Dékány, Éva. 2024. Constituency in classifier expressions: Ch'ol and beyond. *Glossa: a journal of general linguistics* 9(1). pp. 1–47. DOI: https://doi.org/10.16995/glossa.10874

Open Library of Humanities

Constituency in classifier expressions: Ch'ol and beyond

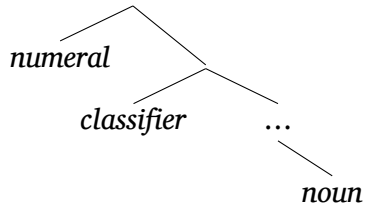
Éva Dékány, HUN-REN Hungarian Research Centre for Linguistics and ELTE Eötvös Loránd University, Budapest, HU, dekany.eva@nytud.hu

This paper considers recent arguments from the literature that have been deemed supportive of the view that in Ch'ol (Mayan) numerals and classifiers form a constituent to the exclusion of the noun, and evaluates proposed diagnostics for probing NP-internal constituency cross-linguistically. I show that the purported diagnostics are unsuitable as adjudicators of constituency both cross-linguistically and internally to Ch'ol. This is because the data involved in the tests are not uniquely compatible with [[Num Cl] N]; they can also be captured on [Num [Cl N]].

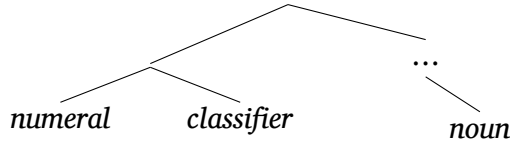
1 Introduction

One of the most significant open debates about the constituency of NP centers around the question of whether the classifier forms a constituent with the noun (1) or the numeral (2).

(1)



(2)



The problem of choosing between the two views has been around for decades, and it has proven to be difficult to solve on empirical grounds. In a recent paper Bale et al. (2019) offer four arguments for the [[Num Cl] N] structure from Ch'ol, a head-marking V-initial Mayan language (Chiapas, Mexico). These are briefly previewed in (3).

(3)
- a. the numeral and the classifier form a phonological word;
- b. phrases undergoing coordination may contain the numeral and the classifier to the exclusion of the noun;
- c. in certain complex numerals a classifier appears inside the numeral, in a position that is non-adjacent to the noun;
- d. the numeral and the classifier can undergo A-bar movement together, stranding the noun.

Importantly, Bale et al. suggest that the way they test NP-internal constituency has applicability beyond Ch'ol as well: their "diagnostics can be used as a template to test the constituency structure in other classifier languages" (p. 1). The goal of this paper is to demonstrate that Bale et al.'s diagnostics are insufficient for distinguishing between [[Num Cl] N] and [Num [Cl N]] even internally to Ch'ol: while all of their data are indeed compatible with [[Num Cl] N], these data can also be captured on [Num [Cl N]] with equal ease. In prior work (Dékány 2021; 2022a; b; to appear) I have supported the universal [Num [Cl N]] hypothesis, so readers are referred to those works for empirical and conceptual arguments for [Num [Cl N]]. In this paper I will not show that the structure of the Ch'ol NP must be [Num [Cl N]], or that [[Num Cl] N] is not viable in this language. My aim is to show that in order to determine constituency, cross-linguistically as well as in Ch'ol, we have to go beyond the tests in Bale et al. (2019).

Before we delve into the discussion, a caveat is in order regarding the structural representation of nominal modifiers. N-modifiers, including the numeral and the classifier, have been approached in two ways. On the one hand, they have been viewed as occupying positions on the main projection line of the extended NP in the form of functional heads or specifiers, taking a projection of the noun as their compelment. On the other hand, they have also been analyzed as adjuncts to a projection of the noun. This paper is primarily concerned with the issue of constituency, which is logically independent of the adjunct/non-adjunct status of N-modifiers. In principle, in (1) the numeral can be either a (specifier of a) functional head or an adjunct, and both options are available for the classifier as well. Similarly, in (2) the constituent that comprises the numeral and the classifier can, in principle, be either an adjunct (as in Bale et al. 2019) or a specifier of a functional head. Since a commitment to the adjunct/non-adjunct status of N-modifies is not necessary for my present purposes, I will remain neutral on this issue. Where labels will be useful in trees, I shall use xNP for 'extended nominal projection'. This is a cover-term for NP, NumP and DP, with no commitment as to the (non)adjunct status of the left-hand daughter. The abbreviations Num and Cl stand for the descriptive terms 'numeral' and 'classifier', respectively, and are not meant to be understood as 'the functional head number' or 'the functional head classifier'. (When I will have to talk about 'the functional head classifier', I will write this out in full as 'the functional head Cl^0'.)

This paper is organized as follows. In Sections 2 through 5 I go through Bale et al.'s (2019) tests one by one and demonstrate why they are not suitable adjudicators of NP-internal constituency. In Section 6 I tackle some further arguments for [[Num Cl] N] made on the basis of Ch'ol, and show that these arguments do not carry force either. I conclude in Sect. 7.

2 Phonological constituency

Bale et al. (2019: 17) argue that "the first piece of evidence" for the [[Num Cl] N] structure is the fact that "the numeral and the classifier form a single phonological word in Ch'ol". This is illustrated in (4), where the hyphen indicates the phonological dependency between Num and Cl.[1]

(4) *Ch'ol* (Bale et al. 2019)
ux-kojty wakax
three-CL cow
'three cows'

Bale et al. suggest that this is easily captured if the numeral and the classifier form a constituent. At the same time, they acknowledge that the alternative [Num [Cl N]] structure can also derive

[1] The examples in this paper come from published materials on Ch'ol. They preserve the orthography of their sources but I unified the glosses to match the conventions used in Bale et al. (2019).

this fact, mentioning that earlier literature, specifically, Borer (2005), proposed that the classifier undergoes head movement to Num. Therefore the phonological facts do not, in and of themselves, constitute an argument for or against either structure.

While this much is acknowledged by Bale et al. (2019), phonological constituency still plays a major – albeit indirect role – in their analysis. This is because the logic of their paper centers around the fact that on Borer's approach the numeral and the classifier form a complex head, while on the [[Num Cl] N] structure the numeral and the classifier form a phrasal constituent (possibly having a complex head inside). Taking this as their starting point, Bale et al. (2019) present three arguments to the effect that the numeral + classifier unit has phrasal status (these will be discussed below), and, in turn, this is taken to directly support [[Num Cl] N] and rule out [Num [Cl N]]. A major problem in this argumentation is Bale et al.'s tacit assumption that the [Num [Cl N]] structure can only capture the phonological facts of Ch'ol via Cl^0-to-Num^0 head movement, and so [Num [Cl N]] is inextricably linked to the complex head analysis. Bale et al. (2019) operate under the assumption that if in a particular environment the complex head analysis can be shown to be incorrect, then we automatically have an argument against [Num [Cl N]]. This is not the case, however. To be sure, syntactic head movement is one of the ways in which morphologically complex words can be derived. However, there are at least two other (non-lexicalist) approaches to such data as well.

On a well-known alternative to syntactic head movement as the (only) purveyor of morphologically complex words, nodes exponed by affixes can remain syntactically independent from their hosts throughout syntax, and may unite with the host, forming a complex head with it, only in the morphological component: via Lowering (Halle & Marantz 1993 et seq.) or post-syntactic Raising (Harizanov & Gribanova 2019). Similarly to syntactic head movmement, these operations are subject to strict locality (they can only target the next lower or next higher head, respectively), but they take place post-syntactically. On this approach parts of a morphologically complex word do form a complex head at some point, but not necessarily in narrow syntax. Therefore the syntactic unit containing the relevant parts might externally look like a constituent without underlyingly being a constituent.

Furthermore, it has been known since at least Marantz (1984; 1988) that mismatches between syntactic and phonological structures do occur, and parts of certain phonological words do not form a complex head either in narrow syntax or in the post-syntactic component. Among others, Marantz mentions the English auxiliary clitic *'ll* (*I'll go now*) and the Saxon Genitive *'s* (*the peacock in the park's feathers*) as examples where a morpheme forms a syntactic constituent with (and in the case of *'ll* scopes over) material to its right, but it forms a derived word with material to its left (without being in a complex head or other constituent with it either in or after sytax). Such data have inspirted Marantz's Morphological Merger. Marantz suggests that phonological

adjacency is associative: a phonological structure such as (5a) can be rebracketed into (5b), if the (morpho-phonological) properties of the lexical items involved motivate this. More recent work in the framework of Distributed Morphology calls (5b) *leaning* (Embick & Noyer 2001; Lipták & Saab 2016).

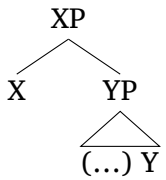
(5) Morphological Merger (leaning)

a. X * [$_{Y}$ Z * W * … (* = adjacency)

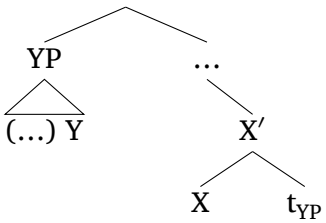
b. [[X Z] W …

(5b) means that in certain cases a syntactic structure will be linearized in such a way that the parts of a complex word end up in just the right configuration at PF for affixation to occur between them. The syntactic representation that will be mapped to such a linear order may be either derived or base-generated; this is immaterial. In the English examples above, the sentential subject and the possessor have arguably ended up in their surface position in front of the clitic via movement from the verbal and nominal thematic domain, respectively. Koopman & Szabolcsi (2000) et seq. have generalized this derivation to essentially all morphologically complex words, and eschewed syntactic head movement. In their approach a morphologically complex word Y-X is formed when a constituent containing Y as its last element moves to a specifier above X, such that there is no overt material between them (7). Y and X never from a syntactic constituent, but linearization makes them string-adjacent, and this can provide the bound morpheme with the host that it requires at PF.

(6) base-generated order



(7) phrasal movement



We should emphasize, however, that leaning has also been been profitably applied to nodes that arguably stay in situ. This is the case, for instance, with the verbal complex in Chichewa and other Bantu languages, where the left-to-right order of prefixes directly corresponds to the order of functional heads in the clause (Cinque 1999; Mchombo 2004).

(8) *ChiBemba* (Cinque 1999: 70)
N-kà-láá-boomba.
1SG-FUT-PROG-work
'I'll be working tomorrow.'

As argued in Cinque (1999), if examples such as (8) were derived by syntacic head movement (or post-syntactic Raising or Lowering), then we would expect the bound morphemes to be post-verbal and in the mirror order, contrary to fact. The attested order is therefore best derived via in-situ leaning.[2]

In addition to the English and Bantu examples mentioned above, cases of so-called suffix-stranding NP ellipsis also provide strong evidence that parts of phonological words do not have to form a syntactic (or post-syntactic) constituent. Suffix-stranding NP ellipsis can be observed in some agglutinating languages. It consists in ellipsis which deletes the noun but leaves some or all of the nominal suffixes unaffected. These stranded suffixes find a phonological host in the nearest overt NP-modifier (e.g., adjective, numeral or possessor) and get affixed to it at PF. There is no constituent at any part of the derivation that contains all and only the overt host and the suffix.

Consider possessive structures in Huallaga Quechua. The possessed noun bears agreement for the ϕ-features of the possessor and case appropriate for its role in the clause, while the possessor is marked with genitive case (9).

(9) *Huallaga Quechua* (Blake 2001: 103)
Hwan-pa wasi-n-ta rika-a.
John-GEN house-3SG-ACC see-1SG
'I see John's house.'

NP-ellipsis from this structure deletes the possessed noun and the agreement marker. The case affix is retained, however, and at PF the possessor serves as is host. This is an effect of linear adjacency rather than a matter of any particular narrow syntactic relationship.

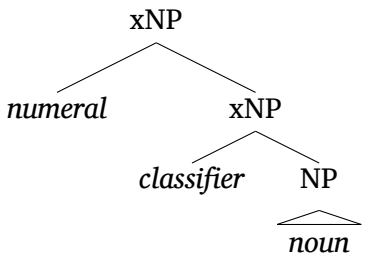
(10) *Huallaga Quechua* (Blake 2001: 103)
Hwan-pa-ta rika-a.
John-GEN-ACC see-1SG
'I see John's (house).'

[2] Recent work on polysynthetic languages (including Branigan et al. 2005; Compton & Pittman 2010; Barrie & Mathieu 2016; Ershova 2020) suggests that certain cases of polysynthesis may also involve leaning in derived or underived structures. The idea is that polysynthetic languages have a syntax very much like other language types, however, they have larger than usual domains of syllabification: rules of prosodification map entire ph(r)ases (e.g., vP, CP or DP) directly to phonological words. On this view polysynthetic words correspond to syntactic phrases rather than complex heads. Importantly, the cases of the English auxiliary clitic and the Saxon Genitive as well as the Bantu verbal complex and suffix stranding NP ellipsis (to be discussed below) show that phonological words can also arise when their components do not form a constituent in syntax, be that on the phrase-level or the head-level.

Further cases of suffix-stranding NP ellipsis, with adjectival and numeral hosts, are illustrated in Appendix A. That section also demonstrates that the stranded suffix becomes part of the vowel harmony domain of its new adjectival, numeral or other host, and thus indeed forms a phonological word with it.

To summarize, parts of a morphologically complex word do not necessarily form a constituent either in syntax or in the morphological component; what is required is a syntax that will be linearized in such a way that these parts end up adjacent at PF. Applying Morphological Merger (5) to the case of the numeral and the classifier in Ch'ol (and other languages where the classifier is a suffix to the numeral), in a structure like (11) these two elements are exactly in a configuration that will lead to their adjacency at PF, where affixation can take place.

(11)



(12) Morphological Merger (leaning)

a.	[$_{xNP}$ Num [$_{xNP}$ Cl NP]]	syntactic structure
b.	Num * [$_{xNP}$ Cl * N	phonological structure
c.	[[Num Cl] N	final phonological representation

The preceding discussion has shown that in principle, phonological constituency between morphemes can arise in three different ways:

(13) a. syntactic head movement
b. post-syntactic head movement
c. Morphological Merger/leaning at PF

It is now prudent to ask if we can determine which of these theoretical possibilities apply to Ch'ol numerals and classifiers. Most previous work takes classifiers to correspond to a functional head in syntax, cf. Tang (1990); Cheng & Sybesma (2005); Simpson (2005); Jenks (2011); Ott (2011); Alexiadou & Gengel (2012); Mathieu (2012), to mention just a few papers covering various language families. In the majority of cases examined in the previous literature, this appears to be well-motivated (e.g., classifiers generally cannot be modified). In Ch'ol, however, this approach would be difficult to maintain. As discussed in Arcos López (2009: 41–52) and Bale et al. (2019), most Ch'ol sortal classifiers are derived forms: they have the shape C_1VjC_2, which is derived from corresponding C_1VC_2 transitive verbal or positional roots. (Due to a phonological restriction on adjacent fricatives, fricative-final C_1VC_2 roots maintain their C_1VC_2 from as classifiers, see Coon

& Preminger 2009: n. 6.) The following examples illustrate the positional root *koty* ‘standing on four legs’ as a basis for the main sentential predicate (14) and for a sortal classifier (15).

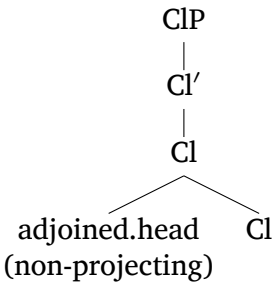
(14) *Ch’ol* (Bale et al. 2019)
Koty-ol jiñi me’.
standing.on.4.legs-STAT DET deer
‘The deer is standing on four legs.’

(15) *Ch’ol* (Bale et al. 2019, glosses adapted)
chäń-ko<j>ty me’
four-standing.on.4.legs<j> deer
‘four deer’

Orthographic <*j*> is IPA [h], and there is disagreement in the literature as to whether it involves infixation of a consonant (Vázquez Álvarez 2011: 65–66), or it reflects a special vowel quality (a lengthened aspirated vowel) whose value is supplied by templatic morphology (Coon 2017b). In either case, a C_1VjC_2 form is the product of a structure that is bigger than just a root/head. This has important repercussions for how to represent Ch’ol classifiers in syntax, which was not discussed in previous literature.

There are two ways in which the complexity of classifiers could be modeled. A base-generated head-adjunction approach (16) would offer the possibility that either the transitive/positional root or <*j*> corresponds to the classifier head Cl^0 projecting a ClP, with the other morpheme being a head-level adjunct to Cl^0 and not projecting its own phrase.

(16) rejected structure



Similar so-called conflation structures, with a non-projecting N being head-adjoined to a projecting V, have been proposed in the verbal domain by Haugen (2009) to account for denominal verbs such as ‘(to) hammer’ and by Mateu (2012) to analyze resultatives.

This cannot work in Ch’ol because neither transitive/positional roots, nor <*j*> is consistently associated to classifier structures. Transitive roots can give rise to transitive verbs and positional roots can give rise to stative predicates (14). <*J*> goes back to a Proto-Mayan nominalizer (Law

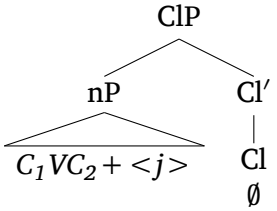
2020) and is used in contemporary Ch'ol as one of the strategies to form passives (Vázquez Álvarez 2011: Chap. 10.2, see (18)).

(17) *Ch'ol* (Vázquez Álvarez 2011: 303)
tyi k-päy-ä-ety te
PRFV A1-call-TV-B2 DIR:toward
'I called you to come.'

(18) *Ch'ol* (Vázquez Álvarez 2011: 303)
tyi pä<j>y-i-ø te
PRFV call<+PASS>-IV-B3 DIR:toward
'He was called to come.'

If neither the transitive/positional root nor $<j>$ corresponds to the functional head Cl^0, then (16) is unviable, and we must model the internal complexity of classifiers in a different way. This leads us to the second logical possibility, namely that Ch'ol classifiers are *phrasal* categories. A specific implementation of this idea can be found in Clausen (2023). Clausen suggests that the $<j>$ of classifiers and passives is a nominalizer, and thus both classifiers and passive predicates are externally nominal phrases of sorts. I will adopt this line of thinking, and for expository purposes, I will use the label nP for structures with $<j>$.[3] I would like to emphasize, however, that what is important for me is that Ch'ol sortal classifiers correspond to phrases; whether these phrases are nominal and what their exact label is is immaterial for the present discussion. I suggest that the phrase comprising the transitive/positional root and $<j>$ merges in the specifier of ClP, and Cl^0 itself is morphologically unrealized.

(19)



The phrasal nature of Ch'ol classifiers is supported by the fact that up until the middle of the 20th century they could be modified by PPs (see (95) in Sect. 6.2).[4]

[3] Bale et al. (2019: fn. 17) note that a small number of classifiers are "nominal forms of intransitive roots, such as *-ñumel*, used to count repetitions". *-Ñumel* comprises the intransitive root *ñum* 'to pass' and the widely used nominalizer *-el*. This is thus a precedent for a classifier that is nominal (and phrasal) in category.

[4] That the $<j>$ of passives and classifiers is indeed the same thing is supported by the fact that it performs very similar syntactic-semantic functions in passives and in classifiers based on transitive roots. In passives, $<j>$ correlates with the demotion of the external argument, with the result being a one-place predicate (18). As for classifiers, the internal argument of the transitive (or positional) root "corresponds to the object being counted in the classifier constructions" (Bale et al. 2019: 12). When a transitive root is used as a basis for a classifier, it is impossible to add an external argument; only the internal argument position can be saturated by the counted noun. This fact can be

This proposal means that cross-linguistically, languages vary as to whether their classifiers spell out the functional head Cl^0 itself or they correspond to a phrase in the specifier of Cl^0. That this variation exists should not surprise us: Dékány (2021) argues that cross-linguistically and even internally to Hungarian, demonstratives show the same two-faced behaviour. That is, some demonstratives spell out the functional head $Deix^0$, while other demonstratives are phrasal categories Merged as specifiers of a phonologically null $Deix^0$.

Let us circle back to the issue of phonological constituency between Num and Cl. We now have a simple explanation for the fact that the classifier and the numeral together do not have the distribution of a complex head: the classifier itself is phrasal to begin with. Notice that this does not correlate with NP-internal constituency in any way: (19) could, in principle, either first-Merge with NP ([Num [Cl N]]) or with the numeral ([[Num Cl] N]). We can now also answer the question of which option in (13) best characerizes Num-Cl phonological constituency in Ch'ol. Both syntactic head movement and post-syntactic head movement (Raising or Lowering) are excluded: these are operations on heads, and Ch'ol classifiers are phrasal constituents. This leaves Morphological Merger (leaning) as a model for the phonological constituency, and as we have seen, this is compatible with both [[Num Cl] N] and [Num [Cl N]] as the underlying syntax.

Summarizing this section, [Num [Cl N]] does not force us to assume that Num and Cl form a complex head in syntax. As a result, ruling out the complex head analysis does not provide an argument in favour of [[Num Cl] N]. In the following sections I will go through Bale et al.'s remaining three diagnostics in detail and show that a) they are unsuitable as adjudicators of the [Num [Cl N]] vs. [[Num Cl] N] debate, and b) the data they are built on are even compatible with the view that in languages where the classifier spells out the functional head Cl^0, the numeral and the classifier form a complex head in syntax via head movement.

3 Coordination

Bale et al. (2019) observe that a syntactic unit containing the numeral and the classifier can undergo asyndetic disjunctive coordination.[5]

(20) *Ch'ol* (Vázquez Álvarez 2011: 255)
cha'-tyikil ux-tyikil kixtyañuj
two-CL three-CL SP:person
'few people'

straightforwardly captured if $<j>$ performs the demotion of the external argument (or the absorption of ergative case, as in Coon & Preminger 2009) in classifier structures as well. This would make (at least transitive root based) Ch'ol classifiers passive nominals of sorts. Just as passives function as predicates in clausal structures, the classifier is likely a predicate in spec, ClP, with the functional head Cl^0 mediating the predicational relationship between the classifier and the noun.

[5] Although in the previous section we saw that most Ch'ol classifiers are bimorphemic, for ease of readability I will gloss them in the rest of the paper as CL, as Bale et al. (2019) do.

(21) *Ch'ol* (Vázquez Álvarez 2011: 256)
ux-tyikil chäñ-tyikil kistyañuj tyi majl-i-y-ø-ob
three-CL four-CL SP:people PFV go-IV-EP-B3-PL3
'Few people went.'

They suggest that this would be difficult to derive on an approach where Num and Cl form a complex head via head movement, and they conclude from this that the Num + Cl unit has phrasal status. This, in turn, is taken as evidence for a structure in which the numeral and the classifier form a constituent to the exclusion of the noun: [[Num Cl] N].

The underlying assumption behind this line of argument is that the relevant examples have a WYSIWYG representation: the coordinated constituent contains only the numeral and the classifier.

(22) [[cha'-tyikil] & [ux-tyikil]] kixtyañuj
two-CL three-CL person
'few people'

This is indeed a logically possible source for examples like (20) and (21). However, there are two alternative structures, not at all considered in Bale et al. (2019), which can derive the data equally well. Specifically, the data also fall out from a structure in which entire extended NPs (complete with a numeral, a classifier and a noun) are coordinated, and the noun is either elided from the first conjunct (23) or it undergoes extraposition from both conjuncts via Right Node Raising (24).[6]

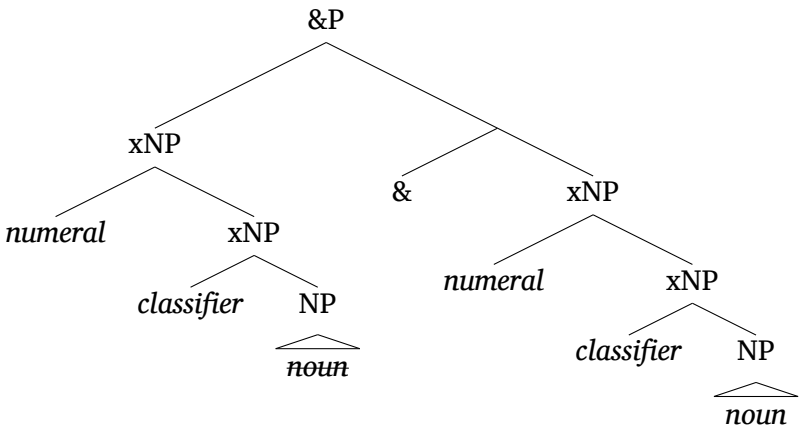
(23) [cha'-tyikil ~~kixtyañuj~~] & [ux-tyikil kixtyañuj]
two-CL person three-CL person
'few people'

(24) [cha'-tyikil ~~kixtyañuj~~] & [ux-tyikil ~~kixtyañuj~~] kixtyañuj
two-CL person three-CL person
'few people'

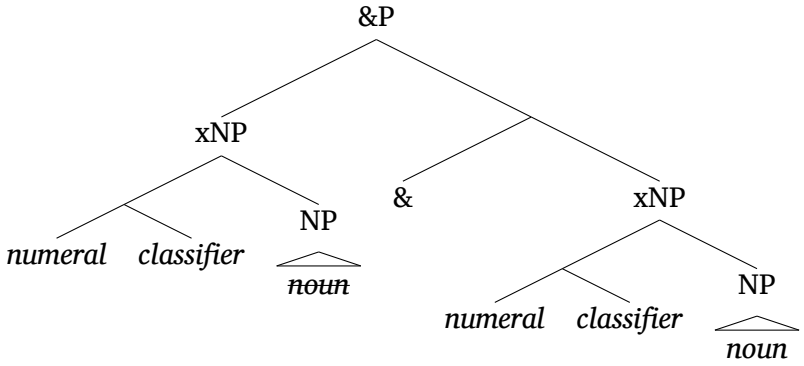
On both of these analyses, the coordinated units are indeed phrases rather than complex heads, but critically, both are neutral as to whether the classifier forms a constituent with the numeral or the noun. The two possible structures on the first, elliptical approach are shown below. Notice that on both representations, it would even be possible to assume that in languages where the classifier spells out the functional head Cl^0, the classifier undergoes head movement to and forms a complex head with Num, though this is not necessary to capture the data.

[6] Ionin & Matushansky (2018) argue that complex coordinative numerals cross-linguistically involve either (23) or (24), and never direct coordination between numerals. We will discuss complex numerals in detail in Sect. 4.

(25)



(26)



Importantly, NP/nP ellipsis in classifier expressions in independently attested in Ch'ol.

(27) *Ch'ol* (Coon 2017a: 666)
Añ cha'-k'ej tyi mesa
LOC two-CL PREP table
'There are two (round flat things) on the table.'

(28) *Ch'ol* (Vázquez Álvarez 2011: 157)
Ya' wa'-a-ø cha'-tyiki ix tyi karetera.
there stand-STAT-B3 two-CL there PREP SP:road
'There are two (men) in the road.'

Moreover, NP/nP ellipsis is also possible specifically in the disjunctive coordination under consideration: in (29) there is no overt noun at all. That reference is made to people is recoverable from the human classifier *tyikil*.[7]

[7] Ellipsis involves deletion under identity, thus the antecedent and the ellipsis site have to have a parallel structure. One may wonder whether this causes a problem for cases where one of the conjuncts contains 'one' and the other contains a higher numeral, where a mismatch in number marking on the elided noun might be expected.

(29) *Ch'ol* (Vázquez Álvarez 2011: 164)
juñ-tyiki cha'-tyiki tyi jul-i-ø
one-CL two-CL PFV arrive-IV-B3
'A few people arrived here.'

The structures in (25) and (26) predict that examples in which low NP-modifiers other than numerals and classifiers appear in either conjunct should be grammatical. (30) shows that this is indeed the case. Here both Num + Cl units are followed by a relative clause marked by the relativizing clitic *-bä*, but the crucial position is the one after the first conjunct.

(30) *Ch'ol* (Vázquez Álvarez 2011: 163, translation adapted)
mi i-mäñ-ø cha'-p'ej alaxax, [jum-p'ej k'äñ-ø = ix = bä] [jum-p'ej
IMFV A3-buy-B3 two-CL orange one-CL ripe-B3 = already = REL one-CL
ch'ok-ø = tyo = bä]
unripe-B3-still = REL
'He buys two oranges, one that is ripe and one that is still unripe.'

Examples like (30) must definitely be derived from (25) or (26), with ellipsis or Right Node Raising, as there is no approach on which numerals, classifiers and relative clauses sit together

---

(i) *Ch'ol* (Vázquez Álvarez 2011: 164)
poj jum-p'e cha-p'ej k'iñ toñe
HON one-CL two-CL day work
'few days of work'

Such examples are unproblematic, for two reasons. Firstly, as shown by (ii), this type of mismatch is tolerated under ellipsis.

(ii) Luna caught one shrew and Simba caught two/three.

Secondly, in Ch'ol it is always possible to achieve full parallelism between the antecedent and the ellipsis site. Inanimate nouns do not take the plural marker in the first place, animate non-human nouns rarely do so, and human denoting nouns optionally do so (Coon 2010a; Vázquez Álvarez 2011; Little 2020b). A plural interpretation is always possible in absence of an overt plural marker. Cf.:

(iii) *Ch'ol* (Bale et al. 2019)
Ta' jul-i-yob ux-tyikil x'ixik.
PFV arrive-ITV-EP-PL three-CL woman
'Three women arrived.'

(iv) *Ch'ol* (Bale et al. 2019)
[Ux-tyikil xk'aläl-ob] ta' y-il-ä-y-ety __.
three-CL girl-PL PFV A3-see-TV-EP-B2
'**Three girls** saw you.'

in a left branch.[8] But if (30) is not indicative of constituency between numerals, classifiers and relative clauses, then (20)–(21) do not supply an argument for [[Num Cl] N] either. This just reinforces the well-known thesis that since conjuncts can contain silenced material, coordination is a very unreliable test for constituency.

(31) a. Lewis loves but Marc hates going for long walks in the snow.
↛ *Lewis loves* is a constituent
b. three old and/or four few schools
↛ *three old* is a constituent

The examples above show that coordinating extended NPs/NumPs with ellipsis in the first conjunct or Right Node Raising is a possibility offered by Universal Grammar, moreover, this possibility is part of Ch'ol grammar as well. With these derivations in the toolbox, it becomes very difficult to exclude (23) or (24) as possible sources of (20)–(21). This, then, obviates the disjunction facts as an argument for the [[Num Cl] N] structure.

4 Complex numerals

Bale et al. (2019) suggest that complex numerals above 20 also provide an argument for [[Num Cl] N]. In the case of Ch'ol simplex numerals, the numeral is immediately followed by a classifier.

(32) *Ch'ol* (Bale et al. 2019: ex. 6a)
cha'-p'ej koya'
two-CL tomato
'two tomatoes'

(33) *Ch'ol* (Bale et al. 2019: ex. 25a)
ux-kojty wakax
three-CL cow
'three cows'

[8] Cf. also parallel data from Hungarian, with adjectives and participles in the conjuncts:

(i) *Hungarian* (own knowledge)
két szem piros és három szem zöld kávé
two CL:small.globular red and three CL:small.globular green coffee
'two red and three green coffee beans'

(ii) *Hungarian* (own knowledge)
három darab egész vagy négy darab [kocká-ra vág-ott] paradicsom
three CL:general whole or four CL:general dice-to cut-PST.PTCP tomato
'three whole or four diced tomatoes'

Again, under no approach to the internal structure of NP do numerals, classifiers and adjectives/participles from a constituent to the exlcusion of the noun.

Ch'ol is a vigesimal language. 20 and its powers (i.e., 400, 8000) have a distinguished role in that they a) form the basis of complex numerals and b) occupy the classifier slot (with their multipliers sitting in the regular pre-classifier numeral slot.)

(34) *Ch'ol* (Vázquez Álvarez 2011; Bale et al. 2019)
juñ-k'al, cha'-k'al, jum-bajk', cha'-bajk', jum-pijk
one-CL.20 two-CL.20 one-CL.400 two-CL.400 one-CL.8000
'twenty, forty, four hundred, eight hundred, eight thousand'
(lit. one-20, two-20s, one-400, two-400s, one-8000)

Since 20 and its powers are in the classifier position themselves, they block the occurrence of a 'regular' classifier. Compare (33), where 'cow' appears with *kojty*, the classifier for animals, and (35)–(36), where *k'al*, the root expressing 20 occupies the post-numeral slot and the appearance of *kojty* is ruled out.

(35) *Ch'ol* (Bale et al. 2019: ex. 25b)
ux-k'al wakax
three-CL.20 cow
'sixty cows'

(36) *Ch'ol* (Bale et al. 2019: ex. 25c)
*ux-k'al-kojty wakax
three-CL.20-CL cow
Intended: 'sixty cows'

This means that with 'round' numerals (i.e., multiples and powers of 20) the information about the shape, size or disposition of the noun that is normally coded in a garden variety classifier (33) is not present: the lexeme that is in the Cl slot is neutral as to whether the counted noun is human, animate, round and flat, long and slender, etc. In this respect, 20 and its powers resemble so-called generic or general classifiers (see Grinevald 2004). These "effectively suspend the classificatory function" while still filling the obligatory classifier slot (Law 2024: around 5'42") and can be found both within the Mayan family and outside of it. Generic classifiers perform the function of connecting numerals and nouns the same way as more specific classifiers do (in the sense of Borer 2005, they portion out 'stuff'); they are just more vague/generic in their lexical semantics than others. In many languages, generic classifies can replace more specific classifiers in NPs, similarly to what happens in (35).

Complex numerals that fall between two multiples of 20 employ a so-called overcounting pattern: they comprise a numeral in the 1–19 range and the next *higher* multiple of 20. For instance, 22 is expressed with the juxtaposition of 2 and 40, 36 involves the juxtaposition of 16 and 40, and 90 is coded by juxtaposing 10 and 100, etc.[9]

[9] See Marcos-Marín (1992); Hurford (2003); Harrison (2007); Ochiai (2014) on overcounting as a method of composing complex numerals cross-linguistically.

(37) *Ch'ol* (Bale et al. 2019: ex. 46a, original paraphrase)
cha'-p'ej i-cha'-k'al
two-CL A3-two-CL.20
'twenty two (lit. two of the group of two-20s)'

(38) *Ch'ol* (Merrifield 1966: 98, glosses added)
wík luhúm-p'ehl i-čáʔ k'ál
six ten-CL A3-two CL.20
'thirty six'

(39) *Ch'ol* (Bale et al. 2019: ex. 46b, original paraphrase)
lujum-p'ej i-jo'-k'al
ten-CL A3-five-CL.20
'ninety (ten of the group of five-20s)'

In these complexes the numeral falling between 1 and 19 is followed by a 'regular' classifier, one that matches the animacy, shape or disposition of the counted noun. Compare (38) and (40): the former is appropriate for nouns that are counted with *p'ej*, the general classifier, while the latter is used with nouns that take *kojty*(*/koht*), the classifier for animals.

(40) *Ch'ol* (Aulie 1957: 282, glosses adapted)
wək-luhun-koht i čaʔ-k'al
six-ten-CL A3 two-CL.20
'thirty six'

The second numeral of the complex, on the other hand, is always a multiple of 20 (or one of its powers), so it has 20 (or one of its powers) in the post-numeral classifier slot regardless of what the counted noun is. This second numeral is preceded by a 3SG prefix (belonging to Set A, which is used to cross-reference ergative arguments on verbs, possessors on possessed nouns and the Ground on relational nouns, Coon 2013).

Bale et al. (2019) argue that since these complex numerals are multi-word expressions, they correspond to XPs rather than complex heads. This is then taken to directly support [[Num Cl] N] over [Num [Cl N]]. As shown in detail in Sect. 2, this conclusion does not follow: the complex head analysis is just one of the ways in which [Num [Cl N]] can model phonological constituency between numerals and classifiers (and it cannot be applied to Ch'ol, where classifiers are syntactic phrases). Therefore ruling out the complex head analysis does not mean that [Num [Cl N]] is also ruled out.

The internal structure of Ch'ol complex numerals is not investigated in Bale et al. (2019), as the Authors consider this topic to fall outside the scope of their paper. However, as the XP status of these numerals does not adjudicate the constituency in and of itself, a more detailed look at

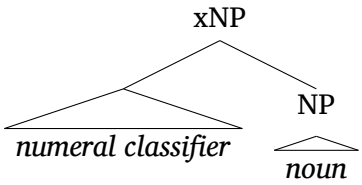
their internal structure is required. As shown above, Ch'ol overcounting numerals are built on the following abstract template:

(41) Num-Cl(regular) A3-Num(multiplier)-Cl.20

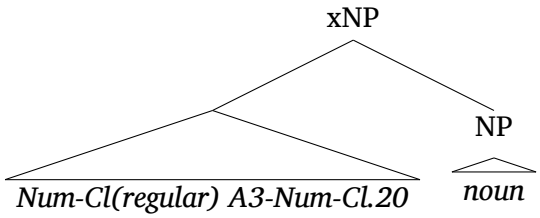
On Bale et al.'s (2019) [[Num Cl] N] approach this simply involves a complex phrase whose internal make-up is indeed immaterial.

(42) [[Num Cl] N] approach

a. simplex numeral



b. complex overcounting numeral



On [Num [Cl N]], the structure of these numerals can be approached in the same way as the coordinated Num + Cl sequences in Sect. 3: the two numerals and two classifiers on the surface are indicative of two Ns at deep structure. That is, overcounting numerals involve a syntactic relation (whose nature will be made more precise below) between two xNPs, both of which contain a numeral, a classifier and an noun. The two xNPs share the same listeme in the N slot, which allows one of the Ns to be deleted under identity (or to undergo Right Node Raising).

(43) [$_{xNP}$ Num-Cl(regular) ~~N~~] [$_{xNP}$ A3-Num(multiplier)-Cl.20 N] ellipsis

(44) [$_{xNP}$ Num-Cl(regular) ~~N~~] [$_{xNP}$ A3-Num(multiplier)-Cl.20 ~~N~~] N RNR

With both the first and the second classifier in a local relationship to a token of N, there is no argument for constituency either way: the data are compatible with both [[Num Cl] N] and [Num [Cl N]]. Moreover, in languages where the classifier corresponds to a functional head, both (43) and (44) are even compatible with the head movement approach to the phonological dependency between the numeral and the classifier.

The analysis in (43)/(44) receives support from Ionin & Matushansky (2006; 2018), who argue that (non-multiplicative) complex numerals, in fact, cross-linguistically involve operations between xNPs (such that each xNP contains an instance of the lexical noun), and components of the complex numeral do not form a constituent to the exclusion of the noun. Ionin & Matushansky (2006; 2018) suggest that additive numerals have the following, invariable structure across languages:

(45) [Num (Cl) N] J/P [Num (Cl) N]

Cross-linguistic variation among additive complex numerals stems from the following parameters: i) whether the relationship of the xNPs is mediated by the functional head J(unction) or by an adposition (P), ii) how the mediating head is exponed (J: ø or *and*; P: locative (*on*) or comitative (*with*)), and iii) which of the identical Ns, if any, is deleted on the surface.

Evidence for (45) is furnished by languages which allow (or require) multiple instances of the lexical noun to be spelled out overtly. Such languages include Biblical Welsh, Luvale (Bantu), Archaic Russian (Ionin & Matushansky 2018: Chap. 5.2.1) and Ainu (Tamura 1988/2000). A Biblical Welsh example is provided below.

(46) *Biblical Welsh* (Ionin & Matushansky 2006: 123)
saith **mlynedd** ac wyth gan **mlynedd**
seven year.PL and eight hundred year.PL
'eight hundred and seven years'

Further evidence for an instance of the lexical noun inside complex cardinals is provided by languages in which the counted noun appears after a conjunct other than the last one (Ionin & Matushansky 2018: Chap. 5.2.2). Examples of this pattern can be found in Biblical Welsh, Scottish Gaelic and archaic English (cf. *threescore years and ten* from Psalm 90:10 of the King James Version of the Bible). Compare (46) and (47):

(47) *Biblical Welsh* (Ionin & Matushansky 2006: 125)
can **mlynedd** a phymtheng **mlynedd** a thri ugain
hundred year.ADN and fifteen year.ADN and three twenty
'one hundred seventy-five years' (Genesis 25:7)

Ionin & Matushansky (2018) argue that languages like (modern) English also build their complex numerals as in (45); where they differ from Biblical Welsh and similar languages is that the ellipsis of the noun in the first conjunct is obligatory.

Ionin & Matushansky (2006) argue that complex subtractive numerals (e.g., Latin 19: *un-de-viginti* lit. 'one-from-twenty') likewise involve a relationship between two full xNPs (48).

(48) [Num (Cl) N] P [Num (Cl) N]

Subtractive numerals involve a P as a connector, and cross-linguistic variation again comes down to i) the exponence of P (caritive or ablative) and ii) whether the lexical N is ellipted from the subtrahend or the minuend (see Ionin & Matushansky 2006 for refinements). The noun can again be observed inside the complex numeral in some cases:

(49) *Welsh* (Ionin & Matushansky 2006: 144)
onid un **mlwydd** cant
unless one year.of.age hundred
'ninety-nine years'

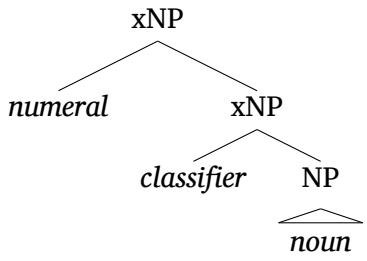
Ionin & Matushansky (2006) do not investigate the Ch'ol-type overcounting numerals. I suggest that similarly to additives and subtractives, these also involve a syntactic relationship between full xNPs, as in (43)–(44). As in the case of additives and subtractives, the components of the complex numeral do not form a constituent to the exclusion of the noun. Support for this position comes from theoretical and empirical considerations. On the theoretical side, this provides a unified representation for additives, subtractives and overcounting numerals: it would be difficult, if not impossible, to motivate why the first two involve relationships between xNPs while the last one has a different structure, with a relationship between just numerals. As for the empirical side, if the noun was represented just once, in its surface position after the whole complex numeral, then it would remain a mystery how the first classifier, deeply embedded inside a complex phrase (42b), could co-vary with the noun: see (38) and (40). This fact is straightforwardly captured, however, if the first classifier is also in a local relationship to an instance of the noun. Finally, overcounting numerals in Ainu (a Northeast Asian isolate with a vigesimal numeral system) provide explicit evidence for a nominal after the smaller numeral.

(50) *(Southern Hokkaido) Ainu* (Tamura 1988/2000: 255, glosses mine)
wan pa e-tu-hot-ne pa
ten year toward-two-20-COP year
'thirty years', lit. ten years toward forty years

I conclude from the foregoing discussion that it is indeed the schema in (43)/(44) that underlies complex numerals in Ch'ol. This means that it is not possible to place any theoretical burden of constituency on these numerals: they can be easily accounted for on [Num [Cl N]] as well.

(51) [Num [Cl N]] approach

a. simplex numeral



b. complex overcounting numeral: [[Num Cl ~~N~~] J/P [Num Cl N]]

In order to give a full-fledged account of the syntax of Ch'ol overcounting numerals, both (42) and (51) need to address the questions of i) how exactly the two components of the complex numeral are connected in syntax, and ii) why the A3 morphology appears on the second numeral of the complex. Since these issues are, strictly speaking, independent of constituency, I will not take them up here. In Dékány (to appear) I provide a cross-linguistic survey of overcounting numerals and conclude that the most frequent overt connector they feature is a lative ('toward') adposition. In the same work I suggest that Ch'ol overcounting numerals are also connected by a silent, latively interpreted P (which undergoes P-drop), thus the Ch'ol numeral 'twenty two' is essentially 'two towards forty'.

(52) *Ch'ol* (Bale et al. 2019: ex. 46a)
cha'-p'ej i-cha'-k'al
two-CL A3-two-CL.20
'twenty two'

The plausibility of this analysis is supported by the fact that an overt P connector is also attested in overcounting numerals in the Mayan family: in 18th century Yucatec Maya *t*, a reduced from of the all-purpose preposition links the two parts of the complex numeral.

(53) 18th c. *Yucatec Maya* (Beltran 1746: 196, gloss based on Yasugi 1995: 307)
ca t-uy-ox-kal
two PREP-A3-three-CL.20
'forty two'

I account for the A3 morphology with a possessive structure, the particulars of which need not concern us here. I refer the interested reader to Dékány (to appear) for detailed discussion.

5 A-bar extraction

The final, and on the face of it, perhaps strongest argument for the [[Num Cl] N] structure in Bale et al. (2019) is that in Ch'ol it is possible to front the numeral and the classifier together to the focus position, leaving the noun in the postverbal domain.

(54) *Ch'ol* (Bale et al. 2019)
Ux-tyikil ta' jul-i-yob [__ x'ixik].
three-CL PFV arrive-ITV-EP-PL woman
'**Three** women arrived.'

Bale et al. (2019) make a strong case that examples such as (54) involve extraction from the Noun Phrase, as the fronting of the numeral and the classifier is subject to the same restrictions as the fronting of interrogative possessors, the latter being a well-known case of A-bar extraction. In the Tila dialect of Ch'ol, possessor subextraction is possible only out of internal arguments: the subjects of unaccusatives and the objects of transitive verbs (Coon 2009; 2013; Bale et al. 2019).

(55) *Ch'ol* (Bale et al. 2019)
[Maxki i-plato] tyi yajl-I __?
who A3-plate PFV fall-ITV
'Whose plate fell?'

(56) *Ch'ol* (Bale et al. 2019)
Maxki tyi yajl-I [i-plato __]?
who PFV fall-ITV A3-plate
'Whose plate fell?'

(57) *Ch'ol* (Coon 2013: 89)
Maxki tyi aw-il-ä [i-chich __]?
who PFV A2-see-DTV A3-older.sister
'Whose older sister did you see?'

Extraction out of the subjects of transitive verbs (and, as expected, out of adjuncts), on the other hand, is banned.[10]

(58) *Ch'ol* (Bale et al. 2019)
[Majki i-chich] ta' y-il-ä-y-ety __?
who A3-sister PFV A3-see-TV-EP-B2
'Whose sister saw you?'

---

[10] As argued in Coon (2013), unergatives involve a light verb construction with 'do'. The light verb is transitive: it takes a nominal complement and introduces the subject in its specifier. As a result, extraction out of unergatives patterns with extraction out of transitives.

(i) *Ch'ol* (Coon 2013: 90)
*Maxki tyi i-cha`l-e soñ [i-chich __]?
who PFV A3-do-DTV dance A3-older.sister
Intended: 'Whose older sister danced?'

(59) *Ch'ol* (Bale et al. 2019)
*Majki ta' y-il-ä-y-ety [i-chich __]?
who PFV A3-see-TV-EP-B2 A3-sister
Intended: 'Whose sister saw you?'

Num + Cl fronting obeys the same constraints: it is felicitous out of unaccusative subjects (60)–(62) and transitive objects (63)–(65).

(60) *Ch'ol* (Bale et al. 2019)
Ta' jul-i-yob ux-tyikil x'ixik.
PFV arrive-ITV-EP-PL three-CL woman
'Three women arrived.'

(61) *Ch'ol* (Bale et al. 2019)
[Ux-tyikil x'ixik] ta' jul-i-yob __.
three-CL woman PFV arrive-ITV-EP-PL
'**Three women** arrived.'

(62) *Ch'ol* (Bale et al. 2019)
Ux-tyikil ta' jul-i-yob [__ x'ixik].
three-CL PFV arrive-ITV-EP-PL woman
'**Three** women arrived.'

(63) *Ch'ol* (Bale et al. 2019)
Ta' a-mäñ-ä cha'-p'ej alaxax.
PFV A2-buy-TV two-CL orange
'You bought two oranges.'

(64) *Ch'ol* (Bale et al. 2019)
[Jay-p'ej alaxax] ta' a-mäñ-ä __?
how.many-CL orange PFV A2-buy-TV
'How many oranges did you buy?'

(65) *Ch'ol* (Bale et al. 2019)
Jay-p'ej ta' a-mäñ-ä [__ alaxax]?
how.many-CL PFV A2-buy-TV orange
'How many oranges did you buy?'

Extraction out of transitive subjects in ungrammatical; here the only option is to move the entire Noun Phrase.

(66) *Ch'ol* (Bale et al. 2019)
[Ux-tyikil xk'aläl-ob] ta' y-il-ä-y-ety __.
three-CL girl-PL PFV A3-see-TV-EP-B2
'**Three girls** saw you.'

(67) *Ch'ol* (Bale et al. 2019)
*Ux-tyikil ta' y-il-ä-y-ety [__ xk'aläl-ob].
three-CL PFV A3-see-TV-EP-B2 girl-PL
Intended: '**Three** girls saw you.'

Based on these parallels, Bale et al. (2019) conclude that discontinuous NPs such as (54) involve extraction out of NP, and since Num and Cl undergo this movement together, they must form a constituent. This is then taken to support [[Num Cl] N] as the structure of NP over [Num [Cl N]].

Unassailable as this line of argument may seem, these data are far from being conclusive for the constituency of classifier expressions. Discontinuous Noun Phrases of the type in (54) constitute a case of Left Branch Extraction (Little 2020a), a phenomenon thoroughly investigated in Slavic languages. The similarities between Ch'ol and Slavic LBE are extensive: i) various types of NP modifiers such as possessors, numerals and adjectives may be extracted (though the extraction of possessors and adjectives is more restricted in Ch'ol than in Slavic and Ch'ol demonstratives don't extract at all, see Little 2020a for discussion), ii) the fronted constituent receives a contrastive interpretation (focus in Ch'ol and Serbo-Croatian, contrastive topic or focus in Russian), and iii) the fronted constituent may contain more than one NP modifier. These properties are illustrated below for Russian.

(68) *Russian* (Lena Borise, p.c.)
Čja u-pa-l-a tarelk-a?
whose PFV-fall-PST-SG.F plate-NOM.SG.F
'Whose plate fell?'

(69) *Russian* (Lena Borise, p.c.)
Maš-in-a u-pa-l-a tarelk-a.
Masha-POSS-NOM.SG.F PFV-fall-PST-SG.F plate-NOM.SG.F
'**Masha's** plate fell.'

(70) *Russian* (Lena Borise, p.c.)
Skol'ko ty/Maš-a kup-i-l-a apel'sin-ov?
how.many 2SG/Masha-NOM.SG.F buy-TH-PST-SG.F orange-GEN.PL
'How many oranges did you/Masha buy?'

(71) *Russian* (Lena Borise, p.c.)
Tr-i ty/Maš-a kup-i-l-a apel'sin-a.
three-ACC 2SG/Masha-NOM.SG.F buy-TH-PST-SG.F orange-GEN
'You/Masha bought **three** oranges.'

(72) *Russian* (Lena Borise, p.c.)
Tr-i bolš-ix Maš-a kup-i-l-a apel'sin-a.
three-ACC.PL big-GEN.PL Masha-NOM.SG.F buy-TH-PST-SG.F orange-GEN
'Masha bought **three big** oranges.'

Given the close correspondence between LBE in Slavic and Ch'ol, it is worth considering what could be profitably used from the extensive literature on the former in the analysis of the latter.

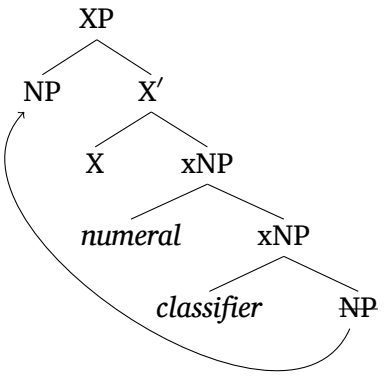
LBE in Slavic has received three different analyses: direct subextraction of the modifier from the NP (Ross 1967; Borsley & Jaworska 1988; Corver 1990; Bošković 2005), remnant movement of the NP after evacuation of the N head (Franks & Progovac 1994; Abels 2003; 2012; Bašić 2008) and scattered deletion (Fanselow & Ćavar 2002; Pereltsvaig 2008; Fanselow & Féry 2013; Bondarenko & Davis accepted). For Tila Ch'ol, scattered deletion can be safely excluded: there is nothing in the theory that could restrict its application to all and only complements. This restriction is straightforward to capture on a movement-based account, however. As pointed out in Little (2020a), the fact that Tila Ch'ol bans discontinuous NPs which correspond to specifiers can be easily understood in light of the fact that specifiers are islands in a variety of languages.[11] This thus leaves two options on the table for Tila Ch'ol: direct subextraction and remnant movement.

Direct subextraction of the left branch from [[Num Cl] N] yields (54) in a simple and straightforward manner. Bale et al. (2019) put aside the possibility of deriving (54) from [Num [Cl N]] with remnant movement in a footnote, citing two reasons why they do not find it a plausible account of their data. Firstly, they do not see independent motivation for the first step of the derivation which moves NP/nP out of the nominal phrase. Secondly, they suggest that the remnant movement approach cannot capture the parallels between Num + Cl movement and other types of subextraction (see (55)–(67)). I will take up these concerns in turn, and show that they can, in fact, be easily eliminated.

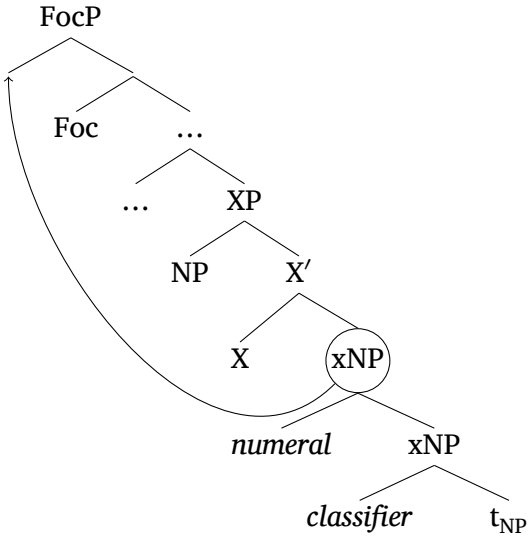
Let us start with the issue of what might motivate the evacuation of the noun from the extended NP (73).

[11] If it turns out to be the case that languages which appear to allow subextraction from specifiers make use of covert resumption, in fact, then it is possible to make the stronger statement that specifiers are islands cross-linguistically. Exploring this possibility is beyond the scope of this work.

(73) step 1: evacuation



(74) step 2: remnant movement



As discussed in Bale et al. (2019) and Little (2020a), examples with full NP movement and with LBE are discourse-pragmatically different. While pied-piping the noun results in focus on the entire Noun Phrase, LBE focuses the quantity alone. Thus (75) is appropriate as a reply to 'who arrived?', while (76) can be used as a reply to 'how many women arrived?' or to correct a previous claim that one woman arrived.

(75) *Ch'ol* (Bale et al. 2019)
[Ux-tyikil x'ixik] ta' jul-i-yob __.
three-CL woman PFV arrive-ITV-EP-PL
'**Three women** arrived.'

(76) *Ch'ol* (Bale et al. 2019)
Ux-tyikil ta' jul-i-yob [__ x'ixik].
three-CL PFV arrive-ITV-EP-PL woman
'**Three** women arrived.'

The evacuation step in (73) is thus motivated by information-structural considerations. Specifically, it serves to background the noun and thus allows focus to be on the quantity only. (LBE of numerals in Russian and Serbo-Croatian has the same effect.)

Turning to the parallels between separated possessors ((55)–(58)) and separated numeral modifiers ((60)–(67)), these fall out from the remnant movement account without further ado. If the first movement that creates the remnant (73) takes the noun outside of the extended noun phrase, then the evacuating step *is* a case of subextraction from xNP, therefore the same constraints that apply to subextraction of other material (e.g., possessors) from xNP are expected to apply here. Since Tila Ch'ol disallows subextraction from specifiers, the evacuating step from the subject of transitive verbs will be blocked, and so the remnant cannot be created. On the other hand, if the landing site of the noun is somewhere high *within* the extended Noun Phrase, then it is the movement of the remnant to the focus position (74) that constitutes a case of subextraction from xNP, and the restrictions on subextraction are predicted to apply at this point. In both scenarios, it is fully predicted that specifiers and adjuncts require the whole xNP to be pied-piped to the left periphery.[12]

(77) *Ch'ol* (Bale et al. 2019)
[Ux-tyikil xk'aläl-ob] ta' y-il-ä-y-ety __.
three-CL girl-PL PFV A3-see-TV-EP-B2
'**Three girls** saw you.'

(78) *Ch'ol* (Bale et al. 2019)
*Ux-tyikil ta' y-il-ä-y-ety [__ xk'aläl-ob].
three-CL PFV A3-see-TV-EP-B2 girl-PL
Intended: '**Three** girls saw you.'

It is worth pointing out here that remnant movement has been profitably applied in the analysis of the Ch'ol extended VP as well in Coon (2010b) and Little (2020a).

To summarize, Bale et al. (2019) make a successful argument for subextraction being involved in the derivation of (54). They do not, however, make a convincing case for [[Num Cl] N] being the base-generated structure in the extended NP, as NP/nP subextraction from [Num [Cl N]] followed by remnant movement achieves exactly the same empirical coverage. Note again that for languages in which the classifier corresponds to a functional head, the derivation in (73)/(74) is fully compatible with Cl^0-to-Num^0 head movement in syntax.

[12] In Tumbalá Ch'ol possessors can extract from external subjects to some degree, and so can the Num + Cl unit (Little 2020a). This is again expected: if subextraction from a specifier is not blocked, then neither (73) nor (74) will create a problem.

Future work will no doubt uncover compelling arguments for one analysis over the other. But even if it turns out to be the case that the direct subextraction approach is on the right track, we should be careful with drawing conclusions for the underlying structure of the NP from the grammaticality of (54). This is because as shown by data across a variety of Slavic languages, it is possible for the LBE-fronted constituent to contain multiple elements in such a way that they certainly do *not* form a deep-structure constituent. This phenomenon has been termed as 'extraordinary LBE' (Bošković 2005). Illustrative examples are provided below.

(79) *Serbo-Croatian* (Bošković 2015)
Onu staru prodaje kuću.
that old sells house
'He is selling that old house.'

(80) *Polish* (Citko 2006)
Z którymi rozmawiałaś studentami?
with which you-talked students
'Which students did you talk to?'

(81) *Russian* (Lena Borise, p.c.)
Za ét-imi bolš-imi Maš-a pri-š-l-a apel'sin-ami.
for these-INS.PL big-INS.PL Masha-NOM.SG.F PFV-come-PST-SG.F orange-INS.PL
'Masha came for **these big** oranges.'

(82) *Russian* (Lena Borise, p.c.)
Tr-i bolš-ix Maš-a kup-i-l-a apel'sin-a.
three-ACC.PL big-GEN.PL Masha-NOM.SG.F buy-TH-PST-SG.F orange-GEN
'Masha bought **three big** oranges.'

This pattern is, of course, straightforwardly captured (in fact, predicted) under the remnant movement analysis, but proponents of the direct subextraction analysis of LBE have also found ways to address such data. These involve some sort of re-arrangement or restructuring of the constituency within the NP in one way or another before the extraction takes place (see Borsley & Jaworska 1988; Corver 1990; Radkevich 2010; Talić 2018; Martinović 2019 for various implementations of this idea). In either case, the constituent that undergoes A-bar fronting arises only at an intermediate step of the derivation.[13] As the constituents targeted by LBE can be

[13] One might wonder what would motivate this reanalysis. Radkevich (2010) shows that in Russian only clitic Ps participate in the pattern in (81), so the reanalysis makes reference to the morpho-phonological properties of the items involved. In this respect, it is interesting to recall that Ch'ol numerals and the classifier form a phonological word. This morpho-phonological property may be responsible for restructuring of the original constituency before extraction, and in languages where the classifier spells out the functional head Cl^0, this restructuring may even consist in complex head formation.

derived, LBE is not a reliable tool for probing into deep structure constituency in the NP.[14] This, in turn, disarms a key argument for [[Num Cl] N] in Bale et al. (2019).

Before closing this section, it should also be noted that the general logic of my argument does not depend on the fronting movement being specifically a case of LBE. *What for* split in Germanic provides independent evidence that the fronted part of Discontinuous Noun Phrases is not necessarily directly extracted from the NP. Consider the following examples from German.

(83) *German* (Leu 2008)
[**Über was**] der alles **für Sachen** Bescheid weiss!
about what he all for things information knows
'(It's amazing) how much the guy knows!'

(84) *German* (Leu 2008)
[**Mit was**] hast du den **für Leut-en** gerechnet?
with what have you PRT for people-DAT reckoned
'What kind of people did you expect?'

As Germanic languages do not allow LBE, the fronting of the bracketed constituent does not involve LBE. But as discussed in Abels (2003: Chap. 4); Leu (2008) and Leu (2015: Chap. 6), it cannot involve direct extraction in general, because *was* does not form a constituent with the adposition in the underlying representation. (See the referenced papers for a remnant movement analysis of these data.) This further strengthens the general point made in this section: there is more than one way to arrive at a split NP, and the fronted phrase may be a derived constituent.

6 Further arguments

In this section I turn to two additional arguments for [[Num Cl] N] that have been made in the literature on the basis of Ch'ol, and I show that these arguments do not carry force either.

6.1 The numeral-effect

Bale & Coon (2014) show that in Ch'ol (and Mi'gmaq; Eastern Algonquian), the choice of the specific numeral in the xNP has a direct influence on whether a classifier appears or not. Today many Ch'ol speakers use native numerals only for 1–6, 10, 20, 40, 80, 100 and 400; in other cases they use numerals borrowed from Spanish (Vázquez Álvarez 2011: 160; Bale et al. 2019). Important for our purposes is the fact that while native numerals require a classifier (85), those borrowed from Spanish block classifiers (86).

[14] Bošković (2016) explores the possibility of multiple focus fronting to account for the relevant data, but Bošković (2015) discusses several problems with this approach. Needless to say, there would be no argument for [[Num Cl] N] on the multiple fronting approach either.

(85) *Ch'ol* (adapted from Bale & Coon 2014)
ux-*(p'ej) tyumuty
three-CL egg
'three eggs'

(86) *Ch'ol* (adapted from Bale & Coon 2014)
nuebe-(*p'ej) tyumuty
nine-CL egg
'nine eggs'

Bale & Coon (2014) take this as evidence for Krifka's (1995) approach to noun phrases, whereby numerals first combine with a classifier and the resulting phrase modifies the noun: [[Num Cl] N].

In Dékány (2022b) I called the pattern in (85)–(86) the 'numeral-effect'. I showed that this effect can be observed beyond Ch'ol and Mi'gmaq, and demonstrated that in a wide array of languages it is the noun rather than the numeral that influences the appearance of the classifier (which I called the 'noun-effect', see also Simpson & Ngo 2018). I further argued against a view whereby the 'numeral-effect' correlates with [[Num Cl] N] and the 'noun-effect' correlates with [Num [Cl N]] (e.g., Little et al. 2022b). I showed that in Nivkh (Paleosiberian isolate) the 'numeral-effect' and the 'noun-effect' are observable within a single xNP. Since the classifier cannot form a constituent with Num and N at the same time, the 'numeral-effect' and the 'noun-effect' cannot reasonably be viewed as indicators of constituency in the fashion described above. Detailed argumentation can be found in Dékány (2022b); I refer the reader to this work for in-depth discussion.

6.2 Mutual dependency between Num and Cl

Little et al. (2022a) make the claim that in Ch'ol numerals and classifiers *never* appear without one another (see also Bale et al. 2019: 20 and Herrera's 2022 discussion of similar data in Nahuatl). This is most spectacular in the counting sequence, where the general classifier *p'ej* must accompany numerals.

(87) *Ch'ol* (Little et al. 2022a: ex. 45)
jum-*(p'ej), cha'-*(p'ej), ux-*(p'ej)
one-CL two-CL three-CL
'1, 2, 3, …'

Little et al. (2022a) suggest that in their use in the counting sequence, numerals are not accompanied by a noun. If there is no noun, but the classifier still appears, then it must form a constituent with the numeral: [[Num Cl] N]. I have a three-pronged rebuttal of this claim.

Firstly, any account of (87) must take into consideration not only syntactic but also morpho-phonological factors. As discussed in Sect. 2, Ch'ol numerals and classifiers form a phonological word. If it is indeed the case that they do not occur without one another, then it is extremely difficult to decide which of them is the bound element that needs a morpho-phonological host: the numeral, the classifier, or they are both bound elements that can, together, make up a phonologically independent word.[15] If numerals themselves are bound morphemes, then it follows that they cannot appear on their own, and should be augmented with some other element, e.g., a classifier, to be able to form independent words. This would then be akin to *do*-support in English. (Below I will provide evidence that numerals are indeed bound morphemes in Ch'ol.) While Little et al. (2022a) heavily downplay the importance of the bound nature of numerals, morpho-phonology – via the Stray Affix Filter – can fully explain the pattern in (87); no conclusions follow for syntactic constituency.

Secondly, I take issue with the claim that in (87) there is no noun. Little et al. (2022a: 26) characterize the Ch'ol counting sequence as follows: "In these instances, we propose that the generic classifier is counting abstract points on a number line […] Even though there is no overt noun, it is inferred that there is still some abstract object being counted (e.g., abstract points on a number line)." I could not agree more with Little et al.'s (2022a) assessment of (87), but I draw quite different conclusions from this than they do.

The literature distinguishes between two different ways of counting: abstract counting (as in the counting sequence in general and in abstract mathematical operations: *one, two, three*; *two plus two is four*) and concrete/object counting (where numerals quantify over a noun: *two cats*), see Greenberg (1978; 2000); Caha & Wągiel (2020); Wągiel & Caha (2021); Žoha et al. (2022). Little et al.'s (2022a) claim that (87) counts abstract points amounts to saying that it is an instance of object counting rather than abstract (serial) counting, which means that a(n elided or Kaynean silent) noun (denoting the abstract points) must also accompany the numeral. This, in turn, renders the 'no noun, hence, [[Num Cl] N]' argument moot.

My approach amounts to saying that Ch'ol does not make use of abstract counting at all: what appears to be abstract counting in (87) is object counting is disguise. Independent evidence that it is possible for a language to use object counting (rather than abstract counting) in the counting sequence comes from Ainu. In Southern Hokkaido Ainu object counting the lower numerals, namely 1–4, are juxtaposed to the counted noun, e.g., *sine pa* 'one year'. What is interesting about 1–4 is that they cannot appear on their own in the counting sequence either: here they are juxtaposed to a generic, morphologically bound noun, *-p(e)* 'thing' (Tamura 1988/2000). That

[15] Most (simplex) numerals and classifiers appear to be monosyllabic, but there are bisyllabic numerals (*waxäk-* 'eight', *bolom-* 'nine') and there is at least one bisyllabic classifier as well (*-tyikil* for people), thus the number of syllables cannot help to decide this matter.

is, the counting sequence in Ainu quite literally begins with ‘one thing, two things, three things, four things’.[16]

(88) *(Southern Hokkaido) Ainu* (Tamura 1988/2000: 254–256)

	serial counting	object counting with *pa* ‘year’
one	*sine-p*	*sine pa* (‘one year’)
two	*tu-p*	*tu pa* (‘two years’)
three	*re-p*	*re pa* (‘three years’)
four	*íne-p*	*íne pa* (‘four years’)

Taking the above quote from Little et al. (2022a) seriously, I hypothesize that Ch’ol uses the same strategy in the entire counting sequence as Ainu does for 1–4, with the surface-difference that the noun in Ch’ol is silent (via surface silence, i.e., ellipsis, or deep silence, i.e., a Kaynean inherently silent noun).

Finally, contrary to the claims made in the previous literature, numerals and classifiers do appear without one another in Ch’ol. Since their relation is not as tight as some would have it, the argument for Num + Cl constituency does not go through. Let us start with numerals without classifiers. So far, we have seen three types of complex numerals in Ch’ol: multiplicatives (34), disjunctive-approximative numerals (89) and overcounting numerals (90).

(89) *Ch’ol* (Vázquez Álvarez 2011: 255)
cha’-tyikil ux-tyikil kixtyañuj
two-CL three-CL SP:person
‘few people’

(90) *Ch’ol* (Bale et al. 2019: ex. 46a)
cha’-p’ej i-cha’-k’al
two-CL A3-two-CL.20
‘twenty two’

I take it to be uncontroversial that as suggested in Ionin & Matushansky (2018: 3), complex cardinals “are built using standard syntactic means […] as well as standard principles of semantic composition”. With their complex structure, (89) and (90) support this hypothesis (which, as far as I can tell, is also tacitly assumed in Bale et al. 2019).

In light of this it is interesting to note that in addition to the three types of complex numerals mentioned above, Ch’ol aslo has additive numerals for 11 through 19.

[16] On the higher numerals and an in-depth analysis of the Ainu numeral system, see Dékány (2022c).

(91) *Ch'ol* (Aulie 1957: 282)
wək-luhum-p'ehl
six-ten-CL.general
'sixteen'

Crucial to us is the fact that while in (89) and (90) there are two classifiers for the two simplex numerals that make up the complex numeral, in (91) the addend and the augend are simply juxtaposed: there is only one classifier on the surface, leaving the addend (*wək-*) without a classifier of its own.

The cornerstone of Bale et al.'s (2019) argumentation in connection with Ch'ol disjunctive and overcounting numerals was a WYSIWYG approach, whereby only the morphemes that we see on the surface are assumed to be present in the syntactic structure. Maintaining this approach for (91) means that the addend is a case of a numeral without a classifier. Bale et al. (2019) are right that numerals cannot appear as stand-alone morphemes. On the WYSIWYG approach (91) seems to indicate, however, that while numerals do need a morphological host, they are not very choosy regarding the category of host: either a classifier or another numeral will do, as long as it is not the case that a simplex numeral is mapped to a phonological word on its own (see also Appendix D).

The other logically possible approach to (91), of course, is to assume that the addend does have a classifier of its own, but it has been elided. This is more in line with my own assumptions about complex numerals. As indicated in Sect. 4, I would, in fact, suggest that the addend has both its own classifier and its own local instance of the counted noun: *wək-p'ehl-N-luhum-p'ehl N* (cf. Ionin & Matushansky 2018). But if the strictly WYSIWYG approach to Ch'ol numerals is abandoned, then (89) and (90) can no longer be leveraged for [[Num Cl] N] either.

Let us now turn to classifiers without numerals. As pointed out earlier, 20 and its powers occupy the Cl slot in Ch'ol (93) and they block regular classifiers (94).

(92) *Ch'ol* (Bale et al. 2019)
ux-kojty wakax
three-CL cow
'three cows'

(93) *Ch'ol* (Bale et al. 2019)
ux-k'al wakax
three-CL.20 cow
'sixty cows'

(94) *Ch'ol* (Bale et al. 2019)
*ux-k'al-kojty wakax
three-CL.20-CL cow
Intended: 'sixty cows'

Recall that 20 and its multiples are like general classifiers in that they do not provide information about the animacy, shape, etc. of the noun. Aulie (1957) points out, however, that there is a way to include this information in the Noun Phrase: for clarity, the speaker may follow up *-k'al* CL.20 by the morpheme *ti* and a regular classifier, for instance, *kojty/koht* for animals (95).

(95) *Tumbalá Ch'ol* (Aulie 1957: 282)
hun-k'al ti koht
one-CL.20 PREP CL
'twenty' (animals)

(96) shows a similar construction with a measure term rather than a classifier. This example is from the New Testament (3rd, online edition, first edition published in 1977). It should be taken into consideration that this is a translation into Tumbalá Ch'ol by the American Wilbur Aulie, the author of Aulie (1957).[17]

[Context: They measured the depth of the water…]

(96) *Ch'ol* (Wycliffe & La Liga Bíblica 2010: 246, Acts 27:28)
An jun-c'al ti jajl.
EXT one-CL.20 PREP arm.span
'It was twenty arm spans [deep].'

Aulie (1957) does not comment on the nature of the linker *ti*, but the same pattern is reported for related Tsotsil in Fleck (1981), Haviland (1981: 165–166) and de León Pasquel (1988: 64), and Fleck and de León Pasquel clearly identify this element as the all-purpose preposition of Tsotsil. The gloss for (96) provided by Carol Rose Little (p.c.) also suggests that the *ti* in question is the all-purpose preposition in Ch'ol as well (written in modern grammars as *tyi*). Based on this, I suggest that (95) is literally 'twenty in animal units'.

Carol Rose Little (p.c.) informs me that (95) is not productive any more for speakers born in or after the 1960's. This might be because as mentioned above, native numerals are increasingly replaced by Spanish borrowings, and as Little (p.c.) points out, numerals with *-k'al* are themselves not common any more. Nevertheless, diachronic data can provide important insights into grammar, and (95) clearly shows that classifiers could, at least into the 1950's, occur without numerals, and could even form stand-alone morphemes. Notice that it is not possible to assume some sort of covert numeral for 20 after *ti*. Firstly, the concept of 20 is expressed in a classifier rather than a numeral (93). Secondly, the lexeme for 20 cannot occur locally to a regular classifier: (94). The only numeral in (95) is *hun* 'one', but (95) means 'twenty animals' rather than 'one animal', thus a covert 'one' would not make sense for (95) (and I am not aware of independent support for classifier-stranding ellipsis of numerals in Ch'ol in the first place).

[17] I thank Carol Rose Little (p.c.) for pointing this example out to me.

That classifiers could be mapped onto phonological words on their own suggests that numerals and classifiers form a phonological word because of the needs of the numeral: Ch'ol numerals are prosodically dependent and lean to the right (i.e., they are prefixes or proclitics). This reinforces that it is worth thinking of the Num-Cl phonological constituency in the same way as the phonological constituency between Bantu prefixes and verbs (i.e., cliticization at PF rather than syntactic complex head formation, cf. Sect. 2 and the references cited there) and also lends additional credence to the idea that numerals and classifiers occur together in the counting sequence (87) because of the PF-requirements of numerals.

To summarize, the mutual dependency between numerals and classifiers is not as strong as previously thought, and where this dependency is observable, it is due to the morpho-phonological properties of numerals. As such, it is not indicative of constituency in the NP.

7 Conclusions

This paper considered four arguments for [[Num Cl] N] from Bale et al. (2019). The argument based on phonological constituency was recognized as inconclusive already in Bale et al. (2019). Two further arguments, namely those based on disjunctive coordination and A-bar extraction, have no probative value. These tests are known to produce false positives cross-linguistically, thus they cannot be pressed into service to adjudicate controversial cases. As for the argument based on overcounting numerals, here everything depends on what kind of structure one adopts for complex cardinals in general. Overcounting numerals can be captured on [Num [Cl N]] with Ionin & Matushansky's (2018) theory of complex numerals, developed independently of the issue of classifiers. The plausibility of this approach was supported with the shape of overcounting numerals in Ainu. An analysis along these lines is worked out in more detail for Ch'ol in Dékány (to appear).

At the end of the paper I also addressed a further constituency test, based on the mutual dependency between Num and Cl, put forth in Little et al. (2022a). I showed that the dependency between numerals and classifiers is (or at least until the 1950's was) unidirectional rather than bidirectional: classifiers can occur without numerals, but numerals cannot occur without classifiers. I argued that numerals are prosodically dependent and need classifiers (or some other suitable host) for phonological reasons. This alone can explain the occurrence of classifiers in the counting sequence; no conclusions can be drawn for syntactic constituency.

In my previous work I supported a universal [Num [Cl N]] structure for classifier expressions. In Dékány (2021) I showed that the standard of comparison is determined compositionally from the semantic contribution of the noun and the classifier. This semantic constituency aligns with a [Num [Cl N]] syntax but produces a syntax-semantics mismatch on [[Num Cl] N]. In Dékány (2022a) I argued that so-called autoclassifier (aka repeater) constructions support the [Num

[Cl N]] constituency, and in Dékány (to appear) I suggested that so does the semantic selection between the classifier and the noun.

In this paper, I did *not* show that the structure of xNPs in Ch'ol is [Num [Cl N]]. I also did *not* show that [[Num Cl] N] is definitely excluded for Ch'ol. The structure of classifier expressions, in UG as well as in specific languages, remains subject to an ongoing debate. What I hope to *have* shown is that going forward, we should not rely on the diagnostics proposed in Bale et al. (2019).

Appendix A: Suffix stranding NP-ellipsis

This appendix illustrates suffix stranding NP-ellipsis in more detail with data from Hungarian, and demonstrates that via this process numerals, too, can end up forming a morphological word with another morpheme with which they do not form a syntactic constituent. As shown in (97), numeral modifiers in Hungarian are prenominal and occur in the Dem > Num > Adj > N order. Participles can appear at multiple places within the extended NP; (97) features one in the immediately post-numeral position.

(97) *Hungarian* (own knowledge)
ez-t a három [tavaly örökbefogad-ott] aranyos vörös cicá-t
this-ACC the three last.year adopt-PTCP cute red cat-ACC
'these three cute red cats adopted last year'

NP ellipsis may scope over all of part of the post-numeral domain, but it leaves the case suffix intact. If only the nominal head is elided, then the color adjective bears the case suffix that would be appropriate for the noun if it was overt (98). We can be sure that the case suffix belongs to the elided noun rather than the adjective because adjectives, numerals and participles do not show concord (97), and the case suffix can only appear on the NP-modifier which is linearly closest to the stranded suffix.

(98) *Hungarian* (own knowledge)
ez-t a három [tavaly örökbefogad-ott] aranyos vörös-et
this-ACC the three last.year adopt-PTCP cute red-ACC
'these three cute red ones adopted last year'

If NP ellipsis affects the noun and the color adjective together, then the case suffix is hosted by the quality adjective (99). If both adjectives are elided with the noun, then the case affix will appear on the participle (100), and if the entire post-numeral domain is elided, then it will be affixed to the numeral (101). All examples below can be understood to be elliptical versions of the full DP in (97), with progressively bigger and bigger structure undergoing ellipsis.

(99) *Hungarian* (own knowledge)
ez-t a három [tavaly örökbefogad-ott] aranyos-at
this-ACC the three last.year adopt-PTCP cute-ACC
'these three cute ones adopted last year'

(100) *Hungarian* (own knowledge)
ez-t a három [tavaly örökbefogad-ott]-at
this-ACC the three last.year adopt-PTCP-ACC
'these three adopted last year'

(101) *Hungarian* (own knowledge)
ez-t a három-at
this-ACC the three-ACC
'these three'

Note that the appearance and quality of the linking vowel on the stranded suffix is determined by its newly found host: the adjective in (98) and (99), the participle in (100) and the numeral in (101). That the stranded suffix is in the vowel harmony domain of its new host is unambiguous indication that they are part of the same phonological world. But there is no doubt that in (98)–(101) the suffix does not form either a complex head or any other type of syntactic constituent with its host: it is merely adjacent to the host at PF.

The same phenomenon can be observed with the plural marker: it is left behind by NP ellipsis, and it leans onto the last remnant preceding the elliptical gap. This is illustrated below with an adjective and a participle. (The plural marker is in complementary distribution with numerals, so numerals will never act as a host for this type of affix.)

(102) *Hungarian* (own knowledge)
a [tavaly örökbefogad-ott] vörös cicá-k
the last.year adopt-PTCP red cat-PL
'the red cats adopted last year'

(103) *Hungarian* (own knowledge)
a [tavaly örökbefogad-ott] vörös-ek
the last.year adopt-PTCP red-PL
'the red ones adopted last year'

(104) *Hungarian* (own knowledge)
a [tavaly örökbefogad-ott]-ak
the last.year adopt-PTCP-PL
'the ones adopted last year'

Kester (1996) discusses similar data from the related North Sámi language. Outside of the Uralic and Quechua families (see Sect. 2), suffix stranding NP-ellipsis has been observed in Persian, Turkish and Basque (Lipták & Saab 2010) as well as the Northeast Caucasian languages Archi and Kubachi (Kibrik 1995).

Appendix B: Discontinuous NPs

In the main text I discussed three ways to separate the noun from the numeral and the classifier on the surface: direct extraction, remnant movement and scattered deletion. There is a further logical possibility as well. One may assume that in Ch'ol a fully extended NP with an empty

noun can be base-generated in the focus position, such that the empty noun is co-indexed with a clause-internal overt noun (which has no modifiers). Similarly to scatterd deletion, though, this approach could not capture the fact that discontinuous NPs in Ch'ol are restricted to complements.

(105) $[_{FocP}\ [_{NumP}$ Num Cl ~~N~~$_i]$ Foc [V N_i] rejected structure for Ch'ol

However, Singhapreecha & Sybesma (2015) persuasively argue that a similar base-generated relationship between Num + Cl on the one hand and a separated N on the other hand does materialize in Thai. In the relevant Thai examples the noun appears linearly first and the numeral plus the classifier occur at the end of the sentence, but the interpretation – similarly to Ch'ol and Slavic LBE – is focus on the quantity (Singhapreecha & Sybesma 2015).

(106) *Thai* (Singhapreecha & Sybesma 2015)
mii dèk taay săam khon
there.be child die three CL
'three children died' (children died, namely three)

Singhapreecha & Sybesma (2015) establish that such examples are bi-clausal, involving coordination between two CPs. The overt noun is in the first CP. The numeral and the classifier are in the second conjunct. They are fronted together with an empty noun (co-referential with the overt N in the first conjunct) to a left-peripheral focus position, and the rest of the clause is elided.

(107) $[_{CP} \ldots [_{xNP}\ noun_i\ ] \ldots\ ]$ & $[_{CP}\ [_{xNP}$ numeral classifier ~~noun~~$_i\ ] \ldots\ t_{xNP} \ldots\ ]$

Since the phrase containing the numeral and the classifier has been moved to the left periphery from a clause-internal position, the construction is subject to island constraints (e.g., the complex DP island). Critically, though, this does not mean that the Num + Cl unit has been subextracted from the NP of the overt noun. An approach that eschews subextraction from the Thai nominal phrase entirely (including the direct and NP/nP-evacuating type) is supported by two arguments: there is no restriction to complements (the subjects of all types of verbs, objects as well as arguments in a PP can occur in the split construction) and some apparently 'split' NPs cannot be 'put back together' into a single grammatical NP. This provides an interesting parallel with Ch'ol: even though the numeral and the classifier are separated from the noun on the surface, and island constraints are in effect in both cases, there is no argument for direct subextraction in either case.

Appendix C: Classifiers in Ch'ol complex numerals

Comparing disjunctive-approximative (108), overcounting (109) and additive (110) numerals, one may wonder why it is precisely additives that call for deletion of the classifier after the first numeral.

(108) *Ch'ol* (Vázquez Álvarez 2011: 255)
cha'-tyikil ux-tyikil kixtyañuj
two-CL three-CL SP:person
'few people'

(109) *Ch'ol* (Bale et al. 2019: ex. 46a)
cha'-p'ej i-cha'-k'al
two-CL A3-two-CL.20
'twenty two'

(110) *Ch'ol* (Aulie 1957: 282)
wək-luhum-p'ehl
six-ten-CL.general
'sixteen'

I believe that this question is best approached from the other direction, however: we should be asking why disjunctive-approximative and overcounting numerals would *not* surface-delete the classifier together with the noun after the first numeral.

For overcounting numerals, I suggest that this is because the two classifiers are not identical. The first classifier conveys information about the shape, size or disposition of the counted noun, while the second one contributes a numerical value (20, 400, etc.). With the two classifiers making different contributions to the semantic composition of the complex numeral, eliding either of them would lead to a loss of semantic information. With additives, the two classifiers are identical. Therefore deleting one of them does not lead to any loss of information. The situation here is quite similar to how English speakers prefer fragment answers to full sentences as answers to interrogatives: this is the more economical option in terms of the number of surface exponents. Additives like (110) are conventionalized numerals; this fact leads to a certain frequency of use, which probably further aids the reduction of the surface form. Conventionalized numerals also appear with reduced morphology in English: while coordination of ordinary NPs requires an overt coordinator (*John *(and) Mary came to visit*), not all coordinative numerals allow it (e.g., *twenty and one* is phrasal coordination rather than a conventionalized numeral).

The two classifiers are also indentical in the case of disjunctive-approximative numerals. I suggest that there is no deletion of the first classifier here because these are occasionally coined phrases rather than conventionalized 'real' complex numerals (cf. English *two-three people showed up* vs. *twentythree people showed up*), thus frequency of use has a much smaller role to play and cannot help reduction of the surface-form. The difference between disjunctive-approximative and additive numerals in classifier retention also aids processing. Neither type of numeral has an overt connector in the form of a conjunction or an adposition. The difference in the surface-realization of the first classifier is an overt cue as to how speakers should interpret the relation between the two numerals: as conjunction or disjunction.

Appendix D: Numerals without a classifier?

In the main text I showed that classifiers can (or up until the recent past could) occur without a numeral. Numerals, however, generally do not occur without a classifier. (The only potential exception is the first numeral of additive numerals, but here it is debatable if there really is no classifier after the first numeral in the underlying structure).

In his in-depth study of nominal classification in Tumbalá Ch'ol, Arcos López (2009: 36) discusses one context, however, in which numerals may perhaps occur without a classifier. He mentions that "the position of the numeral classifier can be occupied by" *ya*, a morpheme that counts repetitions (my translation).

(111) *Ch'ol* (Arcos López 2009: adapted from Chap. 2., ex. 9)
cha'-ya; chäm-ya
two-ya four-ya
'twice', 'four times'

I find it significant that Arcos López does not include *ya* in his detailed tables of classifiers and that in the text quoted above, he does not claim that *ya* is a classifier (instead, he just makes a statement about its position). In addition, he also remarks that "This morpheme also has its own characteristics that separate it from the classifiers" (p. 36, my translation). Although he does not specify what these stand-out properties are, it is clear that *ya* does not show the morphological trappings of classifiers. In the Ch'ol variety investigated by Arcos López, all classifiers (sortal and mensural alike) are built on C_1VC_2 roots with the help of *<j>* (see Sect. 2). The only exceptions to this are i) roots that already end in *j* (arguably a case of haplology), and ii) the sortal classifier *tyikil*, which is also exceptional in that it is specific to human-denoting nouns and it is the only bisyllabic classifier (among both sortals and mensurals). A handful of sortal classifiers are not built on C_1VC_2 roots but have an unknown etymology. With the exception of *tyikil*, these also all contain a *<j>* in their phonological make-up. Whether *ya* also has other properties that set it apart from classifiers is a topic for further study. However, both Arcos López' remark and the morphological dissimilarity to other classifiers seriously raises the possibility that *ya* is not a classifier, and so in (111) a numeral occurs without a Cl. That numerals are morphologically bound elements that need a host still stands, but (111) potentially shows that numerals are more promiscuous as far as the possible hosts are concerned than we thought before.

Abbreviations

A: set A person marking (~ ergative, genitive), ACC: accusative, ADN: adnumerative, B: set B person marking (~ absolutive), CL: classifier, COP: copula, DAT: dative, DET: determiner, DIR: directional, DTV: derived transitive, EP: epenthetic insertion, EXT: existential, F: feminine, FUT: future, GEN: genitive, HON: honorific, IMFV: imperfective aspect, INS: instrumental, ITV:

intransitive verb suffix, IV: status marker for intransitive verb in imperfective, LOC: locative, NOM: nominative, PASS: passive, PFV: perfective, PL: plural, POSS: possessive, PREP: preposition, PRFV: perfective, PROG: progressive, PRT: particle, PST: past, PTCP: participle, REL: relative, SG: singular, SP: Spanish borrowing, STAT: stative suffix, TH: theme vowel, TV: transitive verb suffix

Funding information

This work was funded by the János Bolyai Scholarship of the Hungarian Academy of Sciences, the ÚNKP-23-5 New National Excellence Program of the Ministry for Culture and Innovation as well as NKFIH grants KKP 129921 and FK 145985.

Acknowledgements

I am grateful to Jessica Coon for discussion regarding the distribution of C_1VjC_2 forms and to Lena Borise for providing the Russian examples. I also wish to thank Carol Rose Little, the reviewers, the handling editor and the audience of SinFonIJA 16 for their constructive feedback. All errors are mine.

Competing interests

The author has no competing interests to declare.

References

Abels, Klaus. 2003. *Successive cyclicity, anti-locality, and adposition stranding*. Storrs, CT: University of Connecticut dissertation.

Abels, Klaus. 2012. *Phases. An essay on cyclicity in syntax* (Linguistische Arbeiten 543). Berlin: Mouton de Gruyter. DOI: https://doi.org/10.1515/9783110284225

Alexiadou, Artemis & Gengel, Kirsten. 2012. NP ellipsis without focus movement/projections: The role of classifiers. In Kučerová, Ivona & Neeleman, Ad (eds.), *Contrasts and positions in information structure*, 177–205. New York: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511740084.010

Arcos López, Nicholás. 2009. *Los clasificadores numerales y las clases nominales en ch'ol*. Ciudad de México: CIESAS MA thesis.

Aulie, Wilbur. 1957. High-layered numerals in Chol (Mayan). *International Journal of American Linguistics* 23(4). 281–283. DOI: https://doi.org/10.1086/464420

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

Bale, Alan & Coon, Jessica & López, Nicholás Arcos. 2019. Classifiers, partitions, and measurements: Exploring the syntax and semantics of sortal classifiers. *Glossa* 4(1). Article 77. DOI: https://doi.org/10.5334/gjgl.752

Barrie, Michael & Mathieu, Eric. 2016. Noun incorporation and phrasal movement. *Natural Language and Linguistic Theory* 34(1). 1–51. DOI: https://doi.org/10.1007/s11049-015-9296-6

Bašić, Monika. 2008. On nominal subextractions in Serbian. *Balkanistica* 21. 1–56.

Beltran, Pedro. 1746. *Arte del idioma maya reducido a sucintas reglas*. Mérida, Yucatán, Mexico: Imprenta de J. D. Espinosa 2nd edn. 1859, José D. Espinosa (ed.). DOI: https://doi.org/10.5479/sil.266870.39088005811658

Blake, Barry J. 2001. *Case* (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press 2nd edn. First published in 1994.

Bondarenko, Tatiana & Davis, Colin. accepted. Concealed pied-piping in Russian: on Left Branch Extraction, parasitic gaps, and beyond. *Syntax*.

Borer, Hagit. 2005. *Structuring sense*, vol. I. New York: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199263905.001.0001

Borsley, Robert D. & Jaworska, Ewa. 1988. A note on prepositions and Case marking in Polish. *Linguistic Inquiry* 19(4). 685–691.

Bošković, Željko. 2005. On the locality of left-branch extraction and the structure of the NP. *Studia Liguistica* 59. 1–45. DOI: https://doi.org/10.1111/j.1467-9582.2005.00118.x

Bošković, Željko. 2015. On multiple Left-Branch Dislocation: multiple extraction and/or scattered deletion? In Szajbel-Keck, Małgorzata & Burns, Roslyn & Kavitskaya, Darya (eds.), *Proceedings of FASL 23*, 20–35. Ann Arbor: Michigan Slavic Publications.

Bošković, Željko. 2016. Getting really edgy: On the edge of the edge. *Linguistic Inquiry* 47(1). 1–33. DOI: https://doi.org/10.1162/LING_a_00203

Branigan, Phil & Brittain, Julie & Dyck, Carrie. 2005. Balancing syntax and prosody in the Algonquian verb complex. In Wolfart, H. Christopher (ed.), *Algonquian conference*, vol. 36, 56–69. Winnipeg: University of Manitoba.

Caha, Pavel & Wągiel, Marcin. 2020. Universal semantic features and the typology of cardinal numerals. *Catalan Journal of Linguistics* 19. 199–229. DOI: https://doi.org/10.5565/rev/catjl.296

Cheng, Lisa L.-S. & Sybesma, Rint. 2005. Classifiers in four varieties in Chinese. In Cinque, Guglielmo & Kayne, Richard S. (eds.), *The Oxford handbook of comparative syntax*, 259–292. New York: Oxford University Press. DOI: https://doi.org/10.1093/oxfordhb/9780195136517.013.0007

Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective*. New York: Oxford University Press. DOI: https://doi.org/10.1093/oso/9780195115260.001.0001

Citko, Barbara. 2006. The interaction between across-the-borad wh-movement and left-branch extraction. *Syntax* 9(3). 225–247. DOI: https://doi.org/10.1111/j.1467-9612.2006.00092.x

Clausen, Mica. 2023. Classifiers as complex elements. Handout of a talk delivered at NYU, 5 Oct 2023. https://gitlab.com/mica-linguistics/classifiers-as-complex-elements/-/blob/main/lecture.pdf.

Compton, Richard & Pittman, Christine. 2010. Word-formation by phase in Inuit. *Lingua* 120(9). 2167–2192. DOI: https://doi.org/10.1016/j.lingua.2010.03.012

Coon, Jessica. 2009. Interrogative possessors and the problem with pied-piping in Chol. *Linguistic Inquiry* 40(1). 165–175. DOI: https://doi.org/10.1162/ling.2009.40.1.165

Coon, Jessica. 2010a. *Complementation in Chol (Mayan): a theory of split ergativity*. Cambridge, MA: MIT dissertation.

Coon, Jessica. 2010b. VOS as predicate fronting in Chol. *Lingua* 120. 354–378. DOI: https://doi.org/10.1016/j.lingua.2008.07.006

Coon, Jessica. 2013. *Aspects of split ergativity*. New York: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199858743.001.0001

Coon, Jessica. 2017a. Ch'ol. In Aissen, Judith & England, Nora C. & Maldonado, Roberto Zavala (eds.), *The Mayan languages* (Routledge Language Family Series), 648–684. London and New York: Routledge.

Coon, Jessica. 2017b. Little-v^0 agreement and templatic morphology in Ch'ol. *Syntax* 20(2). 101–137. DOI: https://doi.org/10.1111/synt.12135

Coon, Jessica & Preminger, Omer. 2009. Positional roots and case absorption. In Avelino, Heriberto & Coon, Jessica & Norcliffe, Elisabeth (eds.), *New perspectives in Mayan linguistics: Proceedings of SSILA 2008 (The Society for the Study of Indigenous Languages of the Americas)*, 35–58. Cambridge, MA: MIT Working Papers in Linguistics. DOI: https://doi.org/10.1515/THLI.2009.004

Corver, Norbert. 1990. *The syntax of left branch extractions*. Tilburg: Tilburg University dissertation. DOI: https://doi.org/10.1515/9783110849998-004

de León Pasquel, María de Lourdes. 1988. *Noun and numeral classifiers in Mixtec and Tzotzil: A referential view*. Sussex: University of Sussex dissertation.

Dékány, Éva. 2021. *The Hungarian nominal functional sequence*. Cham: Springer. DOI: https://doi.org/10.1007/978-3-030-63441-4

Dékány, Éva. 2022a. Auto-classifiers in Thai: A post-syntactic reduplication account. In Du, Yuqiao & Chen, Zhuo & Li, Xiangyu & Xu, Zetao & Pan, Victor Junnan (eds.), *Proceedings of GLOW in Asia XIII*, 48–58. Hong Kong: Department of Linguistics and Modern Languages, The Chinese University of Hong Kong.

Dékány, Éva. 2022b. Classifiers for nouns, classifiers for numerals. In Pratley, Breanna & Bakay, Özge & Neu, Eva & Deal, Peyton (eds.), *Proceedings of the Fifty Second Annual Meeting of the North East Linguistic Society*, 245–258. Amherst, MA: GLSA.

Dékány, Éva. 2022c. The syntax of numeral modification in Ainu. Ms., HRCL.

Dékány, Éva. to appear. Anatomy of a complex numeral: overcounting, with special attention to Ch'ol. In Gribanova, Vera & Clemens, Lauren & Scontras, Gregory (eds.), *Syntax in uncharted territories: Essays in honor of Masha Polinsky*, eScholarship.

Embick, David & Noyer, Rolf. 2001. Movement operations after syntax. *Linguistic Inquiry* 32(4). 555–595. DOI: https://doi.org/10.1162/002438901753373005

Ershova, Ksenia. 2020. Two paths to polysynthesis: Evidence from West Circassian nominalizations. *Natural Language and Linguistic Theory* 38(2). 425–475. DOI: https://doi.org/10.1007/s11049-019-09455-w

Fanselow, Gisbert & Ćavar, Damir. 2002. Distributed deletion. In Alexiadou, Artemis (ed.), *Theoretical approaches to universals*, 65–107. Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/la.49.05fan

Fanselow, Gisbert & Féry, Caroline. 2013. A comparative perspective on intervention effects on left branch extractions in Slavic. In Sulym, Wolodymyr & Smolij, Mychajlo & Djakiw, Chrystyna (eds.), *Non progredi est regredi. Festschrift für Alla Paslawska*, 266–295. Lviv: Pais.

Fleck, Margaret M. 1981. Tzotzil numeral root morphology. *Journal of Mayan linguistics* 3. 5–24.

Franks, Steven & Progovac, Ljiljana. 1994. On the placement of Serbo-Croatian clitics. *Indiana Linguistic Studies* 7. 69–78.

Greenberg, Joseph. 2000. Numeral. In Booij, Geert & Lehmann, Christian & Mugdan, Joachim & Skopeteas, Stavros & Kesselheim, Wolfgang (eds.), *Morphologie/morphology. Ein Internationales Handbuch zur Flexion und Wortbildung. An international handbook on inflextion and word-formation*, vol. 1, 770–783. Berlin and New York: Walter de Gruyter.

Greenberg, Joseph H. 1978. Generalizations about numeral systems. In Greenberg, Joseph H. (ed.), *Universals of human language*, vol. 3: Word structure, 249–295. Stanford, CA: Stanford University Press.

Grinevald, Colette. 2004. Classifiers. In Booij, Geert & Lehmann, Christian & Mugdan, Joachim & Skopeteas, Stavros & Kesselheim, Wolfgang (eds.), *Morphologie/Morphology. Ein Internationales Handbuch zur Flexion und Wortbildung. An international handbook on inflection and word-formation*, vol. 2, 1016–1031. Berlin and New York: Walter de Gruyter.

Halle, Moris & Marantz, Alec. 1993. Distributed morphology and the pieces of inflection. In Hale, Ken & Keyser, Samuel Jay (eds.), *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger*, 111–176. Cambridge, MA: MIT Press.

Harizanov, Boris & Gribanova, Vera. 2019. Whither head movement? *Natural Language and Linguistic Theory* 37(2). 461–552. DOI: https://doi.org/10.1007/s11049-018-9420-5

Harrison, K. David. 2007. *When languages die: The extinction of the world's languages and the erosion of human knowledge*. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780195181920.001.0001

Haugen, Jason D. 2009. Hyponymous objects and late insertion. *Lingua* 119(2). 242–262. DOI: https://doi.org/10.1016/j.lingua.2007.10.008

Haviland, John Beard. 1981. *Sk'op sotz'leb: El Tzotzil de San Lorenzo Zinacantán*. México D.F.: Universidad Nacional Autónoma de México, Instituto de Investigaciones Filológicas, Centro de Estudios Mayas.

Herrera, Rafael. 2022. Numeral constructions with classifiers in sixteenth-century central Nahuatl. In Lee, Seung Suk & Song, Yixiao (eds.), *Proceedings of Semantics of Under-Represented Languages in the Americas (SULA) 11*, 87–96. Amherst, MA: GLSA.

Hurford, James R. 2003. The interaction between numerals and nouns. In Plank, Frans (ed.), *Noun Phrase structure in the languages of Europe*, 561–620. Berlin and New York: Mouton. DOI: https://doi.org/10.1515/9783110197075.4.561

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

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

Jenks, Peter. 2011. *The hidden structure of Thai Noun Phrases*. Cambridge, MA: Harvard University, Cambridge, MA dissertation.

Kester, Ellen-Petra. 1996. Adjectival inflection and the licensing of empty categories in DP. *Journal of Linguistics* 32(1). 57–78. DOI: https://doi.org/10.1017/S0022226700000761

Kibrik, Aleksandr E. 1995. Direct-oblique agreement of attributes in Daghestanian. In Plank, Frans (ed.), *Double case: Agreement by suffixaufnahme*, 216–229. New York: Oxford University Press. DOI: https://doi.org/10.1093/oso/9780195087758.003.0005

Koopman, Hilda Judith & Szabolcsi, Anna. 2000. *Verbal complexes* (Current Studies in Linguistics 34). Cambridge, MA: MIT Press. DOI: https://doi.org/10.7551/mitpress/7090.001.0001

Krifka, Manfred. 1995. Common nouns: A contrastive analysis of English and Chinese. In Carlson, Greg N. & Pelletier, Jeff F. (eds.), *The generic book*, 398–411. Chicago, Il: University of Chicago Press.

Law, Danny. 2020. Pattern borrowing, linguistic similarity, and new categories: Numeral classifiers in Mayan. *Morphology* 30. 347–372. DOI: https://doi.org/10.1007/s11525-020-09361-7

Law, Danny. 2024. The history of numeral classifiers in Mayan. In Gutiérrez Bravo, Rodrigo & Skopeteas, Stavros & Verhoeven, Elisabeth (eds.), *Glottothèque: Mayan languages*, Berlin/Göttingen/Mexico City: online resource. DOI: https://doi.org/10.5281/zenodo.xxxx

Leu, Thomas. 2008. *What for* internally. *Syntax* 11(1). 1–25. DOI: https://doi.org/10.1111/j.1467-9612.2008.00113.x

Leu, Thomas. 2015. *The architecture of determiners* (Oxford studies in comparative syntax). New York: Oxford University Press.

Lipták, Anikó & Saab, Andrés. 2010. Movement and deletion after syntax. Talk delivered at the Workshop on Morphosyntax-Phonology Interface Theories, Universiteit Leiden, December 08, 2010.

Lipták, Anikó & Saab, Andrés. 2016. Movement and deletion after syntax: Licensing by inflection reconsidered. *Studia Linguistica* 70(1). 66–108. DOI: https://doi.org/10.1111/stul.12039

Little, Carol-Rose. 2020a. Left branch extraction, object shift, and freezing effects in Tumbalá Ch'ol. *Glossa* 5(1). Article 26. DOI: https://doi.org/10.5334/gjgl.988

Little, Carol-Rose. 2020b. *Mutual dependencies of nominal and clausal syntax in Ch'ol*. Ithaca, NY: Cornell University dissertation.

Little, Carol Rose & Moroney, Mary & Royer, Justin. 2022a. Classifiers can be for numerals *or* nouns: Two strategies for numeral modification. *Glossa* 7(1). 1–35. DOI: https://doi.org/10.16995/glossa.8437

Little, Carol Rose & Moroney, Mary & Royer, Justin. 2022b. Classifying classifiers. In Lee, Seung Suk & Song, Yixiao (eds.), *Proceedings of the Eleventh Conference on the Semantics of Underrepresented Languages in the Americas*, 145–160. Amherst, MA: GSLA.

Marantz, Alec. 1984. *On the nature of grammatical relations* (Linguistic Inquiry monographs). Cambridge, Mass: MIT Press.

Marantz, Alec. 1988. Clitics, morphological merger, and the mapping to phonological structure. In Hammond, Michael & Noonan, Michael (eds.), *Theoretical morphology*, 253–270. San Diego, CA: Academic Press. DOI: https://doi.org/10.1163/9789004454101_017

Marcos-Marín, Francisco A. 1992. Typology of numerals: Minority languages of China. In Luksaneeyanawin, Sudaporn (ed.), *Pan-Asiatic linguistics. Proceedings of the Third International Symposium on Language and Linguistics*, vol. II, 754–763. Bangkok: Chulalongkorn University Printing House.

Martinović, Martina. 2019. Interleaving syntax and postsyntax: Spellout before syntactic movement. *Syntax* 22(4). 378–418. DOI: https://doi.org/10.1111/synt.12169

Mateu, Jeume. 2012. Conflation and incorporation processes in resultative constructions. In Demonte, Violeta & McNally, Louise (eds.), *Telicity, change, and state: A cross-categorial view of event structure*, 252–278. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199693498.003.0010

Mathieu, Eric. 2012. Flavors of division. *Linguistic Inquiry* 43(4). 650–679. DOI: https://doi.org/10.1162/ling_a_00110

Mchombo, Sam. 2004. *The syntax of Chichewa* (Cambridge Syntax Guides). Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511486302

Merrifield, William R. 1966. Number names in four languages of Mexico. In Corstius, H. Brandt (ed.), *Grammars for number names* (Foundations of Language Supplementary Series 7), 91–102. Dordrecht: Springer. DOI: https://doi.org/10.1007/978-94-010-3449-4_6

Ochiai, Izumi. 2014. Overcounting in North-Central Vanuatu languages and its distribution. *Kyoto University Linguistic Research* 33. 229–252. DOI: https://doi.org/10.14989/196272

Ott, Dennis. 2011. Diminutive-formation in German: Spelling out the classifier analysis. *Journal of Comparative Germanic Linguistics* 14(1). 1–46. DOI: https://doi.org/10.1007/s10828-010-9040-x

Pereltsvaig, Asya. 2008. Split phrases in colloquial Russian. *Studia Linguistica* 62(1). 5–38. DOI: https://doi.org/10.1111/j.1467-9582.2007.00141.x

Radkevich, Nina. 2010. *On location: The structure of case and adpositions*. Storrs, CT: University of Connecticut dissertation.

Ross, John Robert. 1967. *Constraints on variables in syntax*. Cambridge, MA: MIT dissertation.

Simpson, Andrew. 2005. Classifiers and DP structure in Southeast Asia. In Cinque, Guglielmo & Kayne, Richard S. (eds.), *The Oxford handbook of comparative syntax*, 806–838. New York: Oxford University Press. DOI: https://doi.org/10.1093/oxfordhb/9780195136517.013.0018

Simpson, Andrew & Ngo, Bingh. 2018. Classifier syntax in Vietnamese. *Journal of East Asian Linguistics* 27(3). 211–246. DOI: https://doi.org/10.1007/s10831-018-9181-5

Singhapreecha, Pornsiri & Sybesma, Rint. 2015. An ellipsis analysis of split nominals in Thai. *Cahiers de Linguistique Asie Orientale* 44(1). 57–88. DOI: https://doi.org/10.1163/19606028-00441p03

Talić, Aida. 2018. Upward P-cliticization, accent shift, and extraction out of PP. *Natural Language and Linguistic Theory* 37(3). 1103–1143. DOI: https://doi.org/10.1007/s11049-018-9424-1

Tamura, Suzuko. 1988/2000. *The Ainu language* (ICHEL Linguistic Studies 2). Tokyo: Sanseidō. Originally published in Japanese under the title "Ainugo" (The Ainu language). In Kamei, Takashi & Kōno, Rokurō & Eiichi, Chino (eds.), *Gengogaku daijiten (Encyclopaedia of linguistics)*, 6–94, Tokyo: Sanseidō.

Tang, Chih-Chen Jane. 1990. *Chinese phrase structure and the extended X-bar theory*. Ithaca, NY: Cornell University dissertation. http://ishare.iask.sina.com.cn/f/15214444.html.

Vázquez Álvarez, Juan Jesús. 2011. *A grammar of Chol, a Mayan language*. Austin, TX: The University of Texas at Austin dissertation.

Wągiel, Marcin & Caha, Pavel. 2021. Complex simplex numerals. *Acta Linguistica Academica* 68(4). 470–515. DOI: https://doi.org/10.1556/2062.2021.00460

Wycliffe & La Liga Bíblica. 2010. *Jini nuevo testamento (el nuevo testamento en el ch'ol de Tumbalá) [The New Testament in Tumbalá Ch'ol]*. La Liga Bíblica versión electrónica edn.

Yasugi, Yoshiho. 1995. *Native Middle American languages: An areal-typological perspective* (Senri ethnological studies 39). Osaka: National Museum of Ethnology.

Žoha, Lukáš & Wągiel, Marcin & Caha, Pavel. 2022. The morphology of compound numerals: A cross-linguistic case-study. *Linguistics Beyond and Within* 8. 200–217. DOI: https://doi.org/10.31743/lingbaw.14966