Chapter 13

Deconstructing base numerals: English and Polish 10, 100, and 1000

Heidi Klockmann

University of Agder

Base numerals differ from other simplex numerals in that they license mathematical operations like multiplication and addition. This paper investigates the syntactic status of base numerals in two languages, Polish and English, focusing on three numerals: 10, 100, and 1000. It concludes that these numerals instantiate three types of bases, nominal bases, syntactic bases, and lexicalized bases. A nominal base is a noun used as a base, as is the case with Polish 1000. A syntactic base involves the use of a morpheme to create basehood, as is proposed for English 100 and 1000. Finally, lexicalized bases, English 10 and Polish 10 and 100, are the result of grammaticalization, i.e. the reduction of a numeral base into a morpheme. This paper speculates that the three types of bases form a grammaticalization cline, suggesting that more types of bases are possible morphosyntactically, depending on the grammaticalization path.

Keywords: numeral, base, category, syntax, Slavic

1 Introduction

Developed numeral systems are characterized by serialization (von Mengden 2008): the ability to combine numerals together to create reference to unlexicalized quantities. The quantity 304, for example, is expressed via a combination of the numerals 3, 100, and 4 in English. Crucial to serialization, or complex numeral formation, are the base numerals, e.g. English 100 and 1000. Base numerals license mathematical operations like multiplication or addition, which are central to complex numeral formation (e.g. 304 = 3×100+4). This property distinguishes base numerals from other simplex numerals, which do not license mathematical operations, e.g. *two seven* and *seven and one*.

Base numerals have been observed to show morphosyntactic differences from other simplex numerals. Corbett (1978), for example, argues that crosslinguistically, higher numerals differ from lower numerals in being more noun-like, his higher numerals generally corresponding to base numerals. This trend is evidenced with English and Polish base numerals. English 100, for instance, requires an indefinite article when no other material is present (e.g. a determiner, demonstrative, or other numeral), while non-base simplex numerals do not:

(1) a. *(a) hundred books
     b. (*a) two books

Polish 1000 in subject position can trigger gender and number agreement on the verb, while a non-base simplex numeral like 5 does not:

(2) a. Cały   tysiąc   dziewczyn   spał.
        whole.M.SG.NOM 1000.M.SG.NOM girls.F.PL.GEN slept.M.SG
        'A whole thousand girls slept.'
     b. Pięć   dziewczyn   spało.
        five.NV.NOM/ACC girls.F.PL.GEN slept.N.SG$_{(default)}$
        'Five girls slept.'

That higher numerals in Polish and other Slavic languages differ from other numerals has been recognized in various places in the literature, where such numerals are suggested to be (more) nominal, e.g. Rutkowski (2002) and Miechowicz-Mathiasen (2014) on Polish, Neidle (1988) and Franks (1995) on Russian, Giusti & Leko (2005) on Bosnian/Serbian/Croatian, and Veselovská (2001) on Czech.

Morphosyntactic differences between base and non-base numerals have led some to propose a deeper difference between the two numeral types. Kayne (2005), for instance, proposes a (silent) nominalizing suffix -NSFX which attaches only to base numerals and in effect allows them to act as bases. In his approach, being a base is a matter of whether something combines with silent -NSFX or some overt equivalent, and only bases have this property. From another perspective, Rothstein (2013) proposes that base numerals (or in her terminology, lexical powers, e.g. 100, 1000, but not 10), have a different semantic type than non-base numerals; she relates this to their need for some kind of multiplier, this being built into the semantic type of the base numeral, and the ability of these bases to form approximatives (e.g. *hundreds, thousands*).

Ionin & Matushansky (2018) take an opposing approach, arguing that Rothstein's (2013) base/non-base dichotomy is insufficient empirically and theoretically. Instead, they develop an account in which all numerals are of the same

semantic type but have varying morphosyntactic properties, which they argue do not clearly correlate with basehood. They suggest that what is and is not a base is extralinguistic, and make use of a diacritic to identify those numerals that can function as bases. In essence, they reject the idea that bases differ from non-bases semantically, but accept that bases may differ morphosyntactically, though not in any way systematic enough to suggest a special status for bases.

The present paper is concerned with the morphosyntactic status of base numerals. Like Kayne (2005) and Rothstein (2013), it explores the hypothesis that the morphosyntactic differences observed between base and non-base numerals are meaningful, but in line with Ionin & Matushansky (2018), it accepts that a simple dichotomy of base/non-base is insufficient empirically and pursues a more nuanced approach. A conclusion of this paper is that bases can differ syntactically from non-bases, and furthermore, that there are at least three types of base numerals among Polish and English 10, 100, and 1000: nominal bases, syntactic bases, and lexicalized bases; this paper speculates that these may represent steps along a grammaticalization cline, leaving the potential open for even more bases morphosyntactially. Whether morphosyntactic differences between bases and non-bases also relate to semantic differences along the lines of Rothstein's (2013) analysis remains beyond the scope of this paper.

This paper is structured as follows. It begins in §2 by introducing some recent literature on the internal structure of simplex numerals, adopting a root analysis of numerals. It then turns to numerals 10, 100, and 1000 in Polish and English in §3, arguing that they instantiate three types of bases. §4 explores how this might relate to a grammaticalization cline, drawing on historical evidence presented in previous literature, and finally §5 concludes.

2 The internal structure of a numeral

Recent research on complex numeral formation has adopted the view that complex numerals are constructed in the syntax (see especially Ionin & Matushansky 2004, 2006, 2018). According to Ionin & Matushansky (2018), complex numerals are formed using existing syntactic means in a language, e.g. complementation or potentially adjunction for multiplication and coordination or adpositional structures for addition. In most approaches, the numerals involved in complex numeral formation are atoms and have no internal structure themselves. However, there has been a trend in recent research to decompose even apparently atomic words into pieces of structure, starting with approaches in the late 1980s and early 1990s which isolate inflection (tense, agreement, number) from the verb or noun (e.g. Pollock 1989, Ritter 1991), to the relatively recent sub-field of nano-syntax

(e.g. Starke 2010), which decomposes individual words into features, even without clearly identifiable morphemes corresponding to those features. This general line of thought has been applied to numerals, with some researchers suggesting that individual simplex numerals can be internally complex. In this section, I will briefly highlight a few analyses, and discuss how they motivate an extended decomposition for base numerals.

Fassi Fehri (2018: Ch. 3) makes the claim that a Distributed Morphology (DM)-style approach is appropriate for simplex numerals. He proposes that numerals correspond to an acategorial root embedded under functional structure, a premise which is also adopted in Klockmann (2017) and Wągiel (2020, forthcoming) for Polish numerals. Fassi Fehri (2018: 61) points out that numerals are *polycategorial*, meaning that they take the form of a variety of categories crosslinguistically (see e.g. Ionin & Matushansky 2018: section 3.4 for examples of nominal, adjectival, verbal, and mixed numerals), and furthermore, that numerals are *polysemous*, meaning that they can express a variety of numerosity-related senses: cardinals, ordinals, fractionals, etc. The proposal that numerals contain a root at their core which is embedded under functional structure provides the needed flexibility for capturing the differing but related senses that are found (presumably via different functional structures above the root), as well as the numerous idiosyncrasies and category types associated with various numerals (e.g. the notorious case and agreement patterns found with Slavic numerals).

There is a further reason to treat numerals as containing roots: numerals can be considered to form a (semi-)open class of elements. The distinction between open and closed class is often taken to correlate with being a lexical or functional category (e.g. Abney 1987), where lexical categories like nouns and verbs are open class, and functional categories like tense or number are closed class. If being lexical corresponds to containing a root (as argued for in Klockmann 2017: Ch. 2), then presumably the correlation relates to it being easier to add new roots to the lexicon than new functional items. As such, the ability to add new numerals to a numeral system would argue in favor of its treatment as open class, and hence as being lexical and containing a root. Fictitious numerals and high numerals provide such evidence. While in a language like English the most useful quantities have already been named (*hundred, thousand, million, billion, trillion*), new lexical items have been created to name very high quantities, e.g. *quadrillion, quintillion, sextillion, vigintillion, centillion, googol, googolplex*. Likewise, numeral-like lexical items also exist to describe fictitious quantities, e.g. *zillion, gazillion, bajillion*. These lexical items are presumably numerals,[1] and as such, suggest that

---

[1] For example, the definition of *googolplex* on Wikipedia (https://en.wikipedia.org/wiki/Googolplex) clearly makes use of *googol* as a numeral: "Written out in ordinary decimal notation, it is 1 followed by 10^{100} zeroes, that is, a 1 followed by *a googol zeroes*." (my emphasis)

the set of numerals is not closed class. Further in favor of this view are less developed numeral systems. Comrie (2013) provides examples of languages with very limited sets of numerals, e.g. Mangarayi (Australian) with numerals for 1–3, Ydiny (Australian) with numerals for 1–5, and Hixkaryana (Cariban, Brazil) with numerals for 1–5 and 10; Bowern & Zentz (2012) also provide substantial data on a large number of numeral systems on the Australian continent, where the majority of language varieties (n = 139) have numerals maximally up to 3 or 4. Serialization (complex numeral formation) is dependent on the numerals available in a system, and thus, for these numeral systems to grow beyond the limits of serialization, new numerals must be added to the lexicon. This suggests that the development of a numeral system is in line with its members being open class. I adopt the view that a numeral contains a root.

If numerals contain roots, the next question is what functional structure dominates that root as a cardinal numeral, i.e. in a structure such as Figure 1 what is the identity of F(P)?
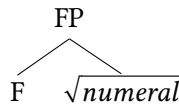
$$FP$$

F $\sqrt{numeral}$

Figure 1: Numeral functional structure

According to Wągiel (2020), who discusses the semantics of Polish numerals but also considers their internal structure, the answer is a classifier operator called Card which gives the numeral its properties as a cardinal. In Wągiel (2020), this classifier is silent and combines on top of a gender marker for numerals like 5 in Polish (which distinguish virile and non-virile gender, e.g. *pięci-u*$_V$ vs. *pięć-*$\emptyset_{NV}$); in Wągiel (forthcoming), which aligns more closely with the role of gender argued for in Fassi Fehri (2018), he adjusts the analysis and connects the classifier to the overt realization of virile gender, maintaining a silent classifier with non-virile gender. In sum, structurally the numeral 5 looks as in Figure 2 and 3, i.e. as a virile and non-virile numeral (semantic formulas omitted).

The use of a classifier in the structure of the numeral relates to Sudo (2016), who considers Japanese numeral classifier constructions. Sudo (2016) argues against the predominant view that classifiers occur in numeral constructions to make nouns count, and instead proposes that they act to convert the numeral into a modifier. This view is further consistent with the findings of Bale & Coon (2014), who show that in Chol, a Mayan language with mixed sources for numerals, the need for a classifier in a cardinal-noun construction is dependent on the
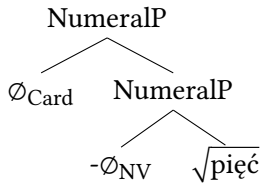
NumeralP
∅_Card   NumeralP
-∅_NV   √pięć

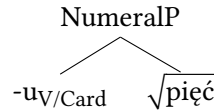Figure 2: Non-virile numeral 5

NumeralP
-u_V/Card   √pięć

Figure 3: Virile numeral 5

source language of the numeral: native Chol numerals require a classifier while imported Spanish numerals do not. Their conclusion is that the classifier occurs for the numeral. The work of Wągiel (2020, forthcoming), Sudo (2016), and Bale & Coon (2014) suggests a potential identity for the F(P) in Figure 1 – a classifier-like element which gives the numeral root its cardinal properties. For now, I will simply assume a head Card in the functional structure of a cardinal numeral.

The present discussion has focused on the decomposition of simplex numerals, and in particular, non-base simplex numerals. While the presented analyses give us a handle on what the functional structure of non-base numerals might look like, it's not immediately clear that they translate to the base numerals. The base numeral 1000 in Polish, for example, does not distinguish virile and non-virile gender like its non-base counterpart 5, and as we shall see shortly in §3.1, it has a number of other properties that make it incompatible with the structures in Figures 2 and 3. Despite this, the general approach, i.e. decomposing numerals into roots and functional structure, is just as plausible for base numerals as for non-base numerals, and it may turn out that they contain different or additional structure from what we've seen above. In the next section, we turn to morphosyntactic data for numerals 10, 100, and 1000 in Polish and English, which give clues into their syntactic representation.

3 Three types of base numerals

Polish and English numeral systems are centered around 10, with multiples of 10 acting as bases. In both languages, the lexical items for 10, 100, and 1000 are considered to be base numerals, given that they each seem to license addition and multiplication, e.g. in English, *six-ty* (= 6 x 10), *six hundred* (=6 x 100), and *six thousand* (= 6 x 1000) and in Polish, *sześć-dzisiąt* (= 6 x 10), *sześć-set* (= 6 x 100), and *sześć tysięcy* (= 6 x 1000). In this section, I argue that these numerals can be classified into three types of bases: nominal bases, syntactic bases, and lexicalized bases. Polish 1000 is an example of a nominal base, and it involves

the use of what is morphosyntactically a noun as a base numeral. English 100 is an example of a syntactic base, and along the lines of what was proposed by Kayne (2005), it involves a silent BASE morpheme which gives the numeral root its basehood. Finally, English 10 and Polish 10 and 100 are lexicalized bases. These are not active bases in the language, but grammaticalized morphemes (and it may not be appropriate to call them bases); the approach pursued is similar to what is proposed in Wągiel (2020).

3.1 Nominal base numerals

The numeral 1000 in Polish behaves morphosyntactically like a noun. This can be seen in its morphosyntactic paradigm and in how it interacts with other elements in the sentence. I will start by illustrating the paradigm of the numeral, and then turn to its case and agreement properties. Examples which are extracted from the National Corpus of Polish are marked as NKJP.

Polish is a language which distinguishes case, number, and gender. The numeral 1000 inflects for case and number using the same morphology as a masculine inanimate noun; this suggests it carries masculine inanimate gender. The paradigm is illustrated in Table 1, which compares the numeral 1000 *tysiąc* to the masculine inanimate noun *miesiąc* 'month'.

Table 1: Paradigm of Polish numeral 1000 and noun *miesiąc* 'month'

	SG		PL	
	'thousand'	'month'	'thousands'	'months'
NOM/ACC	tysiąc	miesiąc	tysiąc-e	miesiąc-e
GEN	tysiąc-a	miesiąc-a	tysięc-y	miesięc-y
DAT	tysiąc-owi	miesiąc-owi	tysiąc-om	miesiąc-om
LOC	tysiąc-u	miesiąc-u	tysiąc-ach	miesiąc-ach
INST	tysiąc-em	miesiąc-em	tysiąc-ami	miesiąc-ami

Simplex numerals and even numerals 10 and 100 in Polish inflect for the gender of the quantified noun, either virile (= grammatically masculine, biologically male, and human, see Rappaport 2011) or non-virile (= everything else) in the plural. The numeral 1000 does not. This is illustrated in (3).

(3) a. pięć   dziewczyn, pięciu chłopców
    five.NV girls.GEN   five.V boys.GEN
    'five girls, five boys'

    b. dziesięć dziewczyn, dziesięciu chłopców
       ten.NV girls.GEN   ten.V   boys.GEN
       'ten girls, ten boys'

    c. sto   dziewczyn, stu   chłopców
       hundred.NV girls.GEN   hundred.V boys.GEN
       'a hundred girls, a hundred boys'

    d. tysiąc   dziewczyn, tysiąc   chłopców
       thousand girls.GEN   thousand boys.GEN
       'a thousand girls, a thousand boys'

The numeral 1000 does not show agreement with the quantified noun for gender. Instead, numeral 1000 seems to have its own gender value, masculine inanimate, as suggested by its paradigm in Table 1 above. That numeral 1000 can carry its own gender feature is further evidenced by adjectival and verbal agreement: pre-modifiers (e.g. demonstratives, adjectives) and verbs can both surface with masculine singular agreement, in agreement with the numeral itself.

(4)   Cały   tysiąc   dziewczyn spał.
      whole.M.SG.NOM 1000.M.SG.NOM girls.F.PL.GEN slept.M.SG
      'A whole thousand girls slept.'

Furthermore, when plural, as in approximatives (5a) or when quantified by another numeral (5b), the numeral surfaces as plural, and verbal agreement likewise can surface as non-virile plural. The examples below use virile masculine nouns to exclude any possibility that agreement could somehow be with the genitive noun. Verbal agreement is necessarily with the plural numeral.

(5) a. Tysiące   Polaków   opuszczały obozy i
       1000s.M(NV).PL.NOM Poles.M(V).PL.GEN left.NV.PL   camps and
       więzienia.
       prisons
       'Thousands of Poles left camps and prisons.'   (NKJP)

    b. Cztery tysiące   widzów   dopingowały
       four.NV 1000s.M(NV).PL.NOM spectators.M(V).PL.GEN cheered.NV.PL
       Polaków przez   całe   spotkanie.
       Poles   through whole meeting
       'Four thousand spectators cheered Poles throughout the meeting.'
       (NKJP)

What the paradigm of the numeral and its ability to control agreement in its singular and plural form show is that the numeral carries phi-features, number and gender, like any noun in the language. These are nominal properties, and argue for its treatment as a noun.

Note that numeral 1000 can also trigger default agreement in each of the example types in (4–5) above, as illustrated below:[2]

(6) Tysiąc dziewczyn spało.
1000.M.SG.NOM girls.F.PL.GEN slept.N.SG*(default)*
'A thousand girls slept.'

Klockmann (2017) attributes this to an optional absence of gender in the representation of the numeral, a conclusion also found in Ionin & Matushansky (2018). The absence of gender leads to failed agreement on the probe, with default features as the result.

Like a noun, numeral 1000 also triggers genitive case on the quantified noun, as can be observed in previous examples; see also (7). This occurs in all case environments. This property distinguishes 1000 from other numerals like 5, 10, and 100, which only trigger genitive in structural case environments (nominative, accusative), e.g. (3) above, but not oblique case environments, e.g. (8).

(7) a. z tysiącem ptaków
with thousand.INST birds.GEN
'with a thousand birds'

b. z kluczem ptaków
with key(flock).INST birds.GEN
'with a flock of birds'

(8) a. z pięcioma ptakami
with five.INST birds.INST
'with five birds'

b. z dziesięcioma ptakami
with ten.INST birds.INST
'with ten birds'

---

[2]Pre-modifiers add more to this picture – they can optionally surface with non-virile plural default features (see Klockmann 2017: 121–122 for evidence that these are default features in the nominal domain), which appears to be failed agreement with the numeral, or as genitive plural, in agreement with the quantified noun.

c. ze   stoma   ptakami
with hundred.INST birds.INST
'with a hundred birds'

The case and agreement properties of numeral 1000 speak towards its positioning in the language as a noun. This suggests that the functional structure dominating the root of the numeral is nominal in nature. Depending on the theory of nominal functional structure adopted, this would imply some position for gender and number, which I will call GenderP and NumberP, respectively.[3] I would also propose that the numeral allows a quantificational layer in its functional structure, QP, as host to other cardinality expressions, such as numerals or quantifiers like *kilka* 'a few' or *wiele* 'many'. A numeral in this position would create a complex numeral, with 1000 as the base and the numeral in QP the multiplier (see also (9) below). Together, this gives a rough structure as below.[4]
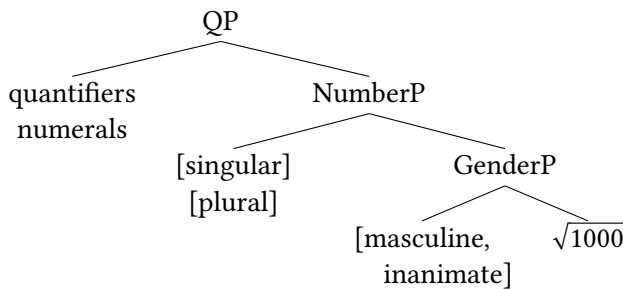

Figure 4: Structure of Polish numeral 1000

As a multiplicand in a complex numeral, the numeral 1000 inflects in the same way as a noun modified by that numeral would, i.e. numeral 2 agrees with the plural noun or numeral in gender and case (9a), while numeral 5 assigns genitive to the plural noun or numeral (9b):

(9)   a. dwa   ptaki,   dwa   tysiące
two.M.NOM birds.M.PL.NOM two.M.NOM thousands.M.PL.NOM
'two birds, two thousand'

b. pięć   ptaków,   pięć   tysięcy
five.NV.NOM/ACC birds.M.PL.GEN five.NV.NOM/ACC thousands.M.PL.GEN
'five birds, five thousand'

---

[3]In the absence of successful agreement, GenderP is absent from the structure, see (6).

[4]I will not address how the numeral combines with the noun as this takes us too far afield. There are various views on this, but most assume a numeral with no internal structure.

This is in line with the nominal status of 1000, since quantificational material combines in the same way with 1000 as with other nouns. Further in favor of this view is the behavior of 1000 with modifiers. A noun allows for an adjective between the quantifier and the noun, e.g. *trzy piękne psy* 'three beautiful dogs'; the same is true for quantified 1000, as illustrated below:

(10) a. … kosztował 8 tys.   zł.   Trzy kolejne tysiące
          cost   8 thousand złoty(currency) three next   thousands
          wydano na autokary.
          spent   on coaches
          '[It] cost 8000 złoty. The next three thousand (złoty) was spent on
          coaches.'   (NKJP)

      b. … i   pewnie jeszcze z   paroma   innymi   tysiącami
          and probably still   with a.few.INST other.INST thousands.INST
          ludzi
          people.GEN
          'and probably with another few thousand people'   (NKJP)

Modifiers are permitted internal to a complex numeral as in (10a), which is consistent with the numeral having the functional structure of a noun, even to the QP layer. Together, this argues for numeral 1000's status as a noun in Polish.

Numeral 1000 is both a noun and a base. This implies that it is possible for a base numeral to have the morphosyntax of a noun. Note that the structure in Figure 4 is not immediately compatible with the structures presented above in Figures 2 and 3, as it is not clear where a Card head would belong (Is it in QP? Is there a piece of structure above the nominal functional structure of the numeral? Is it absent?). I leave the status of Card with 1000 aside, and conclude that the nominal properties of 1000, in combination with its ability to act as a base, illustrates that base numerals can be morphosyntactically nouns.

3.2 Syntactic base numerals

The English numerals 100 and 1000 show some nominal properties, but not enough to be classified as a noun as Polish 1000 was. While like nouns they can surface with an indefinite article (*a hundred people, a thousand people*) and also allow a plural form (as an approximative: *hundreds of people, thousands of people*), they differ from nouns in many crucial ways. I will briefly compare them to nouns by considering some of the properties nominal Polish 1000 had, before turning to what makes them a syntactic base. Examples which are extracted from the Corpus of Contemporary American English are marked COCA.

Polish 1000 could control verbal and pre-modifier agreement. English 100 and 1000 cannot; both verbs and demonstratives are plural in agreement with the quantified noun:

(11) a. A {hundred/thousand} books {were/*was} stolen.
     b. {these/*this} {hundred/thousand} books

Polish 1000 required case marking on the quantified noun; no comparable *of* surfaces with English 100 and 1000:

(12) * a {hundred/thousand} of books

Likewise, Polish 1000 behaved as a noun would in a complex numeral: it surfaced as plural and it allowed intervening modifiers between it and the quantifier/numeral. Nothing comparable occurs with English 100 and 1000:

(13) a. * two {hundreds/thousands}
     b. * two {other/good/extra} {hundred/thousand}

Many of the nominal properties we might expect to find with English 100 and 1000 were they nominal bases are not present. Instead, what we do find that is "nominal" is the indefinite article *a*, which occurs when no other element is present (e.g. a determiner, demonstrative or other numeral). Given this, I would suggest that the presence of *a* is not a nominal property at all, but instead marks the presence of a morpheme BASE, which is absent with non-base numerals. I turn now to evidence in favor of this reinterpretation of the role of the article; note that the proposal below is not intended to apply to the indefinite use of *a* (as in *a cat*). I direct readers to Klockmann (2020) for a fuller discussion of the article in English cardinality expressions, and its relation to the indefinite article.

A crucial difference between English numerals 100 and 1000 and lower numerals, including 10, is the apparent indefinite article:

(14) a. one book
     b. two books, ten books
     c. a hundred books, a thousand books

However, this difference disappears when a pre-numeral modifier is included. Modification of all numerals, from simplex *one, two, ten* to complex *one hundred, two hundred* and even plural numerals, requires an article if an adjective precedes it. This is a phenomenon which has been observed in a number of works (e.g. Honda 1984, Keenan 2013, Ionin & Matushansky 2018, among others) many of which assume *a* to be an indefinite singular article.

(15) a. One property? One property? A measly one property?

   b. Maybe it will be a full two terms, maybe it won't.

   c. The animals stopped a respectful ten paces away and bowed their heads.

   d. There were more than a thousand of the latter alone, representing a good hundred journals.

   e. Sinan's best efforts had raised a bare two hundred warriors to combat the fiends.

   f. Yet there are records a mere thousands of years ago of Perseid storms

(all from COCA)

The inclusion of the article does not make the construction singular; verbal agreement remains plural, targeting the quantified noun:

(16) A further 18 women were diagnosed with ovarian cancer in the five-year period that followed.   (COCA)

I propose that the article we see is a lexicalization of the Card head (see §2), or some other more general head related to quantification. If we adopt some form of phrasal spell-out, then we can assume that the Card head is not necessarily silent, but spelled-out together with the numeral root for those numerals that do not usually show an article (e.g. *one, two, ten*). This is illustrated in Figure 5 below.

$$\text{CardP} \Rightarrow \textit{seven}$$
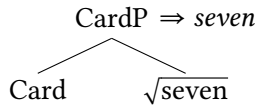
Card $\sqrt{\text{seven}}$

Figure 5: Spell-out of *seven*

When a modifier is included in the structure, it interrupts the adjacency between the numeral root and the Card head, leaving Card stranded and unlexicalized. The article *a* is used as a last-resort spell-out of this head (comparable to *do*-support in the clausal domain; we might call this Card- or Q-support). See Figure 6. Use of a modifier, then, forces this rescue operation of inserting an article, due to a requirement that Card/Q have a phonological realization. In that sense, the article is neither indefinite nor singular, and should be termed a default cardinality marker instead, as suggested by Lyons (1999).[5]

---

[5]Analyses of this type face questions about how the article disappears in the presence of D-level

CardP
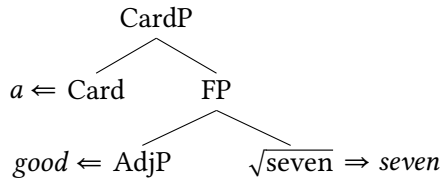a ⇐ Card   FP
good ⇐ AdjP   √seven ⇒ *seven*

Figure 6: Spell-out of modified *seven*

Returning to English 100 and 1000, even in the absence of a modifier, the article is needed. I propose that the motivation for said article is the same. There is an intervener, and it prevents the numeral from spelling out with Card. Given that what distinguishes these numerals from the others is their basehood, I propose that the intervener is a silent morpheme BASE. BASE blocks phrasal spell-out of the numeral and Card and instead, Card must be realized by the article *a*.

CardP
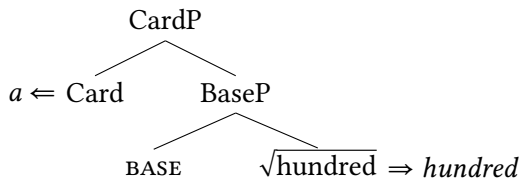a ⇐ Card   BaseP
BASE   √hundred ⇒ *hundred*

Figure 7: Spell-out of *hundred*

In unmodified multiplicative complex numerals (e.g. *seven hundred*) no article occurs, suggesting the spell-out issue has been resolved. Under the analysis presented in Figure 5 above, non-base simplex numerals spell-out CardP in addition to the numeral root; thus, we can assume that the use of a multiplier provides CardP with a spell-out, alleviating the need for the article. This is depicted in Figure 8. Note that introduction of a modifier (*a good seven hundred*) reintroduces the need for the article, similarly to Figure 6.

Note that the analysis in its current form places different spell-out requirements on BASE and Card; BASE can lack phonological content while Card cannot. This could imply that Card has a special status over BASE; alternatively, it may suggest that *hundred* and *thousand* phrasally spell-out BASE as well, but not Card. I leave this open for now.

---

material like determiners and demonstratives if it is not a determiner itself (e.g. *the (*a) hundred books*). There are various possibilities – there may be a phonological constraint preventing their co-occurrence (Lyons 1999), *the* might also have quantificational properties which obviates the need for the article (Borer 2005), or they might indeed co-occur if what is in D is only *th-*.
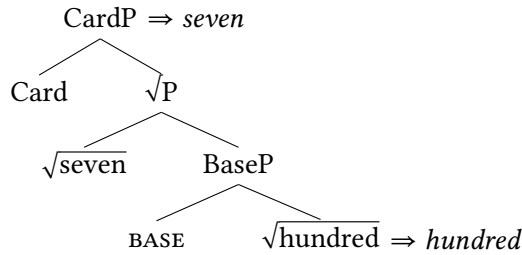
Figure 8: Spell-out of *seven hundred*

The use of a morpheme BASE to give the numeral roots *hundred* and *thousand* their basehood is what I refer to as a "syntactic base"; these become bases via the syntactic structure. Note that the final proposal, i.e. of a silent morpheme BASE which combines with the numeral, is not very far from what was proposed by Kayne (2005) and adjusted in Kayne (2019); in both cases a silent morpheme combining with bases is assumed: -NSFX in Kayne (2005) and SET in Kayne (2019) (though Kayne's SET combines with a wider range of numerals than BASE).

3.3 Lexicalized base numerals

I reserve the term "lexicalized base" for numerals which appear to license mathematical operations, but do not do so in a transparent or productive way. Instead, I propose that there are lexicalized morphemes, distinct from the numerals they are bases of, which fulfill the base function that the root and its functional structure previously filled. In this sense, these numerals are not true bases. This analysis applies to English 10 and Polish 10 and 100.

English 10 appears to have two allomorphs when functioning as a base, *-ty* and *-teen*. The morpheme *-ty* is a multiplicative base occurring only with multipliers (e.g. *thir-ty, for-ty, fif-ty, six-ty*) and the morpheme *-teen* is an additive base occurring only with additives (e.g. *thir-teen, four-teen, fif-teen, six-teen*). I propose that *-ty* and *-teen* are not allomorphs of *ten*, but instead are distinct morphemes which express multiplication by 10 and addition by 10, respectively (see von Mengden 2010 for a similar approach to *-ty* and *-teen*). This is the approach taken by Wągiel (2020) for Polish, who encodes multiplication and addition in the semantics of the morpheme. These morphemes augment the value denoted by the simplex numeral they combine with (which he takes to be of type *n*). The structures and formulas in Figures 9–10 are borrowed from Wągiel (2020) and adjusted for English and the present paper.[6]

---

[6]CardP with English lexicalized bases is not realized as the article *a* unless a modifier is present

CardP
Card   NumeralP
-teen   √fif-
$\lambda n.\text{INTEGER}(n)$
$[n + 10]$

CardP
Card   NumeralP
-ty   √fif-
$\lambda n.\text{INTEGER}(n)$
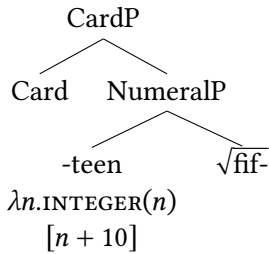$[n \times 10]$
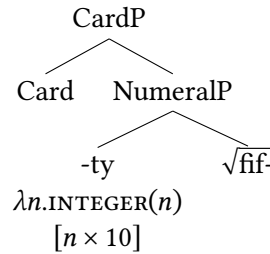
Figure 9: English additive *-teen*   Figure 10: English multiplicative *-ty*

Presumably, contextual allomorphy adjusts the phonological form of the multiplier, e.g. *five* to *fif-* and *three* to *thir-* in the context of a multiplicative or additive base morpheme. Under this analysis, *ten* is a non-base simplex numeral, while *-ty* and *-teen* are functionalized morphemes, grammaticalized from a previous stage in which *ten* was a base. In this sense, *ten* is not a base, but *-ty* and *-teen* are. This captures the fact that *ten* does not need an article (**a ten*) and that it cannot pluralize on its own as an approximative (**tens of people*) (for this, it requires the presence of a base numeral, e.g. *tens of thousands of people*).

Polish 10 and 100 are likewise lexicalized base numerals. As with English 10, the multiplicative and additive base morphemes for Polish 10 and 100 are distinct from the lexical items for 10 and 100. The forms of 10 and 100 are given in Table 2. The NOM/ACC forms are used with non-virile nouns in nominative and accusative case contexts, while the OBL forms are used with virile nouns in all case contexts and with non-virile nouns in oblique case contexts. An additional instrumental form (with *-oma* instead of *-u*), not depicted here, also exists for all numerals except 500–900.[7]

A few words regarding Table 2 are in order here. Firstly, the multiplicative and additive forms of 10 and 100 are not consistent with the forms of the lexical items for 10 and 100 (e.g. the first row vs. all other rows). In the nominative/accusative columns, the forms are fully distinct, while in the oblique columns, they are partially distinct (10 shows regularity with multipliers 5–9, while 100 shows regularity with multipliers 2–4). The distinct forms are frozen, from a stage in which 10 and 100 were transparent, productive bases. For example, *-ście* (in 200) and *-sta* (in 300, 400) are historical nominative dual and plural forms for 100, while *-set*

(e.g. *fifteen minutes* vs. *a good fifteen minutes*); this suggests that lexicalized bases are not interveners for spell-out (unlike BASE) and can spell-out CardP in combination with the numeral root.

[7]The absence of a form with *-oma* correlates with the positioning of the gender/case marker, which for 500–900 occurs on the multiplier and for all other numerals, on the multiplicand.

Table 2: Morphological form of Polish 10 and 100. *Note*: The form of the multiplier/additive differs for 40, 15, and 19.

	10		100	
	NOM/ACC	OBL	NOM/ACC	OBL
	dziesięć	dziesięci-u	sto	st-u
2 ×	-dzieścia	-dziest-u	-ście	-st-u
3–4 ×	-dzieści	-dziest-u	-sta	-st-u
5–9 ×	-dziesiąt	-dziesięci-u	-set	-u-set
1–9 +	-naście	-nast-u		

(in 500–900) is a historical genitive plural form of 100 (Dziubała-Szrejbrowska 2014). These forms are in line with historical (and modern) properties of 2–4 and 5–9, which showed agreement (2–4) or genitive case assignment (5–9) with subjects. This pattern is repeated in the frozen forms of 10 (Miechowicz-Mathiasen 2014); similarly, 10's additive forms are historically derived from a prepositional construction *na dęsete* 'out of ten' (Dziubała-Szrejbrowska 2014: 86). Thus, we see a lack of transparency in the modern multiplicative and additive forms of these numerals.[8] Secondly, in terms of their morphosyntactic behavior, Polish 10, 100 and their multiples and additives behave identically to non-base numerals like 5; this was already shown in (3), which illustrated their gender agreement and genitive case assignment properties, and in (8), which illustrated their case agreement properties in oblique environments. We can add to this their pattern of triggering default agreement, given below:

(17)   a. Pięć dziewczyn   spało.
          five girls.F.PL.GEN slept.N.SG(*default*)
          'Five girls slept.'

   b. {Dziesięć / dwanaście / dwadzieścia} dziewczyn   spało.
         ten   twelve   twenty   girls.F.PL.GEN slept.N.SG(*default*)
        'Ten / twelve / twenty girls slept.'

---

[8]Further evidence can be found with numeral 12. In Modern Polish complex numerals, the additive component determines the case properties of the quantified noun, e.g. in subject position, 22–24 have nominative quantified nouns, while 25–29 have genitive quantified nouns (a pattern repeated in the 30s, 40s, etc.). In modern Polish, 12 requires genitive on the noun, but Dziubała-Szrejbrowska (2014: 96–97) reports that in Old Polish it also allowed nominative. This shows a different status of the 2-component in modern numeral 12.

    c. {Sto   / dwieście}   dziewczyn   spało.
      hundred   two.hundred girls.F.PL.GEN slept.N.SG$_{(default)}$
      'A hundred / two hundred girls slept.'

There is not the space to attempt a full analysis of the properties of these numerals in this paper, but what we see is that (a) the forms of 10 and 100 as simplex numerals and bases are distinct and (b) both 10 and 100 pattern with non-base numerals morphosyntactically, as do their multiples and additives. Under a lexicalized base analysis, this is because the lexical items for 10 and 100 are not bases in the language, but there are corresponding morphemes which are.[9] There are three lexicalized base morphemes, with allomorphs conditioned by the numeral root and case: × 10 (*-dzieścia, -dziesiąt, -dziestu, -dziesięciu*), + 10 (*-naście, -nastu*), and × 100 (*-ście, -sta, -set, -stu*). These morphemes augment the value of the root they combine with, and furthermore, assign it the morphosyntax of a numeral like 5, 10 and 100. In Wągiel's (2020) analysis of Polish, the root combines with the base morpheme, a gender node, and Card. I will omit gender from the structure for now, pending further analysis on the case and agreement properties of these items; what is crucial here is the status of base morpheme.[10] See Figures 11 and 12.
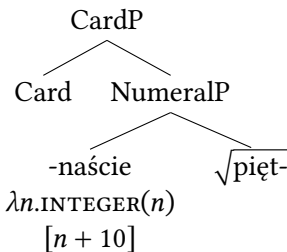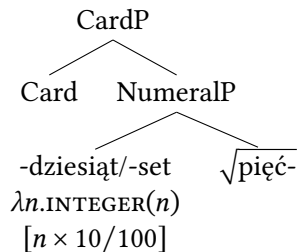

Figure 11: Polish additive *-naście*

Figure 12: Polish multiplicative *-dziesiąt/-set*

    English 10 and Polish 10 and 100 are lexicalized bases. In the context of this paper, this implies that there are grammaticalized morphemes, distinct from the lex-

---

[9]Something more needs to be said about 100, which does not permit multipliers, e.g. **jedno sto*, but does allow additives, e.g. 101 (*sto jeden*) to 199 (*sto dziewięćdziesiąt dziewięć*). This may suggest it remains an additive base, but not a multiplicative base, in contrast to 10 which is neither.

[10]Differences in the position of the gender/case morpheme in these complex numerals may also suggest that gender/case has a different position with respect to the base morpheme in different numerals: gender/case seems to sit between the root and the base morpheme for 500–900, but above the base morpheme for 11–19, 20–90, and 200–400. Such a low position with 500–900 might explain their lack of a dedicated instrumental form, as mentioned in footnote 7.

ical items for these numerals, which combine with the root of a simplex numeral and create basehood. These base morphemes have a very restricted distribution in that they only augment roots for 1–9 and certain quantifiers.

4 Grammaticalization

I would like to suggest that the three types of bases identified in this paper, nominal bases, syntactic bases, and lexicalized bases, represent stages along a grammaticalization path from noun to morpheme. This section will explore this hypothesis and possible evidence in favor of it.

Nominal bases involve the functional structure of a lexical noun; lexicalized bases are morphemes that give basehood by augmenting the value of the numeral root. These appear to be initial and final stages of a grammaticalization path for base numerals, a hypothesis which is supported by Polish 10 and 100. As mentioned in §3.3, historically numerals 10 and 100 combined transparently with other simplex numerals to form complex numerals (see Miechowicz-Mathiasen 2014 and Dziubała-Szrejbrowska 2014); this is because they were both nominal bases (see also Miechowicz-Mathiasen 2014). This is supported by the examples below, illustrating their ability to control verbal agreement[11] and to trigger genitive case assignment even in an oblique case environment; these are properties which modern-day Polish 1000 carries (see 4, 5, and 7), but modern-day 10 and 100 have lost (see 8 and 17b).

(18) a. Jako minęła   dziesięć lat.
        as   passed.F.SG ten.F.SG years.GEN
        'As ten years passed.'
                    (Siuciak 2008 as cited in Dziubała-Szrejbrowska 2014: 103)

   b. ku   trzydzieści i   ku   stu   lat
      towards thirty₂   and towards hundred.DAT.SG years.GEN.PL
      'to a hundred and thirty years'   (Miechowicz-Mathiasen 2013: 99)

This data is suggestive of the nominal base status of Polish 10 and 100 in earlier stages. With regards to English, the picture is less clear, as additive and multiplicative 10 had already fossilized in Old English (and therefore formed a lexicalized base) (von Mengden 2010). However, von Mengden (2010) argues that the grammaticalization relation between *tyn* '10' and *tyne* '+ 10' remained visible in

---

[11]Though, see Miechowicz-Mathiasen (2014) for a fuller discussion of the intricacies of agreement with Old and Middle Polish numerals.

Old English, *tyne* being an inflected form of *tyn* in a previous stage of English; no such obvious connection is visible with multiplicative *(hund-)-tig* ' × 10', though von Mengden (2010) suggests a similar earlier grammaticalization process.[12]

English 100 and 1000 may have had a more nominal status than they do today. Von Mengden (2010) reports that Old English numerals higher than 20 often participated in a "partitive construction," namely, the use of genitive on the quantified noun without a subset interpretation. This could also be accompanied by singular agreement on the verb. These patterns are reminiscent of what we see in modern Polish 1000, a nominal base. Example (19) illustrates the use of genitive case with 100 and 1000 but not 10, and (20) illustrates the use of a singular verb with a multiple of 10.

(19) a. tyn colt-um
        10   colt-DAT.PL
        '10 colts'   (von Mengden 2010: 219)

     b. hund cne-a   werþeod-a
        100   generation-GEN.PL people-GEN.PL
        '100 generations of men'   (von Mengden 2010: 220)

     c. ðusend ge-wæpn-od-ra   cemp-ena
        1000   CIRC-arm-PTCP-GEN.PL fighter-GEN.PL
        'a thousand armed warriors'   (von Mengden 2010: 220)

(20) wear-ð   [...] fiftig mann-a   ofsleg-en
     become.PRS-3SG [...] 50   man-GEN.PL slay.PTCP-PTCP
     'there were 50 men killed'   (von Mengden 2010: 224)

I suggest the following grammaticalization process. A nominal base begins grammaticalization by shedding some of the projections that make it nominal (see Miechowicz-Mathiasen 2014); this seems plausible for Polish 10 and 100 and English 100 and 1000, and likewise, may be an ongoing process for modern Polish 1000, specifically with regards to a loss of gender (see 6). This results in a reduced functional structure above the numeral root, and I suggest that at some point, this reduced functional structure is reanalyzed as a BASE morpheme, the result being a syntactic base. As a final step, the numeral root and base morpheme coalesce into a single functional morpheme, acting as an additive or multiplicative base. The structures in Figures 13–15 illustrate these three stages (omitting the Card projection).

---

[12]The morpheme *(hund-) -tig* was a suffix on 2–6 (20–60) and a circumfix on 7–12 (70–120).
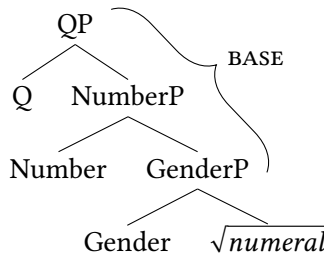
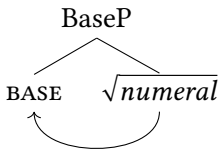Figure 13: Stage 1 – nominal base

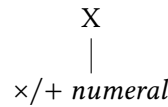

Figure 14: Stage 2 – syntactic base

Figure 15: Stage 3 – lexicalized base

There seems to be clear evidence that Polish 10 and 100 have grammaticalized from nominal bases to lexicalized bases (see 18) – however, it remains to be seen whether they underwent a syntactic base stage, as is predicted under the hypothesis above. I leave this for future work, along with the question of whether English 10 and 100 were indeed nominal bases. Altogether, this hypothesis gives us a handle on why we see three types of bases: these are developmental stages from noun to morpheme.

As a final note, this hypothesis predicts substantial variation in the morphosyntax of base numerals cross-linguistically. If a base numeral grammaticalizes from a noun to a morpheme, then its morphosyntax will depend on how noun-hood is realized in the language, how grammaticalization proceeds, and how functional projections are spelled-out. For example, Polish is a rich case and agreement language, with gender on nouns, but no definite/indefinite determiner distinction; English is the reverse, with a rich system of determiners, no gender on nouns, and a morphologically poor system of case and agreement. The consequence is that the properties of nouns in Polish and English differ (e.g. gender or no gender, triggering agreement on something or not, etc.), and thus, nominal bases are likewise expected to differ between the languages. The process of grammaticalization is also important, both regarding the language as a whole and the individual lexical item. Changes in the language, such as the loss of case on Old English nouns or the introduction of a new gender distinction in Old Polish (Miechowicz-Mathiasen & Dziubała-Szrejbrowska 2013), could affect the realization of a numeral and its grammaticalization path. Likewise, the changes

317

that a numeral undergoes, such as gender loss (ongoing for Polish 1000), might differ between numerals, predicting more variation among bases. Finally, how functional projections are spelled-out (for example, if a language has an overt BASE morpheme or not) can create further differences between base numerals. In sum, we expect dramatic differences between base numerals cross-linguistically, but we also expect those differences to be in line with the properties of nouns, defective nouns, and morphemes in that language, diachronically and synchronically. This could mean that we find many "types" of base numerals, but under this hypothesis, they are constrained by the grammaticalization path from noun to morpheme and the spell-out of functional projections.

5 Conclusion

This paper has proposed that there are three types of bases: nominal bases, syntactic bases, and lexicalized bases. This analysis has built on the idea that numerals can be internally complex, and in particular, that they consist of a root which is dominated by functional structure. For nominal bases, that functional structure is nominal in nature; for instance, Polish 1000 consists of a root, number and gender features, and a quantificational layer. For syntactic bases, that functional structure involved a morpheme BASE which gave the numeral root its basehood. Lexicalized bases do not have internal structure, because they are grammaticalized morphemes, distinct from the numerals they are bases of (those numerals being non-bases synchronically). It was also proposed that these bases form steps along a grammaticalization path from noun to morpheme.

The present proposal is limited empirically to Polish and English numerals. However, the general spirit of it may be applicable to other languages, since it predicts a wide array of variation cross-linguistically, constrained by the noun-to-morpheme grammaticalization path and spell-out. How noun-hood is realized and how grammaticalization proceeds can lead to very different looking numerals cross-linguistically; furthermore, how functional projections are spelled-out (e.g. CardP, BaseP) may lead to other differences. Exploring the diachronic and synchronic properties of bases in other languages may provide further evidence for the base types proposed above and the grammaticalization path. Finally, the patterns discussed here are relevant for base numerals which grammaticalize from nouns. It may be possible that base numerals grammaticalize from other categories, in which case more types of base numerals could exist cross-linguistically.

Abbreviations

3	third person	NOM	nominative
ACC	accusative	NV	non-virile
DAT	dative	OBL	oblique
F	feminine	PL	plural
GEN	genitive	PRS	present tense
INST	instrumental	PTCP	participle
LOC	locative	SG	singular
M	masculine	V	virile

Acknowledgements

References

Abney, Steven. 1987. *The English noun phrase in its sentential aspect.* Cambridge, MA: Massachusetts Institute of Technology. (Doctoral dissertation). http://www.ai.mit.edu/projects/dm/theses/abney87.pdf.

Bale, Alan C. & Jessica Coon. 2014. Classifiers are for numerals, not for nouns: Consequences for the mass/count distinction. *Linguistic Inquiry* 45(4). 695–707. DOI: 10.1162/LING_a_00170.

Borer, Hagit. 2005. *Structuring sense Volume I: In name only.* Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199263905.001.0001.

Bowern, Claire & Jason Zentz. 2012. Diversity in the numeral systems of Australian languages. *Anthropological Linguistics* 54(2). 133–160. DOI: 10.1353/anl.2012.0008.

Comrie, Bernard. 2013. *Typology of numeral systems.* Ms. Max Planck Institute for Evolutionary Anthropology.

Corbett, Greville G. 1978. Universals in the syntax of cardinal numerals. *Lingua* 46(4). 355–368. DOI: 10.1016/0024-3841(78)90042-6.

Dziubała-Szrejbrowska, Dominika. 2014. *Aspects of morphosyntactic constraints on quantification in English and Polish*. Poznań: Adam Mickiewicz University. (Doctoral dissertation).

Fassi Fehri, Abdelkader. 2018. *Constructing feminine to mean: Gender, number, numeral, and quantifier extensions in Arabic*. Lanham: Rowman & Littlefield.

Franks, Stephen. 1995. *Parameters of Slavic morphosyntax*. Oxford: Oxford University Press.

Giusti, Giuliana & Nedzad Leko. 2005. The categorial status of quantity expressions. *Lingvisticki Vidici* 34. 121–183. http://hdl.handle.net/11707/106.

Honda, Masura. 1984. On the syntax and semantics of numerals in English. *Journal of Osaka Jogakuin 2-year College* 14. 97–115.

Ionin, Tania & Ora Matushansky. 2004. A singular plural. In Vineeta Chand, Ann Kelleher, Angelo J. Rodríguez & Benjamin Schmeiser (eds.), *Proceedings of WCCFL*, vol. 23, 399–412. Somerville, MA: Cascadilla Press.

Ionin, Tania & Ora Matushansky. 2006. The composition of complex cardinals. *Journal of Semantics* 23(4). 315–360. DOI: 10.1093/jos/ffl006.

Ionin, Tania & Ora Matushansky. 2018. *Cardinals: The syntax and semantics of cardinal-containing expressions*. Cambridge, MA: MIT Press. DOI: 10.7551/mitpress/8703.001.0001.

Kayne, Richard S. 2005. A note on the syntax of numerical bases. In Yubun Suzuki, Keizo Mizuno & Ken-Ichi Takami (eds.), *In search of the essence of language science: Festschrift for Professor Heizo Nakajima on the occasion of his sixtieth birthday*, 21–41. Tokyo: Hituzi Syobo.

Kayne, Richard S. 2019. Some thoughts on 'one' and 'two' and other numerals. In Ludovico Franco & Paolo Lorusso (eds.), *Linguistic variation: Structure and interpretation*, 335–356. Berlin: de Gruyter. DOI: 10.1515/9781501505201.

Keenan, Caitlin. 2013. 'A pleasant three days in Philadelphia': Arguments for a pseudopartitive analysis. *University of Pennsylvania Working Papers in Linguistics* 19. 87–96. https://repository.upenn.edu/pwpl/vol19/iss1/11/.

Klockmann, Heidi. 2017. *The design of semi-lexicality: Evidence from case and agreement in the nominal domain* (LOT Dissertation Series 460). Utrecht: LOT. https://www.lotpublications.nl/the-design-of-semi-lexicality.

Klockmann, Heidi. 2020. The article *a(n)* in English quantifying expressions: A default marker of cardinality. *Glossa* 5(1). DOI: 10.5334/gjgl.1151.

Lyons, Christopher. 1999. *Definiteness*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511605789.

Miechowicz-Mathiasen, Katarzyna. 2013. On agreement with numeral phrase subjects including ≥ 5 in Old and Middle Polish. *Scando-Slavica* 51(1). 80–107. DOI: 10.1080/00806765.2013.800727.

Miechowicz-Mathiasen, Katarzyna. 2014. Numeralization of numeral nouns in Polish. In Ludmila Veselovská & Markéta Janebová (eds.), *Nominal structures: All in complex DPs*, 48–68. Olomouc: Palacký University.

Miechowicz-Mathiasen, Katarzyna & Dominika Dziubała-Szrejbrowska. 2013. The role of gender in the rise of numerals as a separate category. *Journal of Historical Syntax* 1(1). 1–39. https://journals.linguisticsociety.org/elanguage/jhs/article/view/2787.html.

Neidle, Carol. 1988. *The role of case in Russian syntax*. Dordrecht: Kluwer. DOI: 10.1007/978-94-009-2703-2.

Pollock, Jean-Yves. 1989. Verb movement, universal grammar, and the structure of IP. *Linguistic Inquiry* 20(3). 365–424. http://www.jstor.org/stable/4178634.

Rappaport, Gilbert. 2011. Toward a multi-level theory of morphology: How Polish gender works. In Piotr Bański, Beata Łukaszewicz, Monika Opalińska & Joanna Zaleska (eds.), *Generative investigations: Syntax, morphology, and phonology*, 166–197. Newcastle upon Tyne: Cambridge Scholars Publishing.

Ritter, Elizabeth. 1991. Two functional categories in noun phrases: Evidence from Modern Hebrew. In Susan Rothstein (ed.), *Perspectives on phrase structure: Heads and licensing*, vol. 25 (Syntax and semantics), 37–62. San Diego: Academic Press. DOI: 10.1163/9789004373198_004.

Rothstein, Susan. 2013. A Fregean semantics for number words. In Maria Aloni, Michael Franke & Floris Roelofsen (eds.), *Proceedings of the 19th Amsterdam Colloquium*, 179–186. Amsterdam: ILLC.

Rutkowski, Paweł. 2002. The syntax of quantifier phrases and the inherent vs. structural case distinction. *Linguistic Research* 7(1). 43–74.

Siuciak, Mirosława. 2008. *Kształtowanie się kategorii gramatycznej liczebnika w języku polskim [The development of grammatical category of numerals in Polish]*. Katowice: Wydawnictwo Uniwersytetu Śląskiego.

Starke, Michal. 2010. Nanosyntax: A short primer to a new approach to language. *Nordlyd* 36(1). 1–6. DOI: 10.7557/12.213.

Sudo, Yasutada. 2016. The semantic role of classifiers in Japanese. *Baltic International Yearbook of Cognition, Logic and Communication* 11(1). 1–15. DOI: 10.4148/1944-3676.1108.

Veselovská, Ludmila. 2001. Agreement patterns of Czech group nouns and quantifiers. In Norbert Corver & Henk van Riemsdijk (eds.), *Semi-lexical categories: The function of content words and the content of function words*, 273–320. Berlin: de Gruyter. DOI: 10.1515/9783110874006.

von Mengden, Ferdinand. 2008. The grammaticalization cline of cardinal numerals and numeral systems. In María José López-Couso & Elena Seoane (eds.),

Rethinking grammaticalization: New perspectives, 289–308. Amsterdam: John Benjamins. DOI: 10.1075/tsl.76.14men.

von Mengden, Ferdinand. 2010. *Cardinal numerals: Old English from a cross-linguistic perspective*. Berlin: de Gruyter. DOI: 10.1515/9783110220353.

Wągiel, Marcin. 2020. Several quantifiers are different than others: Polish number-denoting indefinites. In Franc Marušič, Petra Mišmaš & Rok Žaucer (eds.), *Advances in formal Slavic linguistics 2017*, 323–353. Berlin: Language Science Press. DOI: 10.5281/zenodo.3764869.

Wągiel, Marcin. Forthcoming. Grammatical gender meets classifier semantics: Evidence from Slavic Numerals. In Vera Gribanova, Sabrina Grimberg, Erika Petersen O'Farrill, Eva Portelance & Brandon Waldon (eds.), *Formal Approaches to Slavic Linguistics 27: The Stanford Meeting 2018*. Ann Arbor, MI: Michigan Slavic Publishers.