


# Arabic Numeral Noun Constructions: Some Preliminary Thoughts

Louisa Sadler(University of Essex,UK)<sup>2</sup>

March 4, 2010

---

<sup>2</sup>This work arises out of research on agreement and on numerals supported by the AHRC. I gratefully acknowledge the support from this body under research grants APN17606/AN10939 and AH/E006868/1. 

# Objectives

To investigate the behaviour of numeral noun constructions (of various types) in some vernacular dialect of Arabic, with a particular focus on agreement patterns within NP, interaction with placement of (other) nominal modifiers and nominal complements, interaction with coordination, and arguments for or against particular notions of headedness. To propose an LFG analysis of the results.

However the starting point for generating questions and collecting vernacular data is existing descriptions of numeral noun constructions in MSA. The discussion here is wholly about the system pertaining for numerals 3-10.

## Postnominal Numerals

Adjectives show NP internal concord in NG, DEF and CASE.

- (1) bayt-u-n                      ġamīl-u-n  
house-NOM-INDEF beautiful-NOM-INDEF  
a beautiful house (Kremers 2003: 167:3a)
- (2) al-bayt-u                    -l-ġamīl-u  
the-house-NOM the-beautiful-NOM  
the beautiful house (Kremers 2003: 167:3b)
- (3) al-riġal-u                    -l-ṭiwāl-u  
the-men-NOM the-tall.PL-NOM  
the tall men
- (4) al-nisā-u                    -l-ṭawīl-āt-u  
the-women-NOM the-tall-F.PL-NOM  
the tall women (ibid 58:39)

As is well-known, plural non-human Ns trigger FEM.SG forms of adjectives, other modifiers, verbs and for pronominal anaphora.

(5) al-kitāb-u            -l-'aḥmar-u  
the-book.M-NOM the-red.M-NOM (Kremers 2003: 58:38b)

(6) al-kutub-u            l-ḥamrā'-u  
the-books.MPL-NOM the-red.FSG-NOM  
(Kremers 2003: 58:38c)

Ordinals appear to behave in a fashion similar to adjectives, sharing their position (postnominal) and agreement behaviour. 1 and 2 appear postnominally like ordinals: they show standard adjectival agreement as found with adjectival modifiers to SG and DU.

- (7) risālatāni                      ʔiṭnatāni  
letter.F.DU.NOM.IND two.F.DU.NOM.IND  
(two) letters (ElSaadany, 2008, 20)

## 3-10: Definiteness

According to Ryding (2005), for 3-10, with an indefinite counted item the numeral precedes the counted noun. With definite counted nouns the numeral is required to follow the noun : it shows *CASE* concord and takes the definite article. (Note that it is clear from the examples provided that a bare noun in the *idāfa* construction constitutes a definite counted noun.)

Cardinals 3-10<sup>9</sup>, show polar behaviour w.r.t GEN: feminine plural nouns (both human and non-human) take masculine cardinal numbers, and masculine plural nouns (both human and non-human) take feminine cardinal numbers:

(8) al-buyūt-u                    -l-ḥamsa-t-u  
the-houses.M-NOM the-five-FEM-NOM  
the five houses (Kremers 2003 62:46b)

(9) al-jihaat-u                    l-ʔarbaʔ-u                    l-ʔaṣliyyat-u  
the-directions.F-NOM the-four.M-NOM the-cardinal-NOM  
the four cardinal directions (Ryding, 2005, 338)

---

<sup>9</sup>and also the units in 13-19

Adjectives continue to show agreement with the Noun in the presence of a numeral (within both definite and indefinite NPs) “when a definite counted noun is modified by an adjective, the adjective follows the numeral and agrees with the noun in gender, case and definiteness. For nonhuman nouns, the plural form of the adjective is feminine singular, for human nouns, the adjective is plural in form.” (Ryding, 2005, 338) Any additional adjectives show the normal pattern of adjectival agreement with the head noun (and follow the numeral). Note that if the head noun is definite by virtue of participation in the construct state, then the possessed element precedes the numeral.

- (10) bayna ʔalwān-i l-ṭayf-i l-sabʕat-i  
 among colours-GEN the-spectrum-GEN the-seven.F-GEN  
 l-maʕrūfat-i  
 the-known.F-GEN  
 among the seven known colors of the spectrum  
 (Ryding, 2005, 338)



These cardinals (3-10) therefore place unusual (ie polar) gender requirements on heads that they modify.<sup>12</sup>

---

<sup>12</sup>A preliminary sketch of this view was given at the Christchurch Winter School in LFG in 2005 and has been available on the web in handout form since that time at <http://privatewww.essex.ac.uk/~louisa/agr/finalday3.pdf> By extraordinary coincidence, some virtually identical lexical entries appear in ElSaadany (2008), where they are (incorrectly) characterized as implementing a CONCORD/INDEX mismatch analysis. An analysis of Welsh numeral noun constructions positing an INDEX/CONCORD mismatch is presented in Mittendorf and Sadler (2005).

# Separating Morphological Form from Concord Features

Constraint-based approaches to syntactic agreement in HPSG and LFG typically adopt a distinction between two sets of agreement features within the NP (CONCORD and INDEX features).



In HPSG Wechsler and Zlatić (2000) (see also Kathol (1999) for a related view) CONCORD features are closely related (but distinct from) morphological (inflectional) classes and INDEX features are related to (but distinct from) features of the semantics.<sup>15</sup>

---

<sup>15</sup>For Wechsler and Zlatić, CONCORD features are typically relevant for NP-internal concord (between nouns, determiners and adjectives), while INDEX features are typically relevant to subject-verb agreement and for pronominal anaphora.

King and Dalrymple (2004) make a related but nonetheless distinct proposal to distinguish between `CONCORD` and `INDEX` features associated with nominal f-structures in `LFG`. The key difference here is distributivity. King and Dalrymple (2004) show that either or both of `INDEX` and `CONCORD` can be relevant for NP-internal agreement in a given language.

(11) For any *distributive* property  $P$  and set  $s$ ,  
 $P(s)$  iff  $\forall f \in s.P(f)$ .

For any *nondistributive* property  $P$  and set  $s$ ,  
 $P(s)$  iff  $P$  holds of  $s$  itself.

Dalrymple and Kaplan (2000)

$$(12) \quad \text{NP} \rightarrow \begin{array}{c} \text{N} \\ \uparrow = \downarrow \end{array} \quad \begin{array}{c} \text{AP} \\ \downarrow \in (\uparrow \text{ADJ}) \end{array} \quad 19$$

<sup>19</sup>This highly over-simplified view of the structure of the MSA NP is, curiously, also adopted in ElSaadany (2008) Here however, the following extremely curious explanation is given for the f-description associated with the AP “the ‘ $\downarrow \in (\uparrow \text{ADJ})$ ’ equation beneath AP indicates that the feature of that node is shared with both the adjacent head N and the higher node dominating it”. (ElSaadany, 2008, 18) I am completely at a loss to understand what this is supposed to mean or to give it any interpretation within the formalism.


It's currently an open issue whether MSA NP-internal agreement follows the IND, CONC or IND/CONC system: for concreteness we assume CONC at this point.

- (13) *-l-ğamīl-u* ((ADJ ∈ ↑) DET) = DEF  
 ((ADJ ∈ ↑) CASE) = NOM  
 ((ADJ ∈ ↑) CONC GEND) = MASC  
 ((ADJ ∈ ↑) CONC NUM) = SING  
 (↑ PRED) = 'BEAUTIFUL'

On this view, the lexical entry for *morphologically* feminine *ḥamsa-t* ‘five’ contains the following:<sup>22</sup>

- (14) *ḥamsa-t* (( ADJ ∈ ↑ ) CONC GEND) =<sub>c</sub> MASC  
 ( (ADJ ∈ ↑) CONC NUM) =<sub>c</sub> PLUR  
 (↑ PRED) = FIVE

---

<sup>22</sup>The lexical entry in (15) is reproduced from Dalrymple and Sadler (2004). By extraordinary coincidence, a virtually identical lexical entry (but for *sabʿ-at* ‘seven’) appears in ElSaadany (2008, 27). While the text below the lexical entry correct asserts that this displays a form/CONCORD mismatch (see directly below), that above very curiously asserts “in such cases, the numeral modifier contributed the INDEX NUM and the nominal codtributes the CONCORD NUM”. It is difficult, hence, to avoid the speculation that the author of ElSaadany (2008) does not entirely grasp the content of his own (?) lexical entry. 





Similar view of plurals also: we might propose that *morphologically* feminine singular adjectives and demonstratives constrain the containing f-structure to have *either* fem singular concord, *or* a plural concord feature and be marked non-human. Morphologically feminine singular pronouns and relative pronouns constrain their antecedent's index to be either fem singular, or to be plural non-human.

# Specifying agreement relations between CONCORD features

CASE	NOM				
HUM	—				
PRED	'HOUSES'				
DET	DEF				
IND		[ PERS 3 ]			
		[ NUM PL ]			
		[ GEN MASC ]			
CONC		[ NUM PL ]			
		[ GEN MASC ]			
ADJ		{ [ PRED 'FIVE' ]	[ PRED 'NEW' ]		
		[ CONC [ GEN FEM ] ]	[ CONC [ NUM SG ] ]		
			[ GEN FEM ]		

- low numerals show gender polarity
- FSG adjectives combine with FSG or MPL nonhuman nouns

(15) *hamsa-t* (( ADJ ∈ ↑ ) CONC GEND) =<sub>c</sub> MASC  
 ( (ADJ ∈ ↑) CONC NUM) =<sub>c</sub> PLUR  
 (↑ CONC GEND) = FEM  
 (↑ PRED) = FIVE

(16) *jamil̄-at* (( ADJ ∈ ↑ ) CONC GEND) =<sub>c</sub> MASC  
 ( (ADJ ∈ ↑) CONC NUM) =<sub>c</sub> PLUR  
 ( (ADJ ∈ ↑) HUM) =<sub>c</sub> –  
 (↑ PRED) = BEAUTIFUL

## Summary

(17) Morphological form of noun, adjective, number:

noun	feats	adj	card (3-10)
man	MS	MS	
book	MS	MS	
woman	FS	FS	
car	FS	FS	
men	MPL	MPL	F
books	MPL	FS	F
woman	FPL	FPL	M
cars	FPL	FS	M

- Postnominal numerals are (simple or complex) modifiers/adjectives, some of which show gender polarity in their agreement behaviour
- My current working hypothesis is that the mismatch is between MORPHOLOGY and CONCORD (ie what Sadler and Spencer (2001) call m-features and s-features).

## Prenominal Numerals

Again according to Ryding, if the counted noun is indefinite that the numeral (3-10) precedes: she notes “the case marker on the numeral varies according to its role in the sentence and it is considered definite because it is in an *ʔiḍāfa* relationship with the noun “ (so it does not take nunation). “The counted noun itself is plural, indefinite and in the genitive case.” (Ryding, 2005, 335)

Parallels with the CONSTRUCT STATE (ʔiḏāfa) and the behaviour of prenominal quantifiers might suggest a Head-Argument style of analysis. We know from other constructions that adjectives can appear in the CS.

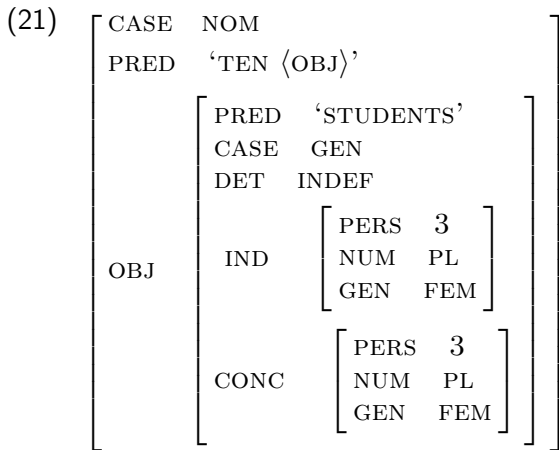


- (18) 'arba<sup>c</sup>a-t-u kutub-in  
four-F.NOM books.M.GEN  
four books (Kremers 2003 62: 47b)
- (19) ?ashrat-u riḡal-i-n  
ten.F-NOM men.M-GEN-INDEF  
ten men
- (20) ?ashr-a talebat-i-n  
ten.M-ACC students.F-GEN-INDEF  
ten (female) students

Dalrymple and Sadler (2004) tentatively proposed (using OBL for the function given here as OBJ):<sup>33</sup>

---

<sup>33</sup>See also the largely identical f-structure also “tentatively shown” in ElSaadany (2008, 27). Oddly this two-tier f-structure first appears in a section allegedly concerned with INDEX/CONCORD mismatches. In previous LFG work, Mittendorf and Sadler (2005) argue that Welsh Num Noun constructions (in which the Noun is obligatorily SING and controls Adjective (but not demonstrative) agreement) involve an INDEX/CONCORD mismatch: the numeral provides a PL INDEX but requires a Noun with SING CONCORD: the crucial element of this (other) approach is that the prenominal numeral is an adjunct/modifier of the head noun, rather than a selecting head.



A (morphologically) feminine numeral requires its complement to be masculine and vice versa - polar concord behaviour:

- (22) *?ashra-t-u* (↑ PRED) = ‘TEN⟨OBJ⟩’  
(↑ CASE) = (NOM)  
(↑ OBJ CASE) =<sub>c</sub> GEN  
(↑ OBJ DET) =<sub>c</sub> INDEF  
(↑ OBJ CONC GEND) =<sub>c</sub> MASC  
(↑ OBJ CONC NUM) =<sub>c</sub> PLUR

The prenominal (quantifier?) use of the numeral and the postnominal (adjectival) use of the numeral show the same agreement information

*3-10.Fem.Sg (Prenom)* (↑ OBJ CONC GEND) =<sub>c</sub> MASC  
(↑ OBJ CONC NUM) =<sub>c</sub> PLUR

*3-10.Fem.Sg (Postnom)* (( ADJ ∈ ↑ ) CONC GEND) =<sub>c</sub> MASC  
( (ADJ ∈ ↑ ) CONC NUM) =<sub>c</sub> PLUR

## Generalizing over Patterns

Common Template:

3-10.FSG.ADJ  $\equiv$  ( $\uparrow$  AGRPATH) = %CONTROLLER  
(%CONTROLLER CONC GEND) =<sub>c</sub> MASC  
(%CONTROLLER CONC NUM) =<sub>c</sub> PLUR

Defining Agreement Path (Otoguro, 2006)

(AGRPATH) = { OBJ | ADJ  $\in$   $\uparrow$  }

Arguments in favour of the two tier approach maybe come from the commonality with the other uses of the genitive construction.

## Polar Quantifiers

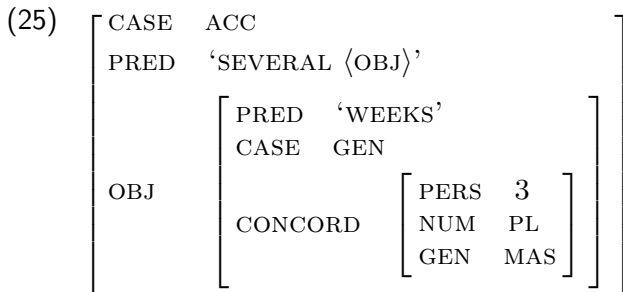
At least one prenominal quantifier (*biḍḥ*, *biḍḥat* ('several') also shows morphosyntactic polarity (and appears in *iḍāfa* construction with the following head). This quantifier often co-occurs with numeral expressions such as 'hundreds'.<sup>39</sup>

- (23) ya-taṭallab-u                      biḍḥat-a              ḥasābīḥ-a  
 3SM-require.IMPERF-3SM several.F-ACC week.MPL-GEN  
 It requires several weeks (Ryding, 2005, 232)

- (24) baḥd-a biḍḥ-i                      daqaaḥiq-a  
 after    several.M-GEN minute.FPL-GEN  
 in a few minutes (Ryding, 2005, 232)

<sup>39</sup>EISaadany (2008) gives a lexical entry for *biḍḥ-at-a* which is very similar to (26) and identical (pace typos) to that appearing in Dalrymple and Sadler (2004), though again it is curiously referred to as involving an INDEX/CONCORD mismatch.

F-structure for (23):





- (26) *bid<sup>c</sup>-at-a* (↑ PRED) = ‘SOME⟨OBJ⟩’  
(↑ CASE) = ACC  
(↑ OBJ CONC GEN) =<sub>c</sub> MASC  
(↑ OBJ CONC NUM) =<sub>c</sub> PLUR  
(↑ OBJ DET) =<sub>c</sub> INDEF  
(↑ OBJ CASE) = GEN

## Index Transparency

Note that externally, the whole construction must have an INDEX value. I assume that externally a numerally quantified phrase such as *?ashrat-u riğal-i-n* will behave (ie control anaphora) as a human MPL and *?ashr-a talebat-i-in* as a human FPL and *'arba<sup>c</sup>a-t-u kutub-in* as a nonhuman MPL, indicating that the INDEX features (and HUM) are available in the outermost f-structure. As a first approximation:<sup>43</sup>

$$(27) (\uparrow \text{INDEX}) = (\uparrow \text{OBJ INDEX})$$

---

<sup>43</sup>But note work to be done on analysis of complex numerals.

## Some Issues for Future Work

- What is the treatment of various types of complex numerals occurring postnominal as ADJ, and prenominally (complex numerals do not share the behaviour of 3-10 prenominally)? The analysis of complex numerals appears to present challenges for LFG even without the additional complication of complex agreement facts.
- what further evidence for or against various 'locations' for the polar mismatch will come from the rest of the numeral system?
- what are the facts in a specific vernacular?

## Specific Data Issues

[1] Ryding's description suggests that, with respect to simple numerals 3-10, in indefinite NPs the numeral is prenominal, and postnominal in the definite NP (where it then shows definiteness agreement). Are these grammatical and what are the optionality facts for polar agreement?

- (28) banāt-un                                  sabʕ-un  
      girls.FEM.PL-NOM.IND seven.MAS-NOM.IND  
      /sabʕat-un  
      /seven.FEM-NOM.IND  
      seven girls (ElSaadany, 2008, 33)

[2] What is the status of the following data?

Ryding suggests that the material following *biḍḥ* would be GEN.

(29) *biḍḥ-u*            *ʔašra-ta*    *safīn-at-an*  
some-M-NOM ten-F-ACC ship-F.SG-ACC  
some ten ships (ElSaadany, 2008, 33)

(30) *biḍḥ-u*            *ʔašra-ta*    *qārib-an*  
some-F-NOM ten-M.ACC ship-M.SG-ACC  
some ten boats (ElSaadany, 2008, 33)

[3] if it is correct that a prenominal numeral (3-10) forms an indefinite *iḍāfa* with the noun it is in construction with, is an example with a demonstrative impossible or not? How do the agreement facts work out for postnominal adjectives in combination with Num Noun constructions?

[4] These examples have somehow got into my data and need checking.

(31) bid<sup>c</sup>-at-u                    ṭullāb-i-n  
several-FEM-NOM students.MASC-GEN-INDEF  
some students

(32) bid<sup>c</sup>-u                    ṭālib-āt-i-n  
several-MASC-NOM student-FEM.PL-GEN-INDEF  
some(female) students

[5] Beyond the simple numerals, there will be some interesting interactions with NUM as the overall structure is PL despite the occurrence of a SG noun as the following indicate:

- (33) haʔulāʔi    ʔal-ḥamsa-ta    ʕašara  
this.PL.OBL the-five-FEM.ACC ten-MAS.ACC  
ṭālib-an                            ʔal-judud-i  
student.MAS.SG-ACC.IND the-new.MAS.PL-GEN  
these fifteen new (male) students (ElSaadany, 2008, 34)
- (34) ʔal-ḥamsa-ta    ʕašara    ṭālib-an  
the-five-FEM.ACC ten-MAS.ACC student.MAS.SG-ACC.IND  
ʔal-judud-ihaʔulāʔi  
the-new.MAS.PL-GEN this.PL.OBL  
these fifteen new (male) students (ElSaadany, 2008, 35)



- Dalrymple, Mary and Ronald M. Kaplan. 2000. Feature indeterminacy and feature resolution. *Language* 76(4):759–798.
- Dalrymple, Mary and Louisa Sadler. 2004. Advanced Topic in LFG: Agreement 3 Index and Concord. Handout for Winter School University of Christchurch.
- ElSaadany, Kamel. 2008. Toward an LFG account of agreement mismatches of numerals within Arabic NPs. In D. Parkinson, ed., *Perspectives on Arabic Linguistics: Papers from the annual symposium on Arabic Linguistics*, pages 13–42. Amsterdam: John Benjamins.
- Kathol, Andreas. 1999. Agreement and the syntax-morphology interface in HPSG. In R. Levine and G. Green, eds., *Studies in Current Phrase Structure Grammar*, pages 223–274. Cambridge, UK: Cambridge University Press.
- King, Tracy Holloway and Mary Dalrymple. 2004. Determiner agreement and noun conjunction. *Journal of Linguistics* 40(1):69–104.
- Mittendorf, Ingo and Louisa Sadler. 2005. Nouns, Numerals and Number in Welsh NPs. In M. Butt and T. H. King, eds., *Proceedings of the LFG05 Conference*. Stanford, CA: CSLI Publications.
- Otoguro, Ryo. 2006. Agreement and path specification in Icelandic. In *Proceedings of CLS 2005*. Chicago Linguistic Society.

- Ryding, Karin. 2005. *A Reference Grammar of Modern Standard Arabic*. Cambridge, UK: Cambridge University Press.
- Sadler, Louisa and Andrew Spencer. 2001. Syntax as an exponent of morphological features. *Yearbook of Morphology 2000* pages 71–97.
- Wechsler, Stephen and Larisa Zlatić. 2000. A theory of agreement and its application to Serbo-Croatian. *Language* 76(4):759–798.