Decomposing cardinals

Heidi Klockmann
heidi.e.klockmann@uia.no
University of Agder, Norway

Numerals in Grammar and Beyond Workshop
Leiden University, Netherlands
October 17-18, 2019

1 Introduction

Proposal: Numerals are internally complex. Their similarity to nouns, verbs, and adjectives is due to shared functional structure with noun, verb, and adjective categories in their respective languages.

Gist of the proposal:
- Numerals contain a root.
- That root is embedded in functional structure.
- It is the root + functional structure that makes the numeral.

Implication: There is no universal category “numeral”; numerals are constructed categories (a la Wiltschko 2014).

This approach is in line with recent work on numerals:

Ionin & Matushansky (2018: 5-6): “Simplex cardinals are lexical rather than functional, and they do not form a special linguistic class; depending on the language, and the type of cardinal, they may be nouns or adjectives (or more rarely, verbs). At the same time, they do not necessarily exhibit all properties of nouns or adjectives, respectively.”

Fassi Fahri (2018: 60): “I assume that numerals have no pre-established “lexical” or “functional” category (be it n, a, q, d, or v). It is more plausible to think of numerals as born in the computation first as acategorial roots. Roots presumably associate first with a general numerosity sense (say n), and they are later compositionally “molded” with various categories, positions, and inflections, to derive specific senses (e.g., single terms, cardinals, ordinals, etc.).” (italics original)

The structure for the talk is as follows:
1. Numerals have an inconsistent morphosyntax cross-linguistically, resembling nouns, verbs, and adjectives but showing no apparent unified category between languages.
2. What a root + functional structure analysis of numerals would look like: semi-lexicality
3. The roots and functional structure of numerals
4. Implications for the universality of numerals
2 Numeral variation

Cross-linguistically, numerals have been found to show a resemblance to each of the major lexical categories (nouns, verbs, adjectives) (see e.g. Donohue 2005). I give examples of such variation here. I focus on cardinal numerals.

2.1 Slavic numerals: Adjectival and Nominal

Slavic numerals are famous for ranging between adjectives and nouns in their morphosyntax. References to “adjective-hood” or “noun-hood” is found repeatedly in the literature (e.g. from early work, Corbett 1978, to more recent work, Ionin and Matushansky 2018).

Diagnostics vary, but:
- Adjectival numerals show patterns of agreement (like adjectives do)
- Nominal numerals show patterns of case assignment (like nouns do)

Polish serves as a good example of this, but it’s also visible in other Slavic languages:

(1) Adjectival numeral 1 (gender, number, and case agreement)
   a. Jeden ptak, jedna dziewczyna
      One.M.SG.NOM bird.M.SG.NOM, one.F.SG.NOM girl.F.SG.NOM
      ‘One bird, one girl’
   b. Jedne drzwi
      One.NV.PL.NOM door.pluralia.tantum.NV.PL.NOM
      ‘One door’

(2) Adjectival/nominal(?) numerals 2-4 (gender & case agreement, gender-dependent case)
   a. Dwa ptaki, dwie dziewczyny
      Two.M.NOM birds.M.PL.NOM, two.F.NOM girls.F.PL.NOM
      ‘Two birds, two girls’
   b. Dwaj chłopcy, dwóch chłopców
      Two.V.NOM boys.MH.PL.NOM, two.GEN boys.MH.PL.GEN
      ‘Two boys, two boys’ (same meaning)

(3) Adjectival(?)/nominal numerals 5+ (gender & some case agreement, case assignment)
   a. Pięć ptaków, pięciu chłopców
      Five.NOM/ACC birds.GEN, five.V.GEN/ACC boys.GEN
      ‘Five birds, five boys’
   b. Z pięcioma ptakami, z pięcioma chłopcami
      With five.INST birds.INST, with five.INST boys.INST
      ‘With five birds, with five boys’

(4) Nominal numerals 1000+ (case assignment)
   a. Tysiąc ptaków, tysiąc chłopców
      1000.M.SG.NOM birds.GEN, 1000.M.SG.NOM boys.GEN
      ‘1000 birds, 1000 boys’
Similar patterns obtain in other Slavic languages, e.g. Russian (Corbett 1978, Franks 1994, Pesetsky 2013, Ionin and Matushansky 2018), Serbo-Croatian (Franks 1994; Saric 2015), Slovenian, Czech (Veselovská 2001; Caha 2013), though the details differ.

2.2 Koasati numerals: Verbal

Data here is based on Donohue (2005) taken from Kimball (1991). Koasati is a Muskogean language from the south-eastern United States. Numerals can only occur as predicative verbs.

Diagnostics for verb-hood:\(^1\)
- Participation in the switch reference system: -n indicates two predicates do not share a subject (and the connective -t is used for predicates that share a subject).
- Use of stative (or active sometimes) verbal inflection when indicating group number (e.g. we are three).

(5) Verbal numerals, switch reference marker -n (Kimball 1991: 358; Donohue 2005: 7)

   man-PL ten and two-SW see-1SG.PAST
   ‘I just saw twelve men.’

b. [(the) men (were) twelve] and [(I saw (them)]
   SUBJ1 SUBJ1≠SUBJ2 SUBJ2

(6) Verbal numerals, connective marker -t (Kimball 1991: 358)

a. Náni-ha-k pokkó:l awáh tóklo-t ca-hí:ca-toho:-li-mpa:-s
   man-PL-SUBJ ten and two-CONN 1SG.OBJ-see-REALIS-DEDUC-HRSAY-1PST
   ‘They say that twelve men have just seen me.’

b. [(the) men (were) twelve] and [(they) have just seen me]
   SUBJ1 SUBJ1=SUBJ2 SUBJ2

(7) Verbal numerals, inflection (Kimball 1991: 359)

a. Ostá:-hílk
   Four-1PL.SUBJ(3A)
   ‘We are four (by choice).’

b. Ko-ostá:k
   1PL.STAT.SUBJ-four
   ‘We are four (by chance).’

\(^1\) According to Kimball (1991: 354), all Koasati numerals are verbs, except 100 and 1000, which are nouns. No further information is given.
2.3 Austronesian numerals: Verbal and more

Data here is based on Donohue (2005), taken from Mosel and Hovdhaugen (1992) for Samoan.

Donohue presents the following table, categorizing predicative and modificational numerals in four Austronesian languages:

(8) Behavior of (some) numerals in Austronesian languages (Donohue 2005: 8)

<table>
<thead>
<tr>
<th></th>
<th>Samoan</th>
<th>Maori</th>
<th>Tukang Besi</th>
<th>Indonesian</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicative</td>
<td>as verb</td>
<td>as verb</td>
<td>as verb</td>
<td>as verb?</td>
</tr>
<tr>
<td>Modificational</td>
<td>in RC</td>
<td>(special)</td>
<td>as ‘adjective’</td>
<td>as noun?</td>
</tr>
</tbody>
</table>

(9) Samoan (GENR = general tense-aspect-mood particle)

a. E tolu tusi na mua Independent clause
   GENR three letters PAST get
   ‘(She/he) got three letters. (lit: The letters she/he got are three)’

b. Ua mavae vaiaso e tolu… “Attributive numeral clause”
   PERF pass weeks GENR three
   ‘Three weeks passed…’ (Mosel & Hovdhaugen 1992: 318)

c. …i tamaloloa [ e toa=lua ] [o lo o eli le vai=eli]²
   LOC/DIR men GENR HUM=two PROG dig ART water=dig
   ‘…to the two men who were digging a well.’ (Mosel & Hovdhaugen 1992: 319)

Diagnostic for verb-hood in Samoan: Presence of general tense-aspect-mood particle e

2.4 Inflectionless numerals: Adjectival?

Languages like English, Dutch, Norwegian, and Spanish have numerals which are mostly inflectionless (though this might differ in the predicative use or for the numeral 1).

(10) The four dogs English
(11) De vier hond-en Dutch
   The.PL four dog-PL
   ‘The four dogs’
(12) De fire hund-ene Norwegian
    DET.PL four dog-PL.DEF
    ‘The four dogs’
(13) Los cuatro perro-s Spanish
    The.M.PL four dog.M-PL
    ‘The four dogs’

Evidence for category is then often taken based on their position, leading to a categorization as adjectives, or alternatively “quantifiers” or “numerals.”

² There may be typos due to poor visibility in the original source.
Adjectives occur between the determiner and noun in English and Dutch, like numerals:

(14) The (four) brown dogs  
(15) De (vier) bruine honden  
    English  
    Dutch  
    The (four) brown-PL dog-PL  
    ‘The (four) brown dogs’

In Norwegian, adjectives trigger the need for an extra determiner in an NP-initial position (the phenomenon of double definiteness), like numerals:

(16) a. Hund-ene  
    Dog-PL.DEF  
    ‘The dogs’  
    b. De (fire) brun-e hund-ene  
    DET.PL four brown-PL dog-PL.DEF  
    ‘The (four) brown dogs’

In Spanish, numerals precede the noun while qualifying adjectives follow the noun:

(17) Los (cuatro) perro-s marrone-s  
    The.M.PL four dog.M-PL brown-PL  
    ‘The four brown dogs’

Possible analysis:
- English numerals = adjectives (A) or quantifiers (Q)
- Dutch, Norwegian numerals = adjectives without inflection (A) or quantifiers (Q)
- Spanish numerals = quantifiers (Q) or a special type of adjective (pre-nominal, without inflection) (A)

2.5 Summary

- Numerals show a wide variety of behaviors cross-linguistically, patterning with verbs, adjectives, and nouns.
- Diagnostics to determine category generally involve a comparison with the morphosyntax of existing lexical categories in the language.
- Generalization: Numerals usually share some inflectional or distributional properties with lexical categories.
- In short – numerals can be very different!

3 A semi-lexical approach to numerals

Semi-lexical is a label applied to lexical items which show a mix of lexical and functional properties. Such items are usually difficult to categorize, hence the label. Some examples:

Psuedopartitive nouns are often labeled semi-lexical (van Riemsdijk 1998; Vos 1999; Stavrou 2003; Alexiadou, Haegeman and Stavrou 2008; Hankamer and Mikkelsen 2008; Klockmann
2017; (and more!)); in English, this semi-lexicality is realized through a nominal morphosyntax that is contradicted by an invisibility for verbal agreement:

(18) A lot of books were destroyed in the fire.
(19) Lots of water was used to put out the fire.

In Corver and van Riemsdijk (2001), the label semi-lexical is applied to:

- One (N), so (A), with (P), be (V) (Schutze 2001)
- Numerals/Quantifiers+Classifiers in Bangla (Bhattacharya 2001)
- Czech numerals (Veselovská 2001)
- Light verbs in Urdu and Hindi (Butt and Geuder 2001)
- Motion verbs (I’ll go buy bread) (Cardinaletti & Starke 2001)
- Dutch and German postpositions (Zeller 2001)

Let us apply the label semi-lexical to numerals because:
(a) They often resemble lexical categories, but rarely show full identity.
(b) Their function is arguably a functional one? (quantifying)
(c) They show a mix of lexical and functional properties, by Abney’s (1987) diagnostics:
   i. open(?) class (lexical),
   ii. phonologically and morphologically independent (lexical),
   iii. take a single category as complement (?) (functional),
   iv. do not assign theta roles (?) (functional),
   v. inseparable from complement (?) (functional)

**Semi-lexicality**

**Basic set of assumptions ("the syntax"):**

Assume a syntax where lexical categories such as nouns, verbs, and adjectives correspond to roots dominated by functional structure which together form an “extended projection”.

The content of the functional structure in the extended projection differs between approaches, e.g. in Distributed Morphology, categorizing heads determine lexical category, while in exo-skeletal approaches (Borer 2005), category is epiphenomenal:

(20) $\left[ aP \ n \ [\text{ROOT}] \right] = \text{noun}$  
* Distributed Morphology

(21) $\left[ aP \ # \ [\text{ClP} \ Cl \ [\text{ROOT}] ] \right] = \text{noun}$  
* Exo-skeletal model

In terms of **derivation**, syntax builds functional structure using Merge, following some predefined functional sequence (which may or may not be universal – opinions differ).

(22) Possible functional sequence:

$$D > Q > \text{Num} > n > \text{root}$$

---

3 With accompanying analyses to explain the origins of that semi-lexicality
Note that some kind of functional sequence is necessary to capture systematic ordering relations between functional projections cross-linguistically (e.g. \( C > T > v > V \)) (see also Cinque 2005, Abels and Neeleman 2006).

A syntax for deriving semi-lexicality

Klockmann (2017): Semi-lexicality originates in the lexical specification or lexical entry of an item. The gist of the analysis is as follows:

(a) Lexical categories are specified in the lexicon purely as roots (and shades of meaning are built in the syntax via functional structure, i.e. the mass-count distinction)

(b) Functional categories are specified in the lexicon purely as syntactic features (DEF, INDEF, SG, PL, PRES, PAST, etc.); these are the components of functional structure, the functional morphemes

(c) Semi-lexical categories are specified in the lexicon with both roots and syntactic features → the syntactic features put constraints on the functional structure that can be built around the root.

Example: pluralia tantum nouns in English are obligatorily plural:

(23) My pants are ripped. (The one pair of pants I’m wearing now.)

- The root is specified for a PL feature in the lexicon: \([\sqrt{\text{pants}}, \text{PL}]_{\text{lexical_entry}}\)
- This PL feature acts as instructions for the syntax: the root pants must co-occur with a PL number head, and hence, Merge combines the root with a PL number head.
- This PL number head is what determines verbal and demonstrative agreement (plural). Syntactically (but not necessarily semantically), pants is plural.

Back to numerals:

- Assume a semi-lexical analysis for numerals.
- The lexical entry for the numeral contains both a root and syntactic features.
- The syntactic features in the lexical entry of the root determine how functional structure will be Merged.
- It is this functional structure which (a) makes the root into a numeral, and (b) can give the numeral its resemblance to other lexical categories.

4 Decomposing cardinals

Let’s look at how a decomposition of cardinal numerals would work. I will argue that numerals have (a) roots and (b) functional structure, some of which is lexically specified.

4.1 Numeral roots

What reason do we have to believe that numerals might have roots?
I give two reasons here: (a) an open class nature of numerals; (b) grammaticalization patterns.

### 4.1.1 Numerals as open class

Closed class: no new elements can be added; functional categories are closed class usually and represent only a handful of distinctions (e.g. SG/DU/PL, PAST/PRES, DEF/INDEF, etc.)

Open class: new elements can be added; lexical categories are open class usually and can easily add words with new meanings (*to google, xerox, computer, Internet* – all recent words)

Proposal: Numerals are open class. This relates to them being partly lexical, and in this theory, having roots.

Reason #1: In many languages, it is possible to add new elements to the set of numerals:

- High numerals in English: *million, billion, trillion, quadrillion, quintillion, sextillion, vigintillion, centillion, googol, googolplex,* ...
- Fictitious numerals in English: *bajillion, zillion, gazillion, gajillion*
- High and fictitious numerals in other languages?

Reason #2: In languages with less developed numeral systems (e.g. where only a few numerical values are named), the numeral system can *only* develop by adding new numerals.

- Comrie (2013)
  - Mangarayi (Australian): numerals for 1, 2, 3 only
  - Ydiny (Australian): numerals for 1, 2, 3, 4, 5 only
  - Hixkaryana (Cariban, Brazil): numerals for 1, 2, 3, 4, 5, and 10 only
  - Haruai (Piawi, New Guinea): numerals for 1 and 2 only; 3 and 4 built by concatenation of 2+1 or 2+2

Related to this: Closed class elements often represent only a handful of distinctions, but numerals can be used to refer to an infinity of values. They do so by combining *syntactically*, as has been argued extensively by Ionin and Matushansky (2006, 2018). This property is more characteristic of something which is lexical rather than functional.

In sum: the open class nature of numerals is consistent with them having a root, given that roots can be freely added to the lexicon.

### 4.1.2 Nominal origins (grammaticalization)

Grammaticalization of numeral systems according to von Mengden 2008:

- Step 1: Subitization (naming subitizeable quantities, 1-4)
- Step 2: Extending beyond subitizeable quantities (possibly using body part nouns?)
- Step 3: Serialization (introduction of complex numerals)
- Step 4: Grammaticalization (the grammaticalization of some complex numerals)
- Step 5: Adding new numeral bases to the system
Basic argument: If numeral words have their origins in non-numeral words, then we can expect there to have been some “lexical” stage to the numeral where it had a root and the syntax of the non-numeral word.

Such lexical origins are clearly visible in some languages:

(24) Hixkaryana (Comrie 2013)

<table>
<thead>
<tr>
<th>Numeral</th>
<th>Etymology?</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 kamori irakay (o) me</td>
<td>‘half of our hands’</td>
</tr>
<tr>
<td>10 kamothi tktxehkaxe ro</td>
<td>‘our hands completely’</td>
</tr>
</tbody>
</table>

(25) Hup (Naduhup, Brazil) (Epps 2005: 262, 264)

<table>
<thead>
<tr>
<th>Numeral</th>
<th>Etymology?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 ɨyɨp (TD, B) ɨɗp (UN)</td>
<td>compare demonstrative ɨyɨp ‘that (intangible)’</td>
</tr>
<tr>
<td>2 ko ɨyɨp (B) ka ɨyɨp (TD, UN) kawɛk-ɨɨp (NF?)</td>
<td>‘eye-quantity’ (body part)</td>
</tr>
<tr>
<td>3 mɨwawɨp (B) mɨnka ɨp (TD) mɨk-wig-ɨp (NF?) bab’ pâ (UN)</td>
<td>‘rubber.tree-seed-quantity’ ‘sibling NEG:EX’ (‘without sibling’)</td>
</tr>
<tr>
<td>4 hi-bab’-ni (TD, B) bab’-ni (TD, UN)</td>
<td>‘(FACT)-have.sibling/accompany.NMZ’ (deverbal) OR ‘sibling exists’</td>
</tr>
<tr>
<td>5ɨyɨp dâp (TD, B) ɨdâp (TD) nəp’tih (variant B) ɨəp d’apə (UN)</td>
<td>‘one hand’</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Numeral (and variants)</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 cûp cob cako’et ɨyɨp</td>
<td>‘other finger stands up one’</td>
</tr>
<tr>
<td>ɨyɨp cob cako’et ɨyɨp</td>
<td>‘one finger stands up’</td>
</tr>
<tr>
<td>cûp cob popo’g</td>
<td>‘other finger RED.big (=thumb)’</td>
</tr>
<tr>
<td>7-9 cûp cob cako’et (2,3,4) cob cako’et</td>
<td>‘other finger stands up (2,3,4)’</td>
</tr>
<tr>
<td>(2,3,4) cob popo’g</td>
<td>‘(2,3,4) fingers stand up’</td>
</tr>
<tr>
<td>10 cûp cob cako’et ɨdâp ɨdâp</td>
<td>‘other finger stands up five’</td>
</tr>
<tr>
<td>cako’et cob cako’et</td>
<td>‘five fingers stand up’</td>
</tr>
<tr>
<td>nûh?</td>
<td>‘finger be.finish’</td>
</tr>
<tr>
<td>d’apə nûh?</td>
<td>‘hand be.finish’</td>
</tr>
<tr>
<td>d’apə=d’əh pël</td>
<td>‘hand-PL both’</td>
</tr>
<tr>
<td>11-14 (another, 2, 3, 4) j’ib (popo’g) cako’et</td>
<td>‘… toes/foot stand up’</td>
</tr>
<tr>
<td>(Note same forms as 16-19)</td>
<td></td>
</tr>
<tr>
<td>15 ɨyɨp j’ib lu?</td>
<td>‘one foot finish’</td>
</tr>
<tr>
<td>16-19 (another, 2, 3, 4) j’ib (popo’g) cako’et</td>
<td>‘… toes/foot stand up’</td>
</tr>
<tr>
<td>(Note same as 11-14)</td>
<td></td>
</tr>
<tr>
<td>20 j’ib nûh? j’ib=d’əh pël</td>
<td>‘feet be.finish’</td>
</tr>
<tr>
<td></td>
<td>‘feet-PL both’</td>
</tr>
</tbody>
</table>
This suggests that in some languages, nouns symbolic of a cardinality (eyes, rubber tree seeds, hands) may serve as a source word for numerals: in the initial stages of the grammaticalization process, the numeral/noun would presumably contain a root.

This also already suggests an answer to the resemblance of numerals to lexical categories: if the source word is lexical, then we might expect evidence of that to remain as the numeral grammaticalizes.

For example: In Polish, there seems to be a grammaticalization from noun to “numeral” for e.g. 5, 10:

(26) Old Polish: 5, 10 as i-stem nouns triggering F.SG verbal agreement:
   a. Pięć lat minęła.
      Five.F.SG years.GEN passed.F.SG
      ‘Five years passed.’
   b. Jako minęła dziesięć lat
      As passed.F.SG ten.F.SG years.GEN
      ‘As ten years passed.’ (Dziubała-Szrejbrowska 2014: 103)

(27) 15th-16th Polish: 10 could pluralize; 5 and 10 could assign case:
   a. Z piąciądziesiąt synów
      With[INST] five-SG.INST-ten GEN PL sons GEN
      ‘With fifty sons’
   b. Z piącią synów
      with[INST] five.SG.INST sons GEN
      ‘with five sons’ (Miechowicz-Mathiasen 2012)

This suggests a previously nominal origin for the numeral (based on its morphosyntax), which has mostly grammaticalized away, but remains visible for e.g. case assignment in nom/acc.

In sum: The potentially lexical origins of numerals suggests a lexical stage, where presumably numerals quantified while still containing a root.

4.1.3 The roots of numerals

Fassi Fahri (2018): What if the different numeral senses (cardinal, ordinal, fractional, etc.) are connected via a root, with functional structure differentiating between these senses?

Do these share a root?

(28) a. two, twice, twofold, twos
   b. three, thrice, threefold, third, threes

A root+functional structural analysis of numerals can give us a handle on the cross-linguistic variation we find, and also a base for constructing different senses of numerals.
4.2 Functional structure of numerals

What makes a numeral a numeral?

Under our approach, we expect there to be minimally some piece of functional structure (FP) which turns the numeral root into a cardinal numeral, e.g.:

(29) \[ [\text{FP} \ F \ [\text{NUMERAL ROOT}]] \]

What is the identity of F? Possible solutions:

- A categorizer numeral which categorizes the root as a numeral? (but this isn’t so informative – what is a category “numeral”? and how are differences between numerals in the same language captured?)
- A piece of functional structure which corresponds to some semantic operation that converts the numeral root into the proper semantic type to act as a cardinal? (e.g. Sudo 2016; Wągiel t.a.)
- Something else?

Classifier constructions:

(30) [ichi-rin]-no hana \hspace{1cm} \text{Japanese}
    one-CL-GEN flower
    ‘one flower’ (Sudo 2016: 4)

Bale & Koon (2014): Classifiers occur for numerals, not for nouns

- Chol (Mayan language of southern Mexico) speakers use both numeral words native to Chol (1-6, 10, 20, 40, 60, 80, 100, 400) and numeral words borrowed from Spanish.
- Native Chol numerals require classifiers, while borrowed Spanish numerals do not:

(31) Numeral use in Chol
    a. ux*-((p’ej) tymuty \hspace{1cm} \text{Classifier required (native Chol numeral)}
        three-CL egg
        ‘three eggs’
    b. nuebe-(*p’ej) tyumuty \hspace{1cm} \text{Classifier not permitted (borrowed Spanish numeral)}
        nine-CL egg
        ‘nine eggs’

- They conclude that the classifiers support the numeral rather than the noun, given that it is dependent on numeral choice.

Sudo (2016): Classifiers form a constituent with the numeral and convert it from type \( n \) to the type needed to modify nouns.

Wągiel (t.a.): The head Card converts numeral roots from type \( n \) to the type needed to modify nouns; this head is invisible in English and Polish, but realized as a classifier in Japanese.
Variation: \([FP \ F [ \ Root \ ]]\) is either spelled-out together by a numeral (English, Polish, Chol Spanish-numerals) or individually as a classifier + numeral (Japanese, Chol native numerals).

The behavior of modified English cardinals can be used to argue in favor of this: insertion of an adjective with numerals triggers an obligatory indefinite article:

(32) A good twenty boxes  
(33) An estimated hundreds of thousands of dollars

If insertion of the adjective interrupts the spellout of \(F+\text{ROOT}\), we might expect some default realization of \(F\), that presumably being \(a\) (see Bylinina, Dotlačil, and Klockmann 2016 for an implementation of this; as well as Ionin & Matushanski 2018).

Question: Do all numerals in all languages carry an \(F\)? Or is it possible to “be a numeral” in other ways?

4.3 Lexically-specified numeral functional structure

How do we account for the similarity of numerals to lexical categories?

4.3.1 Nominal Polish numeral 1000

Polish numeral 1000 is morphosyntactically a noun:

(34) Paradigm of 1000 \(\text{tysiąc}\), compared to \(\text{miesiąc}\) ‘month’

<table>
<thead>
<tr>
<th></th>
<th>SG</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOM</td>
<td>tysiąc</td>
<td>tysiąc-e</td>
</tr>
<tr>
<td>ACC</td>
<td>miesiąc</td>
<td>miesiąc-e</td>
</tr>
<tr>
<td>GEN</td>
<td>tysiąc-a</td>
<td>tysiąc-y</td>
</tr>
<tr>
<td>GEN</td>
<td>miesiąc-a</td>
<td>miesiąc-y</td>
</tr>
<tr>
<td>DAT</td>
<td>tysiąc-owi</td>
<td>miesiąc-om</td>
</tr>
<tr>
<td>LOC</td>
<td>tysiąc-u</td>
<td>tysiąc-ach</td>
</tr>
<tr>
<td>LOC</td>
<td>miesiąc-u</td>
<td>miesiąc-ach</td>
</tr>
<tr>
<td>INST</td>
<td>tysiąc-em</td>
<td>tysiąc-ami</td>
</tr>
<tr>
<td>INST</td>
<td>miesiąc-em</td>
<td>miesiąc-ami</td>
</tr>
</tbody>
</table>

(35) Has a plural form which occurs in complex numerals or to indicate abundance

a. Tysiące ptaków
   Thousands.NOM/ACC birds.GEN
   ‘Thousands of birds’

b. Dwa tysiące ptaków
   Two.NOM/ACC thousands.NOM/ACC birds.GEN
   ‘Two thousand birds’
(36) Assigns case like a noun (compared to *klucz* ‘key’)

a. Tysiąc ptaków / Klucz ptaków
   Thousand.NOM birds.GEN / Key.NOM birds.GEN
   ‘A thousand birds / A flock of birds (flying in a V)’

b. Z { tysiącem ptaków / kluczem ptaków }
   WithINST { thousand.INST birds.GEN / key.INST birds.GEN }
   ‘With a thousand birds / a flock of birds (flying in a V)’

(37) Allows modification before quantification (even internal to a complex numeral)

a. Gdyby panu, nie daj Bóg, coś się stało,
   If you NEG give.IMP God something SIĘ happened.N.SG
   to z trzydziestoma dodatkowymi tysiącami …
   DEM.N.SG with[INST] thirty.INST additional.INST thousands.INST
   ‘If God forbid anything were to happen to you, (then) with additional thirty thousand [złotych (Polish currency)]’ (NKJP)

b. Tych kilkanaście nędznych tysięcy dolarów
   DEM.PL.GEN tens(11-19) miserable.PL.GEN thousands.GEN dollars.GEN
   ustawiało ich finansowo …
   set.up.N.SG them financially …
   ‘Those miserable several/many thousands of dollars ($11,000-$19,000) set them up financially.’ (NKJP)

But it optionally lacks gender which leads to two patterns of agreement:

(38) Pattern 1: Full agreement (numeral has masculine gender)

a. Cały tysiąc ptaków spali.
   Whole.M.SG thousand birds.GEN slept.M.SG
   ‘A whole thousand birds slept.’

b. Te tysiące ptaków spaly.
   DEM.NV.PL thousands birds.GEN slept.NV.PL
   ‘Those thousands of birds slept.’

(39) Pattern 2: Default agreement (numeral lacks gender)

a. Cale / calych tysiąc ptaków spalo.
   Whole.NV.PL / .GEN thousand birds.GEN slept.N.SG
   ‘A whole thousand birds slept.’

b. Te / tych tysiące ptaków spalo.
   DEM.NV.PL / .GEN thousands birds.GEN slept.N.SG
   ‘Those thousands of birds slept.’

We can assume a lexical entry along the lines of:

(40) [√tysiąc, ¬(M)]_lexical_entry

The nominal properties (plurality, adjectival modification, case assignment, agreement target) fallout from the nominal functional structure of the numeral.
Does Polish 1000 have an FP?

4.3.2 Adjectival/nominal Polish numerals

Numerals 2-4 and 5+ (5-10, 100) in Polish are both “adjectival” and “nominal”, though to differing degrees. The proposal here is that some of these properties are simply encoded in the lexical entry of the numeral, leading to “adjectival” or “nominal” effects in the syntax.

Wągiel (t.a.), for example, analyzes the 5+ numerals as follows:

(41) \[ [\text{NumeralP CARD} \ [\text{NumeralP Numeral} \ [ \text{ROOT} ] ] ]\]

Certain morphosyntactic characteristics of the numeral are captured by the Numeral head, which is realized as -u in virile contexts and -0 in non-virile contexts (virility determined by the gender/number of the noun).

(42) Pięć ptaków, pięciu chłopców
‘Five birds, five boys’ (nominative or accusative case context)

Here, presumably the need for a gender-marking Numeral with 5+ is encoded in the lexical entry of the numeral.

In Klockmann (2017), I propose that the difference between numerals 2-4 and 5+ in Polish is the presence of a gender probe on 2-4, which is optionally present in virile contexts, and never present with 5+. This gender probe is part of the lexical specification of 2-4.

(43) Non-virile contexts (nouns which are not human + masculine gender)
   Dwa ptaki
   ‘Two birds’

(44) Virile contexts (nouns which are human + masculine gender)
   Dwaj chłopcy, dwóch chłopców
   ‘Two boys, two boys’ (same meaning)

Effects of a gender probe:
- Agreement of numerals (2-4) with the noun for gender creates a unification of features between the numeral and noun, which bleeds genitive case assignment to the noun.
- This interacts with verbal agreement, allowing for successful agreement:

(45) Verbal agreement when 2-4 have a gender probe:
   a. Dwa ptaki spały.
      ‘Two birds slept.’
b. Dwaj chłopcy spali.
   Two.V.NOM boys.MH.PL.NOM slept.V.PL
   ‘Two boys slept.’

In the absence of a gender probe:
- There is no unification of features between the noun and numeral.
- As a result, they are treated as separate for the computation of case, leading to genitive assignment to the noun.
- This interacts with verbal agreement, since it bleeds verbal agreement of a possible target:

\[(46)\] Default verbal agreement:

a. Dwóch chłopców spało.
   Two GEN boys.MH.PL GEN slept.N SG
   ‘Two boys slept.’

b. Pięć ptaków / pięciu chłopców spało.
   Five NV birds GEN / five V boys GEN slept N SG
   ‘Five birds / five boys slept.’

- Something must still be said about the pięć / pięciu distinction: structurally conditioned allomorphy?

Under this approach: differences in the lexical specifications of the numerals lead to idiosyncratic differences between numerals. Furthermore, the presence of a gender probe makes some numerals look adjectival.

5 The non-universality of cardinals

The implications of the present analysis:
- There is no single category which unifies numerals cross-linguistically.
- Instead, numerals are constructed from roots and functional structure.
- The content of this functional structure can vary from language to language, likely depending on the source word for the numeral and grammaticalization patterns.
- An expectation is that while quantification and cardinality may very well be universal, numerals may not be – there might exist languages without any numerals (e.g. Pirahã).

6 References


Wagiel, M. (t.a.). Several quantifiers are different than others: Polish indefinite numerals.
