Numeral Gender Variation in Semitic and Slavic at the Interfaces

Abdelkader Fassi Fehri Mohammed V University, Linguistic Society of Morocco, Rabat

Interfacing (natural) numbers and numerals (their linguistic expressions) appears to be more complex than just making use of 'smooth' Merge (as in Watanabe's 2017 implementation of Chomsky's 2008 set-theoretic approach to number). Mediating processes, functions, or category projections are necessary, including e.g. CLASSIFIER, SET, NUMBER (Kayne 2016, etc.). The locus of variation (micro-parametric) is identified as (a) the lexicon of morphosyntactic features, or (b) the vocabulary properties at PF (pronounced or silent; Kayne 2005, Cinque 1999, Borer 2005. But the variation seems also to be macro-parametric, when it concerns a group of varieties 'genealogically' related to a language macro-variety, e.g. Semitic or Slavic here. The paper will focus on Gen(der) distribution in these varieties. It identifies three distinct behaviours and properties of Gen in Semitic numeral constructions, associated with three distinct uses or kinds of numerals: *n-numerals*, in e.g. counting number sequences, *c-numerals* in cardinal dPs, and *o-numerals* in ordinal dPs. Gen is also shown to play an important role in Slavic numeral structures, typically in so-called *collective numerals*. While Semitic and Slavic vary quite freely along micro-scales, it is shown that at least the 'gender polarity' pattern, typically known to occur in Semitic construct state (CS) numerals, has no obvious parallel in Slavic. Our analysis is implemented in a root-category syntax model (Marantz 1997, 2006, Borer 2005, 2013, Harley 2014, etc.).

1. Polycategorial and specific numerals, roots and categories

1.1. Numerals are often analysed as nouns or adjectives (Hurford 1975, Greenberg 1978, Kayne 2003, Ionin & Matushansky 2006, Zweig 2005, Stavrou & Terzi 2011, etc.), but in fact they (almost) exhaust the inventory of 'lexical' or 'functional' categories, i.e. *p*, *v*, *adverb*, *q*, *d*, etc. (Selkirk 1977, Jacendoff 1977, Barwise & Cooper 1981, Corver & Zwarts 2006, etc.). They are also exhibiting specific behaviour when associated with 'derivational' or 'inflectional' properties (number, gender, definiteness, case), enter *agreement* or *case* relations (as goals or probes), occupy various syntactic positions (arguments, predicates, modifiers, determiners), and acquire various *senses* (and 'flavours') as *number terms*, *cardinals, ordinals, fractions*, etc. (Ruckowski 2002, 2007, Zabbal 2005, Le Vieillard 2011, Miechowicz-Mathiasen 2012, Ouwayda 2014).

1.2. Polycategorial, specific, and diverse derivations can be captured if numerals are first born as *acategorial roots*. They are then compositionally 'molded' with various positions, categorizations, and inflections, to yield specific senses. In constructions (1), from Standard Arabic (= SA), all the numeral forms are reasonably derived from the same root <u>*tlt*</u>: (1) a. <u>*talaat*</u> 'three (cardinal)'; b. <u>*taalit*</u> 'third (ordinal)'; c. <u>*tulut*</u> 'a third (fraction)'; d. <u>*taaluut*</u> 'triad' (noun) ; e. <u>*tulaat-a*</u> 'by three (distributive adverb)' f. <u>*mutallat*</u> 'tripled' (deverbal adj.). (2) a. <u>*Patlata l-qawm-u*</u> 'the people became three' (<u>inchoative</u>)

b. *tallat-a r-rajul-u l-?amr-a* 'the man made the act three times';

"The man tripled the matter" (multiplied it by three; <u>causative</u>)

In (1), the common numeral base is carrying the core meaning of cardinality 'three-some, three-hood', represented by the root <u>*t*lt</u>. The infixal vowel pattern (plus the prefix in 1f) is providing categorical information (*n*, *a*, *adv*), but also specification of the derivative sense dealt with (ordinal, fraction, cardinal, ...). In (2), the verbal forms IV and II have a 'nominal' *n* source, and the morphological form contributes category specification as *v*, and information about aditicity and aspect (as in Hale & Keyser' 2002 L-syntax).

2. Gen variation in Semitic and kinds of numerals

2.1. N-numerals are used in counting sequences to designate or *count numbers, not objects.* Their interpretation does not depend on any (over or silent) nP. They are rather bare (unlike c-numerals), of type *n*, and behave like proper names. In Semitic, they are distinguished

typically by their Gen. In Arabic varieties, they are normally formed by suffixing a Gen mark (identical to the feminine -at, a, or i, depending on the variety) obligatorily:

(3) a. Lebanese LA *tlat-i* 'three', *arbas-i* 'four', *xams-i* 'five' b. Moroccan MA *sett-a* 'six' *sabs-a*'seven', ... c. SA: *talaat-at*, *?arbas-at*, *xams-at*, ... *sasr-at* 'ten'

Bare un-suffixed counterparts, which lack the feminine mark, are not suitable for such use. The Gen mark functions as *group* (or set) *classifier* (Fassi Fehri 2016). When constructed with a predicate, the latter obligatorily agrees with the n-numeral in feminine (singular): (4) *talaat-at* hiya 2+1; 'three-fem-nom she 2+1'; "Three is 2+1".

Note that Hebrew appears to have made the opposite choice by taking un-suffixed forms to be the suitable norm for n-numerals, although they are still distinguished from c-numerals: (5) *šaloš, arba, xameš, šeš, šev-a, šmone, teš-a,* ...

2.2. In ordinal nP constructions, ordinals behave essentially like adjective modifiers, and the feminine affix is a probe, rather than goal, valued through the Gen value of the counted nP: (6) a. MA: *taleb-a talet-a* 'student-fem third-fem'; "a third female student"

b. SA: *taalib-at-un taalit-at-un* 'student-fem-nom third-fem-nom'; "a third female student" Any Gen mismatch is excluded (Shlonsky 2004, Sichel 2012; Fassi Fehri 1999).

2.3. As for c-numerals, they have a 'mixed' status. In addition to the obligatory presence of a nP in their syntax and interpretation, they are characterized by *Gen polarity* (or incongruent) agreement (Hetzron 1987, Lecarme 2002) when in a CS configuration. The predicate typically agrees in Gen (and Num) with the nP (hidden or overt), rather than with the cardinal: (7) a. <u>talaat-u</u> (fatay-aati-n) ji?-na b. <u>talaat-u</u> (?awlaad-in) jaa?-uu

. three-nom (girls-gen) came-fem.pl three-fem-nom (boys-gen) came- pl . Three (girls) came. Three(boys) came.

Rather than thinking about the distribution of Gen here as sort of *Gen switch* agreement (Halle 1994), it is more plausible that the Gen of the numeral is simply 'missing' in (7a), as strongly suggested by the distribution of Gen in c-numerals in dialectal Arabic (Cowan 1972), or Modern Hebrew (Bolozky & Haydar 1986). In the latter, the feminine occurs on the cardinal only in the free state (b), and is missing in the CS (a), the Gen of the noun being irrelevant:
(8) LA. a. *xams wlaad (banaat)* "five boys/girls"; b. *wlaad (banaat) xams-i* "five boys/girls"
(9) MA. a. *sett ayyam* "six days"; b. *ayyam sett-a* "six days" (the fem. mark in bold) **3. Slavic.** The triple taxonomy of numerals has its counterpart in Slavic. O-numerals are more 'adjectival' than c-numerals. Ordinals normally agree in Gen, whereas cardinals do so only marginally. N-numerals are found with their own characteristic Gen. In Polish, non-virile unmarked forms are used (in contrast with o-numerals and c-numerals; Wagiel 2014, p.c.): (10) *jeden* 'one', *dwa* 'two', *trzy* 'three', *cztery* 'four', *pięć* 'five', etc.

In Czech, the feminine (marked) forms of '1' and '2' are strongly preferred in the counting sequence (Gen not being transparent with high numerals): *jedna* 'one.f', *dvě* 'two.f', ... The talk will focus more on so-called 'collective' numerals, however, because they offer a fine-grained taxonomy of c-numerals, based on three distinct kinds of Gen: (a) neutral, (b) fem, and (c) variable. The examples are from the Bosnian-Serbo-Croatian variety (Kim 2009,

Lučić 2015), but comparable contrasts in Polish and Russian will be presented:

(11) *collective numerals*: <u>neutral Gen</u> (singular): *troje ljudi* 'three people' ; *petero putnika* 'five travellers'; *dvoje djece* 'two children'; *troje teladi* 'three calves';

(12) *numerical adjectives:* <u>variable Gen</u> (exclusively plural) : *dvoji*,-a,-e 'two', *troji*,-*a*,-*e* 'three', *peteri*,-*a*,-*o* 'five'; *dvoje čarape* 'two pairs of socks'; *troji svatovi* 'three groups of wedding guests'; *petora vrata* 'five doors'; groups: *svatovi* 'wedding guests';

(13) *numerical nouns*: <u>feminine</u> (singular, ending in *-ica*): *dvojica braće* 'two brothers'; *trojica igrača* 'three players'; *petorica putnika* 'five travelers'.

Some References. Cowan, W. 1972. The Historical Syntax of the Arabic Numbers. *Glossa* 6.2. Fassi Fehri, A. 2016. Semantic Gender Diversity and Its Architecture in the Grammar of Arabic. *BJALL* 8.1. Kayne, R. 2016. Some Thoughts on *One* and *Two* and other Numerals. Ms. NYU. Zabbal, Y. 2005. The Syntax of Numeral Expressions. Ms. Univ. of Amherst.