

Lab 5: Ditransitives and Optional Subjects

Bill McNeill

May 15, 2004

1 The Facts of Farsi

In Farsi, pronominal subjects are optional. Sentences (1) and (2) mean the same thing. Sentence (2) is the more usual way of speaking, while sentence (1) would be used for emphasis.

(1) *man ketabra xaridam*
I the book-ACC bought-1SG-PAST
'I bought the book.'

(2) *ketabra xaridam*
the book-ACC bought-1SG-PAST
'I bought the book.'

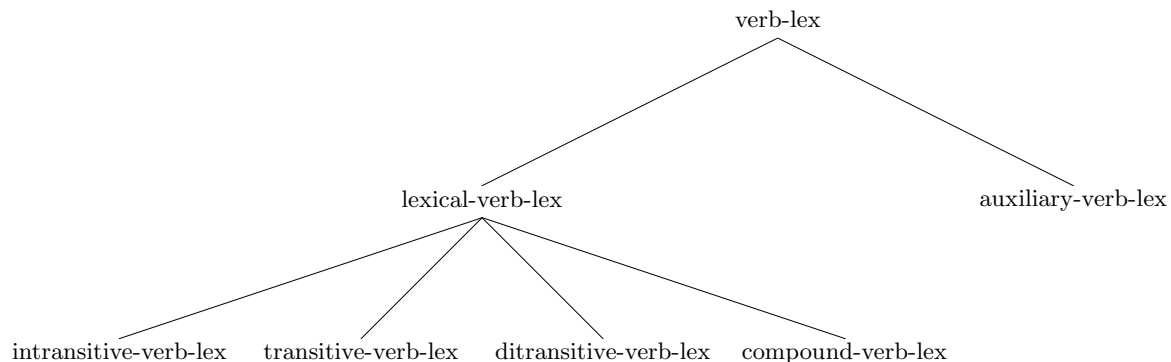
I am not sure how individual Farsi verbs subcategorize for optional arguments because none of my sources come out explicitly and say so and my Farsi reading skills are not to the point where I could skim for examples. For this lab I am treating all transitive verbs as having required objects.

All the ditransitive Farsi verbs I've seen in my sources take a noun as a direct object and a prepositional phrase as an indirect object. These constituents appear in the order subject-direct object-indirect object-verb. I have implemented one ditransitive verb, *dadan* 'give'. A typical usage would be

(3) *shirin ketabra be man dand*
Shirin the book-ACC to me gave-3SG-PAST
'Shirin gave me the book.'

2 Implementation

I've implemented the following verb hierarchy.



Intransitive verbs take no objects. Transitive verbs take a single noun-phrase object. Ditransitive verbs take a noun phrase direct object and a prepositional phrase indirect object as described above. Compound verbs are verbs like *kardan* ‘make’ or *budan* ‘be’ that combine with non-verbal elements to form compound verbs. An example of this kind of construction is

- (4) *dard nemikonad*
 harm doesn’t NOT-PROG-make-3SG-PRES
 ‘It doesn’t hurt me.’

Compound nouns in Farsi may take objects, but currently I’m treating them all as intransitive. To handle the full range of compound nouns I would probably have to change this hierarchy to allow multiple inheritance from compound vs. simple base types as well as the intransitive, transitive, and ditransitive types.

Auxiliary verbs take other verb phrases as objects. An example of an auxiliary construction is

- (5) *mitavanam shishera bexoram*
 PROG-can-1SG-PRES glass-ACC SUBJUNC-eat-1SG-PRES
 ‘I can eat glass.’

Here the auxiliary verb *tavanestan* ‘be able’ takes the verb phrase *shishera bexoram* ‘eat glass’ as its object.

In Farsi, the past and present tenses are formed by inflecting past and present stem forms while the future tense is formed with the auxiliary *xastan* ‘want’. The past stem form can be derived regularly from the infinitive form of the verb, but the present is fairly irregular. Since the two stem forms must undergo the same morphological processes but cannot be derived regularly from one another, I have two lexical entries for each of my verbs, a past tense form and a present tense form. For example the lexical entry for *mordan* ‘die’ is as follows.

```
mord := past-intransitive-verb-lex &
[ STEM <"mord">,
  SYNSEM [ LOCAL.CAT.HEAD.TAM.TENSE past_tense,
           LKEYS.KEYREL.PRED '_die_v_rel' ] ].
```

```
mir := present-intransitive-verb-lex &
[ STEM <"mir">,
  SYNSEM [ LOCAL.CAT.HEAD.TAM.TENSE present_tense,
           LKEYS.KEYREL.PRED '_die_v_rel' ] ].
```

For several verbs I don’t know the irregular present stem, so I have only their past tense in the lexicon. I have not implemented the future tense.

I implemented optional subjects by instantiating a basic-head-opt-subj-phrase from the matrix.

I implemented ditransitive verbs requiring their COMPS list to contain a prepositional phrase and an accusative noun phrase. Prepositions are implemented as preposition-lex items which subcategorize for one argument, a non-accusative noun phrase. The prepositions copy their argument’s LOCAL.CONT.HOOK information copy into their own CONT, i.e. the prepositions have no semantic content themselves.

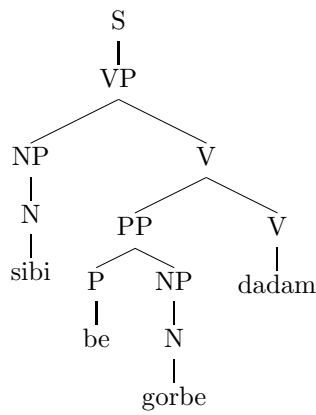
3 Ditransitive Syntax and Semantics

Here is a typical ditransitive sentence.

- (6) *sibi be gorbe dadam*
 an apple to the cat gave-1SG-PAST
 ‘I gave the cat an apple.’

Syntactically (6) parses like so

(7)



It has the following indexed MRS.

```
<h1,e2:TENSE:ASPECT:MOOD:SEMSORT,
{h3:_apple_n_rel(x4:SEMSORT:THIRD:NUMBER:BOOL),
h5:indef_q_rel(x4, h7, h6),
h8:_cat_n_rel(x9:SEMSORT:THIRD:NUMBER:BOOL),
h10:def_q_rel(x9, h12, h11),
h1:_give_v_rel(e2, x13+:SEMSORT:FIRST:SG, x9, x4)},
{h6 qeq h3,
h11 qeq h8}>
```

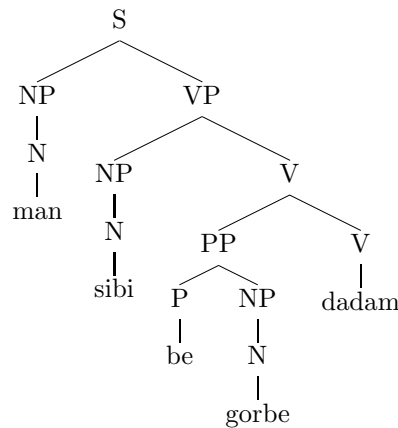
The ‘give’ relationship has three arguments: the subject x13 (not connected to anything else in this MRS because the subject was omitted), the indirect object x9, and the direct object x4. The prepositional phrase just copies up the semantics of its argument, so x9 points directly to *gorbe* ‘the cat’. The indefinite relationship has scope over ‘apple’ (h6 qeq h3) as expected.

Now make the subject overt

- (8) *man sibi be gorbe dadam*
 I an apple to the cat gave-1SG-PAST
 ‘I gave the cat an apple.’

The subject pronoun appears in the syntax.

(9)



The semantics is very similar, except now the subject of the ‘give’ relationship x4 is indexed as a pronoun in this MRS.

Sentence	Description	Result
man ketabra be shirin dadam	'I gave Shirin the book.'	Pass
ketabra be shirin dadam	'I gave Shirin the book.'	Pass
*ketabra dadam be shirin	Incorrect argument order	Pass
*be shirin ketabra dadam	Incorrect argument order	Pass
*be shirin dadam ketabra	Incorrect argument order	Pass
*dadam ketabra be shirin	Incorrect argument order	Pass
*dadam be shirin ketabra	Incorrect argument order	Pass
*be shirin dadam	Missing direct object	Pass
*ketabra dadam	Missing indirect object	Pass
*ketab be shirin dadam	Direct object not accusative	Pass
*ketabra gorbera dadam	Indirect object not a prepositional phrase	Pass
*be shirin be gorbe dadam	Direct object a prepositional phrase	Pass
*ketabra be dadam	Missing prepositional phrase object	Pass
*man ketabra bar shirin dadam	Incorrect preposition	Fail

Table 1: Regression tests for ditransitive verb

```
<h1,u2:SEMSORT,
{h3:pronoun_n_rel(x4:SEMSORT:FIRST:SG:BOOL),
h5:pronoun_q_rel(x4, h7, h6),
h8:_apple_n_rel(x9:SEMSORT:THIRD:NUMBER:BOOL),
h10:indef_q_rel(x9, h12, h11),
h13:_cat_n_rel(x14:SEMSORT:THIRD:NUMBER:BOOL),
h15:def_q_rel(x14, h17, h16),
h18:_give_v_rel(e19:SEMSORT:TENSE:ASPECT:MOOD, x4, x14, x9)},
{h6 qeq h3,
h11 qeq h8,
h16 qeq h13}>
```

4 Testing

To verify basic subject-verb agreement I tested the entire paradigm for the past and present tenses of the transitive verb *xaridan* 'buy' and the intransitive verb *mordan* 'die'. I verified that simple declarative sentences parsed with an omitted subject or with the correct subject and did not parse for a subject that did not agree in person or number.

To verify the coverage of ditransitive verbs I ran the regression tests outlined in table (1). All but the last one of them passed. The sentence *man ketabra bar shirin dadam* should be ungrammatical because *dadam* subcategorizes for the preposition *be* 'to' not *bar* 'on'. However I haven't yet implemented a feature that enforces this subcategorization, so any preposition can go with any ditransitive.

In addition to this failure, there are four other failing tests in my regression suite that are unrelated to the work I did in this lab. The first three are all related to pronouns not getting the correct case marking.

- (10) *ishan ishanra xarid*
they them-ACC bought-3PL-PAST
'They bought them.'
- (11) *gorbe mara xarid*
the cat me-ACC bought-3SG-PAST
'The cat bought me.'

- (12) **gorbe ma xarid*
the cat me bought-3SG-PAST
'The cat bought me.'

In (10) and (11) the pronoun direct objects are given with the correct accusative case marker *-ra*, but the signs *ishanra* and *mara* are not recognized. Something is preventing the lexical rule that strips off the accusative case marker from firing for pronouns. Likewise in (12) the covert non-accusative lexical rule does not fire for the pronoun *ma*, so its CASE feature is left unspecified, which causes the LKB to incorrectly accept it as a valid object of *xarid*. I haven't figured out exactly what is preventing these lexical rules from working properly.

The last example is a case of overgeneration with compound verbs.

- (13) **an nemikonad*
it NOT-PROG-make
'does not make-it'

The compound auxiliary *kardan* forms compound verbs with only certain nouns. I don't have this restriction in the grammar yet, so it can currently form compound verbs with any noun, including the pronoun *an* as in (13).

References

- [1] Ann K.S. Lambton. *Persian Grammar*. Cambridge University Press, 1960.
- [2] John Mace. *Persian Grammar for Reference and Revision*. RoutledgeCurzon, 2003.
- [3] W.M. Thackston. *An Introduction to Persian*. IBEX Publishers, Bethesda, Maryland, 1993.