

AN ANALYSIS OF SIMPLE AND CONSTRUCT-STATE NOUN PHRASES IN MODERN STANDARD ARABIC

Issa AlQurashi
Taif University

1. Data

This paper aims to propose an HPSG analysis for simple and construct-state noun phrases in Modern Standard Arabic (MSA). To the best of my knowledge, there are no major HPSG analyses of MSA noun phrases (NPs).¹ A parallel phenomenon in Hebrew has been discussed quite extensively in the same framework by Wintner (2000).

Simple MSA noun phrases can be definite or indefinite. Definite nouns are prefixed with the definite article (*al-*) -glossed 'DEF' - (see, for example, Ouhalla, 1991; Fassi Fehri, 1993; Ryding, 2005; Benmamoun, 2006, among others), and indefinite nouns are suffixed with the indefinite marker (-n) -glossed 'INDEF' - (see, for example, Ryding, 2005, among others) - as in (1).

- | | | | | |
|-----|----|---|----|--|
| (1) | a. | ʔal-kitaab-u
DEF-book-NOM ²
'The book' | b. | kitaab-u-n
book-NOM-INDEF
'a book' |
|-----|----|---|----|--|

MSA also has construct state nouns consisting of a head noun directly followed by a possessor. The head/construct noun can carry neither the definite article (*al-*) as in (2a), nor the indefinite marker (-n) as in (2b) (Ouhalla, 1991; Fassi Fehri, 1993; Benmamoun, 2006; Ryding, 2005), but the form of a modifying adjective (which follows the possessor) shows that the nouns agrees with the possessor in definiteness.

- | | | | | |
|-----|----|--|--|-------------------------------------|
| (2) | a. | [(<i>*al-</i>)-kitaab-u
DEF-book.SG.MASC-NOM
'the female student's new book' | T-Taaliba-t-i]
DEF-student.SG-FEM-GEN | l-jadiid-u
DEF-new.SG.MASC-NOM |
| | b. | [kitaab-u-(<i>*n</i>)
book.SG.MASC-NOM-INDEF
'a (female) student's new book' | Taaliba-t-i-n]
student.SG-FEM-GEN-INDEF | jadiid-u-n
new.SG.MASC-NOM-INDEF |

Adjectives in MSA agree in definiteness plus other agreement features with the noun they modify.

The form of the adjective in (2a) shows that the noun is definite although it does not bear the definite article, and the form of the adjective in (2b) shows that the noun is indefinite although it does not have the indefinite suffix.

In addition to the attributive adjective and the possessor, the construct-state noun can have a PP or a clause as a complement. Consider the following example showing a PP complement:

- | | | | | | |
|-----|----|---|---|--|---|
| (3) | a. | kitaab-u
book.SG.MASC-NOM
l-qayyim-u
DEF-valuable.SG.MASC-NOM
'the erudite author's valuable book about syntax' | (<i>*l-qayyim-u</i>)
DEF-valuable
fii
in | l-muʔallifa-t-i
DEF-author.SG-FEM-GEN
n-naHw-i
DEF-syntax-GEN | D-Daliiʕa-t-i
DEF-erudite.SG-FEM-GEN
(<i>*l-qayyim-u</i>)
DEF-valuable.SG.MASC-NOM |
|-----|----|---|---|--|---|

Any such complement appears after the possessor and the adjectives. Notice also that an adjective cannot appear either before the possessor or after the complement. This means that the order has to be NP AP PP. If a relative clause is used in an example like the one in (3) above, it will occur after the ordinary complement as in (4) below:

- | | | | | | |
|-----|--|---|--|-----------|----------------------------|
| (4) | kitaab-u
book-NOM
[ʔallaḏii
that.SG.MASC
'Siibawaih's valuable book about syntax which you gave me as a present' | siibawayh-i
Siibawaih-GEN
ʔahdayta-nii
give present-me | l-qayyim-u
DEF-valuable-NOM
ʔiyyaah]
it | fii
in | n-naHw-i
DEF-syntax-GEN |
|-----|--|---|--|-----------|----------------------------|

The examples in (3) and (4) are the most important facts in this paper and hence they will be the central focus of the analysis.

¹ There is some discussion of HPSG and Arabic noun phrases in Borsley (1995) but he mainly discusses Syrian Arabic.

² The nominative case is the citation form in MSA.

As for the complement selection possibilities of definite and indefinite nouns, they both allow a complement (PP) following the attributive adjective (just like construct-state nouns) as shown in the following examples:

- (5) a. qaraʔ-tu kitaab-a-n jadiid-a-n fii n-naHw-i
 read.PAST-1SG book-ACC-INDEF new-ACC-INDEF in DEF-syntax-GEN
 ‘I read a new book about syntax’
- b. qaraʔ-tu l-kitaab-a l-jadiid-a [fii n-naHw-i]
 read.PAST-1SG the-book-ACC DEF-new-ACC in DEF-syntax-GEN
 ‘I read the new book about syntax’

These differences between definite and indefinite nouns on the one hand, and construct state nouns on the other hand will be captured by appropriate constraints in the following section.

2. Analysis

2.1. Basics

I will begin with the treatment of possessors, and the constraints on the three types of noun (def, indef, and construct). After that, I will discuss the status and position of attributive adjectives. Concerning the possessor, I treat it as an extra complement of the head noun rather than a realisation of the SPR feature.³ This position is taken by Borsley (1989, 1995) for Welsh and Arabic, and by Wintner (2000) for Hebrew. Borsley based his arguments on the fact that possessors always follow the associated noun and can be realised as clitics like the objects of verbs and prepositions. With verbs and prepositions clitics realise what is an uncontroversial complement. This suggests they also realise a complement with nouns and hence that possessors are complements. An example where a possessor is realised as a clitic is shown below:

- (6) kitaab-u-hu fii n-naHw
 book-NOM-her in DEF-syntax
 ‘his book about syntax’

Now, I discuss the constraints to which the subtypes of the type *noun* are subject. The subtype *indef-noun* is subject to the following constraint:

$$(7) \quad \text{indef-noun} \rightarrow \left[\begin{array}{l} \text{MORPH} \left[\begin{array}{l} \text{FORM } F_{\text{indef}} [1] \\ \text{I - FORM} [1] \end{array} \right] \\ \text{SYNSEM} \left[\begin{array}{l} \text{HEAD} [\text{DEF} -] \\ \text{ARG - ST } \neg \langle \text{NP} \dots \rangle \end{array} \right] \end{array} \right]$$

The constraint in (7) contains MORPH and SYNSEM features. The MORPH feature has two features: FORM and I-FORM, which are taken from Miller and Sag (1997). The I-FORM is the inflectional form of the noun without the indefinite marker. A noun will have various values for I-FORM depending on its case and whether it is singular or plural. The value of FORM is the noun suffixed with the indefinite marker. The function F_{indef} adds the indefinite marker to the inflectional form of the noun. As for the SYNSEM feature, it has the indefinite marker because it is indefinite. The $\neg \langle \text{NP} \dots \rangle$ stipulation ensures that a noun bearing the indefinite marker does not have an ARG-ST list whose first member is a possessor. This means that the indefinite noun can have an ARG-ST list which may contain other members such as PPs and clausal complements but not a possessor.

The subtype *def-noun* is subject to the following constraint:

$$(8) \quad \text{def-noun} \rightarrow \left[\begin{array}{l} \text{MORPH} \left[\begin{array}{l} \text{FORM } F_{\text{def}} [1] \\ \text{I - FORM} [1] \end{array} \right] \\ \text{SYNSEM} \left[\begin{array}{l} \text{HEAD} [\text{DEF} +] \\ \text{ARG - ST } \neg \langle \text{NP} \dots \rangle \end{array} \right] \end{array} \right]$$

³ Possessors in English are analysed as realisations of the SPR (SPECIFIER) feature.

The features FORM and I-FORM in (8) are not identified. The function F_{def} adds the definite article to a basic form of the noun which marks it as definite. Hence, the value of DEF feature is [+]. The $\neg \langle NP \dots \rangle$ stipulation ensures that a noun bearing the definite article does not have an ARG-ST list whose first member is a possessor. This means that the definite noun can have an ARG-ST list which may contain other members such as PPs, but not a possessor as shown in (5b) above.

The subtype *construct-state-noun* is subject to the following constraint:

$$(9) \quad \text{construct-state-noun} \rightarrow \left[\begin{array}{l} \text{MORPH} \left[\begin{array}{l} \text{FORM [1]} \\ \text{I-FORM [1]} \end{array} \right] \\ \text{SYNSEM} \left[\begin{array}{l} \text{HEAD [DEF [1]]} \\ \text{ARG-ST} \langle \text{NP} \dots \rangle \\ \left[\begin{array}{l} \text{DEF [1]} \\ \text{CASE } gen \end{array} \right] \end{array} \right. \end{array} \right]$$

The constraint in (9) says that the values of the FORM and I-FORM features are identified. This ensures that a construct-state-noun has neither a definite prefix nor an indefinite suffix. Furthermore, the constraint guarantees that the construct-state noun has an ARG-ST list whose first member is a possessor, which is genitive and has the same value for DEF as the head noun. It thus requires definiteness agreement between the head noun and the possessor.

2.2. Attributive adjectives as complements

Turning to attributive adjectives, one way to ensure that appear between possessors and ordinary complements is to assume that they are optional extra complements. Treating adjectives as extra complements is rather like the approach taken to verbal adjuncts (particularly postverbal adverbs) in Bouma, Malouf, and Sag (2001). They argue that in English, postverbal adjuncts are extra complements of the verb.

To ensure that attributive adjectives do not appear as adjuncts modifying N or NP in head-adjunct structures, we could impose a restriction on the type *head-adjunct phrase* excluding a nominal head, as in the following constraint:

$$(10) \quad \text{head-adjunct-ph} \rightarrow \left[\text{HEAD} \neg \left[\begin{array}{l} \text{noun} \\ \text{LEX +} \end{array} \right] \right]$$

This says that a *head-adjunct-ph* cannot be a noun that is [LEX +]. Thus, a nominal head is excluded. However, this will only prevent adjectives from modifying a noun and coming before the possessor. We also need to prevent adjectives from modifying NP and coming after a complement. Probably the best thing to do is to assume that adjectives are [MOD none] and hence they don't modify anything.

There is one important objection to this analysis. Treating attributive adjectives as extra complements makes them different from relative clauses (assuming the latter are adjuncts). However, they are like relative clauses in reflecting the definiteness of the modified noun. Relative clauses modifying a definite NP are introduced by a complementizer whereas relative clauses modifying an indefinite NP lack a complementizer as the following examples show:

- (11) a. raʔay-tu r-rajul-a *(llaðii) qaabal-tu-hu bi-l-ʔams
SEE.PAST.ISG DEF-man-ACC that.SG.MASC meet.PAST-1SG-him in-DEF-yesterday
 'I saw the man whom I met yesterday'
- b. raʔay-tu rajul-a-n *(llaðii) qaabal-tu-hu bi-l-ʔams
SEE.PAST.ISG man-ACC-INDEF that.SG.MASC meet.PAST-1SG-him in-DEF-yesterday
 'I saw a man whom I met yesterday'

In the following section, I will propose a different approach in which a possessor is treated differently.

2.3. Possessors as special complements

A second way to ensure the correct positioning of adjectives is to assume that they modify a noun but to treat possessors as special complements with which the noun combines to form a complex noun. This requires a special type, which might be called a construct-state-noun, subject to the following constraint:

$$(12) \quad c-s-n(oun) \rightarrow \left[\begin{array}{l} \text{HEAD [LEX +]} \\ \text{COMPS[3]} \\ \text{DTRS } \langle [1] \left[\begin{array}{l} \text{HEAD } noun \\ \text{COMPS } \langle [2] \rangle \oplus [3] \end{array} \right], [2] \text{NP[CASE } gen] \rangle \\ \text{HD - DTR [1]} \end{array} \right]$$

The constraint states that a construct state noun is [LEX +], and has a nominal head daughter and a genitive NP non-head daughter which is the first item on the COMPS list of the head and that the COMPS value of the phrase is identical to the remainder of the head's COMPS list. To prevent an adjective modifying the head noun intervening between the head noun and the possessor, we could stipulate that adjectives are [MOD N [COMPS \neg \langle NP, ... \rangle]] so that they can only modify nouns which do not require a nominal complement (possessor). NPs will be modified by relatives as I mentioned above. I also need to prevent possessors being analysed as ordinary complements; this can be done by stipulating that a nominal head of a head-complement-phrase is [COMPS \neg \langle NP, ... \rangle]. This ensures that the first member of the COMPS list is not a possessor. This analysis is quite complex since it not only needs the special treatment of possessors but also needs a stipulation on adjectives to prevent them combining with a noun before it combines with a possessor and a stipulation to prevent possessors being analysed as ordinary complements. So, I reject this analysis, and I will go on to suggest a third approach in 2.4 below.

2.4. Head-adjunct-complement analysis

Kasper (1994) has proposed that heads, adjuncts, and complements may be sisters. This permits a simple account of examples in which a head and a complement are separated by an adjunct.

- (13) a. He [went **last night** to the cinema].
 b. She [talked **incessantly** about syntax].
 c. Sandy [said **yesterday** that he would be here].

In this approach, I will propose that nouns appear in head-adjunct-complement structures, in which the head has both adjuncts and complements as sisters. These require something like the following constraint:

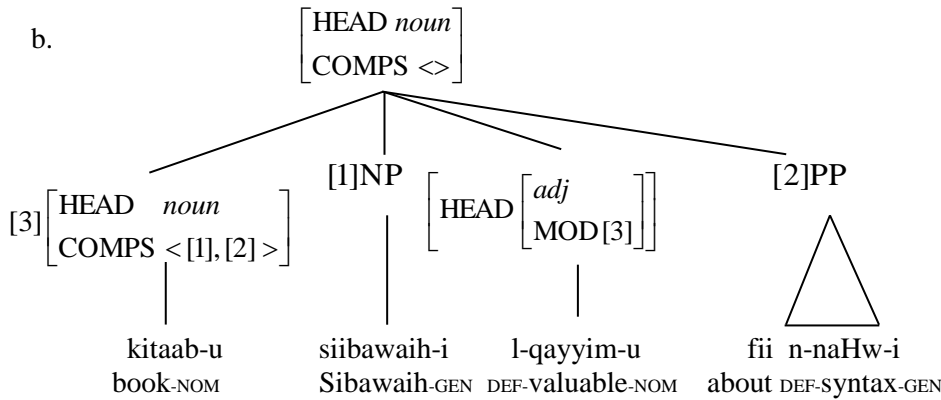
$$(14) \quad \text{head-adjunct-complement-phrase} \rightarrow \left[\begin{array}{l} \text{DTRS } \langle [1] \left[\begin{array}{l} \text{word} \\ \text{SS}[2][\text{COMPS } \langle [3], \dots [n] \rangle] \end{array} \right] \rangle \oplus \text{list}([\text{SS}[\text{MOD } [2]]) \oplus \langle [\text{SS}[3]], \dots [\text{SS}[n]] \rangle \\ \text{HD - DTR [1]} \end{array} \right]$$

This says that the head-adjunct-complement-phrase has a head daughter and two lists of non-head daughters. The first list is optional adjunct daughters whose MOD value is identical to the value of SYNSEM in the head daughter. The second list is complement daughters whose SYNSEM values are identical to the COMPS value of the head daughter. It should be noted that (14) is not only relevant to NPs. Probably it is relevant to VP's as well given examples like (13) for English above and (15) for MSA below.

- (15) a. takallam-tu biwuDuH-i-n ʕani l-muʕkilat-i
 talk.PAST.ISG clearly-GEN-INDEF about DEF-problem-GEN
 'I talked clearly about the problem'
 b. ʔahab-tu bi-l-ʔams ʔilaa l-maʕraD-i
 go.PAST.ISG in-DEF-yesterday to DEF-gallery-GEN
 'I went yesterday to the gallery'

The constraint in (14) will allow structures like the following in (16b) for the example in (16a):

- (16) a. kitaab-u siibawayh-i l-qayyim-u fii n-naHw-i
 book-NOM Siibawaih-GEN DEF-valuable- in DEF-syntax-GEN
 'Siibawaih's valuable book about syntax'



The order NP AP PP will be ensured by LP constraints since these elements are sisters.

Having allowed nouns to appear in head-adjunct-complement structures, we need to exclude them from head-adjunct structures in order to avoid structures where an adjective appears between the head noun and the possessor. The obvious approach to do this is with the following constraint:

$$(17) \quad head-adjunct-ph \rightarrow \neg \left[\begin{array}{l} \text{HEAD} \left[\begin{array}{l} \textit{noun} \\ \text{LEX} + \end{array} \right] \\ \text{COMPS} \langle \text{NP}, \dots \rangle \end{array} \right]$$

This says that a head-adjunct-phrase cannot be a noun that requires an NP complement (i.e. a possessor). It is [LEX +] because we need to allow the head to be an NP (a [LEX -] constituent); this is what we have with relative clauses as they appear after the ordinary complement as in (18).

- (18) kitaab-u siibawayh-i l-qayyim-u fii n-naHw-i
 book-NOM Siibawaih-GEN DEF-valuable-NOM in DEF-syntax-GEN
 [ʔallaḏii ʔahdayta-nii ʔiyyaah]
 that.SG.MASC give present-me it
 ‘Siibawaih’s valuable book about syntax which you gave me as a present’

The analysis in § 2.4. above seems simpler as it only needs one stipulation. The *head-adjunct-complement-phrase* is needed anyway for the examples in (15). We just need to stipulate that nouns cannot appear in head-adjunct structures. Therefore, I conclude that it is the best approach for Arabic.

REFERENCES

- Benmamoun, E. (2006). Construct state. In K. Versteegh, M. Eid, A. Elgibali, M. Woidich & A. Zaborski (Eds.), *Encyclopedia of Arabic language and linguistics* (Vol. I, pp. 477–482). Leiden: Brill.
- Borsley, R. D. (1989). An HPSG approach to Welsh. *Journal of Linguistics*, 25(2), 333–354.
- Borsley, R. D. (1995). On some similarities and differences between Welsh and Syrian Arabic. *Linguistics*, 33(1), 99–122.
- Bouma, G., Malouf, R., & Sag, I. A. (2001). Satisfying constraints on extraction and adjunction. *Natural Language & Linguistic Theory*, 19(1), 1–65.
- Fassi Fehri, A. (1993). *Issues in the structure of Arabic clauses and words*. Dordrecht: Kluwer Academic Publishers.
- Kasper, R. T. (1994). Adjuncts in mittelfeld. In J. Nerbonne, K. Netter & C. Pollard (Eds.) (39–69).
- Miller, P. H., & Sag, I. A. (1997). French clitic movement without clitics or movement. *Natural Language & Linguistic Theory*, 15(3), 573–639.
- Ouhalla, J. (1991). *Functional categories and parametric variation*. London: Routledge.
- Ryding, K. (2005). *A reference grammar of modern standard Arabic*. Cambridge: Cambridge University Press.