Aspects of English morphosyntax

Jinyuan Wu

February 4, 2024

Contents

1 Introduction

1.1 The language and the speakers

The modern English language is a West Germanic language, with heavy Latin and Romance influence.

The English language is arguably the most important language in the world, taking factors like the number of both first and second language speakers, geographical distribution, the amount of written materials in it, and the influence of English-speaking countries (Quirk et al. 1985, p. 3). The wild popularity of English makes it the default language in linguistics study, both as the metalanguage or as the object language. This work is aimed to give a theory-informed and unified overview of English grammar.

Considerable variance exists even between versions of English that are considered standard in main English-speaking communities, summarized as follows:

- A large part of the English vocabulary is never regularized (e.g. *biscuit* (UK) v.s. *cookie* (US)). Some words even have variances across cities. The word *transit* may refer to the bus system or mainly the subway system, for example.
- Words that clearly are
- Variances in some grammatical traits in commonly accepted speech can also be observed; British English is said to have a preference for the present perfect tense for past events. The pronoun *they* in Indian English sometimes is replaced by the phrase *the same*, which probably is a residue from old-fashioned British business English.

Given the reasons listed above, it can be seen that "standard English" can never be well-defined defined. Still, since most of the variance is concentrated on the lexicon, we are able to define "International Standard English", or in other words the shared core grammar accepted by all English-speaking communities as the standard one.

1.2 Theoretical orientation

In short, this note is based on Basic Linguistic Theory (Dixon 2009, 2010, 2012) with generative flavor. It would cost too much space to given an outline of the descriptive framework here; TODO: turn note-1 into a note about formalism

1.2.1 Formalism

In short, this note is based on Basic Linguistic Theory (Dixon 2009, 2010, 2012) with generative flavor. Here by the term *generative* and especially *modern generativism*, I mean Dis-

tributed Morphology and Cartography (but probably not Antisymmetry). I'm fully aware of other frameworks, like classical GB or HPSG, but again, I need to make theoretical commitments.

A well-known (unfortunate) fact about the generative enterprise is people tend to focus on fragments of languages, and writing a systematic reference grammar in terms of modern generativism seem unpractical. I will not go over criticisms towards generativism – readers can find lots of them in Dixon's works. It's however my belief that the descriptive theoretical framework outlined by Dixon is largely – if not completely – compatible with modern generativism. This note is partially intended as a demonstration of this idea.

Haspelmath (2009) argues we need framework-free description of languages, because each language has its own categories. The fact, however, is although it's definitely plausible that languages can vary indefinitely with respect to the grammatical framework required, an empirical observation is that the metalanguage used in grammars published by, say, Language Science Press, are largely commensurate. This note is an attempt to summarize the basic ideas underlying the successful frameworks and apply them to English.

1.2.1.1 Constituency and dependency

I consider three major perceptions of constituency and dependency:

- The generative approach, in which dependency relations and grammatical functions like "subject" are not primitives themselves but are reflected by the structure. In the Minimalist, Distributed Morphology-like and Cartography-like approach favored in this note, concepts like auxiliaries, etc. are realizations of the hierarchy of functional heads, and concepts like subject, object, adverbials are specifiers of these functional heads, and at the core (i.e. complement in the generative sense of the lowest functional XP) of a functional hierarchy lays a noun or verb root; a noun or a verb is the root plus realizations of some functional heads, like Tense, etc.; top-level constructions, like clauses and noun phrases, are essentially "phases" in generative syntax.

- The (more traditional) constituency-based structuralist approach that is recently found in large-scale descriptions like Huddleston and Pullum (2002) and Abeillé and Godard (2021).

- The Basic Linguistic Theory (BLT) approach, advocated in Dixon (2009, 2010, 2012), sometimes referred to as the functional-typological approach, although in pure linguistic description, usually no functional explanations are invoked. This approach is the underlying framework of most modern descriptive grammars, e.g. Jacques (2021); Friesen (2017); Forker (2020), and more.

Since the difference between the flavor of generativism sketched above and BLT is huge, here I give the dictionary between the two roughly in Table 1.1. In this table we see the constituency-dependency duality: complement-specifier relations are replaced by dependency arcs, and c-command relations are replaced by the order of "closeness" or "height" of dependency relations. It should be noted that in BLT, a constituent (word or phrase) can be a tree in generativism, but also a phase layer (e.g. a CP minus all DPs in it); thus we get the definition of the verb phrase as the main verb plus auxiliaries, *excluding* the object, in some works like those of Dixon.

Table 1.1: Correspondence between concepts in generative syntax and BLT

Generativism	BLT
phrase (DP, *v*P, etc.) with specifiers	dependency relations
phrase without specifier	marking of grammatical category
specifier-complement relation of phrase	grammatical relation (subject, object, etc.)
functional head	grammatical category (tense, aspect, number, etc.)
several layers of co-occurring projections	constructions
root in Distributed Morphology	lexical head
a batch of spellout result	a phrase (verb phrase = *v*P-TP-CP span) or word
phase (DP or CP)	top-level constructions like clauses or noun phrases
tree structure	order of "closeness" of dependency relations
movement	dependency network

Besides the constituency-dependency duality, two other parameters are important in the transition from generativism to BLT: headness and coarse-graining.

The alternation of the definition of headness is clear, since in BLT, functional heads are no longer recognized as heads. There however exist ambiguities, like whether some prepositions are lexical heads, i.e. whether they are actually adverbial roots surrounded by their own functional hierarchies, or they are just like case markers that are realizations of functional heads introducing noun phrases into a clause. The definition of headness influences terminological choices and visualization of dependency relations: the term *preposition phrase* for example is not be used if prepositions are recognized as functional heads and not lexical heads, and the dependency relation between a verb and a preposition phrase, under this convention, will not be represented by an arc from the verb to the preposition: an arc from the verb to the central noun will be used instead.

Coarse graining means to zip several dependency relations into one. The concept of subject, for example, contains at least two dependency relations: the verb-agent relation (i.e. the specifier-complement relation in the external *v*P) and the verb-clausal pivot relation (created in TP).

Now we turn to the approach in Huddleston and Pullum (2002). There is a difference between the representation of constituency in traditional structuralism and in generativism. In Minimalism we have bare phrase structure: we don't really attach labels like SpecTP to the specifiers: which constituent is the specifier of a functional projection can be easily read from the tree; and similarly labels like "DP" are also not really needed. But in the traditional structuralism, invisible functional heads that those introducing the external agentive argument or the subject are removed, and thus instead of letting the tree to tell us that a subtree is the SpecTP (or to be accurate, SpecNomP, where Nom is some kind of "nominative head") and it's a DP, we have to attach the label "subject: NP" to the subtree.

Alternation of the definition of the head and coarse graining are also important in transition from generativism to the more traditional structuralist approach: this is exemplified by mentioning "subject" as opposed to "the syntactic pivot and the agentive argument" and "NP" as opposed to "DP" in the last paragraph. The relation between the two formalisms are summarized in Table 1.2.

Going from generativism to structuralism in Huddleston and Pullum (2002) and the like still involves some amount of constituency-to-dependency mapping: in constructions like *a better result than we expected* (Huddleston and Pullum 2002, p. 55), the complement *than we expected* licensed by the adjective *better* is moved to the end of the noun phrase, possibly because of prosodic reasons, but Huddleston and Pullum (2002) is reluctant to use movement analysis when no in-situ forms can be identified in surface forms, so the authors have to man-

ually add a dependency arc between teh adjective and the *than* preposition phrase besides the constituency structure. Another example is in Huddleston and Pullum (2002) we see ternary trees in ditransitive constructions; here the relative height of the two objects can still be decided by careful syntactic tests, but this piece of information is conveyed by discussion on "which object is easier to passivized", etc.

Table 1.2: Correspondence between concepts in generative syntax and more traditional structuralist grammars

Generativism	Structuralism
phrase (DP, *v*P, etc.) with specifiers	phrases taking complements or adjuncts
phrase without specifier	phrases dressed by grammatical category marking
specifier-complement relation of phrase	grammatical function (i.e. sub-tree labels)
functional head	grammatical category
several layers of co-occurring projections	constructions
root in Distributed Morphology	lexical head
a batch of spellout result	a "routine" without constituenthood
phase (DP or CP)	top-level constructions like clauses or noun-phrases
tree structure	tree structure
movement	tree structure change; "indirect complement"

A topic left untouched in the above discussion is the argument-adjunct distinction, or in the terminology of Huddleston and Pullum (2002), the complement-adjunct distinction. (Here the term *complement* doesn't have the generative meaning.) Although some lines of generativism insist on a strict distinction (in Minimalism, the distinction between Merge and Adjoin), this note does not make such a hard distinction. Adjuncts, follows the tradition in Cartography are analyzed as optional specifiers of some functional heads, and arguments are specifiers of functional heads as well. The usual tests, like obligatoriness, closeness to the head, and information structure properties, are then decomposed down to orthogonal syntactic properties. Obligatoriness boils down to morphosyntactic properties like whether the verb root spells out well with or without "transitivity", the functional head introducing the object (Siddiqi 2009, Ch. 10), or pragmatic requirements for the speaker to give a complete event structure (Grimshaw and Vikner 1993) and/or to say something meaningful (Goldberg and Ackerman 2001). [1] The closeness criterion is not clear-cut by definition. The island effect criterion is possibly not really due to the argument-adjunct distinction (McInnerney 2023). Therefore, I choose to regard the complement-adjunct distinction as a shorthand for various configurations of the aforementioned parameters. This viewpoint reflects subtleties in distinguishing adjuncts from complements.

Unfortunately, although the above discussion clearly demonstrates the close relation between the abstract concepts in generativism and actual linguistic description, and tools from formal linguistics can inspire and already have inspired linguistic documentation (Rice 2006), field linguists are typically dubious about the value of generativism, for good reasons: generative field work often produces distorted data, generative linguists often misrepresent the data they acquired and fail to place the data in a proper context, etc. It's my hope that in the future the wall between the two disciplines can be eliminated, since without in-depth analysis, many "exotic" features of less known languages can also be exaggerated. Even for mainstream

[1] The semantic and pragmatic explanations can't replace the morphosyntactic explanation because, for example, some verb frames indeed do not give the full event structure but are still felicitous, like *the dog kills* ('the dog kills some small animals'); this involves syntactically suppressing the internal argument.

languages like Mandarin, misleading claims like "in Mandarin argument structure is not grammaticalized" are still prevalent.

1.2.1.2 Form and function

The form of a construction is its internal makeup, while its function is just the grammatical function like subject or object; in generativism, the form of a construction is its internal functional hierarchy, while its function is the functional hierarchy that contains it. Of course, clear terminological distinction between form and function should be made, but in typological literature this is not always the case. For example, the labels in Fig. 3.1 are about function and not form; but in typology, we use the symbols Dem, Num, A, and N to represent their prototypical syntactic function in NPs: Dem = the determiner-like region, A = the attributive region, etc. Saying English has a Dem Num A N constituent order doesn't mean the attributives are all adjective phrases. This misuse of terminology may also be kind of justified, since part of speech tags are defined by syntactic functions anyway (§ 1.2.2.3); although the ability to be an attributive isn't truly substantial for adjective-ness.

I also accept the term *fused-function construction* in Huddleston and Pullum (2002), and specifically, *fused-head constructions*. In *all of the books*, for example, we may say that *all* collectively realize the noun categorizer and the quantifier head, and thus in the surface-orientated analysis, we say *all* is a fused-head construction where the lexical head and the quantifier function are fused together. What is important is not the terminology, but to notice that fused-head constructions have properties that are different from ordinary ellipsis.

1.2.1.3 Fossilization

Some constructions are only semi-analyzable – they are fossilized. Fossilization has several stages: a fossilized construction created by a fossilized process be totally historical and no longer analyzable in contemporary morphosyntax, or analyzable in contemporary morphosyntax but already with an established meaning (semantic fossilization), or semantically compositionally analyzable but the relevant syntactic process is limited in productivity (syntactic fossilization). The parameter of morphosyntactic fossilization and the parameter of semantic establishment, despite having strong connections to each other, are still two parameters. Mechanisms of syntactic fossilization is plural: it can be grammaticalization (lexical morphemes collapse into grammatical markers, or large constructions collapse into smaller ones), complicated selection and/or incorporation processes developed in the lexicon that block some syntactic processes, or idiomization, which isn't far from "mere" semantic fossilization: to ensure the listener is able to recall what idiom is being used, syntactic operations should already be limited (Nediger 2017). Fossilization is also closely related to wordhood because criteria for wordhood are usually just aspects of fossilization (§ 1.2.2.1).

1.2.1.4 Gradience

Gradience in grammar does exist, usually because of competing analyses of the same form, often as a result of fossilization – the structures before and after fossilization are both present in the current stage, and the status of the form is therefore not completely determined. Sometimes, though, gradience comes from terminological confusion. As an example, sometimes *on the hill* in *on the hill lays a small house* is analyzed as something between an adjunct and a subject. The structure of this construction is quite clear and deterministic: the verb *lay* is an unaccsative one, and therefore the argument – the person or thing that lays down – is not agentive enough and is not in a very external position; so now *small house* stays in-situ and the oblique argument *on the hill* is promoted to a higher position. The term *subject* is usually defined as the constituent that both is in the most external obligatory position in the argument

structure and is the syntactic pivot; *on the hill* isn't obligatory, and therefore is only partially a subject, and the fact that it's in a non-obligatory locative position in the argument structure means it's also partially an adjunct. What is demonstrated here however is the fact that terms like *subject* or *adjunct* are not always useful, since they pack too many concepts into one word, not that *on the hill lays a house* is some exotic construction.

1.2.1.5 Semantics and syntax

Sometimes people say that only semantics is real, and syntax is just the "realization" or projection of the underlying semantic structure. But then some constructions have rather rigid meaning with very different conversational context (as in, say, obligatory control), and in other occasions, the most natural syntactic representations of semantically well-formed meanings are not acceptable. These phenomena can be explained by fossilization of historically semantically motivated constructions, but the question then is what these constructions have fossilized into; hence the differences in construction grammar and the like, lexicalist theories, and the derivational, piece-by-piece Minimalist Distributed Morphology framework adopted here. That's not to say semantics is not important in grammaticality: indeed, the obligatoriness of some clausal dependents may be explained by semantic factors (§ 1.2.4.2).

1.2.2 Lexicography

1.2.2.1 The word

Wordhood can be divided into phonological and grammatical wordhood. The latter can be divided into syntactic and morphological wordhood. A syntactic word is just a mini-constituent, like a compound noun (as opposed to, say, a DP). Thus, in this note I refrain from discussing whether a function item like *a*, *the*, etc. is a grammatical word or not in detail, since it's the realization of one or more functional heads and is not the realization of a maximal projection. A function item licensed in a phrase may be phonologically dependent – we say it's a clitic, while a functional item licensed in a grammatical word, like *-like* in *Distributed Morphology-like*, may sometimes has the status of an independent phonological unit. Thus when talking about wordhood of a function item, it's always better to specify whether we are talking about its licensing position or its phonological independence.

The syntactic wordhood has some vagueness: the meaning of *constituent* is not determined. Recall that some authors use the term to refer to a *phase* in generative syntax, and thus a verbal complex is a "verb phrase" and a inflected verb is a mini-constituent (i.e. an inflectional version of the verb phrase). But if we insist on using the term *constituent* to refer to a constituent in the generative sense, then an inflected verb is not recognized as a word, but always a word plus some auxiliaries – indeed this is the perspective taken in traditional Japanese grammar, even though these auxiliaries are tightly attached to the verb root (see below) and are unable to undergo, say, English auxiliary fronting.

Syntactic wordhood is not enough for wordhood definition: If we only focus on syntactic wordhood we may say *has being eating* is a word, since it's the exponent of the verbal complex; of course we can also say it's a span in the verb phrase. To capture the intuition that *has being eating* has several words, we define the concept of morphological word: if functional items pertaining to a root are tightly attached to that root, then they form a morphological word, while in *has being eating*, we can insert an adverb between the auxiliaries and *eating*, so there are three morphological words. But whether a function item is firmly attached to the root it modifies is sometimes hard to say and observes diachronic changes. An example is whether spoken French has polysynthetic structures. Another example is about Japanese, in which historical auxiliaries seem to already fossilized into inflectional endings, as an inflec-

tional template seems to have already formed (although these auxiliaries still have their own "inflections", just like English auxiliaries).

Thus, we can say the concept of wordhood works like this: first we have syntactic wordhood, which comes from the abstract constituency and dependency structures; derivational morphology is mainly about this step, but this step may also be relevant to inflectional morphology (as in "verb phrase" in the BLT sense). Then some small parts in the structure are attached to larger units, like how tense and person endings are added to the verb; in this step we get morphological wordhood, which is primarily about inflection, but also derivational morphology since derivational affixes may be attached to the lexical head, etc. Finally, the (already flat) morphological structure is rearranged by phonology and divided into phonological words.

Another criterion frequently invoked is that a word has an established meaning. This however is better understood as a tendency: some phrases also have established meaning. It seems words have established meanings not because wordhood enjoys a special position a priori in grammar, but because a word is a small unit and small units tend to be fossilized both in meaning and in structure.

In conclusion, all aspects of wordhood – smallness in size, morphological (though not necessarily syntactic) closeness of the parts inside, morphophonological behaviors, established meanings – are theoretically independent parameters that also appear in clearly non-word constructions and in principle should be analyzed independently (Harley and Noyer 1999). The seemingly close relation between the parameters is more likely due to historic or functional reasons; for example, a perfectly analyzable compound noun with an established meaning will soon be replaced by its acronym (*turn on the [AC]!*) or undergo corruption of its internal structure, and the result is perfectly a "word" no matter how you understand the term.

1.2.2.2 The meaning of "morphology"

As is said above, the term *morphology* involves both the internal structure of the word ("syntax within the word") and how this structure is realized. Traditional grammars usually have a large paradigm with its row and column headers being grammatical categories. (When there are too many categories – and in this case the language in question is usually agglutinative – the paradigm will be unbearably large, and another way – like the School Grammar of Japanese – is needed to cover verb inflection. Still, partial paradigms are useful in this case.) This is a morphosyntactic way to represent the inflection of a word, but if we are talking purely about the *morphological* part (i.e. how grammatical relations and categories are realized), then it's sometimes not necessary to recognize so many forms: If a verb appears exactly the same in two different syntactic environments, then we say there is only one *inflectional form* of that verb. For languages like Latin, the traditional large-paradigm way is handy, while for English, we can zip the paradigm severely (Huddleston and Pullum 2002, Ch. 3.§ 1.2).

Note that the fact that the realization of a mini-phrase may not absolutely transparently reflect its inner structure means that even though the morphology of a word seems quite "concatenative", it's possible that, say, a grammatical category is realized by two or even more affixes, both of which are obligatory, not because of any morphosyntactic reason, but because of a template of the word. In this case we say the language has template morphology; on the other hand, when the affixes reflect the abstract structure of the word clearly and can be seen to be added one by one according to the functional hierarchy, we say the language has layered morphology. Still, it seems that template morphology can be well captured by primitives that are already well known in languages with layered morphology Oxford (2019), so the distinction between the two seems to be smaller than it appears to be.

1.2.2.3 Part of speech tags

The variance of parts of speech observed cross-linguistically is often used to support the idea that concepts like "noun" or "verb" are all language-specific (and also to attack the idea of language universals and "formalists"). Following the idea of lexical decomposition (a certain brand of "formalist" linguistics), the position of this note is to regard "nominal properties" (i.e. being in the center of the nominal system (§ 1.2.3)) and "verbal properties" as universal concepts. That's to say, I assume that we can take noun phrases and clauses with functional hierarchies similar to well-known languages for granted; the part of speech tags like "noun" or "verb" then are about how a root interacts with these syntactic environments, like whether and how a root can be spelt out with the nominal categorizer or the verbal categorizer in Distributed Morphology.[2] Part of speech classes then are epiphenomena of interaction between roots without part of speech labels and functional hierarchies.

Still, part of speech tags are useful in language description. With respect to the parameters of openness, the content-function dichotomy, and whether function items are similar to content items enough, there are roughly four types of parts of speech:

- Open content categories with clear part of speech labels like noun, verb or adjective.
- Close content categories with clear part of speech labels like noun, verb, or adjectives.
- Close grammatical categories which may still be seen as noun-like or verb-like. The prototypical members of this type are pro-forms.
- Close grammatical categories like particles or affixes that don't really need part of speech labels. Grammatical words in this type – the type without much resemblance to prototypical content words – don't really need part of speech tags, but since they are surface realizations of different grammatical categories, and exponents of different values of the same set of categories usually have similar distributions, dividing these items into groups helps us to organize the grammar, though unlike labels like noun or verb, these labels have less substantiality in the mind of native speakers. For example, in a language with case particles, we may set up a word class called "case particles", which falls under the class of "particles". This is the case for traditional Japanese grammar. Then we can look up for an unfamiliar grammatical construction in a "grammar dictionary".

The third and the fourth classes are usually primarily exponents of functional heads. A word in the third class may contain some features making it look like a word from the first two classes, like the categorizer feature or the person/number/case feature, while a word in the fourth class may not.

Gradience occurs between the boundaries of the four types of parts of speech. Fillers of a specifier position, if limited to a few (like adverbs filling certain positions in TP), may be bleached into realization of functional heads, thus coming into the third and fourth types of parts of speech. Competing analyses occur when this change is happening. The boundary between the latter isn't clear, and neither is the boundary between the first type and the second type (because many so-called closed content word classes sporadically accept new members), and the second type and the third type (for example, when the number of verbs is so-limited – say, only a handful – then is it a better idea to regard the verb class as the exponents of different light verbs?), and also the third type and the fourth type (since the criteria of "looking like a noun or a verb" are never clear).

[2] Siddiqi (2009, Ch. 10) identifies the verbal categorizer and the light verb in the argument structure; the concept of "verb" – and hence the categorizer functional head carrying it – however is likely to enter the derivation earlier than the introduction of the first argument, since in derivational morphology, derivational affixes seem to select the part of speech tag of the syntactic tree they are attached to.

It's often said that nouns are about objects, verbs are about events, and adjectives are about properties. This mapping from word class to semantics is a coarse one and should be refined for more systematic description of languages.

It should be noted that actions and processes can be conceptualized as objects: we have, for example, *his playing of the national anthem*, where *play* is nominalized into *playing* (this is not an ING-participle – see TODO: ref). The boundary between objects and properties is also hard to draw: Apart from pronouns or demonstratives that directly refer to the conversation context and pull out a specific object from the listener/reader's memory (and therefore introduce a free variable in the semantic interpretation of the utterance), nouns denote *sets of objects*, and we know we can have a one-to-one mapping between a set A and a predicate in the form of $\cdot \in A$: A noun like *toothbrush* can be immediately mapped to an adjective, like *toothbrush-like*, and therefore whether *toothbrush really* means a set or a predicate becomes an undecidable question. And similarly a property can also be conceptualized as an object: as a set, or maybe as a kind of "essence" (compare *the ice harvesters are [manly]$_{about\ property}$ men* and *the ice harvesters have strong [manliness]$_{object?}$*).

The boundary between actions and properties is also not clear: In traditional grammar they are all called "predicates". When translated into logical expressions, a clause about an event introduces an event argument, but a clause about a property doesn't. The difference between events and properties may be that while a clause about the fact that an object has a property can be simply interpreted as 'property(x)', or if we want to reduce the number of logical predicates (to avoid the necessity of using higher-order logic), '$x \in$ the-set-with-the-property'. But of course having a property is temporal, so *this is beautiful* is to be interpreted as '$\exists e(\text{time}(e, \text{speech-time}) \wedge \in\text{-in-a-time}(x, \text{what-is-beatuful}, e))$'. This doesn't seem quite different from the meaning of a clause about an event, although there may still be some subtle differences like whether the e in clauses about properties can be the invisible topic, which seems to be the reason for the mysterious *wa*/*ga* alternation in Japanese (Heycock 2008).

So in the end, nouns are prototypically about objects, and they may denote events, and whatever they denote, they can be thought as properties in semantic interpretation. What makes a root a noun is essentially its *syntactic environment*. Intuitively, we say nouns are similar to pronouns, demonstratives, etc., which, however, can never be interpreted as properties. That's because these words contain a determiner fused inside and therefore have clear and direct reference (§ 3.7), and the reason we say nouns are like them is that a noun can be placed at the center of a DP, and the DP now has almost the same syntactic distribution with pronouns, etc.

Semantically, verbs are about actions and properties, but again, they are *categorized* as verbs not because of inherent semantic properties, but by the syntactic environment (being immersed in the *v*P-TP-CP projections). Thus, in principle we don't really need content words other than nouns, which is indeed the case in some languages with very limited verb classes.

The semantic function of adjectives can in principle always be realized by nouns or verbs. The role of them is highly language-specific, and they appear when a meaning is hard to convey using existing constructions concerning nouns or verbs. For example, in English, when it comes to gradience of properties (manifested in comparative constructions), nouns are of limited use, so adjectives are indeed necessary. But we still have *he's more a scientist than a public health official*, and the adjective version *he's more scientist-like than public health official-like*, despite being grammatical, is awkward and only appears in language games instead of natural, everyday speech.

Note that the above discussion may be as well found in an introduction to, say, "Radical Construction Grammar", in which it's argued that grammatical categories only make sense in constructions. This note, however, takes the stance that constructions are not routinized structureless strings, but are made of building blocks that are subject to universal constraints. It's my theoretical assumption that a root being categorized into a noun has nothing substan-

tiality different from adding an article before a noun, although when it comes to processing, the first may be more "automated" and "fossilized".

1.2.2.4 Derivation and inflection

The accepted wisdom is "derivation relates one lexeme to another lexeme, while inflection relates one lexeme to its form in the final utterance". This definition however still has intrinsic vagueness. It involves two parameters: structure size and fossilization. Regarding structure size, derivation is on the level of lexemes, which are smaller than phrases, and also smaller than "finished words" – inflected words – that involve influences from the external syntactic environment. For verbs, for example, we may define derivation and inflection like the follows: derivation means everything in the lowest several projections in the *v*P, while inflection means everything in the TP (basically, the morphological counterpart of Dixon's verb phrase). Regarding fossilization, derivation should be less synchronically active than inflection (and therefore less productive). There is certain amount of correlation between the two parameters: small units are easier to lexicalize, so compound words are more likely to have established meanings, and compounding is therefore prototypically derivational. But of course the parameters may not always agree with each other, and both parameters have vagueness. The derivation-inflection distinction is therefore not appropriate in some cases (Dixon 2009, p. 221).

Here are some examples of the vagueness. Concerning structure size, in case stacking, should the inner case markers be considered as inflection? And note that a valency changing device also doesn't apply to all verbs that seem to have an appropriate number of arguments – indeed, Jacques (2021) calls valency changing *derivation*. A further subtlety is the parameter of fossilization should be further split into two: some constructions may appear less frequently but still have largely compositional meaning, while others – like the Latin *com-* prefix – appear everywhere but the meanings of resulting words can hardly be inferred regularly.

1.2.3 The nominal system

1.2.3.1 The nominal or the NumP-like region

There usually is a nominal part in the nominal system, which is basically the central noun plus modifiers and corresponds to the NumP domain. The NumP itself is similar to the role of TP, and the various adjectives are similar to specifiers of AdvPs in the TP domain, and the complementation is similar to the VP layer (Laenzlinger 2017).

1.2.3.2 The determiner-like region

The higher determiner-like region corresponds to the DP domain, which contains D_{det} (the DP version of FinP; § 3.4.1), the quantifier Q position (§ 3.4.2), which is placed over the determiner (Gianollo et al. 2021), the DP version of topic (§ 3.4.3), and the D_{deixis} position (§ 3.4.4), which corresponds to the ForceP in the CP domain and is about whether the DP is referential, etc., quite similar to the way ForceP expresses what the clause is intended for (Laenzlinger 2005a). Note that in English we don't have NP-inside topics and focuses, between D_{deixis} and D_{det}, although some Romance languages allow them.

Though quantifiers are often seen inside NPs, their semantic scopes are definitely larger. Quantification, strictly speaking, is a phenomenon involving the semantics of the whole clause and is not restricted to the NP: The (final, not immediately – see Box 3.4) interpretation of an NP within a clause is like '$\forall/\exists x$ (x is something $\wedge \ldots x \ldots$)', with possible change of the conversational context; the content of the NP adds more branches containing x after the logical

quantifier. introduce a conjunctive branch like 'x is a student' introduce a quantifier to bind the variable, Below is an example of this procedure.

a) The sentence *[students] usually take [at least four courses] [each year]* – which contains three NPs – is to be interpreted as '$\forall x(\text{is-student}(x) \wedge x \in \text{context} \wedge x$ usually take at least four courses each year)'.

b) The part x *usually take at least four courses each year* is in turn interpreted as '$\forall y(\text{is-year}(y) \wedge \exists e(\text{time}(e, y) \wedge \text{frequency}(e, \text{habitual}) \wedge x$ takes at least four courses in the event $e))$'.

c) The part x *takes at least four courses in the event* e can further be interpreted as '$\exists S(|S| \geq 4 \wedge S \subseteq \text{all-courses} \wedge \text{action}(e, \text{take}) \wedge \text{agent}(e, x) \wedge \text{patient}(e, S))$'. The number 4 here can further be expanded into things like '$\exists c_1, c_2, c_3, c_4(c_1, c_2, c_3, c_4$ are all different $\wedge \cdots)$'.

It should be noted that when studying the structure of DP, we should distinguish semantics and syntax. The meaning of *determination* may be realized by something like the English articles, but it may also be realized by something that looks very like an adjective, as is the case in Latin. What is uncontroversially universal is a set of atomic syntactic features and related semantic meanings, not how they are packaged into concrete constructions, and whether NPs and clauses follow the same cross-linguistic template is still a disputed problem (see the info box in note 1).

1.2.4 The verbal system

1.2.4.1 The verbal complex

Roughly speaking, in Fig. 4.1, the verb phrase is the part of TP that is lower than the projection in which the subject is introduced. The subject-predicate structure is roughly the complete TP. The extended argument structure is the VP domain, and the TAM marking part is the TP domain. Layer 3 and layer 4 are about CP.

1.2.4.2 Clausal dependents

Traditionally, argument structure is analyzed in terms of θ-role and (abstract) Case; we may add a further parameter of *purely* semantic roles that are related to the θ-roles in LF. Minimalism changes this landscape by its pursuit of unification in analyses; Case then is considered one feature, while the θ-roles are largely replaced by functional heads in the *v*P shell (Harley 2011) and purely semantic concepts about the event structure. Obligatoriness of arguments is therefore explained by (a) that the verb stem doesn't spellout without the presence of certain light verb functional heads, like the transitivity head that introduces the object, etc. (Siddiqi 2009), and (b) the need of a complete event structure (Grimshaw and Vikner 1993); valency alternation can be fine-tuned by the two parameters: for example, detransitivization (as in *the dog bites*) contains no TRANS light verb and *bite*+TRANS, in LF, is interpreted as a habitual predication, therefore also nullify the necessity of a patient argument. As for the case system, here we take one step further and follow an implicit tradition in Distributed Morphology that has been made explicit in Schäfer and Anagnostopoulou (2020) and reject the notion of an abstract Case system that motivates promotion of the most external argument to the subject position or requires that the verb and the object is close to the verb: the first phenomenon is now explained by, say, the EPP feature of TP (or maybe AgrS) that requires a specifier, and this specifier can't come from external Merge because this makes the event structure incomprehensible (See? No need for a separate θ-system); the second phenomenon is explained by some obligatory movement of the object to presumably AgrO together with the obligatory fronting of the verb, and

there is no position for locational or manner phrases, which are assumed to be specifiers of functional projections, and not adjuncts in the GB sense. The intertwined three-aspect system of argument structure consisting of θ-role, Case and LF is hence replaced by a more neatly organized three-aspect system consisting of purely semantic concepts, what happens in *v*P, and what happens above it. In the rest of the note we use *(syntactically significant) semantic role* to refer to the second and use *clausal dependent position* to refer to the third.

1.3 Previous studies

1.4 Notable features

English has a pretty rigid surface constituent order, directly reflecting the inner structure of NPs.

2 Parts of speech

This note is organized in a bottom-up manner. Form classes of words – or parts of speech, using terms in traditional grammar – in English are recognized (§ 2.1). A detailed discussion on wordhood can be found in § 2.1.1. Basic properties structure of the word, namely roots and derivational processes of them, are introduced in § 2.10.

Since the construction of phrases the roots project into are important in determining parts of speech, the structures of phrases can be covered when discussing their head words, so in some traditional grammars, all morphosyntactic information – i.e. things not about phonology, the writing system, cultural background, etc. – falls under the part about "parts of speech". The structures of the extended phrases of nouns and verbs however are too large to be placed into one chapter: They are the main focus of the following several chapters.

Since characterization of parts of speech and their contexts unavoidably involves information about larger constructions like NPs or clauses, forward references to the following chapters are to be expected in this chapter.

2.1 Distinguishing parts of speech

2.1.1 Wordhood in English

Unlike the case in Chinese or Japanese, the orthographical tradition of English contains the notion of word.

The English orthographical word is a mixture of the phonological word in very slow speech (i.e. the unit between two natural stops in very slow speech) and also some factors in grammatical wordhood. The phonological factors are dominant: *Distributed Morphology-like* is not recognized as a orthographical word, but at least two orthographical words, although its outmost constituent is the derivational suffix *-like* and is not subject to any adjectival modification within the scope of *-like*, and therefore is a sub-phrasal unit that is smaller than a nominal (§ 3.1) and is to be recognized as a grammatical word. English orthography is therefore different from, say, German orthography, where an orthography word usually covers a whole grammatical word even when the latter is very complicated.

The orthographical word is also not always necessarily the phonological word in more fluent and *natural* speech, where some orthographical words, like a personal pronoun immediately after a verb and the article *the*, are essentially clitics when not emphasized.

There seems to be no highly non-trivial disagreements between phonological wordhood morphological or "realizational" wordhood (i.e. existence of an obligatory inflectional template, etc.): if a formative belongs to a grammatical word, it will never be

shifted to the phonological counterpart of another grammatical word in spontaneous speech. This kind of non-trivial disagreements do exist between morphological wordhood and syntactic wordhood (i.e. constituency and dependency relations in a very small scale) in complex words (§ 2.10.3, § 2.10.5),

if a unit doesn't have phonological wordhood, it is always phonologically attached to the lexical head directly related to it.[1]

2.1.2 Open content categories

Open content categories with clear part of speech labels include nouns (§ 2.3), verbs (§ 2.4), adjectives and adverbs.

The noun-verb distinction can be established by different inflectional patterns, syntactic distributions, and also different modification patterns; details can be found in the sections about each part of speech. There exist words with both verb and noun versions. The noun *strike* and the verb *strike*, for example, clearly have semantic relation. There however lacks regular correspondence between the meanings the verb and noun versions of these words, and no "verb-used-as-noun" or similar rules have to be included in a synchronic grammar of English.

The English adjective class – there is only one adjective class, not two or more, which is the case in Japanese – can be distinguished by

2.1.3 Close content categories

Unlike Japanese, no close content categories with clear part of speech labels like noun, verb, or adjectives exists in English. Still, some subclasses of verbs seem to be closed TODO: ref.

2.1.4 Functional items

Close function words that still bear clear part of speech tags include personal pronouns, TODO

Close grammatical categories like particles or affixes that don't really need part of speech labels include prepositions (§ 2.9) TODO: list

Division between the two classes is sometimes hard to make. Blurred cases include prepositions that are adverb-like and may be regard as lexical heads, TODO

TODO: combining forms (Huddleston and Pullum 2002, p. 1661)

Huddleston and Pullum (2002, p. 330) introduces the class of determinatives. A full list of determinatives are given by Huddleston and Pullum (2002, p. 356).

2.2 History of the lexical inventory of English

The sources of the lexicon inventory of English also helps understand English grammar.

Derivational affixes are easier to borrow than inflectional ones, since when borrowing a mini-tree from one language to another, the smaller the tree is, the better

[1] On the other hand, in French, for example, liaison exists between a conjugated form of the copula and its complement, as in *c'est impossible*, and thus the morphological word boundary defined by the inflection template is destroyed by phonology.

the compatibility will be. Indeed, some affixes in English originate from inflectional affixes in other languages, but since then are reanalyzed into derivational devices.

2.3 Nouns

Nouns can be found in

Since English still has some inflectional morphology, the class of **countable nouns** can be easily told from others: When we see the single/plural *-∅/-s* alternation, it has to be a countable noun. The class of **uncountable nouns** can be tell from the fact that they appear in similar syntactic environments to those of countable nouns, including adjectival modification, argument positions, as well as their inability to be the predicator of a clause.

Descriptive parameters of the noun includes its inner structure (§ 3.2), its countability and how the plural number is realized (§ 3.2.6), ability to take complementation and modification (§ 3.3), TODO.

2.4 Verbs

The verb class can be distinguished by its inflection (§ 5.2.1) as well as syntactic distribution, including being only able to receive adverb modification.

Most salient descriptive parameters of the verb include inflection (§ 5.2.1) argument structure (§ 5.3), TODO: full list More subtle – but still important – parameters include compatibility with lexical aspects (TODO: ref), and, for complement clause-taking verb, the subtleties regarding interpretation of TAM categories of the complement clause (TODO: ref).

There are several pieces of evidence suggesting that the internal structures of at least some prefix constructions are still available for synchronic analysis. For example, both *set* and *reset* have irregular past tense, and this can be easily explained by postulating that *reset* actually has two parts, the prefix *re-* and the verb root *set*, and since the realizational morphology component of grammar is highly localized, the root *set* shows no alternation of the surface form when the past tense feature appears, regardless of the existence of *re-*.

This can also be explained alternatively by analyzing *reset* as a synchronically unit without internal structure, and explain the *reset* past tense form by analogy to the past tense form of *set*; however, this analogy is not very likely, or at least not universally true: in chemistry, the verb *thermoset* is sometimes back-formed from the adjective *thermosetting* (although *thermoset* is much more frequently used as a noun), and when the past participle of the verb is truly needed, the form *thermosetted* is elicited (D'Amico et al. 2021). It seem that the contrast between *reset* and *thermosetted* is better characterized by whether the root *set* exists or not.

2.5 Adjectives

2.6 Pro-forms

2.6.1 Personal pronouns

2.6.2 Demonstratives

2.7 Numerals

There are four types of numerals in English: the cardinal numerals, the ordinal numerals, the adverbial numerals, and the multiplicative numerals. There is no affixational derivation to show the rank or quality of something (which is attested in the Latin ordinal numeral plus *-āris* derivation): The meaning is conveyed by

Cardinal numerals prototypically appear in NPs, possibly in a

2.8 Adverbs

Adverbs can be regularly formed by adding *-ly* to adjectives; this derivation is terminal: No further derivation is possible after that.

Contrary to the claim frequently seen in teaching materials, adverbials can't be inserted between any two constituents in a clause. When subject-auxiliary inversion is present, for example, an adverbial can't be inserted between the fronted auxiliary and the subject. We therefore have reasons to believe that adverbials appear at structurally licensed positions in the clause, just like other constituents.

(1) Experts occasionally make mistakes.

(2) Occasionally, experts make mistakes.

(3) Do experts occasionally make mistakes?

(4) *Do occasionally experts make mistakes?

Now we do a surface-oriented analysis and check the positions of possible adverbial appearance. Adverbs appear at the initial (5) or at the end of a clause (6, 7), or in the verbal complex (8, 9). The verbal complex positions can be further divided into the positions before or after the first auxiliary (8) and the position before the main verb (9). Similarly – though less apparently – there are also two types of clause final adverbs, one with a pause in utterance, usually shown by a comma (6), the other without a pause (7). The five positions correspond to several structural positions in the clause (Table 4.2). Adverbs appearing at the very initial or the very end of a clause (usually after a comma) (5, 6) may be in the speech act-related positions, or they may be topicalized. Adverbs appearing directly before the verb or after the arguments without a comma (7, 9) are in the manner-like positions, appearing as alternative forms of peripheral arguments (§ 4.2). Adverbs appearing before or after the first auxiliary (9) are in the TAM-related region (§ 4.2).

(5) [Strikingly], he is a liar.

(6) He is quite smart, [frankly].

(7) He finished the task [quite cleverly].

(8) He might [now] be hoping to skip the test.

(9) If you were me, you would have [smartly] answered the question.

As is observed cross-linguistically, the English adverb class is highly heterogeneous; the only shared structural feature is they primarily modify verbs (and sometimes adjectives). The main reason to recognize an adverb class is that *-ly* adverbs appear in all of the three structural positions listed above, and in each structural position, adverbs that are not derived by *-ly* have similar properties with *-ly* adverbs. The adverb class is thus recognized via family resemblance.

2.9 Prepositions

Prepositional phrases can be modified by adverbs, as in *especially in this area*; question: is it possible to use adverbs to modify a case form, say, in Latin?

The preposition class is a highly heterogeneous class. They usually take NP complements, but *as*, a function word usually regarded as a preposition in English, takes copular complements or even incomplete clauses (e.g. *as is known to all*), and preposition stacking is also possible (Huddleston and Pullum 2002, p. 609), Deciding the level of grammaticalization is often a hard task. A preposition taking something as its complement at the surface level may be analyzed as a transitive adverb or even a noun, or a case particle attached to an NP, or – this may seem exotic but this is prevalent in PIE and many older IE languages – the preposition is actually a *modifier* of TODO: find relevant papers on verb particle construction; Latin *ex*+abl. may be seen as an example: the position of *ex* is much more flexible than that of a real adverb, which usually has to appear before its complement.

For prepositions that are truly extensions of the case system we should do away with the term *preposition phrase*, because in this case prepositions aren't lexical heads, but functional heads i.e. markers of grammatical relations (Dixon 2009, § 5.4). Some English prepositions can indeed be seen as such. The most case-like preposition is probably *by*, which is embedded into the rules pertaining to passivization. The genitive preposition *of* appears in the rules about nominalization: the object in a clause becomes a *of*-phrase after nominalization.

Apart from the prototypical case-like prepositions, we have the system of locational prepositions; a comparable system in languages like Finnish is also regarded as a part of the case system. In English this system can take modifiers (10), and has preposition stacking in limited cases (11),

(10) The spot is [[ten meters]$_{\text{degree}}$ [behind]$_{\text{preposition}}$ the house]$_{\text{PP}}$

(11) The sample is collected [from under] the glacier.

An interesting feature of English prepositions is that there exist a huge inventory of fossilized constructions headed by prepositions, like *in order to* and *for the sake of*.

An interesting phenomenon in English is preposition stranding, which happens in two cases: one being passivization, the other being WH-movements and topicalization. Some verbs – but not all – also allow *pied piping* for their preposition complements. Both phenomena need exploration. Preposition stranding is actually typologically rare outside of Indo-European languages, and even within Indo-European languages, not

all languages allow this; a question then is what licenses preposition stranding in English, and what typological parameter it reveals in English. On the other hand, although in many other languages, pied-piping is a *must*, in English it is strongly controlled by the main verb, and is *not* always possible. The parameters controlling the licensing of pied-piping is also a question.

(12) Who(/whom) am I talking to?

(13) To whom am I talking?

For adjuncts, usually pied-piping is the default choice: preposition stranding sounds strange (Sportiche et al. 2013, p. 265). TODO: now we know why there is obligatory P stranding, but why is there obligatory pied piping?

It's said that Bayer and Bader (2007) solves the problem.

(14) In what way could this be possible?

(15) *What way could this be possible in?

2.9.1 Distributions

TODO: copular PP and adverbial PP

2.9.2 Comparison with adjectives and adverbs

The word *worth* has an exceptional property that makes it similar to prepositions: It takes an NP complement directly. The overall properties however are still adjectival (Huddleston and Pullum 2002, p. 607).

(16) The paintings are worth thousands of dollars.

2.10 Morphological typology and wordhood

2.10.1 Compounding

Derivational morphology can be roughly divided into affixation – derivation in the narrow sense – and compounding, corresponding to grammatical category marking and grammatical relation in syntax.

A compound is interpreted as an indication of "aboutness" when it is created for the first time. The word *sickbed*, for example, of course means a bed related to sickness, but most compounds have gained established meanings: *sickbed* means hospital beds.

Another tendency is compounding happens before derivation. In constructions like *acceptability judgement*, compounding does happen after derivation, but it's likely to come from the fact that compounding here is already in the nominal, not the head noun. An alternative analysis is *-ment* is attached after *acceptability judge-* is formed; this possibility is explored in § 2.10.4.

2.10.2 Derivation

In English a relatively clear derivation-inflection distinction can be established: in English clearly inflection always happens after derivation; this generalization needs some refinement in dephrasal derivation (§ 2.10.5). This is partly because English inflection has already been simplified, and doesn't have the opportunity to entangle with derivational devices anyway. Some recent works argue that the *-ly* suffix for adjective-to-adverb derivation should better be considered as an inflection (TODO: ref).

Here I adopt the terminology in Huddleston and Pullum (2002) and call the most general term for units participating derivation a *base*.[2] A root with no category is a primitive base, and a fully derived, ready-to-inflect unit is a maximal base, the latter sometimes being known as a *stem*.

Derivation frequently involves conversion of word class of a base. In English this part of speech switching usually happens between nouns, verbs and adjectives; derivation into the adverb class is finalized and is unable to be followed by further suffix stacking. Both "symmetric" and "non-symmetric" part of speech conversions exist in English. A symmetric conversion is essentially alternation between two possible categorizations of the same category-free root,[3] while asymmetric conversion involves two part of speech labels, one added *outside of* another one (Huddleston and Pullum 2002, p. 1641; Siddiqi 2009, p. 62, (15)), and the derivational affixes are not merely part of speech tags, as is demonstrated by subtle (yet regularly inferrable, not fossilized) meaning differences between *healthful* and *healthy*.

The symmetric/asymmetric distinction seems to be orthogonal to whether the conversion has zero marking. A symmetric conversion can have explicit marking (*speech* N 'the action of speaking' – note that here I'm not talking about the meaning 'an event in which someone publicly speaks'), and a non-symmetric conversion can have zero marking (*attempt* N 'the action of attempting').

English derivation is highly constrained: attested suffix sequences are much less than possible suffix sequences if they were attached to each other freely. Below are some constraints:

- Most derivational devices select categorized bases, and wrong part of speech tags cause incompatibility. Specifically, some affixes – like the adverbal *-ly* – are terminal ones: The part of speech tags carried by them are never accepted by any other derivational devices. Once they are added, no further derivation is possible.

- Some suffixes can only be added directly to the root. The reason is complicated. It might just be that the suffix has already stopped to be productive, and only simplest root-suffix constructions survive until today. In the case of the

[2]The term *base* may refer to a *form* of base elsewhere. In agglutinative languages like Japanese, attaching a suffix to an existing unit may slightly changes its tail (we may say each suffix carries a morphophonological command at the initial dictating this change, which reflects historical inflection of the suffixes as morphologically independent auxiliaries), and there are a finite number of such changes. Thus, we may say "this unit is conjugated into the 2nd base before accepts that suffix". Here the term *base* in Japanese grammar means a particular type of ending of a base in this note.

[3]Still, many roots participating in symmetric conversion have preferred part of speech, so people may still informally talk about a *noun root* or a *verb root* even knowing it appears without a category label.

denominal or deverbal suffix *-ful*, redundancy seems to be the reason: the complex *??-ness-ful*, for example, is extremely rare, because usually *-ness* is attached to an adjective, and the whole sequence *-ness-ful* therefore adds nothing new semantically. The fact that it does appear in manufactured examples, like *awkwardnessful*, to show a sense of cumbersomeness, confirms the above claim that its rarity is largely semantically motivated.

- Some suffixes are only attached to another suffix, like *-ary* in *-ion-ary*. We may postulate that a scheme *X-ion-ary* is stored in the lexicon of modern English as an idiom in derivational morphology, but the single *-ary* is already wiped away.[4]

- Some combinations of affixes are not possible because of realizational reasons: **-ic-ly* is not possible because this combination produces a form that is phonologically hard in English, and there is no vowel insertion rule pertaining to this configuration to ease the problem. We have to add a *-al* between the two suffixes. On the other hand, although *-ize-tion* is also awkward, *-a-* is inserted, and the resulting *-ization* is completely fine.

- We still have some problematic groups of suffixes, the unacceptability of which can't be reduced to any constraints mentioned above. For them the difference between Latinate and Germanic historical origin seems to be the main reason. The rule that Romance affixes appear in more inner positions seems expected, because the Germanic affixes are integrated parts of the grammar and interact naturally with external environments (TODO: coordination of affixes, etc.), while Romance affixes are not, so Romance affixes are unable to appear at the edge between the inner structures of the lexeme and the external morphosyntactic surroundings. (TODO: cross-linguistic evidence) Thus the combination **-ness-al* is not good, because *-ness* is of Germanic origin, while *-al* is of Romance origin.

The tendency that Romance affixes are usually added before Germanic affixes, makes some to propose that English derivation involves two levels, each of which involves a set of suffixes and a distinct set of phonological rules immediately applied after the level is completed. The observations however can be accounted for using traditional and much simpler selectional constraints outlined above (Fabb 1988): once a suffix added to the root is Germanic, no Latinate suffix can be added since Latinate suffixes tend to appear after other Latinate suffixes, so we get a Germanic suffix level after a Latinate internal suffix level, the latter containing

Certain – although highly limited – degree of recursion exists in English derivational morphology. The sequence *-ize-tion-al* is an example. Another frequently noted example is *nationalizationalization* … (Roeper and Speas 2014, p. xvii). This recursive suffix chain can be applied to various verbs ending with *-ize*, not just *nationalize*: the verb *renormalize* is a term in physics, which means to systematically modify the effective values of constants in a theory when throwing away unneeded degrees of freedom

[4] Fabb (1988) assumes a constituent *[-ion-ary]* and then goes on to argue that the mapping from the constituent structure to semantics is not transparent; although as is said below in the example of *acceptability judgement*, the primitives in this analysis are all well attested, such an analysis is *not necessary*, since assuming a constituent tree stored in the lexicon with holes in it is also not something strange, and it happens for idiomatic phrases as well.

(and a formally similar procedure used to solve divergence problems). Thus we have the following chain of derivation:

a) The noun *renormaliz-ation*.

b) The adjective *renormaliz-ation-al*, 'having something to do with renormalization, or is expressed in the theoretical framework of renormalization'.

c) The verb *?renormalizationalize* is usually not acceptable, but this is possibly due to the absence of a feasible meaning. When the highly marked meaning "how to make a theory that is usually not written down in the renormalizational framework renormalizational" is given, the verb form gains acceptability.

d) The noun *renormalizationalization* or even the adjective *renormalizationalizational*, ...

The reason why people decide that such use of derivation is not authentic English is complicated: it probably involves the tendency to reject long words, rejection of replication, or lack of established meaning, or some other factors.

2.10.3 Mismatch between phonological and syntactic wordhood

Mismatch between the realization and the underlying structure is possible in derivation. The structure *[cross-language]-ic* is realized as *cross-linguistic*. Here *linguistic* is not to be interpreted as "related linguistics", but merely a collective realization of the root *language* and the adjectivizer *-ic*, ignoring the real constituency structure.

2.10.4 Licensing of phrasal dependent positions

The verb *judge* has an manner adjunct (as in *she judges this article scientifically, not according to its literature value*), and after nominalization, the manner adverbial is replaced by a manner attributive, comparable to the fact that the object of a verb becomes a prepositional phrase after nominalization of the verb. So *grammatical judgement* mentioned above (§ 2.10.1) is not a single noun, but a nominal created by the nominal complementation construction. The fact that *grammatical judgement* has its established meaning of "judgement of grammaticality" instead of, say, "a judgement written grammatically correctly", exemplifies the caveat that established meaning is not a deciding factor in determining wordhood.

It's tempting to represent the structure of it as *[[grammati-cal] judge]-ment*, and thus the whole construction is to be recognized as a syntactic word (despite being two phonological words). The problem however is that *[[grammati-cal] judge]* is never attested, either as a phrase or as a base.

2.10.5 Dephrasal derivation and definition of wordhood

Dephrasal derivation (17) exist in English. In this example we also find uncontroversial inflection (the comparative *-er*) happens before uncontroversial derivation, but clearly *holier-than-thou* and the inflection step inside is already completely "sealed" when the compound headed by *attitude* is being built, and therefore this doesn't violate the derivation-before-inflection generalization.

(17) His [holier-than-thou] attitude is annoying. (Huddleston and Pullum 2002, p. 1646)

In compounding the branch attached to the lexical head usually is a sub-phrasal unit, and in this case a base, and the existence of dephrasal derivation in English means that whether a nominal appearing in a complex word structure, it's hard to know whether it has already collapsed into a noun base, weakening a possible test of wordhood. In (18), the part *[Distributed Morphology]* contains two orthographical (and maybe phonological) words and may be considered a nominal; whether it has collapsed into a noun base and hence a syntactic word is not clear. (19) is attested in a discussion between two physicists and demonstrates a structure that is very similar to (18) and is rather hard to analyze if we don't acknowledge the status of *phase transition* as a unit within a grammatical word.

(18) Its allomorphy can be accounted for within a [Distributed Morphology-like] framework …

(19) This [phase transition-ish thing] …

(20) Man, I just need one more [elementary school field day]-type day

Marginal cases like *generative grammarian* have both the analysis *[[generative grammar]-ian]*$_{NP}$ and the analysis in which the adjective *generative* is licensed because *grammar* licenses an adjunct about its content and this adjunct position is still valid with the suffix *-ian*. It's likely that this ambiguity is not a theoretical artifact but a source of variance in English.

3 Noun phrase

3.1 Overview

Typologically speaking, the constituent order of English is Dem Num A N (1). A more articulated template of the noun phrase is given in Fig. 3.1.

(1) He was frightened by [[these]$_{\text{determiner: Dem}}$ [three]$_{\text{Num}}$ [ugly]$_{\text{modifier: A}}$ [bears]$_{\text{N}}$].

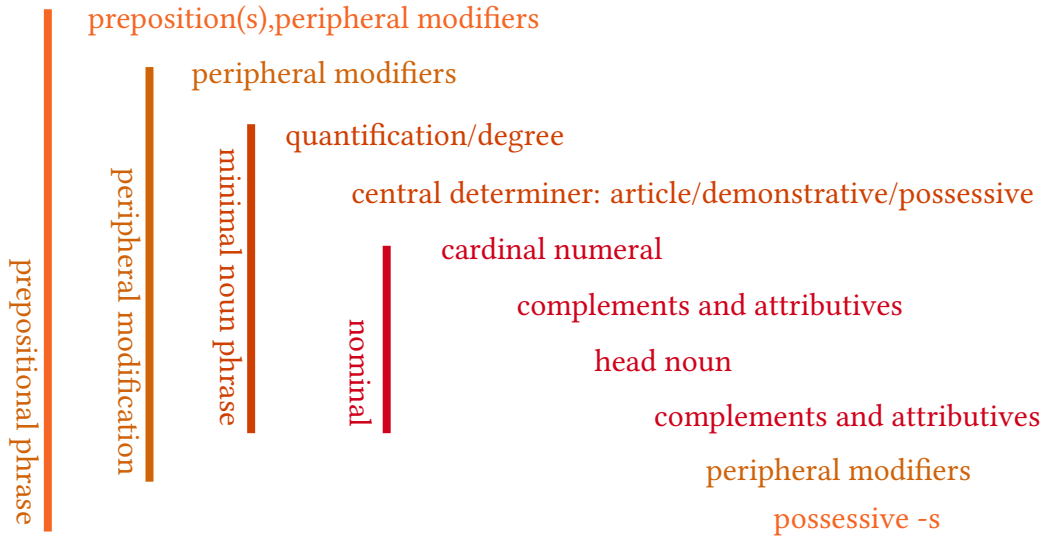

Figure 3.1: The structure of English noun phrase (the indentation means linear order and not constituency relations; the left bars represent constituency relations)

English NPs may have – sometimes must have – one determiner, which may be a demonstrative (*these* in 1), an article (*a* and *the* in 2), or more complicated cases (§ 3.4). The region below the determiner-like region is named the **nominal** (§ 3.3), following the notation in Huddleston and Pullum (2002, p. 329). A nominal plus its determiner(s) (when necessary) is a minimal NP.

(2) This is [[a]$_{\text{article}}$ book about learning Vim in [[the]$_{\text{article}}$ difficult way]$_{\text{NP}}$]$_{\text{NP}}$.

Above the minimal NP region, we still have peripheral modifiers available, like *even* or *along* (TODO: ref). Note that the peripheral modifier *along* appears on the right side of the modified minimal NP. Above peripheral modifiers we have prepositions and the possessive *-s* clitic; a part of them can be conceptualized as a periphrastic case system; other prepositions license subjects and should better be modeled as TODO .

We may also have fused-function constructions (§ 3.7): The demonstrative *these*, for example, may appear as a determiner, but may also appear in a fused-function construction, covering the functions from the head to the determiner (§ 3.7.2).

Although traditionally, complement clauses are regarded as *noun clauses* or *nominal clauses*, this notion is given up in this note, because the inner structure of complement clauses is too different from prototypical NPs, and they are therefore excluded from the class of NPs.

TODO: possibly related literature:

- teacher's desk
- Chapter 1 Reference and Quantification in Nominal Phrases: The Current Landscape and the Way Ahead in Determiners and Quantifiers
- Definiteness By Christopher Lyons
- Layers in the Determiner Phrase by Rob Zamparelli
- The English noun phrase in its sentential aspect by SP Abney
- Laenzlinger (2017): the end part of it!!

Coordination can be applied on the level of the noun (§ 3.2.5), the nominal (§ 3.3.7), and the noun phrase.

3.2 The (extended) noun

3.2.1 Compound nouns

Compound nouns – a compound where the two immediate constituents are all nouns – can be divided into centered an non-centered ones (Huddleston and Pullum 2002, pp. 1646-1648). The semantic relation between the two branches of a centered compound noun is highly diverse; some compounding structures seem to be head-complement or head-modifier constructions, while for others, what relates the two branches is merely "aboutness": when an established meaning is absent, the form A-B means "B that has something to do with A". Non-centered compound nouns (or **dvandva nouns**, which is the term in Sanskrit) are relatively rare compared with the case in Sanskrit.

The noun compounding construction sometimes can be confused with the nominal NP dependent construction. The two are compared in § 3.3.2.2.

3.2.1.1 Conditions for centered noun compounding

The two constituents in a compound noun usually can't bear any inflection. However, irregular plurals are sometimes permitted. This means the threshold of noun compounding has a "lightweight" requirement: The first branch should be categorized as a noun, but *nothing more*. Thus ordinary plurals have a number feature higher than the category feature and are excluded from engaging in noun compounding, while in fused plurals like *mice*, the root, the noun category feature and the plural number feature are all realized into one unit, so compounding is licensed (Siddiqi 2009, § 7.1)

(3) a. mouse-infected
b. mice-infected

3.2.1.2 Complementation and modification

The complementation and modification behaviors of a compound noun sometimes do not clearly originate from its immediate constituents: in (4), the nominal modifier *mango* is clearly not licensed by *ice* or *cream*. Similarly, in (5), the attributive *effective* is not licensed by either *green* or *house*. This doesn't mean the internal structure of *icecream* is already purely historical because *ice* and *cream* are still subject to limited syntactic operations (3.2.5); but it does seem that the subcategorization pattern of a compound noun can be arbitrarily decided *after* the compound is formed and before noun phrasal dependents are attached.

(4) mango ice-cream

(5) a very effective greenhouse

3.2.2 Adjective-noun compounding

An adjective-noun compound is usually a modifier-head structure, but not always: *sick-bed* receives an aboutness interpretation outlined in § 3.2.1, where the adjective *sick* is used as a category-less root, which means *sickness* here.

3.2.3 Verb-centered compound nouns

3.2.4 Deverbal nominalization

Nominalization of a verb either gives a noun about the action or state described by the verb (Huddleston and Pullum 2002, p. 1700), or gives a noun referring to an object or person that is a semantic argument of the verb (Huddleston and Pullum 2002, p. 1697). The two kinds of nominalization are semantically similar to content clauses and relative clauses, respectively. Thus, cross-linguistically, we often see nonfinite forms of a verb serving as its normalized forms, either in the first meaning or in the second meaning mentioned above. In English, the relative clause-like use of participles is rare, but the content clause-like use of participles is highly frequent (6, 7).

(6) [His playing the national anthem]$_{\text{ING-participle complement clause}}$ amazed us.

(7) [His playing of the national anthem]$_{\text{ING-nominalization}}$ is amazing.

3.2.4.1 Zero derivation

the one doing something: *coach*, *spy*
the action of *read*, *go*, *attempt*

3.2.5 Coordination within the noun

Coordination in the noun is rare, but still possible for prefixes (8) and compounding bases (9, 10, 11),[1] with a structure parallel to more frequent coordinated attributives (34). Coordination involving both a constituent in the noun and a constituent licensed

[1] Huddleston and Pullum (2002, p. 449, [24]; p. 450, [27]) claims that such constructions are never possible; but coordination in both branches of the compound noun are indeed attested on the Internet; the low acceptability of these forms in oral English may be due to processing factors: listeners a greedy algorithm and tend to add brackets around, say, *tooth ache*, and thus *back- and tooth-ache* is parsed

in the nominal seems to be much less acceptable (12). Fossilization also reduces the productivity of the in-noun coordination construction (13).

(8) Today we are going to talk about [pre- and post-revolutionary] France.

(9) The chandelier's sleek, spherical frame is draped with dozens of crystal droplets in alternating [tear- and rain-drop] shapes for pure drama …
(from dezignclubstore.com)

(10) …used to treat pain, [back- and tooth ache], amongst other afflictions.
(from knightpiesold.com)

(11) Our sunrise and sunset calculator displays the daily [Sun rise and set] times for the U.S. and Canada.
(from almanac.com)

(12) *two ice- and ten custard-creams

(13) *pre- and description

3.2.6 The number category

The number category is regularly realized as the suffix *-s*, attached to the end of the noun. Some nouns have irregular plural forms, like *foot* (SG) v.s. *feet* (PL),

3.3 The nominal

3.3.1 Internal dependents

As is said in Fig. 3.1, within the nominal region, we first encounter the cardinal numeral region. Below the cardinal numeral we have complements and modifiers;[2] modifiers within the nominal are known as attributives.[3] Complements and attributives may appear before and after the head noun. A generalization is modifiers always appear in more external positions compared with complements; thus a core nominal subregion only containing the complements can be recognized. Relatively fixed orders of the constituents in the pre/post-head complement/modifier regions can also be recognized, although this order may be slightly disrupted by information structure influences (§ 3.3.5).

---

as *back and [tooth-ache]* in the first pass. This effect is reduced in written English and in turn these coordination constructions become more acceptable. Other forms in Huddleston and Pullum (2002, p. 449, [24]) however are indeed not allowed in English (TODO: ref).

Some may explain the existence of this coordination constructions as products of analogy; although this is not wrong, the explanation relying solely on analogy with coordination in the nominal fails to explain why [iii-v] in Huddleston and Pullum (2002, p. 449, [24]) are almost never attested inside the noun, although they are quite frequent in the nominal. Note that this can't be due to semantic reasons (Huddleston and Pullum 2002, p. 450). If the answer is "because *ice-* and *cream* and more tightly linked to each other than *Italian* and *cream* do", then we just get back to the traditional structuralist analysis.

[2] In Liberman and Sproat (1992) the innermost position, which I name as a complement according to Huddleston and Pullum (2002, p. 441), is also called the modifier. They also claim that *store owner* is a compound noun. This analysis is rejected in § 3.3.2.2.

[3] Huddleston and Pullum (2002) seems to define *attributives* as pre-head modifiers in the nominal (Ch. 5, § 14.2); but in other sources we have terms like *postpositive attributives*, indicating attributives appearing after the head noun.

(14) the [three [exotic]$_{\text{modifier}}$ [chess]$_{\text{complement}}$ [boarders]$_{\text{head}}$]$_{\text{nominal}}$

(15) these Chinese store owners

(16) the

Besides its position in the NP, a bare nominal without any determiner can also have limited distribution as a pre-head NP dependent, although not as a full NP (17, § 3.3.2). The cardinal numeral, modifiers and post-head dependents seem to be only licensed in nominals as a part of larger NPs: they are not allowed to appear when the nominal is directly used as a pre-head NP dependent unless the nominal is dephrasalized (§ 3.3.1.1). Both modifier and complement positions can be filled by a nominal, although usually only the latter gains attention (and is frequently confused with noun compounding); when a nominal attributive appears, it still appears before (possibly adjectival) complements (18).

(17) We plan to plant four more [Fuji apple]$_{\text{nominal}}$ trees.

(18) this [[London]$_{\text{modifier (position)}}$ [legal]$_{\text{complement (type)}}$ firm]$_{\text{nominal}}$

3.3.1.1 Recursion complexity

When used as an NP dependent, a nominal can contain another nominal (19, 20), or adjective (21). Several layers of nominal complementation/modification constructions are possible (22). Coordination of nominal NP-dependents is also possible (§ 3.3.7, 36).

On the other hand, it seems when a nominal acting as an NP dependent contains an adjective that is clearly a modifier and not a complement, the whole form becomes unacceptable (23; c.f. 21). Also, there is a strong tendency to rule out post-main noun dependents in an NP-dependent nominal (24). This seems to be a piece of evidence suggesting that the post-head noun region that contains prepositional phrases and relative clauses is only licensed in a full NP.

(19) We planted [an [apple]$_{\text{complement: nominal}}$ tree]$_{\text{NP}}$ yesterday.

(20) [The [[Fuji apple]$_{\text{complement: nominal}}$ tree]$_{\text{nominal attributive}}$ variety]$_{\text{NP}}$ has a reddish-green color.

(21) We are facing a [[medical]$_{\text{complement: adjective}}$ doctor]$_{\text{complement: nominal}}$ shortage.

(22) The [[theoretical]$_{\text{complement: adjective}}$ [[condensed matter]$_{\text{complement: nominal}}$ physics]$_{\text{complement: nominal}}$]$_{\text{compl}}$ community doesn't seem to be willing to write textbooks.

(23) *a [[good apple]$_{\text{complement: nominal}}$ tree]$_{\text{nominal}}$

(24) a. the current [[medical doctor] shortage]
b. the current shortage of medical doctors
c. ???the current [[doctor of medicine] shortage]
d. the current shortage of doctors of medicine

The constraints on post-head dependents and pre-head modifiers are of course not in effect when the nominal in question has been dephrasalized (25). Interestingly, these constraints are also not in effect for nominals appearing as proper names (26, 27); indeed, if *doctor of medicine* in (24c) is regarded as a proper name (we may represent this by capitalizing it), the form is no longer unacceptable, even though it contains a post-head prepositional phrase; if *good apple* in (23) can be interpreted as some kind

of proper names, then the form is also acceptable, despite the pre-head modifier *red*. This fact, together with the fact that constituents within a proper name rarely involve syntactic processes happening outside,[4] is a possible indication that proper names are dephrasalized and have become compound nouns (§ 3.3.2.2) at least in nominal NP dependent constructions.

(25) This state-of-the-art model

(26) [The [Ministry of Defense]$_{\text{modifier: nominal}}$ officials]$_{\text{NP}}$ are having a secret meeting in that room.

(27) Just remember [the [[Red]$_{\text{adjective attributive}}$ River]$_{\text{complement: nominal}}$ Valley]$_{\text{NP}}$, and the cowboy that has loved you so true.

3.3.1.2 The post-head region

The post-head region contains constituents that are too large to appear before the head noun, including prepositional phrases, some adjectives, and relative clauses.[5] Complements licensed by constituents staying in the pre-head region, known as indirect complements, may also appear in the post-head region.

3.3.2 Complements

Nominal complements are much more common than adjective complements, the former in turn is less frequent than one-word nominals in everyday speech.

3.3.2.1 Relation with the head

Generally, if the head licenses a prepositional dependent that has a generic reading

(28) a. management of products
b. product management

(29) quality control

3.3.2.2 Comparison with noun compounding

Unlike the clear distinction between the verb compounding construction and the ordinary verb phrase, the distinction between a core nominal and a compound noun is not straightforward. Semantic criteria are not reliable, since a compound noun can also be semantically compositional, and established meanings are quite frequent for core nominal constructions – either nominal complement construction or adjectival complement construction (Huddleston and Pullum 2002, Ch. 5, § 14.4). The syntactic tests used to divide the two include the follows:

[4] This may also be explained by semantic reasons, i.e. after modification or coordination of their constituents, these forms are no longer recognizable as forms with established meanings. The two explanations are not necessarily contradictory, as the latter explanation provides motivation for the fossilization of these constructions into compound nouns.

[5] Laenzlinger (2005b) contends that the nominal is first assembled, and then prepositional phrases inside are moved out, and then the core nominal – now without prepositional phrases – are fronted, forming the surface constituent order. Interestingly, although this analysis is tedious, Huddleston and Pullum (2002, p. 332, [11]) puts a prepositional complement in a position higher than the position of a modifier, although even at the surface level, complements are always closer to the head noun. This may be understood as an implicit recognition of something similar to Laenzlinger's analysis.

- A compound noun is able to license NP dependents "as a whole" while still having synchronically analyzable internal structures (§ 3.2.1.2), while this doesn't seem to be the case for complex nominals.
- The size constraints in the two constructions are different: noun bases involved in noun compounding can't be further modified, while this is possible for nominal dependent constructions (§ 3.3.1.1). Sometimes however this test is hard to run, since the first element in the construction can't receive any sensible modification or complementation.
- The first base in noun compounding always appears after all dependents in the nominal; but in a nominal complementation construction, the nominal dependent is also very close to the head noun, and therefore sometimes this test can't be run.
- Coordination is much less frequent in the noun compounding than in the nominal dependent construction – but marginal examples do exist (§ 3.2.5).

For prototypical cases, the result of one of these tests implies the result of the rest of the tests; but gradient cases do exist, as is discussed below.

There are several other rather weak criteria:

- Orthographically, a compound noun is often written as one orthographical word; but exceptions are countless: the noun *icecream* is alternatively written as *ice cream*, for example.
- Often, the first element in a compound noun receives the stress. But many speakers pronounce the compound *hotdog* with the stress on the second element (Huddleston and Pullum 2002, p. 451), and most nominal complementation constructions do have the primary stress on the first element.

Some forms fall on the boundary between compound nouns and complex nominals. Proper names like *Ministry of Defense* and *Red River* seem to be dephrasalized at least in the nominal complementation construction (§ 3.3.1.1, 26, 27).

3.3.3 Adjectival attributives

3.3.4 Attributive possessives

Possessive NPs can also serve as attributives. TODO: hierarchical position (Huddleston and Pullum 2002, p. 469)

It's possible – although not obligatory – that the possessor NP is in the singular form, while lacking the article; actually, in this case the article or any other kind of determiner is forbidden to appear. The reason might be semantic: the possessive attributive construction is to express a property of the head noun and therefore the reference of the possessor has to be generic; but a determiner gives the possessor a definite reference, no matter what, and rendering the whole construction invalid. The possessor therefore seems to be the generic usage of the singular form.

(30) There are three [teacher's] books on the desk.

(31) The [poor man's] patent is just a myth.

(32) Informally the role of [teacher] may be taken on by anyone (e.g. when showing a colleague how to perform a specific task). (from Wikipedia)

3.3.4.1 Interpretation of attributives: restrictive and non-restrictive

If an attributive is removed in an NP and the reference of that NP remains the same, we say it's non-restrictive. Non-restrictive attributives are "comments" or "afterthoughts" concerning the reference of the NP. In *members of the music club, [who have developed very close friendship], are all going to the same college*, the relative clause (TODO: ref) is not necessary to decide who the members of the music club are.

If the concept of definiteness were perceived according to Russell's theory of description (Box 3.2), then the distinction between restrictive and non-restrictive attributives would by definition be categorical: Restrictive attributives are about uniqueness of the NP's reference, while non-restrictive attributives are not. The position of this note is however more contextualist, TODO: how we treat Of course, *some* attributives are still bound to be restrictive or non-restrictive for various reasons: If an NP is definite, TODO: why removing some attributives makes a definite NP unnatural?? (TODO: ref).

3.3.5 Ordering and compatibility of attributives

3.3.6 Cardinal numerals

There can be zero or one cardinal numeral in a nominal, and it always appears over all attributives.

(33) I need [ten [green lasers]$_{\text{inside num.}}$]$_{\text{nominal}}$.

3.3.7 Coordination in the nominal

For nominal NP dependents, apart from the coordination constructions already seen in compound nouns (34, 35), greater possibilities are allowed in coordination (36; c.f. 12).

(34) Theoretical and experimental physicists need to work in closer relations.

(35) Experimental physicists and chemists use quite different synthesis techniques.

(36) two Oxford and three Cambridge students

3.4 The determiner and the like

Above the nominal layer, we have a series of determiner-like grammatical functions, filled by various syntactic objects. In this note, I define the prototypical grammatical function of articles, demonstratives, and "determining" possessives as the **central determiner**. I also assume a quantification function over the determiner function to account for expressions like *all the things*, where *all* is the syntactic quantifier and *the* the central determiner.

Box 3.1: About *determiners* and *predeterminers*

Huddleston and Pullum (2002, p. 331) takes a slightly different definition of the term *determiner*, and it also recognizes a syntactic function called the *predeterminer*. The rules seem to be the follows: if there is no article or demonstrative or any other thing that is prototypically a determiner, then the lowest determinative is the determiner; otherwise the article or demonstrative or ... is the determiner, and determinatives lower than it are modifiers. Thus, in *the three cute cats*, *three* is a modifier, while *the* is the determiner (Huddleston and Pullum 2002, p. 356, [4ii]), although in *three cute cats*, *three* is the determiner (Huddleston and Pullum 2002, p. 355, [2ii]; § 3.4.5). In *all vases*, *all* is the determiner, while in *all the vases*, *all* is the predeterminer (Huddleston and Pullum 2002, p. 356, [4i]).

I find the term *predeterminer* unnecessarily complicate the matters, and to say something is a predeterminer tells us nothing about its position in Fig. 3.1 or its possible semantic interpretations. Predeterminers in Huddleston and Pullum (2002, p. 433) all seem to be some kind of quantification, so I just skip the term *predeterminer* in Fig. 3.1, and explicitly inserts a quantifier position.

So, a wise terminology is to replace the term *predeterminer* with *quantifier*. Since some grammars use the term *determiner* to cover all determiner-like grammatical functions, Quirk et al. (1985, p. 253) call the position prototypically filled by articles and demonstratives the *central determiner*. TODO: whether to use this notation

One typological property of English is the prominence position of determiner, as opposed to some other Indo-European languages, like Latin. In Latin the determiner – like a demonstrative – looks just like an attributive: morphologically speaking, it has the adjectival declension pattern (there is no such thing as an article category), and its surface position in the NP is as flexible as other attributives, modulated by the information structure. In English, on the other hand, we have prototypical fillers of the determiner position – articles *a* and *the* – which are inactive in syntactic movements are are usually fixed to their slots in Fig. 3.1.

The description of the determiner region is further complicated by prevalent default semantic configurations. When, for example, there is no syntactic quantifier, the NP still receives some sort of semantic quantification, and TODO: depending on subject or object? Fortunately the organization of the semantic concepts still largely follows the hierarchy in Fig. 3.1 and will be discussed in corresponding sections.

3.4.1 The central determiner and identifiability

The semantic motivation to have a determiner region is to decide what are being referred to by the nominal, the coverage of which can be extremely huge. The speaker may invite the listener to identify the things(s) referred by the NP from the conversational context and/or some uniqueness conditions. The usual syntactic device to send such an invitation is **definiteness**, incarnated in English NP as the central determiner function, and the corresponding semantic concept is called **identifiability**.

3.4.1.1 The articles *the*

There are several degrees of identifiability that may be conveyed by *the*; with low-degree identifiability, the reference of an NP containing *the* is to be decided partially by its inner attributives and partially by the context.

Logical uniqueness is the strongest, but this is usually constrained to mathematical objects (37). Uniqueness from empirical observation or man-made rules is slightly weakened: The article *the* in (38) is almost never replaced by *a*, because for almost all companies, the CEO position is unique.

A further weakened version is uniqueness in the conversational context. In (39), of course the speaker isn't implying that there is only one T-shirt in the world: but if *in the context of the conversation*, there is only one T-shirt, then the sentence makes perfect sense. This gives rise to the famous *a-the* alternation in discourses: When a person or an object first appears, *a* is used, and then the entity is referred to with *the*.

But even this can be loosen: It may be the case that there are several entities in the conversational context that satisfy the conditions, but it's OK to randomly pick up one, and this is still a kind of identifiability.

(37) [The set containing no element] is [the empty set].

(38) The CEO of this company declined to comment.

(39) Pass [the green T-shirt] to me.

Box 3.2: The weakness of the naive theory of description

The fact that definiteness implies identifiability has long been noticed in the study of semantics. Russell's theory of description interprets *the* as a logical symbol ι, and $Q(\iota x P(x))$ means $\exists x(P(x) \wedge Q(x) \wedge \forall y(P(x) \wedge P(y) \rightarrow x = y))$. This is a neat approximation of the definiteness concept, but still has some subtle differences from definiteness in natural languages. When the uniqueness condition is broken, usually we don't say the sentence has a false truth-value: We say it doesn't make sense at all. (Some may develop more delicate logic systems to handle this, but I feel this is not necessary: natural language sentences are more like "commands" in imperative programming than logical expressions, and of course a subroutine can throw an error and give no return value.) The position of this note is theory of description is not one hundred percent correct, and is better replaced by a contextual account (Huddleston and Pullum 2002, p. 368), though the techniques used in the theory of description are of course of great importance: we may, for example, correct the description theory by letting P be the mix of the interpretation of a nominal and contextual information.

3.4.1.2 The abstract generic usage of *the*

(40) I like playing the guitar.

(41) The computer replaces the typewriter.

(42) The kangaroo lives in Australia.

(43) The dermatologist specializes in skin care.

Modification also destroys the possibility of the abstract generic reading. This is also a piece of evidence that there are more fine-grained structure within the nominal, specifically, an "extended noun" (TODO: ref).

(44) *The towel absorbs water.

(45) I like playing the high-quality guitar.
'I like play that specific guitar that has high quality. / *I like play all high-quality guitars.'

3.4.1.3 The subject-determiner possessives

Note that the possessive NP is also a complement or "argument" of the head noun, and in some morphologically rich languages, the possessive NP is indeed marked as an argument (Jacques 2021, § 5.1.2.1). Thus we may say the possessive is the subject in the NP, and it's therefore the subject-determiner (Huddleston and Pullum 2002, p. 467).

3.4.1.4 *We* and *you* as determiners

> **Box 3.3: About *we the people***
>
> Not all personal pronouns appearing at the initial of an NP are determiners. In *we the people*, *we* seems to be TODO: appositive

3.4.1.5 Indefiniteness and numerals as determiners

Indefiniteness is the opposite of definiteness, and occurs whenever identifiability can't be established.

The function of *a* and the function of *one* when there is no other determinative in the NP seems to be the same (Huddleston and Pullum 2002, p. 372), which may be the reason why Huddleston and Pullum (2002, p. 385) claims the numeral sometimes is the determiner. TODO: so what should be my analysis?

It should be noted that there is something asymmetric between definiteness and indefiniteness. A definite NP directly refers to some objects *on its own*, while an indefinite NP doesn't: It may gain specificity from the context (TODO: ref), but itself doesn't have specific reference. Thus, an indefinite NP *always* introduces a bound variable when interpreted and involves quantification, while a definite NP doesn't necessarily involve quantification. The status of indefinite is not the same as the status of definite semantically, and also possibly syntactically (Gianollo et al. 2021; Klockmann 2020)

> **Box 3.4: About the "quantifier-less" interpretation of definite NPs**
>
> The end of Box 3.2 shows that even in the contextualist approach taken in this note, an existential quantifier can still bind the logical variable introduced by a definite NP, so in the first glimpse, definiteness and indefiniteness has nothing different concerning quantification. But for a definite NP, in principle, *immediately after it is interpreted*, it's possible that we don't see any quantifier introduced: The interpretation of a subject NP may just be 'P(concept-to-set(interpretation-of-nominal))', where P is the interpretation of the VP (in the sense of this note). Of course, to *finally* eliminate the function concept-to-set, we still need to use the techniques in Box 3.2 and introduce quantifiers, but it's not the *immediate* result of interpreting the definite NP. On the other hand, for an indefinite NP, its *immediate* interpretation *always* comes with a logical quantifier.

3.4.2 Syntactic quantifiers

Syntactic quantifiers in English are given by Huddleston and Pullum (2002, p. 361, [9]), replicated here:

3.4.2.1 Quantification

In mathematics, quantification is about bound variables, while definiteness is essentially a template which maps a predicate to a set (Box 3.4). In natural languages, we need to note that syntactic marking of quantification can be applied on top of definiteness: We have *all the things I've heard about* and *both the parents*: Quantification can act as a "filter" to further filter what is retrieved from the conversational context. The reverse is not possible, possibly for semantic reasons.

Apart from that, we get the familiar universal-existential distinction (as in mathematics). Note that in natural languages, usually syntactic $\forall$ implicates semantic $\exists$: This may be motivated by the "iterating-over" meaning of $\forall$, where iterating over an empty set throws an error.

3.4.2.2 Collective or not?

One important parameter – although without explicit syntactic coding in English – is the **joint (or collective)-distributive distinction**. The speaker/writer may only talk about a certain part of the objects retrieved from the conversational context (universal), or he or she may talk about all of them (existential). The object denoted by NP may jointly participate in the predication (joint) so that it's not correct to say one object participate in the predicate, or maybe individually (distributive). Ambiguity may occur here. Someone says the student selects four courses – does it mean the student takes the four courses one by one in the course selection system (distributive), or does it mean the student selects the four courses at once (by, say, using the worksheet of the course selection system)? The ambiguity has to be settled by contextual information.

3.4.2.3 The negative polarity items

The *any* family

> **Box 3.5: Is the semantico-pragmatic approach correct?**
>
> One problem is whether we are heading to the wrong direction here. It may just be the case that the distribution of *any*, *anyone*, etc. is determined by syntactic factors. Indeed, usually people consider the misuse of these items a *grammatical* problem instead of a logical or feasibility problem (Zeijlstra 2013, p. 812). *I know the current King of France* is of course wrong, but it's valid; **I know anything about this topic* is grammatically wrong. This criticism however doesn't kill the semantico-pragmatic approach, because the latter focuses on the syntax-semantic interface instead of purely semantic or pragmatic issues. Using the programming language metaphor, misusing negative polarity items doesn't break AST generation, but it does break the compiling procedure because the AST can't be mapped to machine codes. For ordinary people, the machine of grammar contains the syntax proper-LF interface. A comparable example is the English *wa-ga* alternation, which seems to be motivated pragmatically but is usually regarded as a part of the grammar.

3.4.3 Specificity

It should be noted that definiteness doesn't necessarily imply *specificity*, because while any kind of specificity implies identifiability, the weakest sense of identifiability is not

specificity; and specificity – or more generally, identifiability – also doesn't imply definiteness. In *a man was sent to hospital after the shooting*, *a man* usually receives a identifiable reading because of scalar implicature: If I say there is one man sent to hospital, then it's highly unlikely that I mean there are two men sent to hospital. So there is sort of a uniqueness condition concerning *a man*. But still, we use the indefinite article *a*, because the nominal *man* itself isn't enough for us to retrieve who is sent to hospital from the conversational context (while, say, even if Tom's father is unknown to us, we know there is someone – and likely only one – who is his father, so *Tom's father* still fixes the reference).

Box 3.6: Specificity as a syntactic function

Here in English, specificity is purely semantic. Cross-linguistically, the semantic concept of specificity may be realized by a syntactic function higher than the determiner, which also serves as the "NP-inside topic" position (Ihsane and Puskás 2001), which is absent in English.

3.4.4 Referentiality

Another semantic parameter of NPs is whether it's referential, i.e. whether its appearance introduces a new entry in the old information list that can be referred to by a following pronoun. There is no explicit syntactic marking of referentiality in English.

Note that referentiality is not strongly coupled to determination or quantification: Although the reference of an indefinite NP cab never be determined on its own (even with contextual information), indefinite NPs are still referential.

(46) $[\text{Teachers}]_{\text{indefinite NP},i}$ here are expected to be patient. They$_i$ shouldn't give up on a child too quickly.

Non-referential usages of NPs are relatively limited. The cases include negative NPs (TODO: ref), interrogatives (TODO), and meta-linguistic usages as in *["Mary"] is a famous name for girls* (Huddleston and Pullum 2002, p. 400).

Box 3.7: Referentiality as a syntactic function

3.4.5 Is the numeral a kind of determiner?

In Huddleston and Pullum (2002, p. 355, [2ii]) it's said that the cardinal numeral has a determiner function besides its function as a modifier in the nominal when there is no other determiner, because it seems the article *a* is very similar to the numeral *one* without *the*, and therefore a cardinal numeral without a higher determiner appears to be an indefinite determiner. An alternative analysis is to assume that the so-called article *a* in fact a specific cardinal numeral (Lyons 1999, § 2.5).

3.4.6 Ordering and compatibility

The order of all determiner-like elements is highly rigid (Table 3.1). The compatibility between them however shows certain degree of variations.

Table 3.1: Possible values of the determiner-like region

Quant.	Det.	Num.	Nominal
			things
		a/one	*thing*
		two/three/four/...	*things*
all			*things*
all		*three/four/...*	*points*
all	*the/these/those*		*things*
all	*the/these/those*	*three/four/...*	*things*
	we/you		*engineers*

3.5 Peripheral modifiers

3.6 Possessive constructions

Possessive NPs appear in TODO

3.7 Fused-head constructions

A fused-head NP is a nominal constituent in which the function of the main noun – the lexical head – is missing in the surface form, but is undoubtedly an NP with other grammatical relations like complementation still available (47), although usually less variances are allowed compared with ordinary NP: attributives, for example, are not allowed for quantifier-head fusion (48).

(47) a. [[All] of these books]$_{\text{quantifier-head fused NP}}$ tell us nothing about the underlying idea of quantum physics.

b. [All [instances]$_{\text{head}}$ of these statements]$_{\text{ordinary NP}}$ are not valid.

(48) *Interesting all of these books are friendly for beginners.

Fused head constructions have several differences with ordinary ellipsis or TODO

3.7.1 Personal pronouns

(49) *[No car in the race]$_i$ broke down and [it]$_i$ had to be repaired.

> **Box 3.8: Pronouns are not gap fillers**
>
> Pronouns are not residue of NPs moved out, even with coreferential relations. In (49), for example, if the pronoun *it* can be analyzed by the trace left by *no car in the race*, then there is no reason for the unacceptability. This means coreferences are not always generated by movement.

3.7.2 Demonstratives

3.8 Preposition constructions

Also: in Latin it seems the whole English locational preposition system (we may call it a locational case system, if we compare it to, say, Finnish) is non-existent: ablative and dative cases seal the whole projection, and no adjustment to the axial part, etc. Thus in *ex*+abl. actually abl. already contains the full locational case structure, and *ex* is merely here to pin down the complete meaning of *ex*+abl..

Prepositions that are not prototypically locational, like *with* and *of*, are almost never used before other prepositions while they do appear after other prepositions, making them look like a part of the case system, although they themselves govern morphological cases – always the accusative case in modern English – as in *with him*. Note, however, that these prepositions are deeply embedded into the English grammar: *of*, for example, regularly labels the object of a verb in normalization; *with* regularly introduces instruments and manners; *by* also regularly introduces instruments and it introduces the deep A argument in passivization.

The English locational preposition phrase, like similar adposition constructions in many other languages, has a layered (although often unnoticed) structure (Svenonius 2010): *from behind the door* is a legitimate expression, which has two prepositions; forms like *move this from behind the door to in front of the gate* are also attested. This means the preposition phrase can be divided into at least two layers: the PATH layer and the PLACE layer. Furthermore, *behind of* is a frequently attested sequence of prepositions, where *of* may be understood as an analytic case marker (labeling the genitive structural case – TODO), we need an additional case-like layer below the PLACE layer. Between the PATH and the PLACE markers, we may find one degree expression, as in *They came from right between the trees.* Another degree expression can be inserted above the PATH preposition, as in *he goes right to the store.*

We may also stipulate a more complicated internal structure of the PLACE layer: if *in front of*, a PLACE prepositional idiom, is to be analyzed as a compound construction instead of a fully fossilized sequence which is synchronically a single preposition, the final *of* is obviously the case marker (c.f. Latin cases assigned by prepositions), and the sequence *in front* may further be split at least into a LOCATION *in* ('within, not out of') and an AXIAL PART *front* ('a line around the object, here a line before it'). An alternative analysis is to regard the construction as an idiom with a structure similar to *in the name of*; this possibility is explored in § 3.8.1.1. Then similarly, *outside* is $out_{LOCATION}$-$side_{AXIAL\ PART}$. Other possibly explicit realizations of the LOCATION-AXIAL PART-CASE hierarchy include *on top of* (Svenonius 2010).

Yet another position, above the DEGREE layer, is needed to host the relation between the object modified by the whole prepositional construction ("the figure") and the NP complement that the prepositional construction takes ("the ground"): this relation between the figure and the ground is explicitly specified by prepositions like *in* as in *in the silent winter* and *on* as in ; In English this position is never realized individually: whenever the relation between the figure and the ground is focused on, no preposition sequence may appear.

There are particles that seem to be syntactically related to the prepositional construction following it, as in *the boat drifted down from above the dam* and *the boat drifted up from inside the cave*; note that this is not a simple verb-particle construc-

tion like *the soup warmed me up from inside*: in *warm me up from inside*, *from inside* is a manner expression, and *up*, despite being a (metaphorical) directional particle, is not even semantically connected to *from inside*; in *the boat drifted up from inside the cave* though, the particle *up* describes the direction of the motion of the boat which happens in the path given by *from inside the cave*. Hence a DIRECTION layer above the PATH layer is needed; and similarly, another DIRECTION layer above the LOCATION layer is needed, because of *the boat drifted from up above the dam*. It appears that the two DIRECTION particles may appear together, as in *wandered down from up above*.

Some PLACE prepositions like *behind* allow empty complements, while others like *among* don't. Svenonius (2010) proposes a mechanism which, in my opinion, is kind of artificial, to explain this alternation.

It seems that the locational prepositions are combined with the NP more tightly: locational prepositions like *under* can't be stranded at the end of a clause in case of *wh*-movement, but *at*, *on*, etc. can. Huddleston and Pullum (2002, p. 630) summarizes this phenomenon as happening when "the PP is itself complement of a larger PP"; but it seems even when we don't have a construction like *from under which couch* at the end of the clause, stranding is still not permitted.

3.8.1 Prepositional idioms

3.8.1.1 Comparison with *in front of*

Most prepositional idioms lack the ability to be placed after a PATH preposition, and less fossilized idioms also show behaviors that align more with ordinary complex PPs (Huddleston and Pullum 2002, p. 620), highlighting the necessity of a different treatment of *in front of*.

The fact that *in front of* is not really a whole in some cases (the *of* phrase may be omitted sometimes) doesn't need to be explained by stipulating that *in front of* is a real prepositional idiom that still keeps the synchronic *in*-plus-NP structure: the fact can be well captured in the layered preposition construction analysis.

4 The clause structure

Before the following chapters about each part of the clause, this chapter gives a sketch of how these parts are assembled into one.

4.1 High-level concepts

4.1.1 Clause types

We organize this chapter in a top-down manner and start with a classification of clause types in English Table 4.1 is a classification of clauses with one or more morphological verbs; these verbs may include at least one lexical verb or may be zero or more auxiliaries plus a copula. Besides the constructions shown in Table 4.1, there is a further class of clauses – the **verbless clause** (Huddleston and Pullum 2002, p. 1266) – that may be placed into the nonfinite column, but some think it's just a type of sub-clausal phrase. It is largely marginalized when not used as "small-clauses" as in *he considers me strange.*

Table 4.1: Classification of clauses based on independence and finiteness

sentence			embedded clause		
finite				nonfinite	
imperative	"indicative"		subjunctive	infinitive	participle
(1)	(2, 3, 4, 5)	(6)	(7)	(8,9)	(10)

(1) [Go back immediately]$_{\text{imperative}}$!
(2) [Today is a love day]$_{\text{declarative}}$.
(3) [Is it a lovely day]$_{\text{yes-no question}}$?
(4) [What's the weather today]$_{\text{content question}}$?
(5) [It's a lovely day, isn't it]$_{\text{tag question}}$?
(6) The weather forecast says [that today should be a love day]$_{\text{declarative complement clause}}$.
(7) The doctor suggests [that he brush his teeth more carefully]$_{\text{subjunctive complement clause}}$.
(8) I want [to go hiking outside].
(9) I want [for Mary to complete this project].
(10) Do you mind [my opening the window]?

Here we have two top-level grammatical concepts: finiteness (the second title row) and mood (the third title row).

Finiteness The distinction between finite and non-finite clauses is defined here by the ability to be a sentence:[1] a finite clause is either itself a sentence, or in other words, a clause that is an utterance on its own, or is structurally close enough to a sentence, while a nonfinite clause is almost always embedded and structurally deviates from the prototype of sentence structures, and when it does appear as an utterance, it behaves more like an NP that's used as an utterance. Therefore (1, 2) are both finite, and although the bracketed clause in (6) has been subordinated, since it allows the full TAM marking, it is also finite. The subjunctive mood (7; see below for the definition of *subjunctive*) is less typical as a finite construction because it doesn't allow any TAM variation, but since the subject isn't suppressed and is nominative, the structure of it is still "full" compared to prototypical non-finite constructions, and it's also considered a finite construction (Huddleston and Pullum 2002, p. 83).

In (8), on the other hand, the subject is suppressed in the bracketed clause, and if the subject is to appear, *for* appears before the whole complement clause (4) or otherwise we have raising or control constructions; the so-called infinitive clauses (8,9) are therefore not finite. In (10) even the verbal morphology is changed, and it's also nonfinite.

Mood of sentences For English sentences we have the concept of *mood*, a *syntactic* concept that is tightly related to the pragmatic function of the sentence. In English we have the imperative mood (1) and non-imperative moods (2, 3, 4, 5).[2] The distinction between the two can be decided by compatibility with the auxiliary verbs. The former doesn't allow any nontrivial modality and aspect, and the tense is always present, while non-imperative moods interact freely with all TAM categories.

The non-imperative finite mood can be further divided into the declarative mood (2) and several interrogative moods (3, 4, 5). The formation of interrogative clauses only takes two (often skipped) syntactic steps: auxiliary inversion and WH-fronting (§ 4.9.1), which can be attributed to a focus construction which also happens in declarative clauses (TODO: ref). Indeed, some, like Dixon (2005, p. 25), only recognize two moods.

[1] In this note, an *utterance* is a unit spoken by a speaker, while a *sentence* is a "maximal" clause that is an utterance. An utterance doesn't have to be a sentence: It can be an NP, as a concise reply to a question, or even a single word.

Some people use the term *sentence* to cover all utterances. Huddleston and Pullum (2002, p. 45, p. 853) uses the term *sentence* almost as a synonym of *utterance*, and all discussions concerning the syntax in their account of English grammar are about clauses.

[2] Dixon (2009) firmly argues against using the term *mood* for the syntactic marking of modality, while Huddleston and Pullum (2002) uses the term *mood* for the syntactic marking of modality and uses *clause type* to specifically refer to Dixon's *mood*. To avoid confusion (the term *clause type* is too vague), this note follows the definition of Dixon.

The confusion between mood and modality seems to arise from traditional Latin grammar, in which there is no significant difference between a declarative sentence and an interrogative sentence, while there is significant difference between the verbal morphology in indicative and subjunctive clauses. On the other hand, in imperative clauses there is no indicative-subjunctive distinction. Therefore the imperative-non-imperative distinction is fused with the indicative-subjunctive distinction and is named *mood*. This relies on the specificities of Latin grammar and surely is not a universal category for all languages. English also has modal clauses with auxiliaries like *would* or *should*, but that's about modality, not mood. There is indeed a subjunctive clause type in English, but it has already been restricted to complement clauses, and never appear as a full sentence; see the discussion below.

"Mood" of finite embedded clauses Although the term *mood* is prototypical defined as top-level clause structure subtypes that are related to pragmatic speech act, top-level structural variations similar to those are seen in embedded clauses, which are also known as moods in many grammatical works.

In embedded finite clauses the imperative mood is never attested (except as direct reported speech), while we do have embedded non-imperative clauses (6). Beside declarative embedded clauses, we have the category of subjunctive clause (7) which is not nonfinite since its structure is too close to the embedded declarative clause.

Types of nonfinite constructions Nonfinite clauses are deficient in TAM marking and subjecthood (Huddleston and Pullum 2002, p. 1174, [5-7]), and rarely appear as full sentences. In English we have participle and infinitive constructions introduced below, which, if you would like, may also be called moods.

The class of participles contain ED-participles and ING-participles. The class of infinitive clauses can be further divided into TO-infinitives and bare infinitives Huddleston and Pullum (2002, Ch. 14, § 1.4.3). The class of TO-infinitives have three superficial constituent order: *to do sth.*, *sb. to do sth.*, and *for sb. to do sth.* However, the *sb. to do sth.* sequence doesn't correspond to a separate type of infinitive clause: the *sb.* position is always an object position licensed by the verb. Indeed, we never find the *sb. to do sth.* sequence in constructions beside various infinitive complement clause constructions. Thus we only have two types of infinitive clauses: the one without *for* and with a null subject, and the one with *for* and a visible subject.

Nonfinite clauses prototypically appear as complement clauses, but they can also be relative clauses and adverbial clauses (Huddleston and Pullum 2002, p. 1264).

Of course, the inner structure of nonfinite clauses are strongly related to their licensing environments, which we discuss in the next chapter.

4.1.2 Overall template of English clause structure

The template of English verbal clause structure is shown in Fig. 4.1. The figure displays the four rough levels of clause structure. Each layer in Fig. 4.1 as well as justification of them, if not described in chapters above, are described in the rest of this chapter.

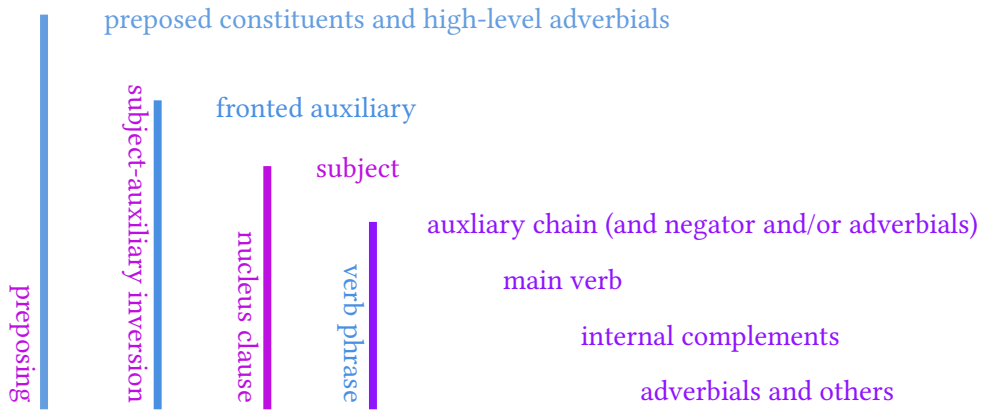

Figure 4.1: English clause structure (the indentation means linear order and not constituency relations)

The highest layer, the layer containing preposed constituents and high-level adverbials, is a "fat one", despite being an optional one: there exist several types of preposing

operations (§ 4.3.3.1), and the layer also contains some high-level adverbials which are about speech force, etc., like *frankly* (§ 2.8). Note that to some extent speech force adverbials can also be seen as a part of TAM marking (Cinque 1999, § 4.4), since they are similar to, say, evaluative adverbials, which usually appear after the first auxiliary and can be seen as a part of the TAM hierarchy. Others consider them close enough to clause type categories and are therefore more peripheral than most TAM adverbs (Van Gelderen 2013, pp. 214-216).

The subject and the position of fronted auxiliary are shown as two separate layers lower than the highest layer shown in Fig. 4.1, because the subject-predicate relation and the subject-auxiliary inversion traditionally gain more attention. A subject plus a verb phrase is a **nucleus clause**, containing the core argument structure, peripheral arguments, TAM marking, and possibly negation; a declarative clause without information packaging operations can just be a nucleus clause without further syntactic operations. Subject-auxiliary inversion happens (§ 5.2.8), which is the case in question formation (§ 4.9.1).

Below the subject is the auxiliary chain possibly mixed with TAM adverbials and at most one negator. After the whole auxiliary chain are manner-like adverbials. Then comes the main verb, which may be a lexical verb or a copula. Internal complements (as opposed to the external argument, i.e. the subject) selected by the verb follow the main verb; prototypical internal complements, like the object without heavy NP shifting, strictly follow the main verb, and no adverbial can be inserted between them and the main verb; but some clausal dependents classified as components indeed can appear after an adjunct (e.g. 61). After the prototypical internal arguments, we find a end-of-clause region containing various clausal dependents, like locational adjuncts (or temporal adjuncts, which are metaphorically locational), heavy objects (so-called **heavy-NP shift**), and complement clauses.

The above auxiliary-main verb-argument-adverbial constituent in Fig. 4.1 containing the verb-argument (core or peripheral) grammatical relations, TAM marking (by inflection, auxiliary construction, or adverbs), and negation is named the **verb phrase** (Ch. 5).[3] The verb phrase can be further divided into two parts: the TAM part, which are discussed in § 4.2, and the extended argument structure, containing all the core arguments and location-, manner- or instrument-like adjuncts (**peripheral arguments** in terms of Dixon (2009)).

The facts listed above about the verb phrase structure may be interpreted as evidence for seeing the post-object region as the residue of the original core verb phrase with possible modification: we can imagine that the object is fronted to a object-only position within the verb phrase (and so is the main verb), leaving a swamp of prepo-

[3]The term *verb phrase* here is used in structuralist tradition as is described in Huddleston and Pullum (2002). Dixon argues against using the term *verb phrase* in the sense of this note; his *verb phrase* is Table 5.2. The two definitions of *verb phrase* are all frequent in modern descriptive grammars. When the term *verb phrase* is used in the sense in Fig. 4.1, Dixon's verb phrase is sometimes called the *verb complex* (Friesen 2017).

Another terminology issue is many people – like Dixon – use the term *predicate* for the syntactic function of the verb complex (i.e. the realization of functional heads), while others use it for the syntactic function of the verb phrase (i.e. a lower part of the TP – see § 1.2.4.1). To avoid this endless confusion, I will just avoid the notion of *predicate* as much as possible. I then confuse *function* (predicate) with *form* (verb phrase). However, English verb phrases – roughly *v*P after case assignment, etc. – almost never appear outside a clause, and it doesn't provide additional information to introduce separate terms for form and function in Fig. 4.1. This is also the practice taken in most works adopting the notion of verb phrase, like Friesen (2017).

sitional phrases, manner-like adverbials and complement clauses behind, which are either not fronted or are too heavy to move, and by nature no adverbials can appear between the main verb and the object (Koizumi 1995, pp. 26-28). Interestingly, a similar analysis is proposed for the clausal word order in Latin (Devine and Stephens 2006, p. 88), although in this analysis, the fronting of the object has information structure motivations.

4.1.3 Clausal dependents and the argument/adjunct distinction

In traditional grammar, clausal dependents in Fig. 4.1, besides auxiliaries, are traditionally divided into arguments (or "complements") and adjuncts; and whatever doesn't look like prototypical subjects and objects are classified as **adverbials**, which may be complements or adjuncts. I respect this tradition as it's a neat way to organize a grammar, and summarize clausal dependents in Table 4.2.

Table 4.2: English clausal dependents

semantic role	syntactic position	
	argument (i.e. complement)	adjunct
prototypical core roles	*[I] loves [that apartment]*	
prototypical peripheral role	*She lives [in that apartment]*	*The machine is fixed with this new tool*
manner, frustrative, etc.	*We were treated [quite badly]*	*He answered the question in a silly manner*
TAM-related adverbials		*I [always] feel tired*
peripheral adverbials		*[Frankly], I think you are fooled by them*

In the extended argument structure part of the table, i.e. in the first three rows, prototypical core arguments, peripheral arguments and oblique arguments representing prototypical peripheral semantic roles allow quite diverse choices when it comes to how to fill them; manner-like phrases – be them oblique arguments or peripheral arguments – allow less variations. It's in this part that we need to discuss the argument-adjunct distinction, or, since the term *argument* is used in Dixon (2009) to refer to all clausal dependents in this region, we may call the distinction the complement-adjunct distinction.

- Prototypical core argument roles, like agent, patient, theme, etc., are always clausal complements and never adjuncts, according to every reasonable standard of the distinction.

- Prototypical peripheral roles, like location or instrument, can also be complements (the light green cell in Table 4.2) or adjuncts (the blue cell in Table 4.2). Temporal adjuncts that give accurate temporal positions or durations which are frequently prepositional phrases are also considered to be in this type at least in English (Cinque 1999, § 1.5).

  They are clearly complements when they are even able to participate in passivization (§ 4.4.3.3), and even if they are unable to participate in passivization, when the preposition introducing them is selected by the verb, it seems they should be classified into the *complement* side. Since the passive *by*-phrase (which is similar to an instrument argument) is directly licensed by the verb, it is also not prototypically peripheral and may be classified as a complement, not an adjunct (§ 4.4.3.5).

But it's also possible to have a location or instrument phrase that's not very related to the verb, which then is to be called an adjunct. It's also possible that for some verbs, a peripheral argument that is structurally not close to the verb is almost obligatory because of the need of a complete event structure (see the discussion below about manner phrases), and the peripheral argument then seems better classified as a complement.

- Manner-like clausal dependents, whether the action in question creates frustration (e.g. *I spent the whole day working on that problem [in vain]*), etc. Following Cinque (1999, § 1.5), they are recognized as a part of the extended argument structure, and not as a part of TAM marking: manner-like clausal dependents, for example, can still be asked about (11), just like prototypical peripheral arguments (12). The main difference between the two types is manner-like clausal dependent positions can be filled by adverbs, like *badly* (c.f. *in a bad manner*); but prototypical peripheral argument positions may be filled by adverbs like *locally* as well.

  They are sometimes like arguments, as in *we were treated [badly]* or *we were treated in a bad way*, without which the clause is not grammatical (the pale green cell in Table 4.2), but more frequently they are adjuncts (the light blue cell in Table 4.2). In English we can't find any *structural* difference between complement-like manner phrases and adjunct-like manner phrases: the obligatoriness is more likely to be a pragmatic or semantic constraint requiring the sentence to have a complete event structure,[4] while at least some locational phrases that are complements do have structural difference with prototypical locational adjuncts (see above).

There are other two types of extended argument structure clausal dependents that include more than one cell in Table 4.2:

- So-called peripheral arguments are either adjuncts with prototypical peripheral semantic role or adjuncts about manner or frustrative expressions like *in a stupid way*, i.e. the first two blue cells in Table 4.2.

- The so-called *oblique arguments* refer broadly to all arguments with oblique case marking, i.e. not nominative or accusative. Their syntactic statuses may appear in all the three green cells in Table 4.2.

There are of course subtleties between the classification of meanings in Table 4.2. The instrument role, for example, may appear in the subject position or as a peripheral argument, while the manner expression is similar to the peripheral instrument argument in their forms (prepositional constructions or "oblique cases") and the ability to be *wh*-extracted (11, 13).

Box 4.1: The term *argument* and the argument-adjunct distinction

Note that here is a terminological confusion: The term *argument* is used sometimes as opposed to more grammatical clausal components like TAM adverbs (as in *peripheral argument*, i.e. any specifier positions that allows large variation with regard to its

[4] c.f. the grammatically important *wa-ga* alternation in Japanese grammar, which comes from information structure constraints.

content), and sometimes as opposed to *adjunct*, i.e. an element that doesn't have very strong relation with the verb. Here we have two descriptive parameters when we talk about arguments: one is the ability of variation (core arguments, peripheral arguments, oblique arguments can be filled by diverse constituents, while TAM adverbials only allow a limited number of adverbs), and the other other is the closeness to the lexical head, which is the main verb here (core arguments, oblique arguments are closely related to the main verb, while TAM adverbials and peripheral arguments are not). The parameter of closeness to the lexical head is the parameter used for argument-adjunct distinction in Table 4.2.

Quirk et al. (1985, p. 732) says we need gradient analysis in cases like *we were treated [quite badly]*. This is correct, but doesn't say much about the essence of English grammar: At a given time, for a given speaker, we can still tell how close the constituent *quite badly* is to the verb. Here, the requirement of gradience comes from the inherent deficiency of the terms *argument* and *adjunct*. The emphasis of the authors on this kind of construction seems to arises from confusion between form and function: AdvPs "should" only occur as adjuncts and not complements, and when they actually appear to be complements, some make-up mechanisms are needed to maintain the generalization that AdvPs are adjuncts.

(11) - [How] did they treat you? - They treated us [quite badly].

(12) - [Where] did they detained you? - They detailed us [in a building near the sea].

(13) - How did you finish this article? - I finished it with LaTeX.

TAM adverbials are adverbials that mark the TAM categories in the way that can be also found in § 5.2 and § 4.2. They are usually quite limited in variation, resembling the tense or aspect system in the verbal complex. As is said above, adverbials like *yesterday* or *in that very moment* (TODO: ref) seem to be peripheral arguments, instead of a part of TAM marking devices, because their syntactic functions allow too much variation and therefore can't be captured by that kind of feature combination ("S before R" or "S=R", etc.) usually seen in TAM devices, and their syntactic distribution also look closer to, say, locational adverbials. Since traditionally they are regarded as a part of the adverb family, I paint them blue in Table 4.2 to show that they are regarded as adjuncts. More peripheral adverbials marking speech acts are also regarded as adjuncts (TODO: coloring). Topics and focuses, which sometimes are considered peripheral adverbials, allow much more variation.

4.1.4 Clause combining

Some adverbials are even higher than the speech act-related adverbials mentioned in Table 4.2, but they are too high to be considered as clausal dependents: many of them are clause linking devices (§ 8.2). Subordinated adverbial clauses may be about cause and result (*now*), concession, and condition (*if* ... *then* ..., with the "reason" clause being semantically irrealis). There are also connective adjuncts like *moreover* or *alternatively*, which refer to *discourse* structures, instead of syntactic structures.

Box 4.2: Adverbial classification in the literature

Different authors have slightly different terminologies concerning adverbials. Huddleston and Pullum (2002, p. 576) put TAM adverbials (except modality adverbials) and peripheral arguments under the class of VP-oriented adverbials, while modality adverbials and speech-act-like adverbials are called clause-oriented adverbials. Dixon (2005, p. 386) on the other hand put all TAM adverbials and speech-act-like adverbials into the category of sentential adverbials (he calls them *adverbs*) and the non-prototypical peripheral arguments are packaged into the class of manner adverbials, though some of them are not really about manner – for example it may be about degree Huddleston and Pullum (2002, p. 576).

Clause combining may happen in each layer represented in Fig. 4.1. (TODO: ref) Linked clauses may appear before or after the main clause. Supplementation and subject-sharing coordination is also not shown in this figure (§ 8.3, § 8.4). Nor is clausal derivation illustrated in the figure (§ 4.3.3), because arguably, some post-internal complement adverbials are likely to be the result of heavy constituent postponing (TODO: ref). Apart from the above constructions, the scheme illustrated in Fig. 4.1 works for all clause types (14, 15, 16), including nonfinite clauses, though for the latter, the properties of the subject and the allowed auxiliaries deviate from the finite case, and this is also the same for allowed preposing constructions. (14) is a fused relative clause, in which there is WH-fronting but no subject-auxiliary inversion (TODO: ref). In (15) we see two preposing constructions, one topicalization (TODO: ref) and WH-movement for question formation (TODO: ref), and the only verb – the copula *is* – is moved out of the verb phrase because of subject-auxiliary inversion. The

(14) $[[\text{What}]_{i,\text{WH-preposed: WH-pronoun}}\ [[\text{Max}]_{\text{subject:NP}}\ [\text{said Liz bought} -_i]_{\text{verb phrase}}]_{\text{nucleus}}]_{\text{WH-preposing}}$

(15) $[[\text{In your opinion}]_{\text{topicalized}}\ [[\text{what}]_{i,\text{WH-preposed}}\ [\text{is}\ [-_i\ \text{the most dangerous}]_{\text{verb phrase}}]_{\text{SAI}}]_{\text{WH-preposing}}]_{\text{topic-preposing}}$

(16) $[[\text{what}]_{i,\text{WH-preposed}}\ [\text{to}\ [\text{do} -_i]_{\text{verb phrase}}]_{\text{nucleus}}]_{\text{WH-preposing}}$

4.2 TAM marking

TAM categories realized in Table 5.2 contain tense, aspect, and modality[5] ; besides the morphological and periphrastic conjugations mentioned in the table, these categories can also be marked by adverbials. Besides these categories, English also has several speech-act related categories[6] that are marked purely by adverbials, as well as lexical aspect categories. Below is an overview of these categories attested in English; note that the scopes of them don't always follow the order in which they are introduced.

4.2.1 Top-level categories

Categories in this section involve the modality slot and the tense inflection in Table 5.2. Apart from the verbal complex, categories in this section also involve adverbs,

[5] A clear classification of TAM categories under labels like tense, aspect, etc. is always controversial and is not attempted at in this note. Common terminological controversies include whether the future category in English should be regarded as a modality category (TODO: ref),

[6] These categories are sometimes known as *moods* (Cinque 1999), and this goes in conflict with the convention to use the term *mood* to refer to clause types like "interrogative".

which seem to have similar height with the subject; the speech act adverbial almost always appear before the subject; but the order between the subject and the rest three adverbial positions is more flexible.

4.2.1.1 Speech act

The category of speech act is about the additional information the speaker attaches to the clause to show the "purpose" of this clause. It is solely marked by adverbs, usually with no modification. The speech act adverb is almost always before the subject (17). Huddleston and Pullum (2002, Ch. 8, § 18) gives a list of speech-act related adverbs. Note that some speech-act adverbs can also appear in the manner position, including *frankly* in (17).

(17) [Frankly], I think you plan will turn out to be a complete diaster.

4.2.1.2 Evaluative adverbials

Evaluative adverbials are attached to the clause which is presented as a fact to be evaluated by the speaker. A list of evaluative adverbs can be found in Huddleston and Pullum (2002, p. 771, [2]).

(18) [Fortunately] they managed to escape the country the day before the coup.

(19) [Obviously], the people involved in this incident were deeply shocked and provided no useful information.

(20) [Quite obviously], many find an English degree almost useless for their careers.

4.2.1.3 Evidentiality

The category of evidentiality marks how the speaker knows the fact being talked about. English lacks verb inflection related to evidentiality; the category is completely marked by adverbs with no modification.

(21) They [allegedly] stole a valuable painting from the local museum.

4.2.1.4 Epistemic modality

Parameters of modality categories include kind (epistemic, alethic, "root" modality) and strength (necessity v.s. possibility). Epistemic modality orients from the judgement of the speaker towards a given event (22, 23), while alethic modality is about possibility of an event (24; § 4.2.1.7), and root modality is about moral obligation, ability, etc. (25). It should be noted that although both alethic modality and root modality are about the tendency for an event to happen, the latter describes intrinsic properties of the subject and to some extents is stative (TODO: really). Thus, in some dialects the two may be stacked, with the possibility category modifying the ability category (27).

(22) The crime boss [should]$_{\text{epistemic: weak necessity}}$ already be in New York City – let's go and catch him.
'We the police have gotten the daily schedule of the crime boss, and from this we expect he to arrive at New York City today …'

(23) You [must]$_{\text{epistemic: necessity}}$ be kidding!
'I'm quite sure that you are kidding.'

(24) You [must]$_{\text{alethic: necessity}}$ combine hydrogen and oxygen in a ratio of 2:1 to form water.
'The only possibility that you form water is that you combine hydrogen and oxygen in a ration of 2:1.'

(25) He knew he [should]$_{\text{obligation}}$ go to New York to testify in a court.
'To arrive at New York City to testify in a court is his legal and moral duty.'

In Standard English these modality categories have no superficial structural differences, and only one modal auxiliary is available in the verbal complex; but in some variants of English modality categories can be stacked, and hence the hierarchy in which the possibility and ability categories are below the tense categories and the epistemic category is explicitly displayed (Cinque 1999, pp. 78-79; 26, 27).

(26) % He [should]$_{\text{epistemic}}$ [can]$_{\text{ability}}$ do it.
'He likely can do it.' (Scots English; Cinque 1999, p. 54)

(27) % He [will]$_{\text{future}}$ [might]$_{\text{possibility}}$ [could]$_{\text{ability}}$ do it for you.
'He might be able in the future to do it for you.' (Hawick Scots; Cinque 1999, p. 79)

The parameter of kind and the parameter of strength are not orthogonal; necessary epistemic modality and possible epistemic modality, for semantic reasons, can't be stacked; things may be different for other kind of modal categories (§ 4.2.1.7). TODO: example of auxiliary stacking

4.2.1.5 Tense: past and future

Vikner (1985, see also Cinque (1999, p. 82)) proposes eight prototypical tenses for a group of unrelated languages, including English, as a modified Reichenbach account of tense. There are three parameters: past or not, future or not, and anterior (i.e. "perfect" in traditional English grammars) or not. The first two parameters are closer to each other Shlonsky (2010), and may be fused into one in some languages, as in Latin. The third parameter is closer to aspectual categories (§ 4.2.4.1).

4.2.1.5.1 Past or not The past category means a past reference time R_1 (not necessarily the event time – see below) is before the "speech time" S, while in the non-past case, S is the same as R_1 (28, 29). S usually is the real speech time, but the following variances exist:

- S is the encoding time, i.e. the time when a speech/written text is made (30).
- S is the decoding time, i.e. the time when a speech/written text is understood (31).
- S is "suppose we are in this case" time. In some novels and historical recounts, all tenses are simple present, which means the time of orientation is placed to the (historical or even imaginary) time points in question (32).

When the S time has the "suppose we are in this case" meaning and indicates a real or imaginary speech time, which is true for all the cases mentioned above, it is said

to be **deictic** and refers to the **deictic time**; when such a deictic time is missing, either because of specific usages of a TAM construction (TODO: ref) or because of the absence of S in a non-finite clause, we say the clause is non-deictic.

(28) I [like]$_{\text{non-past, non-future, non-anterior}}$ this book.
'My liking the book (E = R_2 = R_1) happens spontaneously with the speech time (S)'.

(29) He [went]$_{\text{past, non-future, non-anterior}}$ to London yesterday.
'His going to London (E = R_2 = R_1) happens *before* the speech time (S).'

(30) Today is a nice day. (on a diary page)

(31) You are now leaving New York City. (on a sign board)

(32) My Lady Dedlock has returned to her house in town for a few days previous to her departure for Paris. (*Bleak House* by Charles Dickens)

4.2.1.5.2 Future or not The parameter about the future-present distinction is the relation between R_1 and yet another reference time R_2 (33, 34). In a future TAM category, R_2 is later than R_1 (33, 34), while in a non-future TAM category, R_1 and R_2 are the same, as is shown in examples above.

(33) I [will]$_{\text{non-past, future}}$ [graduate]$_{\text{non-anterior}}$ next year.
'My graduation (E = R_2) happens *after* the speech time (S = R_1)'.

(34) He [would]$_{\text{past, future}}$ probably believe this claim but for some reason he didn't.
'His believing this claim (E = R_2) happens *after* the reference time R_1, which happens before the speech time S.'

4.2.1.5.3 Anterior or not The final parameter is the relation between R_2 – some kind of future time in a future tense, or the same as R_1 in a non-future tense – and the event time E. When E is before R_2, it's an anterior tense; otherwise E is R_2. The interpretation of R_2 and E is heavily entangled with aspectual categories: we may regard R_2 as the "aftermath", and thus an anterior tense gains the usual "present prevalence" reading; but it's also possible to simply regard the relation that E is before R_2 as a realization of the semantic past, and hence the plain anterior (non-past, non-future) is identified with the plain past (non-future, non-anterior).

(35) I have completed this project.

Box 4.3: Alternative to the S, R, E notation

In Huddleston and Pullum (2002), the labels S, R and E are not used; labels like T_o and T_r are used to give the relative relation between a pair of time points; thus in simple past, S is labeled as T_o and R=E is labeled as T_r (p. 141, [4iii]), and in present perfect, when we compare S with R, S is referred to as T_o^1, and R is referred to as T_r^1, and when we compare R with E, R is referred to as T_o^2 and E is referred to as T_r^2 (p. 141, [4i, ii]). When S is missing in non-finite clauses, when R and E are compared, R is then T_o and E is T_r (p. 140, [2]). The main problem of this system is it implies the possibility that we are able to have unbounded number of orientation time-reference time pairs, while in fact we only have three.

4.2.1.5.4 Tense categories in the verbal complex With the parameters above, the tense system of English can be classified by Table 4.3. It should be noted that despite having a seemingly combinatorial tense system, these forms are not really orthogonal to other TAM categories. The English past tenses, for example, occasionally have modal meanings, and the perfect tenses are compatible to continuative aspect of stative verbs, while the imperfect tenses are not. The usage of English verbal tense marking is elaborated in § 4.2.6. The future tense marker *will*/*would* in English has clear usages which belong with other modal auxiliaries, and thus is often discussed not in the tense system, but as a marker of modality (Huddleston and Pullum 2002, p. 209).

Table 4.3: The tense system of English

	not future		future	
	not past	past	not past	past
not perfect	*work*	*worked*	*will work*	*would work*
perfect	*have worked*	*had worked*	*will have worked*	*would have worked*

4.2.1.5.5 Past and future adverbs Adverbials related to the past and the future include are highly limited. The most frequent adverbs include *once* and *then*. TODO

4.2.1.6 The realis-irrealis distinction

A top-level realis-irrealis distinction can be observed in languages like Latin, where a clause with a strong modal meaning has different morphosyntactic appearance compared with a clause about this world, regardless of the kind of modality. This distinction can also be seen in the subjunctive complement clause construction in English (§ 7.1). Still, in English we have adverbs marking

(36)

4.2.1.7 Modality about possibility and necessity

This section is about alethic modality, i.e. modality about possibility and necessity of an event.

In some languages alethic possiblity and alethic necessity can appear together cross-linguistically, as well as moral obligation and moral permission (Cinque 1999, p. 80); in

(37)

4.2.1.8 Compatibility

To summarize, top-level TAM categories in English involve both the first slot in the verbal complex, namely the modal slot (Table 5.2), and TAM adverbs. Speech act-related adverbs, evaluative adverbs and evidential adverbs combine freely with the rest.

Categories in this section are not available in all non-finite constructions in English. TODO: really?

4.2.2 Habitual, repetitive and frequentative aspects

In English, lexical aspects, i.e. the natural organization of an event,[7] don't have explicit marking, but can still be told by licensing of temporal adverbials (TODO: ref), interpretation of TAM markings (TODO: ref) and TODO: full list

TODO: for some reasons, these categories seem to be (at least mildly) incompatible with the perfect tense and the like

4.2.2.1 The habitual

Note that although the habitual aspect is used without other TAM markings, that's not always the case. Habitual aspect plus perfect tense, for example, is attested.

(38) I $[\text{usually}]_{\text{habitual}}$ walk to school.

(39) Authors of English textbooks have usually preferred other terms.

4.2.2.2 The repetitive and the frequentative

TODO: position of the volition modality

4.2.3 TODO: delayed, predispositional, volition, celerative

My guess would be that *quickly* is actually below *still*

4.2.4 The anterior (perfect) tense and related aspect categories

There are several aspect categories related to the relation between the point-like event time E and the complete situation time.

4.2.4.1 The anterior or "perfect" tense

4.2.4.2 Terminative

The terminative aspect is marked by the adverb *no longer* in English. It means that the event time E is slightly before the situation time: the situation is no longer the case at E. TODO: relation with perfect

4.2.4.3 Continuative

(40) I [still] don't know why Yale does so awfully in science.

TODO: relation with perfect

4.2.4.4 Perfect/imperfect aspect?

TODO: relation with the perfect

[7] Note that the term may be misleading: the lexical aspect of a clause isn't completely determined by the lexical properties of the main verb; changing of the lexical aspect, just like changing of valency, is possible.

4.2.4.5 Retrospective

The retrospective aspect is a

(41) I just finished my homework!

4.2.4.6 Proximative

(42) They soon will be regretful.

4.2.4.7 Progressive or imperfective

A restricted version of the concept of composition – how the inner makeup of an event is represented (Dixon 2012, § 19.10) – is marked by the so-called plain-progressive distinction in English. A clause representing a state instead of an event is never progressive (TODO: ref: lexical aspect), even if it's usually semantically imperfective.

> **Box 4.4: Perfectivity**
>
> Whether perfectivity should be listed as a separate level in Cinque's hierarchy of TAM categories is not clear. If perfectivity is to be listed as a separate category, its position should be similar to that of English progressive Pearce (2015), but since the imperfective aspect can also be regarded as a catch-all term for the habitual, the inchoative and the progressive (Van de Vate 2011, p. 255), it seems equally possible to regard the imperfective aspect as a secondary concept. I will not discuss which analysis should be taken.

4.2.4.8 Perfect's interaction with other TAM categories

Although at the first glance, the auxiliary *have* is just a marker of the perfect tense (Table 4.3), it's intermingled with other TAM categories in the follows.

The perfect tenses sometimes implicitly have a continuative (§ 4.2.4.3) reading (43; Huddleston and Pullum 2002, p. 141). This implicitness seems to be *obligatory*: appearance of a continuative adverb is not acceptable (44); it seems the continuative aspect has already been integrated into the perfect tense in this case. When the continuative reading is correct, the clause is perfective in its meaning and has virtually no meaning change when the progressive aspect legally appears (45).

(43) a. She has lived in Berlin ever since she married.
'She was in Berlin and still is in Berlin until today.'

b. She has lived in Berlin.
'She was in Berlin and this event probably has some relevance today; but I don't know whether she's still in Berlin.'

(44) *She has still lived in Berlin.

(45) She has been being living in Berlin.

On the other hand, the *terminative* aspect (§ 4.2.4.2) can also be fused into the perfect tenses (46). Again, the terminative adverb is forbidden to appear (47).

(46) They have lived!
'They have lived their lives and now they are no longer live.' (translation of Cicero's exclamation after suppressing the Catiline conspiracy)

(47) *They have no longer lived.

In this usage the English PERFECT is a distorted version of the completion concept – whether the time of an event is before the time of narratives defined by tense (Dixon 2012, § 19.7). Since it's also possible that the event time in the English PERFECT is the starting time of the event in question (43), not the finishing time, the term *completion* is not generally good for English PERFECT.

The retrospective adverb and the proximative adverb can also appear with the perfect.

(48) The conversation had soon moved on.

(49) I have just finished my homework!

4.2.5 Root categories

4.2.5.1 Obligatory modality

4.2.5.2 Frustrative

4.2.5.3 Completive

4.2.6 Details about English verbal primary tense system

4.2.6.1 The perfect tenses

The PRESENT PERFECT is the only available tense category in all non-finite forms, if it's recognized as a tense at all; in this usage it always has a non-deictic meaning.

4.2.6.2 Back-shift usage of past tenses

(50) a. He said "I need to go to New York immediately".
b. He said that he needed to go to New York immediately.

4.2.7 Special usages of modal auxiliaries

4.2.7.1 *Would like*

4.2.7.2 *Should have*

Since *have* almost never receives focus and is usually reduced to *-ve* in *should've* and *could've*, many native English speakers tend to misspell *should have* as *should of* when paying no attention to the orthographical convention.

4.3 Copular clause

The claim that *be* as the copula only realizes the tense information and can be further supported by the fact that TODO: *?as known to all*, etc.

4.3.1 Clausal continuation

TODO: *not even* construction, heavy NP shift, etc.

4.3.2 Logic and default information structure

4.3.2.1 Descriptive parameters

TODO: position of quantifiers for NPs with or without determiner, and the difference between *all*, *every*, *some*, *any*; the position and scope of negation; how the position of NPs (subject or object) influences quantification; the relation with information structure (if a subject NP is never given, it tends to be read as a representative of its kind and is therefore bound by ∀; but it's never the case for objects)

4.3.3 Clausal derivations

(TODO: heavy NP shift, final adverbial with a pause)

4.3.3.1 Preposing

(§ 4.5)

4.4 Clausal dependents in verbal clauses

In this section I discuss the first three rows in lower clausal dependents in Table 4.2. Although there can be multiple clausal dependents appearing after the main verb, detailed observation reveals that the innermost complements are highly restricted: when the verb takes internal complements, either we have an object, or we have a "small clause" which may be predictive, directional, locational, or possessive; internal complements, including particles, direct and indirect objects and copular complements, are all first licensed in the small clause construction. We do have particles that are not a part of a small clause, but they seem to form a compound with the main verb, and strictly speaking, are closer to the verb than prototypical clausal dependents are.

4.4.1 Object, and "unaccusativity"

4.4.1.1 Descriptive predicative constituents

(51) She ate the eggs raw.

4.4.2 Dative and beneficiary ditransitive constructions

Most ditransitive verbs are about giving and receiving, and therefore they have A, G, T arguments (G = goal-like, T = theme-like), where the subject in active voice is the A argument. In English there are two major verb frames for this: the *give sb. sth.* structure and the *give sth. to sb.* structure. Although it's tempting to analyze this alternation using transformational rules, or in other words, we have a clearcut G-to-G and T-to-T correspondence in the two constructions, a deeper look reveals that the two verb frames have their differences rooted in the deep argument structure. Both G and T have some similarities with the monotransitive object. The question is whether G and T show relatively uniform syntactic behavior between the two verb frames (so that they can be regarded as labels of clausal complements, rather than just labels of semantic roles).

Box 4.5: Mismatch between semantics and syntax

On the other hand, in English, the contrast between the experiencer role and the stimulus role: doesn't have much syntactic significance, since some verbs' experiencers work like some other verbs' stimulus. TODO: ref

4.4.2.1 Passivization and the extended argument approach

Dixon (2009) analyzes the English ditransitive construction with labels O and E (i.e. the less object-like internal complement) because of the behaviors of the arguments in passivization. (6) and (7) in Dixon (2009, § 3.3) with semantic role labels replaced by ones in (Huddleston and Pullum 2002, § 4.2.2) are shown in (52, 53). The NPs *all his goods* and *his facorite students* are assigned the label O, and the PP *to charity* and the NP *some books* are named E arguments. The first two cannot be promoted to the subject position in passivization, while the latter two can.

(52) John gave [all his goods]$_{\text{O, theme}}$ [to charity]$_{\text{E, goal}}$

(53) John gave [his favorite student]$_{\text{O, goal}}$ [some books]$_{\text{E, theme}}$

The NP *all his goods* is a theme while *his favorite student* is a goal, but they have similar syntactic properties in passivization. So it seems the division between O and E is useful in English while the division between G and T is not if passivization is all we are curious about: no goal-like and theme-like semantic role classes with stable syntactic appearance can be established.

Another phenomenon supporting the claim that *all his goods* in *give all his goods to charity* is the direct object is from the constituent order. It's possible, although rare, to swap the order of the goal argument and a manner phrase, effectively inserting the manner adverb between the two arguments.

(54) All I did was take every ounce of energy I had and give it unwillingly to you. (from https://tomblog.rip/a-thank-you-letter-to-the-unfaithful-ex-boyfriend/)

4.4.2.2 The direct object/indirect object approach

However, there is also evidence for a stable G-T contrast in syntax and not just semantics. This is the approach taken in Huddleston and Pullum (2002). Applying criteria in Huddleston and Pullum (2002) § 4.4.3, (52) and (53) are to be labeled as (55, 56). (55) is labeled according to the rule that indirect objects are nowhere to be found in a clause without a direct object, and (56) is labeled according to Huddleston and Pullum (2002, § 4.3 [8]). If these labeling rules stand, then a stable syntactic contrast between G and T can be established: the T argument is always the direct object, while the G argument is less object-like: it is either an indirect object or a prepositional complement.

(55) John gave [all his goods]$_{\text{O}^{\text{d}}\text{, theme}}$ [to charity]$_{\text{prepositional complement, goal}}$

(56) John gave [his favorite student]$_{\text{O}^{\text{i}}\text{, goal}}$ [some books]$_{\text{O}^{\text{d}}\text{, theme}}$

Huddleston and Pullum (2002, Ch. 4, § 4.3) provides ample argumentation for the analysis shown in (55) and (56). Though *all his goods* and *his favorite student* have the passivization and constituent order properties of typical monotransitive objects, only *all his goods* enjoys the full range of monotransitive object properties. In the

ditransitive example (56), *some books* behaves like the monotransitive object in object postposing and preposing, the predicative adjunct construction, and controlling. Thus the position of *some books* has more object properties than the position of *his favorite student*.

4.4.2.3 Small clause analysis

John gave all his goods to charity contains a directional small clause: *all his goods (be) to charity*. A similar usage can be seen in Latin, as in *Deo gratias* 'thanks (be) to God'. On the other hand, *John gave his favorite student some books* may be analyzed as containing a small clause like *his favorite student (having) some books*. This leads to an interesting subtlety of the meaning of the two verb frames. Consider (57, 58) from Devine and Stephens (2006, p. 81): (58) is somehow unnatural in its meaning, because its small clause is *a new nose (be) to the film star*, as if the nose is already somewhere before the surgery.

(57) The plastic surgeon gave the film star a new nose

(58) The plastic surgeon gave a new nose to the film star

Another piece of evidence supporting the small clause analysis is the interaction between particles and ditransitive constructions: a particle that usually appears in directional constructions is never licensed in the *give sb sth* environment; it however appears in the *give sth. to sb.* environment (§ 4.4.5).

4.4.2.4 Control in purpose infinitive

The second argument in the *give sb. sth.* construction is both semantically and structurally patient-like in the possessive small clause construction (*I give you this* = 'I cause you to have this'); this is reflected in the control relation in the purpose infinitive construction (Huddleston and Pullum 2002, p. 250).

(59) She gave him$_i$ it$_j$ to $-_i$ spend on his children $-_j$.

4.4.2.5 Descriptive secondary predicates

See Depictive secondary predicates, light verb give, and theories of double object constructions which proposes that the small clause analysis is wrong.

4.4.2.6 The typology of G and T

To summarize, no complement type can be merged with another in the three types of VPs having been considered now, namely V-O, V-G-T, and V-T-pG, except the T argument in a V-T-pG clause and the monotransitive object. Fig. 4.2 visualizes the resemblance of these complement types with the prototypical monotransitive direct object according to the six criteria listed in The Cambridge Grammar of the English Language (CGEL) § 4.4.3. The T argument in V-T-pG has almost identical properties with O in a V-O construction. Both of them, therefore, are put under the category of monotransitive object. For the rest three complement types, G in the V-G-T construction is similar to the monotransitive object in terms of passivization and constituent order, though when the G is the beneficiary, passivization is marginally acceptable

(CGEL § 4.3.3, [10ii]), so in Fig. 4.2 I slightly lower the G in V-G-T point. (And for the same reason, since T in V-G-T can be marginally passivized – see CGEL § 4.3.3, [10iii] – the T in V-G-T point is slightly raised.) If passivization and constituent order are considered as the decisive factors to distinguish complement types, then we get the classification in § 4.4.2.1. On the other hand, if the rest four factors are emphasized, then the classification in § 4.4.2.2 works, where G arguments have uniform behaviors and so do T arguments, and we find T arguments are largely similar to O arguments. When all the six factors are considered, the total distance between the monotransitive object (i.e. the O argument) and T in V-G-T is smaller than the distance between the monotransitive object and G in V-G-T, and the former two are clustered into the *direct object*. Finally, the direct object is clustered with G in V-G-T – or the *indirect object*, correspondingly – and we get the final object concept. G in V-T-pG is far from the rest four points and hence is placed out of the range of objects.

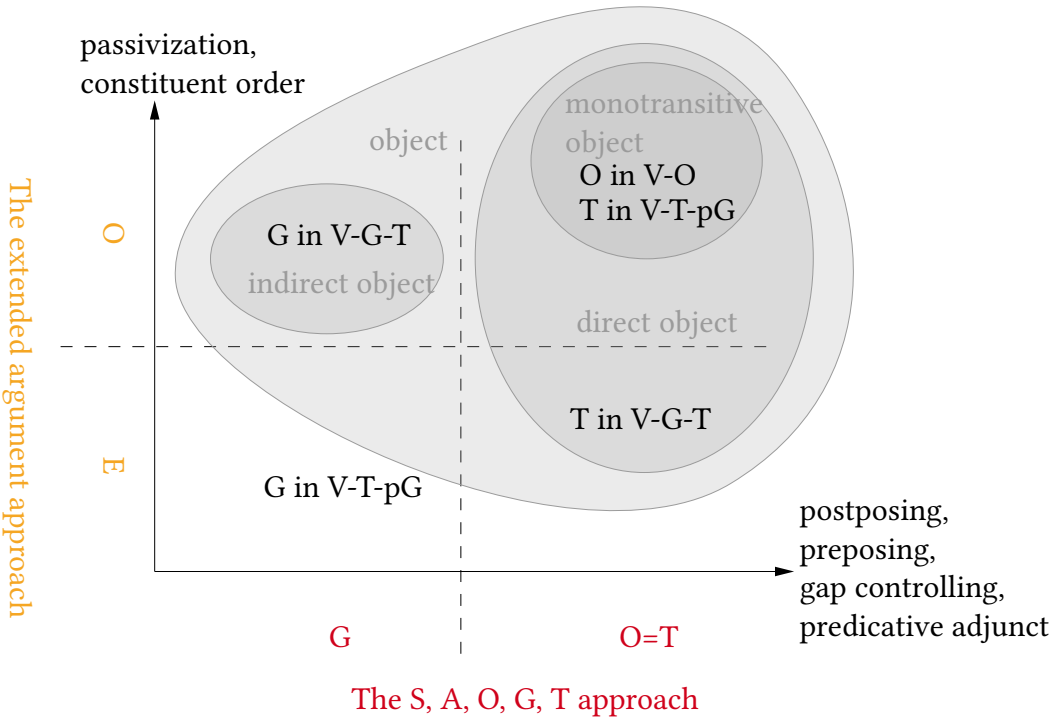

Figure 4.2: Classification of internal complements in English V-O, V-G-T, and V-T-pG clauses. The orange labels are discussed in § 4.4.2.2, and the red labels are discussed in § 4.4.2.1. The grey blobs indicate clustering of the points.

So here we see why *object* is a useful concept in English grammar, at least among V-O, V-G-T, and V-T-pG clauses: it can be defined both via form (i.e. no preposition and not predicative) and via function (the aforementioned six factors), and the two definitions happen to include the same complement types. If we try to doing away with the abstract concept of *object* and only keep notions like monotransitive object, then the coincidence has the following equivalence formulation: a complement slot prototypically filled by non-predicative NPs always share some or all grammatical properties about passivization, canonical constituent order, preposing and postposing, gap controlling, and being able to be modified by a predicative adjunct.

There are other constructions in which an object position can be distinguished,

including PPs and clauses with both object and predicative complement. Whether in these constructions the term *object* still hints certain resemblance with the prototypically monotransitive object is a question to be answered when these constructions are discussed about.

Typological significance for the two possible analyses is discussed in fn. 26 in CGEL Ch. 4. Some languages assign the direct object position to the goal-like argument. Were English such a language, it would follow that (53) would be the correct analysis, and since the subcategorization frame of (52) prohibits assignment of the object status to the goal argument (the object cannot be oblique), in (52), *all his goods* would be the subject, and hence both (52) and (53) would be correct, and we see a split G role. If, however, English belongs to the type of languages that assign the direct object position to the theme-like argument, we immediately get (55) and (56). Since the position of *some books* has more object properties than the position of *his favorite student*, (55) and (56) are preferable, but (52) and (53) also make sense to some extents. This means English belongs to the class which identify T with the monotransitive O, but it is not a clear-cut member.

Clausal complementation of ditransitive verbs, therefore, is summarized in Fig. 4.3. Note that in the diagram there are four – not five – internal complement types, because the O in V-O is identified with the T in V-G-T.

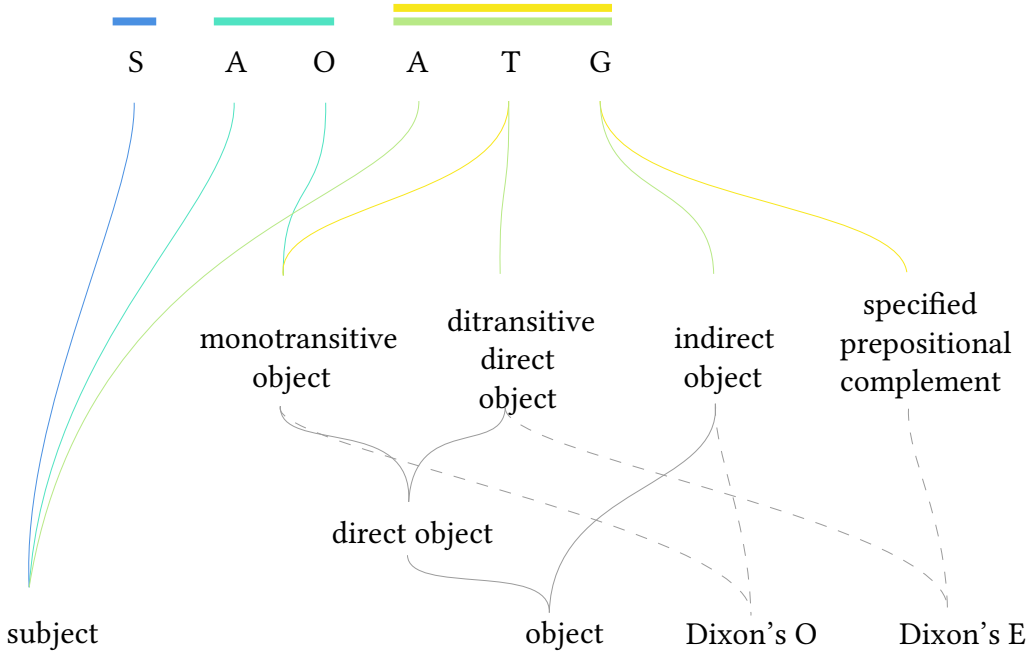

Figure 4.3: English alignment concerning S, A, O, G, T. Each color of lines means one canonical clause which codes a type of argument structure.

TODO: control c

4.4.3 Prepositional complements

Besides the prototypical transitive, intransitive and ditransitive constructions, English also has what is often broadly referred to as **phrasal verbs**, i.e. verb frames contain-

ing at least one preposition or particle.[8] There is ambiguity in what the term means. Dixon (2005, p. 289), for example, refuses to call a verb selecting a prepositional complement with the preposition specified a phrasal verb, keeping the name *phrasal verb* for verbal idioms in Huddleston and Pullum (2002). Huddleston and Pullum (2002) on the other hand rejects the term *phrasal verb* because according to their definition of constituency, what is a phrase in *play upon the fact* is not *play upon* but the whole verb phrase; if we change the terminology and define the verb phrase a la Dixon (2005, p. 41) the term *phrasal verb* can still be used.

Verb-preposition constructions and verb-particle constructions that contain only one preposition/particle can be classified according to the following parameters (Huddleston and Pullum 2002, pp. 272-274):

- whether it's a transitive preposition or a particle (an intransitive preposition, or something else),
- whether the construction can be interpreted in a compositional way or has already gained an established (idiomatic) meaning,
- how the choice of preposition/particle is restricted by the verb,
- the word order between the particle/prepositional phrase and everything else that may appear in the post-verbal region, like adverbs,
- the mobility of the preposition phrase/particle in WH-movements, and topicalization,
- the mobility of the NP following the preposition (this parameter makes no sense for verb-particle constructions), and
- complement-related properties of the associated NP coming with the preposition/particle, like whether it can be passivized.

I postpone verb-single particle constructions to § 4.4.4. This section is about verbs coming with a single preposition.

4.4.3.1 Mobility of preposition

The mobility of the preposition is about whether the prepositional phrase can be WH-fronted and topicalized (it seems the two movements are either both possible or both impossible), and whether an adverb can appear between the verb and the preposition, or, in other words, whether the prepositional phrase can exchange its position with the adverb in the clause-final region. A prototypically verb-preposition construction with a mobile preposition should allow all means of separating the preposition from the main verb.

(60) I refer to that book as a reference.

[8] Some *particles* here are intransitive prepositions, if the term *preposition* refers to a form class; the term *preposition* then is used to cover transitive prepositions. In the terminology of Huddleston and Pullum (2002), this confuses form and function (prepositions are a word class, and can be used intransitively in some cases), I choose to do so to keep the notation consistent with the current grammar writing practice.

(61) I referred frequently to that book in my presentation.

(62) To what do you refer?

Although it seems tempting to analyze a preposition fixed immediately behind the verb as a preposition incorporation into the main verb (and indeed this is an appropriate analysis for verb-preposition constructions in other languages, like Mandarin Chinese), this line of analysis doesn't fit with modern English well: if, for example, the preposition is really incorporated into the main verb, then when the complement of the preposition is long, we should be able to shift it behind lighter constituents, but this is not possible (Huddleston and Pullum 2002, p. 277). Furthermore, if the preposition is truly incorporated into the main verb, then we shouldn't expect marginal acceptability of insertion of an adverb between the preposition and the verb; examples of this usage appear even in academic papers (63). Long-range movement of the WH-phrase however is not possible, and I failed to find any incurrence of **across what did you come*. The problem of forbidden heavy NP shift could be explained by stipulating that the post-syntactic incorporation has to take place when the verb and the preposition is close enough to each other. The difference of acceptability between adverb-PP inversion and WH-movement however is hard to explain.

(63) Throughout the analysis of the world we come frequently across a juxtaposition of…(from Biemel and Emad (1980))

A more natural account seems to be based on whether it's possible to semantically make sense of the topicalized construction: *come across* and other syntactically "fossilized" verb-preposition constructions in Huddleston and Pullum (2002, p. 278) are not compositionally understandable, and hence topicalization or even merely inserting an adverb between the verb and the preposition can disrupt comprehension of the construction (Nediger 2017). But clearly inserting an adverb still isn't as disruptive as topicalization and WH-movement

A small number of seemingly compositional verb-preposition constructions like *fuss around* are still less syntactically flexible. This phenomenon – seemingly compositional idioms being syntactically rigid – is also observed in other types of idioms (Nediger 2017, § 5.7.2). This phenomenon is unlikely to have a unified analysis; real preposition incorporation effects might be one explanation for some. Variance in the population is also considerable, as *fuss happily around* is at least attested once.

(64) All night she'd watched the bankers get drunker and drunker, watched Bryan fuss happily over them, …(from COCA)

4.4.3.2 Selection relation with the verb and mobility

If a verb doesn't specify the preposition following it, a reasonable expectation is that the preposition is always mobile; anyway if the verb stem can appear with more than one preposition but one preposition is not mobile when used with that verb, we will still say the verb selects the preposition and forms an idiom with it. We therefore have a three-fold classification regarding the parameter of the selectional relation between the verb and the preposition and the mobility of the preposition in WH-fronting and topicalization:

- verbs with non-specified prepositions,

- verbs with specified but mobile prepositions (**preposition verbs with mobile prepositions**; Huddleston and Pullum 2002, p. 273), and
- verbs with specified and fixed prepositions (**fossilized preposition verbs**; Huddleston and Pullum 2002, p. 277).

It's possible that a non-specified prepositional phrase is still a complement (Huddleston and Pullum 2002, p. 273): we may say it's an oblique argument, although when passivization of it is available, it's actually object-like (see the end of this section; TODO: but could a non-specified prepositional phrase be passivized?).

4.4.3.3 Passivization

Passivization is completely not predictable from the classification made above (Huddleston and Pullum 2002, p. 276 [11]). Fossilized verb-preposition constructions are usually idioms, but some, like *break with*, still have largely inferrable meaning; the same applies for verbs with specified prepositions (indeed, the presence of a specified preposition introduces a sense of directed volition (Dixon 2005, p. 293)); verbs with non-specified prepositions usually are less idiom-like, but this is because if they are idiomatic enough, we will recognize them as verbs with specified prepositions.

By the criteria for object above (Fig. 4.2), we can find that the complement introduced by the preposition of a preposition verb is definite object-like when passivization of this complement is available; in this case the preposition verb is essentially transitive (Dixon 2005, p. 291, p. 297; Huddleston and Pullum 2002, p. 277), although dictionary editors may not think so; hence the title of this section. Unlike ordinary prepositional constructions, the preposition after the verb is likely to be a somewhat defunct one: the preposition may still control the semantic role of the NP, and the NP receives its accusative case from the verbal system, not the prepositional construction and is comparable to a direct object of a prototypical transitive verb. If we further stipulate that (a) passivization in English is suppressing the agent argument and moving the object to the subject position *regardless* of its semantic role, and (b) it seems reasonable for the nominative case to override the accusative case, but it seems strange for the nominative case to override an inherent case, then a defunct preposition allows prepositional passivization, while a full preposition likely disallows it, but may still allow preposition stranding caused by topicalization, as in *the thing I came across this morning…* (Richards 2017).[9]

This then implies that it's likely not possible to insert an adverb between the preposition and the verb in this case: no adverb position is available between the main verb and the object, and similarly, since now the NP after the preposition is the object of the main verb, no adverb should appear between the two (Richards 2017). This claim has some seemingly counterexamples (65); such constructions however *can't* be passivized (66), meaning that the structure of (65) is different from that of (67) which can be passivized (68). The *sleep in* sequence therefore has two underlying structures: one with a defunct preposition and hence is transitive, the other is an oblique argument construction (see next paragraph). For a dictionary editor, therefore, the "no adverb between V and P" constraint should be interpreted as a "no adverb between V and P in passive form" constraint.

[9]Passive constructions in some languages are more about the semantic role; these constructions may be better called "experiencing constructions".

(65) He slept comfortably in this bed.

(66) *This bed is slept comfortably in.

(67) He slept in this bed.

(68) This bed is slept in.

(69) ?This bed is comfortably slept in.

(70) He slept in this bed when the incident happened.

When passivization is not available, we may still say the NP following the preposition is an object, but now it's an object of the preposition only and is not object-like in the verb frame; in the verb frame, the prepositional complement is just an oblique argument, and is comparable to the optional complement of a particle in a verb-particle construction (§ 4.4.4).

4.4.3.4 Conclusion

The classification of verb-preposition constructions, therefore, is given in Table 4.4, where the ability of passivization, the existence of established non-compositional meaning, and the status of the preposition gives us ten classes in total. The examples used in the table is based on Huddleston and Pullum (2002, p. 278, [17]). The passivization parameter and the preposition mobility parameter are considered independent to each other; see § 4.4.3.3 for relevant discussion.

Table 4.4: Classification of verb-single-preposition constructions

passivization of NP after preposition	non-compositional meaning	non-specified preposition	specified preposition	
			mobile	fixed
yes	yes		*call on?*	*see to*
	no	*sleep in*	*refer to*	*fuss over*
no	yes		*stand for*	*come across*
	no	*fly to/from*	*feel for*	*come into*

Beside the classification given by Table 4.4, another parameter – not talked in the classification in Huddleston and Pullum (2002) – is the distribution of the verb part in the preposition verb constructions. Some of them have no other subcategorization patterns (just like verbs governing oblique cases found in languages with rich case morphology, like Latin), like *refer to*. For others, like *see to*, the verb part of the construction (usually a simple, monosyllabic one) does appear alone or with other prepositions. In the first case, the parameter of the existence of an established, non-compositional meaning is still relevant semantically (Dixon 2005, p. 291) but is less important in dictionary editing, because the whole verb-preposition sequence is to be treated as a single lexeme anyway (but it's possible that we see limited alternation, like *arrive at* v.s. *arrive in*). In the second case, the parameter of established meaning is important, because *stand at the door* is of course not idiomatic, and omitting it in a dictionary is acceptable, while *stand for* has an established meaning and deserves its own lexical entry. Taking this distinction into account, now we have 10 types of verb-preposition-object verbal idioms, and 4 types of verbs with inherent prepositions ignoring the parameter of non-compositional meaning for the latter.

4.4.3.5 The *by* complement in passivization

(71) He was instructed [secretly] to report what he saw in the jail.

4.4.3.6 Double prepositional complements

A natural generalization of what is discussed above is the V-pO-pO construction.

4.4.4 Verb-particle constructions

Similar to the descriptive parameters in § 4.4.3, the particle in the English verb-particle construction may be free, or it may be specified by the verb as a part of the lexical entry and possibly is subject to syntactic or semantic fossilization. There are two cases in which the verb-particle construction has identical surface linear order with the verb-preposition phrase construction. In the first case a verb licenses both a particle and an object, resulting in a superficial V-p-O linear order. This construction is introduced in TODO: ref and is different with the real verb-prepositional phrase construction in aspects listed in Huddleston and Pullum (2002, Ch. 4, § 6.2). The second case is about the possibility that the particle takes a complement itself and becomes a prepositional phrase; in this case we get a real verb-prepositional phrase construction, but the complement after the particle can't be passivized and is an oblique argument and is not of particular interest.

The exact syntactic position of the particle is a question; it may be a modifier of the verb stem and therefore forms a compound with the verb, or it may be a directional complement of the verb. When there is a direct object of the verb, either the object forms a directional small clause with the particle (but the object still undergoes the implicit movement to a higher object position making it closely attached to the verb; § 4.1.2), or the particle is a direct modifier of the verb stem; the latter case is frequently from the syntactic collapse of the former case. The presence of two large internal complements – one directional complement and one direct object – seems impossible for English.

Of course, if the V-p-O construction doesn't semantically support a small clause analysis, then only the latter analysis can be right, but a viable small clause semantics doesn't necessarily mean the existence of a small clause in syntax. There can also be alternation between the two structures for a V-p sequence. This is hard to directly tell from the surface order, and we will explore into this topic by examining the interaction between the particle and other constituents.

Surface-oriented descriptive parameters for verb-particle constructions include the follows (Huddleston and Pullum 2002, Ch. 4, § 6.3):

- its linear position compared with other complements, mainly the object,
- it ability to be moved out of the VP and topicalized, and
- its linear position compared with adverbs.

Regarding the first parameter, we can distinguish two basic linear orders: *mess everything up* and *mess up everything*. Most of the time the alternation is free and is controlled by the same factors in heavy NP shift. Some sort of verb-particle incorporation is clearly present in all V-p-O constructions: no coordination targeting the

particle is possible, and we have examples like *re-plug in* (Farrell 2005).[10] This may reflect the nature of the particle as a modifier of the verb stem but could also be from post-syntactic operations possibly due to prosodic requirements as in heavy NP shift. The V-O-p constituent order may be a small clause construction but can also be due to the particle being left behind to satisfy other linear order constraints. There are however cases where the particle has to appear before the NP: for example *carry out his threat* is fine while *carry his threat out* is much less felicitous (Huddleston and Pullum 2002, p. 285); this seems to be results of fossilization and likely means that the verb-particle compound analysis is right, or otherwise it's hard to imagine that a somehow non-prototypical version of the small clause construction gets fossilized.

Inability to topicalize is a strong hint at idiomization and hence the verb-particle compounding structure.

The relation between a non-fossilized particle and manner-like adverbs should be relatively free for a small clause-type verb particle construction, because a non-fossilized particle is not an object and there is no requirement for it to appear directly after the main verb (§ 4.1.2).

In conclusion, we have three types of verb-particle constructions: the small clause type, the semi-idiomized verb-particle compound, and the fully idiomized verb-particle compound. TODO: is it possible to have V-O-p construction but p is purely a directional phrase (not a compound modifier)?

4.4.5 Ditransitive constructions with particles

Up to now all types of complements of the verb are introduced. Now we discuss under which condition a particle appear in a ditransitive construction. It turns out that the ditransitive constructions and particles have entangled relations, and it seems that the Latinate verb prefix needs to be treated as a particle.

It is possible to have both particle and a dative *to* complement, as in *Mary showed her painting off to John*, because the particle *off* may be seen as a part of the directional small clause *her painting off to John*, but no particles like this are available in the double object construction (**Mary showed off John her painting*), whose small clause construction is more like *John (has/sees/etc.) her painting*, which is not directional and does not license the particle (Harley 2007). The impossibility for a large family of Latinate verbs with prefixes like *ex-* or *in-* to appear in the double object construction seems to be evidence for the prefixes being not synchronically embedded into the stem yet. Some Latinate verbs still allow the appearance of *away* as in *explain the incident away*, but here *away* is more likely to be a goal argument, instead of a preposition-like particle like *up* as in *up in the sky*, and therefore it can't be modified by adverbs like *right*, which, on the other hand, modify real particles like *he finished his job right up*. The fact that not all Latinate verbs allow this use of *away* may suggest at a structural difference between the argument structures of Latinate verbs and Anglo-Saxon verbs; one explanation is that the small clause of an Anglo-Saxon verb is really a structurally independent constituent, serving as a complex complement of the verb, while the arguments in the "small clause" of a Latinate verb are attached to the verb one by one, not as a whole, and local selection is able to happen (Punske 2013).

[10] This paper rejects the notion of small clauses mainly because of absence of evidence; when the interaction between the particle and more well established small clause constructions is considered, we do have evidence.

There exist another class of verb particle constructions in which the particle is not a part of a small clause but rather a modifier of the verb stem, proven by the easiness for the verb-particle consequence to undergo morphological derivation (Farrell 2005); due to the closeness between the particle and the verb stem, this construction is subject to much more idiomization than the small clause-type of particle verbs (Wurmbrand 2000). Examples include *mess up*. Besides semantic feasibility, we also have syntactic tests listed in Wurmbrand (2000) to distinguish between the two breads of verb particle constructions: including whether the particle can be coordinated (*he nervously took his things [in and out]* v.s. **he took this over and out*) and whether the particle can be alternated. The particle position of *up* in *mess up* and the position of *away* as in *explain the incident away* seem to roughly be the same one, or at least mutually exclude each other; this means in a verb frame we should expect to see at most two particles, one is similar to the aforementioned *up* in *mess up*, another is a part of a directional small clause. It seems in modern English actually the case with two standalone particles is never attested (Dixon 2005, p. 293, Huddleston and Pullum 2002, p. 286): the only allowed case is the *[ex]-plain it [away]* case.

The particle *away* however is different from fully fossilized verb-particle constructions (TODO: CGEL citation) in that it can still be moved around by various syntactic constructions: In *mass it up*, *it* seems to be a pronoun critic; does it mean some verb particle constructions have particles that are aspectual markers, etc.?

Conclusion: one external complement, one big internal complement (possibly a small clause), one verb compound

4.4.6 Transitive construction with additional preposition/particle

In this section I discuss transitive verbs that also have a particle and/or a prepositional complement in their verb frames.

When only one prototypical object is licensed, the verb frames can be divided into three types:

- V-O-p (TODO: ref),
- V-O-pO (TODO: ref),
- and V-O-p-pO.

In all of these constructions, the prepositional complement is unable to be passivized. Thus the descriptive parameters are reduced to TODO

The V-Oi-p-Od construction is also possible TODO

4.4.7 Constructions involving both preposition/particle and copular complement

V-p-CC, V-O-p-CC, V-p-O-p-PC, V-p-(as)-PC, V-O-p_as-PC

4.5 Topicalization

Several sub-constructions that fall under the general category of topicalization can be recognized. Object or copular complement preposing is much less frequent and seems to be dying out in today's "general-purpose" formal writing, although it's still quite common in more artistic registers.

(72) Subtle is the Lord, but [malicious] he is not

4.6 Cleft constructions

4.6.1 *It*-cleft

A *it*-cleft construction contains an expletive or dummy subject *it*, a finite form of *be*, a focused constituent, and a THAT-clause in which there is a gap (73, 74). Note that here *it* never changes into, say, *she* or *they*, and *is* doesn't show any agreement with the focused constituent. This is the expected behavior: Note that it's pretty fine for the syntactic numbers of the NPs before and after *be* to be different in non-cleft clauses, and *be* always agrees with the subject.

(73) It is $[\text{him}]_{i,\text{focused}}$ that he wanted to murder $-_i$!

(74) It is $[\text{by this new method}]_i$ that we have achieved such success $-_i$.

The range of constituents able to be focused can be found in Huddleston and Pullum (2002, pp. 1417-1419). We seems to have a generalization: If the focused constituent is an adverbial, it has to be able to be an answer to a *how* question – and therefore can appear after *be*.

The *it*-cleft construction seems to be a clausal idiom: The dummy *it* can be raised when the *it*-cleft construction is in an infinitive form and embedded into a complement clause construction. Note that the *it*-cleft construction can't be a verbless clause: This makes the focusing reading inaccessible.

TODO: *that* or *who*??

Thus, although the *it*-cleft construction is already an established construction with a fixed meaning (i.e. focusing), it's still analyzable as a bi-clausal construction following the usual syntactic constraints found elsewhere in English.

4.6.2 *Wh*-cleft

4.7 Expletive subject constructions

4.7.1 *There be* existential construction

(75)

4.7.2 *It seems that ...*

4.8 Focus constructions

The English focus construction involves subject-auxiliary inversion and preposing of the focused constituent. Note that the fronted constituent can be a copular complement, a locative adverbial, TODO: list but never an object in a prototypical transitive construction. This seems to be motivated for functional reasons, because in the latter case it's impossible to correctly restore the meaning.

(76) [On the top of the mountain]$_{\text{locative PP}}$ lays [a small church]$_{\text{subject}}$.

> **Box 4.6: Focusing is not valency changing**
>
> This note follows the analysis in Huddleston and Pullum (2002, p. 244) and regard this as a type of information packaging but not a voice construction. Some grammars, like Quirk et al. (1985, p. 736), analyze the fronted constituent as the subject. This is not the position of this note TODO: why?

4.9 Interrogative moods

4.9.1 Overview

There are two movements involved in forming a canonical interrogative clause: the subject-auxiliary inversion, and fronting of the WH-phrase, if any. Both operations can be omitted in casual speech.

The usual landing site of the WH-phrase in interrogative constructions seems to be the clausal focus: the scope of the landing site is narrower than that of the topic, comparable to the position of a usual focus (77), and interrogative formation is in conflict with another focus possibly because in English a sentence can only have one focus (78). Fronting of the WH-phrase before the topic to the beginning of the sentence is also possible (79), but this construction is still in contradiction with another focus, meaning that (79) is better analyzed as a derivative of (77).

(77) In your opinion, who did Watson saw on Baker Street last night?

(78) *What never again will you do?

(79) Who, in your opinion, did Watson saw on Baker Street last night?

4.9.2 Yes-no questions

4.9.3 Open questions

4.9.4 Tag questions

(80) The car is broken, isn't it?

4.10 WH-movement

WH-fronting seems to be subject to several constraints (Cable 2012): extracting a peripheral part of a complete NP is categorically not allowed in English,[11] and at least in formal English, the WH-phrase should be near enough to the initial of the clause, presumably to mark the interrogative nature of the clause. These two constraints together explain why WH-phrases like *whose book* are moved as a whole.

There are more constraints needed to rule out examples like like **See what do you?* (intended meaning: 'what do you see?'), which seem to be unacceptable in all known languages; other constraints are language-specific and are hard to capture neatly (Cable 2013).

4.11 The imperative mood

The imperative mood is only compatible with active clauses. Passive imperatives are never possible, and the intended meaning may be alternatively expressed by

4.12 The TO-infinitive

4.13 Nominalization

The distinction of the two genitive constructions in English now is clearly displayed: the subject-determiner *'s* clitic is usually used to introduce the NP corresponding to the subject in the clause before nominalization, while *of* is usually used to introduce an internal complement.

(81) a. He plays the national anthem.
b. his playing of the national anthem

[11] Cable (2012) captures this constraint by stipulating that the leftmost part of a noun phrase can't be moved out; we could, on the other hand, capture the same constraint in a more natural way: WH-words are function words after all and might even be the collective realization of several grammatical categories, and moving them alone seems absurd.

5 Verb phrase

5.1 Introduction

This chapter is the first chapter about the verb phrase, which contains "core" clausal dependents (§ 4.1.2, Fig. 4.1). It contains several subsystems: first the core arguments i.e. internal complements (§ 5.3), then peripheral arguments (TODO: ref), then the auxiliary chain and the main verb i.e. the verbal complex (§ 5.2, § 5.2.4) with complicated interaction with negation, and also adverbials about TAM marking (§ 4.2). The alternation of the argument structure involves the category of voice (TODO: ref).

5.2 The verbal complex

5.2.1 Regular lexical verbs

Modern English has already lost most of its verb inflection. Following the analysis of Huddleston and Pullum (2002, Ch. 3.§ 1.1), for lexical verbs, there are six remaining inflectional forms: the past form, the plain present form, the 3sg present form, the plain form, the ING-participle, and the ED-participle. The two present forms and the past form appear solely with trivial aspectual values and trivial modality. They are **primary** forms: They already have all TAM categories marked on them. The plain form and the two participles are **secondary** forms: They usually appear after auxiliaries in a periphrastic construction to have full TAM marking, though a subjunctive clause may sometimes get rid of any auxiliary verb, as in *he suggests that she [complete] this task first* (§ 7.1).

Examples of these forms are illustrated in Table 5.1. This is a copy of [1] in Huddleston and Pullum (2002, § 1.1). It can be noticed that the plain form is usually the same as the plain present form. However, since modal verbs (see below) have no plain form, and that the syntactic environments of the plain form and the present plain form are too different, if Table 5.1 is to be regarded as a paradigm – that is, to be incorporated with morphosyntactic information – then the two forms should occupy two cells.

Table 5.1: Paradigms of lexical verbs

			take	*want*	*hit*
Primary	past form		*took*	*wanted*	*hit*
	present form	3sg	*takes*	*wants*	*hits*
		plain	*take*	*want*	*hit*
Secondary	plain form		*take*	*want*	*hit*
	ING-participle		*taking*	*wanting*	*hitting*
	ED-participle		*taken*	*wanted*	*hit*

Box 5.1: The name of the forms

Here I deviate from the practice in (Huddleston and Pullum 2002, Ch. 3) and pick up the more common names for some of the forms.

The ING-participle is frequently called the *gerund*, because it now has the function of both a gerund and an active participle. Huddleston and Pullum (2002) call it the *gerund-participle*. Some grammars use the term *present participle*. Since in Modern English, the ING-participle no longer carries any tense information, the historical term *present participle* is abandoned in this note.

The traditional name *past participle* for the ED-participle makes more sense, because it's morphologically related to the past form for regular verbs and it still has some sense of "past": It is strongly related to the PERFECT and therefore has some sense of the past, though it doesn't carry the past tense. A better term would be the one in Latin grammar: the *perfect passive participle*, but this is in conflict with the name of the *having been done* construction.

A usual name for the plain form is the infinitive form, which I reject here because the morphological marking of the main verb after modal auxiliary verbs (*would [like]*), the verb in a subjunctive clause (*he suggests that she [complete] this task first*), and the verb in a real infinitive clause are all the same, and therefore it makes no sense to use the term *infinitive* to cover the *morphological* form of all the three.

The ING-participle is regularly formed by adding *-ing* to the end of the plain form (TODO: -tt- in splitting). The ED-participle and the past form are usually obtained by adding *-ed* to the end of the plain form, but for irregular verbs they can't be inferred from the plain form. Thus English verbs have three **principal forms**: the plain form, the past form, and the ED-participle. We may also say there are three stems in English: the plain form, the past form, and the ED-participle, with only the first one being productive for further morphological processes.

5.2.2 Types of irregular verbs

As is mentioned above, for a number of irregular verbs, the ED-participle and the past form can't be inferred from the plain form. Whether there are still some patterns between the three, or in other words, the formation of the principal parts, is investigated in detail in Quirk et al. (1985, pp. 105-120).

5.2.3 Auxiliary verbs

English also has a number of auxiliary verbs. All auxiliary verbs have tense-dependent forms, because all of them may appear as the first word in an auxiliary chain, and the tense category is to be marked on the highest i.e. the first of them (§ 5.2.4). Thus, we say English auxiliaries also have primary forms. Modal auxiliaries don't have a separate 3sg present form, but *do*, *have* and *be* (when used as auxiliary verbs) do. It should be noted that the past forms of many auxiliary verbs don't just appear in past clauses: They may have distinct meanings (§ 4.2.7).

Modal auxiliaries don't have secondary forms, probably because they never appear after another auxiliary verb or in nonfinite clauses, but *do*, *have* and *be* do.

English auxiliary verbs also have negative forms, which are obtained by attaching *-nt* to the end of auxiliary. The *-nt* is a contraction form of the negator *not*, but in modern English the negative suffix moves together with the auxiliary in subject-auxiliary inversion (§ 5.2.8). Thus, it's recognized as a part of the auxiliary (Huddleston and Pullum 2002, p. 91). This seems to be purely about phonetic realization: There seems to be no large morphosyntactic differences between auxiliary-*not* and the negative auxiliary besides subject-auxiliary inversion. All auxiliaries don't have secondary negative forms, though *do*, *have* and *be* have primary negative forms.

Since auxiliary verbs are a part of the grammar, here I list the paradigms TODO

Box 5.2: Auxiliary constructions are single-clause ones

Huddleston and Pullum (2002) treat auxiliary verbs as verbs taking complement clauses (as in, say, [11] in p. 782). This is not the position of this note: Here I follow the standard practice in generative syntax (probably also American structuralism) and assume auxiliary verb constructions are always single-clause constructions. *Historically*, auxiliaries may origin from complement-taking verbs, but now *synchronically*, they have the same function of inflectional affixations. Complement clause constructions may (or may not) have the same *semantics* of auxiliary verb constructions and inflectional affixations, but they never have the same *structure*.

The main reasons Huddleston and Pullum (2002) analyze auxiliary verbs as complement-taking verbs or *concatenative verbs* in their terms are shown in their Ch. 14.§ 4.2.2. They argue that we have constituency trees like *[would [like to do]]* – which I also agree on. They then argue that this constituency structure means *like to do* is a complement clause – which then is not always true. This second part of their argument seems to come from confusion between lexical heads and (PF realization of) functional heads. But then they are inconsistent when they argue that the complementizer *that* isn't a head. In this note, I follow the standard definition of (lexical) headhood in the descriptive literature while fully being aware of the generative functional head analysis.

Evidences supporting my claim that auxiliary constructions are indeed single-clauses ones can be obtained by observing how auxiliaries interact with clausal dependents. If Huddleston and Pullum (2002) are correct on their claim that English auxiliary verbs take bare infinitive clauses, then we expect the verbal phrase after an auxiliary verb to receive any modification that's acceptable for a bare infinitive clause. However, as we see in § 5.2.6, there is a strong tendency for adverbs to appear after the first auxiliary, which can be easily explained by assuming the first auxiliary undergoes some kind of fronting (§ 5.2.7), or after all auxiliaries and before the main verb, and the functions of adverbs in the two positions have clear correlation with the positions. This pattern are hard to account for when we assume auxiliary verb constructions are complement clause constructions, because nothing motivates it. If, on the other hand, auxiliary verb constructions are

single-clause constructions, then we can say the distribution of adverbs and auxiliaries show is just the surface reflection of a deep functional hierarchy, just like the subject is somehow higher than the object.

5.2.4 Minimal auxiliary chain

In a declarative finite clause, the order of auxiliaries is constantly given by Table 5.2. Table 5.2 is a part of the larger picture of clause structure: The auxiliary *do* (§ 5.2.5), adverbs (§ 5.2.6) and the negator (§ 5.2.7) may be inserted into somewhere between two auxiliaries. Other types of clauses still largely follow the scheme but may undergo subject-auxiliary inversion (§ 5.2.8).

The auxiliaries positions can be filled by the corresponding auxiliaries or be just left blank, without creating ungrammatical constructions. The MODAL slot may be filled by a modal auxiliary. The PERFECT slot may be filled by the auxiliary version of *have* with the correct inflection, and the PROGRESSIVE and PASSIVE slots may be filled by the auxiliary version of *be* with the correct inflection.

The rules of inflection are the follows.

- *Marking of tense.* The tense category is always marked on the first auxiliary (not necessarily one of the slots in Table 5.2 – it may be an inserted *do*), and when there is no auxiliary, it's marked on the main verb. Note that if the first auxiliary is in the past form, it's possible to *not* refer to a past event (§ 4.2.7).

- *Concatenating rules.* The modal auxiliary is always followed by a plain form, and the progressive marking *be* is always followed by an ING-participle, and the perfect marking *have* is always followed by an ED-participle, and so is the passive marking *be*.

- *Agreement.* When the clause is finite and the tense is PRESENT, and the MODAL slot is empty, if the subject is 3sg in number, then the first non-empty slot in Table 5.2 is in the 3sg present form, which means for verbs other than *be*, the *-s* suffix is attached to it; for *be* the correct form is *is*. This is the only case subject-verb agreement happens in English other than the case of *be* (1). For *be*, the tense is still TODO: subjunctive

The English tense is the only category that is marked purely morphologically.

In nonfinite forms, the MODAL slot has to go; the rest are still there, following the same inflectional pattern as is described above (2). Note that the subject-verb agreement is missing in all nonfinite clauses, be it the third person singular *-s* or inflectional forms of *be*.

Table 5.2: The order of auxiliaries and some examples

MODAL	PERFECT	PROGRESSIVE	PASSIVE	main verb
				takes
			am/are/is/was/were	*taken*
		am/are/is/was/were		*taking*
	have/has/had			*taken*
	have/has/had	*been*	*being*	*taken*
will/would	*have*	*been*	*being*	*taken*

(1) a. I [like] this.
b. He [likes] this.

(2) The award is reported [to have been being taken]$_{\text{complement clause: TO-infinitive}}$

(3)

5.2.5 *Do* insertion

5.2.5.1 Obligatory *do* insertion

Do insertion happens in two circumstances. The first is we need an auxiliary but there isn't one. This is the case when we negate a clause with no auxiliary verb (§ 5.2.7), and the case when subject-auxiliary inversion happens but there is no auxiliary verb (§ 5.2.8). In both cases, *do* is inserted before the main verb, and is regarded as an auxiliary, which carries the tense feature and the subject-verb agreement information and is inflected accordingly (4, 5).

We may say the *do* is the default realization of the tense category and the agreement when these can't find an appropriate host. It's roughly in the same position of MODAL in Table 5.2. Then, expectedly, adverbs can be inserted between *do* and the main verb (6).

(4) I do not like the gift. I don't like the gift.

(5) Did he enter the room that night?

(6) I do not particularly like that kind of flower.

5.2.5.2 *Do* for emphasis

Unlike (4, 5, 6), we can also just insert *do* to emphasize on the action, and in this case the inserted *do* receives stress. The morphology of *do* is the same as the obligatory *do* insertion, and so is the distribution of adverbs.

(7) Your company [*do*]$_{\textit{do}\text{ insertion}}$ [have]$_{\text{main verb}}$ lots of rules!

5.2.6 Adverbs in the auxiliary chain

For TAM-related adverbs (§ 4.1.2), also known as sentential adverbs in Dixon (2005), when there is at least one auxiliary, adverbs are prototypically inserted after the first auxiliary and before the rest of the verbal complex. Inserting an adverb after the first auxiliary (possibly, fronting other auxiliaries following the first auxiliary) is only

marginally acceptable (Dixon 2005, p. 389). This linear order holds regardless of the correspondence between the function of adverbs and the function of auxiliaries, indicating some sort of auxiliary fronting rules: we may say that there is an invisible slot at the beginning of the auxiliary sequence in Table 5.2, which is about tense; whenever an auxiliary is present, the highest auxiliary – the first auxiliary – is lifted to the tense position, before negation and the default position of many adverbs; when no auxiliary is present, the tense feature is lowered to the main verb.

Manner-like adverbs and some adverbs coding peripheral arguments prototypically appear after the auxiliary chain, either immediately before the verb, or in the region after the verb (Dixon 2005, p. 386).

but when the negator *not* appears, it immediately follows the first auxiliary and before any adverbs appearing in the verbal complex. (TODO: really?)

TODO: alternation in order

(8) He [is]$_{\text{PROGRESSIVE}}$ [vigorously]$_{\text{TODO:}}$ [doing]$_{\text{main verb}}$ [his job]$_{\text{object}}$.

5.2.7 Negation in the auxiliary chain

The rule of the negator *not* is close to the rule of adverbs: If *not* is used, it is *always* after the first auxiliary (while adverbs can appear before the first auxiliary in marked cases), which may be the inserted *do* (9). Any auxiliary-*not* sequence may be replaced by the negative form of that auxiliary if there is one (10, 11).

(9) He [does]$_{\textit{do}\text{ inserted, pres, 3sg}}$ [not]$_{\text{negation}}$ love his job.

(10) He doesn't love his job.

(11) He isn't vigorously doing his job.

It should be noted the surface position of the negator doesn't determine the scope of negation (Huddleston and Pullum 2002, p. 668). See, for example, the ambiguity of (12). Here the ambiguity is an indicator that there are at least two available syntactic position of the reason clause (TODO: ref). Another ambiguity arises when negation appears together with modality (13, 14). This means the negator-after-first-auxiliary rule is about *realization* and not about the underlying syntactic structure (§ 2.1.1), if we assume the semantic difference has structural significance. This, together with the fact that auxiliaries have negative forms and that the existence of *not* blocks subject-auxiliary inversion of the main verb, may lead to the conclusion that the negator *not* is a quasi-verbal clitic which is always attached after the highest verbal element. We, however, shouldn't rush to such a conclusion, because it's also possible that the rule is actually the highest verbal element is always moved *before* the negator. Note that

TODO: the Tense - Negation - Modality - Perfect - ... sequence

(12) I don't appoint him because he is my son.
'I appoint him, but because of his talent, not because his relation with me. / I don't appoint him, because he's my son and I don't want to appoint him and leave a bad impression on my colleagues.'

(13) He shouldn't play football in the streets.
'It's required that he doesn't play football in the streets./ *It's not required that he plays football in the streets, but he can if he wants to.'

(14) He can't play football.
'It's not possible/permitted that he plays football./ *He can suppress the desire to play football.'

5.2.8 Subject-auxiliary inversion

In interrogative sentences and in other cases (§ 4.1.2), the first auxiliary in the chain undergoes leftward movement, often to the initial position but may be preceded by preposed constituents (§ 4.1.2). This is called **subject-auxiliary inversion**. When there is no auxiliary, the correct form of *do* carrying the tense and agreement features is inserted.

> **Box 5.3: Subject-auxiliary inversion and ordinary constituent movement**
>
> Note that this is head movement and are often attributed to post-syntactic operations in Distributed Morphology, making the operation kind of "realizational", and is comparable to realizational factors in morphology, instead of the underlying structure (§ 2.1.1), while ordinary constituent movement like topicalization or wh-movement is comparable to the latter.

(15) $[\text{Do}]_{\text{inverted auxiliary}}$ $[\text{you see my umbrella}]_{\text{nucleus}}$

(16) Only then do we cook

5.2.9 Semi-auxiliaries

5.2.10 Comparability with moods

5.3 Clausal dependents and verb frames

5.3.1 Overview

5.3.2 The subject

5.3.3 Prototypical transitive and intransitive verbs

5.4 Agreement

If you take a closer look to how native speakers of English do subject-verb agreement, you'll find some more subtle details than the textbook rule that when the tense is present and the subject is 3sg, *-s* is added to the first auxiliary or the main verb (Huddleston and Pullum 2002, Ch. 5, § 18).

5.5 Voice

One thing that happens in the verb phrase and strongly influences the structural building process is the category *voice*. English doesn't have a rich set of valency changing devices, and the active-passive distinction is the only regular valency changing mechanism. There are other alternations of verb valency, but they are much more strongly determined by the lexicon.

5.5.1 Passivized argument

Passivization depends on - though not in a very apparent way – the properties of the main verb. TODO: passivized argument

5.6 Peripheral arguments and manner-like adverbials

5.6.1 Structural classification

5.6.2 Mean and manner

5.6.3 Adverbials related to modality

5.6.3.1 Relative order and compatibility

speech act before evaluative (Cinque 1999, p. 106)

5.6.4 Adverbials and clause-linking constructions

TODO: Huddleston and Pullum (2002, Ch. 8, § 12); position in Cinque hierarchy

6 Adjectives and adverbs

6.1 The structure of the adverb

6.1.1 The *-ly* derivation

TODO: terminal derivation; only *-c-al-ly* is acceptable, except *puclicly*.

6.2 Comparative constructions

Comparative constructions appear to fill a "degree" position of adverbs, adjectives, and sometimes, nouns and verbs. TODO: compare with *how old* constructions

In some languages, like Mandarin, the seemingly counterpart of *than* has a different role: the NP it introduces is not the thing or person that is directly compared, but the *possessor* of the thing or person that is directly compared (1). Speakers of these languages may transfer this usage to the way they use *than*, leading to the rise of the ungrammatical form (2).

(1) 我 的　成绩 [比] 你 好
1 POSS score PREP 2 good
My score is better than yours. (lit. my score is better [than the counterpart of what's compared of] you.)

(2) *My score is better than you.

7 Complement clause constructions

Complement clauses or *content clauses* (Huddleston and Pullum 2002) are clauses embedded as arguments of certain verbs. English adverbial clauses have the same form of complement clauses, and therefore Huddleston and Pullum (2002) use the term *content clause*. Here I'll just stick to the more common terminology in linguistic description.

According to Dixon (2010, § 18.4), there are usually three types of complement clauses:

- the Fact type, which looks like a full sentence and usually express a fact,
- the Activity type, which looks like an NP but still keeps key features of clauses and usually express an ongoing activity without specifying the time (and thus with deficient TAM marking), and
- the Potential type, which describes the potential or plan to do certain things and also has deficient TAM marking, but is formally less similar to NPs.

In English, the three types correspond to precisely the finite complement clause, the participle clause, and the infinitive clause. Note that the classification is pseudo-semantic: it is similar to the "generalized semantic role" labels S, A, P, etc., which suggest their prototypical – but not unique – semantic function. Indeed, a Fact clause – a finite complement clause – is able to express a potential, like *It's possible [that I will spend my summer in Paris]*, and an Activity clause can express a fact (*[Being an engineer], he has a sharp mind when solving practical problems*).

7.1 Subjunctive clauses

The subjunctive complement clause construction here excludes many other constructions that may be called "subjunctive" (TODO); only in subjunctive complement clause introduced here do we have modality (or "situation"?) fused to the main verb.

7.2 Infinitive constructions

7.2.1 Infinitives as VP internal complements

Infinitive clauses within the VP have richer behaviors and are probably the most frequent type of infinitive clause constructions in English. This section discusses infinitive clause constructions with the constituent order shown in Fig. 7.1, where the object position and the infinitive complement clause position are all internal complements shown in Fig. 4.1, and the object position may be absent. Before going into details

of each construction, I list several parameters and syntactic tests to classify infinitive complement clause constructions.

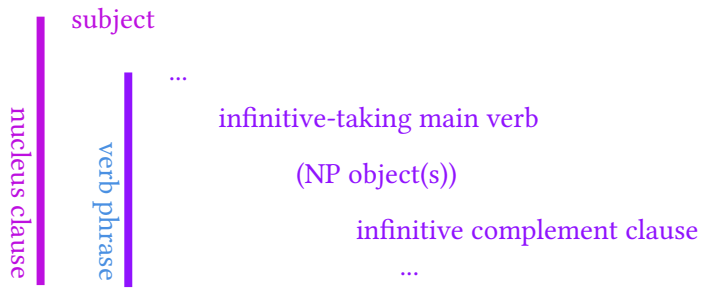

Figure 7.1: Infinitive clause following a verb

An infinitive in an internal complement position may be a "sealed" one: it comes either with its explicit subject and has nothing particularly different from a typical finite complement clause (1, 2), or it comes without a subject but the reference of the subject is solely determined by the conversational context. The second case seems lacking for internal complement infinitives in English: (3) in Huddleston and Pullum (2002, p. 1269) as an example of an infinitive whose subject is missing in the surface form and has contextual reference: in most conversational environments, the subject of the bracketed infinitive would be Max, but by changing the context we can easily change the interpretation: the subject of the bracketed infinitive can be, for example, a friend of Max, who is submitting a paper but realizes the time left for revision is too short, and Max tried to defined him before his angry advisor but failed.

(1) I want [for her to complete this task tomorrow]$_{\text{sealed infinitive}}$.

(2) I promise [you]$_{\text{base-generated object}}$ [for John to come here]$_{\text{sealed infinitive}}$.

(3) Max admitted it had been a mistake [to leave so little time for revision].

In English, when the bracketed infinitive in (3) is an internal complement, it no longer accepts such contextual interpretation. The behavior of superficially subject-less infinitives appearing within the core VP more or less deviates from the behavior of a complete, "sealed" infinitive. When the subject of an internal complement infinitive clause is missing in the surface form, there are two possibilities (Huddleston and Pullum 2002, pp. 1194-1197): either it can be explained by so-called raising (4), or it is an instance of so-called controlling (5). In both cases, the subject of the infinitive is *obligatorily* coreferential with one or more antecedent argument. In the follows, we list several parameters of internal complement infinitives and analyze whether they can divide superficially subject-less infinitives neatly into two types.

(4) [The student]$_{i,\text{subject}}$ seems [$-_{i,\text{raised}}$ to be cheerful]$_{\text{infinitive}}$.

(5) [I]$_{i,\text{subject}}$ want [$-_{i,\text{controlled}}$ to join your group]$_{\text{infinitive}}$.

7.2.1.1 The weight of the infinitive

One piece of evidence suggesting variance of infinitives comes from passivization. A finite complement clause can be passivized, although usually the clause will not be moved to the subject position, and a dummy *it* is inserted. Some verbs taking infinitives support a similar passivization process of the infinitive (7; Huddleston and

Pullum 2002, p. 1196, [7]), suggesting that these infinitives are comparable to finite clause and "sealed" infinitves.

(6) a. He suggested that this proposal was doable.
b. It was suggested that this proposal was doable.

(7) a. We hope to return to this issue.
b. It is hoped to return to this issue.

When passivization of the clause is not supported, it doesn't necessarily mean that the infinitive is deficient compared with sealed infinitives. Another test with a broader scope is comparing the alternation between the infinitive and a sealed infinitive.

(8) What she really wanted was for him to apologize.

(9) *What she really wanted was him to apologize.

(10) What she really wanted was to apologize.

(11) I will arrange [to see a specialist] and [for my child to see one at the same time]

There is, on the other hand, one syntactic test suggesting the openness of the infinitive to the matrix clause, and hence the deficiency of the infinitive. If the verb of the infinitive taken by the matrix verb supports the *there be* construction, then this property sometimes propagates to the matrix clause (12): note that the *there be* form is definitely licensed not by the matrix verb but by the infinitive verb (13).

(12) There seems to exist an all-in-one plugin for your functionality request.

(13) *There seems to sing a man.

In principle the structure of the infinitive can be even more open to the matrix clause: in German, for example, a certain subtype of infinitive clause constructions allows the object in the infinitive clause to be passivized, while clearly the matrix clause verb is a lexical verb (Wurmbrand 2002). This however is not possible in English.

We can see a hierarchy of the structural closeness of infinitives. The German long passivization-compatible infinitive clause is nothing more than a VP and may even have no syntactic subject at all, as it refers to a type of action and not a complete event; it is the most open to the matrix clause, because it is almost disqualified as a clause. Infinitives appearing in environments like (12) are less open but is still not as sealed and independent as a finite complement clause.[1] Finally, although the null subject of *to return to this issue* in (7) has to refer to the matrix clause subject, the infinitive does show features belonging to an independent, sealed clause.

7.2.1.2 Core argument structure

There is one rather intriguing phenomenon in a large class of infinitive-taking verbs: no semantic role is actually assigned to the subject of the infinitive by the matrix main verb. This distinction – whether the thing referred to by the null infinitive subject has a semantic role in the matrix clause – leads to an array of syntactic tests (Huddleston and Pullum 2002, p. 1196, [8], except the matrix passivization and dummy subject tests).

[1] A WH-phrase can be moved across clausal boundaries, but it is more a topic and is more peripheral than the subject.

This seems to mean that the infinitive is more independent when its subject receives no semantic role from the matrix clause; note, however, that if the matrix verb licenses only one argument, which is the infinitive, its subject can't be filled, and in this case the subject of the infinitive has to be extracted to the matrix clause subject position. This means the infinitive can't be a sealed one, or otherwise extracting the subject into the matrix clause will not be possible.[2] Similarly, when an object position is licensed in the matrix clause but the matrix verb has no patient argument, the only viable way is to extract the subject of the infinitive to the matrix clause object position. The resulting subjects and objects have nothing different with ordinary subjects and objects: the object which is the raised subject of the infinitive doesn't allow an adverb appearing between it and the matrix verb, for example.

(14) *He expects confidently you to finish the project within a week.

In theory, even if the matrix verb licenses enough arguments to fill all clausal dependent positions, the infinitive can still be deficient. But this never happens in English. Therefore we defined the class of verbs that take deficient infinitives and license less-than-enough arguments so that the subject of the infinitive has to be raised out of the infinitive and to the subject or object position in the matrix class the **raising verbs**. The rest of the verbs taking seemingly subject-less infinitives are **control verbs**, because the matrix clause subject/object, which is obligatorily coreferential with the infinitive null subject, "controlls" the latter. The exact structure of the control construction, i.e. how it compares to the raising construction and to the sealed infinitive construction, is discussed in § 7.2.1.4.

TODO: explicit subject control??

7.2.1.3 Some constraints

Some generalizations further narrow down the number of possible infinitive clause constructions. English has a strong tendency to spell out only one copy of a moved constituent, so if the subject of the main clause is linked to the subject of the infinitive, then the latter is not visible, and the object of the main clause, if any, has to be base-generated; and if the object of the main clause is linked to the subject of the infinitive, the latter is also not visible, and the subject of the main clause has to be base-generated.

The subject can also be a dummy matrix clause subject in other complement clause constructions. When the complement clause is an infinitive, however, this seems impossible (15).

(15) a. It seems that he is mad.
b. *It seems for he to be mad.

Subject raising (i.e. raising the subject of the infinitive to the subject of the matrix clause) is not compatible with the object position. This seems to have a semantic motivation: If a verb doesn't have an agentive role – which is always true in the case of subject raising, by the definition mentioned above – then it also doesn't have a patientive role, and therefore the object position is not licensed. Subject control, on the other hand, allows the object position.

Thus, possible infinitive clause constructions with the infinitive in the VP are summarized in Table 7.1.

[2] Again, there is no WH-movement to the peripheral of the clause.

Table 7.1: Infinitive constructions with the infinitive being in VP

subject	object	example
subject raising	no object	(4)
subject control	no object	(5)
	object base generated	(16)
subject base generated	no object	(1)
	object base generated	(2)
	object control	(18)
	object raising	(17)

(16) [I]$_{i,\text{subject}}$ promise [her]$_{\text{base-generated object}}$ [—$_{i,\text{controlled}}$ to go away]$_{\text{infinitive}}$.

(17) I expect [you]$_{i,\text{object}}$ [—$_{i,\text{raised}}$ to do this tomorrow].

(18) I want [him]$_{i,\text{object}}$ [—$_{i,\text{controlled}}$ to complete this task tomorrow]$_{\text{infinitive}}$.

The above classification has direct consequence in the form of the infinitive clause. (1, 2) have no essential difference with (25). (2) is less frequent but is still attested (Dixon 2005, p. 243).

The next step is to analyze the *form* of these infinitive clauses appearing in raising/control/no-raising-or-control environments. Whenever raising or control appears, *for* is absent; and if an infinitive clause isn't meant to be involved in raising or control, then *for* is usually present, because otherwise the construction is interpreted as a control or raising construction (19). Thus, we conclude that a) The word *for* in an infinitive clause seals the clause and turns it into an NP-like construction with invisible inner structure in the eys of the syntactic environment, and b) In non-control-or-raising infinitive constructions listed in Table 7.1, the infinitive clauses are the NP-like infinitives just mentioned and are put in object-like positions. (compare (20) and (2))

(19) I want [to join your group].
'*I want someone else to join your group.'

(20) The great powers promised the Jews [an independent nation].

TODO: bare infinitives, let sb. to, make sb. to, see sb. do Huddleston and Pullum (2002, p. 1236,p. 1254)

7.2.1.4 Control compared to similar constructions

We have shown that the infinitive following a control verb is not structurally very different from a sealed infinitive. The question then is why we have the obligatory coreferential relation between the null subject and one or more of the antecedent. One analysis reasons that as usually one NP has only one semantic role,[3] the null subject of the infinitive is just a null pronoun in the infinitive, and there is no structural relation between it and its antecedent; but then the question is where the obligatoriness comes from. Another theory is that both control and raising come from movement (Hornstein 1999); but then it has to explain why control constructions seem comparable to constructions where a movement analysis is clearly impossible. The real problem

[3]The so-called θ-criterion.

then is not whether control *can* be accounted for by a movement theory or by a null pronoun theory, but whether it's appropriate to do so globally. Also, a caveat is control is not truly a homogeneous phenomenon: the long passive in German may also be classified as control, but it's definitely quite different from control infinitives that are comparable to sealed infinitives; although English doesn't have this construction, heterogenity can still be expected.

One heterogenity that can be noticed in English is the reference of the anaphoric *it* (Wurmbrand 2002). In (21), *it* has two possible references: one is the full event that Ezio fools around, the second is the action of fooling around; if Mimi doesn't like the latter, than she doesn't like fooling around *herself*, but may find her husband fooling around entertaining. In (22) we find the semantics of the matrix verb sometimes prohibits the 'Watson makes Holmes start to read case reports' reading, but the 'Watson starts to do the same thing' reading is always available. This suggests that the determination of the reference of the null subject of the control infinitive at least has something to do with the semantics, as the 'someone is doing it but I don't know who' reading is available for an anaphora.

(21) Ezio likes fooling around, but I would bet that his wife Mimi doesn't like it at all.

(22) Holmes starts to read case reports, and Watson starts it too.

This phenomenon however can also be well captured by the control-as-raising-with-semantic-role analysis (Hornstein 1999). We do note that evidence for the structural difference between control and raising constructoins is thin. The clausal passivization test doesn't really say much because passivization depends on lexical properties of the verb. The fact that a sealed infinitive and a control infinitive can both appear after a verb also says nothing definitive because the possibility of coordination of a raising infinitive and a sealed infinitive is eliminated by the requirement that a raising verb has to extract the subject of the inifinitive out, so we don't really have platforms where coordination between a raising infinitive and a sealed infinitive is not prohibited by principles other than their structural complexities. The only motivation for us to reject the analysis of control as raising with semantic role assignment to the infinitive subject is that the prototypical control is too similar to constructions that can't be easily analyzed using a movement account.

7.2.1.5 Passivization with an infinitive

Consider the below tests:

(23) a. I expect you$_i$ $-_i$ to finish this in the following week.
b. You$_i$ are expected $-_i$ to finish this in the following week.

(24) a. I want you to complete this task tomorrow.
b. *You are wanted to complete this task tomorrow.

7.2.2 Infinitive appearing in subject

It's possible for an infinitive clause to appear in the subject position (25). However, this use of infinitive clauses has nothing particularly interesting: The complement clause has almost parallel behaviors with an NP (26).

(25) [To try your best] also includes to ask for help when it's necessary.

(26) [The skill to ask for help]$_{\text{subject:NP}}$ is a strength.

TODO: Is this control? We may analyze this as (semi-)control: The null subject of *to travel a lot* seeks reference and can only has coreference with the object. However, this doesn't seem like a purely syntactic process compared with raising: Similar coreference requirements can be found in several constructions with little structural resemblance (TODO: CGEL chapters). The coreference here may be better analyzed as a semantic effect: The only "active" NP in the clause is the object *him*, and therefore the null subject has to refer to the object. This creates another problem: Whether the classical object control is also from the same mechanism.

(27) To travel a lot annoys him.

(28) To travel a lot sometimes is annoying for him.

(29) For her, to travel a lot is never a burden.

TODO: see Huddleston and Pullum (2002, p. 1269)

However, the subject of an infinitive is not always coreferential with an argument within the clause:

(30) To smoke around babies is dangerous.

If we try hard enough (actually not that hard, because what is done below clearly has semantic motivation), the *annoy* clause can receive an analysis very similar to the object control: Here *him* is the experiencer, and may be higher in the *v*P structure, and thus in a certain stage of the syntactic derivation, *him* indeed controls the null subject in *to travel a lot*. This, however, can also be done at the syntax-semantic interface: We may say *him semantically* controls the null subject of the infinitive clause when interpreted.

Syntactic versus semantic control S Wurmbrand - Linguistik aktuell/linguistics today, 2 is also useful

7.2.3 Semantic classification of infinitives

Semantically, an infinitive clause either expresses a potential situation, or a subjective judgement (Dixon 2005, p. 245). A judgement infinitive clause is always in a subject-raising construction or an object-raising construction, probably for semantic reasons: A verb taking a judgement infinitive clause takes a cognitor semantic argument, TODO: but why can't there be a *for* infinitive clause? **I find for this food to be bad*

7.2.4 Interpretation of the null subject

The *to do sth.* type of infinitive has null subject. What the null subject refers to is sometimes decided by structural factors, as in the cases of control and raising, and sometimes by semantic and pragmatic feasibility.

TODO: semantic subject interpretation

7.2.4.1 Subject-raising

(31) The boy seems unhappy.

7.2.4.2 Object raising

(32) I wanted them to start.

7.2.4.3 Control

(33) I ask them to be helpful.

Although Dixon (2005, p. 15), Dixon (2010, p. 388) argue that it's not necessary to introduce the concept of object raising in English, TODO

7.3 Quoted speech

It should be noted that direct quoted speech is not as simple as a sequence of sound: It's possible for a VP to appear as a quoted speech.

(34) In closing, they said they "stand ready and willing to help you win Michigan in 2024."

8 Relative clauses

The relative clause construction is formed by

8.1 Types of relative clauses

It should be noted that the WH-movement in relative clauses is not structurally the same as the WH-movement in interrogative constructions. Consider the pair in (1): It clearly demonstrates that the relative WH-phrase is structurally higher than the topic, while the opposite is true for interrogative constructions. This may have a semantic motivation Radford (2009, p. 330): In question formation, the WH-movement is just a marking strategy of the *focus*, which appears below the topic, while in the formation of relative clauses, WH-movement happens *last*, marks the whole clause as a relative clause, and "seals" the whole relative clause, separating its content and the matrix clause.

(1) a. [In you opinion]$_{\text{topic}}$, [what]$_{\text{focus:WH}}$ [is]$_{\text{fronted auxiliary}}$ our most urgent task right now?

b. [[What], [in his opinion]$_{\text{topic}}$, is our most urgent task right now]$_{\text{relative clause}}$ still remains unknown for the listeners.

8.1.1 Purpose relative clause

A rare type of

(2) I need [a house [to live]$_{\text{purpose}}$]$_{\text{object: NP}}$

(3) I need [a house [to live in]$_{\text{purpose}}$]$_{\text{object: NP}}$

8.2 Clause linking: subordination

8.2.1 The comparative correlative

The comparative correlative construction, also known as the *the more, the more* construction,

(4) The more you wait, the better the outcome will be.

(5) The more you wait, the better.

The comparative correlative construction is sometimes analyzed as a sui generis construction, having internal structures completely different from the rest of the grammar. This however is more likely to be illusional: it's highly like that the English

comparative correlative has a shared structure with, say, the Latin *tam … quam …* construction, and the Hindi *jo … vo …* relative correlative; in all of these constructions, first a gap is left in the main clause (in Latin, the gap is the degree phrase and is filled by *tam*; in Hindi relative correlative it's the subject and is filled by *vo*), and then a relative clause modifying the gap is created and attached to the main clause (Den Dikken 2005). Correlative constructions like these are no longer frequent in English, but some comparative constructions can still be recast into uncontroversial correlative constructions, especially in more poetic registers. Applying this analysis to the English comparative correlative, the interpretation of (4) is 'the degree how more you wait is the degree how better the outcome will be'.

(6) My plan is so/as good as yours.

(7) As above, so below.
'The below is what the above is.'

(8)

The remaining question is why *the* is used in English comparative correlative.

8.3 Clause linking: coordination

Parameters concerning variation of coordination include number of coordinates, number of coordinators, and whether we have correlative items like *both* in the coordination construction.

TODO: Huddleston and Pullum (2002, p. 1276)

This section talks about FANBOY TODO: subject extraction

8.4 Supplementation

9 Prosody, punctuation, and spelling conventions

Box 9.1: The notion of *thought group*

It should be noted that so-called *speech groups* and *thought groups* are neither morphosyntactic or semantic concepts. When the subject is light (for example, when it's a personal pronoun), the subject and the verbal complex may be grouped into one "thought group", but that doesn't mean the subject and the verb constitute a constituency: it merely arises from prosodical considerations. In real world utterances, we may also see subject-only sentences with the object omitted, but this may be about production of utterance: I already have some templates of clause structures in our mind (though with inner hierarchical structures, as in Tree-Adjoining Grammar), and after we fill the subject position of one of the templates, somehow – possibly because I forget what to say – I just stop filling the template and pour the half-finished sentence to the one I'm talking with. But still, though groups have something to do with stress allocation, which indeed is influenced syntactically (Kahnemuyipour 2009, p. 7).

10 Notable variations

10.1 Early Modern English

Dialects, etc.

References

Anne Abeillé and Danièle Godard. *La grande grammaire du français*. Éditions Actes Sud, 2021.

Josef Bayer and Markus Bader. On the syntax of prepositional phrases. 2007.

Walter Biemel and Parvis Emad. The development of heidegger's concept of the thing. *The Southwestern Journal of Philosophy*, 11(3):47–66, 1980.

Seth Cable. Pied-piping: Introducing two recent approaches. *Language and Linguistics Compass*, 6(12):816–832, 2012.

Seth Cable. Pied-piping: Comparing two recent approaches. *Language and Linguistics Compass*, 7(2):123–140, 2013.

Guglielmo Cinque. *Adverbs and functional heads: A cross-linguistic perspective*. Oxford University Press, 1999.

Francesco D'Amico, Maurizio E Musso, Raphael JF Berger, Nicola Cefarin, Giovanni Birarda, Gianluca Tondi, Durval Bertoldo Menezes, Andreas Reyer, Letizia Scarabattoli, Thomas Sepperer, et al. Chemical constitution of polyfurfuryl alcohol investigated by ftir and resonant raman spectroscopy. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 262:120090, 2021.

Marcel Den Dikken. Comparative correlatives comparatively. *Linguistic Inquiry*, 36 (4):497–532, 2005.

Andrew M Devine and Laurence D Stephens. *Latin word order: Structured meaning and information*. Oxford University Press, 2006.

RM Dixon. Basic linguistic theory. volume 3: Further grammatical topics, 2012.

Robert MW Dixon. *A semantic approach to English grammar*. OUP Oxford, 2005.

Robert MW Dixon. *Basic linguistic theory volume 1: Methodology*, volume 1. OUP Oxford, 2009.

Robert MW Dixon. *Basic linguistic theory volume 2: Grammatical topics*, volume 2. Oxford University Press on Demand, 2010.

Nigel Fabb. English suffixation is constrained only by selectional restrictions. *Natural language & Linguistic theory*, pages 527–539, 1988.

Patrick Farrell. English verb-preposition constructions: Constituency and order. *Language*, pages 96–137, 2005.

Diana Forker. *A grammar of Sanzhi Dargwa*. Language Science Press, 2020.

Dianne Friesen. *A grammar of Moloko*. Number 3 in African Language Grammars and Dictionaries. Language Science Press, Berlin, 2017. doi: 10.5281/zenodo.824016.

Chiara Gianollo, Klaus von Heusinger, and Maria Napoli. Reference and quantification in nominal phrases: The current landscape and the way ahead. In *Determiners and Quantifiers*, pages 1–28. Brill, 2021.

Adele E Goldberg and Farrell Ackerman. The pragmatics of obligatory adjuncts. *Language*, pages 798–814, 2001.

Jane Grimshaw and Sten Vikner. Obligatory adjuncts and the structure of events. In *Knowledge and Language: Volume II Lexical and Conceptual Structure*, pages 143–155. Springer, 1993.

Heidi Harley. The bipartite structure of verbs cross-linguistically (or: Why mary can't exhibit john her paintings). In *Write-up of a talk given at the 2007 ABRALIN Congres in Belo Horizonte, Brazil*, 2007.

Heidi Harley. 427 A Minimalist Approach to Argument Structure. In *The Oxford Handbook of Linguistic Minimalism*. Oxford University Press, 03 2011. ISBN 9780199549368. doi: 10.1093/oxfordhb/9780199549368.013.0019. URL https://doi.org/10.1093/oxfordhb/9780199549368.013.0019.

Heidi Harley and Rolf Noyer. Distributed morphology. *Glot international*, 4(4):3–9, 1999.

Martin Haspelmath. 341 Framework-Free Grammatical Theory. In *The Oxford Handbook of Linguistic Analysis*. Oxford University Press, 12 2009. ISBN 9780199544004. doi: 10.1093/oxfordhb/9780199544004.013.0014. URL https://doi.org/10.1093/oxfordhb/9780199544004.013.0014.

Caroline Heycock. 54 Japanese - WA, - GA, and Information Structure. In *The Oxford Handbook of Japanese Linguistics*. Oxford University Press, 11 2008. ISBN 9780195307344. doi: 10.1093/oxfordhb/9780195307344.013.0003. URL https://doi.org/10.1093/oxfordhb/9780195307344.013.0003.

Norbert Hornstein. Movement and control. *Linguistic inquiry*, 30(1):69–96, 1999.

Rodney Huddleston and Geoffrey K. Pullum. *The Cambridge Grammar of the English Language*. Cambridge University Press, 2002. doi: 10.1017/9781316423530.

Tabea Ihsane and Genoveva Puskás. Specific is not definite. *GG@ G*, 2:39–54, 2001.

Guillaume Jacques. *A grammar of Japhug*, volume 1. Language Science Press, 2021.

Arsalan Kahnemuyipour. *The syntax of sentential stress*. OUP Oxford, 2009.

Heidi Elizabeth Klockmann. The article a (n) in english quantifying expressions: A default marker of cardinality. 2020.

Masatoshi Koizumi. *Phrase structure in minimalist syntax*. PhD thesis, Massachusetts Institute of Technology, 1995.

Christopher Laenzlinger. French adjective ordering: Perspectives on dp-internal movement types. *Lingua*, 115(5):645–689, 2005a.

Christopher Laenzlinger. Some notes on dp-internal movement. *Generative Grammar in Geneva*, 4:227–260, 2005b.

Christopher Laenzlinger. A view of the cp/dp-(non) parallelism from the cartographic perspective. *Languages*, 2(4):18, 2017.

Mark Liberman and Richard Sproat. The stress and structure of modified noun phrases in english. *Lexical matters*, 131181, 1992.

Christopher Lyons. *Definiteness*. Cambridge University Press, 1999.

Andrew McInnerney. The argument/adjunct distinction does not condition islandhood of pps in english. *Linguistic Inquiry*, pages 1–13, 2023.

Will Nediger. *Unifying structure-building in human language: The minimalist syntax of idioms*. PhD thesis, UMich, 2017.

Will Oxford. Fission in algonquian and the status of morphological templates. In *Proceedings of the 23rd Workshop on the Structure and Constituency of Languages of the Americas*, pages 78–92, 2019.

Elizabeth Pearce. Completing and terminating: on aspect marking in unua. *The Languages of Vanuatu: Unity and Diversity*, pages 23–52, 2015.

Jeffrey Punske. Three forms of english verb particle constructions. *Lingua*, 135:155–170, 2013.

R. Quirk, R. Quirk, S. Greenbaum, D. Crystal, L. Geoffrey, Pearson Education, Pearson Longman, D.S.E.U.S. Greenbaum, G. Leech, J. Svartvik, et al. *A Comprehensive Grammar of the English Language*. General Grammar Series. Longman, 1985. ISBN 9780582517349.

Andrew Radford. *Analysing English sentences: A minimalist approach*. Cambridge University Press, 2009.

Keren Rice. Let the language tell its story? the role of linguistic theory in writing grammars. *TRENDS IN LINGUISTICS STUDIES AND MONOGRAPHS*, 167:235, 2006.

Norvin Richards. Multiple case assignment and the english pseudo-passive. *A Schrift to Fest Kyle Johnson. Linguistics Open Access Publications*, pages 313–319, 2017.

Thomas Roeper and Margaret Speas. *Recursion: Complexity in cognition*, volume 43. Springer, 2014.

Florian Schäfer and Elena Anagnostopoulou. Case and agreement in distributed morphology. *Handbook of DM*, 2020.

Ur Shlonsky. The cartographic enterprise in syntax. *Language and linguistics compass*, 4(6):417–429, 2010.

Daniel Siddiqi. *Syntax within the word: Economy, allomorphy, and argument selection in Distributed Morphology*, volume 138. John Benjamins Publishing, 2009.

Dominique Sportiche, Hilda Koopman, and Edward Stabler. *An introduction to syntactic analysis and theory*. John Wiley & Sons, 2013.

Peter Svenonius. Spatial P in English. In *Mapping Spatial PPs: The Cartography of Syntactic Structures, Volume 6*. Oxford University Press, 08 2010. ISBN 9780195393675. doi: 10.1093/acprof:oso/9780195393675.003.0004.

Marleen Susanne Van de Vate. Tense, aspect and modality in a radical creole: The case of saamáka. 2011.

Elly Van Gelderen. *Clause structure*. Cambridge University Press, 2013.

Sten Vikner. Reichenbach revisited: One, two, or three temporal relations? *Acta linguistica hafniensia*, 19(2):81–98, 1985.

Susi Wurmbrand. The structure (s) of particle verbs, 2000.

Susi Wurmbrand. Syntactic versus semantic control. *Linguistik aktuell/linguistics today*, 53:93–127, 2002.

Hedde Zeijlstra. *Negation and negative polarity*, page 793–826. Cambridge Handbooks in Language and Linguistics. Cambridge University Press, 2013. doi: 10.1017/CBO9780511804571.027.