

Number names: internal structure and morphological marking

Authors:

Norbert Corver (Utrecht Institute of Linguistics-OTS, Utrecht University,
n.f.m.corver@uu.nl)

Yuta Tatsumi (Meikai University, ytats0074@meikai.ac.jp)

Abstract:

Knowledge of number names is a subpart of our knowledge of language. This chapter discusses various components of this knowledge with a special focus on the internal structure of number names and their morphological marking. The following issues are addressed: the categorial nature of number names; the computational system underlying the formation of complex number names; the structural position of (components of) number names; the various morpho-phonological and morpho-syntactic operations that are operative on numerals, such as suppletion, contraction, and (c)overtness of functional heads; the formation of ordinals and the formation of counting numerals. The phenomenon of number names is considered from a comparative-linguistic perspective: number names from both classifier languages and non-classifier languages (macro-variation) are discussed. Furthermore, certain formal phenomena related to number names are discussed from a cross-dialectal (micro-variation) and intra-linguistic perspective. At a more general level, this chapter aims to show that there is a system of rules (a grammar) underlying the formation of number names in human language.

Key words:

simple cardinal; multiplicative complex cardinal (MCC); additive complex cardinal (ACC); ordinal; ordinal suppletion; classifier; ‘dressed’ numeral; counting numeral; collective numeral.

1. Introduction

Knowledge of language and knowledge of the number system are two cognitive capacities characteristic of human beings. An important question at the interface of these two capacities is the following: how is knowledge of the number system represented linguistically in human language? This general research question raises a variety of sub-questions, including the following: What is the categorial nature of number names? In what ways can simple number names such as ‘four’, ‘eight’ and ‘hundred’ be combined into more complex number names, as, for example, in the English number name ‘four hundred and eight’. Also, what kinds of linguistic constructs result from such combinatorial operations? Specifically, do they constitute complex words (morphology), or do they constitute composite syntactic constituents (syntax)? Another question regards the issue of cross-linguistic uniformity and variability: to what extent are languages similar or different from each other regarding the ways in which they compose number names. Although it seems fair to say that the grammar of number names has not figured very prominently on the formal-linguistic research agenda, it is also clear that the interest in this empirical domain of human language has increased substantially in recent years. The aim of this chapter is to give an overview of the major questions, empirical findings, and theoretical insights that can be found in studies on number names from the last couple of decades.

This chapter is organized as follows: Section 2 addresses the question about the categorial nature of simple (i.e. non-composite) cardinals and their locus in larger linguistic configurations. Section 3 discusses the formation and structure of ‘multiplicative complex cardinals’, such as English *six hundred* (6×100). In section 4, the same is done for ‘additive complex cardinals’, such as English *twenty five* ($20 + 5$). While Sections 2-4 focus on the grammar of number names in non-classifier languages, Section 5 shifts the focus to the grammar of number names in so-called classifier languages, addressing the same kinds of questions that are posed in Sections 2-4. In Section 6, the inner structure of ordinal numerals is discussed, paying attention to issues such as: (i) cross-linguistic strategies used for the formation of ordinals, (ii) the categorial nature of ordinals, (iii) their derivation from cardinals, and (iv) the derivation of ordinals from complex cardinals, including the phenomenon of ordinal suppletion. Section 7 presents a discussion of several constructions featuring a numeral in combination with a bound morpheme, specifically: composite forms such as English *six-ty* (section 7.1), so-called “dressed numerals” in certain Germanic varieties (section 7.2), numerals used in counting (section 7.3), and collective numerals featuring plural morphology (7.4). Section 8 concludes this chapter.

2. Simple cardinals: their categorial nature and distribution

Numerals are notorious for being difficult to classify (Corbett 1978, Donohue 2005). In traditional grammar, numerals have been identified as a separate class (i.e., category) within the parts of speech. The picture that emerges from more recent studies of numerals, however, is one in which there is not something like a category ‘numeral’. Rather, classes of elements exist that share the function of indicating the cardinality of the set denoted by the noun. By examining the morphosyntactic behavior of numerals, both intra- and interlinguistically, researchers have tried to identify the types of categories to which so-called numerals belong. This line of research has led to two major hypotheses about the categorial status of numerals: (i) the *lexical* hypothesis, which takes numerals to be lexical categories, more specifically nouns (N) or adjectives (A)—see, among others, Jackendoff (1977), Corbett (1978), Hurford (1998), Zweig (2005), Ionin and Matushansky (2006, 2018)—and (ii) the *functional* hypothesis, which analyzes numerals as functional categories, such as, Q(uantifier)—cf. Selkirk (1977), Barbiers (1990), Cardinaletti and Giusti (1992)—Num(erical)—cf. Ritter (1992), Rutkowski and Szczegot, (2001). There is a third, ‘in-between’ hypothesis, the *semi-lexical* hypothesis (Klockmann 2017), which treats numerals as categories that display both lexical (e.g., nominal or adjectival) properties and functional ones (e.g., its quantificational function). The three hypotheses are schematically represented in (1), where the hierarchical relationship between the numeral and the following noun is left unspecified.:

- (1) a. [N/A numeral] + noun
- b. [Q/Num numeral] + noun
- c. [[L+F] numeral] + noun (L = lexical property, F = functional property)

In the literature on the grammar of numerals, arguments have been given in support of each of the above approaches. Evidence in support of the adjectival analysis typically comes from the morphosyntactic property of agreement; that is, the numeral displays agreement with the quantified noun. Examples (2a,b), for instance, show that the Romanian (cardinal) numeral ‘two’ can have two forms, depending on the gender property of the plural noun with which it combines. The Polish patterns in (3) exemplify agreement in Case between the numeral and the quantified noun.

- (2) a. doi copii (Romanian)
 two children_{Masc.Pl.}
- b. două case_{Fem.Pl}
 two houses
- (3) a. Trzej lingwisci spali (Polish; Rutkowski and Maliszewska 2007)
 three.NOM linguists.NOM slept.PLUR.VIR
 ‘Three linguists were sleeping.’
- b. Ona tanczyła z trzema lingwistami.
 She danced with three.INSTR linguists.INSTR
 ‘She danced with three linguists.’

Besides morphosyntactic agreement with the quantified noun, the distributional behavior of the numeral may also be indicative of the adjectival status of a numeral. As shown in (4), for example, the Arabic numeral *wahid* ‘one’ (but also *'itnān* ‘two’), which agrees with the numerated noun in definiteness, case, and gender, must occur in post-nominal position, which is the position in which attributive adjectives also typically appear (Al-Bataineh and Branigan 2020).

- (4) a. rajul-u-n wāhid-u-n
 man-NOM-N one.MASC-NOM-N
 ‘one man’
- b. 'imra'at-u-n wāhid-at-u-n
 woman-NOM-N one-FEM-NOM-N
 ‘one woman’

Arguments that have been adduced in support of the nominal status of numerals typically relate to the combinatorial behavior of the numeral; that is, numerals can cooccur with elements that typically combine with nouns. As shown in (5), for example, numerals can be modified by an attributive adjectival modifier in certain languages; see Jackendoff (1977:128–130) for English, and Corver and Zwarts (2004) for Dutch.

- (5) a. We had a *good* two hours for visiting the museum. (English)
- b. Ik had een *kleine* tien minuten om me voor te bereiden. (Dutch)
 I had a small ten minutes for me PRT to prepare
 ‘I had close on ten minutes to prepare myself.’

In these examples the attributive adjectives *good* (English) and *kleine* (Dutch) qualify the cardinality expressed by the numeral.

A second co-occurrence phenomenon that hints at the nominal status of certain numerals comes from patterns featuring a determiner-like element in combination with a numeral. The examples in (5) already gave an illustration of this combinatorial pattern. The singular indefinite article *a/een* in (5a,b) cannot possibly belong to the plural noun *hours/minuten*. From this one might draw the conclusion that the indefinite article combines with the numeral (*two/tien*).¹ Given the fact that definite articles typically combine with

¹ See Kayne (2005a), though, for an alternative analysis according to which the singular indefinite article combines with a silent singular noun NUMBER, which takes the numeral (but also an attributive AP like *good*) as a satellite constituent [*a [good] [two] NUMBER hours*]. See also Zamparelli (2004) and Ionin and

nouns, one might conclude that the numerals in (5) are nouns categorially. Similar illustrations of this argument are given in (6), where the quantifier *every/iedere*, which typically combines with a singular noun, seems to belong to the numeral (*two/twee*), and not to the plural noun (*seconds/seconden*):

- (6) a. [Every two seconds] someone in the U.S. needs blood.
 b. [Iedere twee seconden] kwam er een auto langs. (Dutch)
 every two seconds came there a car along
 ‘Every two seconds a car passed by’

A third type of co-occurrence argument in support of the nominal status of numerals comes from patterns featuring a numeral that is followed by a numerated noun bearing genitival case. Examples (7)-(8) illustrate this case pattern; (7) is drawn from Al-Bataineh and Branigan (2020), (8) from Klockmann (2017); see also Ionin and Matushansky (2004a,b).²

- (7) a. talāt-at rijāl-i-n b. talāt nisā'-i-n
 three-FEM men-GEN-N three.MASC women-GEN-N
 ‘three men’ ‘three women’
- (8) Pięć ptaków spało
 five.NOM birds.GEN slept.N.SG
 ‘Five birds slept.’

Since the ability to assign genitival case is characteristic of nouns, one may draw the conclusion that the numerals in (7)-(8), which are followed by a noun bearing genitival case, are actually nouns.

As has been shown in the literature on numerals, many languages display internal variation as regards the properties of numerals, some featuring adjectival characteristics, others nominal characteristics. This language-internal variation does not seem to be totally arbitrary, though: if a language has varying numerals, we typically find adjectival properties with lower numerals, and more nominal properties with higher numerals, where the ‘adjective-noun’ threshold may vary across languages (see Jespersen (1969), Corbett (1978), Hurford (1987; 2003), Zabbal (2005), Zweig (2005), Klockman 2017). In Arabic, for example, the numerals *wāhid* ‘one’ and *'itnān* ‘two’ display adjectival behavior while the numerals ‘three’ to ‘ten’ display nominal behavior (Al-Bataineh and Branigan 2020). The former display agreement with the numerated noun (in gender, case and definiteness) and occur in postnominal position, just like attributive adjectives; see (4) for *wāhid* ‘one’ and (9) for *'itnān* ‘two’. The latter display noun-like behavior; as shown in (7a,b), for example, the numeral ‘three’ does not follow but precedes the counted noun, which has to take the genitive plural form. Arguably, this genitival case is assigned by the noun-like numeral.

- (9) a. al-rajul-ā-n al-'itn-ā-n

Matushansky (2018) for discussion of these nominal constructions featuring attributive APs in combination with numerals.

² The Arabic examples in (7) display reverse gender agreement (‘gender polarity’), that is, the numeral takes the masculine marker when defining a feminine noun, and vice versa. See Al-Bataineh and Branigan 2020.

	the-man-DUAL-N	the-two-DUAL-N
	‘two men’	
b.	'imra'at-ā-n	'itn-at-ā-n
	woman-DUAL-N	two-FEM-DUAL-N
	‘two women’	

A second illustration of language-internal variation regarding the categorial status of numerals comes from Luganda, a Bantu language. As Zweig (2005) points out, numerals lower than a certain threshold (often ‘five’ or ‘ten’) agree with the counted noun, featuring adjectival or enumerative agreement prefixes; see Ashton *et al.* (1954). In (10a), this is exemplified for the numeral ‘two’. Higher numerals such as ‘seven’ in (10b) do not display agreement with the counted noun. Instead, the numeral carries its own nominal class prefix.

- (10) a. emi-dumu e-biri
 MI-jug AGR_{MI}-two
 ‘two jugs’
 b. emi-dumu mu-sanvu
 MI-jug MU-seven
 ‘seven jugs’

As noted in Zabbal (2005), English also displays variation as regards the behavior of numerals. As exemplified in (11)-(12), low numerals such as *three*, *eight* and *nine* cannot appear with plural inflection (11), but high numerals such as *hundred*, *thousand* and *million* can (12); see also Zweig (2005). On the basis of this morphosyntactic contrast, Zabbal concludes that English low numerals are adjectival, whereas high numerals are nominal.

- (11) a. *John ate threes (of) eggs.
 b. *Eights (of) people gathered in the stadium.
- (12) a. John ate hundreds of eggs.
 b. Thousands of people gathered in the stadium.

Zabbal makes the same distinction for French on the basis of the morphological behavior of numerals, specifically, the attachment behavior of the suffix *-aine*, which refers to a vague number. As shown in (13), the high numeral *cent* ‘hundred’ can carry the suffix *-aine*, the low numeral *sept* cannot.

- (13) a. Jean a acheté des centaines de livres
 J has bought INDEF.PL hundreds of books
 ‘John bought hundreds of books’
 b. *Jean a acheté des septaines de livres
 J has bought INDEF.PL sevens of books

Besides addressing the issue about the categorial nature of numerals —that is, what kind of lexical atoms are they?— the study of numerals has also addressed the question as to what kind of syntactic position they occupy within the noun phrase in which they are embedded. Two major approaches can be distinguished in the literature on numerals: (I) the ‘NUM(eral) = head’ analysis, and (II) the ‘NUM = phrase’ analysis. According to the former, the numeral is a head that takes the counted noun as its complement; according to the latter the numeral is a phrase that occupies a specifier (or adjunct) position within the (extended) nominal

projection. Within the ‘NUM = head’ analysis, two variants can be identified: (a) the ‘NUM = *lexical* head analysis’ (e.g., Ionin and Matushansky 2004, 2018), and (ii) the ‘NUM = *functional* head analysis (e.g., Ritter 1991, Zamparelli 1995, Zabbal 2005). Within the ‘NUM = phrase’ analysis, we can also make a distinction between two approaches: (a) the approach according to which phrasal numeral is a satellite constituent (specifier or adjunct) within a larger lexical phrase (e.g. Selkirk 1977, Jackendoff 1977), and (ii) the approach that states that the phrasal numeral is a satellite constituent within a larger functional phrase (e.g. Corver and Zwartz 2004, Kayne 2005a). The four approaches to the syntactic placement of numerals within the noun phrase are represented in (14), where English is used as a language of illustration.

- (14) a. [LP specifier [L' [L three] [NP cars]]] (NUM = Lexical head)
 b. [FP specifier [F' [F three] [NP cars]]] (NUM = Functional head (Num/Q/Card)
 c. [NP [YP three] [N' [N cars]]] (NUM = phrase within lexical projection)
 d. [FP [YP three] [F' F [NP cars]]] (NUM = phrase within functional projection)

Corver and Zwartz (2006) try to give evidence for the phrasal analysis of (Dutch) numerals, specifically representation (14d), on the basis of noun phrases containing numeral constructs with an approximative meaning. Some of these ‘approximative numerals’ are given in (15a-c), where the noun phrase in which they are embedded functions as a direct object:

- (15) a. Ik heb [[zo'n veertig] vrienden] uitgenodigd.
 I have so-a forty friends invited
 ‘I invited approximately forty friends’
 b. ...[[iets van veertig] vrienden]
 ...something of forty friends...
 c. ...[[tegen de veertig] vrienden]...
 ...against the forty friends...

In (15a), the numeral *veertig* combines with the singular indefinite demonstrative pronoun *zo'n* (so-a/such-a, meaning ‘approximately’). Since singular *zo'n* cannot combine with plural *vrienden* ‘friends’, Corver and Zwartz conclude that *zo'n veertig* forms a phrasal unit within the extended nominal projection. In (15b), the numeral *veertig* is part of the nominal construct *iets van veertig*. The approximative pattern *iets van X* is also attested with regular nouns, as in *Hebben jullie [iets van pils] in huis?* (have you something of beer in house; ‘Have you got any beer or beer-like liquor?’). Given the phrasal status of the nominal expression *iets van pils*, one may conclude that the sequence *iets van veertig* in (15b) also constitutes a phrasal constituent. Consider, finally, the pattern in (15c), which contains the approximative pattern *tegen de veertig*, which is comparable to English *about forty*, as in *about forty friends*. Corver and Zwartz, who label these numeral constructs ‘prepositional numerals’, propose that these constructs are prepositional phrases that are embedded as pre-nominal satellites within the extended nominal projection.

- (16) [DP [PP tegen de veertig] vrienden].

One argument for saying that these prepositional numerals are prepositional *phrases* is the fact that they can be modified by modifiers that typically occur in prepositional environments:

- (17) Er waren [DP [PP ruim/ver boven de veertig] vrienden] op het feest
 there were amply/far above the forty friends at the party

‘There were far more than forty friends at the party.’

In sum, certain numeral patterns with approximative meaning seem to support the idea that cardinal numerals have a phrasal status when they act as satellites within the extended nominal projection.³ If this conclusion is correct, then this suggests that (certain) numeral *expressions* are not heads but rather phrases.

Our discussion so far has briefly touched on issues regarding the categorial nature and syntactic placement of simple numerals. In recent years, various formal-linguistic studies of the grammar of numerals have appeared. For more detailed and in-depth analyses of the categorial status and distributional behavior of (simple) numerals, we refer the reader to the following (incomplete) overview, which is organized around a number of language families:⁴

(18) *Studies on the grammar of numerals:*

- *Germanic*: Barbiers (1990), Bhatt (1990), Booij (2010), Coppen (1991), Corver (2000; 2012), Corver & Zwarts (2006), Delsing (1993), Jackendoff (1977), Kayne (2010, 2016), Löbel (1989), Selkirk (1977), Van Katwijk (1965);
- *Romance*: Cardinaletti & Giusti (1992), Giusti (1991), Doetjes (1997), Zamparelli (1995);
- *Slavic*: Asinari (2019), Bailyn (2004), Franks (1994), Ionin and Matushansky (2006; 2018), Klockman (2017), Leko (1999), Miechowicz-Mathiasen (2012), Rappaport (2002), Rutkowski and Szczegot (2001), Witkoś and Dziubała-Szrejbrowska (2018);
- *Celtic*: Hurford (1975), Kane (2014);
- *Finno-Ugric*: Norris (2018), Csirmaz and Dekany (2018);
- *Semitic*: Al-Bataineh and Branigan (2020), Alqarni (2015), Borer (2005), Danon (2009; 2012), Fassi Fehri (2018), Marcin (2016), Ritter (1991, 1992), Shlonsky (2004), Zabbal (2005);
- *Bantu*: Zweig (2005), Zerbian and Krifka (2008);
- *Niger-Congo*: Aboh (2004), Ndimele and Chan. (2016), Perekhvalskaya and Vydrin (2019), Makeeva and Shluinsky (2020), Pozdniakov (2018);
- *Sino-Tibetan*: Matisoff (1997), He (2015), He *et al* (2017), Her (2017);
- *Japanese*: Downing (1984, 1996);
- *Korean*: An (2018), Ahn (2019);
- *Mayan*: Keller (1955), Little (2018);
- *Australian Languages*: Bowern and Zentz (2012);
- *Austronesian/Papuan Languages*: Kramer and Kratochvíl (2014);
- *Sign Languages*: Ktejik (2013), Zeshan *et al* (2013), Sagara (2016), Zeshan and Sagara (2016), Safar et al (2018);

³ For further discussion of prepositional numerals, see Plank (2004a), Wege (1997), Pankau (2018). Pankau, basing himself on prepositional numerals in German (see (i)), argues that the PP-analysis proposed in Corver and Zwarts (2006) does not extend to German. Instead of analyzing *an die siebzig* as a prepositional phrase, he proposes that the sequence *an die* is an adverb (i.e. a single lexeme) that functions as a modifier adjoined to the numeral *siebzig*. This means that the lexeme *{an die}* patterns with an adverb like *ungefähr* ‘approximately’, as in *ungefähr siebzig* ‘approximately seventy’.

(i) An die siebzig Flüchtlinge sind erstickt
 on the seventy refugees are suffocated
 ‘Up to seventy refugees suffocated.’

⁴ For typological studies that take a cross-linguistic perspective on the structure and distribution of numerals, see Greenberg (1972; 1978; 2000), Corbett (1978), Comrie (2013; no year).

3. Multiplicative complex cardinals

The numeral *six hundred* in *six hundred cars* is a composite numeral expression. It consists of two meaningful components —*six* and *hundred*— which together express the arithmetical relation of multiplication: 6×100 . Since *six hundred* expresses multiplication, it is called a ‘Multiplicative Complex Cardinal’ (henceforth: MCC). The numeral *six* is the multiplier, the numeral *hundred* the multiplicand. Besides *hundred*, numerals such as *thousand*, *million*, *billion*, *et cetera* can be used as the lexical base for multiplication, as in the MCCs *six thousand* and *six million*. An MCC can itself function as a multiplier, as in *six hundred thousand*, where *six hundred* is the multiplier and *thousand* the multiplicand. As shown in (19a-d), MCCs occur in typologically different languages:⁵

- (19) a. tri stotine studenata (Serbo-Croatian; Indo-European, Slavic)
three hundred.ACC.F student.GEN.PL.M
'three hundred students'
- b. ?arba ?-u mi?at-in rajul-in (Arabic, Semitic, Zabbal 2005:86)
four-NOM hundred-GEN man-GEN
'four hundred men'
- c. san byaku-nin no gakusei (Japanese)
three hundred-CLS GEN student-ACC
'three hundred students'

If MCCs have a composite structure, the question obviously arises as to what it looks like. A question to start with is the following: is MCC a complex *word* having an inner structure or is it a (syntactic) *phrase* with an inner structure? Notice that the complex word analysis would be compatible with those approaches that take numerals to be syntactic heads selecting the counted noun as its complement (see (14a,b)). Schematically: [[word *three hundred*] + NP (*cars*)]. In Booij (2002:165-66; 2010), it is proposed that Dutch MCCs such as *driehonderd* (three-hundred) and *zesduizend* are morphological compounds, which means they have the representation in (20), where *Num(eral)* is used as a label for the sake of illustration (see also De Haas and Trommelen 1993).⁶

⁵ It should be noted that, cross-linguistically, languages may differ as regards the lexical base that is used for the formation of MCCs. Mandarin, for example, uses a decimal base (e.g. *wǔ-shí*, five ten, ‘50’), French uses a vigesimal base for the expression of the number 80 (e.g. *quatre-vingts*, [4 x 20], ‘80’), and Ekari (Papuan) uses ‘60’ as a base for the formation of MCCs, as in *muto wii* (60 x 4, ‘240’); see Comrie (no year) for further discussion.

⁶ That numerals can be components of compound words is shown by the Dutch examples in (i); see Booij (2002:142-15), Ackema and Neeleman (2004:83).

- (i) a. drie-hoek (three-corner, ‘triangle’), drie-klank (three-sound, ‘triptong’), duizend-poot (thousand-leg, ‘centipede’)
- b. drie-hoek-ig (three-corner-ed), vijf-kopp-ig (five-head-ed), negen-deel-ig (nine-part-ig, ‘nine-fold’)
- c. drie-kamer-flat (thee-room-apartment), vier-landen-punt (four-countries-point, ‘quadripoint’)

Some illustrations from English:

- (ii) a. four-square, four-score, five-finger, six-pence, ten-pin, three-coat, three-master
- b. four-legged, four-barrelled, five-toothed, two-handed, two-angled, six-leaved, eight-sided
- c. three-bedroom apartment, four-country meeting

- (20) [Num [Num drie] [Num honderd]]

Potential support for the compound analysis of MCCs may come from patterns in which the MCC is part of a larger word.⁷ Consider, for instance, the following examples from Dutch and English:

- (21) In the last 60 years, there has been a {three-fold / hundred-fold / three-hundred-fold} increase in the amount of pesticide used in farming.

- (22) De {drie-voud-ig / honderd-voud-ig / drie-honderd-voud-ig} internationale wordt
the three-fold-*ig* / hundred-fold-*ig* / three-hundred-fold-*ig* internationale becomes

team-manager van het Nederlands honkbalteam.

team-manager of the Dutch baseball-team

‘The threefold/hundredfold/threehundredfold international player will become team manager of the Dutch baseball-team.

In the English example (21), the simple numerals *three* and *hundred*, but also the MCC *three hundred*, are followed by the suffixal element *-fold*. This may be interpreted as evidence in support of a compound analysis of MCCs. We find the same patterns in the Dutch example in (22), the only difference being that the numeral is followed by two overt bound morphemes, namely the suffix *-voud* ('-fold') and the suffix *-ig* ('-y').⁸

Besides morphological approaches, which take MCCs to be complex words (i.e., compounds), there have also been proposals that MCCs do not constitute complex words (morphological compounds) but rather complex syntactic constructs (e.g. Hurford (1975), Ionin and Matushansky 2018, Zabbal (2005), Tatsumi (2021)). One argument for saying that multiplier and multiplicand form a syntactic object comes from phenomena displaying syntactic interactions between the multiplier and the multiplicand. In Luganda, for example, the number 20 is expressed by the MCC ‘ten two’, where ‘ten’ (*kumi*) is the multiplicand and ‘two’ (*biri*) the multiplier (Zweig 2005). As shown in (23), the multiplier *biri* shows class agreement with the multiplicand *kumi*, and not with the counted noun *dumu* ‘jug’.

- (23) emi-dumu ama-kumi a-biri (Luganda)
 MI-jug MA-ten AGR_{MA}-two
 ‘twenty jugs’

⁷ We say ‘potential support’ because syntactic approaches have been proposed in which “derivational” suffixes are part of syntax. See, for example, *Kayne (2005b)* for arguments that so-called derivational suffixes are functional heads that trigger computational operations (e.g. head movement) in syntax. *Kayne*, for example, lists the following bound morphemes as functional elements relevant to syntax: (i) *-th*, as in *two hundred and fiftieth*, (ii) *-ish*, as in *John looks thirty-five-ish*, (iii) French *-aine*, as in *une vingt-aine d’articles* ‘a twenty-aine of articles, ‘about twenty articles’). Under this approach, Dutch *-voud* and *-ig*, and English *-fold*, would be treated as functional heads that are active in a syntactic structure. See also section 7.

⁸ The morpheme *-voud* seems to be nominal given the fact that it can be part of a head displaying nominal behavior, as shown in (i):

- (i) Je moet de brochure [pp in [NP drievoud / honderdvoud / tweehonderdvoud]] inleveren.
you must the brochure in three-fold / hundred-fold/twohundred-fold hand.in
'You must hand in {three/ hundred / two hundred} copies of the brochure.

The lexical items ending in *-voud* function as nominal complements (NPs) of the preposition *in*. If *drievoud* is a compound noun, then *drievoudig* in (22) has the following structure: $[_A /_{NuM} \text{drie}] + [_N -\text{voud}] + [-ig]$.

Another argument against morphological approaches comes from the so-called approximative inversion (AI) construction in Russian. In Russian, when a head noun precedes a cardinal, the resulting phrase has an approximate interpretation, as shown in (24b).

- (24) a. pjat' knig
five books.GEN
'five books'
b. knig pjat'
books.GEN five
'approximately five books'

Crucially, when AI is applied to a noun modified by a multiplicative complex cardinal that contains multiplicands such as *tysjač* 'thousand' as its multiplicand, the multiplicand must be fronted, instead of the head noun. In (25), the cardinal *sorok* 'forty' functions as a multiplier and *tysjač* 'thousand' as a multiplicand. (25c) shows that approximative inversion of the head noun is blocked by an overt multiplicand. (Δ stands for the original position of the inverted element.)

- (25) a. sorok tysjač mašin [Russian]
forty thousand.PL.GEN car.PL.GEN
'forty thousand cars'
b. tysjač₁ sorok Δ_1 mašin
thousand.PL.GEN forty car.PL.GEN
'some forty thousand cars'
c. *mašin₁ sorok tysjač Δ_1
car.PL.GEN forty thousand.PL.GEN
Int. 'some forty thousand cars' [Ionin and Matushansky (2018: 118-119)]

There are several analyses of the relevant inversion process (see Pereltsvaig (2006) for a head movement analysis and Matushansky (2015) for a phrasal movement analysis), but we will not go into this issue in detail here. What is important for the current discussion is that a part of the MCC (i.e. the multiplicand in (25)) can be the target of the syntactic inversion process. This pattern would be unexpected under the morphological approaches.

It is important to note here that we are not claiming that all MCCs are formed syntactically. For instance, Russian has two MCCs with the multiplicand 'hundred'. As shown in (26), the non-fused MCC allows for AI. On the other hand, the fused MCC in (27a) resists approximate inversion, as shown in (27b). This contrast shows that the fused MCC in (27a) is opaque to the syntactic inversion process. The fused MCC is a morphological complex word as predicted by the morphological approaches.

- (26) Russian: Pasha Kovel, p.c.
a. pjat' soten knig b. soten pjat' knig
five hundred book hundred five book
'five hundred books' 'approximately five hundred books'
(100 \times n, where n is ± 5)
(27) a. pjat'-sot knig b. *sot pjat' knig
five-hundred book hundred five book
'five hundred books' 'approximately five hundred books'
(100 \times n, where n is ± 5)

Another argument is related to the so-called Lexical Integrity (LI) Hypothesis. There are several ways of stating the relevant phenomena. The LI hypothesis has been stated in different ways in the literature. Lapointe (1980) refers to the property in question as Generalized

Lexicalist Hypothesis, Selkirk (1982) as Word Structure Autonomy Condition, and Di Sciullo and Williams (1987) as The Atomicity Thesis. In this chapter, we set aside the exact nature of this property located between syntax and morphology. What is important for the current discussion is that there is a tendency that some syntactic operations cannot be applied within a part of a word formed before syntax. Bresnan and Mchombo (1995) provide several tests of LI. Keeping the LI hypothesis in mind, let us now consider (28a) below (similar data are also noted by Roeper and Siegel (1978: 252)).

- (28) a. *taxi, truck driver (\approx taxi driver or truck driver)
 b. six, seven thousand (\approx around six thousand or seven thousand)

In (28a), the base element of the compound (i.e. *driv(-er)*) is associated with two elements preceding it. (28a) is unacceptable under the disjunctive reading.⁹ In contrast to (28a), the multiplicand *thousand* (28b) can be associated with two multipliers, yielding the approximative reading. If all MCCs are generally analyzed as involving morphological compounding, as in (28a), the contrast in (28) is not expected.¹⁰ Although it has been reported that there are some exceptions to LI (cf. Lieber and Scalise 2005, especially regarding synthetic compounds), the contrast in (28) shows that at least a strong version of the LI

⁹ It is possible though to interpret (28a) as "a driver of a taxi truck", or "a truck driver who works as a taxi driver". We thank Nicolaus Schrum for bringing our attention to this point.

¹⁰ It should be noted that MCCs and compounds like *truck driver* behave alike for some other tests of the LI hypothesis. For instance, the first element can be questioned by using a *wh*-element both in the compound and the MCC in (i).

- (i) a. [What driver] did he see yesterday?
 b. [How many thousands] did he earn last month?

Moreover, when the overt disjunction particle is present, the contrast in (28a) disappears, as shown in (ii).

- (ii) a. taxi or truck driver (\approx taxi driver or truck driver)
 b. six or seven thousand (\approx around six thousand or seven thousand)

These examples can be analyzed in different ways as represented below.

- (iii) Phrases within words
 a. [[taxi or truck] driver] (\approx taxi driver or truck driver)
 b. [[six or seven] thousand] (\approx around six thousand or seven thousand)
 (iv) Gapping
 a. [taxi ~~driver~~] or [truck driver] (\approx taxi driver or truck driver)
 b. [six ~~thousand~~] or [seven thousand] (\approx around six thousand or seven thousand)

The synthetic compound in (iia) itself poses an interesting challenge to the LI hypothesis, and we have to put aside these similarities between synthetic compounds and MCCs in this chapter. (But see section 4 below regarding nominal ellipsis and related ellipsis constructions appearing in MCCs.)

Approximative numerals also give a piece of evidence that MCCs are opaque to some phrasal modification. As shown in (v), when the comparative modifier *more than* attaches to an MCC, the resulting interpretation is something like ' n books, where $n \geq 300$ ', but not ' $n \times 100$ books, where $n \geq 3$ '. This indicates that the "more than" phrase modifies the whole MCC as in (vb), and cannot be a modifier of the multiplier alone, as represented in (vc). (Notice that the intended reading in (vc) is available in the approximative inversion in Russian, as shown in (26b).)

- (v) a. John read more than three hundred books. (n books, where $n \geq 300$, $*n \times 100$ books, where $n \geq 3$)
 b. [more than] [three hundred] books. (n books, where $n \geq 300$)
 c. *[more than] three hundred books. ($n \times 100$ books, where $n \geq 3$)

hypothesis does not hold for complex cardinals. This is not expected under the morphological approaches to complex cardinals.

If MCCs do not constitute (complex) word units but are part of a larger syntactic structure, the question arises as to how they are syntactically integrated into that structure. From a constituency perspective, two major approaches can be distinguished, each of which is schematically represented in (29) on the basis of English:

- | | | |
|------|-----------------------------|--------------------------|
| (29) | a. [three [hundred [cars]]] | MCC is not a constituent |
| | b. [[three [hundred]] cars] | MCC is a constituent |

According to analysis (29a), MCC (*three hundred*) does not form a single constituent that combines with the counted noun *cars*. Rather, the multiplicand (*hundred*) and the main noun (*cars*) form a constituent, which subsequently combines with the multiplier *three*. In other words, the multiplicand and the multiplier independently enter into a syntactic (and semantic) relationship with their sister constituent. Consider next (29b). According to this analysis, MCC (*three hundred*) does form a single constituent. It is the entire MCC which combines with the main noun.

In their analysis of MCCs, Ionin and Matushansky (2018) argue that there are two ways in which a (simple) numeral can be combined with a sister constituent: firstly, a numeral is a nominal head selecting an NP; secondly, a numeral is an attributive AP adjoined to NP.¹¹ The two strategies are illustrated in (30), where English is simply used as a language of illustration.

- | | | |
|------|--|------------------------|
| (30) | a. [_{NP} three _N [_{NP} cars]] | (Head-Complement) |
| | b. [_{NP} three _{AP} [_{NP} cars]]] | (Adjunction Structure) |

The pattern in (30a) is available in languages in which the numeral is involved in case assignment, the latter in languages in which the numeral displays (adjectival) agreement with N(P).

The two strategies in (30) yield the following ‘cascade structures’ for MCCs (where English is used again as a language of illustration):

- | | | |
|------|---|-----------|
| (31) | a. [_{NP} three _N [_{NP} hundred _N [_{NP} cars]]]] | H-C & H-C |
| | b. [_{NP} three _{AP} [_{NP} hundred _{AP} [_{NP} cars]]]] | AS & AS |

Ionin and Matushansky (2018:76) note that they “are not aware of any language in which all multiplicands are adjectival;” that is, the pattern in (31a). It thus seems that multiplicands tend to be nominal. Why this is so remains unanswered.

Importantly, Ionin and Matushansky (2006, 2018) assume that the multiplier and the multiplicand in each of the representations in (31) are syntactically (and semantically) alike. Other researchers adopting the non-constituent MCC-analysis (for certain languages), however, have argued that multiplier and multiplicand should be treated differently, both syntactically and semantically. In Tatsumi (2021), for example, it is proposed that the multiplicand is a (functional) syntactic *head* whose specifier position is occupied by a *phrasal* multiplicand. Schematically, based on the sequence *three hundred cars*:¹²

¹¹ Note that pattern (30a) instantiates the lexical head analysis of numerals (see (14a)), and pattern (30b) the phrasal satellite analysis (see (14c)).

¹² According to Tatsumi, multiplicands such as *hundred* and *thousand* occupy the same syntactic position as classifiers in classifier-languages such as Japanese and Chinese (see section 5).

- (32) [FP [YP three] [F' hundred [NP cars]]]

Tatsumi argues that the structural representation in (32), in which the multiplier (*three*) and the multiplicand (*hundred*) occupy structurally different positions, receives support from a number of phenomena, one of them being NP-ellipsis. Consider, for example, the Spanish patterns in (33b,c), which are reduced variants of (33a); that is, a subpart of the direct object noun phrase has been elided:

- (33) a. Juan tomó [seiscientas fotos], y Maria tomó [trescientas fotos].
 Juan took six.hundred pictures and Maria took three.hundred pictures
 ‘Juan took six hundred pictures, and Maria took three hundred pictures.’
 b.y Maria tomó [trescientas Δ].
 c.y Maria tomó [tres Δ].
 *...and Maria took three hundred pictures
 ..and Maria took three pictures

The NP-ellipsis pattern *trescientas* in (33b) has the same interpretation as *trescientas fotos* in (33a). Thus, the elided part in *trescientas* corresponds to the constituent *fotos*. Consider next example (33c), where *tres* ‘three’ is the remnant of NP-ellipsis. Importantly, the reading in which the elided part corresponds to *cienas fotos* is excluded. Thus, *tres* in (33c) cannot be interpreted as *trescientas fotos*. It can only be interpreted as ‘three photos’. According to Tatsumi, this follows from the structure in (32) — *in casu*, [FP [YP tres] [F' cienas [NP fotos]]] — in combination with the assumption that only maximal phrases (i.e. XPs, but not the intermediate-level category X') can be input to a deletion operation such as NP-ellipsis. In other words, the NP *fotos* can be deleted, but not the X'-constituent *cienas fotos*. As Tatsumi points out, the fact that *trescientas* can be a remnant of NP-ellipsis, but not *tres*, seems unexpected if one adopts a cascade structure like those in (31): under this analysis, the ellipsis site (viz., NP) and the licenser of the ellipsis site (viz., the numeral) would be the same for (33b) and (33c).

When we look at NP-ellipsis in English MCCs, we see more or less the same pattern as in Spanish. Specifically, many speakers feel a contrast between (34b) and (34c). For some speakers, though, it is not completely impossible for *two* in (34c) to be interpreted as “two hundred books”. In (34c), */OK means that there is speaker variation regarding the acceptability of the intended reading. (Ionin and Matushansky (2006: 338) also report a similar observation in a footnote.)

- (34) a. John read three hundred books, but Mary read two hundred books.
 b. John read three hundred books, but Mary read two hundred Δ.
 c. */OK John read three hundred books, but Mary read two Δ.

The speaker variation depicted in (34c) suggests that at least for some speakers, English MCCs may have the cascading structure in (31a). For those who do not accept (34c) under the intended reading (i.e. *two* Δ = “two hundred books”), MCCs seem to have the structure in (32).

Interestingly, also for English it can be observed that NP-ellipsis becomes easier when the remnant of ellipsis is a higher multiplicand. Even for the speakers who cannot obtain the intended reading of (34c), nominal ellipsis with higher multiplicands such as *thousand* and *million* is more acceptable, as in (36) and (37) (Nicolaus Schrum, p.c.).

Clearly, this contrast behavior between lower multipliers and higher multipliers in NP-ellipsis contexts is in need of more detailed cross-linguistic investigation.

Another phenomenon that illustrates the asymmetric behavior of multipliers and multiplicands comes from the phenomenon of pronominalization/substitution. In Dutch, for example, so-called ‘quantitative *er*’ (litt.: there, ‘of them’) can substitute for part of a noun phrase leaving the cardinal as a remnant (Blom 1977, Bennis 1986):

- (38) Jan heeft gisteren [driehonderd mensen] gevaccineerd
Jan has yesterday three hundred people vaccinated ...
'Jan vaccinated three hundred people yesterday....'

a. ..en hij heeft *er* vandaag [tweehonderd *e*] gevaccineerd.
..and he has *er* today two-hundred vaccinated
..'and he vaccinated two hundred Δ today.' (Δ = people)

b. ..en hij heeft *er* vandaag [twee *e*] gevaccineerd.
..and he has *er* today three vaccinated
..'and he vaccinated three Δ today.' (Δ = people; *hundred people)

In (38a), quantitative *er* substitutes for the nominal expression *mensen* ‘people’. Example (38b) can only receive the reading in which *er* substitutes for ‘people’, which implies that Jan has vaccinated just two people. Thus, it cannot receive a reading in which *er* stands for *honderd mensen* ‘hundred people’. Also here, one might argue that this follows from an analysis in which *twee* ‘two’ and *honderd* ‘hundred’ occupy structurally different positions, as in the structural representation in (32); specifically, $[_{FP} [_{YP} \text{twee}] [_{F'} \text{honderd} [_{NP} \text{mensen}]]]$. Under the assumption that quantitative *er* can only substitute for maximal projections (*in casu* NP) and not for intermediate projections (*in casu*, the F'-constituent *honderd mensen*), it immediately follows that (38b) can only have the reading ‘two people’, and not the reading ‘two hundred people’. Also in this case, a cascading structure like (31a)—*in casu*, $[_{NP} \text{twee}_N [_{NP} \text{honderd}_N [_{NP} \text{mensen}]]]$ —faces the question as to why only the NP *mensen* can be substituted for by quantitative *er*, and not the NP *honderd mensen*.¹³

¹³ It should be noted that the impossible readings of Spanish (33c) and Dutch (38b) follow directly from a compound analysis of complex numerals: the right hand member of the compound-numeral and the counted noun that follows, do not form a unit. Schematically for Spanish: [noun phrase [Num [Num tres] / Num cientos]] fotos].

Besides the analyses in (31) and (32), in which the string ‘multiplier + multiplicand’ does not form a constituent, there have also been proposals which treat a numeral expression like *three hundred* as a constituent, as in (29b) above, repeated here as (39):

- (39) [[three [hundred]] cars] MCC is a constituent

Zweig (2005), for example, building on Kayne’s (2003, 2005a) analysis of quantifiers *few* and *many*, proposes that MCCs such as *three hundred* constitute noun phrases (NP), with the high numeral *hundred* being the head of the noun phrase, and the low numeral *three* being an adjectival modifier, as depicted in (40a). When the numeral expression consists only of a simple numeral, as in *three cars*, the numeral expression has the same inner organization, the only difference being that the head of the noun phrase is a silent noun (NUMBER). This yields the structure in (40b):

- (40) a. [[NP [AP three] hundred] cars]
b. [[NP [AP three] NUMBER] cars]

Just like Zweig (2005), Zabbal (2005) proposes that an MCC like *three hundred cars* constitutes a single phrase that combines with the counted noun *cars*. He also argues that the low numeral *three* is adjectival and that the high numeral *hundred* is nominal. Zabbal’s position as regards the inner structure of the phrase *three hundred* is different, though, from Zweig’s. Specifically, he claims that it is not the high numeral (*hundred*) which constitutes MCC’s head, but rather the low numeral (*three*). Since the latter is taken to be adjectival, the MCC constitutes an adjective phrase. The high numeral (*hundred*) is taken to be a nominal complement of the adjectival numeral *three*. This gives us the structure in (41):¹⁴

- (41) [[AP [A three] [NP hundred]] cars]

Zabbal argues that the adjectival nature of the MCC *three hundred* is supported by the fact that it displays exactly the same morphosyntactic behavior as the simplex adjectival element *three*. For example, both *three* and *three hundred* cannot inflect for the plural, as exemplified in (42a,b).¹⁵ In this respect they differ from the numeral *hundred*:¹⁶

- (42) a. *John bought threes (of) eggs.
b. *John bought three hundreds (of) eggs.
c. John bought hundreds of eggs.

It is further assumed that the adjectival head *three* in (41) assigns structural (i.e. genitival) case to the nominal complement *hundred*. Importantly, genitival case does not surface overtly as adpositional *of* in English. Thus, *three* does not behave like *proud*, which typically takes an *of*-complement (*very proud of Mary*), but more like *near*, as in *very near the shore*, where

Consequently, it cannot be input to an ellipsis (Spanish) or pronominalization (Dutch) operation.

¹⁴ In (41), we abstract away from the syntactic position that the adjectival MCC occupies within the extended nominal projection. Zabbal proposes that the MCC occupies the specifier position of a functional head NuM within the extended nominal projection. This structure, slightly adapted on the basis of Zabbal’s bare phrase structure analysis, is given in (i)

(i) [DP D [NumP [MCC] [Num Num [NP N]]]]

¹⁵ See our earlier discussion of examples (11) and (12).

¹⁶ Note that this analysis raises a question as regards the Head Final Filter (Williams 1982); see also Emonds’s (1976, 1985) (Right) Recursion Constraint. According to HFF, an attributive adjective cannot take a complement to its right when it is in prenominal position: *a proud (*of Mary) man*.

near is followed by the noun phrase *the shore* (Maling 1983, Kayne 2005a). As Zabbal points out, the case assignment pattern ‘Adjective + noun phrase_[+genitive]’ is familiar from adjectival construct state constructions in Semitic languages (Hazout 2000, Siloni 2002). An example, drawn from Thackston (2000, 188), is given in (43):

- (43) ?ar- rajul-u l- ḥasan-u l-wajh-i
 the man-NOM the handsome-NOM the face-GEN
 ‘the man handsome of face’

In (43), the adjective *ħasan* ‘handsome’ immediately precedes the noun phrase *l-wajh-i*, which carries genitive case. Note that *ħasan* agrees with *?ar-rajul-u* in definiteness, gender and case.

Tatsumi (2021) presents a third implementation of the constituent analysis of MCCs. Importantly, he assumes that MCCs do not receive a uniform structural analysis cross-linguistically. Specifically, there are languages (e.g. English) in which noun phrases containing an MCC have the structure in (32), repeated here as (44a), but there are also languages in which such noun phrases have the structure in (44b), where NUMBER is a silent constituent.

- (44) a. [FP [YP three] [F' hundred [NP cars]]]
 b. [[FP threeYP [F' hundred [NP NUMBER]]] [NP cars]]]

Note that structure (44b) is somewhat different from Zweig's (2005)—see (40a)—though structurally related to it. The representations in (44b) and (40a) share the syntactic placement of the multiplicand and the multiplier. The former is the head of the phrase *XP*, while the latter functions as a phrasal satellite within *XP*. The two structures differ from each other, however, as regards the selectional behavior of the multiplicand *hundred*: In Zweig's representation (40b), the multiplicand *hundred* does not select a complement; in Tatsumi's representation (44b), on the contrary, the multiplicand *hundred* selects the silent constituent *NUMBER*, which is interpreted as the property of being a number. Thus, it is the string *three+hundred+NUMBER* which forms a phrasal constituent. As indicated in (44b), this constituent is adjoined to *NP*.

Tatsumi legitimates the distinction between the non-constituent analysis of MCCs (i.e. (44a)) and the constituent analysis of MCCs (i.e. (44b)) by considering computational operations that take constituents as their input. Specifically, he points out that Slavic languages such as Serbo-Croatian permit left branch extraction of the entire MCC, as exemplified in (45). The possibility of fronting the entire MCC suggests that it is a syntactic unit. In Tatsumi's analysis, this unit corresponds to the FP consisting of *three+hundred+NUMBER*.

- (45) a. Ivan je pozvao [tri stotine studenata].
Ivan is invited three hundred.ACC.F students.GEN.M
'Ivan invited three hundred students.'

b. [Tri stotine]_k je Ivan pozvao [t_k studenata]
three hundred.ACC.F is Ivan invited students.GEN.M
'Three hundred students, Ivan invited.' [Aida Talić (p.c.)]

In a language like English, subextraction of an MCC like *three hundred* is impossible (see (46)), because the string *three hundred* in (44a) does not form a syntactic unit; it only forms a unit in combination with the counted noun *cars*.

- (46) *Three hundred Ivan invited [three hundred students].

Notice that, in Zweig's analysis and Zabbal's, English MCCs such as *three hundred* do form a single constituent. For those analyses, the contrast between the well-formed Serbo-Croatian pattern in (45b) and the ill-formed English one in (46) must receive a different explanation, possibly in terms of the 'NP (= Serbo-Croatian) versus DP (= English)' parameter, discussed in Corver (1990; 1992) and Bošković (2005). Notice at this point that Zabbal (2005) tries to give evidence in support of the constituency of English MCCs on the basis of anaphoric dependencies. Starting from the assumption that only constituents can function as antecedents in anaphoric dependencies, he points out that an English MCC like *three hundred* can be the antecedent for anaphoric phrase-like (adjectival) expressions such as *that many* and *that few*:¹⁷

- (47) a. John bought *three hundred_i* marbles and Mary bought *that many_i* chocolates.
 b. *Two hundred_i* people came to Pollini's concert. I was surprised that only *that few_i* people wanted to see the famous pianist.

In summary: this section discussed various types of structural analyses of MCCs. Essentially, three major types of analyses can be distinguished: (i) MCCs are complex words, more specifically compounds, (ii) MCCs represent syntactic components of the noun phrase but do not constitute a single syntactic object, (iii) MCCs consist of syntactic components that form a single syntactic object (i.e. a constituent).

4. Additive complex cardinals

The numeral *twenty six* in *twenty six cars* is a composite numeral expression. It consists of two meaningful components —*twenty* and *six*— which together express the arithmetical relation of addition: $20 + 6$. Since *twenty six* expresses addition, it is called an 'Additive Complex Cardinal' (henceforth: ACC). The numeral *twenty* is the augend —that is, the number to which successive numbers are added (e.g. *twenty one*, *twenty two*, *et cetera*)— the numeral *six* the addend. In English numerals such as *sixteen* ($6 + 10$) and *nineteen* ($9 + 10$), the bound form *-teen* is the augend, while *six* and *nine* function as addends. Notice that the ACC *sixteen* has the reverse word order of *twenty six*: in the former, the addend precedes the augend, while, in the latter, it is the augend that precedes the addend.¹⁸

In three-member ACCs such as *three hundred and twenty five* (i.e. $[300 + [20 + 5]]$), *twenty* is the augend for the addend *five*. The MCC *three hundred* functions as the augend for the (complex) addend *twenty five*, because *three hundred* is a number to which successive MCCs are added (e.g. *three hundred twenty nine*, *three hundred thirty*, *et cetera*). Notice that the augend *three hundred* and the addend *twenty five* are connected by means of the conjunction-like element *and*, which is sometimes pronounced as a reduced form (i.e. /ən/) or even left out (*three hundred twenty five*).¹⁹

¹⁷ (47a) is taken from Zabbal.

¹⁸ In a language like Dutch, the MCCs *zestien* 'sixteen' and *zesentwintig* 'twenty six' display the same internal word order: the addend precedes the augend. Notice, however, that in *zesentwintig* the addend *zes* is separated from the augend *twintig* by the conjunction-like element *en* 'and', which is often pronounced as a reduced form: /ən/. Notice that the linking morpheme *en* must be present: **zesentwintig*. Interestingly, while English *and* can follow an augend like *hundred*, as in *a hundred and six*, it cannot follow the augends *twenty* to *ninety*: *twenty (*and) six*.

¹⁹ Dutch displays the same phenomenon, the number 106 can be realized as *honderd en zes*, which features the connecting element *en*, which can have a reduced form (/ən/). A sequence without *en* is also possible: *honderd zes*.

As shown in (48a-d), ACCs occur in typologically different languages:

- (48) a. Ivan je video [dvadeset pet] devojaka. [Serbo-Croatian]
Ivan is seen twenty five girls
'Ivan saw twenty five girls.'
- b. ?arba?at-u ?aalaaf-in wa- xams-u- mi?at-in rajul-in [Arabic; Zabbal 2015 (87a)]
four-FS-NOM thousands-GEN and five-NOM hundred-GEN man-GEN
'four thousand five hundred men'
- c. Taro-ga [gakusei ni-zyuu go]-nin-o mita. [Japanese]
Taro-NOM student two-ten five-CLS-ACC saw
'Taro saw twenty five students.'

Given the composite form of ACCs, the question arises as to what their inner organization is like. First of all, the question needs to be addressed as to whether ACCs are complex *words* (i.e. morphological compounds) having an inner structure or whether they are (syntactic) *phrases* with an inner structure. In view of their coordinate-like meaning (i.e., *twenty five = twenty + five*), one may propose that ACCs are coordinative compounds (also called 'copulative compounds'); see Booij and Van der Wouden (2020, May 14). This type of compound consists of two (or more) elements of the same syntactic category, which bear equal semantic weight (Olsen 2001). Some English examples of coordinate compounds are the following: *child star* (someone who is both a child and a star), *player-coach* (someone who is both a player and a coach), and *gardener-chauffeur* (someone who is both a gardener and a chauffeur). Even though the two coordinated elements are semantically equal, it is clear that the final noun—for example, *star* in *child star*—functions as the head; it is namely the final noun (i.e. the right-hand head) that carries plural inflection: *child stars*, **children star*, **children stars*. So, if ACCs are treated as coordinate compounds, then the second numeral (i.e. the addend) must be the head of the complex numeral: *[[x twenty] [x five]]*. One would expect then that this right-hand numeral determines morphosyntactic interactions (e.g. case/gender agreement) with material external to the numeral expression. This pattern is actually observed in Slavic languages. In (49), the right-hand numeral exhibits gender agreement (feminine) with the main noun.

- (49) Ewa zobaczyła dwadzieścia dwie dziewczyny. [Polish]
Ewa saw twenty.NV.ACC two.ACC.F girl.ACC.PL.F
'Ewa saw twenty-two girls.'

The coordinative compound analysis of ACCs possibly extends to ACCs featuring the bound morpheme *-teen* as augend. The fact that, in certain patterns, the addend has an irregular form hints at a compound analysis: *thirteen* (**three-teen*), *fifteen* (**fiveteen*). It has further been suggested that coordination-like patterns such as *a hundred and ten*, which feature the connecting element *and*, are not syntactic phrases but actually coordinative compounds. Booij (2002: 166), for example, makes that claim for Dutch ACCs like *honderd en tien* (i.e., 110), which he characterizes as "a morphological construction that originated through the morphologization of a syntactic construction". According to Booij, two phenomena hint at this morphologization: firstly, the phonetic weakening of the Dutch coordinator *en* 'and'; that is, instead of being pronounced as /en/, it is pronounced as /ən/. Secondly, the main stress of an ACC like *honderd en tien* is located on the last constituent (i.e. *honderd en TIEN*), whereas the parts of a syntactic coordination are equally stressed. Both claims are disputed; see Van der Wouden (2020) for the weakening of *en*, and Corver (2014) for discussion of the stress pattern of syntactic coordinations.

Further support for the compound analysis of ACCs may come from patterns in which the ACC combines with a derivational suffix, this way creating a complex word. Consider, for example, the Dutch examples in (50), where the nominal suffix *-er* is attached to a numeral creating the meaning ‘someone of age X’. Most common bases are multiples of 10; see (50a). Formations based on ACCs (see (50b)) and MCCs (see (50c)) are also possible, however, and with the right context they will immediately receive the correct interpretation (see Van der Wouden (2020)).²⁰

- (50) a. tien-er (ten-er, ‘teenager’), twintig-er, dertig-er, veertig-er, vijftig-er *et cetera*
 (a person in his twenties/thirties/forties/fifties,...)
 b. zestien-er (sixteen-er, ‘someone of sixteen years old’), achttien-er (eighteen,
 ‘someone of eighteen years old’)
 c. vijf-en-twintig-er (five-and-twenty-er, ‘someone of twenty five years old’),
 negen-en-dertig-er²¹ (nine-and-thirty-er, ‘someone of thirty nine years old’)

Besides morphological approaches that take ACCs to be complex words (i.e., coordinative compounds), there have recently been a number of proposals according to which ACCs do not constitute complex words but rather complex syntactic constructs. One such proposal is Ionin and Matushansky’s (2018), for example, who claim that ACCs are coordinate structures whose conjuncts consists of NPs. A noun phrase like *twenty five cars* has the base structure in (51a). As indicated, both the augend (*twenty*) and the addend (*five*) form a conjunct together with the counted noun (*cars*). The surface pattern *twenty five cars* in (51c) is obtained by application of NP-deletion to the first conjunct, as in (51b):²²

- (51) a. [[[NP twenty [NP cars]] & [NP five [NP cars]]]] (NP coordination)
 b. [[[NP twenty [NP ears]] & [NP five [NP cars]]]] (deletion)
 c. twenty five cars

The English ACC *twenty five* does not feature the overt coordinator *and*: *twenty (*and) five cars*. In other words, it is an asyndetic coordination. In other languages, for example Dutch, the coordinator must be overt: *vijf *(en) twintig auto’s* (five (and) twenty cars, ‘twenty five cars’). It should be noted that, in Dutch as opposed to English, the addend (*vijf*) precedes the augend (*twintig*).

- (52) a. [[[NP vijf [NP auto’s]] en [NP twintig [NP auto’s]]]] (NP coordination)
 b. [[[NP vijf [NP auto’s]] en [NP twintig [NP auto’s]]]] (deletion)
 c. vijf en twintig auto’s

In noun phrases featuring a three-member ACC, such as *three hundred and twenty five cars*, we have a (matrix) coordinate structure whose right conjunct is also a coordinate structure, as

²⁰ As has already been noted in note 8, alternative, more syntax-based, approaches are imaginable; see Kayne (2005b); see also section 7.

²¹ Words such as *negen-en-dertig-er* do not necessarily refer to a person of a particular age. Such expressions can also refer to the year in which someone/something was born/built (e.g. *een negenendertiger* (litt.: a nine-and-thirty-er) can be someone born in the year 1939, or a car that was built in that year) or to a number-based characteristic of an entity (e.g. *een vijf-en-vijftig-er* (litt.: a five-and-fifty-er) can be a battleship with 55 canons); see Van der Wouden (2020, May 14).

²² In earlier analyses, Ionin and Matushansky argued that the gap in the left conjunct results from Right node raising. Schematically:

(i) [[NP twenty {NP cars}] & [NP five {NP cars}] cars]

in (53). The surface pattern is derived by applying NP-deletion in the two most deeply embedded conjuncts, as in (53b).

- (53) a. [[[three hundred [NP cars]] and [[[NP twenty cars] &_{coordination} [NP five cars]]]]]
- b. [[[three hundred [NP ears]] and [[[NP twenty ears] &_{coordination} [NP five cars]]]]]
- c. three hundred and twenty five cars

The strength of Ionin and Matushansky's analysis of ACCs lies in its ability to explain morphological case patterns that are dependent on the right-most numeral in a numeral expression. Consider, for example, the following pattern from Russian (drawn from Ionin and Matushansky 2018):

- (54) a. dva šagA
two-NOM step-PAUC
'two steps'
- b. dvadcat' šagov
twenty-NOM step-GEN.PL
'twenty steps'
- c. dvadcat' dva šagA
twenty-NOM two-NOM step-PAUC
'twenty-two steps'

As shown by (54a), the paucal cardinal *dva* 'two' agrees with the lexical NP it combines with. The numeral *dvadcat'* 'twenty', on the contrary, must be followed by a lexical NP carrying genitive case; see (54b). Notice now that the ACC *dvadcat' dva* 'twenty two' in (54c) must be followed by an NP carrying paucal morphology. This is exactly what one expects if the final numeral takes an overt NP as its complement.

Evidence for Ionin and Matushansky's proposal that both the augend (e.g. *twenty*) and the addend (e.g. *five*) of an ACC combine with a noun (e.g. *cars* in (51)) comes from languages in which the counted noun optionally surfaces in more than one position. Based on English, we would have the (non-existent) form: *twenty cars five cars*. This "reduplicative" pattern is attested in typologically different languages such as Biblical Hebrew (Semitic; (55a)), Luvale (Bantu; (55b)) and Toqabaqita (Austronesian; (55c)); see Zweig (2005), Tatsumi (2021), Lichtenkerk 2008:296)

- (55) a. meah *shana* ve-esrim *shana* ve-sheva *shanim* (Genesis 23:1)
hundred year and-twenty year and-seven years
'one hundred and twenty seven years'
- b. *mikoko* makumi atanu na-*mikoko* vatanu
sheep ten five and-sheep five
'fifty five sheep'
- c. teqe toqoni *imola* ma roo talange-qe *imole* ma lima akwale-qe *imole*
one thousand person and two hundred-ASC person and five tensome-ASC person
'one thousand two hundred fifty people'

Another potential argument in support of the analysis depicted in (51) comes from patterns of NP-deletion. If both the augend and the addend combine with a noun in the base structure, one would expect there to be languages in which deletion does not apply to the noun that combines with the augend but rather to the noun that combines with the addend. Based on English, such a pattern would look as follows: *twenty cars and five* (meaning: 'twenty five

cars'). As noted in Ionin and Matushansky (2018), there are indeed languages that display this deletion pattern, for example Celtic languages (see Hurford (1975) for Biblical Welsh, Paterson (1989) and Hurford (2003) for Scottish Gaelic, and Press (1986) for Breton).²³

- (56) a. tri *fear* dheug 's da fhichead (Scottish Gaelic)
 three man ten plus two twenty
 ‘fifty three men’
- b. dek *den* ha kant (Breton)
 ten person and hundred
 ‘hundred and ten persons’

In (56a), the addend *tri* is followed by the counted noun *fear*, while the complex augend

²³ Modern Welsh has the adposition *ar* as a connecting elements in MCCs (cf. Ionin and Matushansky 2018). Tatsumi (2021) reports that the following patterns featuring *ar* all have the same meaning: ‘their three countries’. Notice that the counted noun (*gwlad*) can surface in different positions:

- (i) a. [tair *gwlad*] *ar* ddeg *ar* hugain. (Modern Welsh, David Willis p.c.)
 three country on ten on twenty
 ‘thirty three countries’
- b. tair *ar* [ddeg *gwlad*] *ar* hugain
 three on ten country on twenty
- c. tair *ar* ddeg *ar* [hugain *gwlad*]
 three on ten on twenty country

Modern Welsh also gives support for the NP-coordination analysis. The suppletive ordinal *cyntaf* 'first' can optionally undergo soft mutation (*cyntaf* > *gyntaf*) when the main noun is singular feminine as in (iia,b) (cf. Borsley *et al* 2007). What is important is that ACCs in Welsh can trigger soft mutation regardless of the position of the main noun. The data in (iia,b) show that the soft mutation process is not sensitive to the linear adjacency. On the other hand, soft mutation is blocked when the main noun is plural as in (iic) even though the main noun is linearly adjacent to the ordinal in (iic). These data indicate that soft mutation in Modern Welsh depends on the number feature on the main noun.

- (ii) a. y tair wythnos ar ddeg {^{OK}gyntaf / ^{OK}cyntaf} (Modern Welsh, David Willis p.c.)
 the three week.SG.F on ten first first
 ‘the first thirteen weeks’
- b. y tair ar ddeg wythnos {^{OK}gyntaf / ^{OK}cyntaf}
 the three on ten week.SG.F first first
 ‘the first thirteen weeks’
- c. y tair ar ddeg o wythnosau {^{*}gyntaf / ^{OK}cyntaf}
 the three on ten week.PL.F first first
 ‘the first thirteen weeks’

The NP-coordination analysis naturally expects the similarity between (iia,b) above regarding the availability of soft mutation. Regardless of the surface linear adjacency, the ACC contains the main noun in each conjunct, as shown in (iii). There should be no contrast between them.

- (iii) a. y [tair wythnos] ar [ddeg *wythnos*] {gyntaf | cyntaf} (= (a))
 b. y [tair *wythnos*] ar [ddeg wythnos] {gyntaf | cyntaf} (= (b))

Romanian is another language that features an adposition (*de*) as a linking element in certain complex numerals. Consider, for example, the following MCC-pattern, where *sută* is the multiplier and *mii* is the multiplicand (see Corver (2001; 2012)).

- (iii) o sută *de* mii *de* oameni
 a hundred of thousand of people
 ‘a hundred thousand people’

dheug 's da *fhichead* ('fifty') is not followed by any overt noun. In (56b), we have the same pattern, with *dek* being the addend and *kant* being the augend. In both examples, NP-deletion applies to the second conjunct, as exemplified in (57) for the Breton pattern in (56b):

- (57) a. [[[NP *dek* [NP *den*]] ha [NP *kant* [NP *den*]]]] (NP coordination)
 b. [[[NP *dek* [NP *den*]] ha [NP *kant* [NP ~~den~~]]]] (deletion)

The patterns in (51) and (57) obviously raise the question as to what kind of deletion operation is involved in their derivation. In Ionin and Matushansky (2006), it is suggested that the pattern in (51), which involves deletion in the first conjunct, could be analyzed in terms of Right Node Raising. Schematically:

- (58) [[[NP twenty *ears*] & [NP five *ears*]] *cars*]

Ionin and Matushansky (2018), however, reject such an analysis on the basis of the deletion pattern in the Breton example (57b), where the non-deleted noun (*den*) clearly is *in situ* and has not undergone any displacement. As an alternative analysis, one may propose that deletion of the counted noun results from the application of NP-ellipsis, which is familiar from examples such as (59a). Under an NP-ellipsis analysis, the noun following the augend (*twenty*) would be deleted, its content being recoverable on the basis of the noun (*cars*) that follows the addend (*five*); see (59b).

- (59) a. John bought [five cars] and Mary bought [twenty *ears*].
 b. [[[NP twenty *ears*] & [NP five cars]]]

It should be noted, however, that there are languages which do not allow NP-ellipsis with numeral remnants but nevertheless do have nominal expressions featuring an ACC. Dutch, for example, does not allow NP-ellipsis, as shown in (60a). It only permits the pronominalization strategy featuring the quantitative pro-form *er* (litt.: there, 'of them'), which has been moved out of the noun phrase; see (60b).²⁴ As shown in (61), however, Dutch does permit deletion of the counted noun in ACCs, if one adopts Ionin and Matushansky's approach to ACCs.

- (60) John heeft [vijf auto's] gekocht
 John has five cars bought....
 a. *...en Marie heeft [twintig *auto*'s] gekocht.
 ...and Marie has twenty cars bought

²⁴ The extraction pattern in (60b) raises an interesting question for Ionin and Matushansky's approach to ACCs. The pro-form *er* can also be extracted out of a noun phrase featuring an ACC, as exemplified in (i):

- (i) Marie heeft *er* gisteren [vijf_{ADDEND} en twintig_{AUGEND} --] gekocht.
 Marie has *er* yesterday twenty and five bought
 'Marie bought twenty five (of them).'

In Ionin and Matushansky's approach, both the addend (*vijf*) and the augend (*twintig*) are followed by a noun in the underlying representation, where the noun following the addend is deleted. For the extraction pattern in (i), the question then arises as to why we don't get a violation of Ross's (1967) Coordinate Structure Constraint, since *er* is only extracted from the right conjunct, as depicted in (ia). A possible way out would be to say that in extraction patterns like (i), extraction takes place in an Across-The-Board fashion, as in (iib).

- (ii) a. ...*er*....[[vijf *ec*] en [twintig *er*]] (extraction from right conjunct)
 b. ...*er*....[[vijf *er*] en [twintig *er*]] (ATB-extraction)

- b. ...en Marie heeft *er* [twintig ~~er~~] gekocht.
..and Marie has of-them twenty bought

- (61) Jan heeft [vijf ~~boeken~~ en twintig boeken] gekocht.
Jan has five books and twenty books bought
‘Jan bought twenty five books.’

If neither right node raising nor NP-ellipsis is the strategy used for creating the silent noun right after the augend, then some other mechanism must be responsible for creating the gap, possibly backward conjunction reduction.

In line with their analysis of MCCs, Ionin and Matushansky (2018) assume that nominal expressions featuring ACCs have an inner structure in which the ACC does not constitute a single constituent. Consider again the coordinate structure of *twenty five cars* in (51b), which is repeated here as (62):

- (62) [[[NP twenty [NP ~~cars~~]] & [NP five [NP cars]]]]

In this representation, the sequence consisting of the augend *twenty* and the addend *five* does not form a constituent. It is the counted noun *cars* that the addend *five* forms a unit with. Adopting the representation in (62) leads to the prediction that ACCs such as *twenty five* or *hundred and five* cannot be the target of computational operations such as displacement. It turns out, however, that in certain languages ACCs behave like syntactic units.²⁵ Tatsumi (2021), for example, notes that, in a language like Serbo-Croatian, ACCs can undergo left branch extraction out of a noun phrase:

- (63) a. Ivan je video [[dvadeset i pet] studenata].
Ivan is seen twenty and five students.GEN.PL.M
‘Ivan saw twenty five students.’
b. [Dvadeset i pet] je Ivan video [~~dvadeset i pet~~ studenata]

Given this left branch extraction of the entire ACC, Tatsumi concludes that Serbo-Croatian noun phrases featuring ACCs have an inner organization which is different from, for example, English noun phrases featuring ACCs. Specifically, he proposes the following structure:²⁶

- (64) [NP [&P [dvadaset NUMBER] [& i [pet]]] [NP studenata]]

According to this analysis, ACC is a coordinate structure. It differs, however, from the English coordinate structure in (62) in a number of ways. Firstly, the augend *dvadeset* ‘twenty’ does not combine with the counted noun ‘students’ but rather with a silent noun NUMBER. Secondly, the addend *pet* ‘five’ is a “bare” numeral functioning as the complement of the coordinator *i* ‘and’. As indicated in (64), the entire ACC is adjoined as a left branch constituent to the NP *studenata*. Being a single constituent, ACC can be input to left branch extraction.

There are other phenomena that seem to suggest that ACCs can function as single constituents that combine with a single counted noun; that is, *[[twenty five] cars]* instead of *[[twenty ~~cars~~] & [five cars]]*. These phenomena typically involve the interaction between the addend (e.g. *five* in *twenty five*) and the counted noun (*cars*). For instance, Kayne (2019) points out the contrast in (65).

²⁵ See also Zweig (2005) and Zabbal (2005) for a phrasal constituent analysis of ACCs.

²⁶ In section 3, a similar structure was proposed for MCCs in Slavic languages such as Serbo-Croatian.

- (65) a. We've been there twenty-*{one | two}* times.
 b. *We've been there *{twenty-once | twenty-twice}*.

English has the suppletive form of adverbial cardinals; *once* and *twice*. Adverbial cardinals of numerals over *two* are formed by using *time* (e.g. *three times*, *four times*, *et cetera*). Note that the suppletive adverbial form *thrice* has become archaic). However, the suppletive form is disallowed in ACCs, as exemplified in (65b). In this case, the adverbial numerals are expressed by *time* with a numeral. If the relevant suppletion process of adverbial numerals is related to the constituency of a given ACC, then the unacceptability of (65b) can be taken as support for the single-constituent analysis (*[[twenty two] times]*), instead of the coordination structure (*[[twenty times] & [two times]]*). The coordination structure needs to explain why suppletion is disallowed in the second conjunct.

Support for the single-constituent analysis of ACCs also comes from Polish. In Polish, the cardinal “one” is normally adjectival, and shows agreement with a noun in number, case and gender, as shown in (66a). However, when the cardinal “one” appears in an ACC, the agreement is blocked, as shown in (66b).

- (66) a. Jan zobaczył *{*jeden | ^{OK}jedną}* dziewczynę.
 Jan saw one.M one.ACC.F girl.ACC.SG.F
 ‘Jan saw one girl.’
 b. Jan zobaczył dwadzieścia *{^{OK}jeden | *jedną}* dziewczynę.
 Jan saw twenty.NV.ACC one.M one.ACC.F girl.GEN.PL
 ‘Jan saw one girl.’

Again, the contrast in (66) may be accounted for if the relevant agreement between the right-hand cardinal and the main noun depends on a certain type of structural relationship. In (66b), the agreement relation is blocked because the right-hand cardinal is trapped within in the ACC. If Polish ACCs have the coordination structure, the contrast in (66a,b) is not expected.²⁷

The opacity in ACCs is also observed in Dutch. As shown in (67b), when the counted noun has been replaced by so-called quantitative *er* (meaning ‘of them’), attachment of the diminutive morpheme after *een* is possible (the numeral *een* ‘one’ is actually the only numeral that permits attachment of the diminutive morpheme). (67c) shows that quantitative *er* is also possible in combination with the numeral *honderd*. It is, however, impossible to use the diminutive morpheme *-tje* when it follows the ACC *honderd en een* (hundred and one), as shown in (67d).

- (67) a. Ik heb gisteren [een(*-tje) leeuw] gezien.

²⁷ Ionin and Matushansky (2018:133) hypothesize that the contrast in (66) arises for the following reason: when the numeral ‘one’ appears in an ACC, it inherits the nominal features of the higher cardinal. The question then arises as to when the relevant inheritance process happens. When numerals other than ‘one’ appear in ACCs, they keep agreeing with the main noun as in (ii). Moreover, in other Slavic languages the opacity of the numeral ‘one’ in ACCs is not present. (ii) is an example from Serbo-Croatian. In (ii), the right-hand numeral ‘one’ agrees with the main noun.

- (i) Jan zobaczył [dwadzieścia dwie] dziewcząt. (Polish: Paulina Lyskawa, p.c.)
 Jan saw twenty.NV.ACC two girl.ACC.F girl.GEN.PL.F
 ‘Jan saw twenty two girls.’
 (ii) Peter je video [dvadest jednu] djevojku. (Serbo-Croatian: Ivana Jovović)
 Peter is seen twenty one.F.ACC girl.F.ACC.SG
 ‘Peter saw twenty one girls.’

- I have yesterday one(-DIM) lion seen
 ‘I saw one lion yesterday.’
- b. Ik heb er gisteren een(-tje) gezien
 I have there yesterday one(-DIM) seen
 ‘I saw one (e.g. lion) yesterday.’
- c. Ik heb er gisteren honderd gezien.
 I have there yesterday hundred seen.
 ‘I saw hundred (e.g. lions) yesterday.’
- d. Ik heb er gisteren honderd en een(*-tje) gezien.
 I have there yesterday hundred and one(DIM) seen
 ‘I saw hundred and one lions yesterday.’

Under Ionin and Matushansky’s analysis, the ill-formedness of (67d) seems unexpected. If *een* takes its own NP-complement, then it is unclear why it cannot be followed by the diminutive *-tje*. Suppose that the noun in (67b) has the structure given in (68a) below. Here we assume that the diminutive *-tje* is a functional head taking an NP-complement (see Borer (2005), Wiltschko (2007), De Belder (2011), Ott (2011)). In (68a), the NP-complement has been substituted for by quantitative *er*. Under Ionin and Matushansky’s analysis, it seems unclear why *eenthje* is not allowed in (67d), because it is analyzed as having the structure in (68b). Here, the numeral *een* appears in the same configuration as in (67b) above, and the contrast between (67b,d) is not expected.

- (68) a. ...er_i[FP een [F' -tje [NP er]]] (= (67b))
 b. ...er_i ... [ConjP [NP honderd t_i] [Conj' en [NP een -tje t_i]]]
 (= (67d) under I&M's analysis)

If the ACC *honderd en een* forms a constituent as in (69) below, it may be expected that the diminutive cannot be available due to the fact that the numeral *een*, which is the only numeral that can license the diminutive *-tje*, is too deeply embedded in the specifier position of the phrase headed by the diminutive.

- (69) *...er_i[FP [honderd en een] [F' -tje [NP er]]]... (= (67d))

In summary: this section discussed various types of structural analyses of ACCs. Just as with MCCs, three major types of analyses can be distinguished: (i) ACCs are complex words, more specifically coordinative compounds, (ii) MCCs represent syntactic components of the noun phrase but do not constitute a single syntactic object, (iii) MCCs consist of syntactic components that form a single syntactic object (i.e. a constituent).

5. Numerals in classifier languages

So far, the discussion has focused on numerals in non-classifier languages in which numerals directly follow or precede the main noun. This section shifts the focus to the grammatical behavior of numerals in classifier languages in light of the discussion so far.

5.1 Simple numerals in classifier languages: their categorial nature and distribution

Unlike some non-classifier languages, numerals in classifier languages tend to not show the N(oun)/A(djective) distinction. This may be partly because of the presence of classifiers. As

shown in (70), in some classifier languages, numerals are separated from the main noun by an intervening classifier.

- (70) a. ba cuốn sách [Vietnamese (Num-CLS-N): Nguyen (2004: 10)]
 three CLS book
 ‘three books’
- b. weubam bo at [Abun (N-CLS-Num): Berry&Berry 1999: 93]
 pineapple CLS four
 ‘four pineapples’

Greenberg (1972) reports that only four combinations of noun (N), quantifier (#), and classifier (CLS) are attested in his sample. The two unattested combinations in Greenberg’s sample are #-N-CLS and CLS-N-#. His observation is summarized in (71). Here, several languages from our sample have been added; see also Jones (1970), Aikhenvald (2003).²⁸

(71)

Num-CLS-N	frequent	Bangla [Indic], Chinese [Sino-Tibetan], Dolakha Newar [Sino-Tibetan], Vietnamese [Austro-Asiatic], Hmong [Hmong-Mien], Indonesian [Austronesian], Uzbek [Turkic], Hungarian [Uralic], Assamese [Indic], Toqabaqita [Austronesian]
N-Num-CLS	frequent	Burmese [Sino-Tibetan], Japanese [Japanese], Khmer [Austro-Asiatic], Lahu [Sino-Tibetan], Mal [Austro-Asiatic], Mokilese [Austronesian], Thai [Tai-Kadai], Loven/Jruq [Mon-Khmer]
CLS-Num-N	rare	Ibibio [Niger-Congo]
N-CLS-Num	rare	Abun [Papuan], Bodo [Sino-Tibetan], Taba [Austronesian], Tetun [Austronesian], Garo [Sino-Tibetan]
Num-N-CLS	*	unattested
CLS-N-Num	*	unattested

In most of the classifier languages above, adjectival modifiers do not require the presence of a classifier to modify the main noun. This indicates that numerals in these classifier languages are not adjectival syntactically. (It should be noted, however, that, in some classifier languages such as Vietnamese and Thai, classifiers appear in relative clause constructions.)

5.2 Multiplicative complex cardinals in classifier languages

²⁸ Bangla allows for post-nominal numeral classifier constructions (i.e. N-#-CLS). However, it seems that post-nominal numeral classifiers are derived by NP-movement (Bhattacharya 1999). Japanese also has pre-nominal and post-nominal numeral classifiers. Japanese numeral classifier phrases are classified as N-#-CLS because, in pre-nominal position, a numeral classifier phrase must be followed by the genitive linker *no* (e.g. *hon san-satsu*, ‘book three-CLS, ‘three books’ versus *san-satsu-no hon*, three-CLS-GEN book, ‘three books’).

Like non-classifier languages, multiplicative complex cardinals (MCCs) are consisting of a multiplier and a multiplicand. An example of a MCC in Japanese is given in (72).

- (72) Taro-ga [gakusei [**go** **hyaku**]-nin]-o syootai-sita. [Japanese (N-Num-Cl)]
 Taro-NOM five hundred-CLS GEN student-ACC invitation-did
 'Taro invited [five hundred students].'

In (72), the numeral *go* 'five' functions as a multiplier and the numeral *hyaku* 'hundred' as a multiplicand. In Japanese, MCCs generally require an appropriate numeral classifier when they are used as a modifier of nouns.

However, there are some classifier languages that use multiplicands as classifiers. Let us first consider multiplicative complex cardinals in Mokilese [Austronesian]. Mokilese is a classifier language where noncomplex cardinals (1 through 9) are followed by a classifier, as in (73) (Harrison 1976: 99).

- (73) a. amwje pah-men b. jiloa pah-w [Mokilese]
 mosquito four-CLS clam four-CLS
 'four mosquitos' 'four clams' [Harrison (1976: 95-96)]

In contrast to non-complex cardinals, multiplicative complex cardinals can modify a noun without a numeral classifier, as shown in (74).

- (74) Mokilese: Harrions (1976: 99)
 a. puk ei-jek b. suhkoa pah-pwki
 book one-ten tree four-hundred
 'ten books' 'four hundred trees'

In Mokilese, numeral classifiers are not needed in MCCs. A similar pattern is also observed in Toqabaqita [Austronesian]. As shown in (75a), the noun (*q*)*alo* 'taro' occurs with the classifier *fa* when it is modified by a noncomplex numeral. However, the classifier is not used in the presence of multiplicands such as "hundred" and "thousand", as in (75b).

- (75) Toqabaqita: Lichtenberk (2008: 293)
 a. ulu fa qalo b. ulu toqoni alo
 three CLS taro three thousand taro
 'three taro corms' 'three thousand taros (corms or whole plants)'

Dolakha Newar is another classifier language which uses multiplicands in the position of numeral classifiers. Most numerals in Dolakha Newar must occur with a classifier to modify a noun (Genetti (2007: 217)). Multiplicative complex cardinals, however, do not take numeral classifiers, as shown in (76).

- (76) Dolakha Newar: Genetti (2007: 218)
 a. nis-sar mi b. *nis-sar mā mi
 two-hundred people two-hundred CLS person
 'two hundred people' 'two hundred people'

A similar pattern is also observed in Nung [Tai-Kadai]. In Nung, the presence of a classifier is obligatory when classifiable nouns are modified by a noncomplex cardinal, as in (77a). However, a classifier becomes optional when a multiplicand is present, as shown in (77b).

- (77) a. lēo jgà [slóng tú má luhc]. b. mi [slám pác (áhn) hon].
 then kill two CLS dog child have three hundred CLS house
 'Then kill two puppies' 'There are three hundred houses.'

The typological data discussed above lead us to the conclusion that there are several types of classifier languages regarding the co-occurrence of numeral classifiers and multiplicands, as summarized in (78).

- (78) a. Type 1: Languages in which multiplicands and classifiers always co-occur
 b. Type 2: Languages in which multiplicands and classifiers cannot co-occur
 c. Type 3: Languages in which multiplicands and classifiers can co-occur

As noted in section 3, Ionin and Matushansky (2018) propose the cascading structure for MCCs in non-classifier languages. The structure is represented in (79).

- (79) [three [hundred [_{NP} cars]]]

To accommodate the crosslinguistic pattern in (78), Ionin and Matushansky's analysis would have to assume that in the Type 2 classifier languages, classifiers and multiplicands appear in the same position, yielding the structure in (80).

- (80) [three [{hundred | CLS} [_{NP} cars]]]

According to this structure, the classifier head and the multiplicand occur in the same position and compete with each other, yielding complementary distribution.

In this context, it is worth noting that the numeral classifier sequence and MCCs in Japanese behave alike regarding sequential voicing (a.k.a *Rendaku*). As shown in (81), some Sino-Japanese classifiers can exhibit sequential voicing when they are preceded by the numeral *san* 'three', but not by other numerals.

- (81) a. san-bon 'three-CLS_{long.object}' (hon → bon)
 san-bai 'three-CLS_{container}' (hai → bai)
 b. yon-hon 'four- CLS_{long.object}' (*yon-bon)
 yon-hai 'four- CLS_{container}' (*yon-bai)

The same pattern is observed with the multiplicands *hyaku* 'hundred' and *sen* 'thousand', as shown in (82).

- (82) a. san-byaku 'three-hundred' (hyaku → byaku)
 san-zen 'three-thousand' (sen → zen)
 b. yon-hyaku 'four-hundred' (*yon-byaku)
 yon-sen 'four-thousand' (*yon-zen)

This is another piece of evidence that multiplicands are morho-syntactically similar to numeral classifiers.²⁹ It should also be noted that Greenberg (1972, 1978) claims that in classifier languages, the order between numerals and classifiers harmonize with the order between multipliers and multiplicands (see also Her (2017) for a recent discussion of this

²⁹ See Tatsumi (2021) for additional cross-linguistic data about similarities between classifiers and multiplicands.

harmonization effect). In other words, the (im)possible constituent orders of multipliers, multiplicands, and the main nouns pattern with Greenberg's observation summarized in (71) above. The two orders where the main noun intervenes between a multiplier and a multiplicand (i.e. multiplier-N-multiplicand or multiplicand-N-multiplier) are not observed cross-linguistically (see also Cinque (2020) for a more recent survey on this issue).

Under Ionin & Matushansky's (2018) analysis, the Type 1 classifier languages like Japanese may also be accommodated by assuming that the structure in (83) is available in this type of classifier languages.

- (83) [three [hundred [CLS [NP cars]]]]

The structure in (83), however, faces some problems. For example, it is expected that the constituent consisting of the classifier and the main noun can be the target of syntactic operations, such as, for example, nominal ellipsis. In Japanese, however, a numeral classifier is required to license the elliptical construction, as shown in (84).

- (84) a. Atsushi-wa *hon nana-hyaku-satu-o katta*.
 Atsushi-TOP book seven-hundred-CLS-ACC bought
 'Atsushi bought seven hundred books.'
 b. Iori-wa {*go-hyaku-satsu* | **go-hyaku*}*-o katta*.
 Iori-TOP five-hundred-CLS five-hundred-ACC bought
 'Iori bought five hundred.'

If Japanese numeral classifiers have the structure in (83), it is unexpected that the constituent consisting of the classifier and the main noun can be the target of nominal ellipsis.

Numerical classifiers in Mandarin Chinese and Vietnamese also provide data which are problematic for the analysis given in (83).

- (85) Mandarin: Shengyun Gu, p.c.
- a. Qiang mai le [wu tiao xianglian].
 Qiang buy ASP five CLS necklace
 'Qiang bought five necklaces.'
 - b. xianglian₁ Qiang mai le [wu tiao Δ₁].
 necklace Qiang buy ASP five CLS
 - c. *[tiao xianglian]₁ Qiang mai le [wu Δ₁].
 CLS necklace Qiang buy ASP five

- (86) Vietnamese: Thuy Bui, p.c.
- a. Khanh mua [năm cuốn sách].
 Khanh bought five CLS book
 'Khanh bought five books.'
 - b. sách₁ Khanh mua [năm cuốn Δ₁].
 book Khanh bought five CLS
 - c. *[cuốn sách]₁ Khanh mua [năm Δ₁].
 CLS book Khanh bought five

As shown in the b-examples of (85) and (86), the main noun can move to the sentence initial position, leaving the cardinal and the numeral classifier in situ. However, it is impossible to move the numeral classifier and the main noun together, as in the c-examples in these classifier languages. This pattern is not expected for the structure given in (83). The classifier

and the main noun form a constituent to the exclusion of the MCC and it is thus not clear why the constituent cannot be the target of the relevant extraction operation.

Another approach to MCCs in classifier languages is to assume that MCCs and classifiers form a constituent to the exclusion of the main noun, as represented in (87).

- (87) [[three hundred CLS] [NP cars]]

Under this analysis, the unacceptability of in (85c) and (86c) can be explained. The classifier head and the noun do not form a constituent and cannot undergo the relevant operation.

In summary: this section investigated some morphosyntactic properties of MCCs in classifier languages. MCCs in classifier languages have not been studied in great depth and are often assumed to be similar to MCCs in non-classifier languages. It is important to keep in mind, however, that MCCs in classifier languages could possibly have notable morphosyntactic properties that are not observed in non-classifier languages.

5.3 Additive complex cardinals in classifier languages

This section discusses additive complex cardinals (ACCs) in classifier languages. In Section 4, it was shown that there are three major types of analyses of ACCs in the literature: the compounding analysis, the non-constituent analysis and the single constituent analysis.

To our knowledge, the coordinative compound analysis of ACCs has not been extensively discussed in the literature on classifier languages. However, the coordinative compound analysis of ACCs can be possibly applied to ACCs in classifier languages, too. In Mandarin Chinese, for example, ACCs seem to form a morpho-phonological unit to the exclusion of the main noun. As shown in (88), the numeral *yi* 'one' in Mandarin has the high tone (T1) when it appears as the final syllable of a word boundary. In the number counting context, it is thus pronounced as in (88).

- (88) yi¹ 'one', er⁴ 'two', san¹ 'three', si⁴ 'four', wu³ 'five', ...

When the numeral *yi* is followed by another element, it bears the falling tone (T4), as shown in (89b). What is important is that this change of tones is not observed in ACCs. In (90), the ACC contains the numeral *yi* as its right-hand cardinal. In this case, *yi* cannot bear the falling tone. Instead, it maintains the citation tone as in (90a).

- (89) Mandarin: He (2015): 197

a.	yi ¹ ben ³ shu ¹	⇒	b.	yi ⁴ ben ³ shu ¹
	one CLS book			one CLS book
	'one book'			'one book'

(90)	a.	wu ³ shi ² yi ¹ ben ³ shu ¹	b.	*wu ³ shi ² yi ⁴ ben ³ shu ¹
		five ten one CLS book		ten one CLS book
		'eleven books'		'eleven books'

Her and Tsai (2020) gives these kinds of contrasts between simplex cardinals and ACCs in Mandarin as an argument against Ionin and Matushansky's (2018) syntactic coordination analysis of ACCs.

Note also that in ACCs the right-hand numeral is opaque to certain morphological processes. For example, Mandarin Chinese has a contracted form consisting of *san* ‘three’ and the general classifier *ge*, as shown in (91b).³⁰

- (91) a. san-ge xuesheng
three-CLS student
'three students'
b. sa xuesheng
three.CLS student
'three students'

He (2015) points out that the contracted form cannot appear in ACCs. The relevant examples are given in (92).

- (92) a. si-shi san-ge xuesheng
four-ten three-CLS student
'forty three students'
b. *si-shi sa xuesheng
four-ten three.CLS student
'forty three students'

The contrast in (92) would be unexpected under Ionin and Matushansky's (2018) coordination analysis. According to this analysis, the ACC in (92a) would have the structure in (93).

- (93) [four-ten CLS student] & [three CLS student]

First of all, it is controversial whether it is possible to elide the classifier *ge* and the main noun in the first conjunct in (93). The second issue here is that even if such a deletion process is available, it is expected that the numeral *san* and the classifier *ge* in the second conjunct should behave exactly like the numeral classifier in (91b), and hence the contracted form *sa* should occur in (92a), just like in (91a,b), contrary to fact. As an alternative analysis, He (2015) proposes that ACCs form a constituent to the exclusion of the main noun, as represented in (94).

- (94) [[four-ten three] CLS] student [He (2015)]

He (2015) posits a sisterhood requirement onto the fused form of numeral classifiers. In (94), the numeral ‘three’ and the classifier are not in a sisterhood relation, and hence the fused form is blocked.

Although these data from Mandarin Chinese can be seen as counter-examples to the coordination analysis proposed by Ionin and Matushansky (2018), the coordination analysis still receives support from the cross-linguistic pattern of ACCs. ACCs are analyzed as including a coordination phrase. Mandarin Chinese does not allow for the presence of a coordination particle in ACCs, as shown in (95) (Recall that as discussed in the previous sections, there are languages which can contain an overt coordination particle in ACCs. Also, see He *et al* (2015) for a more detailed observation about the presence of an overt coordination particle in ACCs observed in languages in South China.)

- (95) *sishi he san-ge xuesheng [Mandarin]

³⁰ Similarly, the numeral *liang* ‘two’ has the fused form *lia*, and the fused form basically behaves like the fused form of *san*.

(i) a. liang-ge xuesheng
two-CLS student
'two students'
b. lia xuesheng
two.CLS student
'two students'

Here, we focus only on the fused form of *san* because the numeral *liang* is prevented from appearing inside ACCs independently. As will be discussed, *er* ‘two’ must be used in Mandarin ACCs instead.

forty and three-CLS student
 ‘forty three students’

Mandarin examples like (95) do not mean that ACCs in classifier languages generally lack an overt coordination particle. There is actually a certain amount classifier languages which exhibit the coordination pattern. One example comes from Toqabaqita [Austronesian]. As shown in (96), the main noun (*ngali*) is repeated in each digit of the ACC in Toqabaqita.

- (96) [teqe taafalu fa ngali] [roo fa ngali] [Toqabaqita: Lichtenberk (2008: 296)]
 one ten CLS year two CLS year
 ‘twelve years’

In (96), the augend and the addend contain the main noun preceded by the classifier and the numeral. Moreover, ACCs in Toqabaqita can be optionally coordinated by the conjunction particle *ma* ‘and’, as shown in (98).

- (97) Toqabaqita : Lichtenberk (2008: 293)
 [teqe toqoni imola] *ma* [roo talange-qe imole] *ma* [lima akwale-qe imole]
 one thousand person and two hundred-ASC person and five tensome-ASC person
 ‘one thousand two hundred fifty people’

These patterns of ACCs in Toqabaqita can be seen as support for Ionin and Matushansky's coordination analysis, in which ACCs are analyzed as involving the coordination of NPs.

In this respect, ACCs in Japanese give us a clue about the nature of ACCs. As shown in (98), Japanese ACCs are normally used without an overt coordination particle.

- (98) yon-zyuu go-nin nogakusei [Japanese]
 four-ten five-cls-gen student
 'forty five students'

If Japanese ACCs are analyzed in the same way as Mandarin Chinese ACCs, it is expected that the right-hand numeral shows certain types of opaqueness to morphosyntactic processes. This turns out to be the case. Japanese has two classifiers for common nouns referring to human beings: *nin* and *ri*. Crucially, the classifier *ri* has a contextual restriction regarding the type of cardinal it combines with. It co-occurs with the native Yamato cardinals *hito* ‘one’ and *huta* ‘two’, as in (99a), but not with the Sino-Japanese cardinals *ichi* ‘one’ and *ni* ‘two’, as shown in (99b).

- (99) a. gakusei {hito | huta}-ri b. *gakusei {ichi | ni}-ri
 student one two-CLS student one two-CLS
 '{one | two} student(s)' '{one | two} students'

However, the selectional relationship between the human classifier and the numeral ‘one’ and ‘two’ disappears when the numerals ‘one’ and ‘two’ are used as right-hand cardinals in ACCs, as shown in (100a,b). In order to use the numeral ‘one’ or ‘two’ as the right-hand cardinal of ACCs, the classifier *nin*, which is the elsewhere exponent of the classifier head dedicated to human beings (Watanabe 2010, 2014), must be used together with the Sino-Japanese cardinals, as shown in (100c).

- (100) a. *gakusei [yon-zyuu {hito | huta}]-ri

- b. student four-ten one two-CLS
**gakusei [yon-zyuu {ichi | ni}]-ri*
 student four-ten one two-CLS
- c. gakusei [yon-zyuu {ichi | ni}]-nin
 student four-ten one two-CLS
 'forty {one | two} students'

The blocking effect of Japanese ACCs is quite similar to the ban on the fused form of the numeral in Mandarin Chinese. Japanese, however, is different from Mandarin in one important respect. Although it is slightly degraded for some speakers, it is possible to overtly realize the coordination particle *-to* 'and' in ACCs.

- (101) gakusei yon-zyuu *to* go-nin [Japanese] (Compare Mandarin Chinese (95))
 student four-ten and five-CLS
 'forty five students'

What is important here is that when the coordination particle is overtly realized, the blocking effect observed in Japanese ACCs disappears, as shown in (102), which is adopted from Hiraiwa (2016).

- (102) gakusei [yon-zyuu *to* {hito | huta}]-ri
 student four-ten and one two-CLS
 'forty {one | two} students'

The contrast between (100) and (102) can be accounted for by assuming that when a Japanese ACC contains the overt coordinator, the coordination pattern involves classifier phrases, as in (103).

- (103) gakusei [[yon-zyuu-CLS] *to* [{hito | huta}-CLS]]

In (103), the classifier in the first conjunct is deleted, on a par with the deletion process suggested by Ionin and Matushansky (2018).³¹ What is important is that in the second conjunct, the numeral and the human classifier stand in a local relationship; consequently, the selectional relationship between them is preserved.

Recall that Ionin and Matushansky (2018) propose that additive complex cardinals generally involve coordination of NPs. In some languages, a coordinator can be overtly realized in ACCs. Importantly, the absence of an overt coordinator does not imply that the coordinating head is also syntactically absent in their analysis. In other words, a coordinator can be present in the syntactic structure but absent phonologically, as in English *three hundred {and/Ø} twenty five*. Japanese ACCs, however, turn out to have different syntactic structures, depending on the presence/absence of an overt coordinator. This contrast as regards the (c)overtness of the coordinator has a strong influence on the morphosyntactic behavior of ACCs in Japanese.

In summary: Section 5 focused on numerical expression in classifier languages based on the discussions in sections 2, 3 and 4. The categorial status of cardinals in classifier languages is relatively fixed in comparison to the categorial status of cardinals in non-classifier languages. This might be related to the presence of classifiers in cardinal constructions. Regarding MCCs, it was shown that multiplicands behave like classifiers in

³¹ The exact nature of the deletion process deserves further investigation. See Section 4 for discussion of potential accounts of the deleted element in non-classifier languages.

several respects. Classifiers thus also have influence on the morphosyntactic behavior of MCCs. This Section also discussed ACCs in classifier languages. As was clear from Section 4, different types of structural analyses have been pursued for ACCs in the literature on non-classifier languages. In this section, it was shown that ACCs in classifier languages may give potential support for those analyses. Further investigation of these numerical expressions in classifier languages is much recommended.

6. The inner structure of ordinal numbers

This section discusses the inner organization of ordinal numbers (henceforth: ordinals). In various studies on ordinals, it has been argued that they are derived from cardinals cross-linguistically (Stampe (1976), Veselinova (1997), Stoltz and Veselinova (2005), and Zabbal (2005)). In many languages, numerical expressions become ordinals by means of certain markers. For instance, the suffix *-th* is used as an ordinal marker in English.³²

6.1 Strategies of expressing ordinals

There are various ways of expressing ordinals in non-classifier languages. (104) presents the major strategies, based on Tatsumi's (2021) survey, which also includes data from Veselinova (1997), Stump (2010), and Slotz and Veselinova (2013).³³

- (104) *Strategies for expressing ordinals in non-classifier languages*

a. Particular affixes/particles:

(i) Post-numeral (# < ORD, 27 languages, 19 genera (cf. Dryer (1989))):

Basque [Basque], Bilua [Papuan], Breton [Celtic], Dutch [Germanic], English [Germanic], Estonian [Uralic], German [Geman], Greek [Greek], Hunzib [Nakh-Daghestanian], Hindi [Indic], Italian [Romance], Kashmiri [Indic], Kazakh [Turkic], Khanty (Earstern) [Uralic], Kisi [Niger-Congo], Koromfe [Niger-Congo], Kurmanji [Iranian], Lezgian [Nakh-Daghestanian], Luwo [Niloitic], Mam [Mayan], Maninka [Bantu], Oko [Niger-Congo], Russian [Slavic], Serbian [Slavic], Span- ish [Romance], Sumerian [Sumerian], Welsh [Celtic]

(ii) Pre-numeral (ORD < #, 10 languages, 5 genera):

Ahàñ [Niger-Congo], Choctaw [Muskocean], Crow [Siouan], Kokota [Austronesian], Kula [Alor-Pantar], Pangasinan [Austronesian], Rapa nui [Austronesian], Sawila [Alor-Pantar], Seediq [Austronesian], Tagalog [Austronesian]

(iii) Circumfixal (ORD # ORD, 6 language, 5 genera):

Garifuna [Arawak], Georgian [Kartvelian], Kanuri [Saharan], Macushi [Carib], Chamula [Mayan], Tojobal [Mayan]

³² It is important to note that, in this chapter, we do not count the definite article *the* as an English ordinal marker because it can be omitted in some environments while keeping the meaning of ordinals (i.e. specifying the rank of a given item in a certain order). For example, English ordinals can appear without *the* in the adverbial use (e.g. *Second*, ..., *Third*, ..., *Fourth*, ...), some idiomatic expressions (e.g., *First of all*, *second to none*), and dates (e.g. *January ninth*).

³³ The names of language families used in this thesis are based mainly on *The World Atlas of Language Structure Online*.

b. Relative clauses (6 languages, 5 genera):

Akan [Niger-Congo], Biak [Austronesian], Gaahmg [Nilo-Saharan], Leti [Austronesian], Tamil6 [Dravidian], Turkana [Sudanic]

c. Postposition (3 languages, 3 genera):

Cahuilla [Uto-Aztekan], Navajo [Athapaskan] ('in, inside'), Yup'ik [Eskimo-Aleut]

d. Attributive/Associative construction (3 languages, 2 genera):

Meithei [Sino-Tibetan], Kikuyu/Gĩkũyũ [Niger-Congo], Swahili [Niger-Congo]

e. Overt “number”: English [Germanic], and many other languages

f. Definiteness (1 language, 1 genus): Maltese [Semitic]

Importantly, the same strategies for the formation of ordinals can be found in classifier languages. The overview in (105) is based again on Tatsumi (2021).

(105) *Strategies for expressing ordinals in optional/obligatory classifier languages*

a. Particular affixes/particles:

(i) Post-numeral (# < ORD, 8 languages, 8 genera):

Atong [Tibeto-Burman], Boko [Mande], Hungarian [Uralic], Japanese [Japanese], Korean [Korean], Lai/Chin [Sino-Tibetan], Toqabaqita [Austronesian], Turkish [Turkic]

(ii) Pre-numeral (ORD < #, 14 languages, 8 genera):

Anong [Sino-Tibetan], Bih [Austronesian], Hakha Lai [Sino-Tibetan], Helong [Austronesian], Indonesian [Austronesian], Japanese [Japanese], Kana [Niger-Congo], Lamjung Yolmo [Tibeto-Burman], Mandarin Chinese [Sino-Tibetan], Mokilese [Austronesian], Metting Belait [Austronesian], Nung [Tai-Kadai], Telugu [Dravidian], Xong [Hmong–Mien]

(iii) Circumfixal 2 languages, 1 genus):

Batak (Karo) [Austronesian], Batak (Toba) [Austronesian]

b. Relative clauses (4 languages, 4 genera):

Abun [Papuan], Kiribati/Kiribatese [Austronesian], Thai [Tai-Kadai], Vietnamese [Austro-Asiatic]

c. Preposition (1 language, 1 genus):

Achenese [Austronesian] (preposition ‘to’)

d. Possessives (4 languages, 3 genera):

Abui [Alor-Pantar], Chontal Maya [Mayan], Itzaj [Mayan], Kove [Austronesian]

e. Overt “number” (2 languages, 2 genera):

Bislama [Creole in Vanuatu], Japanese [Japanese]

f. Definiteness (2 language, 2 genera):

Nuosu [Sino-Tibetan], Samoan [Austronesian]

In (106)-(112) illustrations of each strategy are given. For reasons of space, each strategy is illustrated by a single example. A schematic representation of each strategy has been added.

(106) Particular affixes/particles strategy

- a. Post-numeral: Num-X_{ORD}
seven-th (English)
- b. Pre-numeral: X_{ORD}-Num
ól-íro ashí (Àhàn, Ogunmodimu 2015: 69)
ORD-eight dog
'eight dog'
- c. Circumfixal: X_{ORD}-Num- X_{ORD}
me-ekvs-e (Georgian, Hewitt 1995: 56)
ORD-six-ORD
'sixth'

(107) The relative clause strategy

- a. [[RelP ... N_{PRED} ...] N_{MAIN}]
- b. [[RelP ... Num ...] N_{MAIN}] (Ordinals)
- c. rum [ve-ve-suru] (Biak, Steinhauer 2005: 802)
house REL-QUA-two
'second home'

(108) The postpositional strategy

- a. [PP N P]
- b. [PP Num P] (Ordinals)
- c. naaki góne' (Navajo, Young and Morgan 1980: 37)
two inside
'second'

(109) The attributive/associative strategy

- a. [N_{MAIN} (ASSOC) N_{MOD}]
- b. [N_{MAIN} (ASSOC) Num] (Ordinals)
- c. kitabu cha tatu (Swahili, Polomé: 1967: 133)
book LINK three
'the third book'

(110) The possessive strategy

- a. [NPOSSESSOR POSS NPOSSESSEE]
- b. [NPOSSESSOR POSS Num] (Ordinals)
- c. moe hua ai-a (Kove, Sato 2013: 197)
sleeping.mattress two 3.SG.POSS-A.POSS
'the second sheet of sleeping mattress'

(111) The overt "number" strategy

- [N NUMBER Num]
'the question number four' (English)

(112) The definiteness strategy

- a. [DEF N]
- b. [DEF Num] (Ordinals)

- c. 'o le tolu (Samoan, Neffgan 1918: 38)
 the.SG three
 'third'

In the relative clause strategy, ordinals can be analyzed as the nominal predicate in a relative clause. In the postpositional strategy, numerals are nominal in the sense that they combine with a postposition. In the attributive/associative strategy, ordinals also behave like nominal modifiers. The possessive relation typically holds between two referents denoted by nouns. In the possessive strategy, the ordinal appears in the same position as a possessed phrase (i.e. possessee). This indicates that ordinals in this strategy are also nominal. The overt “number” strategy includes an overt nominal element meaning ‘number’. The overt “number” itself can be the source of noun-hood of ordinals. Lastly, the definiteness strategy also quite clearly shows the nominal nature of ordinals. The definiteness marker typically appears with a nominal expression.

6.2 The categorial nature of ordinals

In each of the above-mentioned strategies for expressing ordinals, numerical expressions appear in a position where nominals can occur. The following cross-linguistic generalization can thus be formulated:

- (113) Numerals used for expressing ordinality appear in the position where nominals can occur.

The generalization in (113) expresses that numerals are inherently nominal and acquire the ability to function as ordinals as a result of morphosyntactic derivations. Support for this view may come from ordinals in Serbo-Croatian. As shown in (114), some higher cardinals in Serbo-Croatian show the so-called genitive of quantification.

- (114) Ivan je pročitao pet knjiga. (Serbo-Croatian, Ivana Jovović, p.c.)
Ivan CL read five books.GEN
'Ivan read five books.'

However, unlike higher cardinals, ordinals generally do not assign the genitive of quantification to the main noun, as shown in (115). Instead, they agree with the main noun, similarly to adjectives.

- (115) a. Ivan je pročitao petu knjigu. (Serbo-Croatian, Ivana Jovović, p.c.)
Ivan CL read five.ACC book.ACC
'Ivan read the fifth book.'
c. *Ivan je pročitao petu knjiga.
Ivan CL read five.ACC book.GEN

According to Ionin and Matushansky (2018), the case-assignment property is a hallmark of the nominal status of a given cardinal. Moreover, agreeing numerals are classified as adjectives in their analysis. The cardinal in (114) is consequently analyzed as a noun. The ordinal in (115a), on the contrary, must be adjectival in view of its agreement relationship with the main noun. In short, numerical nouns in Serbo-Croatian become adjectival modifiers in the ordinal construction.

6.3 Derivations of ordinals

As various studies have shown, ordinals are derived from numerals (Stampe (1976), Veselinova (1997), Stoltz and Veselinova (2005)). Zabbal (2005) gives an explicit structural implementation of this idea. He proposes that ordinals and cardinals are derived from bare numerals by means of the Num head. According to Zabbal's analysis, numerals uniformly appear in Spec,NumP, and the function of numerals in Spec,NumP is determined by an operator generated in the Num head, as shown in (116). As a side note, it should be mentioned that Zabbal adopts Chomsky's (1995) Bare Phrase Structure view; so all the labels in (116) are N or A#, where A# stands for 'adjectival numeral'.

- (116) a. Cardinals: *two hundred five soldiers* (Zabbal 2005: 39)
 [D [D the] [Num [A# two hundred and five] [Num [Num COUNT] [N soldiers]]]]
 a. Ordinals: *two hundred fifth soldier* (Zabbal 2005: 44)
 [D [D the] [Num [A# two hundred and five] [Num [Num -th] [N soldier]]]]

Zabbal's analysis predicts that cardinals and ordinals are in complementary distribution because they occur in the same position, namely Spec,NumP. Zabbal gives the examples in (117) to show that this prediction is borne out.

- (117) a. *John carried the tenth hundred bags.
 b. *John carried the hundred tenth bag. (Zabbal 2005: 37)

It should be noted, however, that ordinals can co-occur with cardinals in some cases. The example in (118) is taken from Rothstein (2017:42).

- (118) Mary won three second prizes in last year's running season.

Moreover, it is possible to iterate ordinals in the single nominal domain. Suppose that we have a line of balls in the context, as illustrated in (119).

- (119) OOOO OOOO OOOO OOOO OOOO

Now, the speaker can utter (120) to refer to the underlined ball in (119).

- (120) Pick up the fourth third ball. [English: Nicolaus Schrum, p.c.]

Although Zabbal's analysis allows us to treat cardinals and ordinals in a uniform way, it is not clear how to deal with the examples in (118) and (120) under his analysis.

In light of this, it is worth noting that ordinal markers exhibit the same pattern as classifiers or multiplicands regarding constituent orders. Consider for this (121), which is a summary of a cross-linguistic survey of constituent orders involving nominals (N), numerals (Num) and ordinal markers (ORD); the data are taken from Tatsumi (2021).³⁴

³⁴ The constituent orders in (121) should not be confused with constituent orders of cardinals and ordinals. For example, English ordinals have the Num-ORD-N order because *-th* is identified as an ordinal marker (ORD) (the six-th book = Num-ORD-N). In Tatsumi (2021), the definite article *the* in English is not counted as an ordinal marker because it can be omitted in some cases, while keeping the meaning of ordinals (e.g. *First of all*).

Regarding orders between cardinals and ordinals, it has been observed that ordinals tend to be located higher than cardinals (Fassi Fehri 1999, Shlonsky 2004) (e.g. *the first five chapters* = Ordinal-Cardinal-Noun). Fassi Fehri (1999) reports

(121)

Orders	Languages
✓ Num-ORD-N	Basque, Bilua, Breton, Dutch, English, German, Greek, Hindi, Italian, Kashmiri, Khanty (Earstern), Lezgian, Mam, Meithei, Russian, Serbian, Spanish, Welsh
✓ N-Num-ORD	Sumerian, Koromfe, Kurmanji, Luwo
✓ N-ORD-Num	Biak, Crow, Kikuyu/Gĩkũyũ, Sawila
✓ ORD-Num-N	Àhàn, Maltese
* Num-N-ORD	None
* ORD-N-Num	None

There are six mathematically possible combinations of N, Num, and Ord (factorial 3 = $3 \times 2 \times 1 = 6$). The last two combinations in (121), however, are not attested in Tatsumi's (2021) sample. The question, obviously, arises as to why these two constituent orders are unattested. Importantly, the two unattested combinations in (121) correspond to Greenberg's (1972) unattested constituent orders of nouns, numerals, and classifiers, namely *Num-N-CLS and *CLS-N-Num. Moreover, the constituent orders harmonize with word order patterns involving a multiplicand, a multiplier and a noun. Recall from Section 5.2 that the following patterns are unattested: *Multiplier-N-Multiplicand, *Multiplicand-N-Multiplier. A simple way to explain the similar word order behavior of ordinal markers, classifiers and multiplicands is to assume that they occur in the same syntactic position in the nominal domain. This possibility is pursued by Tatsumi (2021), but further careful study is needed. The syntactic status of ordinal markers will also influence the (un)availability of ordinal suppletion, as will be discussed in the next section.

6.4. Ordinals derived from complex cardinals and ordinal suppletion

The previous sections discussed some cross-linguistic properties of simplex ordinals. This section examines the structure of ordinals that are based on complex cardinals.

Stump (2010) investigates derivations of ordinals containing ACCs on the basis of a wide range of languages. His conclusion is that the rules of morphological composition of

that in Standard Arabic, cardinals must precede ordinals as shown in (i).

(i) a. ?awwal-u xams-i muhaadaraat-in
first-NOM five-GEN lectures-GEN
'The first five lectures'
b. * xams-u ?awwal-i muhaadaraat-in
five-NOM first-GEN lectures-gen
'The first five lectures'
(Standard Arabic, Fassi Fehri 1999: 113)

However, it seems that some languages allow both Cardinal-Ordinal and Ordinal-Cardinal orders. For example, English allows the Cardinal-Ordinal order as shown in (118), in addition to the Ordinal-Cardinal order. Antonio Fábregas (p.c.) also noted that Spanish also shows a similar pattern allowing both Cardinal-Ordinal and Ordinal-Cardinal orders. Although we do not have space here to discuss this issue, it clearly deserves further investigation.

ordinals varies across languages, and the simple morphological rule in (122) alone cannot capture the whole range of variations.

- (122) $\text{Deriv}([\text{X } \text{Y}]) = [\text{X } \text{Deriv}(\text{Y})]$

Here, $[\text{X } \text{Y}]$ is an ACC consisting of X and Y. *Deriv* stands for a pertinent morphological rule for deriving ordinals. According to (122), the ordinal formation rule applies only to the right-hand numeral in ACCs.

Stump (2010) observes that there are six patterns of ordinal marking in his sample, as summarized in (123):

- (123) External ordinal marking:

An ordinal formation applying to a complex numeral as an unanalyzed whole.

E.g. Kanuri [Nilo-Saharan], Old Georgian [Kartvelian]

Internal right-peripheral type:

Only the right-peripheral numeral exhibits ordinal marking

E.g. English

Internal left-peripheral type:

Only the left-peripheral numeral exhibits ordinal marking

E.g. Anywa [Nilo-Saharan; Eastern Sudanic], Nobiin [Nilo-Saharan; Eastern Sudanic]

Extended internal type

a. Operation-sensitive:

(i) Limited: E.g. Czech (applicable to a multiple of ten less than 100)

(ii) Unlimited: E.g. Portuguese (all addends are eligible for ordinal formation)

b. Operation-insensitive

E.g. Finnish, Modern Greek

In the external ordinal marking, an ordinal marker does not appear inside the target complex cardinal. The ordinal marking in Kanuri, for example, is circumfixal, and it combines with an ACC as a noncomplex unit, as in (124); the examples are taken from Stump (2010).

- (124) Kanuri: Cyffer 2007: 1106–07; Hutchison 1981: 76–77, 202–03

a.	findin (lúkko) tilôn	b.	<i>kón</i> -findin (lúkko) tilôn- <i>mi</i>
	twenty and one		ORD-twenty and one-ORD
	'twenty one'		'twenty first'

In other ordinal marking strategies, ordinal markers appear inside ACCs. English is an example of the internal right-peripheral type, in which the ordinal formation applies to the right-peripheral numeral of a given ACC, as in $[[\text{twenty}] [\text{four-th}]]$. Stump (2010) assumes that suppletive ordinals are derived by applying an ordinal formation rule (*Deriv(X)*). The suppletive ordinal *first* appears in ordinals derived from ACCs, as shown in (125a). English does not belong to the left-peripheral type because (125b) is unacceptable as an ordinal number.

- (125) a. Please open the twenty first book. (English)
b. *Please open the twentieth one book.

Stump (2010) gives Anywa [Nilo-Saharan] as an example of a left-peripheral type of language. The relevant examples are given in (126).

- (126) Anywa: Reh 1996: 286–88.

a.	āpáar kúr cíél ten and one 'eleven'	b.	páaJ-gī kúr cíél ten-ORD and one 'eleventh'
----	---	----	---

As shown in (126b), the ordinal marker attaches to the left-peripheral numeral in the ACC.

In extended internal ordinal marking languages, ordinal markers attach to two or more numerals in a given ACC.³⁵ An example of this type of ordinal marking is given in (128).

- (127) por favor abre el vig-ésimo (??y) *primer* libro. (Spanish, Gabriel Martínez Vera, p.c.)
please open the twenty-ORD and first book
'Please open the twenty first book.'

In (127), the left-hand numeral is marked by the ordinal marker *-ésimo*, and the right-hand numeral also undergoes ordinal suppletion (*uno* > *primer*). In contrast to Spanish, the English ordinal marker *-th* cannot appear on the left-hand numeral in ACCs. (128a) below is unacceptable and cannot mean "please open the twenty first book". When the main noun is plural as in (128b), the sentence is acceptable, but the object is interpreted as involving coordination of two nouns modified by the different ordinals (i.e. "open the twentieth book and the first book").

- (128) a. *Please open the twentieth and first book.
b. Please open the twentieth and first books.

It should be noted here that it is unclear how to account for the unacceptability of (128a) under the NP-coordination analysis proposed by Ionin and Matushansky (2018). The NP-coordination analysis allows for the structure in (129) and does not predict the unacceptability of (128a).

- (129) the [twentieth bœok] and [first book]

The NP-coordination analysis also needs to explain why Spanish ordinals display the pattern that is unavailable in English.

The discussion so far showed that there are languages where suppletive ordinals can appear inside ACCs. As noted in Stump (2010), however, there are languages which do not allow suppletive ordinals to occur inside ACCs. One such language is Italian. As shown in (130a), Italian has the suppletive ordinal *primo* 'first'. This suppletive ordinal is blocked, however, when it is derived from an ACC, as shown in (130b); compare (130b) with the English ordinal in (125a). Example (130c) shows that the ordinal marker attaches to the entire ACC: *[[vent un]-esimo]*.

- (130) a. Per favore apri il *primo* libro. (Italian: Pietro Cerrone, p.c.)
please open the one.ORD book.SG.M
'Please open the first book.'
b. *Per favore apri il *vent(i)* *primo* libro.
please open the twenty one.ORD book.SG.M
'Please open the twenty first book.'
c. Per favore apri il *vent un-esimo* libro.

³⁵ The distinction between the operation-insensitive version and the operation-sensitive version is put aside here for reasons of space. See Stump (2010) for further information.

please open the twenty one-ORD book.SG.M
 'Please open the twenty first book.'

Unlike English and Spanish, the suppletive ordinal cannot be used as a part of an ordinal derived from ACCs in Italian. The blocking effect of the suppletive form may be partly related to morphosyntactic composition of Italian ACCs. As shown in (131), Italian ordinals derived from ACCs never allow for the presence of an overt coordinator, in contrast to Spanish ordinals; see (127).³⁶

- (131) *per favore apri il *vent e un-esimo* libro (Italian: Pietro Cerrone, p.c.)
 please open the twenty and one-ORD book.SG.M
 'Please open the twenty first book.'

Although more careful investigation is needed, the blocking effect might be explained in terms of the morphosyntactic composition of ACCs. As discussed in Section 4, three major types of analyses can be distinguished: (i) ACCs are coordinative compounds, (ii) ACCs do not constitute a single syntactic object, (iii) ACCs form a single syntactic object. Given this, it would not be unreasonable to assume that Italian and Spanish (and maybe English as well) have different morphosyntactic compositions, as represented in (132).

- (132) a. il [vent un]-esimo libro (Italian)
 b. el [&P vig-esimo (??y) primer] libro (Spanish)
 c. the [&P twenty & first] book (English)

Recall from Section 4 that, when the cardinal "one" in Polish appears in ACCs, it cannot agree with the main noun; see (66). It was suggested that the lack of agreement in Polish results from the right-hand cardinal's being trapped within the ACC. The blocking effect of the suppletive ordinal in Italian may be seen as a result of the same effect.³⁷

It should be noted that the morphosyntactic status of what we have called suppletive ordinals is controversial. In some studies, the term 'suppletion' has been used only for inflectional morphology, but not for derivational morphology, because semantic correlations are more transparent in inflectional morphology. According to this definition, ordinal suppletion may not be a true instance of suppletion. Markey (1985) and Mel'čuk (1994), however, include derivational morphology as the domain of suppletion. In this chapter, we do not commit ourselves to this debate. Notice, however, that Buck (1949:939) mentions that "the words for 'first' have no connection with the cardinal 'one'" in his language sample. Veselinova (1997) also observes some non-numerical sources for the suppletive ordinal 'first', specifically a spatial source (e.g. 'front' or 'foremost') and a temporal source (e.g. 'beginning' or 'first time/early'). This could be related to the crosslinguistic tendency that suppletive ordinals are limited to low numbers (mostly "one" or "two"). Possibly then, suppletive ordinals are non-numerical lexical entries and do not undergo any morphosyntactic processes.

Suppletive ordinals also occur in classifier languages. Mokilese, which is an obligatory classifier language, uses the pre-numeral ordinal marker *ka-* to express ordinals. In (133), the ordinal marker co-occurs with a numeral classifier (*pas*). As regards the ordinal of the numeral 'one', Mokilese uses the suppletive form *keiow*, as shown in (133b).

³⁶ It should be noted that the presence of the coordination particle *y* is somewhat awkward in (127), if not unacceptable. We put aside here the cause of the awkwardness in (127).

³⁷ For discussion of ordinal suppletion in Dutch, see Barbiers (2007) and Sleeman (2017). The latter study examines ordinal formation in Dutch from a cross-dialectal perspective (micro-variation).

- (133) Mokilese: Harrison (1976: 98)
- a. *Ka-rah-pas* in pinjellok kajda.
ORD-2-CLS of pencil cracked
'The second pencil is cracked.'
 - b. *Koah ioar keiow* in armaj ma suhoang ngoahi nehn imwen-wini.
you FOC 1.ORD of person who visit me in house.of.medicine
'You're the first person to visit to me in the hospital.'

Note that *armaj* 'person' appears with the classifier *-men* in the cardinal construction (e.g. *armaj pah-men*; person four-CLS, 'four people'). It is thus unclear why the classifier does not appear in (133b). One way to account for the absence of the classifier in (133b) is to assume that *keiow* is not a numerical expression. Instead, one may analyze *keiow* as a suppletive form of the numeral "one", an ordinal marker and a classifier. Clearly, this tentative proposal needs further investigation.

In summary: this section discussed the morphosyntactic behavior of ordinals in both classifier languages and non-classifier languages. After a presentation of cross-linguistic strategies that are used for the expression of ordinals, it was shown that, cross-linguistically, numerals used for the expression of ordinals typically appear in the structural position where nominals can occur. Besides a discussion of simple ordinals, this section also discussed ordinals derived from ACCs. This was based on Stump's (2010) typological work. Although there still remain some controversial issues regarding the nature of suppletive ordinals, several patterns of ordinal marking have been identified in this section.

7. Numerals and bound morphology: other cases.

As was shown in the preceding sections, numerals can have a composite structure, as in *three hundred* and *hundred and three*. In these examples *hundred* and *three* are free morphemes, that is, they can occur on their own, as in *We are three* or *We are hundred*. Besides composite numerals consisting of free morphemes, there are also ones derived by means of a bound morpheme, as in *sixty*, where *-ty* is a bound morpheme attached to the numeral *six*. This section discusses several constructions featuring a numeral in combination with a bound morpheme, specifically: composite forms such as *six-ty* (section 4.1), so-called "dressed numerals" in certain Germanic varieties (section 4.2), numerals used in counting (section 4.3), and collective numerals featuring plural morphology (4.4).

7.1 Suffixes deriving cardinal numbers

In Indo-European languages such as English, Dutch, French and Italian, the formation of numerals from 20 to 90 that correspond to multiples of 10—that is, 20, 30, 40, *et cetera*—involves the use of a bound morpheme. Some examples are given in (135):

- (134) a. for-ty, fif-ty, six-ty seven-ty English
 b. veer-tig, vijf-tig, zes-tig, zeven-tig³⁸ Dutch
 c. quar-ante, cinqu-ante, soix-ante, sept-ante³⁹ French
 d. quar-anta, cinq-anta, sess-anta, sett-anta Italian

³⁸ Dutch *-tig* is pronounced as /təχ/.
³⁹ The form *septante* is used in Switzerland and Belgium (Wallonia). In France, the ACC *soixante dix* (six-ty ten, 'seventy') is used.

The bound morphemes *-ty* / *-tig* / *-ante* / *-anta* represent ‘ten’. A form like *six-ty* (and its equivalent in the other languages) has the meaning of an MCC, that is, 6 (multiplier) x 10 (multiplicand). In this respect, it is similar to the MCC *six hundred*, the only difference being the morphemic status of the multiplicand: in *six-ty*, the multiplicand is a bound morpheme, in *six hundred* a free morpheme.

Some of the forms in (134) are regular derivations from simple numerals. For example, English *sixty* and Dutch *zestig* feature the simple numerals *six* and *zes*, respectively, and French *cinqu-ante* and Italian *cinq-anta*, the simple numeral *cinque*. There are, however, also less regular formations. For example, English *fifty* features the phonological form /fif/, which differs from the numeral *five* (i.e. /faɪv/); Dutch *veertig* features *veer-*, which differs from the numeral *vier* (/vi:r/); French *soixante* and Italian *sessanta* feature the components *soix-* and *sess-*, which differ from French *six* ‘six’ and Italian *sei* ‘six’, respectively.

In view of the existence of irregular forms of some of the composite numerals in (134), one may come to the conclusion that these composite forms are complex *words* that result from a morphological rule that attaches a derivational morpheme (the bound-morphemic multiplicand *-ty* / *-tig* / *-ante* / *-anta*) to the numeral base (the free-morphemic multiplier); see, for example, Booij (2020) for Dutch. Schematically:

- (135) [word [word *six*] [suffix *-ty*]]

A more syntactic approach may emphasize the uniformity of meaning of MCCs such as *six hundred* (i.e. 6 x 100) and MCCs such as *six-ty* (i.e. 6 x 10), and lead to an analysis in which the two MCC-patterns are structurally the same. If, for example, the noun phrase *six hundred cars* has the structure in (136a) —see section 3—, then *six-ty cars* would have the structure in (136b):

- (136) a. [FP [YP *six*] [F' hundred [NP *cars*]]]
b. [FP [YP *six*] [F' -ty [NP *cars*]]]

An analysis of Dutch *-tig* along the lines of (136b) can be found in Kranendonk (2010). Interestingly, Dutch *tig* (pronounced /tɪχ/) can be used independently, as shown in (137a). Importantly, in that specific use, it does not represent ‘ten’ but has an indefinite meaning corresponding to ‘a great many/umpteen’. As shown in (137b), *tig*, just like the indefinite quantifier *veel* ‘many’ and the numeral *drie* ‘three’, can be used in combination with quantitative *er* (litt.: there, ‘of them’), which arguably has been extracted out of the noun phrase.

- (137) a. Jan heeft toen [tig postzegels] verzameld.
Jan has then umpteen stamps collected
‘Jan collected a great many stamps at the time.’
b. Jan heeft er toen [tig er] verzameld.
Jan has of-them then umpteen collected
‘Jan collected a great many of them at the time.’

Kayne (2005b) also takes the position that certain “derivational” suffixes are actually functional heads that play a role in syntax. He proposes, for example, that the French nominal suffix *-aine* (feminine in gender) is a functional head with approximative meaning, which has a numeral (e.g. *dix* ‘ten’, *douze* ‘twelve’, *quinze* ‘fifteen’, *vingt* ‘twenty’, *trente* ‘thirty’, *quarante* ‘forty’, *cent* ‘hundred’) in its specifier position. Schematically:⁴⁰

⁴⁰ The element *de* is taken to be a realization of the (genitival) case, assigned by the nominal suffix *-aine*.

- (138) [DP une [_{FP} vingt [_{F'} -aine [_{NP} d' articles]]]]]
 a twenty -aine of-articles
 ‘about twenty articles’

It seems fair to say that a more elaborate cross-linguistic investigation of bound-morphemic number elements is needed before any firm conclusions can be drawn about the question as to whether these are elements belonging to the domain of morphology or the domain of syntax.

7.2 ‘Dressed’ numerals

The study of numerals generally focusses on numeral expressions in standard languages. This holds for both single-language studies and comparative-linguistic studies. It is important to include micro-comparative data in the study of numerals, that is, data from closely related varieties of a single language, such as dialects. Micro-differences may throw light on the internal organization of numeral expressions and the way they are integrated into larger syntactic structures. This subsection focuses on a phenomenon attested in certain Dutch and German dialects, which we theory-neutrally characterize as the phenomenon of ‘dressed numerals’.⁴¹

In Standard Dutch, cardinal numerals are typically morphologically bare, both in their attributive use and in their “independent” (i.e. substantive) use:

- (139) a. Ik heb [**vier** koeien]. (Standard Dutch)
 I have four cows
 b. Ik heb [vijf koeien] en Jan heeft er [**vier**].
 I have five cows and Jan has *er* four
 ‘I own five cows and Jan owns four.’
 c. [**Vier** van de koeien] zijn ziek.
 four of the cows are ill

In (139a), *vier* is used attributively; it is followed by the noun *koeien*. In (139b), *vier* is used substantively; it is not followed by a phonologically overt noun. The nominal part that goes with *vier* is represented by the quantitative pro-form *er* (litt.: there, ‘of them’). In (139c), finally, *vier* is part of a partitive noun phrase, which may be paraphrased as ‘four cows of the set of cows’.

Consider now the same examples in Giethoorn Dutch, a dialect spoken in the north-eastern part of the Netherlands; see Corver and Kranendonk (2008), Kranendonk (2010), Corver (2010).

- (140) a. Ik heb [**vier** koeien]. (Giethoorn Dutch)
 b. Ik heb [vijf koeien] en Jan heeft er [**viere**].
 c. [**Viere** van de koeien] zijn ziek.

These examples show that, just like in Standard Dutch, the cardinal remains morphologically bare when it is used attributively. In their substantive use, however, Giethoorn Dutch cardinals display different behavior: they must be “morphologically dressed”, namely by *-e* (/ə/).⁴²

⁴¹ Dutch dialects featuring dressed numeral are typically found in the North-eastern part of the Netherlands (roughly the Saxonian area) and the South-west of the language area (the province of Zeeland, and Flanders (Belgium)). See Taeldeman (1984), Kranendonk (2010).

⁴² Another nominal construction that features a dressed numeral is the one given in (i); see Kranendonk (2010:89).

Lehmann (2010) observes similar contrasts between Standard German and colloquial German (see also Drosdowski *et al* 1984:278). In Standard German, the cardinal is morphologically bare both in its attributive use and its substantive use:

- (141) a. Ich habe [sechs Eier] gekauft.
I have six eggs bought
'I bought six eggs.'

b. A: Wie viele Eier willst du?
how many eggs want you
'How many eggs do you want?' B: Gib mir mal sechs.
give me PRT six
'Give me six.'

In colloquial German, *sechs* ‘six’ must be morphologically bare when followed by an overt noun, just as in Standard German. When it is used substantively, however, as in (142B), it must be morphologically “dressed”:

- (142) A: Wie viele Eier willst du?
how many eggs want you
'How many eggs do you want?' B: Gib mir mal **sechse**.
give me PRT six
'Give me six.'

The question obviously arises as to what the grammatical status of the bound morpheme *-e* following the numeral is. Both Van Kranendonk (2010:80-84) and Lehmann (2010) claim that *-e* should not be analyzed as an adjectival inflection. Attributive adjectives are typically inflected, as in Dutch *grot-e koeien* (big-INFL cows) and German *klein-e Eier* (small-INFL eggs), attributive cardinals, on the contrary, are not, as is shown by Dutch *vier(*-e) koeien* ‘four cows’ and German *sechs(*-e) Eier*. The two authors propose that the bound morpheme *-e* is a substantive-like element. Lehmann observes that, besides having the pattern *sechse* in (142B), colloquial German can also have the pattern *sechs Stück* (six piece, ‘six items’) in (141B). In Wiltschko (2007), *Stück* in German expressions such as *12 Stück Vieh* (12 piece cattle) is analyzed as a numeral classifier. Given the similar role of *-e* on *sechs-e* and *Stück*, one may analyze them as realizations of the same structural position. Kranendonk (2010:92-102), who also notes the parallel between classifier nouns and the *-e* on dressed numerals, analyzes *-e* as a realization of a classifier. Schematically:

- (143) [NumP vier [ClasP -e [Ø]]] (Ø = silent NP)

As noted in Taeldeman (1984:54), *-e* does not appear on cardinals ending in *-tig*. Nor does it appear on the cardinal numerals *honderd* (hundred) and *duizend* ‘thousand’; see also Kranendonk (201:77). Van Kranendonk proposes that *tig* ‘-ty’ is a classifier, which means that numerals like *vijftig* and *negentig* are composite forms, whose members occupy syntactically different positions (see section 7.1). If both *-e* and *-tig* instantiate a classifier-head, their non-cooccurrence is expected: they compete for the same syntactic position. From the fact that numerals such as *honderd* and *duizend* cannot occur as dressed numerals one may draw the conclusion that these cardinals occupy a classifier position. Such an analysis would be compatible with Tatsumi’s (2021) proposal that high numerals such as ‘hundred’ and

(i) boek vier-e (Giethoorn Dutch)
book four-*e*
'book number four/the fourth book'

‘thousand’ should be treated differently from low numerals such as ‘five’ and ‘nine’. The former are syntactic heads, while the latter are phrases in a specifier position. Interestingly, Tatsumi points out that the cardinals ‘hundred’ and ‘thousand’ occupy the same syntactic position as classifiers in classifier languages. In short, comparative-linguistic research at the micro-level and at the macro-level possibly converges here.

7.3 The form of “counting” numerals

Numerals can be used for counting. The surface form of the “counting” numeral is often similar to that of the cardinal numeral in its attributive use. Some illustrations are given in (144).

- (144) a. one, two, three, four, five, ..., ten, eleven, twelve, fifteen, sixteen, ... (English)
 b. un, deux, trois, quatre, ..., dix, onze, douze, quinze, seize, ... (French)

As noted in Hurford (1998, 2001), the surface form of the counting numeral sometimes differs from the surface form of the (attributive) cardinal numerals. The table in (145) gives some illustrations of this phenomenon:

	Number	Attributive	Counting
German ⁴³	1	ein(e)	eins
Maltese	2	żweg	tnejn
Mandarin	2	liang	er
Hungarian	2	két	kettő
Basque	2	bi	biga

Thus, in Maltese, for example, *żweg* is used when the numeral is followed by a noun —hence in an attributive environment— while *tnejn* is used in counting contexts. Importantly, the form of the counting numerals in (145) is not restricted to reciting a counting sequence. The counting forms can also be used in other structural environment. In Payne (1996), it is noted, for example, that Maltese *tnejn* occurs in nominal ellipsis constructions (146B) and partitive constructions (147a). Note the difference between the partitive pattern in (147a), where *tnejn* is used, and the nominal pattern in (147b) where the attributive form *żweg* is followed by the counted noun *djar* ‘houses’.

- (146) A: Kemm kien hemm irġiel?
 how many be.PST there man.PL
 ‘How many men were there?’
 B: Tnejn.
 two
 ‘Two (i.e. two men).’

⁴³ The numeral *zwei* ‘two’ is sometimes replaced by *zwo* in counting sequences.

- (147) a. Tnejn mid djar ta' missieri
 two from-the house.PL of father.my
 ‘two of my father’s houses’
 b. iż-żweg djar ta' missieri
 the-two house.PL of father.my
 ‘the two houses of my father’s’

Interestingly, dialectal varieties of Dutch that feature a schwa after the numeral when the latter is used substantively (see (140b)) or is part of a partitive construction (see (140c)), also use a schwa after “counting” numerals. Examples of “counting numerals” featuring *-e* (= /ə/) are given in (148a). These examples are taken from Ruinen Dutch (Sassen 1953:100):

- (148) a. iin-e (1), tweej-e (2), driej-e (3), veer-e (4), vie:v-e (5), ..., tien-e (ten), datien-e (13), viefien-e (15), achtien-e (18)
 b. da-teg (30), veer-teg (40), fief-teg (50), hondert (100), duu:zent (1000)

As shown in (148b), simple numerals (e.g. *veer*, ‘four’) and ACCs having the bound morpheme *-tien* ‘ten’ as augend carry *-e* (/ə/) when they are used as “counting numerals”. Interestingly, composite numerals featuring the bound morpheme *-teg* (English *-ty*) and high numerals such as ‘hundred’ and ‘thousand’ cannot be augmented with *-e*. One may try to account for this distributional behavior of *-e* in terms of complementary distribution; that is, *-e* occupies the same structural position as the bound morpheme *-teg*, and the free morpheme *hondert*. Notice, finally, that the phenomenon discussed in this section is arguably closely related or maybe even identical to the phenomenon of dressed numerals discussed in section 7.2.

Also in classifier languages, the surface form of counting numerals differs from that of cardinal numerals. In classifier languages, cardinals are expressed by numeral classifier phrases, but not by numerals alone. Therefore, the relevant distinction between counting numerals and cardinals often corresponds to the distinction between free forms and bound forms of numerals. Free forms tend to be used as counting numerals. Bound forms are often followed by numeral classifiers, and hence used as cardinals. For instance, Japanese has two different series of counting numerals, as exemplified in (149).

- (149) a. Counting forms in Native Japanese numerals
hii, huu, mii, yoo, itsu, muu, nana, yaa, kono/koko, too
 one, two, three, four, five, six, seven, eight, nine, ten
 b. Counting forms in Sino-Japanese numerals
ichi, ni, san, yon, go, roku, nana, hachi, kyuu, zyuu
 one, two, three, four, five, six, seven, eight, nine, ten

The first series contains Yamato Japanese numerals; see (149a). These forms are less commonly used in the Modern Japanese. The second series uses Sino-Japanese numerals; see (149b). In most cases, numeral classifier phrases contain the Sino-Japanese numerals instead of the Native Japanese ones, except for the default classifier *-tsu* 'CLS ENTITY', which still co-occurs with the Native Japanese numerals. The numeral "one" and "two" also preserve the Native Japanese forms when they are used with the human classifier *-ri* (e.g. *hito-ri* 'one-

$\text{CLS}_{\text{PERSON}} = \text{one person}'$, *huta-ri* 'two- $\text{CLS}_{\text{PERSON}}$ = two people').⁴⁴ When they appear in numeral classifier phrases, they must be realized as bound forms: *hito*, *huta*.

It should be noted that many languages do not make any morpho-phonological distinction between counting numerals and cardinals. As shown in (150), for example, English does not make such a distinction. Mandarin Chinese, on the other hand, distinguishes counting numerals and cardinals only in the number “two”. The numeral *er* ‘two’ is used in counting numbers, as shown in (150a), but not in the cardinal context (151b). As indicated, the cardinal *liang* must be used in the cardinal context; see (151b).

(150) English

- a. one, two, three, etc.
- b. one book, two books, three books, etc.

(151) Mandarin

- | | | | | | |
|----|--|--|-----------------------------------|--|--|
| a. | yi, {er *liang}, san, etc. | | | | |
| b. | [yi ben shu] / [{*er liang} ben shu] | | / [san ben shu], <i>et cetera</i> | | |
| | one CLS book / {two two} CLS book | | three CLS book | | |

What is noteworthy is that the counting form *er* appears in other numerical constructions in Mandarin. The distribution of the numeral *er* is summarized in (152).

(152)

	Multiplier	The right-hand cardinal in ACCs	Counting numbers	Attributive use	Ordinal
liang	*	*	*	OK	*
er	OK	OK	OK	*	OK

As shown in (152), *liang* is exclusively used for attributive use (i.e. in the cardinal context). The numeral *er*, on the other hand, can be used as a multiplier in MCCs (e.g. *er-bai ben shu*, two-hundred CLS book, ‘two hundred books’), ACCs (e.g. *si-shi er ben shu*, four-ten two-CLS book, ‘forty two books’), and ordinals (e.g. *di er-ge xuesheng*, ORD two-CLS student, ‘the second student’).

7.4 Collective numerals

Cardinal numerals can sometimes be used in combination with [+human] personal pronouns, as, for example, in English *the four of us*. In those nominal constructions, the numeral indicates the cardinality of the group referred to. Numerals that have this function are

⁴⁴ See also the discussion in Section 5.3 for the distribution of the classifiers *-tsu* and *-ri*.

sometimes called ‘collective numerals’.⁴⁵ In languages such as English and French, these collective numerals are morphologically bare:⁴⁶

- (153) a. *The four of us* went to the beach just after sunset.
b. Everyone but *us four* went crazy.

- (154) a. *Quatre d'entre nous* y sont allés.
four of-between us there were gone
'Four of us went there'
b. Pourrions-nous garder cela entre *nous quatre*?
could-we keep that between us four
'Can we keep this information to the four of us?'

In a language like Dutch, it is common to add the suffix *-en* to the numeral (Broekhuis and Den Dikken (2012:889)):

- (155) Ik heb [jullie vier/vier-en] niet herkend.
I have you four/four-*en* not recognized
'I didn't recognize the four of you.'

The numeral form derived by means of *-en* is also used in nominal constructions like (156):

- (156) We komen morgen met [z'n/ons vier-en].
we come tomorrow with his/our four-*en*
'We will come with four tomorrow.'

Booij (2005), taking a Construction Grammar perspective on the Dutch collective numeral constructions in (155) and (156), argues that these constructions are not formed in syntax but are stored as complex lexical units in the lexicon. Corver and Kranendonk (2009) propose an alternative analysis according to which these constructions are built in syntax. In the spirit of Kayne (2002, 2003), they argue that these constructions feature a silent noun, more specifically PERSON. The suffix *-en* is analyzed as the plural suffix that is typically found after overt nouns, as in *boek-en* (book-PL, ‘books’). In (155) and (156), the plural suffix is attached to the silent noun PERSON and expresses its plurality. Thus, although *-en* appears attached to the numeral at the sound surface, this plural suffix has a ‘regular’ distribution at the level of morphosyntax, i.e. it appears attached to a (silent) noun, as represented in (157a,b):

⁴⁵ The notion ‘collective numeral’ has also been used to refer to numerals measuring sets of entities, like English *pair*; see Corbett (2000), Ojeda (1997).

⁴⁶ These collective numerals are also possible with higher numerals:

- (i) He then marched *the twenty-five of us* two blocks down the road and made us stop in front of an office building. (Larry, Wade Livingston (2012), *Short stories from my heart*)
(ii) *The hundred of us* had been sent to live in a large warehouse in tight quarters with metal bunk beds, of three levels, same straw mattresses, [...] (Monique Sisich (2009), *A nightingale in the storm*)

- (157) a. [DP *jullie* [NumP vier [NP *PERSOON-en*]]]
 b. [PP met [DP *z'n* [NumP vier [NP *PERSOON-en*]]]]]

As shown in (158), it is also possible to have the diminutive suffix *-(t)je* as a component of the collective numeral constructions in (155) and (156):

- (158) a. *jullie vier-tje-s*
you_{PL} four-DIM-PL
 ‘you four / the four of you’

b. *met [z’n/ons vier-tje-s]*
 with his/us four-DIM-PL
 ‘with the four of us’

Following Wiltschko's (2005) suggestion that (Germanic) diminutives are classifier-like elements, one may assign the following structure to the nominal constructions in (158); see also Ott (2011):

- (159) a. [DP *jullie* [NumP *vier* [FP -tje [NP *PERSOON-s*]]]]
 b. [PP *met* [DP *z'n/ons* [NumP *vier* [FP -tje [NP *PERSOON-s*]]]]]]

It should, finally, be noted that the numeral component of the construction in (156) can correspond to an MCC or an ACC:

- (160) a. We komen morgen met [z'n *drie*^{honderd}-en]. MCC
 we come tomorrow with his three-hundred-PL
 ‘We will come with three hundred people tomorrow.’

b. We komen morgen met [z'n *vijf-en-twintig*-en]. ACC
 we come tomorrow with his five-and-twenty-PL
 ‘We will come with twenty five people tomorrow.’

These collective numeral constructions, which have not been investigated in great depth, deserve further cross-linguistic study.

8. Conclusion

Knowledge of number names is a subpart of our knowledge of language. As was shown in this chapter, this knowledge comprises various components, including: (i) knowledge of the categorial nature of simple numerals; (ii) knowledge of the computational system (e.g. morphological, syntactic) that underlies the formation of complex number names; (iii) knowledge of the structural positions that (components of) number names occupy; (iv) knowledge of the morpho-phonological or morpho-syntactic rules operative on numerals, including operations such as suppletion, contraction, and spell-out of (functional) heads in ellipsis contexts; (v) knowledge of the combinatorial behavior of classifiers and number names in so-called classifier languages; (vi) knowledge about the formation of ordinals on the basis of cardinals; and (vii) knowledge about the form of “counting” numerals.

The “morphological puzzles” that were addressed in this chapter can be summarized as follows: (i) What kind of lexical atoms are simple numerals: lexical or functional?; (ii) what kind of linguistic expressions are complex numerals: complex words (compounds) or complex phrases?; (iii) What is the derivational relationship between different kinds of

numerals (e.g. ‘cardinal’ in relation to ‘ordinal’, ‘cardinal’ in relation to ‘count numeral’)?; (iv) What is the interaction between classifiers and number expressions, and to what extent is the formation of complex numerals in classifier languages and non-classifier languages similar or different?

Clearly, there is a system of rules (i.e. a grammar) underlying the formation of number names in human language. Cross-linguistically, these rule systems display similarities (uniformity) but also differences (diversity). Dimensions of diversity are found both at the macro-level (e.g. classifier versus non-classifier languages) and at the micro-level (e.g. among Dutch varieties). Variation as regards the grammatical behavior of number names was even attested intra-linguistically: within one and the same language, number names can behave differently (e.g. low numerals versus high numerals, cardinal numerals versus count numerals). Various studies on the grammar of numerals have greatly contributed to our understanding of this part of our knowledge of language. It goes without saying that, for increasing our depth of theoretical understanding and broadening our cross-linguistic scope, we will need to continue crunching number names.

Related Articles (See Also)

Article ID
morphcom010
morphcom013
morphcom019
morphcom020
morphcom032
morphcom040
morphcom042
morphcom052
morphcom077
morphcom080
morphcom081

References

- Aboh, Enoch. 2004. Topic and Focus within D. *Linguistics in the Netherlands* 21: 1-12.
- Ackema, Peter and Ad Neeleman. 2004. *Beyond Morphology. Interface conditions on word formation*. Oxford: OUP.
- Ahn, Dorothy. 2019. Korean classifier-less number constructions. Uli Sauerland and Stephanie Solt (eds.), *Proceedings of Sinn und Bedeutung* 22. 23–38. ZAS, Berlin.
- Aikhenvald, Alexandra Y. 2003. Classifiers: A typology of noun categorization devices. London: Oxford University Press.
- Al-Bataineh, Hussein and Phil Branigan. 2020. *The syntax of (complex) numerals in Arabic*. Memorial University of Newfoundland. Ms.

- Alqarni, Muteb. 2015. *The morphosyntax of numeral-noun constructions in Modern Standard Arabic*. PhD dissertation, University of Florida.
- An, Duk-Ho. 2018. On The Word Order Of Numeral Quantifier Constructions. *Studia Linguist*, 72: 662-686. <https://doi.org/10.1111/stul.12084>
- Ashton, Ethel. O., E. M. K Mulira, E. G. M. Ndwula, and A. N. Tucker (1954). *A Luganda grammar*. London: Longmans, Green and Co.
- Asinari, Sarah. 2019. Case Percolation in Russian Numeral Constructions. *University of Pennsylvania Working Papers in Linguistics*, 25(1).
- Bailyn, John Frederick. 2004. The case of Q. In Olga Arnaudova, Wayles Browne, Maria Luisa Rivero & Danijela Stojanovic (eds.), *Proceedings of Annual Workshop on Formal Approaches to Slavic Linguistics* 12, 1–35. Ann Arbor, MI: Michigan Slavic Publications.
- Barbiers, Sjef. 1990. *Telwoorden, Adjektieven en Lege NPs*. Master's thesis, Leiden University.
- Barbiers, Sjef. 2007. Indefinite numerals one and many and the cause of ordinal suppletion. *Lingua* 117. 859–880.
- Bennis, Hans. 1986. *Gaps and Dummies*. Dordrecht: Foris.
- Bhatt, Christa. 1990. Die syntaktische Struktur der Nominalphrase im Deutschen. Tübingen: Narr.
- Blom, Alied. 1977. Het kwantitatieve 'er'. *Spektator* 6-7, 387-395.
- Booij, Geert. 2002. *The Morphology of Dutch*. Oxford: OUP.
- Booij, Geert. 2005. Compounding and derivation: evidence for Construction Morphology. In W. U. Dressler, D. Kastovsky, O. E. Pfeiffer, & F. Rainer (Eds.), *Morphology and its demarcations*. 109-132. Amsterdam: John Benjamins.
- Booij, Geert. 2010. *Construction Morphology*. Oxford: Oxford University Press.
- Booij, Geert, and Ton van der Wouden. 2020, May 14. Coordinative compounds. *Taalportaal*. Retrieved from <https://taalportaal.org/taalportaal/topic/pid/topic-13998813295676713>. (accessed 14 May 2021).
- Borer, Hagit. 2005. *In Name Only*. Oxford: Oxford University Press.
- Bošković, Željko. 2005. On the locality of LBE and the structure of NP. *Studia Linguistica* 59:1–45.
- Bresnan, Joan and Sam Mchombo. 1995. The lexical integrity principle: Evidence from Bantu. *Natural Language and Linguistic Theory* 13. 181–252. DOI: <https://doi.org/10.1007/BF00992782>

- Broekhuis, Hans and Marcel den Dikken. 2012. *Syntax of Dutch: Nouns and Noun Phrases*. Volume 2. Amsterdam: Amsterdam University Press.
- Borsley, R. D., M. Tallerman, and D. Willis. 2007. *The Syntax of Welsh*. Cambridge: Cambridge University Press.
- Bowern, C., and Zentz, J. 2012. Numeral Systems in Australian Languages. *Anthropological Linguistics* 54: 130–66.
- Buck, Carl Darling: 1949, *A Dictionary of Selected Synonyms in the Principal Indo-European Languages*, The University of Chicago Press, Chicago.
- Cardinaletti, Anna and Giuliana Giusti. 1992. Partitive *ne* and the QP-Hypothesis. A case study. *Proceedings of the XVII meeting of Generative Grammar*, Torino: Rosemberg & Sellier, 121-142.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- Cinque, Guglielmo. 2020. Word order variation: towards a restrictive theory. Ms., University of Venice.
- Comrie, Bernard. 2013. Numeral bases. In: *The world atlas of language structures online*, edited by Matthew S. Dryer, Martin Haspelmath, Leipzig: Max Planck Institute for Evolutionary Anthropology, <http://wals.info/chapter/131>.
- Comrie, Bernard (no year). Typology of numeral systems. Handout.
- Coppen, Peter-Arno. 1991. *Specifying the Noun Phrase*. Ph.D. Dissertation. University of Nijmegen.
- Corbett, Greville G. 2000. *Number*. Cambridge Textbooks in Linguistics. Cambridge University Press.
- Corbett, Greville G. 1978. Universals in the syntax of cardinal numerals. *Lingua*: 46:355–368.
- Corver, Norbert. 1990. The syntax of left branch extraction. PhD dissertation, Tilburg University.
- Corver, Norbert. 1992. On deriving left branch extraction asymmetries: A case study in parametric syntax. In *NELS 22: Proceedings of the North East Linguistic Society*, ed. by Kimberley Broderick, 67–84. Amherst: University of Massachusetts, Graduate Linguistic Student Association.
- Corver, Norbert. 2000. Degree adverbs as displaced predicates. *Italian Journal of Linguistics/Rivista di Linguistica* 12: 155–191.
- Corver, Norbert. 2001. On Predicate Numerals. In Ton van der Wouden and Hans Broekhuis (eds.). *Linguistics in the Netherlands*, 65-76. Amsterdam: John Benjamins.

- Corver, Norbert. 2010. Dressed numerals and the structure of universal numeric quantifiers. In J.-W. Zwart & M. de Vries (eds.), *Structure Preserved. Studies in Syntax for Jan Koster*. 91-99. Amsterdam: John Benjamins.
- Corver, Norbert. 2012. *Constructies en Cardinalia*. *Nederlandse taalkunde*, 17.2: 299-309.
- Corver, Norbert. 2014. Recursing in Dutch. *Natural Language and Linguistic Theory* 32(2): 423–57.
- Corver, Norbert and Huib Kranendonk. 2009. *Collective Numeral Constructions in Dutch: Remarkable Plurals, Regular Syntax and Silent Nouns*. In J. van Craenenbroeck & J. Rooryck (eds.), *Linguistic Variation Yearbook* 8: 233-268. Amsterdam: John Benjamins.
- Corver, Norbert and Joost Zwarts. 2004. *Prepositional numerals*. Paper presented at the Workshop on Numerals in the World's Languages, Leipzig, 30 March.
- Corver, Norbert and Joost Zwarts. 2006. Prepositional numerals. *Lingua* 116: 811-835.
- Csirmaz, Aniko & Eva Dekany. 2018. Numerals and quantifiers In: Gábor Alberti and Tibor Lazckó (ed.), *Syntax of Hungarian: nouns and noun phrases*, Volume 2. Amsterdam: Amsterdam University Press. 1044-1150.
- Danon, Gabi. 2009. Grammatical number in numeral-noun constructions. *CGG-19*, April, 1–3.
- Danon, Gabi. 2012. Two structures for numeral-noun constructions. *Lingua*, 122(12), 1282–1307. <https://doi.org/10.1016/j.lingua.2012.07.003>
- De Belder, Marijke. 2011. *Roots and affixes: Eliminating lexical categories from the syntax*. Doctoral dissertation, Universiteit Utrecht.
- Delsing, Lars-Olof. 1993. The internal structure of noun phrases. PhD disssertation, University of Lund.
- Di Sciullo, Anna Maria, and Edwin Williams. 1987. *On the definition of word*. Cambridge, MA: MIT Press.
- Doetjes, Jenny. 1997. *Quantifiers and selection*. PhD dissertation. Leiden University.
- Donohue, Mark. 2005. Numerals and their Position in Universal Grammar. *Journal of Universal Language* 6.1-37.
- Downing, Pamela. 1984. *Japanese Numeral Classifiers: a syntactic, semantic, and functional profile*. Ph.D. dissertation. University of California, Berkeley.
- Downing, Pamela. 1996. Numeral Classifier Systems: The Case of Japanese. Amsterdam & Philadelphia: John Benjamins.
- Drosdowski, Günter (ed.). 1984. *Duden Grammatik der deutschen Gegenwartssprache, Dudenverlag*, Mannheim/Wien/ Zürich.

- Dryer, Matthew. 1989. Large linguistic areas and language sampling. *Studies in Language* 13:257–292.
- Emonds, Joseph. 1976. *A Transformational Approach to English Syntax*. New York: Academic Press.
- Emonds, Joseph. 1985. *A Unified Theory of Syntactic Categories*. Dordrecht: Foris.
- Fassi Fehri, Abdelkader. 1999. Arabic modifying adjectives and DP structures. *Studia Linguistica* 53: 105–154.
- Fassi Fehri, Abdelkader. 2018. *Constructing Feminine to Mean: Gender, Number, Numeral, and Quantifier Extensions in Arabic*. Lanham: Rowman & Littlefield.
- Franks, Steven. 1994. Parametric properties of numeral phrases in Slavic. *Natural Language and Linguistic Theory* 12:597–674
- Genetti, Carol. 2007. *A grammar of Dolakha Newar*. Berlin: Mouton de Gruyter.
- Giusti, Giuliana. 1991. The Categorial Status of Quantified Nominals. *Linguistische Berichte* 136, 438–452.
- Greenberg, Joseph H. 1972. Numeral classifiers and substantival number: problems in the genesis of a linguistic type. In *Working papers on language universals*, 1–39.
- Greenberg, Joseph H. 1978. Generalizations about numeral systems. In *Universals of human language. Vol. 3. Word structure*, edited by Joseph H. Greenberg, Charles A. Ferguson, Edith A. Moravcsik. 249–95. Stanford: Stanford University Press.
- Greenberg, Joseph H. 2000. Numeral. In *Morphology: An international handbook on inflection and word-formation*, ed. Christian Lehmann Geert Booij and Joachim Mugdan. 770–783. Berlin: Mouton De Gruyter.
- Haas, Wim de and Mieke Trommelen. 1993. *Morfologisch Handboek van het Nederlands*. 's-Gravenhage: SDU Uitgevrij.
- Harrison, Sheldon. 1976. Mokilese reference grammar. Honolulu, HI: University Press of Hawaii.
- Hazout Ilan. 2000. Adjectival genitive constructions in Modern Hebrew: A case study in coanalysis. *The Linguistic Review* 17:29–52
- He, Chuansheng. 2015. Complex numerals in Mandarin Chinese are constituents. *Lingua* 164:189–215
- He, Chuansheng, One-Soon Her, Xiaoshi Hu, and Weijing Zhu. 2017. Overt Coordination in Additive Numerals of Minority Languages in South China. *Syntax* 20:3. 292–316.

- Her, One-Sooon. 2017. Deriving classifier word order typology, or Greenberg's Universal 20A and Universal 20. *Linguistics*, 55(2), 265–303.
- Her, One-Sooon, and Hui-Chin Tsai. 2020. Left is right, right is not: on the constituency of the classifier phrase in Chinese. *Language and Linguistics*. 21(1). 1-32.
- Hiraiwa, Ken. 2016. Numerals and their syntactic structures in Dagaare. Ms.
- Hurford, James R. 1975. *The linguistic theory of numerals*. Cambridge: Cambridge University Press.
- Hurford, James R. 1987. Language and number: The emergence of a cognitive system. Oxford: Blackwell.
- Hurford, J. R. 1998. The interaction between numerals and nouns. In F. Plank (Ed.), *Noun Phrase Structure in the Languages of Europe*. 561–620. Berlin: Mouton de Gruyter.
- Hurford, J. R. 2001. Languages treat 1-4 specially. *Mind & Language* 16(1), 69–75.
- Hurford, James. 2003. 'The interaction between numerals and nouns.' In *Noun Phrase Structure in the Languages of Europe: Typology of Languages in Europe*, ed. Frans Plank. 561-620. The Hague: Mouton de Gruyter.
- Ionin, Tanya, and Ora Matushansky. 2004a. A healthy twelve patients. Paper presented at GURT 2004.
- Ionin, Tanya, and Ora Matushansky. 2004b. The non-existence, syntax, and semantics of plurals. USC & CNRS/Universit Paris 8/ENS Ms.
- Ionin, Tanja & Ora Matushansky. 2006. The composition of complex cardinals. *Journal of Semantics* 23: 315-360.
- Ionin, Tania, and Ora Matushansky. 2018. *Cardinals: The syntax and semantics of cardinal-containing expressions*. Cambridge, MA: MIT Press.
- Jackendoff, Ray. 1977. *X' syntax: A study of phrase structure*. Cambridge: MIT Press.
- Jespersen, Otto. 1969. Analytic Syntax. New York: Holt, Rinehart and Winston.
- Jones, Robert. 1970. Classifier constructions in Southeast Asia. *Journal of the American Oriental Society* 90:1–12.
- Kane, Frances. 2014. The fine structure of the Irish NP. Doctoral dissertation, University of Ulster.
- van Katwijk, A. 1965. A grammar of Dutch number names. Foundations of language 1 (1): 51-58.
- Kayne, Richard. 2002. On Some Prepositions that look DP-internal: English *of* and French *de*. *Catalan Journal of Linguistics* 1:71-115.

- Kayne, Richard S. 2003. Silent years, silent hours. In *Grammar in focus*. Festschrift for Christer Platzack, ed. Lars-Olof Delsing, Gunlög Josefsson, Halldór Sigurdsson, and Cecilia Falk, volume 2, 209–226. Lund: Wallin and Dalholm.
- Kayne, Richard S. 2005a. A note on the syntax of quantity in English. In *Movement and silence*, 176–214. New York: Oxford University Press.
- Kayne, Richard 2005b. Some notes on comparative syntax, with special reference to English and French. In: G. Cinque and R. Kayne (eds.), *Handbook of Comparative Syntax*. New York: Oxford University Press.
- Kayne, Richard, 2010. A note on the syntax of numerical bases. In: *Comparisons and Contrasts*. Oxford University Press, Oxford, 57-72.
- Kayne, Richard S. 2016. Some thoughts on one and two and other numerals. Ms., New York University.
- Kayne, Richard. 2019. *Questions of Syntax*. Oxford University Press, Oxford.
- Keller, Kathryn C. 1955. The Chontal (Mayan) numeral system. *International Journal of American Linguistics* 21. 258–275. DOI: <https://doi.org/10.1086/464339>
- Klockman, Heidi. 2017. *The design of semi-lexicality. Evidence from Case and Agreement in the nominal domain*. PhD dissertation. Utrecht University.
- Klamer, Marian and František Kratochvíl (eds.). 2014. *Number and Quantity in East Nusantara*. Canberra: Pacific Linguistics.
- Kranendonk, Huib. 2010. Quantificational constructions in the nominal domain. Facets of Dutch microvariation. PhD dissertation. Utrecht university.
- Ktejik, Mish. 2013. Numeral Incorporation in Japanese Sign Language. *Sign Language Studies* 13: 186-209.
- Lapointe, Steven. 1980. A Theory of Grammatical Agreement. PhD dissertation, Amherst, University of Massachusetts.
- Lehmann, Christian. 2010. “On the function of numeral classifiers”. Floricic, Franck (ed.), *Essais de typologie et de linguistique générale. Mélanges offerts à Denis Creissels*. Lyon: École Normale Supérieure; 435-445.
- Leko, Nedžad. 1999. Functional Categories and the Structure of the DP in Bosnian. In Topics in *South Slavic Syntax and Semantics*, edited by Mila Dimitrova-Vulchanova and Lars Hellan, 229–252. Amsterdam: John Benjamins.
- Lichtenberk, Frantisek. 2008. *A Grammar of Toqabaqita*. Berlin & New York: Mouton de Gruyter.

- Lieber, Rochelle and Sergio Scalise. 2005. The Lexical Integrity Hypothesis in a New Theoretical Universe. In *Proceedings of the Fifth Mediterranean Morphology Meeting* (<http://mmm.lingue.unibo.it>). eds. G. Booij, L. Ducceschi, B. Fradin, E. Guevara, A. Ralli and S. Scalise. Università degli Studi di Bologna. 1-24.
- Little, Carol-Rose. 2018. Possessed numerals in Ch'ol. *Proceedings of the Tenth Conference on the Semantics of Under-Represented Languages in the Americas*, Kimberly Johnson and Alex Goebel (Eds), 95–116. Amherst, MA: GLSA.
- Löbel, Elisabeth 1989. Q as a functional category. In Bhatt, Christa, Löbel, Elisabeth and Schmidt, Claudia (eds.), *Syntactic Phrase Structure Phenomena*, 133-158. John Benjamins, Philadelphia.
- Makeeva, Nadezhda and Andrey Shluinsky. 2020. Numerals in Akebu. *Open linguistics* 6.1: 344-356.
- Maling, Joan. 1983. Transitive adjectives: A case of categorial reanalysis. In F. Heny & B. Richards (eds.) *Linguistic categories: Auxiliaries and related puzzles vol.1*, 253-289. Reidel.
- Marcin, Michalski. 2016. Agreement in Modern Standard Arabic constructions with cardinal numerals over ten. *Lingua Posnaniensis*, 58(1), 69–88. <https://doi.org/10.1515/lipo-2016-0005>
- Markey, Thomas. 1985. On suppletion. *Diachronica*, 2, 51–66.
- Matisoff, James. A. 1997. *Sino-Tibetan numeral systems: prefixes, protoforms and problems*. Pacific Linguistics series B-114. Canberra: Australian National University
- Matushansky, Ora. 2015. On Russian approximative inversion. In Gerhild Zybatow, Petr Biskup, Marcel Guhl, Claudia Hurtig, Olav Mueller-Reichau & Maria Yastrebova (eds.), *Slavic Grammar from a Formal Perspective. The 10th Anniversary FDSL Conference*, 303–316. Bern: Peter Lang.
- Mel'čuk, Igor. 1994. Suppletion. *Studies in language* 18: 2. 339-410.
- Miechowicz-Mathiasen, Katarzyna. 2012. Licensing Polish higher numerals: An account of the Accusative Hypothesis. *Current issues in generative linguistics* 81: 58-75.
- Ndimele, Ozo-mekuri, and S.L. Chan. 2016. *The Numeral Systems of Nigerian Languages*. M & J Grand Orbit Communications.
- Norris, M. J. 2018. The morphosyntax of number in Estonian numeral-noun constructions. *Proceedings of the Linguistic Society of America*, 3(1), 45-1–13. <https://doi.org/10.3765/plsa.v3i1.4337>
- Ojeda, Almerindo. 1997. A Semantics for Counting Numerals in Latin. *Journal of Semantics* 14: 143-171.

- Olsen, Susan 2001. Copulative compounds: a closer look at the interface between syntax and morphology. *Yearbook of Morphology* 2000: 279–320.
- Ott, Dennis. 2011. Diminutive-formation in German: Spelling out the classifier analysis. *Journal of Comparative Germanic Linguistics* 14: 1-46.
- Pankau, Andreas. 2018. The structure of approximative numerals in German. *Glossa: a journal of general linguistics* 3 (1). 1-48.
- Paterson, John. 1989. *Gaelic made easy: A Gaelic guide for beginners*. Glasgow: Dionnasg Gaidhlig na h-Alba (The Gaelic League of Scotland).
- Payne, John R. 1996. The syntax of Maltese cardinal numerals. *Rivista di Linguistica* 8 (1), 175-182.
- Perekhvalskaya, Elena and Vydrin, Valentin. 2019. Numeral systems in Mande languages. *Mandenkan* 61:47–111.
- Pereltsvaig, Asya. 2006. Passing by cardinals: In support of Head Movement in nominals. *Formal Approaches to Slavic Linguistics 14: The Princeton Meeting*. eds. J. Lavine, S. Franks & H. Filip, 277–292. Ann Arbor, MI: Michigan Slavic Publications.
- Plank, Frans. 2004. From local adpositions to approximative adnumerals, in German and wherever. *Studies in Language* 28, pp. 165-201.
- Pozdniakov, Konstantin. 2018. *The numeral system of proto-Niger-Congo. A step-by-step reconstruction*. Language Science Press.
- Press, Ian. 1986. *A grammar of Modern Breton*. Berlin: Mouton de Gruyter.
- Rappaport, Gilbert. 2002. Numeral phrases in Russian: A Minimalist approach. *Journal of Slavic Linguistics* 10, 329–342.
- Ritter, Elizabeth. 1991. Two functional categories in noun phrases: Evidence from Modern Hebrew. In *Perspectives on phrase structure: Heads and licensing*, ed. by Susan D. Rothstein, 37-62. San Diego, CA: Academic Press.
- Ritter, Elizabeth. 1992. Cross-linguistic evidence for number phrase. *Canadian Journal of Linguistics* 37. 197–218.
- Roeper, Thomas, and Siegel, Muffy E. A. 1978. A lexical transformation for verbal compounds. *Linguistic Inquiry* 9:197-260.
- Ross, John Robert. 1967. Constraints on variables in syntax. Doctoral dissertation, MIT, Cambridge, Mass.
- Rothstein, Susan. 2017. The semantics of counting and measuring. Cambridge: Cambridge University Press.

- Rutkowski, P., & Szczegot, K. 2001. On the Syntax of Functional Elements: Numerals, Pronouns, and Expressions Indicating Approximation. *Generative Linguistics in Poland: Syntax and Morphosyntax*, 187–196.
- Rutkowski, Paweł and Hanna Maliszewska. 2007. On prepositional phrases inside numeral expressions in Polish. *Lingua* 117.5: 784–813.
- Safar, J., O. Le Guen, G. Collí Collí, and M. Collí Hau. 2018. Numeral Variation in Yucatec Maya Sign Languages. *Sign Language Studies* 18: 488-516.
- Sagara, Keiko. 2016. Aspects of number and kinship terms in Japanese Sign Language. *Semantic Fields in Sign Languages*, Ulrike Zeshan and Keiko Sagara, eds., Berlin, Boston: De Gruyter Mouton. 301-332. <https://doi.org/10.1515/9781503429-009>
- Sassen, Albert. 1953. *Het Drents van Ruinen*. Assen: Van Gorcum.
- Selkirk, Elisabeth. 1977. Some Remarks on Noun Phrase Structure. In P. W. Culicover, Wasow and A. Akmajian, eds., *Formal Syntax*. London: Academic Press. 285-316.
- Selkirk, Elizabeth. 1982. The Syntax of Words. Cambridge, MA: MIT Press.
- Shlonsky, Ur. 2004. The form of Semitic nominals. *Lingua* 114.12, 1465-1526.
- Siloni, Tal. 2002. Adjectival constructs and inalienable constructions. In *Themes in Arabic and Hebrew Syntax*, Jamal Ouhalla and Ur Shlonsky (eds.), 161–187. Dordrecht: Kluwer Academic Publishers.
- Sleeman, Ruby. 2017. Ordinal numerals in dialects of Dutch. MA Thesis, Leiden university.
- Stampe, David. 1976. Cardinal number systems. In *Chicago Linguistic Society 12.*, ed. Mila Tasseva-Kurtchieva James Lavine, Steven Franks and Hana Filip, 594–609. Ann Arbor, MI: Michigan Slavic Publications.
- Stoltz, Thomas, and Lyuba N. Veselinova. 2005. Ordinal numerals. In The world atlas of language structures, ed. David Gil Martin Haspelmath, Matthew S. Dryer and Bernard Comrie, 218–221. Oxford: Oxford University Press.
- Stump, Gregory. 2010. The derivation of compound ordinal numerals: Implications for morphological theory. *Word Structure* 3.2:205–233.
- Taeldeman, Johan. 1984. De morfonologie van hoofdtelwoorden in de zuidwestelijke dialecten. *Taal en Tongval* 36:50-55.
- Tatsumi Yuta. 2021. *Linguistic Realization of Measuring and Counting in the Nominal Domain: A Cross-linguistic Study of Syntactic and Semantic Variations*. PhD dissertation. University of Connecticut.
- Thackston, W . M . 2000. *An Introduction to Koranic and Classical Arabic*. Bethesda, Maryland: Ibex Publishers.

- Veselinova, Ljuba. 1997. Suppletion in the derivation of ordinal numerals: a case study. In Proceedings of the 8th Student Conference in Linguistics, ed. Benjamin Bruening, 429–447. Cambridge, MA: MIT Working Papers in Linguistics.
- Watanabe, Akira. 2010. Vague quantity, numerals, and natural numbers. *Syntax* 13:37–77.
- Watanabe, Akira. 2014. 1-deletion: Measure nouns vs. classifiers. In *Proceedings of Japanese/Korean Linguistics* 22, ed. Naonori Nagaya Akiko Takemura Mikio Giriko, Kyoko Kanzaki and Timothy J. Vance, 245–260. Stanford, CA: CSLI Publications.
- Wege, B., 1997. Skalare Präpositionen – grammatisierte Varianten. In: Haumann, D., Schierholz, S.J. (Eds.), Lexikalische und grammatische Eigenschaften präpositionaler Elemente. Max Niemeyer Verlag, Tübingen, pp. 19-45.
- Williams, Edwin. 1982. Another argument that passive is transformational. *Linguistic Inquiry* 13: 160–63.
- Wiltschko, Martina. 2007. Why should Diminutives count. In *Organizing grammar: studies in honor of Henk van Riemsdijk*, eds. Hans Broekhuis, Norbert Corver, Riny Huybregts, Ursula Kleinherz, and Jan Koster, 669–678. Mouton de Gruyter.
- Witkoś, J., & Dziubała-Szrejbrowska, D. 2018. Cardinal numerals and complex numerals as specifiers. *Poznań Studies in Contemporary Linguistics*, 54(4), 605–635.
<https://doi.org/10.1515/pscl-2018-0024>
- Wouden, Ton van der. 2020, May 14. Cardinal numerals. *Taalportaal*. Retrieved from <https://www.taalportaal.org/taalportaal/topic/pid/topic-13998813299109379>. (accessed 27 March 2021).
- Wouden, Ton van der. 2020, May 14. -er (nominal). *Taalportaal*. Retrieved from <https://www.taalportaal.org/taalportaal/topic/pid/topic-14093156822693140>. (accessed 26 March 2021).
- Zabbal, Youri. 2005. The syntax of numeral expressions. MS, University of Massachusetts at Amherst.
- Zamparelli, Roberto. 1995. *Layers in the Determiner Phrase*. Ph.D. dissertation, University of Rochester.
- Zamparelli, Roberto, 2004. Every two days. *Snippets* 9, 19–20.
- Zerbian, Sabine, and Manfred Krifka. 2008. Quantification across Bantu Languages. In *Quantification: A Cross Linguistic Perspective*, edited by Lisa Matthewson, 383-414. North Holland Linguistic Series, Linguistic Variations 64. Bingley, U.K.: Emerald.
- Zeshan, Ulrike, Cesar Ernesto Escobedo Delgado, Hasan Dikyuva, Sibaji Panda & Connie de Vos. 2013. Cardinal numerals in rural sign languages: Approaching cross-modal typology. *Linguistic Typology* 17. 357–396.

Zeshan, Ulrike and Sagara, Keiko. 2016. *Semantic Fields in Sign Languages*, Berlin, Boston: De Gruyter Mouton. <https://doi.org/10.1515/9781501503429>

Zweig, Eytan. 2005. Nouns and adjectives in numeral NPs. In *Proceedings of the 35th the North Eastern Linguistics Society Annual Meeting*, ed. Angelo J. Rodríguez Vineeta Chand, Ann Kelleher and Benjamin Schmeiser. 663–675. Amherst, MA: GLSA Publications.