ON THE PERSIAN COMPOUND VERB

by

BRIAN LEANDER O'NEILL

B. S., Kansas State University, 1972

_____

A MASTER'S THESIS

submitted in partial fulfillment of the
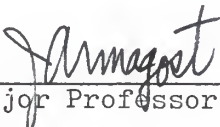
requirements for the degree

MASTER OF ARTS

Department of Speech

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1978

Approved by:

_____
Major Professor

CONTENTS

Introduction

Many authors have noted the preponderance of what we shall be referring to as compound verb constructions in the Persian language. In fact, the major portion of verbal forms in Persian are compounds, composed of some initial non-verbal element and a second purely verbal element. Perhaps because Persian has not been subjected to intensive analysis these constructions have remained poorly described. The aim of this thesis is to examine the compound verb and to determine its status as an element in the grammar of Persian.[1]

In the past few years several analyses of various aspects of Persian have appeared, often employing a transformationally based theoretical framework.[2] Preceding these were a number of normative and descriptive works. Included among the former are the much older works by Hadley (1776), Jones (1771) and an anonymous work published in 1790 that was written for the Persian speaker learning English. More recently, Lambton (1966) and Elwell-Sutton (1963) have written grammars to be employed by students of the language. Additionally, there have been a number of phrase books of the type edited by C. L. Hawker (1937) as well as a volume in the 'Teach Yourself...' series, in this case by Mace (1967). In the more traditional descriptive vein, two noteworthy volumes by Rastorgueva (1964) and Rubinchik (1971) have been translated from their original Russian making them more accessible. It is these latter works and the grammars of Lambton and Elwell-Sutton that will provide us with most of the data with which to confirm or deny the existing interpretations

of the Persian compound verb.

The problem itself is multifaceted. In addition to a lack of terminological agreement among scholars, there have been several 'characteristics' posited concerning the nature of the compound verb construction. The first of these is that the compound verbs are a result of a productive process. Secondly, it has been suggested that these compound forms are idiomatic. Finally, it has been stated that these forms are inseperable. I shall examine each of these areas and demonstrate the inter-relatedness of each of these concerns. Initially, however, some preliminary information will be presented concerning the general characteristics of Persian and the events that have led to its present form.

Preliminaries: Historical

Persian is the national and literary language of Iran. Though there are numerous dialects and other languages spoken in the country, it exists as the language of government and commerce within the borders of the country. Even though other languages have gained prominence in Iran (ie. Arabic, English, French), Persian has managed to maintain its identity.

As a member of the Iranian group of Indo-European languages, Persian is closely related to Tajik, Pushto and Ossetic, to name but a few, and because of its geographic position in the Mid-East it shares with a number of languages a large number of lexical items. Unlike many of its sister languages, Persian has a lengthy written tradition spanning over 2500 years. Of course, it should go without saying that the earlier stages of the language were quite different from the present form. The periods commonly defined in the literature are: Old Persian, Middle Persian, and Modern Persian. It may be of some interest to briefly note the historical circumstances present in each of these periods.[3]

The Old Persian, or Ancient period, lasted from about the 6th century BC until the 3rd century BC. It was during this stage that cuneiform inscriptions were written in honor of various military feats. In addition to these monumental writings, cuneiform was used on various weights and measures and to record inventories of goods. The highly inflected Old Persian language was similar in structure to the language of the Avesta (Avestan) and for our purposes it is interesting to note, as

Sheintuch (1976) has, that there was even in this stage of the language evidence of compound verb constructions. Though not as widespread as now, the compound verb construction appears to be an indigenous phenomenon.

Middle Persian, or Pahlavi[4], lasted from the end of Old Persian until the 8th century AD and, as in the previous period was the language spoken by the inhabitants of the South-Western part of Iran (what is now known as the province of Fars). As Rubinchik (1971) recounts it, 'Middle Persian was the official language of the Sassanid state and of the ruling Zoroastrian priesthood' (p. 18). The case/inflectional system seems to have degenerated during this period. The category of gender, which had been present in the Old Persian language, was lost along with the dual number. At the end of this period we find Arabic assuming the position of the state language. From the 7th century to the 9th no records have been found in Persian; the 'state, literary and written language' was Arabic.

It isn't until the beginning of the 9th century that we find Persian reasserting itself in the Arabic script form that it maintains today in Modern Persian. Because of the intensive contact between the two languages, Persian and Arabic, it's not surprising that a great number of borrowings were made into Persian. However, 'the grammatical structure of the Persian language showed exceptional firmness as regards the Arabic language and suffered hardly any changes' (Rubinchik: 1971 p. 20).

Grammatically, the internal changes that Persian underwent from the Old Persian period to the New Persian period resulted

in an analytic language lacking case and gender, the syntactic
relations of nouns being expressed by different analytic means--
prepositional and postpositional constructions and the ezāfeh
construction.

Among the languages that Persian has come into contact with
are: Elamite, Babylonian, Greek, Turkish, Arabic, and more
recently Russian, French, German, and English.

<u>Preliminaries: Grammatical</u>[5]

Although other orders are possible, as stylistic variants, the unmarked word order in Persian sentences is SOV, and Moyne and Carden (1974) have claimed that this basic SOV order is generated by phrase structure rules. Consider the following examples:

(1)   a.   <u>ali kebāb xord</u>
            Ali kebab ate
            'Ali ate a kebab'

       b.   <u>mansur sag did</u>
            Mansour dog saw
            'Mansour saw a dog'

In both of the above examples the object is indefinite. In sentences where there is a definite direct object it is marked, as the following examples illustrate:

(2)   a.   <u>sārā māh -rā did</u>
            Sarah moon-Obj saw
            'Sarah saw the moon'

       b.   <u>sārā ketāb -e bozorg -rā xand</u>
            Sarah book   big -Obj read
            'Sarah read the big book'

As the following examples illustrate, the direct object may be preposed with no resultant semantic change:

(3)   a.   <u>māh -rā sārā did</u>   (Compare with 2.a)
            moon-Obj Sarah saw
            'Sarah saw the moon'

b. <u>ketāb -e bozorg -rā sārā xand</u> (Compare with 2.b)
   book   big   -Obj Sarah read
   'Sarah read the big book'

In examples (2.b) and (3.b) there is an additional element that
I have allowed to remain unglossed. The <u>ezāfeh</u> construction,
represented here as the suffix <u>-e</u>, is the standard modified-
modifier construction in contemporary Persian. It has been
suggested by Moyne and Carden (1974) that, historically, the
<u>ezāfeh</u> was a reduced relative clause.[6]

Rubinchik (1971) has divided the Persian verbs according
to their structure and types of word building. These categories
are:

(4) a. simple   (ie. <u>didan</u> 'to see', <u>xordan</u> 'to eat')

     b. nominal (ie. <u>nāmidan</u> 'to name'-from the N <u>nām</u>
                    'name', <u>raqsidan</u> 'to dance'-from the
                    N <u>raqs</u> 'dance')

     c. prefixed (ie. <u>bar dāštan</u> 'to lift or take'-from
                    <u>bar</u> 'up' and <u>dāštan</u> 'to have', <u>dar</u>
                    <u>āvardan</u> 'to take out'-from <u>dar</u> 'in'
                    and <u>āvardan</u> 'to bring')

     d. compound (ie. <u>šoru'</u> <u>kardan</u> 'to begin'-from <u>šoru'</u>
                    'beginning' and <u>kardan</u> 'to do',
                    <u>harf zadan</u> 'to talk'-from <u>harf</u> 'word'
                    and <u>zadan</u> 'to hit')

As we shall see, others have combined the third and fourth cat-
egories considering them both to be compound forms. The remain-
ing two verbal forms follow the same pattern of affixation of
the tense marker and negative particle. In general, the form

of the simple and nominal verbs is as follows:

(5) Neg - Prog - V$_{stem}$ - Person/Number

as in:

(6) <u>ne</u> - <u>mi</u> - <u>xor</u> - <u>am</u>
     'I am not eating'

In addition to the above prefixes and verb stems to indi-
cate tense, there are also a small number of auxiliaries that
are combined with the verb to form compound tenses. Among others
are the definite future, the perfect, and the past perfect.
These are formed in the following manner:

(7) definite future:
     Aux + Pers/Num - Verb

as in:

(8) <u>fardā   man   be dāne šgāh   x āh + am   raft</u>
     tomorrow I to university shall + 1sg go
     'Tomorrow I shall go to the university'

In the definite future the Aux <u>xāstan</u> is employed.

(9) perfect:
     past participle   - Aux
     /Vstem past + e/

as in:

(10) <u>kasi   āmad + e   -ast</u>
      someone came   has
      'Someone has come'

In the perfect tense the copula is used as the Aux.

(11) past perfect:

past participle - Aux + Pers/Num

as in:

(12) <u>man dārs -e man -rā tamām kard + e bud + am</u>

I lesson my -Obj finish was + 1sg

'I had finished my lesson'

In the past perfect tense the Aux <u>budan</u> is employed.

It will have been guessed by now that the verb must agree in number and person with its surface subject.

(13) present indicative of <u>xordan</u> 'to eat, drink'

<u>singular</u>   <u>plural</u>

1 pers. <u>man mi - xor - am</u>   <u>mā mi - xor - im</u>

I prog - pres stem - 1sg   We prog - pres stem - 1pl

'I am eating'   'We are eating'

2 pers. <u>to mi - xor - i</u>   <u>šomā mi - xor - id</u>

'You are eating'   'You are eating'

3 pers. <u>u   mi - xor - ad</u>   <u>išān mi - xor - and</u>

'He/ she/ it is eating'   'They are eating'

Though there are different verb stems for past and present tenses, the personal endings for the past tense forms are identical to those of the present tense forms with the exception of the 3rd singular form which is phonetically null. Pronominal subjects are optionally deleted.

(14) (u) māh -rā did -∅   (Compare with 2.a)

      (she) moon -Obj saw
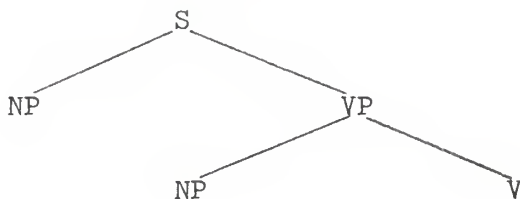
      'She saw the moon'

Notice that in (14) the 3rd singular marker is represented as
∅ and that the subject, in this case the pronominal form u̱, may
be deleted.

   At this point is may be worthwhile to present some basic
phrase structure rules and demonstrate the manner in which they
will operate. I will not attempt a full explication of the phrase
structure rules necessary for the generation of all Persian
sentences, but the following might be considered to be a first
approximation of the set that would derive a large number of
simple sentences:

(15)   S $\longrightarrow$ (Adv$_T$) NP VP

       VP $\longrightarrow$ (NP) V

       NP $\longrightarrow$ $\begin{Bmatrix} \text{N (Adj)} \\ \text{N}_{prop} \\ \text{Pro} \end{Bmatrix}$

It should be noted that the claim of Persian being an SOV lang-
uage is reflected in the above rules. That is to say, the NP
dominated by S may be considered to be the 'subject'; the NP
dominated by VP may be considered to be the 'object'; the V
dominated by VP may be considered to be the 'verb'. The ap-
plication of these phrase structure rules is represented in
the following tree diagram:

(16)

```
                        S
                  ╱         ╲
               NP              VP
                           ╱      ╲
                        NP           V
```

Given the rules as stated in (15) we might consider the deriva-
tional histories of examples already presented. To do this we
will need to specify several transformational rules that will
operate on the output of these base rules. The transformational
rules that have been alluded to so far have been: Subject-Verb
Agreement, Object Marking, and Object Preposing. The following
are the formal descriptions of these rules:

(17)   Subject-Verb Agreement   (obligatory)

$$X \quad - \quad NP \quad - \quad VP \quad - \quad X$$
$$[\alpha \, \text{person}]$$

1	2	. 3	4
1	2	$3 + [\alpha \, \text{person}]$	4

(18)   Object Marking   (obligatory)

$$X \quad - \quad _{VP}[NP \quad - \quad V_{VP}] \quad - \quad X$$

1	2	3	4
1	2 + ra 3		4

Cond: $N = [+ \text{definite}]$

(19)   Object Preposing   (optional)

$$X \quad - \quad NP \quad - \quad X \quad - \quad _{VP}[NP \quad - \quad V_{VP}] \quad - \quad X$$

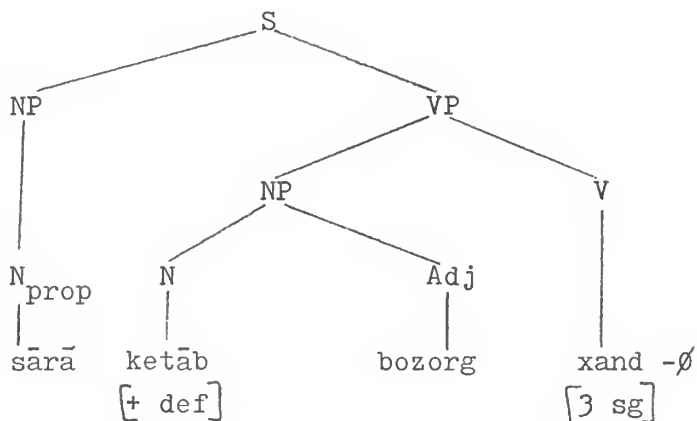| 1	2	3	4	5	6
1	4 2	3	$\emptyset$	5	6

12

Given the phrase structure rules of (15) and the above transformational rules, we can now observe how these would account for several sentences.
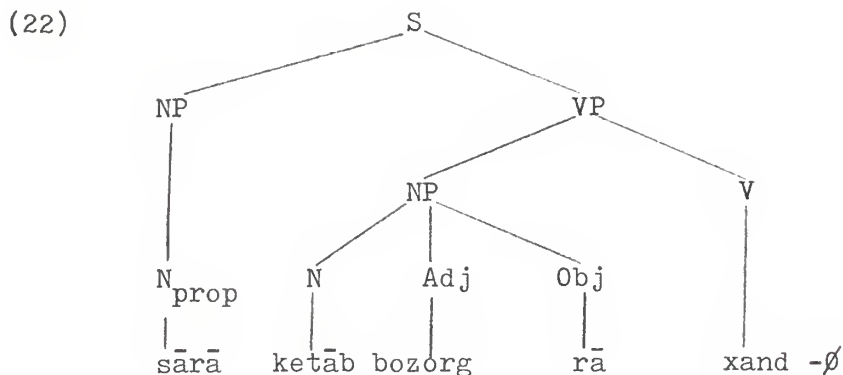
(20) derivational history of (3.b)
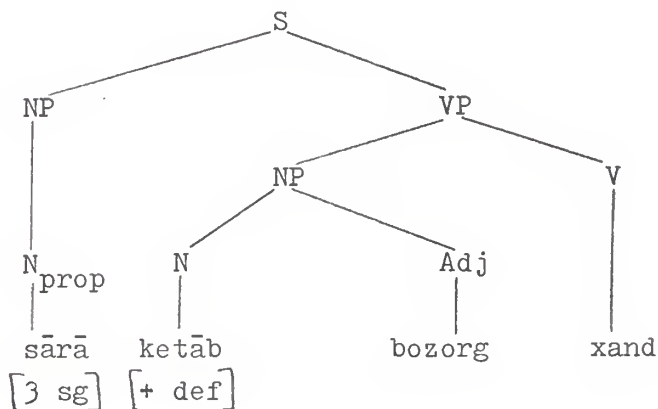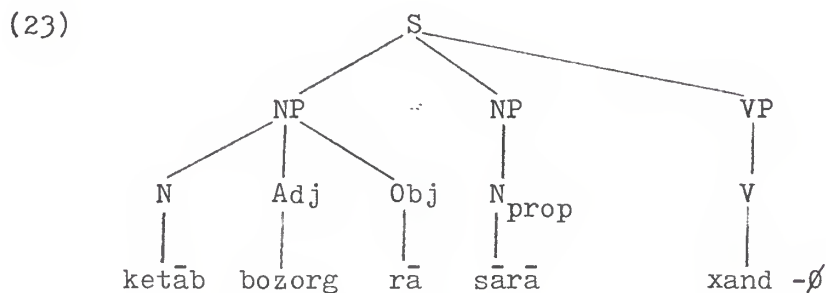

After Subject-Verb Agreement:

(21)

At this point we are in a position to apply the next transformational rule, that of Object Marking.

After Object Marking has applied we would have the following tree diagram:

(22)

```
                          S
              ┌───────────┴──────────┐
             NP                      VP
              │              ┌────────┴──────┐
              │             NP               V
              │        ┌─────┼──────┐        │
          N_prop       N    Adj    Obj       │
              │        │     │      │         │
            sārā    ketāb bozórg    rā      xand -∅
```

With the exception of the insertion of the ezāfeh construction, which is discussed in note 6, the above is a fully grammatical sentence. It is now possible to apply Object Preposing, as the following tree diagram demonstrates:

(23)

```
                          S
              ┌───────────┼──────────────┐
             NP        ··  NP             VP
          ┌───┼───┐        │              │
          N  Adj  Obj   N_prop            V
          │   │    │       │              │
       ketāb bozorg rā   sārā          xand -∅
```

I believe this example should be sufficient to demonstrate how it is expected that the phrase structure rules and accompanying transformations will apply in the following section.

The Compound Verb

The problem of defining the compound verb is complicated
by a lack of terminological agreement among scholars who have
mentioned this type of construction. The following brief re-
marks should indicate the state of affairs.

(a) Rastorgueva indicates at one point that the verb in Persian
is a compound construction consisting of the tense marker
(in present tense forms) followed by the verb stem, which
in turn is followed by the personal ending.[7]

(24) $_V$/ $\langle$tense$\rangle$ - Vstem - Pers/Num /$_V$
                        $\langle$+ pres$\rangle$

At another point he claims that compound verbs are composed
of auxiliaries and verbs of the type above, what we have
previously described as 'compound tenses'. It is only in
an appendix that he alludes to a 'compound or composite'
verb that is 'formed through combining a substantive with
a verb' (p. 75).

(b) Rubinchik (1971), as we have seen, notes four classes of
verbs, one of which he labels the compound verb. This is
'formed by combining simple verbs with nouns and adjectives'
(p. 82). Another class, prefixed verbs, 'are formed by
joining different dependent words (mainly prepositions)...
to some simple verbs' (p. 81).

(c) Other grammarians, such as Lambton (1966) and Elwell-Sutton

(1963), have described compound verb constructions as consisting of a simple verb combined with a noun, adjective, adverb or prepositional phrase. These authors would necessarily include Rubinchik's 'prefixed verbs' in their classification of compound verb.

(d) Sheintuch (1976) prefers the term 'periphrastic verb' to that of compound verb. The structure of this verb form is consistent with the previous group in that it is a stable lexical element consisting of a simple verb (for Sheintuch this is a 'periphrastic Aux') preceded by a NP, Adj P, or Adv P.

There appear to be several common denominators in all of these classifications. The first is the superficial form of the compound verb. That is, there is some non-verbal element followed by a verbal element.

(25) $_{CV}/[-\text{Verb}] + [+\text{Verb}]/_{CV}$

Another common feature is that the non-verbal element ([-Verb]) is, in other contexts, funtioning as a noun or adjective. Third, the verbal element, when not combined with this non-verbal element, continues to function as a verb, a simple verb. Tentatively then, we might want to consider a compound verb as being of the following form:

(26) $_{CV}/\left\{\begin{array}{l}N\\ \text{Adj}\end{array}\right\} + V /_{CV}$

Consistent with the above description, a Persian verb such as:

(27) <u>zamin xordan</u> 'to fall' (lit. 'ground' + 'eat')

would be a compound verb. Others of the type / N + V / would
include:

(28) <u>harf zadan</u> 'to speak' (lit. 'word' + 'hit')
(29) <u>guš kardan</u> 'to listen' (lit. 'ear' + 'do')

Those of the type / Adj + V / would include:

(30) <u>boland kardan</u> 'to raise' (lit. 'tall' + 'do')
(31) <u>dur oftādan</u> 'to be separated' (lit. 'far' + 'fall')

There are two further, inter-related features of the com-
pound verb that have been noted by various authors. Both of
these have been labeled 'characteristics' of the Persian compound
verb. The first is that this type of construction is a very
productive device in Persian. By 'productive' we mean that some
existing paradigmatic scheme is being exploited so that new lex-
ical items may be formed. The second feature is that Persian
compound verbs are, in fact, idioms. The term 'idiom' refers
to a sequence of morphemes of which the 'sense' of the whole is
not equal to the sum of its parts. Let's take a look at the
first of these suggestions.

As was mentioned in the first section, 'Preliminaries: Hist-
orical', there has been a great deal of contact among languages
in the Mid-East, and in particular that area that is now Iran.
That Persian has borrowed extensively from numerous languages,

both living and dead, is a well recorded fact.[8] In fact, Rubin-
chik (p. 43) notes that 'more than half of the vocabulary of the
Persian language consists of Arabic borrowings....' Combining
this fact with the previously mentioned observation concerning
the plethora of compound verb constructions, what we find is that
a large number of borrowings have been systematically incorporated
into Persian through a 'nativization' process.

Sharifi (1976), Sheintuch (1976) and Wexler (1974) have
all suggested that there are several results of the compounding
process. The first of which has already been mentioned; nativ-
ization. The second is that these constructions help to intro-
duce new semantic concepts to the language. And thirdly, they
provide new ways of expressing an already existing semantic con-
cept.

It may be the case that historically there have been at
least two productive methods of forming verbs. Consider the
following:

(32) a. <u>raqs</u>   'dance' (an Arabic loan)

    b. <u>raqsidan</u>   'to dance'

    c. <u>raqs kardan</u>   'to dance'

(33) a. <u>talab</u>   'request' (an Arabic loan)

    b. <u>talabidan</u>   'to request'

    c. <u>talab kardan</u>   'to request'

(34) a. <u>fahm</u>   'understanding' (an Arabic loan)

    b. <u>fahmidan</u>   'to understand'

    c. <u>fahm kardan</u>   'to understand'

The genesis of these forms seems to be an initial borrowing in which the nominal elements are followed by the affixing of the 'infinitive marker' -idan. This process was subsequently followed by the productive process of compounding. No reasons have been given for this last process taking precedence, in recent years, over the former productive verb making process.[9] That this last process has come into its own can be seen by such recent additions as:

(35) telefon kardan 'to telephone'

(36) pārk kardan 'to park'

In neither case do we find attested forms such as *telefonidan or *pārkidan

Persian is not alone in failing to note the morphological peculiarities of the language borrowed from. For example:

(37) a. qalabe 'victory' (Arabic verbal noun)
        qalabe kardan 'to vanquish'
     b. qāleb 'vanquishing' (Arabic present active part.)
        qāleb kardan 'to vanquish'
     c. maqlub 'vanquished' (Arabic passive past part.)
        maqlub kardan 'to vanquish'

Note that all three compound forms are semantically equivalent. A similar situation exists with the equivalent compounds

(38) a. qārat kardan
     b. qanimat kardan
     c. yaqmā kardan

all meaning 'to plunder'. In addition to these synonymous com-
pounds that are formed through the combining of an Arabic loan
and a Persian verbal element (in 38), there exists a simple
Persian verb čapidan which also means 'to plunder'.

As is to be expected, though it hasn't been mentioned yet,
there are but a relatively small number of simple verbs that
act as the verbal elements of compound verbs. Additionally,
these simple verbs are all 'pure Persian' verbs. The most fre-
quently occurring would include:[10]

(39) a. šodan   'become
      b. kardan   'do'
      c. dādan   'give'
      d. xordan   'eat'
      e. zadan   'hit'

Each of these then, may be combined with non-verbal elements to
form compounds.

The other characteristic of compound verbs mentioned above
was that they are idioms. Sharifi and Sheik (1976) have respond-
ed to this by stating that there are altogether too many of them
in the language to warrant that type of analysis. We will re-
turn to this problem.

Mace has claimed 'that the non-verbal element of the com-
pound never changes and never separates itself from the verbal
element' (1967: 131--emphasis his).

Consider the following:

(40)   present indicative of <u>harf zadan</u> 'to speak'

<u>singular</u>   <u>plural</u>

1 pers <u>man harf mi-zan-am</u>   <u>ma harf mi-zan-im</u>

    I   prog-pres stem-1sg   We prog-pres stem -1pl

    'I am speaking'   'We are speaking'

2 pers <u>to harf mi-zan-i</u>   <u>šomā harf mi-zan-id</u>

    'You are speaking'   'You are speaking'

3 pers <u>u harf mi-zan-ad</u>   <u>išān harf mi-zan-and</u>

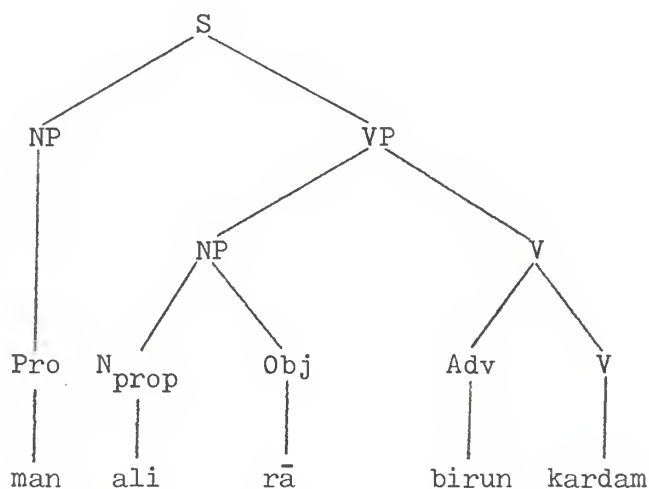   'He/she/it is speaking' 'They are speaking'

The above is reminiscent, as it should be of the description of verbal affixing in (13). It is evident that it is the verbal element that is inflected, in the same fashion as its simple verb counterpart. It appears that the first part of Mace's statement is accurate. That is, the verbal element receives the tense marker and the personal endings, whereas the non-verbal element remains unchanged.[11] It is the second half of Mace's statement, the notion of inseparability, that deserves some investigation. To do this we will need to progress more deliberately. Consider the following sentence:

(41)   <u>man ali-rā birun kardam</u>

    I Ali-Obj /out + did - 1sg/

    'I expelled Ali'

In the above example we find the compound verb <u>birun kardan</u> 'to expel' in a stylistically unmarked sentence. As mentioned
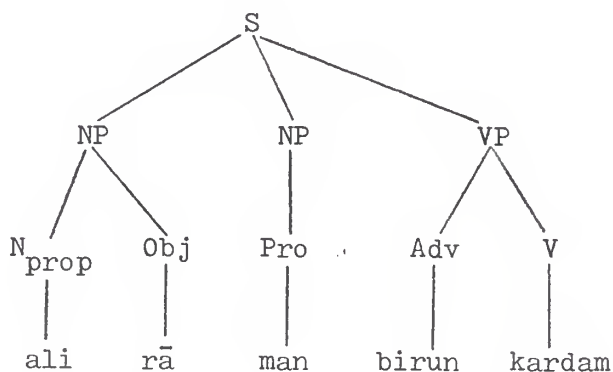
previously, this type of sentence would be derived by phrase structure rules (and certain obligatory transformations). These have been described in (15, 17-19). After the rules of Subject-Verb Agreement and Object Marking have applied, the tree diagram for such a sequence would be of the following shape:

(42)

```
                                S
                  _____|_____
                 NP                            VP
                 |                    _____|_____
                 |                   NP                  V
                 |              _____|_____          ____|____
                Pro          N        Obj          Adv       V
                 |            prop      |            |        |
                 |            |         |            |        |
                man          ali       rā         birun    kardam
```

A number of sentences might be derived, through the application of particular transformations, from this basic phrase structure. By the rule of Object Preposing described in (19) we might derive the sentence:

(43)

```
                                 S
                 _____|_____
                NP              NP                VP
            ____|____            |            ____|____
          N       Obj           Pro          Adv       V
          prop     |             |            |        |
          |        |             |            |        |
         ali      rā            man         birun    kardam
```

'I expelled Ali'

Such rules as Object Marking and Subject-Verb Agreement would presumably precede Object Preposing. Otherwise, the ungrammatical sequences

(44) a. *ali man-rā birun kardam

       b. *ali-rā man birun kard

would occur.

In (44.a) Object Preposing has preceded Object Marking and in (44.b) Object Preposing has preceded the Subject-Verb agreement rule. Next, we might want to observe the effect of a rule of Emphasis Placement, of the form:

(45) Emphasis Placement   (optional)

    X - NP - X

    1   2   3

    1 xod-e + 2 3

Note the occurrence of the ezāfeh construction in the above rule. Consider the following sentences:

(46) a. xod-e man u-rā birun kardam
          EMP   I he-Obj expel -1sg
          'I myself expelled him'

      b. man xod-e u-rā birun kardam
          'I expelled him himself'

      c. xod-e man xod-e u-rā birun kardam
          'I myself expelled him himself'[12]

There is an additional rule of Enclitic Formation that can optionally apply to these forms. Following Moyne and Carden (1974: 208), the output of this rule will give us the following forms:

(47) xod-e man --- xodam   xod-e mā --- xodemān
                   'myself'   'ourselves'

      xod-e to --- xodat   xod-e šomā --- xodetān
                   'thyself'   'yourself'

      xod-e u   --- xodaš   xod-e išān --- xodešān
                   'himself'   'themselves'

By applying this Enclitic Formation Rule to (46.a-c) we would derive the following sentences:

(48) a. xodam u-rā birun kardam
          'I myself expelled him'

      b. man xodaš-rā birun kardam
          'I expelled him himself'

      c. xodam xodaš-rā birun kardam
          'I myself expelled him himself'

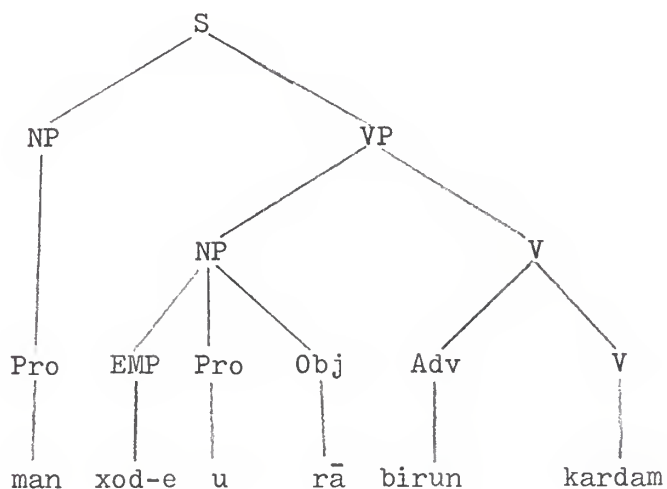Let us return to the original question of 'inseparability'. Consider the following:

(49) man u-rā birun kardam
      'I expelled him'

We have seen how a sentence of this kind may be derived by the

phrase structure rules and obligatory application of Subject-
Verb Agreement and Object Marking transformations.
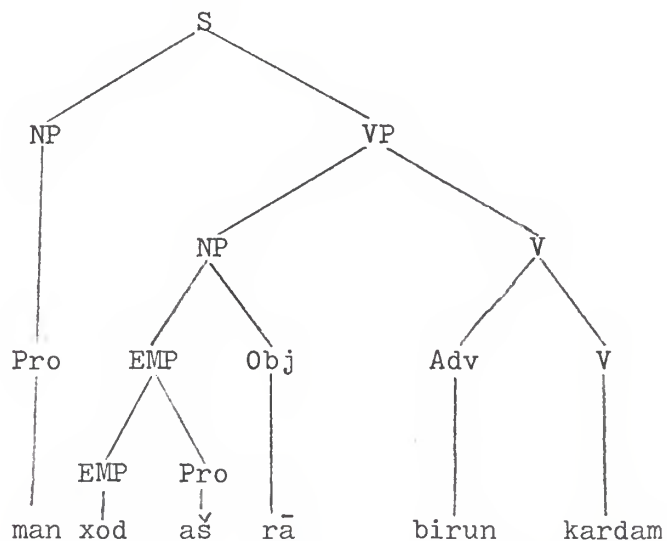
    (50)   <u>man xod-e u-rā birun kardam</u>

          'I expelled him himself'

(50) is derived by the application of EMP Placement.

    (51)   <u>man xodaš-rā birun kardam</u>

          'I expelled him himself'

(51) is derived by the application of Enclitic Formation. But
consider the following:

(52) a. *man birun-e u-ra kardam

b. *man birun-e u kardam

c. man birunaš kardam

Notice the ungrammaticality of (52.a and b) but the accept-
ability of (52.c). It would appear that two process are in-
volved here. Apparently, we need a rule of Object Postposing
in addition to a rule of Object Preposing. The rule of Enclitic
Formation, which was originally specified to be an optional
transformation, is obligatory with postposed pronoun objects.
We could tentatively state this rule as:

(53) Object Postposing

$$X - [\ X - NP - [\ X - v]\ ] - X$$
$$\quad VP \qquad\qquad V \qquad V\ VP$$

1   2   3   4   5   6

1   2   ∅   4+e 3   5   6

That (52.c) is the result of Object Postposing and Enclit-
ic Formation may be seen by the sentences below that follow the
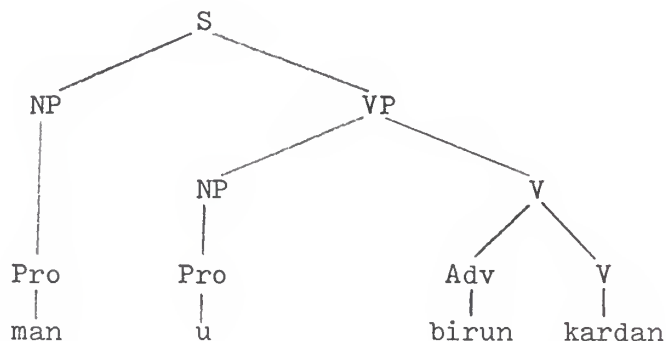paradigm presented in (47).

(54) a. man išān-rā birun kardam
       ·'I expelled them'

b. man birunešān kardam   (by Object Postposing
      'I expelled them'   and Enclitic Formation)

(55)   a.   man šomā-rā birun kardam

'I expelled you'
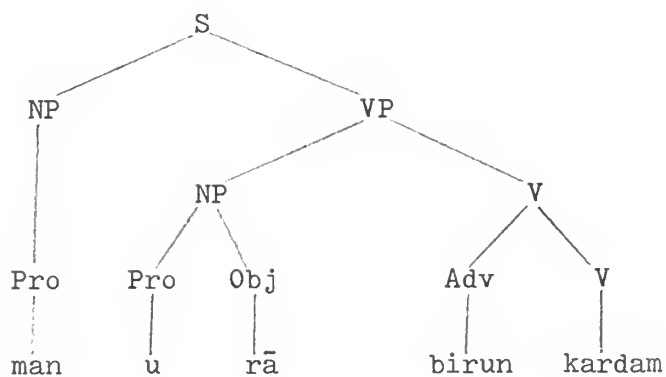
b.   man birunetān kardam

'I expelled you'

As for ordering the rules and accounting for the absence
of the definite direct object marker (-rā) in the forms de-
rived by Enclitic Formation, we have two choices. The more
cumbersome of the two would be to include in Enclitic Formation
a deletion of -rā, making the rule necessarily more complex.
The other choice, and as far as I can tell the better one,
would be to order Object Postposing before Object Marking.
Ordering the rules in this fashion would result in the post-
posed object's remaining unmarked and not requiring special
handling in the rule of Enclitic Formation. At this point it
may be advisable to compare the derivational histories of several
sentences and determine if the above analysis will operate in
the manner that we anticipate. For our purposes, let's com-
pare sentences (49) and (52.c).

(56)   derivational history of man u-rā birun kardam
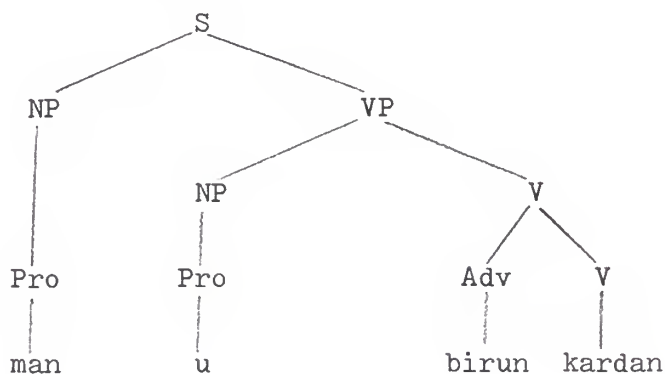
(i) by phrase structure rules:

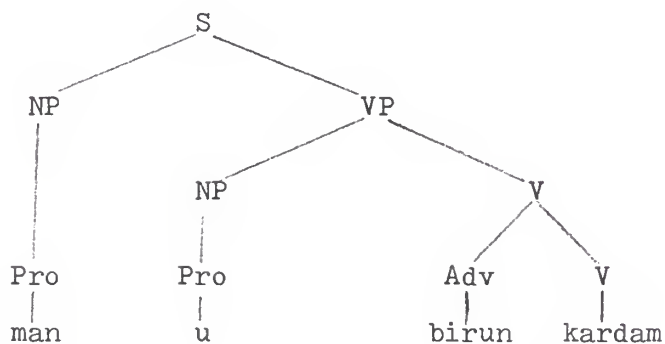(ii) by application of Subject-Verb Agreement and
Object Marking rules:

```
                          S
              _____/ _____
            NP                          VP
            |                _____/    _____
            |              NP                         V
            |           __/ | \__                  _/  \_
           Pro        Pro      Obj              Adv       V
            ┊          |        |                ┊        |
           man         u        rā             birun    kardam
```
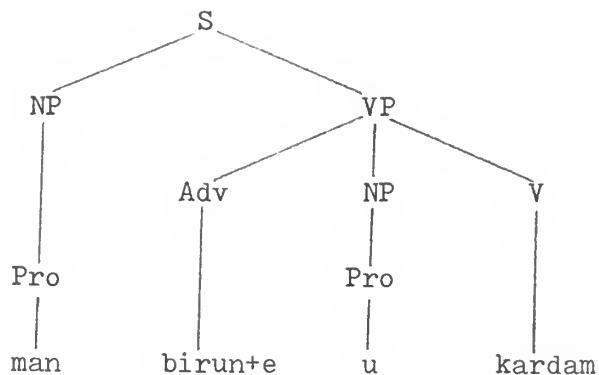
(57)   derivational history of <u>man birunas̆ kardam</u>

(i) by phrase structure rules:

```
                        S
            _____/   _____
          NP                          VP
          |               _____/     _____
          |             NP                          V
          |             ┊                         _/  \_
         Pro           Pro                     Adv       V
          ┊             ┊                       ┊        ┊
         man            u                     birun    kardan
```

(ii) by application of Subject-Verb Agreement:

```
                        S
            _____/   _____
          NP                          VP
          |               _____/     _____
          |             NP                          V
          |             ┊                         _/  \_
         Pro           Pro                     Adv       V
          |             ┊                       |        |
         man            u                     birun    kardam
```

(iii) by application of Object Postposing:

```
                        S
              _____/  _____
            NP                    VP
            |               _____/ |  \_____
           Pro            Adv      NP         V
            |              |       |          |
           man          birun+e  Pro       kardam
                                  |
                                  u
```

(iv) by application of Enclitic Formation:

```
                        S
              _____/  _____
            NP                    VP
            |              _____/  _____
           Pro           Adv                V
            |          __/ |  \__            |
           man       Adv      Pro         kardam
                      |        |
           man      birun      ašˇ
```
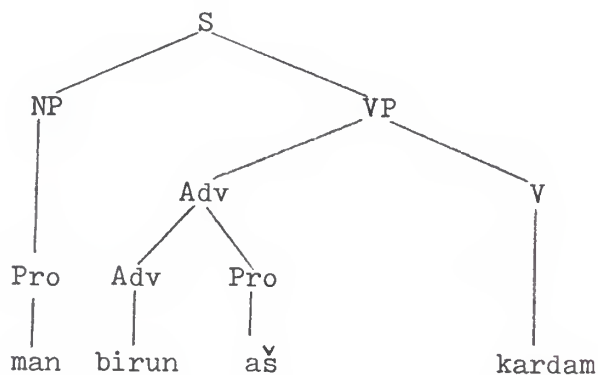
That this solution, the ordering of Object Postposing before Object Marking, seems to be the more appropriate approach can be attested to by the following ungrammatical forms:

(58) a. *<u>man birun-e išān-rā kardam</u>

    b. *<u>man birunešān-rā kardam</u>

    c. *<u>man birun-e išān kardam</u>

We can now account for the ungrammaticality of all three of these cases. In both (a) and (b) the ordering of Object

Postposing and Object Marking has been violated. In the third example the obligatory rule of Enclitic Formation has not applied.

The order of the rules so far would be:

(59) a. Subject-Verb Agreement   (obligatory)[13]

     b. Object Postposing   (optional)

     c. Object Marking   (obligatory)

     d. Enclitic Formation   (obligatory with post-
        posed pronoun objects; otherwise it is optional)

It was stated that the rule for Object Postposing, (53), was tentative. The reason for this 'tentativeness' may be observed in the following example:

(60) a. <u>man u-ra didam</u>

     I he-Obj saw-1sg

     'I saw him'

    b. <u>man didamaš</u>

     'I saw him'

Presumably, we would derive (60.b) in the same manner as we have like sentences involving compound verbs.
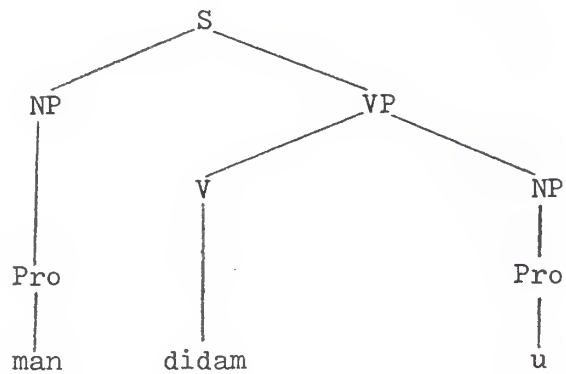
(61) derivational history of <u>man didamaš</u>
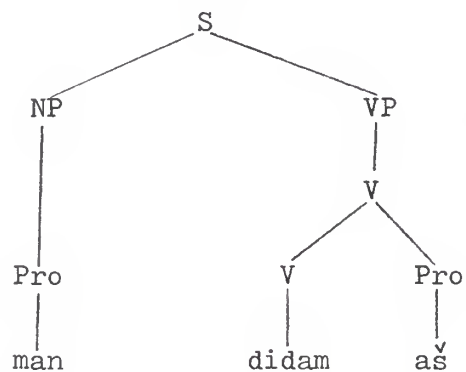
    (i) by phrase structure rules:
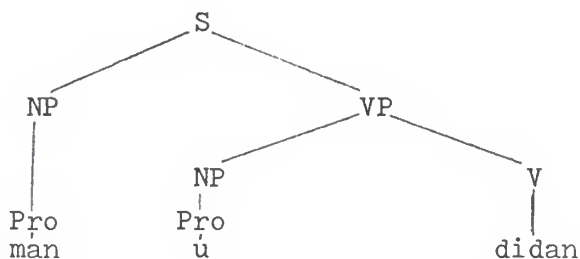
(ii) by Subject-Verb Agreement:
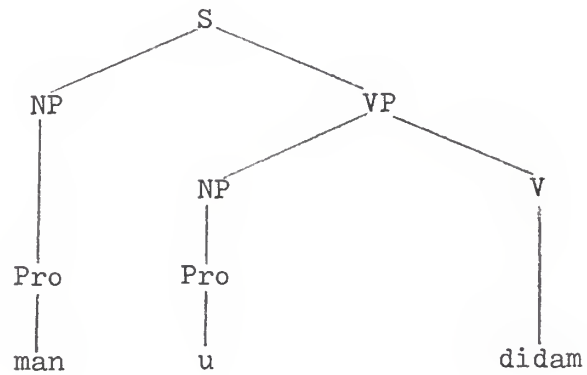
(iii) by Postposing Rule:

(iv) by Enclitic Formation:

From the data presented above, it would appear that we will
have to specify in Object Postposing whether or not the V is
compound or simple. If it is simple, the object follows the
verbal element (it goes without saying that a simple verb does
not have a non-verbal component). If the verb is a compound
construction, the object follows the non-verbal element. In
both instances Enclitic Formation will apply. The form of the
rule needed to account for both instances of Object Postposing
is as follows:

(62)   Object Postposing

$$X - \begin{bmatrix} X - NP - \begin{Bmatrix} V - X \\ [X - V] \\ V \quad\quad V \end{Bmatrix}] \\ VP \quad\quad\quad\quad\quad\quad\quad\quad\quad VP \end{bmatrix} - X$$

1	2	3	4	5		6
1	2	$\emptyset$	4+e	3	5	6

cond: $4 \neq \emptyset$

The condition that 4 not be null reflects the fact that there
are two verb types in Persian, the simple verb and the compound
verb. Where 4 is V we find the occurrence of a simple verb.
Where 4 is X we find a compound verb. This rule accounts for
the placement of the object pronoun after the verbal element
in the case of the simple verb and following the non-verbal
element in the case of the compound verb.

    In all of the examples so far presented the NP that has
been postposed has been a pronominal form. That this is not
exclusively the case for postposing to apply may be seen in the

following examples:

 (63) a. <u>parviz asb -rā savār šod</u>

       Parviz horse-Obj mounted

       'Parviz mounted the horse'

    b. <u>parviz savār-e asb šod</u>

       'Parviz mounted the horse'

 (64) a. <u>išān be tehrān vāred šodand</u>

       they to Tehran arrived

       'They arrived in Tehran'

    b. <u>išān vāred-e tehrān šodand</u>

       'They arrived in Tehran'

The first example, I believe, corroborates the claim that Object Marking should follow Object Postposing. Both examples, because they are grammatical without encliticization, substantiate the analysis that inserts an <u>ezāfeh</u> between the non-verbal component of the compound and the postposed object.

 I believe these last examples demonstrate most conclusively the incompatibility of Mace's claim with the facts of Persian. Namely, compound verb constructions in Persian are seperable. Using this information, we may now return to a previous question, whether or not the compound verb is an idiom.

 Weinreich (1969) discussed in great detail the form and components of idioms. For both Weinreich and Chafe (1970) our previous designation as to what composes an idiom is consistent (p. 16). Additionally, for many idioms there are literal readings

of the elements, accounting for their ambiguity. In English, idioms such as

(65) a. kick the bucket

 b. hit the roof

are ambiguous. (65.a) may either mean 'strike the pail with one's foot' or 'die'. (65.b) may either mean 'strike the house-top (with something)' or 'get angry'. In any event, there are literal readings for these two phrases as well as 'special' or idiomatic readings. It is also the case that literal readings must sometimes be somewhat forced to make sense. Take for example the following:

(66) The boys <u>shot the breeze</u>.

In some fairytale sense, it may be possible to have a literal reading of (66) by substituting for 'the breeze' the character 'North wind'. However, I would maintain that there really isn't a literal reading of (66), but rather a strictly idiomatic one of 'chatted idly'.

In the above examples we find that they may neither be interrupted nor the order of constituents changed, if their idiommatic sense is to be maintained. Consider the application of Passive Movement to the above examples:

(67) a. The bucket was kicked (by someone)

 b. The roof was hit (by someone)

 c. The breeze was shot (by the boys)

The first two examples now have only their literal readings and the last its 'special', fairytale reading. Likewise, if some element interrupts the idiom the result is the literal reading only. For example:

(68) a. John kicked the wooden bucket.
b. John hit the tile roof.
c. The boys shot the fickle breeze.

Weinreich (1969) notes, however, that the idiomatic reading is the only one possible if an element such as 'proverbial' is inserted where we have inserted 'wooden', 'tile', and 'fickle'.
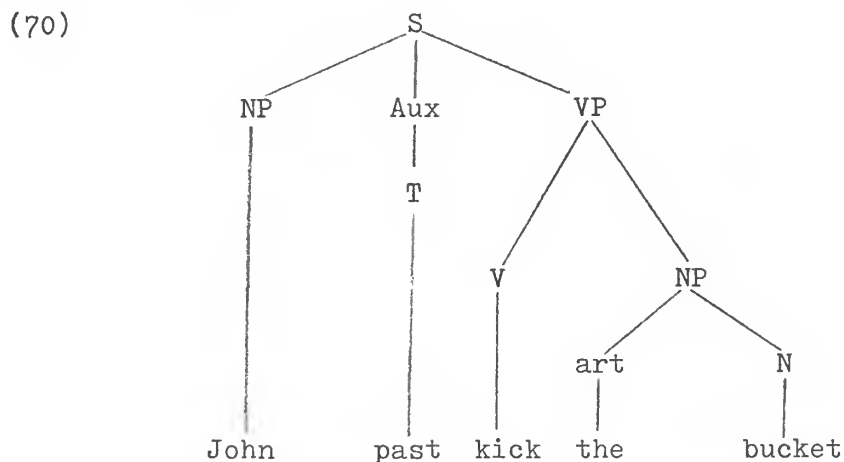
(69) a. John kicked the proverbial bucket.
b. John hit the proverbial roof.
c. The boys shot the proverbial breeze.

In all of these cases it is evident that the only reading possible is the idiomatic one. Weinreich points out that with the insertion of 'proverbial' the idiom is being emphasized or marked.[14] There are other instances in English, compound verbs this time, in which there is no literal reading and in which the insertion of 'proverbial' will not indicate their idiomatic nature.

'Hog-tie' and 'rabbit punch' seem to be closer to the type of compound verb constructions that we have been dealing with in Persian. In both cases we find a non-verbal element followed by a verbal component. Unlike their Persian counterparts, the English compounds may not be separated. However, like all idioms

and the Persian compound verb constructions, it is the verbal
element that is inflected.

An objection could be raised that Persian compounds are
interrupted in the same fashion as English idioms. That is,
some element from outside the idiom has been inserted into the
idiom. In the case of English idioms we find that tense origin-
ates in the Aux and is affixed to the verbal element of the
idiom.

(70)

```
                          S
          ┌───────────────┼───────────────┐
          NP             Aux              VP
          │               │            ┌───┴───┐
          │               T            V      NP
          │               │            │    ┌──┴──┐
          │               │            │   art    N
          │               │            │    │     │
        John            past         kick  the  bucket
```

By Affix Hopping we would derive:

(71)   John kick + past the bucket

       John kicked the bucket.

In Persian we find the same phenomenon of tense affixation to
the verbal element of the compound. What is substantially dif-
ferent in the Persian examples is that we have seen in (63 and
64) that full lexical items may separate the elements of the
compound. Comparing these examples, as well as the encliticized

forms of (52.c, 54.b and 55.b), with the non-idiomatic English sentences of (68) it is obvious that, if we are to consider Persian compounds as idioms, they do not meet the same restrictions placed on English idioms of the type that we've described in (65) or the English compound forms mentioned previously, 'hog-tie' and 'rabbit punch'.

It has been suggested by Weinreich that what is needed for the idioms in example (65) is some kind of 'idiom dictionary' in the grammar that will assign a 'special' reading for those sequences generated by the rest of the grammar that are found to be of the form that is entered in this 'dictionary'. Thus if a sentence such as.

(72)   John kicks a bucket.

is generated, the 'dictionary' would not mark it as containing an idiom (the definite article must be present for the idiomatic reading).

For Persian compounds it seems reasonable to propose that there is a phrase structure rule of the form:

(73)
$$V \longrightarrow (\begin{Bmatrix} Adj \\ Adv \\ N \end{Bmatrix}) \quad Verb$$

This would account for the productiveness of the compound forms as well as their high incidence of occurrence. I would then maintain that an 'idiom dictionary' of the form that Weinreich suggests would be a part of the grammar of Persian to assign

syntactic as well as semantic information to these compound forms. One of the features that would necessarily be marked would be that of transitivity. A compound that was marked [- transitive] would not logically be separated as there would be no object in the first place. Properly speaking, we might want to refer to this 'dictionary' as a 'compound dictionary'.

Because of the limited number of verbal elements that may be combined to form Persian compound verbs and because of the large number of these constructions in the language, it is quite doubtful that these compound verbs are idioms in the same sense that we know of them in English. Additionally, these compound forms in Persian may be interrupted by elements that are not, properly speaking, part of the verb.

Conclusion

We have touched on a number of areas that involve the compound verb. We've seen that it is an indigenous form in Persian and that it continues to be a productive form in the language. The claim that the components of the compound verb are inseperable has been demonstrated to be false as well as the claim that Persian compound verbs should be considered to be idioms.

Recent work by Sharifi (1973) (1976) and Sheik (1976) has attempted to throw additional light on the semantic properties of the compound verb. I've not included their work in this discussion because my knowledge of the language is limited, although it was their work and that of Sheintuch's that roused my interest in the problem.

Notes

1)   I use the term 'Persian' here in preference to the native
term <u>farsi</u> to name the language, simply for the convenience of
the reader unfamiliar with the latter term. It should be pointed
out that the term 'Iranian' would not be technically correct,
as this designates that group of Indo-European languages of
which Persian is a member.

2)   I have seen references to unpublished PhD dissertations
that, at least in the case of Moyne (1970), treat Persian in
a transformational framework. At present I've not had the op-
portunity to look at these works. Rather than arguing for or
against one particular analytic approach, I am here simply sug-
gesting the state of current linguistic inquiry into Persian
grammar.

3)   Though it is peripheral to our discussion, it is important
to understand that there have been vast population movements
through the area in which Persian is spoken today. As we shall
see, Persian was affected by each language that it came into
contact with and borrowed heavily from all. But it showed a
great deal of resiliency.

4)   Pahlavi, properly speaking, refers to the script in use at
this time but scholars have extended the designation to include
the language form as well. See, for example, Browne (1956: 7).

5)   I don't pretend to be a fluent speaker of the Persian

language. I've relied upon the judgements of Iranian acquaint-
ances here at the university for my analysis and for judgements
on other writers' interpretations. Likewise, following the ex-
ample set by Moyne and Carden (1974), I'm making no phonological
claims in my transcription system. It should be noted that with-
in the borders of Iran there are numerous regional dialects
spoken. The dialect presented here might be said to be a 'Stan-
dard Persian' form that one might hear spoken on the radio, a
more deliberate form than one might hear on the street. A great
deal of interesting work could be done in the defining of dialects
in Iran.

6) The word _ezāfeh_, an Arabic loan, means 'something added'
or 'annexed'. The term refers to the use of an unstressed /e/
syllable suffixed to a noun or noun phrase which is followed
by a modifying or qualifying word or phrase. The following
examples illustrate some of the different kinds of phrases in
which the _ezāfeh_ construction occurs:

    a.   _ketāb-e man_
          book   I/me
          'my book'

    b.   _medād-e majid_
          pencil Majid
          'Majid's pencil

    c.   _dȧftar-e barādar-e majid_
          notebook brother Majid
          'Majid's brother's notebook'

    d.   kāgaz va xodkār-e majid

        paper and ballpoint Majid

        'Majid's paper and ballpoint'

    e.   divār-e sāxtemun

        wall   building

        'wall of the building'

    f.   šahr-e tehrān

        city   Tehran

        'the city of Tehran'

    g.   āb-e xordan

        water drink

        'drinking water'

    h.   rāh-e dur

        road   far

        'a long way'

Generally, the ezāfeh construction occurs in situations where the preposition 'of' occurs or where an adjective precedes a noun in English.

7)   This is of course the regular form of the simple verb. Rastrogueva does make a slight error here. It is possible to have mixandam 'I was reading' $_V$/ mi - Vstem past - Pers/Num /$_V$

8)   See for example, Rubinchik (1971) or Wexler (1974).

9)   In her conclusion Sheintuch states, 'it seems that peri-
phrastic verb formation in Persian got great impetus with the
increasing number of Arabic verbs that had to be nativized.
Because of the difference between the structures of the host
and target languages, the direct borrowing of verbs from Arabic
into the Persian declension patterns would have greatly violated
the Arabic structure. PV formation appears to have been the
ideal compromise between the two structures, for it kept the
Arabic component indeclineable' (1976: 153). As we have seen
however, there would have been no 'reason' for compounds as the
verbal suffix would have been sufficient in the then current
Persian verb making paradigm.

10) These are for Sheintuch the most common and Lambton also
includes these in her list. The following is a more complete
list of what Sheintuch refers to as the 'productive periphrastic
Aux's in Persian':

saxtan	'create'	āftādan	'fall'
nemudan	'show'	yāftan	'find'
farmudan	'order'	gastan	'search'
dāštan	'have'	xāndan	'read'
didan	'see'	kandan	'pick'
gereftan	'get'	goftan	'say'
gozāštan	'put'	resāndan	'make reach'
bāxtan	'lose'	gardāndan	'make turn'
bordan	'take'	gardidan	'turn'
bastan	'tie'	duxtan	'sew'
āvordan	'bring'	kešidan	'pull'
raftan	'go'	xāstan	'want'
āmadan	'come'		

11) Of course, I'm assuming that what Mace implies here by 'never changes' is 'never inflected'.

12) This may be a questionable English sentence, but it will have to do to get the sense of the Persian sentence across.

13) At this point there would seem to be little argument for the ordering of Subject-Verb agreement before Object Postposing or vice versa. In fact, at this point we might have said that they were unordered with respect to each other. However, consider (60) and its derivational history (61). If Object Postposing occurred before Subject-Verb Agreement, this may have necessitated some rule that involved infixing of the person/number marker. As it is, with the ordering presented in the text we need only worry about suffixing to the entire V.

(14) I must thank Prof. Armagost for pointing out to me that other lexical items may indeed be found in these idiomatic expressions. For example, we might consider:

John kicked the goddamn bucket.

Other, more pithy, oaths of Anglo-Saxon origin may occur in place of 'proverbial' or 'goddamn'. Unlike the sentences with 'proverbial', however, the above sentence for me is as ambiguous as those without it. What this does go to show is that idiomatic phrases may be separated by other elements besides this emphatic 'proverbial' and still maintain their idiomatic sense as well as their literal reading.

Bibliography

Anonymous
(1970): Persian-English Grammar. The Scolar Press Limited.
Menston, England. (Originally published 1790)

Browne, E.G.
(1956): Literary History of Persia: From the Earliest Times
Until Firdawsi. Vol. 1. Cambridge University Press,
Cambridge.

Chafe, Wallace L.
(1967): Language as Symbolization. Language 34:1:57-91.

(1970): Meaning and the Structure of Language. The
University of Chicago Press, Chicago.

Elwell-Sutton, L.P.
(1963): Elementary Persian Grammar. Cambridge University
Press, Cambridge.

Hadley, George
(1972): Introductory Grammatical Remarks on the Persian
Language. The Scolar Press Limited, Menston, England.
(Originally published 1776)

Hawker, C.L.
(1937): Simple Colloquial Persian. Longmans, Green and Co
Ltd., London.

Jones, William
(1969): A Grammar of the Persian Language. The Scolar
Press Limited, Menston, England. (Originally published 1771)

Lambton, Ann K.S.
(1966): Persian Vocabulary. Cambridge University Press,
Cambridge.

(1967): Persian Grammar. Cambridge University Press,
Cambridge.

Mace, John
    (1967): Modern Persian.   The English Universities Press
    Limited, London.

Manuchehri, Parivash
    (1974): Explaining Problems of Iranian Students by Compar-
    ing English and Farsi Verb Forms.   TESOL Quarterly 8:2:
    171-176.

Marashi, Mehdi
    (1972): Modals and Auxiliaries in Persian.   Orbis 21: 417-
    428.

Masica, Colin P.
    (1976): Defining a Linguistic Area: South Asia. University
    of Chicago Press, Chicago.

Moyne, John A.
    (1970): The Structure of Verbal Constructions in Persian.
    unpublished Doctoral dissertation, Harvard University,
    Cambridge, Mass.

    (1971): Reflexive and Emphatic.   Language 47: 141-163.

    (1974): The So-Called Passive in Persian.   Foundations of
    Language 12: 249-267.

    and Guy Carden
    (1974): Subject Reduplication in Persian.   Linguistic
    Inquiry 5:2: 205-249.

Rastorgueva, V.S.
    (1964): A Short Sketch of the Grammar of Persian.   IJAL
    30:1: pt. 2.

Rubinchik, Yu.A.
    (1971): The Modern Persian Language.   Nauka Publishing
    House, Central Department of Oriental Literature, Moskow.

Sharifi, Hassan

(1973): The Feature Compatibilities of Persian <u>xordan</u> (to eat) with Nouns, Adjectives, etc.: A Metaphorical Process. in <u>Meaning: A Common Groung of Linguistics and Literature</u>. Prodedings of a University of Northern Iowa Conference held April 27-28, 1973 in honor of Norman Stageberg.

(1976): Persian Verbs: A Chafean Analysis. in Ingemann, Francis (ed) (1976): <u>1975 Mid-America Linguistic Conference Papers</u>. Linguistic Department, University of Kansas, Lawrence.

Sheik, Habib and Hassan Sharifi

(1976): The Semantic Structure of the Persian Verb. in Ingemann, Francis (ed) (1976): <u>1975 Mid-America Linguistic Conference Papers</u>. Linguistic Department, University of Kansas, Lawrence.

Sheintuch, Gloria

(1976): Periphrastic Verb Formation in Persian. Studies in the Linguistic Sciences 5:2: 139-156. University of Illinois Working Paper, Urbana.

Weinreich, Uriel

(1969): Problems in the Analysis of Idioms. in Puhvel, Jaan (ed) (1969): <u>Substance and Structure of Language</u>. University of California Press, Berkely.

(1972): <u>Explorations in Semantic Theory</u>. Mouton, The Hague.

Wexler, Paul

(1974): The Cartography of Unspoken Languages of Culture and Liturgy: Reflections on the Diffusion of Arabic and Hebrew. Orbis 23:1: 30-51.

ON THE PERSIAN COMPOUND VERB

by

BRIAN LEANDER O'NEILL

B. S. Kansas State University, 1972

———————————————

AN ABSTRACT OF A MASTER'S THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF ARTS

Department of Speech

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1978

ABSTRACT

This work attempts to shed light on a poorly defined area
of Persian grammar. In the process, several questions are
raised concerning the nature of the compound verb. Particular
attention has been given to three characteristics that have
been alluded to in the literature--(a) the productiveness of
the compound process, (b) the idiomatization of the compound
and (c) the inseparability of the components of the compound
verb.

Introductory sections giving historical and grammatical
background are included.