84. Denumeral categories

1. Introduction
2. Ordinal numerals
3. Fractional numerals
4. Non-strict numeral denumerals
5. Non-numeral denumerals
6. Numeral-based compounds
7. References

Abstract
This chapter describes morphologically complex lexemes that are formed on numerals in European languages. Besides ordinals and fractionals, it sheds light on less well-known categories such as approximatives, exhibitives, etc. A minimal formalisation helps us to establish the degree of complexity of the various denumeral derivations examined in the chosen sample of languages.

1. Introduction
Before embarking on the description, it is worth noting that some numerals can be used in counting and calculi (cardinal, fractional and ordinal numerals), whereas others (distributive, collective numerals, etc.) may lack such uses and appear in constructions where the numerical content they convey functions as a component of a larger description. Numerals of the first type form series without gaps, which are infinite, and their meaning is strictly number-based. Numerals of the second type constitute gapped series, which are finite, and their meaning is generally mixed with other predicates (section 4.). Expressions of the first type are usually called numbers, while those of the second type have various denominations. Following Huddleston and Pullum (2002: 1715), the term numeral will be used for linguistic expressions (five) and number for meanings (‘5’). For reasons of conveniency, I will sometimes use the term Numbers for cardinal numbers and ‘numeral n’ will abbreviate ‘numeral denoting Number n’. The abbreviation of the languages’ names follows norm ISO 639-3 (cf. http://www.ethnologue.com/language_index.asp).

The complex lexemes investigated here can be sorted out into four groups: strict numerals (ordinal, fractional numerals), which belong to numeral denumerals and exist in all the languages under examination; non-strict numerals (distributive, collective, multiplicative numerals), which constitute the second branch of numeral denumerals and do not always exist as derived lexemes; non-numeral denumerals e.g. fra quatrain ‘quatrain’; numeral-based compounds e.g. eng five-storeyed. The last two categories are not attested in all languages. As in other areas of morphology, the expression of a derivational category C, e.g. ‘Ordinal’ is equated with the application of a lexeme-formation rule (LFR) relating a base B of the appropriate sort to a derivative D, whose form ordinarily differs from that of B in that it includes some formal mark correlated to category C e.g. /θ/ in English (cf. Stump 2010: 205). I shall refer to the rules which modify the phonological representation of the base as rules of derivational exponence.

The morphological processes studied here involve LFRs which take as their input whichever numeral is appropriate, be it simple or complex. Generally, the simplest numerals are those expressing simple cardinal numbers. Simple cardinal numerals are those that cannot be formed by the morphological rules creating complex cardinal numerals in the language in question. This view leads us to consider as simple former derivation from older stages of the language which are no longer transparent nowadays, e.g. eng twelve. The way complex numerals expressing higher cardinal numbers are
formed is not our concern. These numerals are syntactic composites. Several linguists consider them as (a subtype of) compounds, but this claim is only partially true. Many are lexicalized phrases as suggested by the fact that they include a coordination e.g. nor hundre-og-en hundred and one ‘101’, a noun e.g. Eastern eus luar-etan hogoí 4-time 20 ‘80’, or a preposition e.g. war ‘on’ in bre unan war-urgent one on-twenty ‘21’. For our purpose, we only need to assume that they constitute identifiable and structured constructs, whose structure can be inspected by subsequent rules of exponence.

2. Ordinal numerals

2.1. Semantic properties

LFRs forming ordinal numerals will be considered functions applying to a numeral base and yielding a lexeme of category A. As for semantics, the interpretation of ordinal numerals is straightforward and transparent. Within the informal presentation adopted here, the function interpreting cardinals is (1a), and the corresponding one for ordinals is (1b). OrdSEM is the function, included in the LFR forming ordinals, which constructs their meaning. The N that ordinal adjectives modify gives us the sort of the series, e.g. (the) third child = ‘the child occupying rank 3 in the series of children’.

(1)  
\begin{align*}
\text{a. } & I(\text{cardinal_numeral}) = \langle \text{number} \rangle, \text{ e.g. } I(\text{three}) = 3 \\
\text{b. } & \text{OrdSEM(ordinal_numeral)} = \text{‘occupying rank } \langle \text{number} \rangle \text{ (in a given series)’ e.g. } \text{OrdSEM(third)} = \text{‘occupying rank 3’}
\end{align*}

In what follows, I will focus on the phonological changes brought about by these LFRs, that is on the way the rules of derivational exponence apply in the sample of languages investigated here. The parameters of variation according to which the ordinals can be classified are akin to those illustrated in Stump’s study. More idiosyncratic properties will be dealt with in section 2.4.

2.2. Checking the parameters

The first parameter is the nature of the exponent. In all the languages of the sample, the exponent is an affix. The second parameter is the nature of the base. In all languages but one, ordinal numerals are based on cardinal numbers e.g. lit trýlik-tas ‘thirteenth’ ← trýlika ‘thirteen’; akv ištu-liič.’ida ‘fifth’ ← ištu-da ‘five-SFX’. In Hungarian however, they are based on fractional numerals as table 1 shows (on suppletive bases cf. section 2.4).

<table>
<thead>
<tr>
<th>Cardinal</th>
<th>Fractional</th>
<th>Ordinal</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>három</td>
<td>harm-ad</td>
</tr>
<tr>
<td>10</td>
<td>tiz</td>
<td>tiz-ed</td>
</tr>
<tr>
<td>30</td>
<td>harminc</td>
<td>harminc-ad</td>
</tr>
</tbody>
</table>

Table 84.1: The bases of ordinals in Hungarian

The third parameter concerns the place of the exponent, the discussion of which is posponed until section 2.3. The mark can be suffixed or prefixed to its host, but the latter situation does not occur in European languages if we exclude Maltese, Albanian and Romanian. In Albanian prefixing the determiner (masculine i, feminine e) onto cardinals is the regular way to construct ordinals for numbers above 5 e.g. sqi pesë ‘five’ → i pestë DEF.M five.ORD ‘5th.M’, tre mijë e njëzet e pesë ‘3,025’ → e tremijenjëzetepesë ‘3,025th.F’. In a parallel way, Romanian prefixes the determiner
(masculine *al, feminine *a) and suffixes *-lea (M) / *-a (F) to form regular ordinals e.g. ron 

If the canonical situation were the rule (Corbett 2010), one would expect (i) that one
and the same exponent should be affixed to all bases, and (ii) that the stem of the base
would be the stem of the free form of the base. Condition (i) is generally met, if we
discard a handful of cases such as bre nəo-ver ‘9-ORD’ instead of the expected *nəo-ved
(cf. sei-ved 7-ORD). Dutch is a first exception however, since ordinals for tens are
formed by suffixation of *-ste e.g. veertig-ste ‘40th’ ← veertig ‘40’ whereas for all other
regular ordinals (but achtste ‘8th’) the suffix is *-de e.g. veertien-de ‘14th’ ← veertien
‘14’. The second exception is Greek, which raises more serious problems inasmuch as
the exponent of the ordinal is –tos, e.g. ell déka-tos ‘ten-ORD.M’, ekatos-tós ‘100/ORD.M’
only if we postulate allomorphic stems for all numerals between 30 and 90; moreover,
for numerals between 4 and 9,000, their first element takes exponent *-akis which differs
from the one used for the corresponding simple numeral e.g. eks-áki hiliostós ‘6,000th’
← éksi hiliádes ‘6,000’ facing ék-tos ‘6th’ ← éksi ‘6’ (cf. section 2.4. on Greek
ordinals). As for condition (ii), it is generally respected for numerals denoting numbers
higher than 10 but often disregarded for lower numbers. For the latter, in all languages
but Turkish, a suppletive form replaces the expected ordinal. This can be illustrated with
German where rules (3) yielding the suppletive forms override the default rule of
dervational exponence for ordinals, namely (2).

\[ \text{Ord}_{\text{PHON}}(X) = X-te \]
\[ \text{Ord}_{\text{PHON}}(ein) = \text{erste}, \text{Ord}_{\text{PHON}}(drei) = \text{dritte} \]

This phenomenon is widespread but the situation is contrasted. In most languages, it is
limited to the first numbers. In Greek it concerns almost all numerals denoting a simple
number and tens e.g. saránta ‘40’ → tessarákos-tós ‘40th.M’, and in Maltese ‘100’ mija
has a suppletive ordinal correlate il/mitt. This type of discrepancies is illustrated in table
2. Cardinals appear on the first row, ordinals on the second one (the lack of space
precludes us to give all inflected forms).

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>ell</td>
<td>éna / mia</td>
<td>déio</td>
<td>tria</td>
<td>téssera</td>
</tr>
<tr>
<td>prótos. i</td>
<td>déferos, -i</td>
<td>tritos, -i</td>
<td>tétartos, -i</td>
<td>ekatostós, -i</td>
</tr>
<tr>
<td>eus</td>
<td>bat</td>
<td>bi(ga)</td>
<td>hiru</td>
<td>laur(r)</td>
</tr>
<tr>
<td></td>
<td>lehen</td>
<td>bigarren</td>
<td>hirugarren</td>
<td>laugarren</td>
</tr>
<tr>
<td>fin</td>
<td>yksi</td>
<td>kaksi</td>
<td>kolme</td>
<td>neljā</td>
</tr>
<tr>
<td>ensimmäinen</td>
<td>toinen</td>
<td>kolmas</td>
<td>neljās</td>
<td>satas</td>
</tr>
<tr>
<td>rus</td>
<td>odin, odna</td>
<td>dva, dve</td>
<td>tri</td>
<td>čětyre</td>
</tr>
<tr>
<td></td>
<td>pervyj, -aja</td>
<td>vtorój, -aja</td>
<td>tretj, tret’ja</td>
<td>četvéryj, -aja</td>
</tr>
<tr>
<td>spa</td>
<td>uno, -a</td>
<td>secundo, -a</td>
<td>tres</td>
<td>cuatro</td>
</tr>
<tr>
<td></td>
<td>primero, -a</td>
<td>secundo, -a</td>
<td>tercero, -a</td>
<td>quarto, -a</td>
</tr>
<tr>
<td>swe</td>
<td>ett</td>
<td>två</td>
<td>tre</td>
<td>fyra</td>
</tr>
<tr>
<td></td>
<td>förste, -a</td>
<td>andre, -a</td>
<td>tredje</td>
<td>fjärde</td>
</tr>
<tr>
<td>tur</td>
<td>bir</td>
<td>iki</td>
<td>üç</td>
<td>dört</td>
</tr>
<tr>
<td></td>
<td>birinci, ilk</td>
<td>ikinci</td>
<td>üçüncü</td>
<td>dördüncü</td>
</tr>
</tbody>
</table>

Table 84.2: Sample of numeral suppletive forms in seven European languages

At this point, the distinction between absolute and conjunct form has to be introduced in
order to account for the behaviour of suppletive forms and their interfering with the
formation of ordinals. Absolute forms can occur as word-forms e.g. deu erst ‘first’ in das erste Wörterbuch ‘the first dictionary’ or can correspond to the stem of a word-form e.g. ita /prim/ in il primo dizionario. Conjunct forms, on the contrary, never show up as such and always occur in larger derived or compound units. In general, suppletive ordinal numerals are absolute forms. In some languages, these absolute forms are the only ones and, consequently, have also to be used in complex ordinals. This is what happens in German as table 3 illustrates.

<table>
<thead>
<tr>
<th>n</th>
<th>Cardinal</th>
<th>Ordinal</th>
<th>Suppletive</th>
<th>Ordinal 100+n</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>eins</td>
<td>*einste</td>
<td>erste</td>
<td>*hunderteinste hunderterste</td>
</tr>
<tr>
<td>2</td>
<td>zwei</td>
<td>zweite</td>
<td></td>
<td>hundertzweite</td>
</tr>
<tr>
<td>3</td>
<td>drei</td>
<td>*dreite</td>
<td>dritte</td>
<td>*hundertdreite hundertdritte</td>
</tr>
<tr>
<td>4</td>
<td>vier</td>
<td>vierte</td>
<td></td>
<td>hundertvierte</td>
</tr>
<tr>
<td>5</td>
<td>fünf</td>
<td>fünfte</td>
<td></td>
<td>hundertfünfte</td>
</tr>
</tbody>
</table>

Table 84.3: Ordinals and suppletion in German (pattern A)

But a language may also have regularly constructed ordinal numerals with the status of conjunct. These forms never appear as such in discourse but are enlisted in the formation of complex ordinal numerals. Hungarian is a case at hand, as table 4 shows.

<table>
<thead>
<tr>
<th>n</th>
<th>Cardinal</th>
<th>Fraction.</th>
<th>Ordinal</th>
<th>Suppl.</th>
<th>Ordinal 100+n</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>egy</td>
<td>—</td>
<td>*egyedik első</td>
<td>százegyedik *százelső</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>két, kettő</td>
<td>ketted</td>
<td>*kettedik masodik</td>
<td>százkettedik *százmásodik</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>három</td>
<td>harmad</td>
<td>harmadik</td>
<td>százharmadik</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>négy</td>
<td>negyed</td>
<td>negyedik</td>
<td>száznegyedik</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>öt</td>
<td>ötöd</td>
<td>ötödik</td>
<td>százötdik</td>
<td></td>
</tr>
</tbody>
</table>

Table 84.4: Ordinals and suppletion in Hungarian (pattern B)

Note that the ungrammatical forms are in complementary distribution when we change pattern. These patterns cut across the Romance and Uralic families, while preserving the others: Pattern A: Germanic, Fennic, Slavic and Baltic languages, Spanish, Portuguese. Pattern B: Hungarian, French, Italian, Romanian, Albanian, Basque, Celtic languages. Let’s examine now the place of the exponent parameter.

2.3. Formal account

Provided the rule of derivational exponence applies to a complex base with n elements i.e. $\text{Ord}_{PHON}([X_1,\ldots, X_n])$, it can be concluded for sure that the mark is internal to the stem in only two cases: when it is suffixed to the first element (4a), and when the exponence is an absolute suppletive form (4b). In cases like (4c), on the contrary, it is impossible to decide whether the marking is external or internal.

(4) a. X-sfx Y bre daou-ved warn-ugen 2-ORD on-20 ‘22nd’ ← daou warn-ugen 2 on-20 ‘22’

b. X Y\$sfx eng hundred first hundred one\$ORD ← hundred one

c. X Y-sfx eus hirur-ogei-garren 3.20-ORD ‘60th’ ← hirur-ogei 3.20 ‘60’

Germanic languages follow pattern (4b), while French, Italian, Basque, Turkish conform to (4c). Celtic languages conform to (4a), with some variations e.g. bre eil warn-ugen or even eil-ved warn-ugen are possible forms for ‘22nd’ (Trépos 1994: 164) cf. cym ail ar hugain ‘22nd’, where eil, ail = 2\$ORD.
Up to now, the category of Ordinal has been realized by one exponent only. But cases of multiple-exponence exist in Finnish, Estonian and Iberian Romance languages. This leads us to introduce a parameter ‘Extension’ into our description, the values of which are {unique, iterated}. On its turn, iteration can be unrestricted or restricted, the restricting factors usually being operation sensitivity and rating. Unrestricted ordinal marking is illustrated by Finnish (5) and Greek (6) (pace Stump 2010).

(5)  \[ \text{Ord}_{\text{PHON}}([2.10^2 + 7.10 + 5]) = [2^{\text{ORD}}.10^2 \text{ORD} + 7 \text{ORD}.10 \text{ORD} + 5 \text{ORD}] \]
Cardinal: kaksi.sataa seitsemän.kymmentä viisi (275)
Ordinal: kahde's.sada's seitsemä's.kymmen'e viide's

(6)  \[ \text{Ord}_{\text{PHON}}([7.10^3 + 8.10^2 + 6.10 + 5]) = [7^{\text{ORD}}.10^3 \text{ORD} 8.10^2 \text{ORD} 6 \text{ORD} 5 \text{ORD}] \]
Cardinal: eptá hiliãdes oktakösia eksínta pénte (7,865)
Ordinal: eptá-kis hilios'tós oktakosios'tós eksíkos'tós pemp'tos

Restricted ordinal marking occurs when the derivational rule of exponence applies only to elements of the complex numeral linked by the operation of addition, and not to those linked by the operation of multiplication (or the other way around). A case combining operation sensitivity and rating is provided by Czech. As illustrated in (7), the exponent is obligatorily suffixed to operands of addition if they denote a multiple of 10 under 100. Higher operands may be left unsuffixed. Polish has the latter option only e.g. 5,237th = pięć tysięcy trzysta trzydzieści siódmy (for masculine) ← pięć tysięcy trzysta trzydzieści siedem.

(7)  \[ \text{Ord}_{\text{PHON}}([2.10^3 + 3.10^2 + 8.10 + 5]) = [2.10^3(\text{ORD}) + 3.10^2(\text{ORD}) + 8.10-\text{ORD} + 5-\text{ORD}] \]
Cardinal: dva tisíce tři sta osmdesát pět (2,385)
Ordinal: dvoutisící třístý osmdesátý pátý or dva tisíce tři sta osmdesátý pátý

Capitalizing on Stump’s study, I now make explicit the rules we need to generate the various types of ordinals we observe in European languages. Turkish is the simplest system since a unique rule of exponence applies to simple as well as to complex bases (Göksel and Kerslake 2005: 205). The distribution of ilk partially overlaps that of birinci, but ilk never occurs in complex numerals. Birinci supposes that a list will follow, a supposition not shared by ilk which implies the lack of any preceding instances e.g. (ilk | birinci) gözlem-imiz NUM observation-1PL.POSS ‘our first observation’ vs. ‘our first observation (in a list of observations)’, also The first man on the Moon: ilk ok, birinci out. Ilk then cannot be considered a suppletive form of birinci and does not override (8a).

(8)  Type I. (partial account)
   a. Simple ordinals: \( \text{Ord}_{\text{PHON}}(X) = X-(l)ncI \)
   b. \( \text{Ord}_{\text{PHON}}(\text{bir}) = \text{ilk} \)
   c. Complex ordinals: \( \text{Ord}_{\text{PHON}}([X_1, \ldots X_n]) = [X_1, \ldots \text{Ord}_{\text{PHON}}(X_n)] \)
   d. Override: (8c) >> (8a)

Type II is slightly more complex, insofar as it involves absolute overriding forms. German (and other Germanic languages or Basque) illustrates this type.

(9)  Type II.
The next type is more complicated for what regards suppletion, since it involves both conjunct and absolute forms. Hungarian illustrates this type.

(10) Type III.
   a. Simple ordinals: \( \text{Ord}_{\text{PHON}}(X) = X\text{-te} \)
   b. \( \text{Ord}_{\text{PHON}}(\text{ein}) = \text{erste}, \text{Ord}_{\text{PHON}}(\text{drei}) = \text{dritte} \)
   c. Complex ordinals: \( \text{Ord}_{\text{PHON}}([X_1, \ldots, X_n]) = [X_1, \ldots, \text{Ord}_{\text{PHON}}(X_n)] \)
   d. Overrides: \( (9c) >> (9a); (9b) >> (9a) \)

The next types concern extended marking. Type IV deals with a case of total extended marking, that of Finnish (cf. section 2.4.). Cases of partial extended marking are subcases of type IV. If we leave aside the conditions governing the nature of the exponent and the absolute suppletives, for cardinal numerals with hundreds or higher addends, Czech has to stipulate either \( \text{Ord}_{\text{PHON}}([X_1, \ldots, X_{n-1} + X_n]) = [\text{Ord}_{\text{PHON}}(X_1), \ldots, \text{Ord}_{\text{PHON}}(X_n)] \) or \([X_1, \ldots, \text{Ord}_{\text{PHON}}(X_{n-1}), \text{Ord}_{\text{PHON}}(X_n)] \). Only the latter condition is allowed in Polish.

(11) Type IV. (partial account)
   a. Simple ordinals: \( \text{Ord}_{\text{PHON}}(X) = X\text{-n} \)
   b. \( \text{Ord}_{\text{PHON}}(\text{ys}i) = \text{ensimmäinen}, \text{Ord}_{\text{PHON}}(\text{kaksi}) = \text{toinen} \)
   c. Complex ordinals: \( \text{Ord}_{\text{PHON}}([X_1, \ldots, X_n]) = [\text{Ord}_{\text{PHON}}(X_1), \ldots, \text{Ord}_{\text{PHON}}(X_n)] \)
   d. Overrides: \( (11c) >> (11a); (11b) >> (11a) \)

The last type is that of Celtic languages, where the structure of the complex numeral determines the place of the exponent. In structure \([X_1 \text{ MRK } X_n] \), where MRK notes a set of elements that the exponence rule is sensitive to (preposition ‘war’ ‘on’ and conjunction ‘ha ‘and’), the exponent is suffixed to \( X_1 \) cf. (4a). In other complex structures, it is suffixed on the last element e.g. \( \text{an daou.ugent-ved 2.20-ORD} \text{ ‘40th} \) (Stump 2010). French ordinals present a mix of type I and III, since ‘deuxième ‘2\textsuperscript{nd}, is both absolute and conjunct and is not overriden by the absolute suppletive second.

2.4. Idiosyncratic cases

The mechanisms conceived of up to now are too simplistic to account for stem allomorphy in certain languages. In Finnish, this phenomenon is pervasive and also occurs in inflection e.g. kieli ‘language’ GEN = kiele-n, PART = kiel-tä. On the model of nouns, cardinal numerals have four stems, which may be distinct e.g. ST0 (basic form) kuusi ‘6’, ST1 (inflectional stem) kuute-en ‘6-poss’, ST2 (inflectional stem) kuude-n ‘6-gen’, ST3 (partitive stem) kuut-ta ‘6-part’, or not e.g. neljä ‘4’ (Karlsson 1987: 115). When they are declined, cardinals take the stem that the case requires e.g. kuude-ssa- toista maa-ssa ‘in sixteen countries’ cf. kuusistoista ‘16’. Ordinal numerals also have four stems the form of which results from regular alternations affecting the ending: /Xs/, /Xne/, /Xne/, /Xu/. So for ‘3\textsuperscript{rd}, we have kolmas (ST0), kolmante (ST1), kolmanne (ST2) and kolmat (ST3), for ‘6\textsuperscript{th} kudes, kuudente, kuudenne, kuudet, etc. When
declined, the ordinal selects the appropriate stem e.g. *joulukuun kahtende-na kymmenente-nā kuudente-na päivä-nā december-gen 26th-suieess day-suieess ‘on the 26th of December’. Note that suppletive forms (11b) may occur as conjunct but only as the rightmost element of complex numerals e.g. ‘202nd, (kahdes | *toinen)sadas. (kahdes | toinen). To cope with the fact that, e.g. kuudestoista is the ordinal corresponding to kuusitoista ‘16’, rule (11) has to be modified and completed taking into account conditions (12), which basically specify that the default stem for complex ordinals is the cardinal’s Stem 2. Hence, Ord<sub>PHON</sub>(kuusitoista) = Ord<sub>PHON</sub>(kuude) ⊕ toista = kuudestoista.

(12) a. Simple ordinals: Ord<sub>PHON</sub>(X) = Ord<sub>PHON</sub>(X<sub>ST2</sub>)
   aa. Ord<sub>PHON</sub>(X<sub>ST2</sub>) = X<sub>ST2</sub>-’s
   c. Complex ordinals: Ord<sub>PHON</sub>([X<sub>1</sub>,…X<sub>n</sub>]) = [Ord<sub>PHON</sub>(X<sub>1</sub>),…, Ord<sub>PHON</sub>(X<sub>n</sub>-1), X<sub>n</sub>] if X<sub>n</sub> = toista, else = [Ord<sub>PHON</sub>(X<sub>1</sub>),…, Ord<sub>PHON</sub>(X<sub>n</sub>)]

Many Slavic languages also show stem alternation. Cardinal numbers decline on the model of As for the first four Numbers and of Ns for the others, following a well-known typological cline. With higher numerals, a lot of idiosyncrasies crop up. Ordinals of tens are formed on the cardinal’s stem e.g. pol pięćdziesiąt ‘50’ → pięćdziesiąty ‘50<sup>th</sup>’ (but dwudziesty ‘20<sup>th</sup>’ ← dwudziesiąta ‘20’), ces padesát → padesátý ‘50<sup>th</sup>’, srp pedeset’ → pedeset’ ‘50<sup>th</sup>’, rus pjaťdesiat’ → pjaťdesiatyj ‘50<sup>th</sup>’. In rus, but not in ces nor srp, a similar phenomenon takes place for hundred and thousands: the stem of the unit looks like a genitive form (so traditional grammars say) e.g. rus triста ‘3,10<sup>st</sup>’ → trëx.sot-yy 3ŽGEN-10<sup>th</sup>ŽGEN-ÁZRR ‘300<sup>th</sup>’. In Polish it occurs when sto ‘100’ is involved e.g. pol sto tysiąc ‘100.10<sup>th</sup>’ → stu.tysieczny 100ŽGEN-10<sup>th</sup>ŽÁZRR ‘100,000<sup>th</sup>’ but cztery tysiące ‘4.10<sup>th</sup>’ → czter-o-tysieczny 4-RFX-10<sup>th</sup>ŽÁZRR ‘4,000<sup>th</sup>’. In any case, forms such as rus trëx, četyrëx, etc. are instances of inherent inflection. It is then less problematic to say that they are morphemes.

Postulating allomorphic stems for Greek ordinals belonging to the series of tens and hundreds seems to be the best way to handle the variations we observe. For the tens, the alternation can be formulated as (i) /Xnţa/-/Xkos/ e.g. trianta ‘30’ → triakos/tós ‘30<sup>th</sup>’, but we have to resort to suppletion for saranta ‘40’ / tessarakotós ‘40<sup>th</sup>’. For hundreds, the alternation is completely regular (ii) /Xsia/-/Xsios/ e.g. tetrakósia ‘400’ → tretrakosis/tós ‘400<sup>th</sup>’. The last alternation regards thousands: the form of their first element is that of the multiplicative numeral (section 4.2.) when it ranges from 4 to 9 e.g. pent-áki instead of ordinal pëmptos for pente ‘5’ (section 2.2.). We need to stipulate that the multiplicative form is used instead of the ordinal in this context.

In languages with extended exponention, it happens that the ordinal exponent does not appear on all elements in long ordinals. In addition to Czech, in Finnish long ordinals may occur with the exponent on the last element only e.g. kolmetuhatta satat kolmekymmentäjänne-n ‘3,134<sup>th</sup>-GEN’ (Karlsson 1987: 119), but such forms are considered ungrammatical by many speakers. This practice is the rule in Spanish and Portuguese, at least for long ordinals. A form such as por décimo milésimo seiscentésimos quinquagésimo quarto ‘10,654<sup>th</sup>’ is extremely formal and because people know only a handful of simple ordinals (all inherited from Latin), they shift to the cardinal e.g. spa el veinticinco aniversario ‘the 25<sup>th</sup> birthday’ instead of el vigésimo quinto aniversario.

The ordinal exponent may be affixed to non-numeral bases, mainly (interrogative) words whose meaning has to do with ranking e.g. hun hännya-dik ‘how-ORD’, fra combien(t)-ième ‘how-ORD’, quel-ième ‘which-ORD’, n-ième ‘n<sup>th</sup>’ C’est la
combiennième inondation dans la région? ‘Which number flood is this in the region?’ (Fradin and Saulnier 2009).

3. Fractional numerals

Fractional numerals are two-part expressions, insofar as fractional numbers involve a numerator and a denominator. While the numerator is always a cardinal numeral, the denominator, which is the part that interests us here, is constructed either on a cardinal or an ordinal. The first option is chosen in Hungarian (X → X-Ad cf. table 84-1) e.g. négy öt-öd ‘4/5’, Basque (X → X-en) e.g. zazipi-r-en bat ‘1/7’ ← zazipi ‘7’, German (X → X-tel) e.g. ein Zehn-tel ‘1/10’ ← Zehn ‘10’, Czech (X → X-ina) e.g. jedna sedm-ina ‘1/7’ ← sedm ‘7’, Welsh (X → X-fed) e.g. tri wythfed ‘3/8’ ← wyth ‘8’. The second option is the norm in most languages as shown by the numerals expressing ‘1/7’ and ‘5/10’ in the following sample: ell én-a évdom-o one.F seventh-NEU, pénê dékat-a five tenth-PL; pol jedna siódm-a one.F seventh-F, pięć dziesiątych five tenth.F.GEN.PL; lit vieną septint-oji one.F seven-ORD.F, peiktinis dešimt-osiūs five ten-ORD.F.PL; swe en sjunde-del one seven-ORD, fem tiondel-ar five ten-ORD.PL; nld een zeven-de one seven-ORD, vijf den-de five ten-ORD; fra un septième one seven-ORD, cinq dixièmes five ten-ORD.PL; ron o septime ‘one seventh’, cinci zecim-i ‘five tenth-PL’; sqi një e shtat-a ‘one seventh’.F, pesë të dhjet-at ‘five tenth-PL’. When the first ordinal numerals are suppletive forms the latter are frequently used to build fractional numerals e.g. ita un quint-o one fifth.M ‘1/5’, tre quart-i three fourth-PL ‘3/4’. But in many languages, the first fractional numerals have specific suppletive forms for ‘½’ e.g. hun fél, eng, ndl half, swe halv, deu Halb, sqi gjysmë, por meio, metad, fra demi, ita metà, tur yarum, rus, srp polovina, mlt nofs, cym haner, eus erdi; for ‘1/3’ rus tret’, cym traean, eus heren, por terço, ita terzo, fra tiers, mlt terz; for ¼ rus çetvert’, eus laurden, fra quart, ita quarto, mlt kwart, sqi çerek. Higher denominators are usually regular numerals, except eus bortzen ‘1/5’. The Slavic languages also have a special numeral for ‘1 ½’ e.g. rus poltora, pol półtora, etc.

The phrases expressing fractional numerals vary from language to language. The following patterns have been observed: (a) CARD FRAC with no variation at all e.g. deu ein Zwanzigs-tel ‘1/20’, dreizehn Zwanzigs-tel ‘13/20’. (b) CARD FRAC:FLX where the two elements agree since FRAC is a plain N governed by CARD e.g. ces jedn-a sedm-ina ‘1/7’, třináct dvacet-in ‘13/20’. (c) CARD ORD:FLX, the cardinal governs the ordinal (a N or an A) which must be inflected e.g. por un terço ‘1/3’, cinco décimo-s ‘5/10’; fra un dixième ‘1/10’, six dixièmes ‘6/10’. (d) The fractional numeral is an NP explicitly headed by the noun part, or its equivalents, which regularly agrees with the ordinal DET (CARD) ORD:FLX part e.g. spa la quinta parte ‘1/5’, las siete doceava-s parte-s ‘7/12’. (e) The fractional numeral is an NP whose head (= part) has been elided but whose constituents nevertheless agree according to the rules in force with numerals CARD:FLX ORD:FLX (part) e.g. rus odn-a pjavaja (čast’) one-F.SG five\AZR.F.SG (part[.NOM.SG]) 1/5’, pjav ’šest’ýx (čast-ef’) five six\AZR.GEN.PL (part-F.GEN.PL) 5/6’. (f) CARD-FLX CARD, the relation between the numerals expressing the numerator and the denominator is marked by a case e.g. tur üç-te iki three-INESS two ‘2/3’. Structure CARD böllö CARD e.g. iki böllö beş 2 on 5 ‘2/5’ is also available.

4. Non-strict numeral denumeral

4.1. Derivation of collectives

Whereas cardinal numerals denote combinations of individual entities, collective numerals denote combinations of groups of entities (Ojeda 1997). Hence constrasts such as isl tveir sokkar ‘two socks’ (cardinal) vs. tvennir sokkar ‘two pairs of socks’
(collective). In addition to counting groups of individuals, collectives may be used to count individuals belonging to the same group (or kind) e.g. srp sedmoro děčeř ‘7.COLL children’. The semantic variation observed with collectives is tied to the way the group of entities is constituted. Besides Icelandic e.g. einn / enir, tveir / tvennir, prír / prennir, fjörir / fernir, collective numerals are found in the Balto-Slavic languages. The derivational nature of collective numerals can be ascertained only if they form long enough series and present a recurrent pattern of affixation. This is the case in Polish, and in Serbian e.g. srp dvoje, troje, četvoro, petoro, šestoro, sedmoro, osmoro, devetoro, desetoro, pedesetoro...; pol dwoje, troje, czworo, pięcioro, sześcioro, siedmioro, ośmioro, dziewiörö, jedenaśćoro, dwanaśćoro... dziesięćoro, dwadzieśćioro, trzydzieńioro, etc. In Russian and Czech, they seem limited to the first ten numbers. It is generally assumed that collective numerals used alone imply the notion of ‘group’ e.g. srp devedesetoro ‘group of 90 people’. It is not so when they modify a N, since their use is almost always grammatically constrained: they must be chosen when the N is a pluralia tantum e.g. rus troje san’ej ‘3:COLL sledge:GEN.PL’ vs. *tri san’ej, ‘3:CARD ~’ or when it is a neuter N denoting humans e.g. pol Widęc pięcioro dzieci vs. *Widzę pięć dzieci ‘I see (five.COLL | **five.CARD) children’. On the basis of such cases, Saloni argues that collective and cardinal numerals constitute one and the same paradigm in Polish since the distribution of the former is strictly conditioned by the (sub-)gender of the N they apply to (Saloni 2010). This view is supported by the fact that collective numerals in Polish, do occur with complex numbers, in contradistinction to Russian, e.g. pol pięćdziesięci-o rzow-o dzieci ‘54 children’ (with exponent on both addends). In Serbian, however, collective numerals are restricted neither to human, nor to pluralia tantum e.g. srp pet-o ro porta ‘5 dogs’. Lithuanian has two series of collective numerals. Series (i) vieneri, dveji, trej, ketveri, penkieri, šešieri, septynieri, aštunieri, devynieri (from 1 to 9) is used with pluralia tantum and pairs, while series (ii) dvejetas, trejtas, kėtvertas, peňketas, šešetas, septynetas, aštuonetas, devynetas is used for groups.

The above mentioned numerals in Polish and Serbian convey the additional meaning that the NPs they occur in refer to sexually mixed groups of persons. In Bulgarian, collectives ending in -mina have the same effect or exclude a female only referent. Serbian also has a series of collective (feminine) nouns, regularly formed by suffixing -ica onto the collective numeral of the first series e.g. srp dvoj-ica ← dvoje ‘2.COLL’, četvrenaštór-ica ← četrnaestaoro ‘14.COLL’, which is used only when the collective number refers to male individuals. These nouns yield an agreement mismatch: their DET is marked feminine singular while the number of their predicate is singular or plural e.g. srp su petor-ica (doš-la / doš-li) (Meillet and Vaillant 1952: 128) all-F.SG be:PRS.3PL 5.COLL-NZR.F.SG (come-F.SG | come-M.PL) ‘all five (men) came’.

4.2. Derivation of multiplicatives

In most languages, the semantic domain of multiplicatives is temporal and their meaning can be roughly represented as (i) Multį(cardinal_numeral) = ‘<number> times’. They are generally derived from (a special stem of) cardinal numerals and their distribution is that of adverbials. Vestigial in English e.g. twice, thrice, multiplicatives exist in other Germanic languages e.g. deu (X → X-mal) ein-mal ‘once’, zehn-mal ‘10 times’, nld (X → X-mail or X → X-voud) drie-mail / drie-voud ‘thrice’, vier-maal / vier-voud ‘4 times’. We also find them in the following languages: ell (X → X-ákis) e.g. pent-ákis ‘5 times’ (slightly old-fashioned), lit (X → X-kart) e.g. kėturiš-kart ‘4 times’, hun (X → X-šorz) e.g. öven-szer ‘50 times’, cym (X → X-waith) e.g. deng-waith ‘10 times’, akv (X → Xče) e.g. k’e-če ‘2 times’, ces (X → X-krát) e.g. tří-krát ‘3
times’, pol (X_{COLL} → X_{COLL}-krotnie, †X-kroć) e.g. trzy-krotnie ‘3 times’ (cf. also wielokrotnie ‘many times’, często-kroć ‘frequently’). The latter forms are semantically related to derived adjectives (X_{COLL} → X_{COLL}-krotny) e.g. pol pięcio-krotny ‘repeated 5 times’. From five onward, these As are formed on the short stem of collective numerals. Some languages have spatial multiplicatives, the meaning of which can be expressed as

(ii) Mult_{CARD}(X) = ‘has <number> similar parts’. Basque illustrates this case (X → X-koit) e.g. hiru-koit as in arrazoinamendu hiru-koit-a reasoning 3-MULT-DEF ‘a reasoning in three parts’. Polish has two processes of this type, one derives adjectives by circumfixation onto the stem of collective numerals (X_{COLL} → po-X_{COLL}-ny) e.g. pol po-czwór-ny ‘with 4 parts’ potrójna porcja ‘triple portion’, the other by suffixation onto the same stem (X_{COLL} → X_{COLL}-aki) e.g. dwoj-aki ‘twofold’, czwor-aki ‘fourfold’, etc.

4.3. Derivation of distributives
Derived distributives exist in Basque and Akhvakh. Their semantics corresponds to (i)
Distr(X) = ‘<number> apiece’ e.g. eus hiru-na ogi three-DISTR bread ‘three (loaves of) bread apiece’ (Hualde and Ortiz De Urbina 2003: 128). In Basque the rule of exposition suffices /na/ to the cardinal’s stem (X → X-na) e.g. hiru-na ← hiru ‘3’ and these forms are DET. In Akhvakh, it involves the reduplication of the cardinal’s stem e.g. akv išt’amendu hiru ‘(give) five apiece (to each)’ ← ištu-da ‘5’.

5. Non-numeral denumerals
5.1. Approximative denumerals
Approximatives are attested in Romance and some Slavic languages. As expected, they are mainly based on cardinals corresponding to tens. Their meaning can be uniformly represented as (i) Appr(X) = ‘<number> ± n’, where n varies in function of the value of ‘number’ e.g. fra une quinzaine de voitures ‘about 15 cars’ i.e. ‘15±2’ vs. une centaine de voitures ‘about 100 cars (±10). In Czech (X → X-ak), they are used to refer only to humans e.g. dvacítka ← dvacet ‘20’, but not in Serbian (X → X-ak) e.g. deset-ak ← deset ‘10’, dvadeset-ak ← dvadeset ‘20’, stotín-ak ← sto ‘100’ petnaest-ak kolača ‘about 15 cakes’. In French (X → X-aine), the series is longer e.g. huit-aine ← huit ‘8’, douze-aine ← douze ‘12’, quinze-aine ← cinquante ‘50’, nonante-aine ← nonante ‘90’ (in Belgium and Switzerland), etc. In non-standard French, approximatives can be formed on many more cardinals e.g. dix-sept-aine ← dix-sept ‘17’, quatre-vingt-aine ← quatre-vingt ‘80’, cinq cent-aine ← cinq cent ‘500’, etc. (Saulnier 2010). Such wealth of forms is observed neither in Ibero-romance, nor in Italian e.g. spa quince-ena ← quince ‘15’, veinte-ena ← veinte ‘20’, cent‘ena | cent/ener ← ciento ‘100’; ita dec-ina ← dieci ‘10’, quindici-ina ← quindici ‘15’, vent-ina ← venti ‘20’, cent-in-aino ← cento ‘100’.

5.2. Expositive denumerals
Exhibitives are derived nouns denoting an entity somehow explicitly correlated with a particular Number. In Slavic languages, names of playing cards are formed on (the stem of the collective) numeral corresponding to the value of the card e.g. bul dvoj-ka karо ‘2 of diamonds’, osmor-ka kupa ‘8 of hearts’, pol trój-ka pik ‘3 of spades’, siódmor-ka trefl ‘7 of club’. The same forms are used for the marks given to pupils at school e.g. bul ūčiteljat mu piša dvoj-ka po matematika ‘the teacher gave him two in mathematics’; pol dostalem czwór-ke ‘I have got a 4’; also deu (Austria) Eins-er, Zwei-er, Drei-er, Vier-er, Fünf-er ‘1’, ‘2’ etc. and ces dvoj-ka, šest-ka ‘2’, ‘6’ etc. The latter also name trams e.g. ces jednička, třináctka ‘tram 1, 30’. In Russian exhibitives are used for cards
and for grades up to 10. In Serbian, pupils of the n\textsuperscript{th} class can be denoted by an
exhibitive N formed on the corresponding ordinal numeral e.g. srp drug-ak, trec-ak
‘pupil of the 2\textsuperscript{nd}, 3\textsuperscript{rd} class’. Finnish developed special forms for school grades e.g.
ysi ‘9’, kymppi ‘10’. These forms also denote entities for which the corresponding
number is a distinctive feature e.g. sata kuutonen ‘bus 106’, kymppi ‘10 euros bill’.

5.3. Appellative denumerals
A crucial property of some entities is their being composed of n parts, repeated n times,
etc. Appellative denumerals precisely provide the entity a name based on the numeral
denoting number n. Groups of people, animals engaged in a common activity are
frequently denoted by appellatives: teams e.g. rus troj-ka ‘(sledge) drawn by 3 horses’;
music groups e.g. hun vonós négy-es ‘string quartet’, fuvós öt-ős ‘wind quintet’, etc.
both derived from cardinal numerals (X → X-AS), ita duetto ‘duet’, quintetto ‘quintet’, etc.
(← ORD-DIM). Activities or artefacts involving n parts, measures, etc. also fall in the
realm of appellatives e.g. sqi nénëshe ‘series of 9 prayers’, fra siz-ain ‘poem of 6
verses’, vingt-deux-ain ‘cloth with 22 hundreds of thread’, neuv-aínes ‘pious exercice
made during 9 days’, ita terz-ina ‘tercet’, quart-ina ‘poem of 4 verses’, srp sedmer-ac
‘free-throw shot from 7 meters (handball)’. In French suffixing -aire onto the cardinal’s
learned stem is a way to form adjectives expressing the age e.g. quaranten'aire
fourty\textsuperscript{th} year old’, which coexists with the older form based on loan translations
from Latin e.g. quadrangénaire ‘40 year old’ < lat quadragenarius. By metonymy,
these forms also denote the person with the corresponding age (Fradin and Saulnier 2009:
216-220).

6. Numeral-based compounds
In all language families but Romance, compounding with numerals is a common way to
form adjectives indicating that the noun they modify possesses what the base-N’s
referent denotes in n examples. The general pattern looks like (i) NUM-N-AZR e.g. deu
fünf-tür-ig five-door-AZR ‘with 5 doors’. The most frequent domain denoted by N are:
age or time e.g. hun négy-év-es four-year-AZR ‘4 years old’, pol dwu-let-ni, rus dvux-let-
njy, nld twee- iar-ig two-year-AZR ‘2 year old’, fin kolme-vuot-ias three-year-AZR ‘3 year
old’, nld drie-daa-ag ‘3 day long’, part of a functional whole e.g. hun négy-lab-ú four-
leg-AZR ‘quadruped’, három-árbóc-os hajó three-mast-AZR ‘three-masted boat’, pol
cztero-list'ný four-leaf\textsuperscript{th} AZR ‘four-leaved’, fin kymmen-ikka-na-inen ten-window-AZR
‘with 10 windows’, ell exa-mel-is six-member-AZR ‘with 6 members’; others : rus dvu-
ja-zyc'nyj two-language\textsuperscript{th} AZR ‘bilingual’, dvoe-muz-ie 2.COLL-husband-NZR ‘biandry’,
kuusi-lapsi-inen six-child-AZR ‘with 6 children’. Derived As also exist e.g. Fr. bis’ann-
ue-el 2\textsuperscript{nd} year-RFX-AZR ‘lasting 2 years’. Some of these compounds have been lexicalized
as nouns e.g. ell exá-psalm-os six-psalm-AZR ‘hymn of 6 psalms’, srp tro/meda
three|border ‘3 border (point)’, spa quince-añ-er-a 15-year-NZR-F ‘15 year old girl’.

Acknowledgements
I gratefully thank my informants, Astrid Alexander-Bakkerus, Peter Arkadiev, Michel
Aurnague, Gabriela Bilbie, Denis Creissels, Hans-Ohaf Enger, Mats Forsgren, Žsuzsana
Géseg, Teresa Giermak-Zielnińska, Asli Göksel, Anna Mańkowska, Alexandru
Mardale, Dragomir Milošević, Kaarina Pitkänen-Heikkilä, Tomorr Plangarica, Svetlana
Sokolova, Dejan Stosic, Jana Strnadová, Sophie Vassilaki, Madeleine Voga.

Languages investigated
akv = Akhvakh lit = Lituanian
bre = Breton mlt = Maltese
bul = Bulgarian nld = Dutch
ces = Czech nor = Norwegian
cym = Welsh pol = Polish
deu = German por = Portuguese
ell = Greek ron = Romanian
ing = English rus = Russian
eus = Basque spa = Spanish
fin = Finnish sqi = Albanian
fra = French srp = Serbian
hun = Hungarian swe = Swedish
isl = Icelandic tur = Turkish
ita = Italian

7. References


Bernard Fradin, Paris (France)