

Utrecht, 17-18 November 2006

**Canonical and non-canonical order in noun phrases
(and where is information structure?)**

Frans Plank

frans.plank@uni-konstanz.de

0. Preliminaries

Main kinds of functions for syntactic forms:

- referring (labeling, describing); • predicating; • referent-modifying; • predicate-modifying;
- propositioning; • proposition-modifying;
- relatedness-indicating; • relationship-identifying (of forms-in-constructions);
- illocutionary-force-giving, • speech-act-anchoring; • context-anchoring (of propositions).

How to express information-structure?

- special (dedicated?) forms, segmental and/or prosodic;
including typically attenuated form (pro, zero) for given/topic vis-à-vis new/focus;
also including form alterations, such as re-ordering.
- special (dedicated?) constructions.

Domains for information-structuring (given – new, topic – focus, ...):

- It minimally needs **two** lexical/content parts, whatever the attendant grammatical(ised)/function machinery.
- Hence, simple referring expressions (NPs) and simple predicating expressions (VPs) aren't domains for information-structuring in this sense.
- They aren't speech-act or context-anchored either, and don't have illocutionary force either, as such, but only by virtue of being parts of propositions. (No matter what overt markers happen to be associated with them, e.g. tense with NPs.)

1. In what terms are noun phrases internally-ordered (and re-ordered)?

(relative to phrase edges, relative to co-constituents, special designated positions)

suggested universal generalisations, in terms of constituent classes and/or functions

Greenberg #18

When the descriptive adjective precedes the noun, the demonstrative and the numeral, with overwhelmingly more than chance frequency, do likewise.

Greenberg #20

When any or all of the items (demonstrative, numeral, and descriptive adjective) precede the noun, they are always found in that order. If they follow, the order is either the same or its exact opposite.

i.e.	Dem Num Adj N	these five large houses (English et al.)
	N Adj Num Dem	houses large five these (favourite when N is initial)
	N Dem Num Adj	houses these five large (less popular, e.g. Kikuyu)

revised by Hawkins 1983: 119-120

... For those that follow, no predictions are made, though the most frequent order is the mirror image of the order for preceding modifiers. In no case does the adjective precede the head when the demonstrative or numeral follow.

to accommodate	N Adj Dem Num
	N Dem Adj Num
	N Num Adj Dem

as attested in Bantu, Nilotic, Cushitic (see further THE UNIVERSALS ARCHIVE, #57, #59)

terms of reference, more generally:

- functions: The more grammatical, the closer to the phrase edge; The more lexical, the closer to the N (the lexical core)
(grammatical: determiners, classifiers (?); lexical: quantifiers (?), modifiers, head)
but when after N, also other orderings!

Dem Num Adj N
N Adj Num Dem
!N Dem Num Adj
!N Adj Dem Num
!N Dem Adj Num
!N Num Adj Dem

- functional harmony: Dependents (whether modifiers, determiners, ...) uniformly before/after heads, in whatever construction
(or Branching Direction, Head Proximity: phrasal categories always on the same side of non-phrasal categories)
- meaning of constituents: The nearer the meaning, the closer to the noun
(evaluation, size/shape, weight, age, colour, material, provenance, purpose ... thing);
but when after N, also anti-iconic ordering!
- scope: Scope-taking elements outside (before/after) elements in their scope
- inherent emphasis/focus: Emphatic (BIG, SO Adj, HOW Adj, ...) earlier than otherwise
- complexity of constituents: Less complex before more complex
- integrity (continuity) of phrasal constituents: Discontinuity to be avoided
(or Early Immediate Constituents, Domain Integrity)
- constituent classes: adjective (phrase), adpositional phrase, adverb, numeral (phrase), article ...: only arbitrary, language-particular regulations, or derivative of other general principles?
e.g., German Modifier after N when an Adv: *die Sitten hier* 'the customs here'

die hiesigen Sitten 'the local customs'

follows from general principles about distribution of inflection:

constituents in between determiner and noun need to agree (gender, number, case, as determined by noun and the whole NP) and to weakly/strongly-inflect (as determined by the determiner).

e.g., German (and other) Modifier after N when an Adpositional Phrase or a Clause:

follows from Less Complex before More Complex.

- And **information structure**??? Given – New info, Topic – Focus, ...
well, it matters in complex NPs, e.g., English genitive (determiner) vs. adpositional phrase (modifier): *the return of the native* vs. *the native's return* [native the by now established topic].

But what with simple NPs with two lexical constituents, i.e., Adj and N, rigidly ordered in English?

Noël [topic of paragraph] *She ... had very little leisure, and in the evenings would often fall asleep in the drawing-room curled up in a large chintz-covered chair. ... That evening after dinner she did not go to sleep in the chair [N old info], but sat under the open window ...*

When you spend a summer holiday in the Horton country you must not fail to make the ascent of Horton Hill. It is an easy climb [N old info] and there is a wonderful view.

Attributive adjectives old info when epithets: e.g., *Henry, my old friend ... my old friend.*

2. Manifest order \neq abstract order: Relative proximity of attributive adjectives of different meaning to their noun, pre- and postnominally

2.1. Concrete representations

- forms-in-constructions, at complexity levels of stems, (syn/phon) words, (syn/phon) phrases, (syn/phon) clauses, sentences, paragraphs, texts;
- (sub-)classes of forms, (sub-)classes of constructions:
form classes, distribution/position classes, meaning classes;
- relation of (immediate) precedence (i.e., linear order);
- relation of (immediate) constituency (i.e., part-whole);
- syntagmatic relationships between manifest constituents of constructions,
not necessarily specific to particular classes of constructions:
 - subjunction, or dependency (head – dependent),
with subtypes: modification, determination, complementation,
 attribution, predication, apposition, ...
 predicate-argument, circumstance
 quantification, classification
 - conjunction,
 - adjunction, ...
- several kinds of influence (preferably local, possibly also at a distance) among manifest constituents of constructions, not specific to particular classes of constructions:
 - exclusion of co-presence,
 - requirement of co-presence,
 - license of co-presence,
 - government, command,
 - referential binding,
 - scope-taking, ...
- paradigmatic relationships;
- kinds of constructional marking:
 - relationship-identifying,
 - relatedness-indicating,
 - linking, separating,
 - speech-act and context anchoring;
- typically concrete MEANING/FUNCTION (“functions” of forms-in-constructions):
reference, reference-modification, predication, predicate-modification;

topic, comment, focus;

assertion, question, command, exclamation, ...;

- typically concrete EXPLANATION
 - expressiveness [allowing/facilitating the expression of thought];
 - simplicity (formal economy, markedness) [facilitating storage, production, processing];
 - clarity (ambiguity avoidance) [facilitating processing];
 - iconicity (function motivating form) [facilitating what? giving pleasure?];
 - inertia (Don't reanalyse! or, Only reanalyse within reason/limits!);
 - extravagance (Be different!);

2.2. How can (syntactic, morphological, phonological) representations be **abstract**?

- first, by virtue of excluding all sorts of things present in the speech signal, but considered irrelevant for grammar;
- second, by virtue of including some sorts of things not present in the speech signal, but considered relevant for grammar.

Abstractness in the “inclusive” sense can involve:

- structurally relevant PRECEDENCE of parts of wholes differing from manifest precedence (movement/displacement);
- structurally relevant CONSTITUENCY differing from manifest constituency (restructuring);
- structurally SOMETHING represented by manifestly NOTHING, with subtypes of manifest nothingness: phonological (pro, PRO, trace), syntactic (ellipsis);
- structurally NOTHING represented by manifestly SOMETHING (including underspecification). [Or is this “exclusive”-type abstractness?]

2.3. A case of manifest \neq structural order in NPs: The facts, summarized:

English et al.

<i>a beautiful big red ball</i>	VALUE SIZE COLOUR N	(1a)
---------------------------------	---------------------	------

Bahasa Indonesia et al.

<i>bola merah besar tjantik</i>	N COLOUR SIZE VALUE	(1b)
---------------------------------	---------------------	------

Maltese et al. (incl. Celtic)

<i>ballun sabiħ kbir aħmar</i>	N VALUE SIZE COLOUR	(2a)
--------------------------------	---------------------	------

Italian et al.

<i>una bella grande palla rossa</i>	VALUE SIZE N COLOUR	(2a')
-------------------------------------	---------------------	-------

unattested (or so it seems)

<i>redd bigg beautifull ball</i>	COLOUR SIZE VALUE N	(2b)
----------------------------------	---------------------	------

In establishing such unmarked orderings all kinds of considerations have to be borne in mind:

- make sure the language permits the stacking of property-concept words to begin with (some/many don't, or only do so very reluctantly);
- the several property-concept words shouldn't be (asyndetically) coordinated (the only option in non-stacking languages);
- property-concept words should be modifying the noun rather than one another (*a beautifully red ball*);
- there should be no intonation breaks between the property-concept words suggesting non-cohesiveness of the modifier construction;
- sometimes inherently emphatic property-concept words (especially 'big') push forward;
- there may be further factors relevant for ordering, such as the length or other phonological properties of property-concept words;
- further semantic classes of property concepts should be taken into consideration, such as weight, age, material, provenance, and purpose, increasingly closer to the noun (*a beautiful big heavy new red woollen Swiss medicine ball*).

2.4. Descriptive generalisation over (1a/b), the crosslinguistically most common orderings:
The relative distance (position class) of semantic classes of property-concept words from the noun is the same, whether the noun comes after or before property-concept words. (That is, the two most common orderings are mirror images of each other.)

This generalisation is richly supported in the relevant literature, most substantially by Hetzron 1978 and Sproat & Shih 1991 (uncharitably or sloppily failing to acknowledge or even mention Hetzron).

2.5. Higher-level generalisation, perhaps explanation of (1a/b):

The nounier a modifier, the closer to the noun.

The nouniness ranking COLOUR > SIZE > VALUE is independently motivated, on language-particular and general grounds. Relevant evidence includes: the nounier modifier words actually are themselves nouns, or are derived from nouns; they can enter a morphological relation with head nouns (compounding); their range of applicability to nouns of different semantic classes is narrower. The less nouny modifier words are verbal forms (participles) or are derived from verbs; they don't compound; their range of application is wider. In the smallish adjective inventories of languages which prefer to express property concepts through nouns and/or verbs, specifically adjectival lexicalisation of property concepts also follows that scale.

2.6. Even higher-level generalisation/explanation, in terms of ICONICITY:

Linear distance mirrors conceptual distance.

2.7. It follows from this account that (2b) should not occur (and it apparently doesn't).

But neither should (2a), which does, even if not so frequently!

Considering the existence of (2a), the obvious question is:

WHY is there no mirror image of (2a), i.e., (2b)?

The answer:

Maltese et al. are really, at a relevant level of syntactic representation where conceptual proximity matters (which is abstract), like English et al., i.e., their NPs are N-final, and N is only fronted in manifest form (half-way fronted in Romance).

That is, iconicity rules ok everywhere!

But for this explanatory account to be possible, syntactic representations need to be assumed which are abstract, in the sense of manifest order \neq structural order.

2.8. Can such abstract syntactic representations be indirectly inferred by a learner from any concrete forms and relations between forms?

In Maltese (possibly like in other relevant languages with abstract precedence differing from concrete precedence), a few adjectives do precede their nouns, but only in highly marked formal and stereotyped constructions:

il-kbir Alla

DEF-big god

‘God Almighty’

l-gharef hu-k

DEF-wise brother-2SG

‘your clever brother’ (only ironic)

Synthetic comparatives, themselves postnominal, move ahead of their noun to express the superlative grade (with one definiteness marker suppressed):

il-belt il-qawwi-ja

DEF-city DEF-beautiful-FEM

‘the beautiful city’

belt aqwa

city beautiful.COMP

‘a more beautiful city’

l-aqwa belt

DEF-beautiful.COMP city

‘the most beautiful city’

Analytic comparatives/superlatives as such stay behind; only the comparative marker on its own may move ahead (again, with one definiteness marker dropped):

il-ktieb l-aktar sabiħ

DEF-book DEF-more beautiful ‘the more/most beautiful book’

l-aktar ktieb sabiħ

DEF-more book beautiful ‘the most beautiful book’

And, being in some ways similar to adjectival modifiers, numerals other than ‘one’ also precede nouns (which are in the singular with numerals above ‘ten’):

għoxrin suldat

twenty soldier (SG) ‘twenty soldiers’

Whatever the right “thick description” (analysis), it seems to me doubtful that the manifest ordering of nouns in very marginal modifier constructions could suffice to inspire learners to set up abstract representations at odds with manifest ordering in ordinary constructions encountered much more frequently.

[To mention an analogous case, abstract uniformly verb-final representations, thus also for main clauses, in Germanic are perhaps easier to infer for learners, on the assumption that verb-final patterns are unmarked, lacking all sorts of things that make a clause “main”.]

2.9. But what other inspiration could there be?

Answer: Universal Grammar prohibiting representations of modifier constructions at odds with iconicity, thereby forcing abstract representations on learners of languages where manifest ordering, implemented through N-fronting, is at odds with iconicity.

2.10. The remaining questions are (i) why iconic representations can be interfered with through N-fronting, and (ii) why there are no abstract representations in line with iconicity ending up with counter-iconic concrete precedence relations through N-backing — deriving manifest (2b) from abstract (1b). Ruled out by UG?

A lead that remains to be explored is that languages which have N-fronting (including Semitic and Celtic, but not in fact Maltese) seem to be ones where V is initial too, at least in abstract representations. If there really is a connection between the construction types of noun phrases and of clauses, how can this be made sense of? Would seem to require abstract representations where noun phrases and clauses are seen to share constructional principles that are perhaps impossible to express in concrete representations. (Inflectional differences in the marking for gender and number have also been claimed to be implicated; but this seems more dubious.)

2.11. Are such ordering divergences between manifest and abstract representations of noun phrases to do with information structure? (N-fronting, but no N-backing)

[Compare the case of abstractly V-final but manifestly V2 clauses; relevant fronting mechanism: cliticization of AUX, generalised to (generally light) V]

3. Scope and the left edge: Indefinite plural noun phrases with an adjective before a cardinal numeral, in English and elsewhere in Germanic

3.1. The puzzling three questions

- (i) What is the morphosyntactic form of the construction(s), illustrated in (1), with an adjective before a cardinal numeral and a singular indefinite article in referentially and morphologically plural NPs, corresponding to the construction(s), illustrated in (2), with an adjective and a cardinal numeral in the canonical reverse order and without an indefinite article?
- (ii) What is its meaning, or — in case (1a), (1b), (1c) differ — what are their meanings?
- (iii) How come English has such a construction(s)? Compare the German counterparts, (3)–(4), differing in the obligatory absence of the indefinite article when the adjective non-canonically precedes the cardinal numeral.

- (1)
 - a. *a sad two days*
 - b. *a remarkable ninety-nine performances*
 - c. *a mere ten millimeters*
- (2)
 - a. *two sad days*
 - b. *ninety-nine remarkable performances*
 - c. *ten mere millimeters*
- (3)
 - a. *(*eine) traurige zwei Tage*
 - b. *(*eine) bemerkenswerte neunundneunzig Vorstellungen*
 - c. *(*eine) bloße zehn Millimeter*
- (4)
 - a. *zwei traurige Tage*
 - b. *neunundneunzig bemerkenswerte Vorstellungen*
 - c. *zehn bloße Millimeter*

3.2. Recent answers to the first question

- Jackendoff (1977: 126-136), concerning (1a)
 - (i) Cardinal numerals in English are — underlyingly — nouns, in particular “group nouns”, which like other group (hence non-mass) nouns (such as *number*, *group*, *bunch*, *score*) (a) require an article, (b) are grammatically singular (though notionally plural),

and (c) occur in “pseudopartitive” constructions:

- (5) a. *a number of days*
 b. *a hundred of days*
 c. *a two of days*
- (ii) At the surface, cardinal numerals want to seem quantifiers; thus, unlike all other group nouns, all cardinal-numeral nouns require the suppression of *of* before their pseudopartitive dependent noun phrase; and unlike the “seminumeral” nouns *hundred*, *thousand*, *million*, *billion*, *dozen*, *few*, *little*, and (for some speakers) *couple*, all other cardinal-numeral nouns also require the suppression of the indefinite article:
- (6) a. *a number of days*
 b. *a hundred ~~of~~ days*
 c. *~~a~~ two ~~of~~ days*
- (iii) However, the INdefinite article, not its definite counterpart, is obligatorily suppressed with these “pleni”-numeral cardinal nouns — except when they are followed by an adjective rather than a noun, in which case they are obligatorily retained:
- (7) a. *a sad number of days*
 b. *a sad hundred ~~of~~ days*
 c. *a sad two ~~of~~ days*
- NB: Cardinal-numeral nouns, being nouns, can be modified by attributive adjectives; it is not the pseudo-partitive nouns that are modified here.
- (iv) It is the cardinal-numeral nouns (taking singular articles) which are the heads of the entire noun phrases, with the pseudopartitive phrases as the dependents: hence external agreement would be expected to be singular — or if plural on semantic grounds, no differences would be expected depending on the presence or absence of an adjective before the cardinal. However, when no adjective is present before the cardinal, and the indefinite article is therefore suppressed, verb and other agreement is normally plural (9a); when there is an adjective, rescuing the indefinite article, both plural and singular agreements are equally possible for most speakers (9b).

(8) *[the/a [(sad) number/hundred/two [of [days]]]]*

- (9) a. *Two sad days lie/*lies ahead: Good Friday and, a little later, All Fool’s Day. These/*This are/*is always two sad days.*
 b. *A sad two days lie/lies ahead: Maundy Thursday and Good Friday. These/This are/is always a sad two days.*

- Ionin & Matushanski (2004), concerning (1a) and (1b)
 - (i) Cardinal numerals are singular nouns, hence can be modified by an adjective and be determined by a singular indefinite article. (Like Jackendoff 1977.)
 - (ii) They are of semantic type, $\langle\langle e, t \rangle, \langle e, t \rangle\rangle$, which is to account for their modificational function; they combine with count nouns (type $\langle e, t \rangle$), but syntactically they are the head of the construction, so that it is the numeral that a determiner and modifiers are most directly associated with:

(10) *[a remarkable [ninety-nine [performances]]]*
 - (iii) The obligatory presence of a (singular) determiner when an adjective precedes a numeral is claimed to instantiate some general tendency for adjectival modifiers to specially encourage determiners where they would not be used otherwise, as variously seen in various languages — for example with French predicate nominals or with English quotation:

*Elle est (*une) physicienne* ‘She is (*a) doctor’,
 but *Elle est *(une) physicienne sportive* ‘She is *(a) sporty doctor’;
*He responded with (*a), “How do you do?”*,
 but *He responded with *(an) absent, “How do you do?”*
- Questions not addressed:
 - (ii) What’s the meaning(s) of (1), as opposed to (2)?
 (Implicit answer: Should be the same.)
 - (iii) How does English compare to other languages in relevant respects?
 (Implicit answer: Should depend on the underlying and surface categorisation of cardinal numerals and the concomitant morphosyntax.)

3.3. Scope matters (concerning question (ii))

- **NUMBER-MODIFICATION** (1b, 1c), as opposed to **noun-modification** (2b, 2c)

(11) *a remarkable ninety-nine utterly unremarkable performances*
 — not contradictory

(1’) b. *a remarkable {[NUMBER of] ninety-nine} performances*
 — noun itself is not in the scope of the modifier (scope indicated by { })
- friends of ellipsis, watch the article, which is definite in the non-elliptic version:

“On All Fours” ran for the remarkable [number of] 99 performances.
- adjectives suitable for number-modification: evaluative, measuring, ...
 adjectives unsuitable for number-modification: colour (**a red five pencils*), ...

- Is there a strict word-class distinction between modifiers of nouns (primarily adjectives) and of verbs and adjectives (adverbs)?

No: bare adjectives are often possible as modifiers of

- verbs: *She ran fast(*-ly)*
- other adjectives: *the dark red flowers, a bitter cold wind, a bloody difficult question, a real nice guy, a good long time, a mighty good obedient thing, an extraordinary bright pupil, an exceeding poor household, a whole different matter*, not to mention productive “compounding” of an adjective and a participle: *easy-going, green-coated*

- (12) a. *a good ten hours/washes* (‘not less than’, ‘complete, thorough’)
 b. *ten good hours/washes* (ditto, but with scope over nouns)

- (13) a. *merely ten millimeters*
 b. *well(-nigh) ten hours/washes*

German equivalents, with adverbialisation showing through non-inflection:

- a. *bloß zehn Millimeter*
 b. *gut zehn Stunden/Wäschen*

- **UNIT-MODIFICATION** (1a), as opposed to **noun-modification** (2a)

- (1') a. *a sad {[RUN of] two days}*
 — noun, together with numeral, is in the scope of the modifier, unlike with number-modification

- (14) *a sad two happy days*
 — contradictory on the most straightforward reading of the adjectives

- adjectives suitable for unit-modification: evaluative, ...
 adjectives unsuitable for unit-modification: colour (**a red ten roses*), open (**an open three questions*; compare *a puzzling three questions*), ...

- ambiguous between **UNIT-MODIFICATION** (a) and **NUMBER-MODIFICATION** (b)

- (15) *He scored a decisive two goals.*
 a. at the score of three-all, he put two further goals past the keeper, and the (new) rules are that you need to lead by two goals to win;
 b. he scored two goals in a 4-3 win (old rules: one-goal lead suffices), the first and the third.

3.4. The singular-plural indefinite puzzle, in a historical light (concerning question (i))

- Is the introductory *a* in (1) really the indefinite singular article?
 - degree adjective phrases can skip ahead of it, just like they can skip ahead of the indefinite article:
- (16) a. *More intelligent a student I've never seen.*
 b. *More arduous a ten miles I've never walked.*
 — but: any grammatical marker introducing an indefinite NP would equally be skipped, such as cardinal numerals:
- (17) *?More intelligent two/a coupla students I've never seen.*
- mutual exclusion between *a(n)* and other determiners in the constructions at issue:
- (18) a. *an insane ten days*
 b. *this (*an) insane ten days* (an attested example from Jespersen 1949: II, 112)
- sandhi alternation: *a* before consonant, *an* before vowel; but not emphatic [eɪ]
- history of indefinite article *a(n)*
 - cardinal numeral 'one' > indefinite pronoun/article;
 — clearly, the introductory word in (1) is not the cardinal numeral.
 - Other semantic extensions: oneness (> collectivity) (> numerical approximation)
- (19) a. *for an-um xii niht-um huhugu swa*
 before one-DAT.PL twelf night-DAT.PL however so
ante dies fere duodecim (Latin source)
 'before nearly twelve [successive] days'
- b. *wel ny an eighte busshel-s*
 well nigh one eight bushel-PL
 '[a collection of] some eight bushels'
- cf. (among others) Dutch *for* (not ONE-based) collective > approximative:
- (21) *een honderd-tal kippen*
 a hundred-number chickens
 'some hundred chickens'
- loss of agreement inflection of 'one' and its derivatives (INDEF, COLLECT, APPROX), with the former agreement categories having been gender, case, and number;

nouns in such constructions frequently in basic singular form:

- (20) *an vif 3er after þis*
 one five year after this
 ‘[a period of] some five years after this’

- *a(n)*, derivative of the indefinite article and/or the numeral ‘one’, was to lose its collectivising force and to become increasingly marginal in approximative function in Modern English (as a remnant, compare *a hundred chickens*, which is presumably approximative (as well as collective), with *(one) hundred chickens*, which is precise (and need not be collective)).

— Thus: introductory *a(n)* in (1) not to be identified SYNCHRONICALLY with the ONE-based collective or approximative marker of old.

- Still, while *a(n)* was still common as an approximative marker, formally indistinguishable from the indefinite article, it began to be combined with the inherently approximative paucal and multal numerals *few* (13th century) and *many* (17-19th centuries) — both of an originally adjectival nature, thereby differing from many other quantifying words of this kind which derive from nouns, such as ‘pair’ or ‘couple’:

- (22) *a few years; a many years*

- contrast between the presence and absence of *a(n)* was probably one of emphasis on approximateness;
- semantic contrast between *few* ‘not many’ and *a few* ‘not none at all’ appears to have been a secondary development;
- the multal numeral also acquired a singular construction, with *a(n)* and numeral in reverse order: *many a needy Greek scholar*.
- Certain evaluational adjectives then began to be added, most commonly *good*, which on their own could also subserve the purpose of numerical approximation — with the presence of such an adjective now necessitating the presence also of *a(n)*, which had only been an optional component without an adjective:

- (23) *a good few years; a good many years;*
a feeble few, a very slender few, a select and dainty few, a single poor few,
a good very few rude things

- Also attested, though rarely: parallel instances with a similar mid-range quantifying word: *a good several*

- On the model of these paucal and multal numerals (23), and for an initial while presumably also encouraged by collective-approximative *a(n)* with precise numerals (19), *a(n)*, followed by an evaluational adjective could also be used with precise numerals, with a preference, at least initially, for round numbers:
- (24) *a good twenty years*
- and with non-round numerals coming to be permitted, too;
 - also adjectives other than those used for approximation, though still all essentially evaluational;
 - thus approaching the modern constructions of type (1), with some indeterminacy in respect of the identity and constituency of *a(n)*:
 - approximative (cum collectivising) marker, forming a syntactic co-constituent with a numeral and eventually perhaps entering into an even closer morphological association with a couple of numerals (*a-few*, *a-many*) or the standard approximative adjective (*a-good*), or
 - singular indefinite article, forming a co-constituent with the entire rest of the NP.
 - Remember that *a(n)*, in whatever uses, had ceased to inflect during the 12-13th centuries: an invariable word, no longer carrying singular inflections, will have been felt to be less of a misfit in plural NPs.
 - plural morphology on nouns in general negotiable?
(if so, a mitigating circumstance for the clash between a plural noun and *a(n)*)
 - plural inflection on nouns now virtually sacrosanct, unlike earlier (20);
 - cardinal numerals on their own continued to be able to be nominalised and thus be made to refer to collectives of the precise cardinality given: *another college eight* [i.e., crew of a rowing-boat]; *The two elevens* [i.e., cricket teams] *are now retiring for tea*; *We reckon by tens*; *Those many hundreds of victims have not been forgotten*;
 - especially: semantics and agreements of NPs with numerals above ‘one’ could increasingly be assimilated to those of singulars through collectivisation (called “unification” by Jespersen, “respecification” in the *Cambridge Grammar*), with a plurality of referents conceived of as a unity:
- (25) a. *There is that ten guineas you were sending to the poor gentleman.*
 b. *What is sixe winters, they [PL!] are quickly gone?*

c. *I could not count upon my forces from one twenty-four hours to another*

3.5. The singular-plural indefinite puzzle, synchronically (concerning question (i))

- Why is the introductory *a(n)* so strictly contingent on an adjective coming before a cardinal numeral? Why not further reanalyses, maintaining the same contrasts?

(26) NON-COLLECTIVE: *two days*

COLLECTIVE: *a two days*

(27) NON-COLLECTIVE: *two sad days*

COLLECTIVE: *sad two days*

- Actual reanalyses dictated by the architecture of NPs in languages where phrasehood is crucially defined through linear order and word-class combinatorics (co-occurrence), but not or not centrally through overt agreements (co-marking):
 - NPs here need a determiner in initial position;
 - an NP being both indefinite and plural, with plural morphology on the count noun, is one of the (set of) circumstances that permit the determiner position to remain overtly unfilled;
 - nothing preceding the first determiner can claim membership in the NP
 - unless given special dispensation:
 - degree adjective phrases skipping ahead of a determiner, licensed by their special introductory marking: *how(ever)/so/too/as/more/less/this/that* ADJ *a(n)* NOUN;
 - certain quantifiers other than cardinal numerals (*all the money; twice the amount*), licensed by their dubious modifier status;
 - regular adjectival modifiers before cardinal numerals?

license for the determiner position to remain overtly unfilled (see above: plural indefinite) is revoked, in order to reinforce an earlier left phrase boundary, and thus to reincorporate the adjectival modifier in the NP (unassociated adjectives in other than predicative function are illicit in Modern English: to form NPs on their own, they need a determiner and/or a pro-noun *one*):

(28) DET MOD QUANT HEAD

a. *the sad two day-s*

b. **∅ sad two day-s*

the best filler, under the circumstances, is the indefinite article:

- first, it is a determiner (a left-edge marker);
- second, it is indefinite;
- third, though otherwise confined to singular NPs, its conflict with a plural noun in the same NP is defused owing to the COLLECTIVE semantics of unit-modification (1a) and of number-modification (1b/c) (which also licenses a single indefinite article with conjuncts forming a conceptual unit, such as *a knife and fork*);
- fourth, there had been NPs around for a while, such as (23) and (24), which could serve as formal as well as semantic models.

3.6. Elsewhere in Germanic (concerning question (iii))

- Modern (Standard) High German:
 - a cardinal numeral and an adjective can be inverted as in English, and for exactly the same scopal purposes; but only in English and not in German is there a need for a determiner in the case of such inversions;
 - phrasehood, at least that of NPs, is not as rigidly defined through positions as it is in Modern English; the left extension of NPs instead is crucially delimited through agreement, with most kinds of words in determiner and modifier functions agreeing with nouns in gender, number, and case; in addition, the choice between weak (definite) and strong (indefinite) inflectional alternatives for these agreement categories on determiners and modifiers is determined left-to-right, starting with whatever determiner appears, or is licensed to be zero, in initial position;
 - as in English, the initial determiner position is permitted to remain empty when an NP is indefinite and plural;
 - even with an adjective moved ahead of a numeral in a plural indefinite NP, as in (29b), it is unmistakably coded as being within the left extension of the NP through its inflection: the adjective's agreement in case and number (with gender distinctions neutralised in plural) shows it to be a member of the same phrase as the following noun; its strong inflection (in contrast with weak as in (29a)) needs licensing from the left, which also goes to show that the adjective is not outside such a licensing domain:

(29)	DET	MOD	QUANT	HEAD
a.	<i>d-ie</i>	<i>traurig-en</i>	<i>zwei</i>	<i>Tag-e</i>

	DEF-NOM/ACC.PL	sad-NOM/ACC.PL.WEAK	two	day-NOM/ACC.PL
b.	Ø	<i>traurig-e</i>	<i>zwei</i>	<i>Tag-e</i>
	INDEF(-NOM/ACC.PL)	sad-NOM/ACC.PL.STRONG	two	day-NOM/ACC.PL

- thus, no functional pressure to mark a left NP boundary and reincorporate an adjectival modifier that has skipped ahead of a cardinal numeral, by whatever additional means, even such a relatively least obtrusive one as a singular indefinite article amenable to a collectivising interpretation;
- on the contrary, trying to overtly fill the leftmost determiner position by the indefinite (cum collectivising) article would create considerable problems, owing to its inflecting for gender, case, and number (SG only), and needing to match the inflectional features of the noun (which is PL, where gender is neutralised) within the same phrasal domain;
- possible ways out of the inflectional dilemma for an indefinite article, all at odds with current grammar:
 - (a) it could either acquire plural endings (30a), which would not only be a morphological innovation, but one that would raise a question of also using an appropriately inflected indefinite article in plural NPs of all kinds or of limiting it to just this present situation;
 - (b) or it could mismatch the noun and the adjective and stay singular (30b), with the possible problem that a plural noun is neutral between genders and the right gender would therefore have to be recovered from the noun's singular;
 - (c) or it could become inflectionally inert (30c), thereby diverging from the indefinite article and disturbing the overall positional and inflectional architecture of NPs.

(30)	a.	<i>*ein-e</i>	<i>traurig-e</i>	<i>zwei</i>	<i>Tag-e</i>
		INDEF-NOM/ACC.PL	sad-NOM/ACC.PL.STRONG	two	day-NOM/ACC.PL
	b.	<i>*ein-Ø</i>	<i>traurig-e</i>	<i>zwei</i>	<i>Tag-e</i>
		INDEF-NOM/ACC.SG.MASC	sad-NOM/ACC.PL.STRONG	two	day-NOM/ACC.PL
	c.	<i>*ein</i>	<i>traurig-e</i>	<i>zwei</i>	<i>Tag-e</i>
		INDEF (?)	sad-NOM/ACC.PL.STRONG	two	day-NOM/ACC.PL

- the same problem of its inflectional accommodation had also been encountered when *ein* was extended to serve as an approximative marker — which never really caught on: most of the attested examples show inflectional inactivation as the preferred way out (31a); and only rarely would regular plural inflections be added to *ein-* (31b), derivative of the numeral ‘one’ and still of the same form.

- (31) a. *Wir lebten ein 8 Tag-e zusammen.*
 we lived a 8 day-NOM/ACC.PL together
 ‘We lived some eight days [i.e., about a week] together’
- b. *Ich will noch ein-e acht Tag-e warten.*
 I shall still wait a-NOM/ACC.PL day-NOM/ACC.PL wait
 ‘I’ll wait about another eight days [i.e., a week]’
- For present purposes, the most immediately relevant manifestations of phrasehood in transition: (i) the agreement of adjectives (and also the weak/strong contrast, determined from the left), above claimed to be crucial for adjectives to be placed at the left edge of an NP, ahead of a numeral, and (ii) the inflection of determiners, notably the indefinite article deriving from the numeral ‘one’, above claimed to cause a problem of number inflection in NPs with a numeral and a noun in the plural.
 - However heavily the overt paradigmatic contrasts have been curtailed, all Germanic languages other than English have retained weak/strong and agreement inflections on adjectives. Like in Modern High German, attributive adjectives agree in gender, number, and case and further distinguish weak and strong in Modern Icelandic and Faroese, except that three genders are here also distinguished in the plural, where they are completely neutralised in German. Yiddish and Pennsylvania Dutch are in relevant respects much like the Upper dialects of High German (Bavarian and Alemannic). In Modern Dutch, and similarly in Frisian and in Afrikaans (with a biggish, phonologically defined subset of non-inflecting adjectives), adjectival inflections are severely reduced to two, *-e* and zero, but these two exponents are distributed (with zero limited to NPs which are indefinite, singular, and neuter) so as overall to reflect contrasts of weak/strong, two genders (singular only), and two numbers; no case contrast has survived. With a similar minimum of exponents (with only *-t* added for indefinite singular neuter) and a similarly heavy share of neutralisations, the same three categories (no case either) are also being distinguished by adjectives in the continental Scandinavian languages, Danish, Norwegian (Bokmål and Nynorsk), and Swedish.
 - On such grounds, given that a cardinal numeral and an adjective can be reversed for unit-modification and/or number-modification in the first place (they probably can’t in Afrikaans), an adjective (a non-determiner) at the left edge of an NP would everywhere, except English, be expected to be inflectionally licensed, à la High German, and no extra

left-edge marker (such as a determiner), à la English, should be required anywhere.
(Icelandic and Faroese do not have inflectional articles in the first place.)

- potential problems: Dutch (32) and (“thick”) Bavarian (33), permitting (though not requiring) initial indefinite singular (Bavarian?) articles in adjective-before-numeral NPs.

(32) a. ^{?een/??}Ø *treurige twee dagen*

b. ^{?een/??}Ø *merkwaardige negenennegentig vertoningen*

(33) a. (à) *draurige zwoà Dog*

‘a sad two days’

b. à *bemeàggnsweàdde siemàsechzg säiddn langweilige Foàschdellungà*

‘a remarkable sixty-seven exceptionally boring performances’

- (34) indefinite article (=atonic form of numeral ‘one’, with certain inflectional peculiarities),
extended to plural (now invariable where retained; originally regularly inflected)

	MASCULINE	NEUTER	FEMININE	<u>PLURAL</u>
NOMINATIVE	à	à	à	à
ACCUSATIVE	à-n	à	à	à
DATIVE	à-n/à-m (àrà/àrà)	à-n/à-m (àrà/àrà)	à (àrà/àrà)	à

4. Double determination through articles: Seven reasons for overmarking

4.1. When a constituent to be marked consists of one or more parts which are of the same kind as the whole, these parts may thus themselves be worthy of the same externally-assigned marking as the whole.

This is how case, for example, may get marked on all conjuncts of a coordinate NP rather than just once, or on both the head and an appositive. The lesson of double articulation in this respect is that the distinction between constructions consisting of two NPs, with one NP loosely added to another as an appositive or an afterthought or with one NP recursively embedded in another, and of only one NP, amplified by an attributive or a noun in closer-knit apposition, is not a categorical one. Judging by the evidence of §§2.1-5 and 2.11-12, articulation can be highly sensitive to even minimal disturbances of the simplicity and integrity of an NP, being ready to reapply whenever an adjectival attributive approaches nominal status or begins to detach itself from its head or leaves its customary position.

- (13) a. Gib mir das/ein Schnitzel, das größere / ein mageres!
give me the/a schnitzel, the bigger / a lean [one]!
b. Gib mir das größere / ein mageres (*das/*ein) Schnitzel!
give me the bigger / a lean (*the/*a) schnitzel!
- (14) Vergiß das Schnitzel(.) das zähe!
forget the schnitzel(.) the tough [one]!
- (15) the river (*the) Thames
- (16) a. o θíos/fílos mu (o) aléksis
the uncle/friend of.me (the) Alexis
'(my) uncle Alexis / my friend Alexis'
b. i póli i aθína
the town the Athens
- (19) o kírios (*o) zórbas; o aléksis (*o) zórbas
the mister (*the) Zorbas; the Alexis (*the) Zorbas
'Mr Zorbas'; 'Alexis Zorbas'
- (135) a. die (bsuffane) Sau die!
the (drunken) swine the!
b. Hast die (bsuffane) Sau die gsehng?
have:you the (drunken) swine the seen?
- (136) a. du/ich (alter) Depp du/ich!
you/I (old) fool you/I!
b. Woher weißt du (alter) Depp du das?
how know you (old) fool you that?
- (135') a. die (bsuffane) Sau die bsuffane!
the (drunken) swine the drunken!

- b. Hast die (bsuffane) Sau die bsuffane gsehng?
 have:you the (drunken) swine the drunken seen?

4.2. Particular categories may be utilized for purposes of indicating the syntagmatic relatedness of parts of well-integrated phrases and clauses, and are accordingly marked, typically in morphologically bound form, on the parts to be related by virtue of rules of agreement or government.

(In)definiteness is a less popular category of relatedness-indication within the NP than number, case, gender/class or person, but there are instances where agreement or government might be held responsible for such double marking in simple NPs (§§2.6 and 2.16.2).

- (61) a. ett hus
 a house
 b. hus-et
 house-DEF
 c. ett stor-t hus
 a large-STRONG house
 d. det stor-a hus-et
 the/that large-WEAK house-DEF
- (152) a. kalbu-n
 dog-INDEF
 b. al-kalbu-Ø
 DEF-dog-DEF

4.3. A category may be marked more than once on the same constituent in order to express a meaning that compositionally results from the combination of this category with itself.

Thus, adding a second plural marker may transform a paucal into a multal; a dual meaning, without a form of its own, may be constructed by combining a singular with a plural; or two local cases may be required for a meaning (such as 'from under') unexpressible by a single case. Simultaneously marking a simple NP as definite and indefinite, or perhaps also as doubly indefinite, is a comparable combinatorial strategy for compensating the lack of a special form for specific reference (§§2.13-14).

- (142) a. si-ktab
 INDEF-book
 'some book (or other)'
 b. l-bent
 DEF-girl
 c. wahed l-bent
 a DEF-girl
 'a (specific) girl'

4.4. There may be what looks like double marking for one category although in actual fact the duplication is more apparent than real because the two identical markers have different, although perhaps related functions, or are even homonyms, whose overt sameness is merely coincidental.

Thus, when a noun carries what are ostensibly two cases this may turn out to be a case of mistaken categorial identity, insofar as one 'case' is really something else, such as a derivational affix or a topic marker. The superlative degree and ordinal numerals are categories with a formal and presumably also functional affinity to definiteness; and if definitely articulated NPs are amplified by superlatives or ordinals what looks like the same marking (but isn't) may appear twice, unless the formally identical but functionally different markers fall prey to reduction (§§2.7-8). What second articles may also turn out to be upon closer inspection is nominalizers or attributivizers, with nominalization and attributive marking being other functional domains that overlap with (in)definiteness (§§2.6 and 2.16.1).

4.5. Like whole words or groups of words, marking may be repeated for a pragmatic rather than a semantic effect, especially for the purposes of emphasizing the category concerned (**focus**) or of lending greater affective force to the entire expression containing it.

Negation, diminution or gradation are categories with a well-known proclivity to emphatic or affective doubling. Re-articulation for analogous reasons, or at least with such reasons as an influence among others, has been amply documented above (§§2.3.1, 2.4, 2.7, 2.9, 2.12, 2.13).

- (34) a. a sheyn-Ø meyd
a pretty-NEUT.SG.NOM.INDEF girl
b. a meyd a sheyn-e / sheyn-s
a girl a pretty-FEM.SG.NOM / pretty-NEUT.SG.NOM.PRED
- (39) a. soldat=ul bun
soldier=the good
b. soldat=ul cel bun
soldier=the the good
- (48) soldat=ul cel-a bun=ul
soldier=the the-MASC.SG.NOM.PRO good=the
- (49) soldat=ul cel bun=ul
soldier=the the good=the
- (99) a. ein recht/?so/ganz/(viel) zu/wie warms Bier
a quite/?such/very/(much) too/how warm beer
b. recht/so/ganz/(viel) zu/wie ein warms Bier
quite/such/very/(much) too/how a warm beer
c. ein recht/so/ganz/(viel) zu/wie ein warms Bier
a quite/such/very/(much) too/how a warm beer

- (98) b'. [Adverb [Adjective]] Article Noun
 (105) a. an awfully big schnitzel
 b. such/what a (big) schnitzel
 b'. so/as/too/how/however big a schnitzel

4.6. A constituent may get doubly marked as a result of the simultaneous realization of two roughly equivalent but syntagmatically mutually exclusive ways of marking it, with the execution of the utterance plan strictly speaking going astray. That is to say, two alternative constructions, with the same elements in different linear order or with different elements in the same position, may get blended or contaminated.

It is by way of blending, for example, that double plural marking sometimes originates from the sequential combination of plural allomorphs, and such complex exponents may subsequently be reanalysed as simple. And blending is sometimes the reason also for double articulation in NPs, bare or amplified, that do not contain another NP but are relatable to an equivalent NP with a single article in a different position or with a different article in the same position (§§2.9, 2.10, 2.15).

- (98) a. Article [Adverb [Adjective]] Noun
 b. Adverb Article Adjective Noun
 c. Article Adverb Article Adjective Noun
 (99) a. ein recht/?so/ganz/(viel) zu/wie warms Bier
 a quite/?such/very/(much) too/how warm beer
 b. recht/so/ganz/(viel) zu/wie ein warms Bier
 quite/such/very/(much) too/how a warm beer
 c. ein recht/so/ganz/(viel) zu/wie ein warms Bier
 a quite/such/very/(much) too/how a warm beer
 (116) en pert das beste
 a horse the best
 (145) a. ketâb
 'book, a (or some) book(s), the book (in question)'
 b. ketâb=i
 book=INDEF
 c. yek ketâb
 a book
 d. yek ketâb=i
 a book=INDEF
 'a book, whatever book, a certain book'
 (146) a. ketâb-e kamyâb=i
 book-EZÂFE rare=INDEF
 b. ketâb=i kamyâb

- book=INDEF rare
- c. *yek ketâb-e kamyâb
 *a book-EZÂFE rare
- d. yek ketâb-e kamyâb=i
 a book-EZÂFE rare=INDEF
- e. yek ketâb=i kamyâb
 a book=INDEF rare

4.7. When a first marker is not really very distinctive or is no longer recognizable as what it used to be, as a result of formal erosion or owing to its having been reanalysed as something else, such as an integral part of its carrier, a constituent may be marked anew for the same category in the interest of transparency.

The lack of distinctiveness is a frequent reason for overmarking for categories of any kind. In the sporadic cases on record where an original article or parts of an article have become amalgamated with a noun (§2.17), the reanalyses responsible are not synchronic processes and the subsequent re-articulation of such nouns therefore does not count as synchronic doubling. Also, when articles age and develop into whatever they are destined to be recycled as, and eventually come to coexist with their successors (§2.16), it is only in a historical sense that the two generations of articles share a categorial identity.

- | | | | | |
|-------|----|---------------|---|---|
| (154) | a. | l-ilma | < | il-ma |
| | | DEF-water | | DEF-water |
| | b. | in-nuccali | < | n- (< l-)uccali (< Sicilian <i>ucc[h]iali</i>) |
| | | DEF-glasses | | DEF-glasses |
| (155) | a. | a newt | < | an ewte |
| | b. | a nickname | < | an ekename |
| (156) | | the/an alcove | < | al-qobbah (Arabic) |
| | | | | DEF-vault |