Reading notes of Cambridge Grammar of English Language

Jinyuan Wu

June 17, 2022

This note is a reconstruction of the contents The Cambridge Grammar of the English Language (Huddleston and Pullum 2002, henceforth CGEL).

1 Theoretical preliminaries

For the discussion on the underlying theoretical framework, see the relevant chapter in this note. Briefly speaking, the framework is coarse-grained Minimalism: invisible functional heads are erased, but dependency relations created by the functional projections are kept (Fig. 1(a)), and the corresponding CGEL tree is obtained by replacing the name of dependency relations by syntactic function-form pairs (Fig. 1(b)). The head is defined as the dominant *lexical* word. This makes CGEL look like old-fashioned X-bar theory, where a lexical word projects into a phrase (CGEL § 5.2 [5]), but there can be several "maximal projections" headed by the same lexical word (CGEL § 5.2 [11]). This is actually consistent with Minimalism (Fig. 2).

The standard of "lexical", however, varies from one author to another. A PP and a CaseP (in generative terms) are similar objects, but in the coarse-grained CGEL framework, the latter is definitely an NP, while the status of the first is kind of controversial: should we recognize the preposition as a lexical word, or as a functional word? In the first analysis, a PP in generative terms is still a PP, while in the second analysis, a PP in generative terms is an oblique NP. Choosing between the two analyses is merely a problem of notation. But it is good practice to choose a notation that hints the readers about certain properties of the construction in question. For example, if a PP is analyzed as a PP, then probably the preposition has some predicative properties: a PP is therefore more like a VP, not an NP with case marking (or "CaseP" in generative terms, especially in theories focusing on morphosyntax and do not make sharp contrast between syntax and morphology). If, however, prepositions in a language work just like case markers, the argumentation raised in Dixon (2009, § 1.11, § 5.4) is then attempting: we should recognize the prepositions as case markers with standalone phonological realizations. Which analysis to choose should be decided according to observed phenomena. This is summarized by the slogan "describing a language in its own terms" – but the slogan does not go into conflict with generativism.

Not all dependency relations can be expressed purely by the context-free phrase structure grammar (Pullum and Rogers, 2008). Movements (or cross-serial dependencies in dependency terms) in Minimalism are represented by a gap (CGEL § 2.2 [5]), or by indirect dependency (CGEL § 5.14.1).

The close relation between CGEL and generative syntax apparently deviates from the common descriptive practice, especially the common descriptive framework extracted and summarized by Dixon (2009, 2010, 2012) and named as Basic Linguistic Theory (BLT) by Dixon, but this distinction is illusory: while BLT assumes a flatter phrase structure, information coded by the binary-branching phrase structure in CGEL is coded by dependency arcs in BLT (Fig. 3), and the hierarchy ordering of the dependency relations are coded by notions like "pipeline ordering" and "scope", e.g. "a prototypical passive construction takes in an AVO argument structure and turns the deep A into the surface S, the deep O into the surface E" (which in generative terms means VoiceP is higher than vP), and "the scope of *the* in Fig. 3 is *fat man*, indicating that the latter is definite, and the scope of *to* is *the fat man*".

Though several disagreements with the mainstream generative syntax is raised in CGEL, and much more severe accusations are made in BLT, it can be seen all the three approaches are describing almost the same complexity class of grammars.

It should be noted that even though CGEL comes close with the lexical decomposition-style Minimalism, sometimes binary branching is still not available in the surface-oriented analysis, and

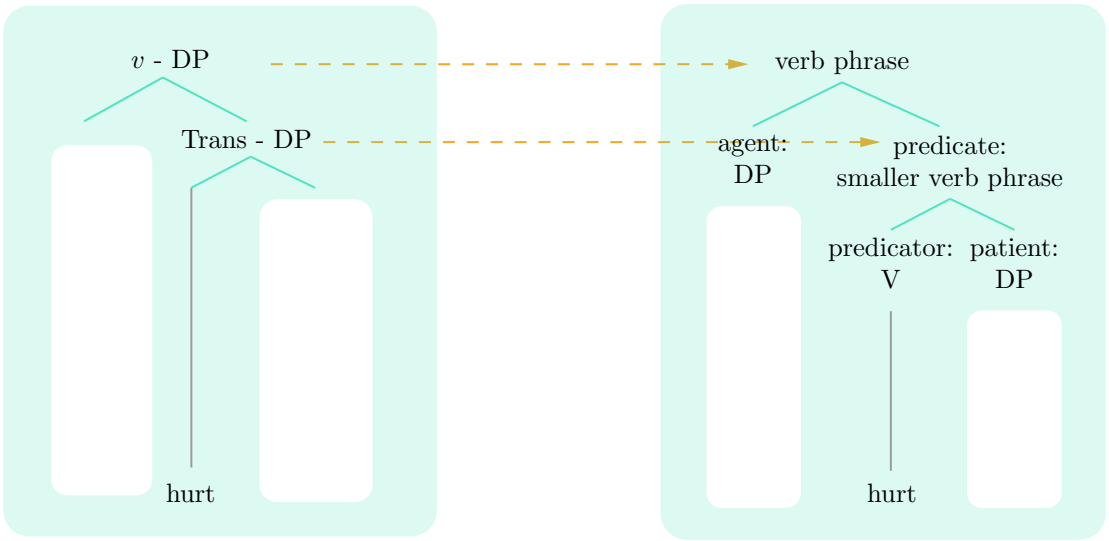

Figure 1: From coarse-grained Minimalist derivational tree to CGEL tree

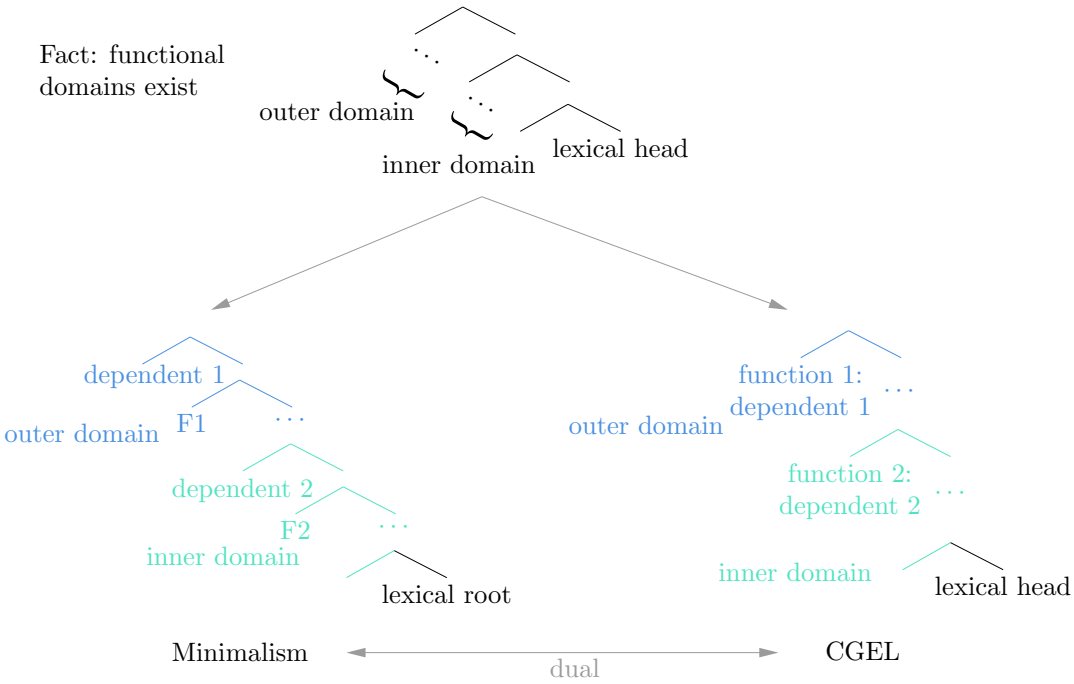

Figure 2: Duality between CGEL and Minimalism

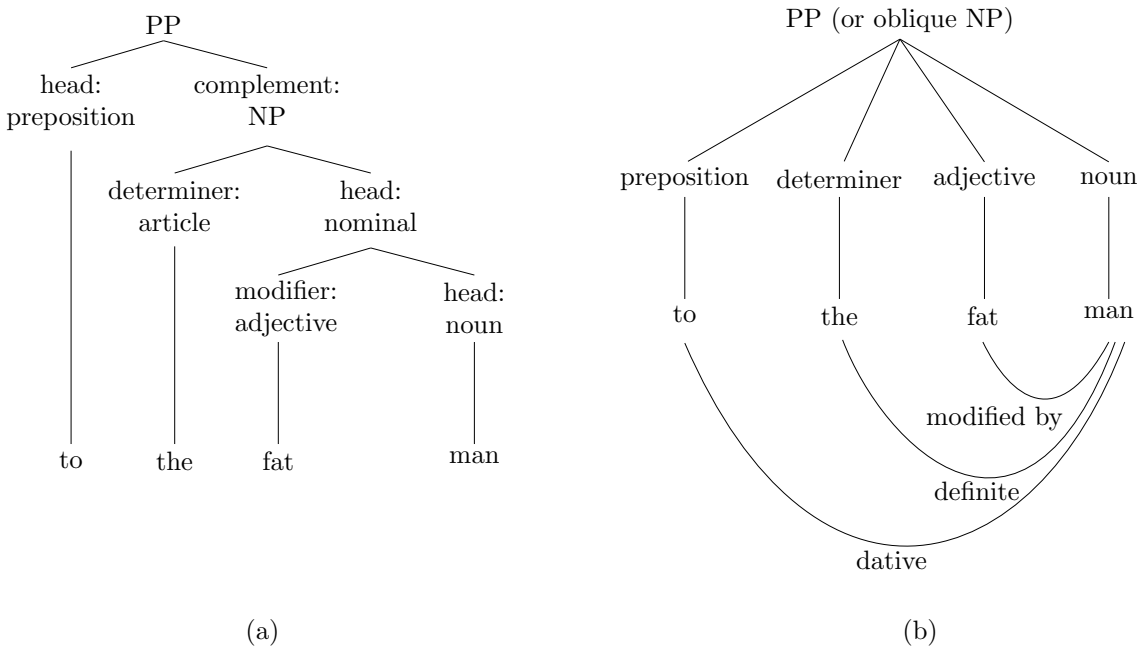

Figure 3: Comparison between the CGEL and BLT analyses of *to the fat man*

some words about the underlying dependency relations put aside the tree is needed, just as in the case of BLT. An example is ditransitive verbs. Even though in generative analysis we have plenty of reasons to assume a *v*P structure like [somebody [something SHOW]] in the clause *I showed him the photo* (possible reasons including binding effects, passivization preference, etc.), the head movement (or its post-syntactic version) breaks the binary branching structure in the surface-oriented analysis of the clause, and hence in CGEL we get a rare ternary tree (CGEL § 12.3.1 [4]).

2 Overview of clause structure

2.1 The pipeline of clause building

This section and several following sections are about the phrase structure of the clause. I will discuss clause dependents, their forms and functions, grammatical systems in the clause, and how everything is put together.

Clausal grammars of nominative-accusative languages can fit more or less into the following paradigm:

1. Clausal complements are fed into the argument structure[1] or in other words *v*P. They may be NPs, adverbials or subordinated clauses.

2. Then several clausal grammatical systems are employed on the verb-complement complex or "small clause", including agreement (or argument indexation more generally), case marking[2]

[1] This is not the term used in CGEL. For discussion on *argument* and *complmenet*, see § 3.1.1.

[2] Technically, case marking involves the argument structure. The nominative-accusative alignment, for example, can be summarized as "the patient-like argument of a transitive verb receives the accusative case, while anything promoted to the subject position receives the nominative case". In more generative terms, we have "the transitive *v*P (or some functional projection lower than Spec*v*P) assigns the accusative case, while T assigns the nominative case". The first half of the generalization means the assignment of the accusative case is directly related to the argument structure. But we can change the wording of the above generalization easily. For example, an equivalent formulation may be "in the TP, any DP with a lower position than some DP else receives the accusative case, while the subject receives the nominative case". Yet we have a third expression: "in the TP, DPs receive the accusative case

and non-spatial settings (or Tense, Aspect, Mood, Evidentiality (TAME) as people call them). The syntactic prominence of the subject is also introduced in this stage (commonly known as TP), but without adjuncts it is hard to see the structural prominence of the subject, and without question formation, etc, it is hard to see the consequences of the prominence. The result of these steps, after including obligatory speech act marking, is a minimal canonical clause.

3. Several adjunctions can be made to the minimal clause, conveying information like time, position, manner, etc. This step may be seen as an additional one, or it may be considered as occurring together with TAME. In the latter analysis, we need an explicit subject promoting analysis, which makes the subject in a higher position than any other complements. The result of these steps, after including obligatory speech act marking, is a canonical clause.

4. Speech act information (or "force" in generative terms, or "mood" in BLT) – indicative, interrogative, imperative – are added to the result of TAME marked and possibly adjoined small clauses. These stages occur in the CP domain.

   After these steps, the clause is now fully assembled. If the speech act information marked meets the standards set in the language (e.g. the clause is finite, there is no subordination markers like a complementizer, etc.), the clause is a qualified *sentence*, i.e. it can occur independently in utterance. But it can be embedded (or subordinated) into other clauses, too, probably with some additional marking. A clause without any subordinated clauses as complements or adjuncts is called a **simple clause**. Otherwise it is a **complex clause**.

   Indicative marking is often (but not always) zero, and in this way there is no obligatory speech act marking for clauses.

5. The canonical clause may further undergo processes like topicalization or focusing. This makes it non-canonical.

   Topicalization and focusing may be analyzed as happening together with speech act marking and/or subordination. In the classical generative analysis, the CP domain is roughly Force-Focus-Topic-Finiteness, with the left feature being introduced after the right feature. Subordination involves both finiteness (gerund clauses can be subordinated, but they are never independent sentences) and force (subordinated clauses often have limited force choices, and the complementizer has to be introduced at a certain position). So in this analysis, there is no strict time ordering between topicalization etc. and subordination.

   But it is also possible to introduce topicalization and focusing with transformational rules, since most grammars are much more surface-oriented. In this case, it still makes sense to say one happens after another.

   Transformational rules are largely abandoned in contemporary generative syntax. In a surface-oriented grammar, transformational rules does not say anything about the plausible generative derivational[3] process in our brain. If construction A and construction B are connected by a transformation, it merely means the derivational process of A and B have some similar stages and there is a uniform correspondence between them.

6. There are usually certain kinds of valency-changing devices. They may be implemented by a Latin-like voice system, or by auxiliary verbs the subject of which undergoes an event, and the complement of which is a clause or small clause indicating the event, or by certain kind of *v*Pconstructions. There are no strict boundary between these strategies. Combination of these devices, like Collins (2005), is also possible. Just like the case of arguments and adjuncts, and the case of topicalization and speed act marking, passivization can be described with a derivational account as in contemporary generative syntax, as well as a transformational approach.

as the default case, while the subject receives the nominative case". It is therefore acceptable to take case marking completely away from the argument structure.

[3]From then on, *derivational process*, etc. all mean Minimalist ones. But *derived* may be associated with surface-oriented transformational rules.

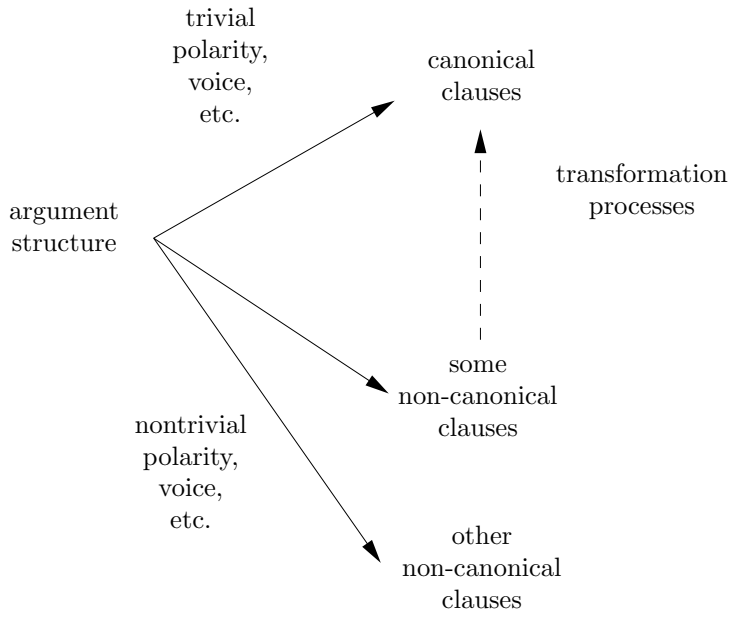

Figure 4: Canonical and non-canonical clauses

7. Each stage of the above processes may be grammaticalized or be realized by a grammaticalized construction. The so-called Chinese passive construction, the *bei*-construction, is just a grammaticalized complement clause construction with the verb *bei* expressing a meaning of "suffer from".

8. Finally, sentences and clauses can be coordinated.

The pipelines of the machine of English clause structure is also organized in this way. Valency is an enduring topic in English grammar. Tense and aspect (usually denoted together as *tense*) appears in every introduction of the English language: do something, did something, be doing something. The subjunctive mood (or "modality" in BLT) is especially necessary in formal or archaic writing. English adverbs are traditionally ignored but are more complicated than many will think. The list can be very long.

2.2 Clause types

CGEL uses the strategy to first describe a canonical clause and then use syntactic processes (or transformational rules in generative terms) to derive non-canonical ones. Several syntactic processes may be used together to derive a clause with multiple non-canonical features.

It should be noted that when the definition of canonical clauses is narrow (a wise decision leading to easier starting), it is possible that some non-canonical clauses cannot be obtained by applying well-recognized syntactic processes (CGEL § 2.2 [3]). This is shown in Fig. 4.

It is already reviewed in § 2.1 that clauses, strictly speaking, are fully marked with speech act information and hence are CPs themselves. Some people use the term *clause* for TPs and *sentence* for CPs. This does not agree with the acceptable terminology in surface-oriented studies, since a pure TP without CP layers rarely occurs. Therefore, in CGEL, no explicit discussion is made about the pipelines of clause building: the approach accepted is to present grammatical relations as *static* ones. Recall Fig. 3. Though (a) has more embedding hierarchies than (b), while (b) uses dependency relations in lieu of the binary branching hierarchies, both of them are static ones: (a) is presented in the grammar *as a whole*, i.e. the grammar does not treat (a) as built by first merging *fat* and *man* and then merging the result with *the* and then *to*. Rather, once a PP is discussed, all grammatical relations involved – attributive modification, definiteness, dative construction – are considered as completed. Though we know a

2.2.1 Canonical clauses

2.2.2 Interrogative

2.3 Clausal grammatical relations

2.3.1 The argument structure

2.3.2 The subject

English is a subject-prominent language. This is to say, in every clause there is a **subject** (which may be omitted in limited cases) which is usually filled by an agent-like argument and is somehow "higher" in the clause structure, and therefore is subject to several types of extraction, including relativization and coordination pivot.

In typological terms, this means English is a syntactic nominative-accusative language, because its syntactic obligatory topic (something somehow "higher" in the clause structure, but lower than the topic in topicalization) is identical to the agent-like argument.[4] In ergative languages the notion of *subject* is more complicated, because in transitive clauses, the two do not coincide. But this is not the case for English.

2.3.3 Agreement

Agreement is only found between the subject and

2.3.4 Case marking

2.3.5 Tense and aspect

3 Dependents in the clause nucleus

3.1 Types of complements

Clausal (or verbal, since the clause is headed by the verb) complements may be NPs and PPs, and less frequently, adverbs (as in *He treated us [kindly]*).

This section lists some criteria of classification of complements. They are all discussed in CGEL § 4.1.1. Though the starting point of the building process of all clauses is the argument structure, CGEL starts chap. 4 with discussion on clausal complements. This is a wise decision for a largely surface-oriented grammar, and also does not obscure the argument structure, because how the arguments (or non-argument complements) of the verb fills the *clausal* complement positions is largely based on the structural distribution of these arguments in the *v*P structure. In other words, the clause complement positions – subject, direct and indirect objects, etc. – are *syntactic* marking of coarse-grained S, A, O, E arguments, the latter being reflection of the relative positions of arguments in *v*P.[5]

How to tell a complement from adjuncts is a question addressed in § 3.5. § 3.5 is delayed after this section.

3.1.1 Core v.s. oblique

One classification standard is the make up of the complement. A **core** complement is a complement with similar morphosyntactic properties of NP complements. A **non-core** complement is a

[4] In languages like Chinese, though when a semantically obvious agent is present, it is always the syntactic obligatory topic, there are quite common clauses in which the subject is not agent-like at all, like the famous *tai shang zuo zhe zhu xi tuan*. But a deeper analysis will show that this can be attributed to the rich light verb inventory of Chinese, which may be understood as "mini-voice" in more descriptive terms, and therefore does not impose any threat to the nominative-accusative status of Chinese.

[5] O is the label of the patient-like or *patientive* argument in BLT. In many modern grammars, the symbol is replaced by P. In this note the symbol O is used to be consistent with the notation in CGEL, though CGEL itself uses O as a symbol of clausal complements.

complement with similar morphosyntactic properties of PP complements. If a non-core complement itself takes an NP complement (or something with similar morphosyntactic properties), the latter is called an **oblique**.

Note that in CGEL, the term *argument* is reserved for purely semantic objects. A clausal complement is therefore the syntactic incarnation of an argument, but itself is not an argument. This is not the way *argument* is used in BLT.

It should also be noted that in CGEL, the terms *non-core* and *oblique* are reserved for clausal complements and the NP part of PP clausal complements. They *do not* include adjuncts with similar forms. On the other hand, in BLT, the term *peripheral argument* covers both complements and adjuncts. The term *oblique* is often associated with *oblique cases*. In traditional Latin grammar, cases other than the nominative and the vocative cases are all called oblique cases. In other usages, oblique cases exclude the accusative case. In English the case system has largely collapsed, and the name *oblique* does not have much morphological consequences: an oblique complement is never nominative as we will see, and that is all.

The prototypical definition of core and oblique complements are based on syntactic forms instead of functions, while the definition is extended by analog with respect to syntactic functions. Whether these terms are useful is a question we need to wait and see.

3.1.2 External and internal

The subject is the **external** complement for obvious reasons. All other complements are **internal**.

In English, internal complements include **objects** and **predicative complements**. The objects split into **direct objects** and **indirect objects**.

3.1.3 Relation with the argument structure

In the typological perspective, both the coarse-grained argument structure and the coding strategy of the argument slots (or semantic roles) as clausal complements are quite straightforward in English.

The argument structure of any verb in English fits into the S, A, O, E paradigm in BLT. Some typological studies have G and T abstract semantic roles, but in English no goal-like and theme-like semantic role classes with stable syntactic appearance can be established. Consider, for example, the example in BLT § 3.3 (6) and (7), with semantic role labels replaced by ones in CGEL § 4.2.2:

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

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

The PP *to charity* and NP *some books* have similar syntactic behaviors, and in a similar manner, *all his goods* and *his facorite students* form another group of arguments with similar syntactic appearance. The first two cannot be promoted to the subject position in passivization, while the latter two can. It is, therefore, reasonable to name the group containing *all his goods* and *his facorite students* with the label O, and name the group containing *to charity* and *some books* with the label E. Note, however, there are some heterogeneity in each group: *all his goods* is a theme while *his favorite student* is a goal, but they have similar syntactic properties. So the division between O and E is useful in English (which is also shown in footnote 26 in CGEL chap. 4), while the division between G and T is not.

In canonical clauses, A and S arguments are consistently coded as the subject, both syntactically and, in the case of pronouns, morphologically. No split of S arguments is easily observable. O arguments are uniformly coded as clausal objects. E arguments may be coded as clausal objects (as in (2)) or obliques (as in (1)).

3.2 Types of minimal canonical clauses

The contents of § 3.2.1 and § 3.2.2 are covered in CGEL § 4.1.1.

3.2.1 Subject and object(s): transitivity and valency

In English every clause has a subject, so the number of subjects is not a parameter. The number of objects may be 0, 1, and 2, which is denoted as **intransitive**, **transitve**, and **ditransitive**. This parameter is named as **transitivity**.

3.2.2 The number of predicative complements

There may be 0 or 1 **predicative complement (PC)**. It is impossible for a ditransitive clause to have a PC. So the parameters of transitivity and PC gives the classification of clauses with respect to their complements as in CGEL § 4.1.1 [9]. Hence we have the notion of **valency** (CGEL § 4.1.1 [10]). The number of PC is not included into the valency.

CGEL § 4.1.1 [9] is just the classification based on S, O, and PC. It does not mean there are no other complements. Nor does it say anything about possible other complements.

3.2.3 Five canonical clauses

3.3 Types of adjuncts

3.4 Semantic roles

Once clausal complement positions are related to coarse-grained argument positions (§ 3.1.3), the question becomes how concrete semantic roles of verbs fit into the paradigm.

3.4.1 The causer-like group

3.5 Distinguish complements from adjuncts

3.6 Summary

4 The verb category

This section deals with subcategories in the verb category, and how the verb changes its form according to the syntactic environment.

4.1 Transitivity and valency

References

Chris Collins. A smuggling approach to the passive in english. *Syntax*, 8:81 – 120, 09 2005. doi: 10.1111/j.1467-9612.2005.00076.x.

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

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.

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

Geoffrey K Pullum and James Rogers. Expressive power of the syntactic theory implicit in the cambridge grammar of the english language. In *annual meeting of the Linguistics Assocition of Great Britain, University of Essex*, 2008.