Two structures for numeral-noun constructions

Gabi Danon

September 12, 2012

Abstract

This paper has two main goals: to argue that crosslinguistically there are two major types of numeral-noun constructions, one in which a projection of the numeral occupies a specifier position and one in which the numeral heads a recursive nominal structure; and to show that the choice between these two structures is partially constrained by the presence of number features and case. It is shown that numerals bearing nominal number morphology display a cluster of properties that often distinguishes them from other numerals in the same language; I claim that presence of morphosyntactic number makes the numeral sufficiently 'noun-like' to be subject to general principles of case theory.

1   Introduction

A common assumption in the literature on cardinal numerals is that, with few exceptions, the structural relation between the numeral and the noun that it combines with is uniform, both cross-linguistically and language-internally. Thus, the debate whether numerals are heads that select a projection of the noun as their complements (see e.g. Ionin and Matushansky, 2006), or specifiers of a (functional) projection of the noun (see e.g. Corver and Zwarts, 2006), has usually been carried out under the assumption that there is just one analysis that applies to (almost) all numerals, with the possible exception of the often-noticed differences between 'adjectival' and 'nominal' numerals (Zweig, 2006). Recently, for instance, Ionin and Matushansky (2006) (henceforth IM) argued convincingly that the case-related properties of (some) numerals in Russian, Finnish and Inari Sami provide strong evidence for viewing them as nominal heads that recursively take another nominal projection as a complement; they then generalize this analysis to represent the universal syntax of numerals. The first goal of this paper is to show that this kind of reasoning is not valid; using data from a variety of languages, I argue that numerals

combine with nouns in at least two distinct ways, and that IM's head-complement analysis is compatible only with some numeral-noun constructions (henceforth, NNCs).[1]

If there are two distinct configurations for NNCs, the question is whether there is any systematic rule governing the choice between them. The second goal of this paper is to argue that these two configurations are restricted by universal constraints related to number features and case. Specifically, I argue that numerals in specifier position may not carry morphosyntactic number which is independent of the number feature of the noun (i.e., non-agreeing number); while numerals as heads are possible only if the numeral or some other element is able to assign case to an embedded nominal projection, which is minimally a NumP. Thus, this paper aims to show that assuming a uniform structure for all numeral-noun constructions is not only wrong but also misses important generalizations regarding the role of features in constraining syntactic configurations.

2 Overview of previous work

Interest in the structural position of cardinal numerals can be traced back to early works on phrase-structure, due to a large extent to what looks like a high degree of irregularity within the relatively small class of numerals and 'determiners' in general. In two influential early works, Jackendoff (1977) and Selkirk (1977) argue for viewing numerals as maximal projections that serve as specifiers of NP (see also Corbett, 1978, who argues in a pre-X-bar framework that cardinals project an NP that is dominated by a larger headless NP, which also dominates a projection of the noun). In later works that adopt various forms of the DP hypothesis, two major approaches have emerged: according to one view (see e.g. Franks, 1994, Giusti, 1997, 2002 and Kayne, 2010), numerals are specifiers of a functional projection dominating NP, an approach adopted and elaborated on in much of the work on the cartography of functional projections within the noun phrase (Cinque, 2005). Alternatively, it has also been proposed (often within the 'QP hypothesis', following Sportiche, 1988) that at least in some cases, numerals (as well as other quantifiers) are heads that select a nominal projection as their complement (Borer, 2005; Cardinaletti and Giusti, 2006; Giusti, 1997; Longobardi, 2001; Shlonsky, 2004). Recently, this debate has received renewed interest, with new facts brought forward to support each of the two approaches. According to Corver and Zwarts (2006), the properties of complex 'prepositional numerals' like *between ten and fifteen* support a specifier analysis; while Ionin and Matushansky (2006) use mostly data from case

---

[1]Abbreviations used in this article: NNC = numeral-noun construction; CS = construct state; NOM = nominative; ACC = accusative; GEN = genitive; INSTR = instrumental; DAT = dative; PART = partitive; ABL = ablative; FEM = feminine; MASC = masculine; SG = singular; PL = plural; DEF = definite; 1 = 1st person; POSS = possessive; PF = perfective

marking with simple and complex numerals in Russian, Finnish and Inari Sami to argue that numerals are nouns that recursively select another nominal projection as their complement.

Many of these works make the implicit assumption that the structure of numeral-noun constructions is mostly uniform, both within a single language and across languages. Variability in the area of cardinals has been discussed mostly in connection to their categorial status: it has often been noted that many languages have both 'adjectival' and 'nominal' numerals, which differ in their morphosyntactic properties (Corbett, 1978; Zweig, 2006). In Modern Hebrew, for instance, the numeral *exad* ('one'), differs from all higher numerals in being post-nominal and showing strict gender (and number) agreement with the noun, even in colloquial speech where gender agreement with other numerals is often not preserved (see e.g. Borer 2005); *exad* thus matches the properties of adjectives, and contrasts with all other, pre-nominal numerals, which do not pattern with adjectives. The division between adjectival and non-adjectival numerals, however, is not always entirely clear, and it seems to have a somewhat gradient nature (Corbett, 1978). Despite this gradience, it has often been assumed that postulating two distinct syntactic analyses for adjectival and nominal numerals is an unavoidable step (but see Zweig, 2006, who argues that the difference in structure is quite minimal); this conclusion appears also in earlier descriptive work, such as Greenberg (1978), who distinguishes between two major types of constructions, 'adjective-noun' constructions and partitive constructions. In this paper I mostly ignore those numerals that display unequivocally adjectival properties; the question that I focus on is whether more than one structure can be justified for numerals that are not prototypically adjectives.

For non-adjectival numerals, the possibility that more than one structure exists has not often been discussed, with the exception of several language-specific analyses that distinguish two types of constructions (see, however, Borer 2005, who explicitly considers both options and argues that both can indeed be found in natural language). Danon (1996, 1998), for instance, argues that some numerals in Hebrew are specifiers and others are heads. This correlates with a morphophonological alternation available in this language: numerals, like nouns, adjectives and participles, can occur either in a 'free' form or in a bound one, where the latter gives rise to what is known as the Construct State (CS).[2] This alternation is illustrated in (1) for numerals, and in (2) for nouns:[3] [4]

---

[2] A construct state is a preposition-less genitival structure in which a phonologically weak element is immediately followed by what is apparently a full embedded DP. See Borer (1999) and Ritter (1991), among others.

[3] Throughout this paper, data for which no source is cited is data that has been collected from native speakers by the author.

[4] While the alternation in (1) is conditioned by definiteness, this is not always the case; we return to this issue in section 3.1.

(1)   a.   *šlošà*   *(sfarim)*
           three(FREE) books

           'three (books)'

       b.   *šlòšet*   **(ha-sfarim)*
           three(BOUND) the-books

           'the three books'

(2)   a.   *minharà*
           tunnel(FREE)

       b.   *minhèret*   **(ha-zman)*
           tunnel(BOUND) the-time

           'the time tunnel'   (Hebrew)

Danon (1996, 1998) argues that the fact that free numerals can be syntactically complex rules out the possibility that these are heads selecting a projection of the noun as their complement. Bound numerals, in contrast, seem to behave in many ways like nominal heads of CS (Ritter, 1991), and hence it is argued that they are heads that give rise to a recursive nominal structure by embedding a full DP and assigning abstract genitive case to it. This analysis relies heavily on the language-specific properties of the construct state, and hence it was proposed specifically for Hebrew (and possibly also Arabic), and not as a universal analysis of numeral-noun constructions.

Another work that argues for more than one structure for numeral-noun constructions in Semitic languages is Shlonsky (2004). Shlonsky discusses cardinal numerals in Hebrew and Arabic, and argues that while both languages have numerals that act as heads, Arabic – but not Hebrew – has also numerals as specifiers; the choice between these two structures, in Shlonsky's analysis, correlates with the choice between prenominal and postnominal numerals.

For Slavic languages, most notably Russian, Bailyn (2004) and Pereltsvaig (2006) have argued that numerals can be either heads or specifiers, and that the choice depends on the kind of case – structural or inherent – that the noun phrase bears. The empirical motivation has to do with the existence of both 'heterogeneous' and 'homogeneous' case patterns with Russian numerals, illustrated in the following examples from Bailyn (2004):
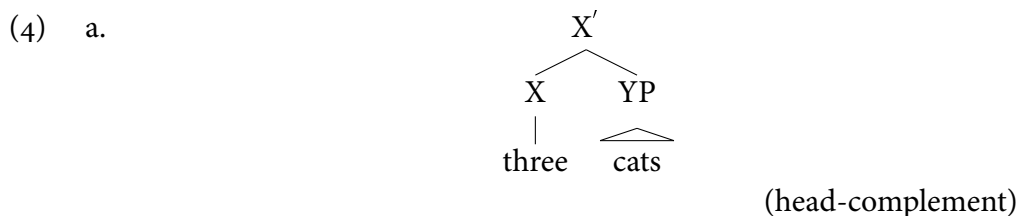
(3)   a.   *Dijana znaet pjat'*   *jazykov.*
           Dijana knows five-NOM/ACC languages-GEN

           'Dijana knows five languages.'

       b.   *Dijana vladeet pjat'ju*   *jazykami.*
           Dijana controls five-INSTR languages-INSTR

           'Dijana knows five languages.'   (Russian, from Bailyn 2004)

According to Bailyn, the numeral in cases like (3a) occupies a specifier position, while the one in (3b) is a head taking NP as its complement.

But other than such anecdotal treatments of numerals in particular languages, the general view in most of the generative literature has been that there is no compelling reason to adopt more than one analysis for numeral-noun constructions. This view has been implicitly taken for granted in recent cross-linguistic studies of numerals, such as Corver and Zwarts (2006); Ionin and Matushansky (2006); Kayne (2010) and Zweig (2006), where data from numerals in one language is taken as representative of the syntax of numerals in all languages. In the next section I argue against this approach, and in favor of extending the dual analysis of Danon (1996, 1998) as a universally-available alternation between two distinct structures made possible by UG. Then, in section 4 I argue that the choice between these two structures is subject to two universal constraints.

3 Against a uniform analysis

While the hypothesis that NNCs are syntactically uniform crosslinguistically might seem to be the null hypothesis, careful examination of the facts shows that there are good reasons to doubt this hypothesis. In this section, I present data that supports the alternative hypothesis that there are at least two distinct types of NNCs made possible by UG. Specifically, I propose that some numerals are heads that select a projection of the noun, while others are maximal projections that occupy a specifier position; schematically, a phrase like *three cats* could therefore, in different languages, have either the structure in (4a) or the one in (4b):[5]

(4) a.

$$
\begin{array}{c}
X' \\
\diagup \diagdown \\
X \quad\ \ YP \\
| \qquad \diagup\diagdown \\
\text{three} \quad \text{cats}
\end{array}
$$

(head-complement)

---

[5]If the XP in (4a) is seen as a functional (or semi-functional) projection, then the lexical head of YP is still the 'head' of the noun phrase in the usual sense of an extended projection. This is, to a large extent, an elaboration on a central theoretical issue that arose with the introduction of the DP hypothesis (Abney, 1987; Hellan, 1986); in what follows I will assume that the choice between (4a) and (4b) raises no major issues regarding headedness per se. As will be discussed throughout this paper, however, numerals in some cases show some properties of lexical, rather than functional, heads. The discussion that follows abstracts away from this issue, with the hope that the lexical/functional behavior of numerals can be reduced to the factors to be discussed rather than being an independent parameter of variation.

b.

```
            YP
           /  \
         XP    Y′
        /      /
     three   cats
```
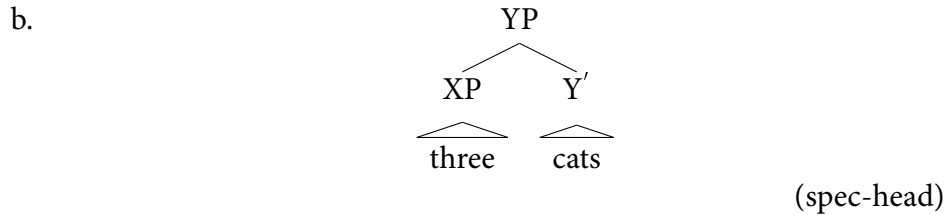
(spec-head)

In what follows, I refer to these two types as the *head-complement construction* and the *spec-head construction*, reflecting the two proposed syntactic relations possible between a numeral and a noun. The choice of construction is argued to vary both cross-linguistically and language-internally, with some languages allowing both options. My proposal is that while UG makes these two options universally available, various factors that have to do with case and grammatical number (to be discussed in section 4) may 'conspire' to constrain the choice in different ways.

This section develops the hypothesis that both structures (4) are indeed possible. In sections 3.1–3.4, I provide morphosyntactic evidence for the existence of two kinds of NNCs. Sections 3.5–3.6 then discuss some issues that this proposed syntactic variability raises for the syntax-semantics interface.

3.1   Languages with two surface types of NNCs

Perhaps the most direct piece of evidence against a uniform analysis of NNCs comes from languages that manifest multiple NNC types at the level of overt morphosyntax. As mentioned above, one such language is Hebrew, in which cardinal numerals may appear in two distinct forms, which I refer to as *free* and *bound*. Free numerals, as shown in (1a) above, are used when no overt noun is present, as well as with indefinite nouns (with some exceptions to be discussed below); while bound numerals, which have the same morphophonological properties as heads of construct state nominals, are used mostly with definite nouns, as in (1b).

If, as I claim, free and bound numerals differ in the syntactic position that they occupy, one immediate prediction is that they will not be able to be coordinated. This prediction is borne out. The numeral 2, unlike most other simple numerals, is used in the *bound* form *šney* even with indefinite nouns; trying to coordinate this numeral with a *free* numeral like *šloša* ('three') preceding an indefinite noun results in ungrammaticality:

(5)   * *šney*   o *šloša*   *sfarim*
       two(BOUND) or three(FREE) books

Instead, either the *free* form of the numeral 2 must be coordinated with the free form of 3, as in (6a); or else the entire Num+N has to be coordinated, with ellipsis of the noun in the second coordinate, as in (6b):

(6)   a.   *šnayim*   o *šloša*   *sfarim*
           two(FREE) or three(FREE) books

    b.   *šney   sfarim o šloša*
           two(BOUND) books or three(FREE)

Crucially, the ungrammaticality in (5) is not due to a general incompatibility of bound numerals with coordination, as a coordination of two bound numerals *is* grammatical, at least in informal speech.[6] This is possible when the coordinated numerals precede a definite-marked noun, an environment where only bound numerals are possible:[7]

(7)   *šney   o šlošet   ha-sfarim*
       two(BOUND) or three(BOUND) the-books
       'the two or three books'

Thus, (5) does not seem to be due to merely morphophonological factors.

    In general, bound numerals, as opposed to free numerals, display many properties characteristic of heads of construct state. Like other construct state heads, and unlike free numerals, bound numerals cannot be separated from the noun that they precede, as shown in (8);[8] and they cannot be part of a syntactically complex 'numeral phrase', as shown in (9).

(8)   a.   *šloša   ve mašehu   sfarim*
           three(FREE) and something books
           'a little over three books'

    b.   * *šney   ve mašehu   sfarim*
           two(BOUND) and something books

(9)   a.   *ben   exad   le šloša   sfarim*
           between one(FREE) to three(FREE) books
           'between one and three books'

    b.   * *ben   exad   le šney   sfarim*
           between one(FREE) to two(BOUND) books

---

[6]Prescriptive Hebrew grammars prohibit a coordination of bound forms; in actual speech, however, speakers regularly use such coordinations.

[7]Trying to construct a coordination of *šney* with another bound numeral preceding an *indefinite* noun is somewhat tricky due to the fact that the only other numerals that allow the bound form with indefinites are vague pluralized numerals like *esrot* ('tens'), discussed in section 3.4; a coordination like *šney o esrot sfarim* ('two or tens of books') sounds highly unnatural, perhaps for pragmatic reasons, and it is hard to elicit reliable grammaticality judgments for such cases.

[8]Some speakers judge (8a) to be somewhat colloquial, and have a preference for placing the modifier after the noun; but even these speakers agree that there is a striking contrast in grammaticality between (8a) and (8b).

Note that the 'modifier' in (8) seems structurally like a coordination; unfortunately, the contrast shown in these examples cannot be reproduced with 'true' modifiers such as adjectives, as no such modifiers can appear immediately after a numeral, not even in the free state.

Finally, like other construct state heads, bound numerals do not allow extraction of the noun, again in contrast to free numerals:

(10)   a.   *ze haya dubim še-ra'iti   šloša.*
            it was bears that-saw.1SG three(FREE)

            'It was bears that I saw three.'

       b.   * *ze haya dubim še-ra'iti   šney.*
            it was bears that-saw.1SG two(BOUND)

Thus, the conclusion at this point is that Hebrew shows quite clearly that analyzing all numerals in the same way cannot be correct. Given the fact that all the properties of bound numerals discussed above are typical properties of heads of construct state, the null hypothesis is that the similarity to construct heads is not only on the surface.[9] Since the influential work of Ritter (1991), the dominant view has been that nouns in the construct state are heads that precede an embedded nominal projection.[10] I propose that an analysis along the same lines also applies to bound numerals, i.e. that these are heads followed by an embedded nominal projection; some of the specifics of this proposal will be discussed later on. As to free numerals, I propose that these occupy a specifier position, which is compatible with the facts illustrated in (8)–(9). We later return to the Hebrew alternation in the context of multiplicative numerals.

In addition to languages like Hebrew, where the numeral itself alternates between two morphological forms, there are languages like Welsh, in which the same morphological form of the numeral can be used in one of two distinct syntactic frames when combined with a noun. Thus, Welsh NNCs with numerals up to 10 may have either a 'simple' or a 'prepositional' structure, where the latter resembles a partitive except for the lack of an embedded definite article (Hurford, 2003; Mittendorf and Sadler, 2005; Sadler, 2000):

(11)   a.   *naw ddyn*
            nine man
            'nine men'

       b.   *naw o ddynion*
            nine of men
            'nine men'   (Welsh; from Hurford 2003)

---

[9]Borer (2005:215), in contrast, stipulates that the bound form *šney* in an indefinite phrase does *not* form a construct state and is syntactically identical to free numerals in indefinite noun phrases, despite having the morphology typical of construct state heads. As this proposal is not backed by any independent evidence, and in light of the data discussed above, I don't see any compelling reason to assume such a morphology-syntax mismatch in the case of the bound numeral *šney*.

[10]The full details of the analysis, which possibly involve multiple movements, will not be discussed here; see Ritter (1991) and much subsequent work.

Somewhat surprisingly, there seems to be little or no semantic difference between these two construction types (Hurford, 2003; Mittendorf and Sadler, 2005, 2010);[11] most notably, (11b) is *not* semantically partitive, i.e. it is not equivalent to the English partitive *nine of the men*. Thus, Welsh provides very clear and direct evidence that there are two distinct means of syntactic composition that can give rise to what looks like the same semantic interpretation for NNCs.

3.2   Case

One of the clearest pieces of syntactic evidence in favor of analyzing numerals as heads involves morphological case. As discussed by IM, in many languages the case of a noun in a NNC depends on the numeral. As shown by Nelson and Toivonen (2000), in Inari Sami a noun following a numeral between 2–6 bears accusative case, while a noun following a higher numeral is partitive. Other languages in which the case of the noun depends on the numeral include Russian (and many other Slavic languages) and Standard Arabic (Shlonsky, 2004; Zabbal, 2005). Under the assumption that case is assigned by a head to a maximal projection under local c-command, these facts seem to rule out any analysis in which the maximal projection of the numeral occupies a specifier or adjunct position within the extended projection of the noun, and support an analysis in which a projection of the noun is a complement of the numeral.[12] Thus, the works cited above argue for the hypothesis that numerals in those languages are heads.

Unfortunately for this analysis, many other languages display case patterns that point in the opposite direction and argue against viewing the numeral as the head. Perhaps the most common pattern is that the case of the noun is not affected by the presence or absence of a numeral, which may either show case concord with the noun, or lack any case morphology. This is illustrated in the following examples,

---

[11]Mittendorf and Sadler (2010) observe that in NNCs containing a coordination of two nouns, a distributive reading where the numeral quantifies over the set denoted by the coordination is possible only in the prepositional structure. One possible hypothesis, which cannot be pursued here for reasons of space, is that this difference might be a side effect of the difference between the two NNC types in terms of hierarchical structure, where only the prepositional NNC involves a numeral that c-commands a number feature and/or a NumP projection. Other than this, the two constructions are apparently synonymous. A question raised by an anonymous reviewer is why should a language allow two different constructions for expressing the same meaning; I leave this as an open question.

[12]But see Bailyn (2004) and Franks (1994), who assume that Russian numerals that seem to assign genitive case are in fact specifiers of a functional projection in which the real case assigner is the abstract functional head. For most of the purposes of the current discussion, we can group such 'proxy head' approaches together with approaches that treat the numeral itself as the head. As pointed out to me by an anonymous reviewer, however, analyzing such Russian numerals as specifiers may be somewhat incompatible with the generalizations and the analysis proposed in section 4.

from Icelandic and Eastern Armenian, respectively:[13]

(12) *Ég þekki þrjá   íslenska   málfræðinga.*
      I   know three(ACC) Icelandic(ACC) linguists(ACC)
      'I know three Icelandic linguists.'   (Icelandic; Þráinsson, 2007:103)

(13) *Menq   tv-el   enq ayd harc-ə   tasə masnaget-i.*
      We(NOM) give-PF.PART AUX that question-DEF.ACC ten specialist-DAT
      'We gave that question to ten specialists.'   (Eastern Armenian)

The prediction of IM's analysis is that facts like this should not be possible, as they would involve non-local case assignment into an embedded DP without the expected intervention of the higher DP. At a minimum, these kinds of facts would require some mechanism of 'case percolation'. No such mechanism is needed if these are not recursive head-complement NNCs, but simple DPs where a projection of the numeral occupies a specifier position.

A further case-related complication with IM's uniform head-complement analysis is that it assumes that in all languages, numerals are case-assigning nouns. However, for languages like English, the standard assumption since the birth of GB Case Theory has been that nouns and adjectives do not have the ability to assign case; it is therefore not clear how English numerals could assign case in a head-complement NNC. This issue is noted by IM themselves, who stipulate that cardinals in English are 'exceptional nouns' that have the ability to assign case. The alternative analysis, according to which not all NNCs have the head-complement structure, does not require such a stipulation; under this approach, numerals that do not occupy a head position and that do not c-command an embedded nominal projection are simply not case assigners. We return to this in section 4.2.

3.3 Word order

Indirect arguments in favor of viewing numerals as heads have sometimes been given based on word order restrictions between numerals, nouns, adjectives and other elements within the noun phrase, which have often been argued to follow from an interaction between the cartography of functional projections and general principles of Head Movement. In the case of numerals, Borer (2005) and Shlonsky (2004) use this kind of reasoning to argue for a head analysis of most numerals in Semitic languages. At the same time, word order has also been central in the development of cartographic approaches in which specifiers are ordered based on a rigid hierarchy of abstract functional projections; see e.g. Cinque (2005).

---

[13]While the Icelandic example in (12) might suggest that numerals in this language are adjectives, there are some syntactic properties that clearly distinguish Icelandic numerals from adjectives; in particular, it has been noted by Sigurðsson (1993) and Þráinsson (2007), among others, that when the noun is marked by the definite article, numerals, but not adjectives, must follow the noun.

At a much more basic level, however, word order facts seem to pose a rather direct problem to the uniform analysis of numerals as heads. Under IM's analysis, there is a clear prediction that in the unmarked case, numerals should precede nouns in head-initial languages, and follow nouns in head-final languages. Thus, the order of numeral and noun is expected under this approach to match the order in a genitive construction, as illustrated in the Supyire examples in (14), taken from Donohue (2005):

(14)   a.   *cyèe   kɛ́*
            women ten:POSSESSED.TONE
            'ten women'

       b.   *kàn-he   mɛ̀-gé*
            village-DEF name-DEF:POSSESSED.TONE
            'the village's name'   (Supyire; from Donohue, 2005)

The prediction that this should be the general pattern, however, turns out to be false. In a large number of head-final languages, such as Amharic, Basque, Hindi, Persian and Turkish, numerals systematically *precede* nouns. According to Dryer (1992), "the two orders of numeral and noun are equally common among OV languages". Furthermore, in many head-initial languages, such as Swahili and Thai, numerals follow nouns. These facts, and especially the patterns shown by head-final languages, seem quite problematic for a uniform head-complement analysis, which would predict a vast majority of OV languages to have N-Num word order. Compared, for instance, to the overwhelming correlation of VO/OV order with Aux-V/V-Aux or with the choice between prepositions/postpositions, it seems problematic to explain the order of numeral and noun under the assumption that this is uniformly a head-complement relation.

This is not to say that there is no correlation between the dominant word order of a language and the order between numerals and nouns. In partitive construction, word order indeed supports taking numerals as heads. This has been noted in Greenberg (1978:284), whose generalization 45 states that "If a language has NG order in the possessive construction, it has QN order in the partitive construction"; Greenberg illustrates this with the English NG *house of the man* which correlates with the QN order in *three of the oranges*. But in head-final languages we often find a contrast in word order between numerals in partitive constructions, which follow the noun, and numerals in non-partitive NNCs, which precede the noun. This is illustrated below for Basque:

(15)   a.   *liburu-eta-tik   bost*
            book-D.PL-ABL five
            'five of the books'

    b.   *bost liburu*
           five book
           'five books'   (Basque)

Thus, under the assumption that languages tend to be consistently either head-initial or head-final, we see that word order does not support a uniform analysis of numerals; specifically, in contrast with the partitive construction, where word order indeed supports analyzing the numeral as the head (contra the analysis in Jackendoff, 1977), word order in non-partitive NNCs in some languages supports viewing the numeral as a specifier and not as a head. Overall, word order does not provide evidence for a unitary analysis of numerals, but instead, it seems to show that certain numerals behave like heads, and others do not.

3.4   Complex numerals

Going beyond the analysis of simple numerals, it might seem that various classes of complex numerals, if composed syntactically, pose a non-trivial challenge to the analysis of numerals as heads. In this context, IM argue that a head-complement analysis of NNCs can actually account quite easily for the distribution of case morphology in complex 'multiplicative' numerals (like *three hundred* or *five thousand*) in Russian and Finnish. Specifically, they propose that in such complex numerals the head-complement analysis applies recursively, such that *three thousand books* has roughly the structure [three [thousand [books]]]. Thus, according to this view, no special machinery is needed in order to account for the class of multiplicative numerals, which is an important advantage of this analysis.[14]

    As appealing as this analysis might seem, the question is to what extent it is supported by the facts of multiplicative numerals in other languages. One language that allows us to test the predictions of this analysis is Hebrew, in which, as was shown in section 3.1, numerals can have one of two forms. The prediction of IM's recursive analysis of multiplicative numerals is that the choice between the free and bound forms for each part of a complex numeral would be subject to the same selectional restrictions on what that part precedes as if it was a simple numeral, as the structure is assumed to be recursively right-branching. This prediction, however, is not borne out.

    Consider, for instance, an indefinite of the type *3000 N*. A structure of [3 [1000(s) [N]]] would lead us to expect the *free* form of '3', as with other indefinites using this numeral: (16a); this should then be followed by the bound form of '1000s', as when '1000s' alone is combined with a noun: (16c). Alternatively, we might perhaps ex-

---

[14]In this paper I will not discuss 'additive' numerals like *a hundred and three*, which have been analyzed both within the numerals-as-heads approach (Ionin and Matushansky, 2006) and within the numerals-as-specifiers approach (Zabbal, 2005).

pect the free form of '1000', as when '1000' alone is combined with a noun: (16d).[15].
However, all these predictions turns out to be false, as shown in (17):

(16)   a.   *šloša   sfarim*
             three(FREE) books

       b.   * *šlošet   sfarim*
             three(BOUND) books

       c.   *alfey   sfarim*
             thousands(BOUND) books

       d.   *elef   sfarim*
             thousand(FREE) books

(17)   a.   * *šloša   [alfey   sfarim]*
             three(FREE) thousands(BOUND) books

       b.   * *šloša   [elef   sfarim]*
             three(FREE) thousand(FREE) books

Instead, the pluralized decimal '1000s' must be in the *free* form, regardless of whether
a noun follows or not, while the non-decimal part '3' is in the *bound* form, as shown
in (18a):

(18)   a.   *šlošet   alafim   sfarim*
             three(BOUND) thousands(FREE) books

       b.   * *alafim   sfarim*
             thousands(FREE) books

Thus, according to IM's analysis, the grammatical DP in (18a) contains an embed-
ded XP, *alafim sfarim*, which is otherwise ungrammatical, as shown in (18b).[16] Fur-
thermore, the fact that the numeral 3 in (18a) is used in its bound form is also
unexpected, given that it is followed by an indefinite. The alternative is to assume
that the complex numeral itself forms a constituent that excludes the noun, *[šlošet
alafim] sfarim*, which might also account for the use of the free form *alafim*, as this
part of the complex numeral is 'intransitive' in this case (since what combines with
the noun is the entire complex numeral rather than its decimal part). This, in turn,
means that numerals cannot be uniformly analyzed as heads that select a nomi-
nal complement, as here we are dealing with a numeral that is syntactically more
complex than a single head.[17]

---

[15] We return to the contrast between '1000' and '1000s' in section 4.

[16] As pointed out by an anonymous reviewer, a similar (but somewhat weaker) argument could
be made for English, where the phrase *three hundred books* would under IM's analysis be claimed
to contain an embedded DP *hundred books* which is ungrammatical on its own.

[17] Another possibility might be that the whole multiplicative numeral is formed in the lexicon and

Note furthermore that lexical idiosyncrasies in the form of multiplicative numerals in Hebrew also argue against IM's derivation of such numerals. First, in multiples of 100 or 1,000, the simple numeral that precedes the decimal part is used in the bound form, and the decimal is morphologically plural (e.g., *meot* 'hundreds'); while in multiples of 1,000,000, the simple numeral is in the free form, and the decimal is morphologically singular:[18]

(19) a. *šloš*   *meot*   *sfarim*
          three(BOUND,FEM) hundreds(FEM books(MASC)
          'three hundred books'

      b. *šlošet*   *alafim*   *sfarim*
          three(BOUND,MASC) thousands(MASC) books(MASC)
          'three thousand books'

      c. *šloša*   *milyon*   *sfarim*
          three(FREE,MASC) million(MASC) books(MASC)
          'three million books'   (Hebrew)

Furthermore, certain multiplicative numerals contain bound forms that are never used in other NNCs; for instance, the bound forms of the numerals 7 and 9 in the complex numerals 700 and 900 are not the regular bound forms *ševa* and *teyša*, but rather the reduced forms *šva* and *tša*, which are unique to complex numerals:

(20) *šva*   / *tša*   *meʻot*   *sfarim*
      seven(BOUND)   nine(BOUND) hundreds books
      'seven/nine hundred books'

(21) a. * *šva*   / *tša*   *ha-dugmaʻot*
             seven(BOUND)   nine(BOUND) the-examples

      b. *ševa*   / *teyša*   *ha-dugmaʻot*
          seven(BOUND)   nine(BOUND) the-examples
          'the seven/nine examples'   (Hebrew)

While capturing such restrictions and idiosyncrasies is straightforward in an analysis of multiplicative numerals as (lexically frozen) constituents, it is much harder to do in an analysis where the simple numeral does not even form a constituent with the decimal part, as in IM's analysis. The main difficulty seems to be in the

---

occupies a single head position in the syntactic representation. This would go against IM's analysis of multiplicative numerals as formed syntactically, thus reducing the appeal of the recursive head-complement analysis as its empirical coverage is reduced. It is beyond the scope of this paper to discuss the evidence for and against a syntactic derivation of complex numerals; in what follows, I concentrate on the consequences if we adopt a syntactic derivation.

[18] It is unclear to me whether the choice of free or bound form for the first part and the choice of morphological singular or plural for the decimal part are independent of each other or not.

directionality of the dependency: what the examples above illustrate is that in Hebrew multiplicative numerals, it is the decimal part that selects a particular form for the part of the numeral which, in IM's analysis, c-commands it (e.g., *me'ot* in (20) selects the form *šva*; selection in the opposite direction would not be enough to exclude *me'ot* from following the regular form *ševa*). This is the opposite of typical selection relations, where a head selects either a complement or a specifier (e.g., a verb selecting a particular type of argument), but not the other way around.

The overall conclusion from Hebrew multiplicative numerals is that they don't seem to support an analysis as heads selecting a projection of the noun as their complement. The significance of this goes beyond the narrow issue of how multiplicative numerals in Hebrew are formed: these facts show that UG must make a non-head position available to numerals, even in languages where some numerals do seem to show properties of heads.

Looking at other kinds of complex numerals, it seems unavoidable that at least some of these cannot simply be analyzed as heads. Corver and Zwarts (2006), for instance, argue that 'prepositional numerals' like *around ten* or *between ten and twenty* are single constituents, which serve as specifiers of a functional projection within the noun phrase. They provide evidence against the alternative view, which might be compatible with the analysis of numerals as heads, according to which the preposition combines with a phrase consisting of the numeral and the noun. Therefore, if we must assume that at least prepositional numerals occupy a specifier position, then there is no immediate reason to reject the possibility of a specifier analysis for some other numerals as well. (See also Danon 1996, 1998 for similar claims, focusing mostly on numerals in Hebrew).

Overall, it looks like even though IM might be correct in arguing that a recursive head-complement analysis provides the best account of multiplicative numerals in some languages, it seems very problematic to try and extend this kind of analysis to all complex numerals; and if UG provides for the option of some complex numerals to occupy a specifier position, there is no good reason not to allow some simple numerals to occupy the same position as well.

3.5   Prepositional NNCs and partitivity

So far, I argued that there are two types of cardinal NNCs in natural language. One issue that arises at this stage is the relation between these two types of NNCs and partitive constructions. This is central since, at least on the surface, there is often a striking similarity between partitive and non-partitive NNCs. In particular, some of the NNCs discussed so far, as well as many of those discussed in the section 4, either employ the same prepositional element as in partitive constructions or involve the same case marking. While it is quite uncontroversial that partitives are semantically distinct from non-partitive numeral constructions, the question is whether this distinction is (necessarily) encoded in the structural position of the

numeral. In this section I wish to briefly comment on this issue.

Some early transformational analyses proposed that partitive and non-partitive NNCs are derived from the same underlying structure; this view was criticized in Selkirk (1977), who showed that there are significant differences between these two kinds of NNCs that are not expected if both are derived from the same underlying structure. Nevertheless, in many languages, essentially the same kind of head-complement structure used in partitives is also used in some non-partitive constructions as well. One particularly clear example is Serbo-Croatian, in which, according to Arsenijević (2006), partitive constructions are identical in surface form to non-partitive NNCs (which will be referred to as 'simple NNCs'); thus, for instance, the following example (from Arsenijević, 2006) is in fact ambiguous between these two readings:

(22)   *Ivan uze   šest miš-eva.*
        Ivan took six mice.GEN

        'Ivan took six mice.' / 'Ivan took six of the mice.'   (Serbo-Croatian;
        Arsenijević, 2006)

It could be that there is a syntactic difference between the partitive reading of (22) and the non-partitive one (e.g., the partitive structure might perhaps contain a phonetically null determiner preceding *miševa*); nevertheless, the identity with respect to case suggests that the syntactic difference, if such exists, is rather minimal.

At the opposite side of the spectrum, it was shown in section 3.1 that there are languages like Welsh, where a non-partitive reading can be achieved using a partitive-like prepositional construction:

(23)   *naw o ddynion*
        nine of men

        'nine men'   (Welsh; from Hurford 2003)

To a limited extent, English also has a partitive-like construction that is used with certain non-partitive NNCs. Specifically, the same form that is used in partitive constructions (*n of…*) is also used with what I refer to as *pluralized numerals*, as in (24b) below (Kayne, 2010);[19] furthermore, like true partitives, partitive-like NNCs allow a limited amount of recursion, as shown in (25):[20]

---

[19] As pointed out by an anonymous reviewer, another kind of construction that bears a similarity to (24b) is the pseudopartitive formed with measure phrases, as in *three kilos of rice*. However, such measure phrases contain both a numeral and a measure noun preceding *of*, unlike the ungrammatical *three hundreds of books*. In contrast with the recent analyses that argue for a structure in which measure phrases occupy a specifier position (see e.g. Landman 2004 and Rothstein 2010), I later argue for an analysis in which the numeral occupies a head position in NNCs like those in (24b). Whether the same kinds of arguments used to justify my analysis can also be applied to the analysis of measure phrases remains as an open question.

[20] This was pointed out to me by an anonymous reviewer.

(24)   a.   five of the books

        b.   hundreds of books

(25)   a.   half of all of the books

        b.   hundreds of thousands of books

It seems that on the surface the only two differences between the partitive in (24a) and the non-partitive in (24b) are in whether or not what follows *of* is a definite/specific DP, and in the form of the numeral itself ('bare' numeral versus plural-marked numeral). While these two differences are probably not 'innocent', I believe it would be wrong to assume that these are two different structures and ignore the similarities between them.[21] We return to the issue of why *of* is needed in cases like (24b) in section 4. For now, the main point is that with respect to the numeral's syntactic position relative to the noun, it is not clear that there is a structural configuration that is unique to partitives; given the surface similarity, the null hypothesis is that the numeral occupies the same type of position in both (24a) and (24b).

Unlike Arsenijević (2006), who takes facts like those in (22) to be characteristic of NNCs in general, I do not claim that all NNCs (or even that all prepositional NNCs) have an *identical* syntactic structure to semantically partitive DPs. Due to space limitations, the remainder of this paper will focus only on non-partitive NNCs, keeping in mind that the fact that these often employ a syntactic structure that bears a strong similarity to that of partitives is probably not a coincidence. What is central to the proposed analysis is that both partitive and partitive-like NNCs involve some sort of DP-internal case-assignment mechanism, which I take as evidence for a head-complement construction and which is not used in specifier-head NNCs. We return to the issue of case in section 4.2.

3.6   Intermediate summary

The conclusion from the discussion so far is that natural language allows both spec-head NNCs and head-complement ones, with both often found within the same language. Of the languages discussed so far (and to the extent that the data given is representative), the patterns suggested by the data surveyed seem to be the following:

---

[21]A very interesting hypothesis raised by an anonymous reviewer is that prepositional NNCs like (24b) are not merely similar to partitives, but that they actually *are* partitives. Under this approach, prepositional NNCs differ from 'ordinary' partitives like (24a) in that instead of quantifying over a specific set, they quantify over instantiations of the kind denoted by the plural that follows *of*. This approach would have the desirable consequence of unifying the two structures in (24) despite the absence of an article in (24b), thus reducing the overall number of distinct construction types that have to be assumed.

As this kind of analysis must be supported by a non-trivial elaboration on the semantics of quantification in general and quantification over kinds in particular, it is beyond the scope of the current paper to pursue this approach. I therefore leave this as a topic for further research.

- Hebrew: head-complement with bound numerals; spec-head with free (or syntactically complex) numerals

- Welsh: head-complement in prepositional constructions; spec-head otherwise

- Russian and Serbo-Croatian; Supyire: head-complement in all NNCs

- English: head-complement with pluralized numerals (using *of*); spec-head otherwise

- Basque, Icelandic, Armenian: spec-head in non-partitive NNCs[22] [23]

Many details of this typology are still in need of a more detailed and explicit analysis; in the next section, we discuss the factors affecting the choice of NNC type and details of the internal structure of each construction.

Before elaborating on the syntactic analysis, it should be noted that the dual analysis argued for so far entails that natural language makes possible two distinct mechanisms for expressing essentially the same semantic relation between a numeral and a noun. At the level of syntax per se, this is expected under a Minimalist perspective in which Merge is not constrained by construction-specific rules and in which neither the numeral nor the noun is lexically designated as the projecting head. As long as no other syntactic principle is violated and a compositional semantics is possible, both structures should be allowed (see also Borer 2005 for a similar line of reasoning). This theoretical prediction, which arises from the hypothesis that there are two distinct means of forming a NNC, receives straightforward support from the Welsh facts noted in §3.1 and illustrated in (11): the two NNC types found in Welsh, which are undoubtedly syntactically distinct, nevertheless give rise to the same interpretation. Similarly, the Hebrew alternation discussed in §3.1 between the Hebrew numerals 2 and 3 (which differ in the kind of NNC that they appear in when the noun is indefinite) shows that despite the syntactic difference between the NNCs formed with these two numerals, both give rise to the same kind of weak indefinite cardinal interpretation.

In what follows, I assume that UG indeed makes it possible to derive the same kind of interpretation for both head-complement and spec-head NNCs. The fact that most languages do not allow free alternation between two NNC types thus has to be accounted for; this is discussed at length in section 4.

---

[22] Partitives seem to involve a head-complement construction crosslinguistically.

[23] Given the very limited amount of data provided for these languages, the conclusion in this case is highly tentative.

4 Constraining the choice of configuration

4.1 The role of number features

Under the proposal put forth in the previous section, an important question is what determines which of the two constructions will be used when. One possible hypothesis might be that the choice of structure is arbitrary and language-specific. Under this approach, we would not expect to see any systematic patterns governing the alternation between the two structures.

There is, however, an interesting observation that suggests that the choice between these two structures is not random: in a wide variety of typologically unrelated languages, there is a strong correlation between morphological number marking on the numeral and the form of NNC being used, such that plural-marked numerals systematically show different syntactic properties than numerals that lack number morphology. This is discussed in section 4.1.1. Furthermore, in some languages there seem to be interactions between the choice of NNC and the presence of overt number marking on the noun; this is discussed in section 4.1.2. The main goal of this section, developed in section 4.1.3, is to provide a systematic account of how number marking restricts the choice of NNC.

4.1.1 Number on numerals

In many languages, overt plural marking on the numeral itself has an effect not only on the interpretation of the noun phrase but also on its syntactic structure. This is very clear in the case of what I will refer to as PLURALIZED NUMERALS, such as *hundreds* and *thousands*. As noted for instance by Kayne (2010) and Schwarzschild (2006), pluralized numerals in English, unlike their unpluralized counterparts, require a prepositional construction containing *of*:

(26) a. (a) million books
      b. millions *(of) books

For other languages, the contrast between pluralized and non-pluralized numerals is even sharper. In Hungarian, a non-pluralized numeral is used in a simple NNC containing the numeral followed by a singular noun, as in (27); pluralized numerals like '100s', on the other hand, give rise to a possessive-like construction, illustrated in (28), which shows the same morphology and word order as in true possessive constructions, illustrated in (29):

(27) *száz   gyerek*
      hundred child(SG)
      'a hundred children

(28) *gyerek-ek száz-a-i*
      child-PL hundred-POSS-PL

'hundreds of children'

(29)   *a   gyerek-ek ház-a-i*
       the child-PL   house-POSS-PL

       'the children's houses'   (Hungarian)

The construction in (28) thus differs from the simple NNC in word order, in having possessive morphology on the numeral (the same as on a possessed noun), and in having plural morphology on the noun. The similarity to possessives suggests that pluralized numerals in Hungarian involve a recursive nominal structure in which a projection of the noun is embedded within an extended projection of the numeral; this kind of analysis will later be argued to be characteristic of pluralized numerals in many languages.

Some other languages allow pluralized numerals with nouns only if the numerals are somehow 'denominalized'. In Turkish, for instance, a pluralized numeral must take the adverbial suffix *-ca* in order to combine with a noun:[24]

(30)   a.   *on   kitap*
            ten book

            'ten books'

       b.   * *on-lar   kitap*
            ten-PL book

       c.   *on-lar-ca   kitap*
            ten-PL-ADV book

            'tens of books'   (Turkish)

Interestingly, Turkish *does* allow pluralized numerals without the suffix *-ca*, as long as these are used in isolation and not as part of a NNC; this is common mostly with larger pluralized numerals, as in (31). The same pattern is also observed in Hungarian.

(31)   *bin-ler   sokaklar-i   doldurdu.*
       thousand-PL streets-ACC filled

       'Thousands filled the streets.'   (Turkish)

While it is clear from the above discussion that pluralized numerals give rise to different structures than those used with other numerals, it is still not clear whether this follows from their morphosyntactic properties or from their semantics. Strong support for the former option comes from Modern Hebrew. Like English, Hebrew uses a different structure for pluralized numerals versus their unpluralized forms.

---

[24]This suffix is used also in manner adverbs; it could be translated roughly as 'in the way of-'.

While most unpluralized numerals take the free form in the presence of an indefinite noun, pluralized ones can only appear in the bound form when followed by a noun (definite or indefinite):[25]

(32)   a.   *esrot   / *asarot   sfarim*
        tens(BOUND)   tens(FREE) books
        'tens of books'

    b.   *asara   / *aseret   (sfarim)*
        ten(FREE)   ten(BOUND) books
        'ten books'   (Hebrew)

What makes the Hebrew facts interesting is the role that plural morphology plays. As noted in Danon (1996, 1998), the numeral *šnayim* ('two') shows the same distribution as pluralized numerals; unlike all other simple numerals (and like pluralized numerals), 2 must be used in the bound form *šney*, and not in the free form, whenever combined with a noun:

(33)   *šney   / *šnayim   sfarim*
    two(BOUND)   two(FREE) books
    'two books'   (Hebrew)

Importantly, what this numeral has in common with pluralized numerals, in addition to having the same syntactic distribution, is that it bears plural morphology: the suffix *-ayim* is one of the several allomorphs of the plural suffix, and is used also with nouns (mostly body parts) such as *yadayim* ('hands') and *šinayim* ('teeth').[26] The contrast in Hebrew between the numeral 2 and other simple numerals thus illustrates very clearly that morphosyntactic plurality on the numeral has an effect on the choice of NNC, even where it is not clear whether this plural marking has any effect on the way in which the numeral is interpreted. A similar conclusion is suggested by Kayne (2010), who provides the following contrast from Romanian:[27]

---

[25]Some speakers also accept (32a) with the free form *asarot*, but this is probably due to the nature of the morpho-phonological difference between the free and the bound form in this case, which in colloquial speech is not entirely clear and productive; judgments are much clearer in the case of numerals like 1000s, where there is no doubt that only the bound form *alfey*, and not the free form *alafim*, can combine with a noun.

[26]Historically, this suffix was once a dual suffix; in Modern Hebrew, however, it is an idiosyncratic plural form and does not imply a cardinality of exactly 2, and hence a phrase such as *šmone raglayim* ('eight legs') is completely well-formed (see also Borer, 2005). Nouns marked with *-ayim* also trigger regular plural agreement; this is true even for the small number of frozen lexical items, such as *yomayim* ('two days') and *pa'amayim* ('twice'), which do have the meaning of exactly 2. In general, there is no evidence for a syntactically-relevant notion of dual in Modern Hebrew.

[27]Kayne, however, goes on to propose that the source of the contrast is not the plural morphology but the presence of an unpronounced nominal suffix, which he claims appears on the second element (the multiplicand) of a multiplicative numeral.

(34)   a.   *zece caiete*
           ten notebooks

       b.   *trei-zeci   de caiete*
           three-tens of notebooks

          'thirty notebooks'   (Romanian; from Kayne, 2010)

In this case, presence of plural morphology on the multiplicative numeral in (34b) forces the use of a prepositional structure, in contrast with the NNC in (34a) in which the numeral has no plural morphology. As in the Hebrew examples discussed above (and unlike the English *hundred/hundreds* contrast), the syntactic alternation here cannot be reduced to a semantic contrast like that between exact and approximative numerals.

One possible objection to the above is that the numerals 20, 30, 40 etc in Hebrew also look as if they are morphologically plural, as they end with the suffix *-im*, the most common masculine plural suffix, yet they do not pattern with pluralized numerals: when followed by a noun, these numerals do not appear in a bound form but in the same form that is used when no noun is present:

(35)   *šlošim (sfarim)*
      thirty books

Unlike other numerals, however, these numerals simply do not have a distinct bound form; hence, the same form is also used in front of definite nouns, an environment where other numerals require the bound form:[28]

(36)   a.   *šlošim ha-mištatfim*
           thirty the-participants

          'the thirty participants'

       b.   *šlošet   / *šloša   ha-mištatfim*
           three(BOUND)   three(FREE) the-participants

          'the three participants'   (Hebrew)

---

[28]The same is true for the numerals 200 and 2000, which end with the same *-ayim* suffix as the numeral 2, but differ from it in not having a bound form:

(i)   a.   *šnayim   — šney*
           two(FREE)   two(BOUND)

       b.   *matayim   — *matey*
           two.hundred(FREE)   two.hundred(BOUND)

       c.   *alpayim   — *alpey*
           two.thousand(FREE)   two.thousand(BOUND)

It therefore seems that numerals like 20, 30 etc in Hebrew are simply morphologically defective: their bound form just happens to be homophonous to the free form.[29] If so, then these are not really a counterexample to the observation that plural morphology on the numeral always triggers a head-complement construction.

4.1.2   Number on nouns

A well known typological fact about NNCs is that in many languages, a morphologically *singular* noun is used in the presence of a numeral (Ortmann, 2000). Some of these languages, however, also allow a *plural* noun to combine with a numeral, using a different morphosyntactic pattern than the one used in 'simple' NNCs. It has sometimes been claimed that the presence or absence of plural marking on nouns in NNCs is somehow vacuous (Ionin and Matushansky, 2006), as it seems to add nothing to the semantic interpretation of the noun phrase as a whole. However, a closer examination reveals that this cannot be entirely correct, as there are systematic regularities in the occurrence of number morphology in languages that usually omit it in NNCs.

The most common pattern is the use of plural morphology on the noun in prepositional and other head-complement constructions. In Welsh, for instance, a noun in an NNC is plural only in the prepositional construction (Hurford, 2003), as was shown in (11). A similar pattern was also illustrated in (15) for Basque partitives; and in (28) for Hungarian pluralized numerals. This pattern thus cuts across different semantic classes of NNCs: it includes ones with a 'simple' cardinal reading (Welsh), a partitive reading (Basque), and a pluralized numeral reading (Hungarian). This is summarized in the following table:

(37)

Language	N[sg]	N[pl]
Welsh	simple cardinal	prepositional cardinal: (11b)
Hungarian	simple cardinal	pluralized numeral ('possessive'): (28)
Basque	simple cardinal	partitive (ablative): (15)

The generalization is thus that even if a language disallows plural morphology on nouns in some 'simple' NNCs, it may still allow (or even require) it in complex NNCs (prepositional or those involving extra case morphology). If the NNCs in which a singular noun is used are of the spec-head type, the generalization that emerges is that in some languages, number morphology on the noun in a NNCs is allowed iff the NNC is a head-complement construction.

4.1.3   The number constraint

The generalization that emerges from the discussion above is that there is a tight correlation between the choice of NNC type and the presence or absence of number

---

[29] A reviewer also pointed out that these numerals, unlike lower ones, do not inflect for gender.

morphology. While correlations involving decimal numerals like *hundred/hundreds* have been noted before (see e.g. Kayne 2010, who refers to these as *numerical bases*), to the best of my knowledge no general principles have been proposed. I propose that underlying this generalization is the following constraint:

(38)   The number constraint: The spec-head construction is not possible if the numeral carries its own morphosyntactic number feature.

Additionally, as shown in §4.1.2, if a language disallows number morphology on the noun in spec-head NNCs, this restriction does not apply to head-complement NNCs. Taken together, the picture that emerges is that the head-complement construction allows the presence of one more number feature than the spec-head construction. In the remainder of this paper, we focus mostly on what is captured directly by the constraint in (38).

It is important to note that what is at stake here is *morphosyntactic* number, not semantic number. While numerals obviously have semantic number associated with them, it is not obvious that this is encoded as a feature that is visible at the morphosyntactic level. Thus, for instance, while it seems reasonable on semantic grounds to associate the feature PLURAL with a numeral like *ten*, there is no morphological motivation for doing so. In fact, morphology argues for the opposite view. First, since the plural in English (and many other languages) is morphologically marked, assuming a PLURAL feature on numerals that lack such plural morphology is a mere stipulation. Furthermore, if *ten* is PLURAL, then it should not be possible to add a plural suffix to it, contra to fact; and it is not clear what feature would distinguish in that case *ten* from *tens*.[30] [31]

Assuming that a numeral like *ten* is singular would also be problematic, however, as this would probably require us to assume that this feature plays no role in the interpretation of the noun phrase. I thus propose that the most reasonable assumption is that at the morphosyntactic level, simple numerals like *ten*, unlike pluralized numerals, simply lack a number feature.

Support for the lack of a number feature on non-pluralized numerals comes from subject-verb agreement in languages like Hungarian and Turkish, in which nouns in NNCs must be in the singular form. As noted in Ortmann (2000), in the

---

[30] Note that this issue arises not only with numerals that are powers of 10, which might perhaps be argued to follow from extra-linguistic properties of our counting system; even though less common, numerals like *two* and *three* can also be pluralized (e.g., *in twos*), which would under my analysis be seen as a 'true' plural.

[31] An alternative hypothesis, which might provide a way around this issue, is that *ten* is ambiguous between a [plural] numeral and a [singular] noun (which forms the basis for the [plural] *tens*). But since what we are dealing with is not merely a language-specific issue but a more general crosslinguistic one, this hypothesis amounts to stipulating the same type of lexical ambiguity in a wide variety of languages and language families, thus making such an ambiguity-based analysis less than optimal.

vast majority of such languages, NNC subjects trigger singular agreement.[32] This is illustrated for Turkish:

(39)   *üç   çocuk geldi   / *gelidiler.*
        three boy   came.SG   came.PL

        'Three boys came.'   (Turkish)

If the numeral contributed a plural feature or agreed in plurality with an abstract Num head, this agreement pattern would have been unexpected. But if we assume that the numeral has no such feature, there is nothing surprising about the ungrammaticality of plural agreement in cases like this, as there is simply no morphosyntactic plurality within the NNC for the verb to agree with.[33] 'Singular' agreement in (39) can then be seen as default agreement.

Under this view, the number constraint can be seen as a restriction against having a number feature (and not just against having plurality) on a numeral that acts as a specifier in a simple numeral-noun construction. As we saw, English allows pluralized numerals only in the prepositional construction that involves *of*; Hebrew allows numerals with plural morphology only in the bound head-complement constructions; etc. In section 4.3 we show that this constraint can be reduced to general principles of case. Apparent counterexamples to the generalization that pluralized numerals require a head-complement construction will be discussed in section 5.

4.2   The role of case

4.2.1   Case in head-complement NNCs

As discussed in section 3.2, it is perhaps the presence of morphological case that serves as the strongest piece of evidence for the recursive head-complement analysis of Ionin and Matushansky (2006); the following example, for instance, illustrates the fact that in Russian, simple numerals 5 and above assign genitive case to an embedded nominal phrase:

(40)   *šest' šagov*
        six   step-GEN.PL

        'six steps'   (Russian; Ionin and Matushansky, 2006)

Other case assignment mechanisms associated with numerals include using a 'dummy' prepositional element, such as English *of*, or the Semitic bound construct state form

---

[32]Ortmann (2000) discusses Basque as one exception, i.e., as a language where NNCs contain singular nouns but trigger plural agreement. As subject-verb agreement in Basque has been reported to be sensitive to semantic factors in other cases (Etxeberria and Etxepare, 2008, 2009), it could be that this is an instance of semantic agreement.

[33]This still leaves open, of course, the question of why the noun in NNCs like (39) cannot be plural.

in Hebrew and Arabic.[34]

But as discussed in section 3.2, not *all* numeral-noun constructions display such case patterns. As was shown in examples (12)–(13), repeated as (41)–(42) below, in many languages numerals are 'transparent' with respect to case, with their presence having no effect on the case of the noun.

(41) *Ég þekki þrjá   íslenska   málfræðinga.*
      I   know three(ACC) Icelandic(ACC) linguists(ACC)

      'I know three Icelandic linguists.'   (Icelandic; Þráinsson, 2007:103)

(42) *Menq   tv-el   enq ayd harc-ə   tasə masnaget-i.*
      We(NOM) give-PF.PART AUX that question-DEF.ACC ten specialist-DAT

      'We gave that question to ten specialists.'   (Eastern Armenian)

I will adopt IM's reasoning that where the noun carries case that is dependent on the numeral, as in Russian, this is indeed evidence for a head-complement construction; but unlike IM, I will also assume that DP-internal case assignment is not involved in all NNCs, and specifically, that no case assignment is involved in forming a spec-head construction. Below I show that this distinction can be derived from general principles of case theory.

4.2.2   The case constraint

The claim above is essentially that if the numeral (or a prepositional element like *of*) assigns case to a projection of the noun, this is a head-complement construction. Under this view, case is a sufficient condition for identifying a head-complement NNC; this is stated in (43a). Furthermore, I propose the stronger hypothesis in (43b), according to which case is not only possible, but necessary, in the head-complement construction.
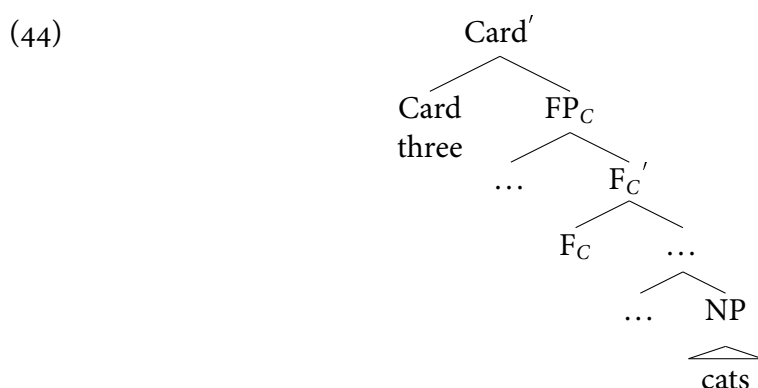
(43) The case constraint:

   a. DP-internal case assignment to a projection of the noun that excludes the numeral is possible only in the head-complement construction.

   b. The head-complement construction *must* involve DP-internal assignment of abstract case to a projection of the noun that excludes the numeral.

Thus, my proposal is that at the level of abstract case, a NNC has the head-complement structure iff it involves DP-internal case assignment.

This raises the question of why case should be so tightly related to the head-complement construction, and whether this constraint can be derived from any

---

[34]In Standard Arabic, the bound form in a construct state assigns morphological genitive case; a standard assumption in the literature on Hebrew construct states is that these involve assignment of abstract genitive; see e.g. Ritter (1991) and Shlonsky (2004).

independently motivated principle. As case is usually taken to be a feature of the extended projection of the noun (rather than of the noun itself), a first step towards deriving (43) is to argue that in a head-complement NNC, the numeral c-commands not merely NP itself but some functional projection, which dominates the noun and which is in some sense 'large enough' to require case; this is a generalization of IM's proposal, who argue for the existence of a full DP embedded below a projection of the numeral. Pending the discussion in section 4.3, where I argue that what counts as 'large enough' is minimally a NumP, this case-requiring functional projection will temporarily be labeled FP_C. Thus, the structure of a phrase like *three cats* in a language that uses a head-complement NNC would be:[35]

(44)

```
                    Card′
                   /     \
               Card       FP_C
               three     /    \
                      …       F_C′
                             /    \
                          F_C      …
                                  /   \
                                …      NP
                                       /\
                                     cats
```
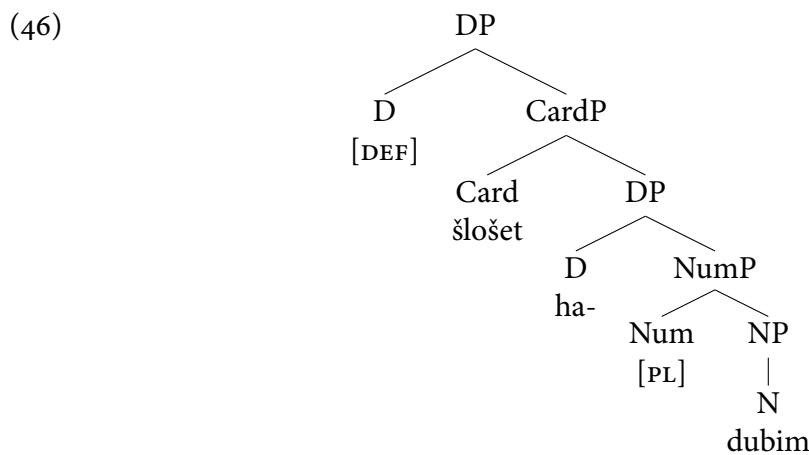
The second part of the case constraint is then nothing more than the requirement that FP_C get case, together with the assumption that in the head-complement construction the NNC as a whole (or some other projection dominating the embedded FP_C) blocks case assignment to the embedded FP_C from outside the entire noun phrase, and hence requires it to receive its case within the NNC. We return to the exact nature of this FP_C in section 4.3, where I argue that it is not necessarily a full DP but it is at least a NumP. First, however, we should establish the claim that there is in fact an embedded phrase larger than NP following the numeral.

Evidence for the claim that the embedded phrase in a head-complement NNC is larger than NP comes from languages where a numeral in this type of NNC can be followed by material that is typically assumed to occupy a relatively high position within the noun's extended projection. In Hebrew, definite NNCs involve a definite article intervening between a numeral in the bound form and the noun, as illustrated in (45) below:

(45)   *šlošet   ha-dubim*
       three(BOUND) the-bears
       'the three bears'   (Hebrew)

---

[35] In (44) and all subsequent trees, I label the category of the numeral as 'Card'.

Shlonsky (2004) discusses data similar to (45) and concludes, under the assumption that articles are heads of DP, that Hebrew must have a position for cardinal numerals above DP in addition to the typical position below DP; this seems unavoidable in a cartographic analysis that does not adopt a recursive structure for such NNCs. A recursive analysis, in contrast, does not require such a stipulation: the numeral, which might occupy a relatively low position in the overall hierarchy of functional projections, merges in (45) (recursively) with an extended projection of the noun, which also contains the definite article. The structure would thus be something like the following:[36]

(46)

```
                              DP
                          ╱        ╲
                     D              CardP
                   [DEF]          ╱       ╲
                            Card            DP
                            šlošet        ╱    ╲
                                      D          NumP
                                     ha-        ╱    ╲
                                           Num       NP
                                           [PL]       │
                                                       N
                                                     dubim
```

Note that the somewhat surprising fact that the definite article in (45) takes scope over the entire NNC is in fact expected if bound numerals involve a construct state; we can then assume that (45), like other Hebrew construct states, involves a process of 'definiteness spreading' in which the definiteness value of the embedded genitive phrase 'spreads' to the embedding phrase. The same kind of 'spreading' is illustrated in the following non-NNC construct state:

(47)   *kufsa'ot ha-karton*
       boxes   the-cardboard
       'the cardboard boxes'   (Hebrew)

---

[36]The tree in (46) is an over-simplification, as it abstracts away from many of the issues involved in the formation of the Semitic construct state. Specifically, it is quite likely that the numeral undergoes head movement into a higher position, as originally proposed in Ritter (1991), who also argued that the embedded DP occupies a specifier position rather than being a complement of the numeral; and it is also debatable whether the definite article *ha-* is indeed the head of D, or is merely a feature of N (Borer, 1999; Danon, 2008). While all these issues mean that the exact structure could be more complex than the one shown here, much of this complexity is independent of the claim made here, that NNCs like (45) provide evidence that what follows the numeral in head-complement NNCs is a nominal structure larger than just NP.

For an explicit analysis of definiteness spreading, see e.g. Borer (1999, 2005) and Danon (2008).

Another piece of evidence in support of the claim that head-complement NNCs involve a larger-than-NP embedded phrase comes from Romanian. In Romanian, certain adjectives that typically occur at the left edge of a noun phrase may come right in front of the noun in prepositional NNCs; this contrasts with non-prepositional NNCs, where such adjectives must precede the numeral:[37]

(48) a.   *trei-zeci   de alte   caiete*
           three-tens of other notebooks
           'thirty other notebooks'

      b.   * *zece alte   caiete*
           ten   other notebooks

      c.   *alte   zece caiete*
           other ten   notebooks
           'ten other notebooks'   (Romanian)

Similarly, in English, pluralized numerals allow *of* to be followed by demonstratives, as in (49a), which gives rise not only to the partitive interpretation but also to the cardinal one (roughly, 'hundreds of books of this kind'); this contrasts with unpluralized numerals, as in (49b), which cannot be followed by demonstratives (unless used in a full partitive construction with *of*, as in (49c)):[38]

(49) a.   hundreds of these books

      b.   * a hundred these books

      c.   a hundred of these books

This supports the view that what follows the numeral in a head-complement NNC (as in partitives) is more than just NP; furthermore, this also strengthens the claim made in §3.5, that the structural difference between head-complement NNCs and partitive constructions might be quite minimal.

In light of facts like this, I assume that the embedded nominal projection in head-complement NNCs is indeed more than just an NP. The second part of the case constraint, in (43b), is thus nothing surprising, and not really an independently stipulated rule: the need for case on extended nominal projections is not specific to those embedded in a head-complement NNC. Being embedded within

---

[37]The adjective *alte* may alternatively precede the numeral in (48a); this, however, gives rise to a different interpretation, 'another thirty notebooks'.

[38]Marginally, the same point can be made using a possessor instead of the demonstrative in (49), to the extent that *hundreds of his books* has not only a partitive reading ('hundreds of the books that are his') but also a non-partitive one, roughly 'hundreds of books that he has written'.

another nominal projection, standard locality principles prevent the phrase following the numeral from getting case from any external case-assigning head. The case constraint thus simply instantiates the general principle of case assignment under local c-command.

The first part of the case constraint, which rules out case assignment in the spec-head construction, is also not very surprising; under the assumption that case is assigned under c-command by a suitable head (or is a side-effect of Agree under local c-command), the numeral in a spec-head construction simply does not c-command the maximal extended projection of the noun. Taking both parts of the case constraint together, it thus amounts more or less to the claim that the head-complement construction, but not the spec-head construction, involves two extended nominal projections. This now also provides a principled explanation for the fact that prepositional NNCs pattern with head-complement ones: both are complex nominals that can be claimed to contain two nominal layers: an external one, which receives its case from the outside, and an internal one, which must receive case DP-internally.[39]

4.3   A unification of the two constraints?

While the case constraint has been argued above to be mostly deducible from general principles of case theory, the number constraint still seems somewhat puzzling and stipulative. The question is whether it is possible to derive the number constraint from the case constraint and/or from other independently motivated principles.

One reason why such a reduction seems promising is the feeling that pluralized numerals like *hundreds* or *thousands* are somehow 'more nominal' than non-pluralized numerals. If indeed pluralized numerals project a structure that is sufficiently similar to that projected by nouns, then it is in fact expected that their projection would require case; this, in turn, forces the use of a head-complement construction according to the case constraint.

From this point of view, a NNC with a plural-marked numeral like (50) could be argued to require a recursive structure with two 'noun phrases' and two case features for the same reason that a phrase like (51), with two nouns, does:[40]

(50)   hundreds *(of) books

(51)   pictures *(of) books

---

[39] One additional question regarding case in NNCs is why the projection of the numeral in a spec-head NNC cannot receive case; in other words, nothing so far blocks the possibility of a prenominal genitive numeral in a spec-head construction (e.g., *a hundred's books*). We return to this in section 4.3.

[40] This is not to say that the two are identical in all respects; for one, there are still obvious semantic differences between *hundreds* and *pictures*.

If non-pluralized numerals, on the other hand, lack a number feature (as argued in section 4.1.3), there is at least one morphosyntactic factor distinguishing such numerals from nouns; this might be enough to make them exempt from the need to get case.[41]

One way to formalize this observation would be to claim that the presence of the nominal functional projection associated with number makes a numeral sufficiently 'noun-like' to require case. With this, we can now return to the question left open in the discussion so far: What is the exact nature of the functional projection that was labeled FP_C, which – by definition – is a nominal extended projection that requires case? Even though case is often seen as a feature of DPs, it might be somewhat odd to claim that the embedded phrase (*books*) in examples like (50) is a full DP, as it is neither referential nor argumental.[42] Instead, I propose that what was tentatively labeled 'FP_C' is minimally the syntactic level in which grammatical number is expressed.

To make this hypothesis explicit, I will adopt the assumption that grammatical number is located in the functional projection NumP which comes between NP and DP; for simplicity, I will abstract away from assuming any other functional projections.[43] The embedded phrase following *of* in (50) would therefore be a NumP, in line with the fact that it carries number but is non-referential. I will further assume that Num may merge not only with NP but also with CardP, the projection of a cardinal numeral. This means that, ignoring *of*, a head-complement NNC like (50) would have a structure like the following (assuming a Bare Phrase Structure analysis):

---

[41]This looks more or less like the mirror image of the claim in Chomsky (2000, 2001) that Case can be valued only as a result of an Agree operation involving a complete set of ϕ-features; here I am claiming that a numeral that doesn't have a number feature doesn't need Case.
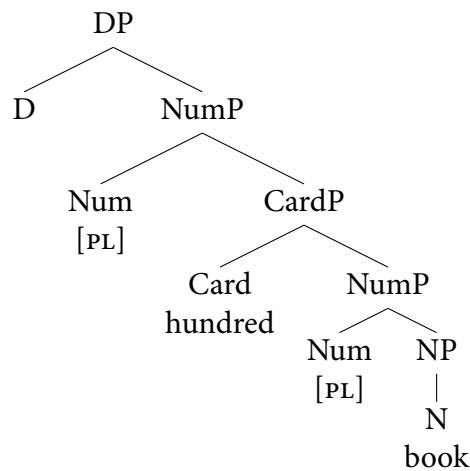
[42]However, if we pursue the proposal in footnote 21, such that (50) is a true partitive where the numeral quantifies over instantiations of a kind, then it might actually be perfectly reasonable to assume that what follows *of* is a full DP. More generally, if it turns out that there is semantic motivation for analyzing the embedded nominal in such cases as containing a DP layer, this would require nothing more than a simple adjustment to the structure proposed below.

[43]Alternatively, we could assume the hierarchy of functional projections proposed by Borer (2005):

- DP: associated with referentiality and definiteness

- #P: associated with counting and measuring

- ClP: associated with classifiers, including plural markers in languages like English

As far as I can see, the analysis proposed below can also be adapted to Borer's framework. For simplicity, though, I will assume the more familiar NumP.
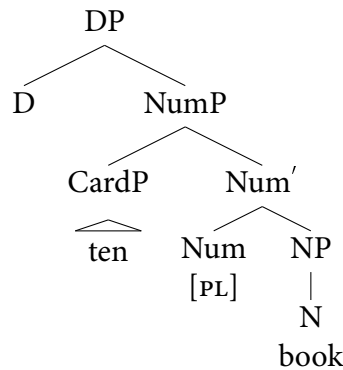
(52)

```
                         DP
                        /  \
                       D    NumP
                           /    \
                         Num     CardP
                         [PL]    /    \
                               Card    NumP
                             hundred   /  \
                                     Num   NP
                                     [PL]  |
                                           N
                                          book
```

If we now assume that case is required if NumP or higher is projected, then (52) would be ungrammatical unless case is somehow available to the lower NumP. In English, presence of *of* provides the necessary case. It seems that this analysis is not dependent on the specifics of the structural analysis of *of* itself, and hence I will remain neutral with respect to the exact status of *of* beyond the more or less uncontroversial assumption that it is associated with the availability of case.

Note that the structure in (52), with two NumP layers, is no more problematic than any other kind of recursive nominal; as the existence of recursive noun phrases is an undeniable aspect of human language, no new machinery is introduced by the structure above. For instance, the only modification needed to make the structure in (52) applicable to partitives is the addition of a DP layer dominating the embedded NumP. Furthermore, from a theoretical perspective, it is important to see that the proposed analysis does not contradict the assumption that the hierarchy of functional projections in a given *non-recursive* noun phrase contains a unique head of any given type.

In contrast to (52), a spec-head NNC like *ten books* involves no more than one NumP: as there is no grammatical number feature associated with the numeral (as discussed in §4.1.1), a non-recursive structure is possible in which the numeral occupies a specifier position, which I assume to be [spec,NumP]:

(53)

```
                          DP
                        /    \
                      D       NumP
                            /      \
                         CardP      Num′
                          /\       /    \
                        ten     Num     NP
                               [PL]      |
                                         N
                                        book
```

Crucially, under the analysis developed above there is no need for DP-internal case assignment in (53) since it contains only one extended nominal projection. A parallel structure with a *pluralized* numeral occupying [spec,NumP] (and hence, with a NumP dominating CardP in addition to the one dominating NP) would be ungrammatical as case is not available in this position. Even though many languages allow prenominal genitives, it seems reasonable to speculate that the position in which such genitive case is available is structurally distinct from the [spec,NumP] position in which I propose that numerals are located, as shown in (53). Given the complexity of the task of characterizing the nature of prenominal genitives even in a single language (see e.g. Cornilescu and Nicolae 2011), it is beyond the scope of the present paper to elaborate on this issue any further.

One consequence of the analysis proposed above is that it entails that case is not tied to one fixed functional projection: On the one hand, 'bare' NumP (such as the embedded phrase in (52)) was argued to require case; but on the other hand, referential mass noun phrases in argument position, which might be argued to lack a NumP layer (see e.g. the related discussion in Borer 2005), nevertheless require case. This means that case is required whenever a functional nominal projection is present (but is not required by NP itself), regardless of whether this projection is NumP or DP.

The above does not yet allow us to fully eliminate the number constraint as a separate principle. The proposal above seems to nicely account for the data regarding number on numerals discussed in section 4.1.1, covered by the number constraint in (38). However, the data regarding number on nouns, discussed in section 4.1.2, is still left unexplained. For instance, nothing so far explains why the literal translation of *five books*, with a plural noun, is ungrammatical in Welsh, while the same in a prepositional head-complement construction is grammatical, as shown in (11). As this requires a more thorough analysis of the nature of number marking on nouns in different languages, I leave it as an open question whether these facts can be reduced to a principled claim about the correlation between the presence of number features and the requirement for a recursive nominal structure.

4.4   Cross-linguistic variation

The constraints proposed above predict several areas of cross-linguistic variation in the use of numerals due to the following:

Case: Languages vary with respect to the availability of case assignment mechanisms in head-complement constructions:

- direct case assignment by the numeral
- prepositional case

Number: Languages vary with respect to the range of numerals that carry a morphosyntactic number feature.

Beyond the restrictions on case and number in a given language, the proposed theory provides an upper bound for the range of allowed constructions, but nevertheless leaves some room for variability that does not follow from the case and number constraints.

If a language has a mechanism for case assignment to a phrase embedded below a projection of the numeral, our analysis makes no general claim about the extent to which this mechanism can be used to form head-complement NNCs. On the one hand, there are languages like Russian and Welsh, which seem to allow such NNCs for all numerals: Russian with the numeral assigning morphological case, and Welsh with a prepositional element. On the other hand, there are languages like Hebrew and English, which seem to use head-complement constructions only as a 'last resort', when forced by the number constraint: Hebrew in a construct state using a bound numeral, and English using *of*. Thus, the fact that English, unlike Welsh, does not allow the prepositional *three of books*; and the fact that Hebrew, unlike Russian or Standard Arabic, does not allow the equivalent prepositionless genitive construction, does not follow from any of the two constraints. It thus seems that the choice whether to allow head-complement constructions pervasively for all NNCs or only for those NNCs that cannot have a spec-head structure due to the number constraint is a language-specific parameter.

5   Extensions and apparent counter-examples

The analysis so far predicts that in languages that use the spec-head construction for unpluralized numerals, pluralized numerals will be systematically different. While in most languages this seems to be correct, there are, however, some languages that don't seem to follow this prediction. In this section I discuss some of these languages, and propose a revision to the number constraint that could account for some of these languages.

5.1 Indonesian, Armenian and Basque

Indonesian is one of the languages that seem to pose a counterexample to the number constraint, as it looks like it uses the same kind of structure with pluralized numerals as with their non-plural counterparts:

(54) a. *se-ratus   pohon*
one-hundred tree

'a hundred trees'

b. *ratus-an   pohon*
hundred-PL tree

'hundreds of trees' (Indonesian)

One thing that might be relevant here is that the plural marking on pluralized numerals is morphologically different from nominal plurality: numerals are pluralized using the suffix *-an*, while nouns are pluralized via reduplication (e.g., *pohon pohon* 'trees'). This raises the possibility that pluralized numerals in Indonesian do not carry the same kind of formal NUMBER feature as nouns do; if so, the number constraint applies vacuously in this case, and the spec-head construction is not ruled out.[44]

If this analysis is correct, it makes the prediction that other languages where pluralized numerals do not bear the same morphological marking as plural nouns will also allow the spec-head construction with pluralized numerals. Two other languages that seem to support this prediction are Armenian and Basque. In Armenian, pluralized numerals do not have the same plural morphology as nouns do, as shown in (55):

(55) *tasn-yak grq-er*
ten-PL   book-PL

'tens of books' (Eastern Armenian)

As pluralized numerals have no effect on the noun's case marking and do not trigger the use of any kind of prepositional element, it seems that such numerals form spec-head NNCs. This fits into the proposed analysis, with the number constraint being satisfied vacuously since the 'plural' on the numeral is not the realization of the same formal number feature as in the nominal plural.

Finally, NNCs with pluralized numerals in Basque also look like those with non-pluralized ones, as shown in the following examples, from Etxeberria and Etx-

---

[44]It might also be the case that the additional morpheme on the pluralized numeral in (54b) is the instantiation of an additional functional head, such that NNCs of this kind in languages like Indonesian contain additional abstract structure beyond what is present in (54a). Even if true, it seems that what matters to the current analysis is simply that these pluralized numerals are formally distinct from pluralized nominals.

epare (2008):[45]

(56)   a.   *Mila   ikasle   etorri dira.*
            thousand student come aux-PL
            'One thousand students came.'

       b.   *Mila-ka   ikasle   etorri dira.*
            thousand-SUFFIX student come aux-PL
            'Thousands of students came.'   (Basque; from Etxeberria and
            Etxepare, 2008)

As Basque is a head-final language, word order supports taking pluralized numerals
in this language, like their unpluralized counterparts, to appear in spec-head con-
structions. As in the case of Indonesian and Armenian, however, plurality mark-
ing on numerals in Basque is clearly distinct from nominal plurality: unlike nouns,
which can be marked as plural only by means of a plural-marked definite article,
plurality of the numeral in cases like (56b) is achieved by using the iterative suf-
fix *-ka*, often used in manner adverbs (cf. the Turkish pluralized numeral in (30c)).
Thus, Basque pluralized numerals are not truly number-marked in the formal sense
relevant for the number constraint.

   We thus conclude that pluralized numerals in Indonesian, Armenian and Basque,
which might at first seem to pose counterexamples to the number constraint, ac-
tually provide surprising evidence to support it. The crucial observation is that
what is at stake is the presence of formal morphosyntactic NUMBER features, rather
than semantic plurality of the numeral. In light of the discussion in section 4.3,
this is expected: if the number constraint is the result of number-marked numerals
containing an extra NumP projection in addition to the one which dominates NP,
then it is not surprising that languages that pluralize numerals in a manner dis-
tinct from that used for nouns do not subject such numerals to the same syntactic
requirements as nouns.

5.2   Finnish

Finnish NNCs display several properties not found in any of the languages dis-
cussed so far. The most common type of NNC in Finnish seems to be the head-
complement construction, where the noun (which is morphologically singular) re-
ceives partitive case from the numeral:

(57)   *kolme kirjaa*
       three book.PART
       'three books'   (Finnish)

When the numeral is pluralized, however, the case on the (plural) noun is no longer dependent on the numeral, but rather is the case assigned to the noun phrase as a whole, which also surfaces on the numeral:

(58)   a.   *sadat   tähdet*
             100.PL.NOM star.PL.NOM
             'hundreds of stars' (nominative)   (from Hurford 2003)

        b.   *satoja   tähtiä*
             100.PL.PART star.PL.PART
             'hundreds of stars' (partitive)   (Finnish)

This pattern seems like a counterexample to the number constraint: on the one hand, both the numeral and the noun carry number morphology, and on the other hand, case concord seems to argue against assuming a head-complement construction.

The solution, I believe, is to analyze this as involving not two independent number features, but one, where multiple realizations are the result of agreement. Finnish shows number agreement between numerals and nouns not only in the construction illustrated in (58), but also in another kind of construction illustrated below (from Hurford, 2003):
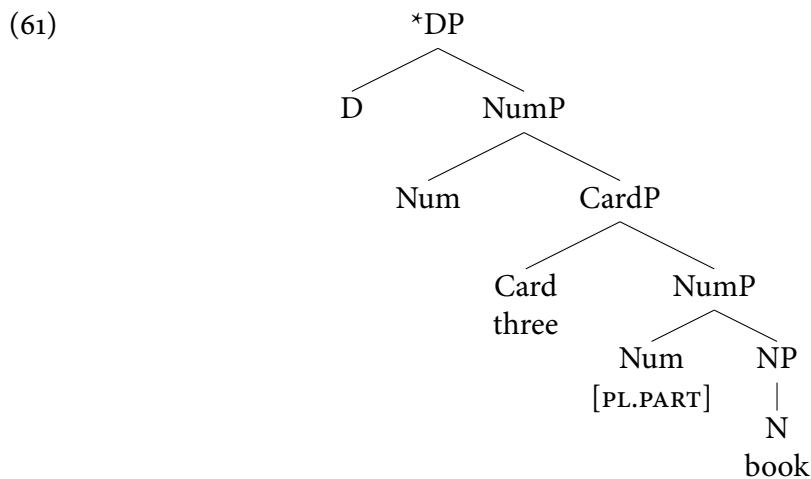
(59)   *neljät   silmät*
        4.PL.NOM eye.PL.NOM
        'four pairs of eyes'   (Finnish; from Hurford 2003)

While morphologically this looks identical to (58), the interpretation is quite different: (59) refers to 4 sets (typically pairs) of eyes, not to an indefinite plurality of sets of 4 eyes each. One way to formulate what (58) and (59) have in common is to say that both involve plurality being interpreted on exactly one node – either on the numeral (in the case of 'hundreds') or on the noun (in the case of 'sets of eyes') – despite being marked on both. This is typical of agreement configurations, and hence the Finnish examples above are different from the cases that were earlier claimed to be excluded by the number constraint due to having number on both the numeral and the noun.

Because of the lack of NNC-internal case assignment, I will assume that the examples in (58) and (59) are spec-head constructions involving agreement in both case and number. To accommodate for these constructions, we could thus reformulate the number constraint as follows:

(60)   The number constraint (revised): The spec-head construction is not possible if both the numeral and the noun carry independent morphosyntactic number features.

Since the two occurrences of the number feature in the Finnish examples above are not independent of each other, Finnish no longer poses a counterexample to this constraint.[46] What is left as an open question is why Finnish NNCs alternate between a head-complement structure in the absence of any plural marking and an agreeing spec-head structure where plurality is marked. One possible explanation is that this has to do with the nature of the Finnish partitive, which is well known to have quantificational and aspectual content (Kiparsky, 1998). Having an embedded NumP which is both plural and partitive might then be blocked either for semantic reasons or due to a structural competition for the Num position, making the following recursive structure ungrammatical:[47]

(61)

```
                    *DP
                   /    \
                  D      NumP
                        /    \
                      Num     CardP
                             /     \
                          Card      NumP
                          three    /    \
                                 Num      NP
                              [PL.PART]    |
                                           N
                                          book
```

This contrasts with a grammatical structure where plurality is missing from the embedded NumP, or with a grammatical spec-head NNC where partitive is absent from Num. For reasons of space we do not develop this tentative hypothesis further.

5.3   Dutch

One last problem, which is not solved by the above discussion, is posed by Dutch and some other Germanic languages. In Dutch, pluralized numerals seem to combine with the noun in the same way that their unpluralized numerals do, which in both cases involves a plural noun following the numeral with no prepositional element:

---

[46]An alternative formulation of the number constraint would be to refer to two *interpretable* number features. Recall, however, that we accounted for the ungrammaticality of the spec-head construction in Hebrew with the numeral *šnayim* ('two') using the fact that it carries plural morphology; while this number feature is clearly independent (as it is inherent on the numeral and is not the result of agreement), it is not clear whether it is also interpretable.

[47]It might be proposed that there is no NumP dominating CardP in languages that use a head-complement NNC with non-pluralized numerals. As far as I can see, this proposal would be compatible with other parts of my analysis.

(62) a.   *honderd koeien*
        hundred cows
        'a hundred cows'

      b.   *honderd-en koeien*
        hundred-PL cows
        'hundreds of cows'   (Dutch)

The fact that plurality on the numeral looks just like nominal plurality rules out the possibility of applying the explanation proposed above for Indonesian; and the fact that this is limited to decimal numerals like 100s seems to rule out an agreement analysis like that proposed for Finnish. This means that, according to the analysis proposed in this paper, we are left only with the option of analyzing this as a head-complement construction, which then raises the question of case assignment. Apparently, the same question arises in Dutch with certain classes of nouns, which may precede other nouns with no preposition similar to English *of*, as in the following:

(63) a.   *een kudde koeien*
        a   herd   cows
        'a herd of cows'

      b.   *een verzameling boeken*
        a   collection   books
        'a collection of books'   (Dutch)

The problem is therefore not specific to numerals; whatever accounts for the apparent lack of DP-internal case in (63) is presumably also what makes pluralized numerals possible in constructions like (62).[48] [49] The general conclusion, once again, is that when numerals carry the same number marking as nouns do, they are subject to the same constraints on case as those that apply to noun phrases.

6   Conclusion

This paper started with the observation that natural language provides at least two distinct syntactic structures for numeral-noun constructions, where both can be found in a wide variety of languages. This raised the question of what constrains

---

[48]One possible hypothesis, that cannot be elaborated on for reasons of space, is that Dutch requires case only on full DPs and not on NumP.

[49]As pointed out by an anonymous reviewer, this might also be somehow related to the optional absence of *of* in English quantificational phrases like *all/both (of) the books*. It seems that many other languages also allow, to different extents, dropping of prepositional genitive markers with certain quantificational or measure-denoting elements; therefore, the Dutch data discussed here might not be as exceptional as it might seem at first.

the choice between these two structures; I have argued that while UG allows some flexibility in this domain, the main two factors in determining the type of NNC to be used are morphosyntactic number and case. It was then argued that these are not two independent factors, as presence of grammatical number is tied to the projection of nominal functional structure, which is subject to general requirements of case.

Even though our starting point was that of analyzing the structural position of numerals while trying to abstract away from the issue of their categorial status, the outcome of the proposed analysis also provides a partial answer to the question of what is the category of cardinal numerals: numerals that are plural-marked (as a result of merging with Num) were shown to behave like nouns, in contrast to numerals that are not plural-marked. We thus end up with a system which fits into the growing body of research that moves some of the burden of categorization from the lexicon to the domain of morphosyntax. This contrasts with approaches that simply classify all numerals as nouns (see e.g. Corver and Zwarts, 2006); and to a smaller extent, my analysis also contrasts with the approach of Kayne (2010) and Stavrou and Terzi (2008), who view numerals as a distinct category which can be nominalized by means of a morpheme which is separate from number.

From the point of view of the syntax-semantics interface, two important conclusions emerge: first, that the syntax of human language may encode the semantic relation between a cardinal numeral and a noun in more than one way; and second, that there is no one-to-one mapping between morphosyntactic number and semantic plurality.

7   Acknowledgements

References

Abney, S. P., 1987. The English noun phrase in its sentential aspect. Ph.D. thesis, MIT.

Arsenijević, B., 2006. Partitivity and reference. In: Dotlačil, J., Gehrke, B. (Eds.), Proceedings of the Second Syntax AiO Meeting. UiL OTS, Utrecht, pp. 48–64.

Bailyn, J., 2004. The case of Q. In: Arnaudova, O., Browne, W., Rivero, M. L.,

Stojanovic, D. (Eds.), Formal Approaches to Slavic Linguistics #12: The Ottawa Meeting 2003. University of Michigan Press, Ann Arbour, Michigan, pp. 1–36.

Borer, H., 1999. Deconstructing the construct. In: Johnson, K., Roberts, I. (Eds.), Beyond Principles and Parameters. Kluwer, Dordrecht, pp. 43–89.

Borer, H., 2005. Structuring Sense, Volume 1: In Name Only. Oxford University Press, Oxford.

Cardinaletti, A., Giusti, G., 2006. The syntax of quantified phrases and quantitative clitics. In: Everaert, M., van Riemsdijk, H. (Eds.), The Blackwell Companion to Syntax. Vol. 5. Blackwell, Oxford, pp. 23–93.

Chomsky, N., 2000. Minimalist inquiries: The framework. In: Martin, R., Michaels, D., Uriagereka, J. (Eds.), Step by Step: Essays in Minimalist Syntax in Honor of Howard Lasnik. MIT Press, Cambridge, Mass, pp. 89–155.

Chomsky, N., 2001. Derivation by phase. In: Kenstowicz, M. (Ed.), Ken Hale: a Life in Language. MIT Press, Cambridge, Mass., pp. 1–52.

Cinque, G., 2005. Deriving Greenberg's universal 20 and its exceptions. Linguistic Inquiry 36 (3), 315–332.

Corbett, G. G., 1978. Universals in the syntax of cardinal numerals. Lingua 46, 355–368.

Cornilescu, A., Nicolae, A., 2011. On the history of Romanian genitives: The prenominal genitive. In: Cornilescu, A. (Ed.), Bucharest Working Papers in Linguistics. Vol. XIII(2). University of Bucharest, Bucharest.

Corver, N., Zwarts, J., 2006. Prepositional numerals. Lingua 116 (6), 811–835.

Danon, G., 1996. The syntax of determiners in Hebrew. Master's thesis, Tel Aviv University.

Danon, G., 1998. Two syntactic positions for determiners in Hebrew. In: Wyner, A. Z. (Ed.), Proceedings of IATL 13. The Israel Association for Theoretical Linguistics, pp. 55–73.

Danon, G., 2008. Definiteness spreading in the Hebrew construct state. Lingua 118 (7), 872–906.

Donohue, M., 2005. Numerals and their position in universal grammar. Journal of Universal Language 6 (2), 1–37.

REFERENCES

Dryer, M. S., 1992. The Greenbergian word order correlations. Language 68, 81–138.

Etxeberria, U., Etxepare, R., 2008. Number agreement with weak quantifiers in Basque. In: Abner, N., Bishop, J. (Eds.), Proceedings of the 27th West Coast Conference on Formal Linguistics. Cascadilla Proceedings Project, Somerville, MA, pp. 159–167.

Etxeberria, U., Etxepare, R., 2009. Measures and counting in Basque. In: Gorrotxategi, J., Lakarra, J., Urgell, B. (Eds.), Proceedings of the Mitxelena Chair.

Franks, S., 1994. Parametric properties of numeral phrases in Slavic. Natural Language & Linguistic Theory 12 (4), 597–674.

Giusti, G., 1997. The categorial status of determiners. In: Haegeman, L. (Ed.), The New Comparative Syntax. Longman, pp. 95–123.

Giusti, G., 2002. The functional structure of noun phrases: A bare phrase structure approach. In: Cinque, G. (Ed.), Functional Structure in DP and IP. Oxford University Press, Oxford, pp. 54–90.

Greenberg, J. H., 1978. Generalizations about numeral systems. In: Greenberg, J. H., Ferguson, C. A., Moravcsik, E. A. (Eds.), Universals of Human Language, Volume 3: Word Structure. Stanford University Press, Stanford, Ch. Generalizations about Numeral Systems, pp. 249–296.

Hellan, L., 1986. The headedness of NPs in Norwegian. In: Muysken, P., van Riemsdijk, H. (Eds.), Features and Projections. Foris, pp. 89–122.

Hurford, J. R., 2003. The interaction between numerals and nouns. In: Plank, F. (Ed.), Noun Phrase Structure in the Languages of Europe. Empirical Approaches to Language Typology / Eurotyp. Mouton de Gruyter, pp. 561–620.

Ionin, T., Matushansky, O., 2006. The composition of complex cardinals. Journal of Semantics 23 (4), 315–360.

Jackendoff, R., 1977. X'-Syntax: A Study of Phrase Structure. MIT Press.

Kayne, R. S., 2010. A note on the syntax of numerical bases. In: Comparisons and Contrasts. Oxford University Press, Oxford, pp. 57–72.

Kiparsky, P., 1998. Partitive case and aspect. In: Butt, M., Geuder, W. (Eds.), The Projection of Arguments: Lexical and Compositional Factors. CSLI Publications, Stanford, pp. 265–307.

Landman, F., 2004. Indefinites and the Type of Sets. Blackwell, Oxford.

Longobardi, G., 2001. The structure of DPs: Some principles, parameters and problems. In: Baltin, M., Collins, C. (Eds.), The Handbook of Contemporary Syntactic Theory. Blackwell, Oxford, pp. 562–603.

Mittendorf, I., Sadler, L., 2005. Numerals, nouns and number in Welsh NPs. In: Butt, M., King, T. H. (Eds.), Proceedings of the LFG05 Conference. CSLI Publications, Stanford, pp. 294–312.

Mittendorf, I., Sadler, L., February 2010. Welsh prenominal numerals: An analytic puzzle. First Regional LFG meeting, SOAS, `http://privatewww.essex.ac.uk/~louisa/numerals/lfgmeet-numho.pdf`.

Nelson, D., Toivonen, I., 2000. Counting and the grammar: Case and numerals in Inari Sami. In: Nelson, D., Foulkes, P. (Eds.), Leeds Working Papers in Linguistics. Vol. 8. University of Leeds, University of Leeds, pp. 179–192.

Ortmann, A., 2000. Where plural refuses to agree: Feature unification and morphological economy. Acta Linguistica Hungarica 47, 249–288.

Pereltsvaig, A., 2006. Passing by cardinals: In support of head movement in nominals. In: Lavine, J., Franks, S. L., Tasseva-Kurktchieva, M., Filip, H. (Eds.), Formal Approaches to Slavic Linguistics 14: The Princeton Meeting. Michigan Slavic Publications, Ann Arbor, MI, pp. 277–292.

Ritter, E., 1991. Two functional categories in noun phrases: Evidence from Modern Hebrew. Syntax and Semantics 25, 37–62.

Rothstein, S., 2010. Counting and the mass/count distinction. Journal of Semantics 27 (27), 343–397.

Sadler, L., 2000. Noun phrase structure in Welsh. In: Butt, M., King, T. H. (Eds.), Argument Realization. CSLI Publications, Stanford, pp. 73–110.

Schwarzschild, R., 2006. The role of dimensions in the syntax of noun phrases. Syntax 9 (1), 67–110.

Selkirk, E., 1977. Some remarks on noun phrase structure. In: Culicover, P. W., Wasow, T., Akmajian, A. (Eds.), Formal Syntax. Academic press, New York, pp. 285–316.

Shlonsky, U., 2004. The form of Semitic noun phrases. Lingua 114 (12), 1465–1526.

Sigurðsson, H. Á., 1993. The structure of the Icelandic NP. Studia Linguistica 47 (2), 177–197.

REFERENCES

Sportiche, D., 1988. A theory of floating quantifiers and its corollaries for constituent structure. Linguistic Inquiry 19 (3), 425–449.

Stavrou, M., Terzi, A., 2008. Types of numerical nouns. In: Chang, C. B., Haynie, H. J. (Eds.), Proceedings of the 26th West Coast Conference on Formal Linguistics. Cascadilla Proceedings Project, Somerville, MA, pp. 429–437.

Thráinsson, H., 2007. The Syntax of Icelandic. Cambridge University Press, Cambridge.

Zabbal, Y., 2005. The syntax of numeral expressions, unpublished paper, UMass, Amherst.

Zweig, E., 2006. Nouns and adjectives in numeral NPs. In: Bateman, L., Ussery, C. (Eds.), Proccedings of NELS 35. GLSA Publications, Amherst, Mass., pp. 663–675.