On the Typology of Number Concord[*]

Sascha Alexeyenko
University of Göttingen

June 23, 2021

1 Introduction: the typological landscape

- In this work, I am interested in the mechanics of **DP-internal agreement** (concord), looked at through the prism of number marking, specifically in connection with **numerals**.

- Given that languages are able to use **plural** or **classifiers** for nouns modified by numerals, there are 4 logically possible types of languages (Rijkhoff 2004: 29):

  ① numeral + noun + plural

  ② numeral + noun

  ③ numeral + classifier + noun

  ④ numeral + classifier + noun + plural   (very few, possibly no languages)

- I won't be concerned with classifier languages here. And I'm only interested in languages that do have plural marking at least on some nouns in the absence of numerals to begin with. In other words, the language types of interest are these:[1]

  ① | noun + plural | and | numeral + noun + plural | **NC**

  ② | noun + plural | and | numeral + noun | **ND**

- An example of a "**number concord**" (NC) language is English:[2]

  (1)   a.   book; book-s
             b.   five *book/book-s

- An example of a "**number discord**" (ND) language is Lezgian:[3]

  (2)   a.   ktab; ktab-ar
             b.   wad ktab/*ktab-ar
                     five book.SG/*book-PL

  (Haspelmath 1993: 81)

---

[1]The reality is more complicated than this. First, languages may be NC wrt some numerals/nouns, but ND wrt others. Second, in many languages, plural marking is optional in the context of numerals, or in general. Note also that a further type of languages is analytically possible, in which plural marking on nouns can/must be used only in the presence of numerals. I don't know if such languages exist.

[2]NC languages correspond to "*singular object noun* languages" in Rijkhoff (2004).

[3]ND languages correspond (at least in principle) to "*set noun* languages" in Rijkhoff (2004).

- Note that ND languages are not a marginal phenomenon: "from a cross-linguistic perspective number marking (with or without an attributive numeral) seems to be the exception rather than the rule" (Rijkhoff 2004: 38).

- The sample of ND languages that I worked with so far:

(3) a. Abkhaz   (NW Caucasian/Abkhazo-Adyghean)
     b. Armenian   (*Eastern*; Indo-European, isolate)
     c. Basque   (language isolate)
     d. Finnish   (Uralic, Finnic)
     e. Georgian   (Kartvelian)
     f. Hungarian   (Uralic, Ugric)
     g. Lezgian   (NE Caucasian/Nakh-Daghestanian)
     h. Turkish   (Turkic, Oghuz)
     i. Welsh   (Indo-European, Celtic)

Questions:

① What **determines** whether a language is NC or ND?
(syntax, semantics, both?)

② Do we have a theory of agreement/concord that **predicts** for NC and ND languages
  - the pattern of DP-external verbal agreement wrt number, and
  - the pattern of DP-internal number agreement on other elements (Det, Adj, Num)?

2 Some existing answers

2.1 The semantics of singular form (Rijkhoff 2004; Bale et al. 2011)

Ingredients:

- variation in the semantics of the SG form:

  $\boxed{\text{NC}}$ **singular**:

  (4) $[\![\text{book}]\!] = \{a, b, c\}$

  $\boxed{\text{ND}}$ **number-neutral**/transnumeral/general number (Greenberg 1974; Corbett 2000):

  (5) a. $[\![\text{kitap}]\!] = \{a, b, c, a \oplus b, b \oplus c, a \oplus c, a \oplus b \oplus c\} = [\![\text{book-s}]\!]$
       b. $[\![\text{kitap-lar}]\!] = \{a \oplus b, b \oplus c, a \oplus c, a \oplus b \oplus c\}$ (Turkish)

(6)   Hüseyin-le Aydın rehber(-ler).
      H.-COM   A.   guide(-PL)
      'Hüseyin and Aydın are guides.'

(7)   Masa-da   tabak(-lar) var.
      table-LOC plate(-PL) exist
      SG: 'There is a plate/are plates on the table.'
      PL: 'There are plates on the table.'

(8)   Ahmet kitap(-lar) oku-du.
      A.   book(-PL) read-PAST
      SG: 'Ahmet read a book/books.'
      PL: 'Ahmet read books.'   (Görgülü 2010, 2012)

- semantics of numerals is defined such that they can only combine with NPs containing pluralities in their denotations (returning pluralities of specified cardinality):

  - only PL in English
  - SG/PL in Turkish and Western Armenian

- numeral denotations are further fine-tuned so that they can (WA) or cannot (T) combine with "pure" pluralities:[4]

  - only SG in Turkish
  - SG/PL in Western Armenian

⇒ The NC-ND distinction is thus primarily about semantics, morphosyntax just tags along.

Problem: (Ionin & Matushansky 2018)

- Not all ND languages have transnumeral SG: Hungarian, Turkish, Western Armenian do, but Finnish (9) and Welsh (10) do not.

  (9) Luin   kirjan   / kirjaa.
       read.1SG book.ACC.SG   book.PART.SG
       = 'I read a/the book.'
       ≠ 'I read (the) books.'   (Ionin & Matushansky 2018: 90)

  (10) Gwelodd   Rhiannon ddraig.
        see.3SG.PAST Rhiannon dragon.SG
        = 'Rhiannon saw a dragon.'
        ≠ 'Rhiannon saw dragons.'   (Ionin & Matushansky 2018: 91)

2.2 The presence/absence of formal agreement (Ionin & Matushansky 2018)

Ingredients:

- both in NC and in ND languages, numerals combine with semantically singular NPs and return plural predicates

- the plural marking on the noun in the presence of a numeral in NC languages is the result of agreement and is not interpreted

Problem:

- "[I]t is not our goal to propose a formal theory of number agreement [...]" (p. 94)

- The question of what determines the NC-ND split gets reformulated as the question of what determines the presence/absence of formal number agreement inside the DP.

---

[4]For languages like Turkish, it could also be said that SG wins because the competing PL form is more marked, while giving rise to the same meaning in the context of numerals. See Ortmann (2000) for a discussion of economy in connection with plural marking.

3 Possible points of variation

① semantics of numerals   (not easily falsifiable)

② semantics of SG: singular vs. transnumeral
   semantics of PL: plural vs. transnumeral

③ syntactic status of numerals: heads vs. phrases/specs   (directly testable?)
   syntactic category of numerals: nominal vs. adjectival
   structural position/height of numerals and PL
   featural specification of numerals: [iPL] vs. [iSG]

While I do believe that some of the points in ③ may be on the right track, the story can't be as simple as that, as I will show below, introducing some more facts about ND languages.

3.1 NC numerals are [iPL], ND numerals are [iSG]   (cf. Danon 2012)

I.e. a system in which

- numerals are heads in the extended NP (c-command with the locus of PL)

- syntactic interpretability doesn't necessarily reflect semantic impact

Problems:

- brute force (but would probably work for Georgian and Turkish)

- DP-internal PL with numerals (Basque, Finnish, Hungarian, Welsh)

(11) y   tri   / pedwar / pum llun   hyn
      the three.M   four.M   five picture.SG DEM.PL
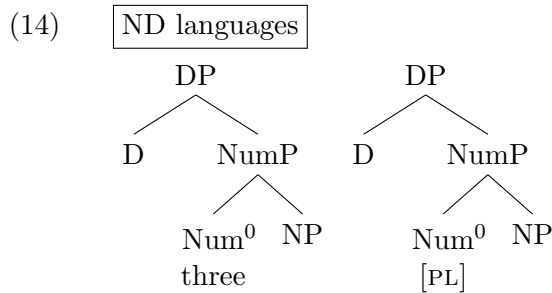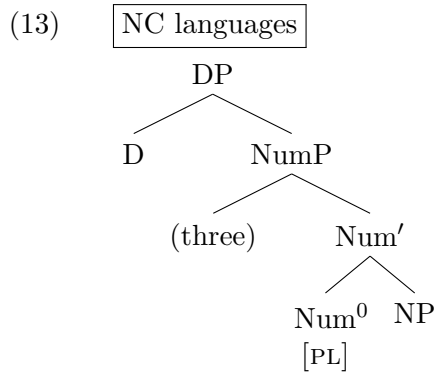   'these three/four/five pictures'   (Welsh; Borsley et al. 2007: 165)

(12) a. Minä odotin |kolme| pitkästyttävä-ä minuutti-a.
          I   waited three.SG boring-PAR.SG   minute-PAR.SG

      b. Minä odotin pitkästyttävä-t |kolme| minuutti-a.
          I   waited boring-ACC.PL   three.SG minute-PAR.SG

      c. Minä odotin ne |kolme| pitkästyttävä-ä minuutti-a.
          I   waited those.PL three.SG boring-PAR.SG   minute-PAR.SG

      d. Minä odotin ne   pitkästyttävä-t |kolme| minuutti-a.
          I   waited those.PL boring-ACC.PL   three.SG minute-PAR.SG
                                           (Finnish; based on Brattico 2010: 60, 66)

   – Where does PL on demonstratives/adjectives come from if numerals are [iSG]?

   – Wherever it comes from, why can't the noun be marked PL in the same way as well?

3.2 NC numerals are specs, ND numerals are heads   (cf. Danon 2012)

(potentially also related to their categorial status)

- Numerals and the Num^0 head responsible for plurality in the absence of numerals are in complementary distribution in ND languages because they compete for the same slot.

(13) NC languages

```
              DP
           ／    ＼
          D      NumP
               ／    ＼
          (three)    Num′
                   ／    ＼
                Num⁰    NP
                [PL]
```

(14) ND languages

```
         DP              DP
      ／    ＼        ／    ＼
     D      NumP     D      NumP
          ／    ＼         ／    ＼
       Num⁰    NP       Num⁰    NP
       three             [PL]
```

Problem:

- independent evidence?

- primarily from case licensing:

    (15) Pekka osti   ne   kolme auto-a.
          P.   bought those.PL three car-PAR.SG
          'Pekka bought those three cars.'

- but numerals in many (most?) ND languages do not license case
  (e.g. Armenian, Georgian, Turkish, Welsh)

- also numerals in Finnish can (must) stop licensing case when structural (inherent) case is
  assigned externally:

    (16) Pekka osti   ne   kolme-t   auto-t.
          P.   bought those.PL three-ACC.PL car-ACC.PL
          'Pekka bought those three cars.'

4 Towards an analysis

4.1 DP-internal agreement

The data from ND languages (and in particular Finnish) may be interpreted in the following
way:

- some abstract element is able to license PL marking on pre-numeral elements (presumably
  the same one that licenses PL marking in the absence of numerals)

- numerals serve as **barriers** for number-feature transmission to lower elements including
  the noun

Possible ways of modeling their barrierhood:

- NumP is a phase

- intervention effect (numerals are nominal)

- intervention effect (numerals have their own number feature)

The intuition behind the latter point is also expressed in Danon's (2012) Number Constraint:

(17) **Number Constraint**   (Danon 2012)
      The spec-head construction is not possible if the numeral carries its own morpho-syntactic number feature.
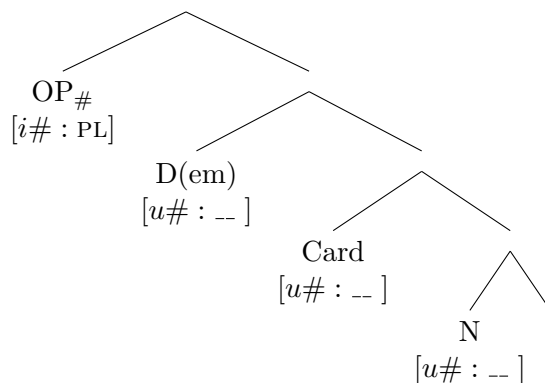
But how do we know that a numeral has its own number feature? Finnish shows very well that numerals stop being barriers for plural-feature transmission when they are themselves marked for plural.

- never marked for number $[\,—\,]$

- marked for number in numeral-noun construction $[u\text{Num: }\_\_\,]$

- can be marked for number, but not in numeral-noun construction[5] $[i\text{Num: }\_\_\,]$

- mixed

(18)   English   (transparent)

OP$_\#$
$[i\# : \text{PL}]$
D(em)
$[u\# : \_\_\,]$
Card
$[-]$
N
$[u\# : \_\_\,]$
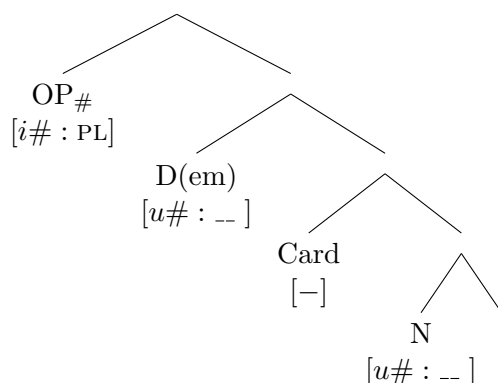
(19)   Icelandic   (transparent)

OP$_\#$
$[i\# : \text{PL}]$
D(em)
$[u\# : \_\_\,]$
Card
$[u\# : \_\_\,]$
N
$[u\# : \_\_\,]$

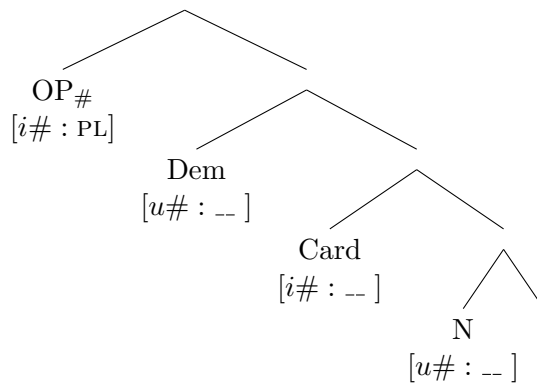[5]Alternatively, the numeral may carry an interpretable and SG-valued number feature in this case. This seems to me to be a more stipulative option.

(20)   Finnish (bare numeral)   (barrier)

```
            OP_#
          [i# : PL]
                  Dem
                [u# : __ ]
                          Card
                        [i# : __ ]
                                  N
                                [u# : __ ]
```

Comments:

- The number feature on N in the last derivation gets checked but not valued, which leads to the insertion of the default value, i.e. the SG marking is default rather than singular number in the case of ND languages (cf. Pesetsky 2013 for default genitive).

Summary of implementational assumptions:

- null number operator OP$_\#$ on top of DP (Sauerland 2003)

- OP$_\#$ may only be inserted if it is involved in checking and valuation of features

- Upward Agree

- the mapping principles between morphological exponence and syntactic features

- unchecked features lead to a crash, unvalued features trigger insertion of default values

4.2 DP-external agreement

Based on my sample, ND languages

- either do not (really) have number agreement on the verb at all (e.g. Welsh),

- or display it only with numeral-less morphologically plural DPs (e.g. Georgian),

- or are Finnish.

(21)   Ne   kaksi   pien-tä   auto-a   seiso-ivat/seiso-i   tiellä.
         those.PL two.SG small-SG.PAR car-SG.PAR stand-PAST.3PL/stand-PAST.3SG road.ADE
         'Those two small cars stood at the road.'   (PL literary F, SG colloquial F)

(22)   Kaksi   pien-tä   auto-a   *seiso-ivat/seiso-i   tiellä.
         two.SG small-SG.PAR car-SG.PAR   stand-PAST.3PL/stand-PAST.3SG road.ADE
         'Two small cars stood at the road.'   (literary/colloquial F)

Preliminary generalization:

- Finnish is the only language in the sample that has number-marked demonstratives in the absence of (definite) articles

- all the other languages are such that either their demonstratives obligatorily require the presence of definite article (e.g. Welsh) or their demonstratives (as well as definite article if available) are not number-marked (e.g. Georgian and Armenian, respectively)

5 Outlook

- further typological work

- interaction with other features (case, definiteness)

- patterns of number agreement on other elements (adjectives)

References