subject that the proposition/situation in the subordinate clause hold:



2a- man mi-xāh-am (ke) pro be-rav-am  
"I want to go."

2b- man mi-xāst-am (ke) ali be-rav-ad  
"I wanted Ali to go."

2c- man az āli mi-xāh-am ke pro be-rav-ad.  
"I ask from Ali pro to go."

In all uses in (2) the subordinate verb must occur in present subjunctive mood in Persian. Here again, the main verb indicates the modal notions desire, wish, and intention of the subject/agent towards the proposition in the subordinate clause. There may be subject control (2a), indirect object control (2c), or no control (2b) between the main clause and the subordinate subject.

**C- The 3<sup>rd</sup> use of *xāst-an*, *xāst-an-3*:** as the aux(iliary), or the modal-like constituent of simple future:

3a- Man xāh-am raft "I will go."  
3b- Mā xāh-im raft "We will go."  
3a- 'ānha xāh-and raft "They will go."

In this use, *xāst-an* is usually considered to be the auxiliary of future tense in Persian. The main verb, i.e., *raft* "went, go" is used in one form only, i.e. past stem form, and is obligatory in third person singular, or neutral. The assumed aux, *xāst-an* "want" is inflected for person, and number.

In this article we argue that the first lexical use of the main verb *xāst-an* "want", is the lexical source from which the 2<sup>nd</sup> and 3<sup>rd</sup> uses are derived. In other words, the 2<sup>nd</sup> and 3<sup>rd</sup> uses of this verb show different degrees of grammaticization along a unidirectional path, which is also attested cross-linguistically in other languages as well. The goal of this article is to provide a syntactic, and semantic account of the verb *xāst-an* and to trace the grammaticization path of this verb from its first use to the second and third uses. I will also argue that *xāst-an* in all the three uses behaves like a main transitive verb taking three different types of categorial complements. This will then lead us to the typological word order studies and the role of *xāst-an* in determining the basic word order in Persian. The article unfolds as follows: In section II,

we discuss grammaticization theory. Section III deals with the evolution and grammaticization of *xāst-an* "want", and discusses Darzi's (1996) analysis of this verb. In section IV, I compare the behavior of *xāst-an* in Persian with restructuring verbs in Italian and show that this verb in simple future tense behaves as a restructuring modal-type light verb. This section provides a syntactic analysis of the three uses of *xāst-an*, contrary to Rosen's (1989) argument structure account. Section V concludes the article.

## II- Grammaticization, or grammaticalization

Grammatization begins with the observation that "grammatical morphemes develop gradually out of lexical morphemes or combinations of lexical morphemes with lexical or grammatical morphemes" (Bybee et al, 1994, p.4.) This includes changes in lexical morphemes by which a lexical morpheme becomes more frequent and general in meaning, and gradually shifting to grammatical status, and developing further after the grammatical status has been attained. That is, grammatical materials or morphemes are the outcome of the evolution of substance from the more specific to the more general, abstract, and relational. Bybee et al argue for "source determination" in their theory of grammaticization according to which the actual meaning of the construction that enters into grammaticization uniquely determines the path that grammaticization follows and, consequently, the resulting grammatical meanings. In other words, grammatical morphemes, as descendants of lexical items, lose most if not all of the specificities of lexical meaning they formerly had; the meaning that remains is very general and is often characterized as abstract or relational. The loss of specificity is also termed "semantic change in grammaticization" or "semantic generalization" which correlates with a generalization of the contexts in which the gram can be used, so that certain components of meaning are lost in this process and so is variously called "semantic reduction", "bleaching", and "erosion". The core of lexical meaning is contained in the



lexical meaning (Givón, 1973). In tracing the origin of grammatical meaning, we must attend to the syntax and morphology of the source construction and not simply to the referential meaning of its lexical items.

In this article, we suggest that the use of *xāst-an* in simple future tense in Persian is an instance of grammaticization along a unidirectional, universal path from its corresponding lexical morpheme/source. We specifically claim that *xāst-an* in this use is a modal-like, agent-oriented verb indicating desire, or volition.

Bybee et al (1994, p. 177) exemplify four types of modality one of which they call "agent-oriented modality" which reports "the existence of internal and external conditions on an agent with respect to the completion of the action expressed in the main predicate", and can be expressed by either lexical or grammatical morphemes. These include notions like obligation, necessity, ability (and root possibility) and desire. The notion we are concerned with is desire which is expressed by the lexical morpheme *xāst-an* "want" in Persian, *would* and *want* to in English, *vouloir* "want" in French, *volère* "want" in Italian, and *quierer* "want" in Spanish. They claim that "in the formation of the future both desire and obligation can come to be used in sentences expressing the intentions of the agent, especially in the first person. They provide examples from Middle English to show that both *will*, from a desire source, and *shall*, from an obligation source, are used to express first person intentions.

### III- The evolution of *xāst-an-3* in future

In this section we argue that *xāst-an-3* in its future use is the grammaticization of its lexical use in *xāst-an-1*. According to Lyons (1968, p. 310) statements made about future occurrences are necessarily based upon the speakers' beliefs, predictions or intentions, rather than upon their knowledge of "fact". He adds that "the expression of 'futurity' in English (and in other languages) is as much a matter of mood as of tense". On the other hand, Bybee et al (1994, p. 244) "regard the focal

use of future as equivalent to a prediction on the part of the speaker that the situation in the proposition, which refers to an event taking place after the moment of speech, will hold." That is, future expresses a prediction about an event, which is yet to occur.

According to Bybee et al's cross-linguistic search, futures evolve from a fairly restricted range of lexical sources from constructions involving movement verbs, from markers of obligation, desire, and ability, and from temporal adverbs (1994, p. 244) which they call "primary futures", and distinguish it from "aspectual futures". In fact, a very common agent-oriented pathway to future begins with 'desire' which they call 'desire future', because they either have 'desire', and/or 'willingness', as another use along with future, or they come from lexical sources with earlier meanings of 'desire'.

The Persian future with *xāst-an* seems to accord with both cases since it not only indicates 'desire' and 'willingness' in its first use, but also its lexical source meaning reveals 'desire'. Persian future is, indeed, a typical instance of desire-future. Bybee et al (1994, p. 256) hypothesize the following pathway (4) for desire future, even though they are unable to consider 'willingness' as a use separate and distinct from 'desire' for most of the languages in their corpus. So they predict that futures from 'desire' will have nuances of 'willingness' at some stages in their development.

4- Desire > willingness > Intention > Prediction

The use of *xāst-an-1* in (1), as its lexical source, indicates the desire of the agent/subject to possess something, i.e.,

5- 'u yek docharxe-ye now mi-xāh-ad

"He wants a new bicycle."

he one bicycle-EZ new want-3S

In (5) 'u "he" has a 'desire, willingness', to possess a new bicycle. This meaning is also available in the historical development of English 'will' (6a) (from a desire source), and 'shall' (6b) (from an obligation source) in Middle English where they are used to express first person intention:



- 6-a. I wyl naugþper grete ne grone ...  
I will not shout or groan.
- 6-b. And I schal ware alle my wyt to wynne me  
þpeder  
and I shall use all my wit to find my way  
there. (Bybee et al., 1994, p. 178)

Desire also gives rise to expression of willingness, as in (7):

- 7-a. I'll help you.
- 7-b. man be to komak xāh-am kard

I to you help want-1S do

"I will help you."

In *xāst-an-2* the sense of 'intention' is clearly inferable from the use of the desire-modal *xāst-an-1*:

- 8- 'u mi-xāh-ad (ke) PRO/pro yek docharxe-ye  
now be-xar-ad  
he want-3S (that) one bicycle-EZ new buy.3S  
"He want to buy a new bicycle."

This sentence clearly indicates the intention of the agent/subject to buy a new bicycle. Accepting 'intention' as the core meaning of *xāst-an-2* in (8), one may hypothesize with Bybee et al (1994, p. 256) that the 'prediction' function, of future, arises from the 'intention' function:

- 9- 'agar ali yek kār-e xub peydā kon-ad, yek xāne-ye  
now xāh-ad xar-id  
if Ali one job-EZ good find do-3S, one  
house-EZ new want.3S buy.ps

"If Ali finds a good job, he will buy a new house."

Not only does (9) express the 'intention' of Ali to buy a new house, given he finds a good job, but also expresses a 'prediction' on the part of Ali. Thus we note that 'intention' comprises an important aspect of the meaning of future, which itself counts a 'prediction'. We note that the primary lexical use of *xāst-an-1*, that lexically specifies the modality of desire, volition, and *xāst-an-2*, that specifies 'intention', evolve into its modal-like function in simple future tense in Persian as a 'prediction' in *xāst-an-3*:

- 10- man fardā be madrase xāh-am raft.  
I tomorrow to school want-1S went (ps)  
"I will go to school tomorrow."

In (10) *xāh-am* is used to express a 'prediction', still a sense of 'intention' lingers in the background sense of the sentence. However, my intention is not only to show that the first lexical source/use of the verb *xāst-an* has evolved along a unidirectional universal path towards its 2<sup>nd</sup> and 3<sup>rd</sup> more restrictive uses, but also to indicate that the verb *xāst-an* in its third use behaves differently from the aspectual aux(iliary) *bud-an* (be) and passive aux *shod-an* (become) in Persian. The Modal-like verb *xāst-an-3* precedes the main verb while *bud-an* "be" and *shod-an* "become" systematically follow the main verb. This difference in distribution denotes the distinct status of the two in the verbal system of Persian. As such, it seems that, contrary to Darzi (1996), the behavior and distribution of *xāst-an-3*, as a simple future marker, does not violate Greenberg's universal 16 according to which in dominantly SOV languages an inflected aux tends to follow the main verb. Not being an aux, and preceding the main verb, *xāst-an-3* does not violate the Universal 16 with regard to SOV order. In order to determine the status of *xāst-an-3* in Persian verbal system, I address Darzi's (1996) arguments regarding *xāst-an-3* as a 'true aux'. I will then show that *xāst-an-3* should not be considered a (true) aux in Persian.

Darzi (1996) in his argument regarding the basic word order in Persian claims that the inflected forms of *xāst-an* in simple future tense, i.e. our *xāst-an-3*, is a true aux, while the inflected forms of *bud-an* "be" and *shod-an* "become" are not. He uses this criterion in order to investigate whether the surface unmarked SOV in Persian main clauses is, or is not, compatible with Greenberg's (1993) universal 16:

Universal 16: In languages with dominant order SOV, an inflected auxiliary always follows the main verb.

By claiming that *xāst-an-3*, in contrast with *bud-an*



and *shod-an*, is a true aux preceding the main verb in Persian, Darzi (1996) attempts to cast doubt on the basic, underlying SOV order in Persian. Given the universal 16, he claims that the distribution of *xāst-an-3* favors an underlying SVO order in Persian. First, following Hashemipour (1988b) and Marashi (1970), he argues that *dārad* "have", *bāyad* "must", *shāyad* "may", that precede the main verb in Persian, are reportedly auxiliaries in Persian; so the inflected form of *xāst-an-3* must be considered an aux as well. (11b) is ill-formed because the aux follows the main verb:

11a- hamid *dārad* *otāq rā rang mi-kon-ad*

11b- \* hamid *otāq rā rang mi-konad dārad*

He presents similar examples with *bāyad* "must", and *xāst-an-3* to indicate their putative status as true auxiliaries preceding the main verb (Darzi 1996, ex. 15-16).

There seems to be little evidence to prove Darzi's above assumptions regarding these defective verbs. *Xāst-an-3* shows much less flexibility and possibility of movement compared with the other defective verbs, i.e. *dārad* "have", *bāyad* "must", *shāyad* "may". The syntactic distribution of *xāst-an-3* is different from the other defective verbs. Historical and cross-linguistic evidences seem to indicate that verbal constituents like want, desire, wish, intention etc. have a tendency to be grammaticized as modal-like verbs, indicating modal notions like desire, intention, prediction and future (cf Bybee et al, 1994).

Comparing (11a-b) with (12a-b), Darzi also argues that (12b) in which the DO intervenes between the main verb and aux is ill-formed but (11a) is not.

12a- mehdi meysam-rā did-e bud

12b- \* mehdi did-e meysam-rā bud

He takes this to show that *bud-an* forms a complex predicate, and a single syntactic unit, with the main verb in (12a), so it cannot be separated from the main verb. In (11a) the (true) aux does not form a single syntactic unit with the main verb so it may be separated from the verb. Given this observation, he claims that *xāst-an-3* "want" (similar

to *bāyad*, *shāyad*, and *dārad*) must be considered a true aux in INFL position, but *bud-an* should not. He then concludes that IP (Inflectional Phrase) in Persian is head-initial which entails that SVO is more plausible than SOV order because the distribution of *xāst-an-3* seems to contradict Universal 16<sup>1</sup>.

Batani (1991, p. 125) considers *xāst-an-3* a defective verb that forms simple future tense in combination with a lexical main verb. Other defective verbs for Batani are *bāyad* "must", *mi-tāvān* "one can", *mi-shavad* "it is possible", in *mi-shavad raft* "one can go", *dāram* "have", *dārad* "he has", in *dārad mi-rav-ad* "he is about to go", *dāsh-t-an* "have". However, Batani notes a distinction between the distribution of *xāst-an-3*, the other defective verbs, and the passive aux *shod-an* "become", and the aspectual aux *bud-an* "be". The former must precede, but the latter must follow the main verb. On this basis, he differentiates the two groups of verbs/auxiliaries and considers *xāst-an-3* a defective verb, rather than a true aux.

Following Marashi (1970, ch III), Karimi (1989, p. 134) claims that *xāst-an-3* is in fact a modal in Modern Persian. She also takes this to argue that since the true auxiliaries *bud-an* and *shod-an* follow the main verb, so Persian respects the Universal 16, indicating that Persian favors SOV order.

In the next subsection, we discuss Rosén's (1989) analysis of restructuring verbs like *volòre* "want" in Italian and *quierer* "want" in Spanish which behave similar to the three uses of *xāst-an* in Persian. In our presentation of her arguments, we will present our analysis of *xāst-an* as well.

#### IV- *Xāst-an-3* as a Modal-type Restructuring Verb

Rosen (1989, p. 159) presents a class of verbs in Italian and Spanish that behave like causatives and perception verbs in these languages. These include the modal-type verbs like the Italian verb *volòre* "want", *cominciare* "begin", *continuare* "continue", *dovere* "have to", and some verbs of motion like *andare* "go" and *venire* "come". She observes a class



of properties associated with restructuring verbs which include clitic climbing, long object preposing, and auxiliary selection by the embedded verb. She then attempts to show that these three properties are a result of argument structure merger in Spanish and Italian. However, she calls this process "light merger" since she notes that restructuring verbs behave like light verbs. My aim is to review her analysis of the light restructuring verb *volere* in these two languages, and show that the Persian verb *xāst-an-3* only partly behaves like the restructuring light verbs of Italian and Spanish. Our search also reveals grammaticization of a similar notion/word in Italian, Spanish, and Persian.

Rosen (1989, pp. 166-171) defines restructuring verbs as modal-like verbs because they are inextricably linked to the embedded verbs, and because they are semantically related to modals in other languages. The argument structure of a restructuring (matrix) verb, i.e., *xāst-an-3* and *volere*, is an empty skeleton, and has no argument of its own. It must compose with another argument-taking item in order to license arguments in syntax.

Gerds (1988) categorizes three types of desideratives that exist cross-linguistically. These include what she calls "structure building" desideratives, in which the verb contributes its own subject argument. This contrasts with "inheritance" desideratives, in which the verb has no argument structure of its own. The inheritance verbs break down into two types, one in which the desiderative imposes selectional restrictions on the matrix subject, (it must be an animate, sentient being), and one in which the desiderative imposes no such selectional restrictions (see Rosen, 1989, p. 170). In short, one type of desideratives have no arguments of their own to contribute to the complex predicate (inheritance desideratives), while the structure building desideratives have (an) arguments to contribute. It seems that the Italian and Spanish restructuring verb *volere*, and Persian *xāst-an-3*, like inheritance desideratives, have an empty skeleton, and have no argument structure of their own to contribute to the complex predicate. They must

compose with an argument-taking item in order to license arguments in syntax.

However, unlike Japanese *suru* "do", and Persian *kard-an* "do", the modal-type restructuring verbs do have some meaning. They also have a heavy counterpart with a complete argument structure which maps into a complete unrestructured matrix clause, and takes either a full NP complement or a full embedded clause complement. These latter two unrestructured uses of *xāst-an* and *volere* are equivalent to our *xāst-an-1*, in *man yek docharxe-ye now mi-xāh-am* "I want a new bicycle", and to our *xāst-an-2*, in *man mi-xāh-am [yek docharxe-ye now be-xar-am]* "I want to buy a new bicycle", and are equivalent to their English translations.

Rosen presents the following lexical conceptual structure (LCS) and argument structures for the verb *volere* "want" in Italian:

#### LCS of *volere* "want"

13- *Volere-1*: "want-1": [X] desires [thing Y] to come to X's possession.

14- *Volere-2*: "want-2": [X] desires [event Y] to occur.

15- Argument Structure of want-1:

( X (Y)) <e>  
Exp Th

Want-1 has a full LCS and a complete argument structure. It is a full unrestructured verb with an experiencer external argument and a theme internal argument. This is equivalent to the Persian *xāst-an-1*, in *man yek docharxe-ye now mi-xāh-am* "I want a new bicycle."

Rosen (1989, pp. 173-174) argues that the other *volere-2* "want-2", with an event internal argument, is associated with two different argument structures in Italian. In one use, *volere* "want-2" takes an experiencer external argument and an event internal argument which is mapped into a full clausal CP complement in Italian:

16a- The 1<sup>st</sup> argument structure of *volere-2*, want-2, equivalent to our *xāst-an-2*



( X ( Y )) <e>  
Exp event

16b- man mi-xāh-am [ (ke) PRO yek docharxe-ye  
now be-xar-am]

I want.1S (that) one bicycle-EZ new buy-1S  
"I want to buy a new bicycle."

16c- I want [ PRO to buy a new bicycle]

The equivalence between the Persian example (16b) and its English translation (16c) is perfect. The Persian example obligatorily takes a subjunctive complement while the English translation must be an infinitival complement. They are both argued to have a big PRO subject controlled by the matrix experiencer subject. The main verbs in (16) are typical desire/volition verbs indicating the desire, volition, or intention of the experiencer subject towards the event indicated by the subordinate clause: a typical agent-oriented modal notion.

The other alternative argument structure of *volère*, want-2, our *xāst-an-3* in simple future use, according to Rosen, has no arguments at all as in (17):

17- The 2<sup>nd</sup> argument structure of *volère-2* "want-2", our *xāst-an-3* (Rosen, 1989, pp. 174, 29b):

( ) <e>

That is, *volère-2*, in the third use, and *xāst-an-3* have an empty, incomplete argument structure. The empty argument structure must then merge with another verb's argument structure in order to project into syntax. She claims that a language like Italian has a mechanism of combining argument structures like those in (17) with the argument structure of another full verb to create a complex argument structure which she calls "light verb merger". This is the restructuring use of *volère-2* "want-2", our *xāst-an-3*, where these verbs behave like a modal-like aux as in simple future tense in Persian, (1 above), and in Italian and Spanish, but not in English.

This means that if the second main verb that combines with want-2, *xast-an-3*, is intransitive, unaccusative, transitive, and di-transitive, then the resulting complex will be intransitive, unaccusative, transitive, and di-transitive respectively as well. The restructuring light verbs contribute no arguments to the resulting complex verbs. This is shown in (18) from Rosen (1989:175):

18a- Transitive verb

volere ( ) <e> } Volere leggere (x (y)) <e> <e>  
leggere(x (y)) <e> } 'want to read' [ ]

18b- Unaccusative verb

volere ( ) <e> } Volere andare ((x)) <e> <e>  
andare ((x)) <e> } 'want to go' [ ]

According to Rosen the <e> role of *volere*, want-2, is identified with the <e> role of the main verb, so that the two verbs express a single event role, a complex predicate.

While this observation might seem to be correct, I suggest the following revision to (17):

19- The 2<sup>nd</sup> argument structure of *volère-2*, want-2, our *xāst-an-3*. "REVISED"

( ( Y ))  
event

This means that *volère-2*, our *xāst-an-3*, in (19) contrary to (16a and 17), lacks an event role <e> and an external argument. That is, the event role is provided by the following lexical verb itself, which is discharged into I (inflection) in syntax. Thus, no single argument structure, or complex predicate, is formed because there is only one <e> role which belongs to the main complement verb.

There seems to be a difference between the restructuring verbs in Italian and Spanish, and the third use of *xāst-an-3* in that the Persian verb does not form a complex predicate with the following lexical verb. In other words, while Rosen argues that restructuring verbs and their embedded lexical verbs form a complex predicate at the level of



argument structure through light verb merger (18), the use of *xāst-an-3* in Persian seems to be a pure syntactic phenomenon, rather than a lexical argument structure one. *Xāst-an-3* semantically selects an event complement, but lacks an <e> role itself as indicated in our revised argument structure (19) which is then projected into syntax in the form of simple future tense in examples like *man be madrase xāh-am raft* in which the second main verb is obligatorily a past stem in neutral third person singular. All the arguments of the future are determined by the second verb. Rizzi (1982), Burzio (1986), and Rochette (1988) also argue for a syntactic account of restructuring verbs in Italian and Spanish, contra Rosen who follows an argument structure approach.

However, they are similar in that both in Italian restructuring verbs and in Persian future *xāst-an-3*, the second lexical verb maps into a VP complement. That is, I am suggesting that in simple future tense in Persian, the second main verb is configurationally a VP2 and a complement of the modal-like verb *xāst-an-3*. *Xāst-an-3* as the super-ordinate verb in the structural, x-bar configuration selects a VP complement. *Xāst-an-3* raises higher in the configurational functional tree diagrams in order to check the tense and agreement features. This, then, confirms Bybee et al's (1994) observation that, cross-linguistically, futures evolve from a fairly restricted range of lexical sources- from

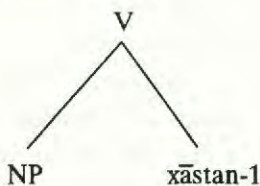
constructions involving movement verbs, from markers of obligation, desire, ability, and from temporal adverbs.

If these arguments are correct, then the three uses of the verb *xāst-an* accords with the universal grammaticization path from more concrete, specific meaning associated with lexical material *xāst-an-1*, volition, desire, to the more general and abstract use in *xāst-an-3*, prediction. Bybee et al (1994, p. 15) observe that "given the source material that enters into grammaticization is similar cross-linguistically, it predicts cross-linguistic similarity in paths of development", or grammaticization. The similarity among the restructuring verbs in Italian and Spanish, the verb *vouloir* in French, and the development of English future seem to confirm the grammaticization of the lexical verb *xāst-an-1* as a marker of future in Persian. Syntactically it is more economical to consider the three uses of *xāst-an* as a single main verb that c-selects three different types of configurational arguments, NP, CP and VP.

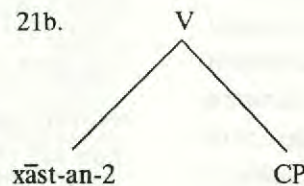
20- *xāst-an*: [ --{NP, CP, VP} ]

As we note, a unified analysis of *xāst-an* "want" makes its description and analysis simpler and more economical. (21a) favors an OV order, while (21b-c) prefer a VO order. The first two are equivalent to the unstructured *volère*, but the latter is equivalent to the restructured *volère* and restructured *xāst-an-3* in Persian.

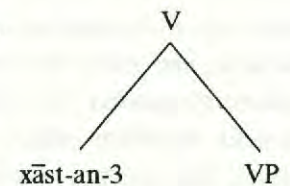
21a.



21b.



21c.





Rosen shows that clitic climbing, auxiliary selection, and long object preposing applies to the restructuring verbs in Italian (21c), but not to their unrestructured counterparts. Persian lacks these syntactic rules, so we will use passivization to see whether there are any differences between the three uses of *xāst-an* in Persian.

Passive structures are rarely used in Persian. Passivization with the stative verb, *xāst-an-1* is bad:

22a- ??? yek docharxe-ye now (tavassot-e man) *xāst-e shod*

one bicycle-EZ new (by me) want.ppr became  
"A new bicycle was wanted by me."

22b- ??? 'in *kālā-hā* (tavassot-e 'ānhā) *xāst-e shod-and*

this goods-pl (by they) want.ppr became.3pl  
"These goods are wanted by them."

These sentences are very odd. I personally consider them unacceptable though they might seem to be grammatically well-formed. The adjectival passive forms of some of these stative passives with *xāst-an-1* are sometimes acceptable:

22c- *kālā-hā-ye xāst-e shod-e resid*  
goods. pl wanted.ppr became.ppr reached

"The required/wanted goods arrived"

In the second use of *xāst-an-2*, passivization still does not seem quite acceptable:

23a- ali mi-*xāh-ad* (ke) PRO be-rav-ad.

Ali want-3S (that) sj-go.3S

"Ali wants to go."

23b- \*/??? ... *xāst-e shod* (ke) ali be-rav-ad

want.ppr became (that) Ali sj-go-3S

"It was wanted that Ali go."

23c- ... az ali *xāst-e shod* (ke) PRO be-rav-ad.

from Ali want.ppr became (that) sj-go-3S

"It was wanted from Ali to go."

In (23a) and (23b) the main verb *xāst-e-2* is

passivized. In sentence (23b), the main passive verb lacks subject, (is impersonal), and sounds very odd to me. Sentence (23c) also lacks subject, (is impersonal), and is ok. Such passive uses of *xāst-an-2* are very rare. We conclude that the passive rule applies to the main verbs *xāst-an-1* and *xāst-an-2*, if it does at all.

Our main concern is the restructuring use of *volè-re-2*, *xāst-an-3*, and their passive uses. Passivization of *xāst-an-3* applies to the lexical verb that follows *xāst-an-3*, and not to *xāst-an-3* itself. Passivization depends on the valency of the complement lexical verb and is always acceptable (subject to usual constraints on passivization in Persian), and is similar to passivization of a simple transitive sentence without *xāst-an* (24c). This is contrary to the observation in Italian where passivization must apply to the restructuring verb itself, and not to the complement lexical verb:

24a- 'u yek docharx-ye now *xāh-ad xarid*

he one bicycle-EZ new want-3S bought (ps)

"He will buy a new bicycle."

24b- yek docharxe-ye now *xaride xāh-ad shod*

one bicycle-EZ new bought (ppr) want-3S became (ps)

"A new bicycle will be bought."

24c- docharxe *xarid-e shod*

bicycle bought (ppr) became

"The bicycle was bought."

It is the lexical, complement verb *xarid* in (24a) that is passivized in (24b), *xarid-e shod*, and not the verb *xāh-ad*. However, as Rosen observes, in Italian the restructuring verb itself, *volè-re* etc, must undergo passivization which is ungrammatical (25). Rosen (1989:206) argues: "in restructuring there is only one argument structure for the two verbs, and the matrix verb takes a VP complement. This leads one to predict that passivization could apply to the matrix verb, taking away the case of the embedded verb, and forcing the embedded object to move to



the matrix subject position ...". However, she notes that the verb *volere* "want" does not/cannot passivize either in its unrestructured use (25b) or in its restructured use (25a), as a complex predicate, in Italian:

25a- \*questo libro é stato voluto leggere (da Giovanni)

this book has been wanted to read (by Giovanni)

25b- ? (?) era voluto come primo attore da tutte le case cinematografiche

he was wanted as a leading actor by all movie producers

Yet, Rosen concludes that passivization is generally available to the restructuring verbs, constrained only by the ordinary constraints on passivization, and accidental gaps. But note the failure of *volere* to passivize in (25). It is so because she notes that some other restructuring verbs like *continuar* "continue" and *cominciar* "begin" do passivize in Italian (see Rosen, 1989, p. 207, ex 83).

Burzio (1986), however, notes that passivization of restructuring constructions is generally bad in Italian and considers the possibility of passivization with some restructuring verbs as accidental (cf. Rosen, 1989, p. 207, ex 83). Clearly, a periphrastic syntactic analysis of passivization is much more attested and plausible over an argument structure account in Italian and Persian.

Thus, we notice a major difference between *xāst-an-3* and the restructuring use of *volere* in Italian. It is the embedded complement verb that passivizes in Persian (24), while it is the restructuring verb itself that must passivize in Italian, and it cannot actually do so as we see in (25). Since Rosen takes passivization a possible operation in Italian restructuring verbs, and an argument structure phenomenon, she concludes that passivization must have applied to the merged, complex argument structure of the two verbs. She then concludes that "a VP complement is only possible when argument structure merger has taken

place." This means that the restructuring verbs take a VP complement and then merge with them to form a single argument structure, and a complex predicate.

Persian data, however, is not compatible with her argument structure merger approach to restructuring and passivization. I suggest that the formation simple future constructions with Persian restructuring *xāst-an-3* is not an argument structure phenomenon, rather it is a pure syntactic phenomenon. No argument structure merger, or complex predicate formation, takes place in Persian simple future tense even though I consider the main lexical verb a VP complement. Passivization, a syntactic phenomenon in Persian, does not apply to *xast-an-3*, rather it applies to its complement verb

Burzio (1986) argues that the selectional properties of the restructuring verbs supports a syntactic derivation of restructuring rather than an argument structure approach. There are three classes of restructuring verbs with respect to the types of subjects that they take, i.e., control verbs, raising verbs, and unaccusative verbs in Italian. A raising verb never selects for its subject. Its subject is selected by the embedded verb while those of the control, (*volere*), verbs and unaccusative verbs are determined by these verbs themselves. That is, the subject position of a raising verb is empty, and the embedded verb's subject moves to that position. So this subject may range from a null expletive, a pro, an expletive or a full animate or inanimate noun depending on the selectional properties of the embedded verb. This is not the case for control and unaccusative restructuring verbs in Italian (cf. Burzio, 1986, pp. 329-330).

Selectional properties of *xāst-an-3* with respect to the subject that it agrees with, in simple future, is very similar to the raising restructuring verbs in Italian that Burzio cites in that the selectional properties of the subject of simple futures with *xāst-an-3* always depends on the lexical, complement verb. That is, *xāst-an-3*, in simple future, does not impose any selectional restrictions on the subject that it agrees with, rather the subject is selected and determined by the complement verb:



26a- ... fardā bārān xāh-ad bār-id  
 "It will rain tomorrow."

26b- ānhā az hayulā xāh-and tars-id  
 "They will get frightened from the dracula."

26c- ali ketāb rā xāh-ad xar-id  
 "Ali will buy the book."

26d- bache-hā xāh-and āmad  
 "The kids will come."

In (26a), the main verb is a weather verb, and lacks an external argument, as a result the subject of xāh-ad-3/clause is a null expletive. In (26b) ānhā "they" is an experiencer noun semantically selected by the psyche-verb tarsid "fear". It is the subject of xāh-and-3 and agrees with it. In (26c) ali is an agentive noun selected by xar-id "bought". Ali agrees with xāh-ad-3 and is its subject. The subject, bache-hā "children" is a theme in (26d). It is selected by the unaccusative verb āmad-an "come" and agrees with the modal, restructuring verb xāh-and-3. It is clear that xāst-an-3 does not provide and select its subject which is always determined by the lexical VP-complement. This observation conforms with Burzio's syntactic analysis of raising restructuring verbs in Italian in which he claims that "the selectional properties, which must be met at d-structure [=syntax], suggest that restructuring verbs must always map into syntax unrestructured, and that the restructuring process takes place after the d-structure representation, where selection is satisfied". This is a clearly syntactic approach to restructuring which we have adopted (cf. Rosen, 1989, pp. 226-230).

Rochette (1988), too, claims that the complement in the type of restructuring constructions in our study is a V complement which projects to the maximal level, the VP. She stipulates that the motivation for restructuring comes from <e> or the lack of <e> role. That is, modal-like verbs each (may) have two counterparts, one with an <e> role, and one without an <e> role. A modal-like verb in its un-restructuring use has its own <e>. This is equivalent to our xāst-an-1 and xāst-an-2.

This is exactly what we note in the three uses of xāst-an in Persian. Xāst-an-1 and xāst-an-2 have an open <e> role in their argument structure (see 1-2 above, and 13-16). In xāst-an-1 there is no embedded verb. Xast-an-1 has an <e> role and is dominated by INFL. As a result xast-an-1 raises to INFL in order to discharge its <e> role in INFL. Here xāst-an-1 is a main un-restructuring verb.

Xāst-an-2 has an <e> role ∅, and has its own INFL independent of the finite verb of the embedded clause. The verb of the embedded clause too has an <e> and its own INFL, e.g. man mi-xāh-am (ke) yek docharxe-ye now be-xar-am (see (2) above). Here too, xāst-an-2 is a main un-restructuring verb.

Xāst-an-3 is, however, a modal-like restructuring verb that lacks an <e> role. As a result, it takes a VP complement where the verb has an open <e> of its own. The argument structure I suggested for xāst-an-3 is repeated below:

19- The 2nd argument structure of volère-2  
 "want-2", our xāst-an-3. "REVISED"  
 ( ( Y ))  
 event

In (19) & (27) the modal-like verb, xāst-an-3 in simple future, has no <e> role, but the complement main verb, raft "went", does. Neither does xāst-an-3 select its external argument. The external argument position is empty, and is determined by "raft".

27- man be madrase xāh-am raft  
 "I will go to school."

A restructuring verb lacks an open <e> role, so it must select a verb/VP with an <e> role. On the other hand, the main verb lacks an INFL, so it must raise to INFL of the restructuring verb in order to discharge its <e> role. Rochette says: "Because the embedded verb may satisfy its <e> role through the matrix INFL, it does not need an INFL of its own. Therefore, the embedded verb only projects up to a VP complement in the restructuring construction" (Rosen 1989, p. 224).

There is no motivation for argument structure



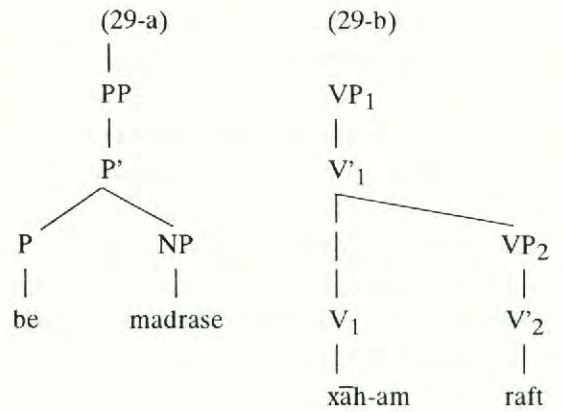
merger and complex predicate formation in restructuring constructions, contrary to what Rosen (1989) claims, in Persian. Restructuring use of *xāst-an-3*, in simple future, is a syntactic phenomenon. In fact the combination of *xāst-an-3* and the following lexical verb in simple future tense never forms a single semantic or morphological unit. However, they may form a syntactically-uniform head, through syntactic head-to-head movement, at some stages of syntactic derivation since the complement verb, *raft* "went" must raise to INFL, i.e., *xāh-am-3* in (27), in order to discharge its <e> role, and to check its NP features. The derivation of a simple future sentence (27) is given below in order to conclude the discussion.

Suppose we select the following items from the lexicon which then project into appropriate X-bar schemas (28):

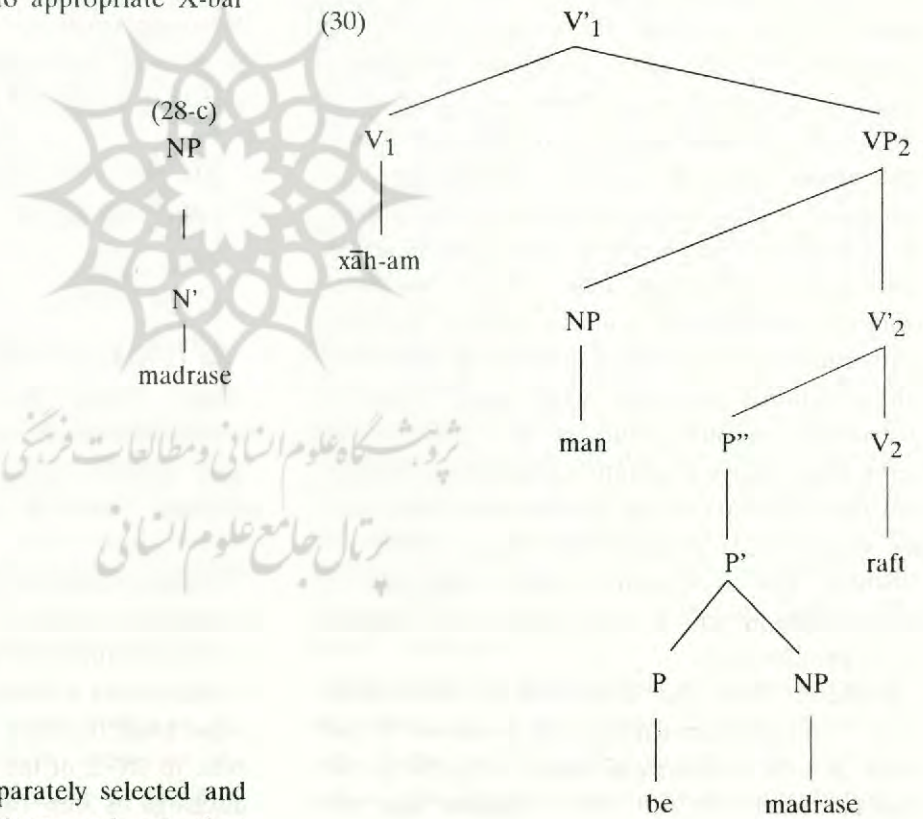
- |        |               |
|--------|---------------|
| (28-a) | (28-b)        |
| NP     | P'            |
|        |               |
| N'     | P             |
|        |               |
| Man    | be            |
| (28-d) | (28-e)        |
| VP/V'  | VP/V'         |
|        |               |
| V      | V             |
|        |               |
| raft   | <i>xāh-am</i> |

Each of these items are separately selected and extended into higher projections, and can then merge into complex x-bar schemas by generalized transformations. Trees below show two of such cases.

First (28e) is selected and merged with (28d), resulting in (29b). In (29b), V<sub>1</sub>, *xāh-am*, is the head



and VP<sub>2</sub>, *raft*, is its complement. Similarly (28a) and (28b) merge into (29a). At a later step (29a) and (29b) merge into (30):

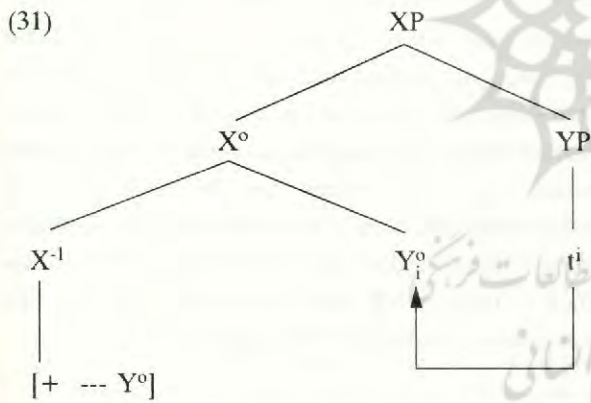


The NP *man* "I" as the underlying subject of *raft* "went" will also merge with V<sub>2</sub>, assuming a version of VP-internal subject hypothesis. In the tree (30), V<sub>1</sub> *xāh-am* is the super-ordinate verb/head, and VP<sub>2</sub>



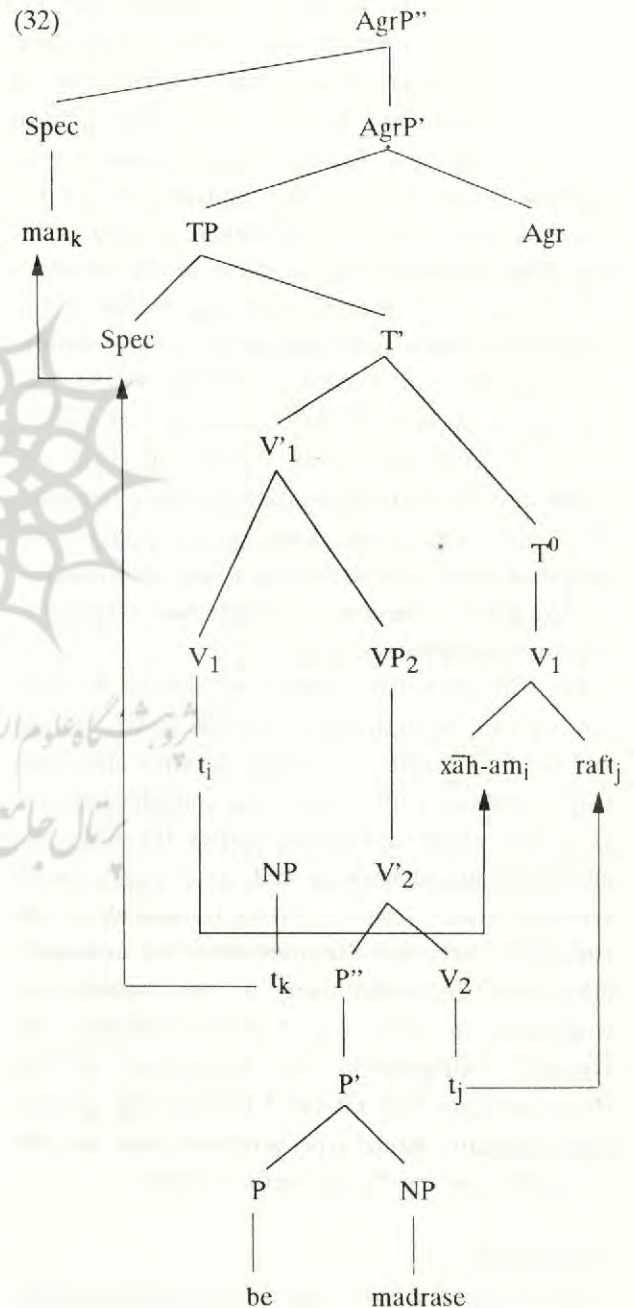
headed by *raft* "went", is its complement. Suppose that in a similar manner  $V_1$  merges with appropriate inflectional heads, i.e. tense phrase (TP), and agreement phrase (AgrP) as in (32):

In this configuration, first the verb  $V_1$ , *xāh-am*, moves up the tree to  $T^0$ , and then to  $Agr^0$  in order to check its morphological features with these heads. However, the subject *man* "I" must also move to Spec-T' and then to Spec-AgrP' to release/check its case and agreement features. At a later step  $V_2$ , *raft* "went", moves up the tree and incorporates into  $V_1$ , *xāh-am*, as an instance of "incorporation by substitution" (cf Roberts, 1991). The tree (31) shows the syntactic configuration for "incorporation by selectional substitution" as argued by Roberts (1991). In (31)  $X^0$  counts the incorporating head and  $Y^0$  is the incorporee. The combination forms a syntactic compound head. The head of this compound is still the super-ordinate head,  $X^0$ , as we note in the tree (31):



The  $V_2$  *raft* has  $\langle e \rangle$  role that it must discharge into T and Agr. It must also be in Spec-head relation with the subject *man* "I" in order to check its nominal NP features and license Nom case on *man* "I". To do so it must move to T and Agr. These head positions are already filled by  $V_1$  *xāh-am*. Incorporation of  $V_2$  *raft* into  $V_1$  *xāh-am* is an instance of syntactic head-to-head movement of the kind "selectional substitution" (Roberts, 1991). This kind of head-to-head movement never leads to the

formation of "morphological compound heads". As a result, the combination of *xāh-am* + *raft* is not considered a morphological compound head. Rather it is a syntactic compound head. In such combinations it is always the incorporating head, i.e. the  $V_1$  *xāh-am*, that counts as the head of the compound, as we note in the trees (31-32). This is shown in (32) below:





The movement of  $V_2$  into  $V_1$  and the combination to  $T^0$  and  $Agr^0$  is motivated by "greed" since it has  $\langle e \rangle$  role and NP features that must be discharged and checked.

Negation in Persian is marked by the verbal prefix *na-* on the main verb only. The distribution of the negative phrase may be most economically described by claiming that the negative prefix *na-* attaches to the verbal heads in the X-bar tree diagrams. This is to say that in all different paradigms of verb inflection in Persian, the negative prefix only attaches to the main verbal heads in X-bar schemas, and never to an auxiliary. It is interesting to note that *na-* also attaches to the syntactic head,  $V_1$  *xāh-am*, and not to  $V_2$  *raft* which does not count as the syntactic head in the tree diagrams (30-32). So we have *na-xāh-am raft* "I will not go", but not *\*xāh-am na-raft*. This means that the negative head/morpheme, *na-*, recognizes the super-ordinate head  $V_1$ , *xāh-am*, as the head and the main verb of the clause, and not the subordinate, complement head  $V_2$ , *raft*. Li (1990) has shown that in the incorporation structure "though a verb morpheme may form a compound with neg(ation), it is impossible for a verb incorporation compound to "contain" a neg that exclusively negates the embedded verb."

The incorporated compound head in (32), *xāh-am<sub>i</sub> raft<sub>j</sub>* (underlined), is headed by  $V_1$ , *xāh-am*, and not by  $V_2$ , *raft*. As a result, negation affects the super-ordinate verb and not the embedded verb as Li (1990) predicts. *Xāh-am<sub>i</sub>* carries the tense and agreement morphemes as well, so it counts as the syntactic head. Negation passes transitively to the embedded verb, *raft*. Negation may not exclusively affect the embedded verb in the incorporated compound in (32), e.g. *\*xāh-am<sub>i</sub> na-raft<sub>j</sub>*. The syntactic configuration and derivations in (32) clearly indicate that *xāh-am* behaves like a main, super-ordinate, modal-type verb/head, and not like an auxiliary in the Persian verbal system.

### Conclusion

In this paper, I have argued for a unified analysis

of the three distinct functions of the verb *xāst-an* "want" in its three different, and frequent uses in Persian. I have shown that the verb *xāst-an-1* in its main lexical use indicates the desire and wish of the agent/subject to possess something which surfaces as an NP complement. This meaning then shifts towards an intention function which takes an event proposition surfacing as a subjunctive complement (CP), i.e. *xāst-an-2*.

In this use, both the main verb, *xāst-an-2*, and the embedded verb are independently inflected for tense and agreement. In a further step of grammaticization *xāst-an-3* is used to indicate prediction in simple future tense. In this use *xāst-an-3* takes a VP complement. While the analyses for the first two uses seem to be straightforward, the last function of this verb has been variously described by linguists, calling it an auxiliary or a modal verb. My main concern in this paper has been this last function. I have presented syntactic analyses and arguments to show that *xāst-an-3* is a modal type main verb that functions as the head of the syntactic construction in X-bar schemas. Our unified analysis observes economy in derivation and computation as well as the economy in the lexicon. Our analysis assumes a single, lexical modal verb that takes an NP, VP, or CP complement. *Xāst-an* is not considered as auxiliary, rather a main modal-type verb. It favors a VO order when it takes a VP and CP complement, but OV order when it takes an NP complement.

### Note

<sup>1</sup> Darzi also provides arguments and examples from gapping constructions with *bud-an* and *xāst-an* to show that the combination of [verb + *budan*] forms a single v-node, hence a single syntactic unit, but the combination of [*xāst-an* + verb] does not form a single syntactic unit (see Darzi, 1996, p. 37, ex 18-19). This observation then leads him to conclude that *bud-an* is not a true aux but *xāst-an* is a true aux. This phenomenon seems to favor an underlying SVO.

I believe that these same arguments can be



adopted to argue for the raising of the main verb to INFL (aux) bud-an in the combination [verb +bud-an]; hence we have a single syntactic node. The lack of a single syntactic node with [xāst-an +verb] denotes separation, and non-raising of the main verb to the INFL (aux) xāst-an. As a result, the main verb, i.e., raft in man be madrase xāh-am raft, does not have to raise to INFL (aux), xāh, to check its features. As a result no single syntactic node is formed. In other words, Darzi's own arguments favor bud-an as a true aux, but xāst-an-3 as a non-aux, modal-type verb.

<sup>2</sup> The complement may surface at the level of a V' or V-zero as well, in the minimalist program terms.

<sup>3</sup> Three question signs, ???, indicate severe oddness, and the star sign \* indicates ungrammaticality.

## References

- Batani, M.R. (1991). *Towsif-e sakhtaman dasturi-ye zaban-e Farsi* ("A Description of the Grammatical Structure of Persian"). Iran, Tehran: Amir kabir Publication.
- Baker, M. (1988). *Incorporation: A Theory of Grammatical Function Changing*. Chicago: University of Chicago Press.
- Burzio, L. (1986). *Italian Syntax: A Government & Binding Approach*. Dordrecht: Reidel.
- Bybee, Joan L., Revere Perkins, & William Pagliuca (1994). *The evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago, The University of Chicago Press.
- Chomsky, N. (1989). Some Notes on Economy of Derivation and Representation. In Itziar Laka and Anoop Mahajan (Eds), *Functional heads and Clause Structure*. MIT Working papers in Linguistics, vol 10. Cambridge, Mass.
- Chomsky, N. (1995). *The Minimalist Program*. MIT Press.
- Comrie, Bernard (1989). *Language Universals and Linguistic Typology*. 2nd ed. Basil Blackwell Ltd.
- Dabir-Moghaddam, M. (1985). *Majhul dar zaban-e Farsi* (Passive in Persian). *Iranian Journal of Linguistics*, Vol 2, No. 1.
- Darzi, Ali (1996). *Word Order, NP-Movement, and Opacity Conditions in Persian*. Ph.D. dissertation. University of Illinois at Urbana-Champaign.
- Fukui, N. (1986). *A Theory of Category Projection and Its Application*. Doctoral dissertation, MIT.
- Gerdts, D. B. (1988). *Semantic Linking and the Relational Structure of Desideratives*. In *Linguistica*, 26, 843-872.
- Ghomeshi, J (1997). *Topics in Persian VPs*. In *Lingua* 102, 133-167.
- Greenberg, Joseph H. (1963). *Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements*. In *Universals of Language*, ed. By Joseph Greenberg, Cambridge, MA: MIT Press
- Higginbotham, J. (1985). *On Semantics*. *Linguistic Inquiry*, 16, 547-593.
- Karimi, S. (1989). *Aspects of Persian Syntax, Specificity, and the Theory of Grammar*. Ph.D. dissertation, University of Washington.
- Karimi, S (1996). *Case and Specificity: Persian Ra Revisited*. In *Linguistic Analysis* 26:3-4.
- Li, Y. (1990). *X0 Binding and verb incorporation*. In *Linguistic Inquiry* 21, 399-426.
- Marashi, M. (1970). *The Persian Verb: A Partial Description for Pedagogical Purposes*. Ph.D. dissertation, University of Texas at Austin.
- Miller, D. G. (1993). *Complex Verb Formation*. *Current Issues in Linguistics*, 95, John Benjamin co.
- Pollock, J.Y. (1989). *Verb Movement, Universal Grammar, and the Structure of IP*. In *Linguistics Inquiry* 20.
- Rizzi, L. (1982). *Issues in Italian Syntax*. Dordrecht: Foris.
- Roberts, I. (1990). *Excorporation and Minimality*. In *Linguistics Inquiry* 22.
- Rochette, A. (1988). *Semantic and Syntactic Aspects of Romance Sentential Complementation*. Doctoral dissertation, MIT.
- Rosen, Sara Thomas (1989). *Argument Structure and Complex Predicates*. Doctoral Dissertation, Brandeis University
- Samiian, V. (1983). *Origins of Phrasal Categories in Persian: an X-bar analysis*. UCLA doctoral dissertation.
- Windfuhr, G. (1979). *Trends in Linguistics: Persian grammar-history and state of its study*. New York: Mouton.