1. Base numerals

Base numerals are those numerals which license mathematical operations (e.g. addition or multiplication), like *hundred* or *thousand*.

(1) two hundred and five (= 2 x 100 + 5)

Languages have a finite set of base numerals, which seem to represent the highest values. The values of base numerals will differ depending on whether the language follows a decimal or vigesimal (etc.) system (see Comrie 2013). Some have noted that base numerals are characterized by an idiosyncratic morphosyntax (see Corbett 1978, who points out that higher numerals are more "nominal"). In this poster, I focus on base numerals in Polish and English, and examine their morphosyntax, building on work in my dissertation (Klockmann 2017).

2. The roots of numerals

I propose that like lexical items (nouns, verbs, adjectives), numerals have roots. However, their functional structure differs from typical lexical items, contributing to their more functional nature (see Klockmann 2017 for a detailed presentation of this idea). Reasons to consider numerals as having roots include: (a) grammaticalization, (b) open class properties, and (c) classes of words related to certain numerals. Fassi Fahri (2018), who makes a similar claim, also points out that numerals are polycategorial (they assume various categories cross-linguistically, see Donohue 2005), this being amenable to a root + functional structure analysis of numerals.

The grammaticalization cline of some numerals can be seen as a development from noun to numeral. Traces of a nominal past are still visible in some Polish numerals today, for instance, Polish 5 and 10, which Miechowicz-Mathiasen (2012) and Dziubała-Szrejbrowska (2014) have shown to have had a nominal syntax in Old Polish; remnants of this is seen in their case patterns in modern Polish. This points to a lexical origin for some numerals, which suggests they may have had a root at some stage.

Numerals could be considered open class. It is possible to create new numerals, for instance, very high numerals (*googol*, *vigintillion*) and fake numerals (*zillion*, *bajillion*); in English, these follow the morphosyntax of *million*. In languages with less developed numeral systems (e.g. Mangarayi or Ydiny (Australian) with numerals for 1-3 and 1-5, respectively, Comrie, n.d.), expansion of the numeral system will require new lexical items to name those higher values. Thus, numeral systems can and do grow, suggesting they are an open class. This places them on the side of a lexical item (as something semi-lexical, cf. Klockmann 2017), and having a root is consistent with this.

Finally, there are sets of words which seem to be related to individual numerals. For instance, the words *two*, *twin*, *twice*, *twofold* and *twos* share a two-onset as well as a meaning of two-ness; the same can be said for *three*, *thrice*, *threeword*, and *threes*. This suggests a root analysis. A similar argument is found in Fassi Fahri (2018: 61), who points out that numerals occur in a variety of senses which share a "common numerosity core"; for instance, they can function as cardinals (*three*), ordinals (*third*), fractions (*third*), the names of natural numbers (*three*), adverbials (by *three*, *thrice*), etc. This is capturable under a root analysis which assumes differing functional structures for the senses.

3. Numeral functional structure

Sudo (2016) and Wągiel (t.a.) make the claim that numerals are of type *n*, and combine with a classifier (in Japanese, cf. Sudo 2016) or silent Card head (in Polish, cf. Wągiel t.a.), which converts them to modifiers, allowing them to quantify nouns. In the morphosyntax, this would imply that it is the functional head Card/Classifier which makes the root a numeral. I adopt this as a starting point for the analysis of cardinal numerals: they consist minimally of a root and Card/Classifier.

4. Types of base numerals

I propose that numeral bases can manifest their basehood in a variety of ways morphosyntactically. The first suggestion is that numerals which are syntactically nouns can function as bases; this holds for numeral 1000 in Polish (see poster ex. 9-12). Such numerals lack a Card head, and are bases by virtue of being nouns which can be modified by numerals with a Card head; this type of numeral falls easily under the analysis proposed by Ionin & Matušanský (2006, 2018). Such numerals might not be considered numerals proper (if Card is a prerequisite for being a cardinal), but they create an effect of numeral basehood.

A second type of base I treat is lexicalized bases, which might not be considered numerals proper either. Following Wągiel (t.a.) and von Mengden (2010), I take these to be grammaticalized
morphemes which combine with a numeral root and augment its value. Thus, -ty in English, for instance, combines with a root, multiplying its value by 10. The grammaticalization path of these morphemes appears to have left them completely independent from the numerals they grammaticalized from; as a result, English 10 and Polish 10, 100 are not bases (they have the morphosyntax of non-bases), but there are corresponding morphemes which have grammaticalized the base function they previously had (see poster ex. 18-19). In Polish, the form of the morpheme is sensitive to the root it augments (and also the case context and gender of the quantified noun), and in English, it is the root which changes form (e.g. three, thir-teen, thir-ty); this can be modeled under a late insertion approach, which assumes context-sensitive allomorphy.

Finally, there are syntactic bases. These bases involve a numeral root which combines with a morpheme -BASE prior to combining with Card/classifier (the semantics of which is yet to be determined). Motivation for -BASE comes from the article a(n) which surfaces in a number of English quantificational constructions, without seeming to have a connection to singularity: a few, a bunch, a lot (see Klockmann 2017, ch. 6 for extensive discussion). When such quantified expressions occur as subjects, agreement is always with the quantified noun, suggesting we are not looking at a typical English indefinite article:

1. A few books were stolen.
2. A bunch/lot of books were stolen.

This same article seems to be occurring with base numerals in English, and in fact, with modified non-bases (ex. 16 of poster), as these all also show verbal agreement with the quantified noun:

1. A dozen/hundred books were stolen.
2. A measly seven books were stolen.

In Klockmann (2017), I develop an analysis of the article found in a lot, a ton, and a bunch, building on the article alternation of non-base numerals (seven, a measly seven). I argue it to be a default cardinality marker, drawing on insights from Lyons (1999); if Card fails to be lexicalized, a default article a(n) is inserted, which is what we find with lot, ton, and bunch.

Turning to numerals, non-base numerals usually spell out both the root and Card, and no article is needed. However, the insertion of an adjective interrupts the adjacency, triggering separate spell-out of the nodes; Card is realized as the default cardinality marker a(n), and the root as the numeral (see poster ex. 17a,b). In the case of base numerals (100, 1000), the article is required precisely in those contexts where no other marker of cardinality occurs (or the definite article: see Lyons 1999 for a solution as to why the definite article precludes the default cardinality marker). Presumably, this article is also the default cardinality marker, and the question becomes: why is it required, even if no intervener is present? A plausible answer is that there is in fact an intervener: the morpheme -BASE. The presence of -BASE blocks lexicalization of Card, triggering the default cardinality marker a(n) (see poster ex. 17c). Furthermore, it is this morpheme which licenses the basehood of the numerals 100 and 1000.

5. Grammaticalization

Grammaticalization plays an important role in the types of base numerals that we find. Lexicalized bases, in particular, are a product of grammaticalization processes, while nominal bases represent a potential “initial state” in the grammaticalization to numerals and numeral bases. With regards to Polish 10 and 100, it seems that they have developed from nominal bases (nouns) into non-base numerals, with the base function being taken over by the lexicalized base morphemes (see Dziubała-Szrejbrowska 2014 on their history). English has similarly developed lexicalized bases, though it is difficult to say yet whether these bases originated as nouns. In general, the development of base numerals seems to be triggered by functional need (see von Mengden 2008, 2010: the introduction of a new base numeral simplifies the counting process), and presumably can take a number of paths depending on the source of the base numeral (and its lexical-functional category). With regards to the innovation of English -BASE, this could be related to the loss of case and gender in Old English, and with it, evidence of a numeral’s functional structure. A morpheme -BASE serves to differentiate the base numerals from non-base numerals, as signaled by the article a(n).

6. References