



# An Analysis of Hierarchy in English Clause Combination

Clarence Green

School of Languages and Linguistics, University of Melbourne

Bld. 139 Parkville, Melbourne 3010, Australia

E-mail: c.green4@pgrad.unimelb.edu.au

Received: 14-05- 2012

Accepted: 08-06- 2012

Published: 01-07- 2012

doi:10.7575/ijalel.v.1n.2p.1

URL: <http://dx.doi.org/10.7575/ijalel.v.1n.2p.1>

## Abstract

The following study presents research into English clause combination that describes how the types of clauses in English are distributed along a hierarchy of grammatical integration, in a much more complex fashion than the traditional coordination/subordination dichotomy suggests. This hierarchy extends and synthesizes previous descriptions of English combined clauses found in the most referenced descriptive grammars of English. A corpus analysis of patterns in the combined clauses that may be a consequence of hierarchy was also conducted. A corpus of 50 examples of each form was coded for tense/aspect continuity, subject continuity and syntactic function. The analysis confirmed that the different formal levels of integration amongst the clauses are reflected in their functional and discourse patterns. The implications of the study are that English clause combination might be beneficially described and taught as hierarchical beyond traditional binary categories, as combined clauses are a range of more or less integrated structures with distinct properties. For teachers and advanced learners the research may help in an understanding of the relationship in English grammar between function, form, discourse and syntax.

**Keywords:** English grammar, clause combination, corpus linguistics, coordination, subordination

## 1. Introduction

The combining of clauses is central feature of the English language. Beyond their importance in theoretical linguistics which has argued clausal recursion is a unique feature of human communication and the language faculty (Chomsky, 2007), knowledge of the patterns of form and function across the different clause structures is one of the more complex aspects of understanding English grammar for ESL students (Sjolie, 2006). It is also of importance for native speakers developing communicative competence appropriate for different genres and registers that require certain styles of language use (Biber et al, 1999). The study of English therefore requires a description of English grammar in which a central component is a precise analysis of the combined clauses of the language. This study aimed to contribute to that goal. It synthesized the discussion of clause types from the most widely referenced descriptive grammars in English language study, the 'bestsellers' of The Longman Grammar of Spoken and Written English (Biber et al, 1999), The Cambridge Grammar of the English Language (Huddleston et al, 2002), The Collins Cobuild English Grammar (Sinclair et al, 1995) and The Comprehensive Grammar of the English Language (Quirk et al, 1985). From this synthesis, the study proposed a description of clause combination that centred on hierarchy amongst the clauses, something lacking in these grammars. For further support of the legitimacy of the proposed hierarchical description of English clauses, a corpus analysis was conducted. This indicated that hierarchy has verifiable consequences for other aspects of English grammar and discourse, such as the typical patterns in combined clauses regarding their syntactic function and their continuity of tense, aspect, and subjects.

## 2. Descriptions of Combined Clauses in the Central English Grammars

The binary categories of coordination and subordination have a long history in descriptions English grammar. However, it may be that this tradition of beginning any description of clause combination with the distinction is all that keeps it central to the study of English, particularly in pedagogical textbooks, rather than any centrality the notions have to the grammatical system itself (Cristofaro, 2005). The notion that there are two classes of clause combination in English grammar, one of grammatical dependency and one of equality, oversimplifies a combined



clause grammar far too complex to be meaningfully categorized into such binary classes (Payne, 2011; Givón, 2001).

The purpose of this section is to review previous descriptions of clause combination in the major grammars of English, and summarize the central complexities and differences amongst them. This section highlights areas in which a synthesis of grammatical analyses of English combined clauses can be created for the current study. The grammars reviewed can be considered the most important descriptions of clauses combination in English language study in that they are the most used and most cited in English linguistics; they are the 'best-sellers', the weighty reference tomes in institutions and offices everywhere. There are four of these major grammars of the English language which present descriptions of clause combination: The Longman Grammar of Spoken and Written English (Biber et al, 1999), The Cambridge Grammar of the English Language (Huddleston et al, 2002), The Collins Cobuild English Grammar (Sinclair et al, 1991) and The Comprehensive Grammar of the English Language (Quirk et al, 1985). For ease of reference, Appendix A provides a list of the combined clause classifications of each of the grammars, indexing all the forms and functions reviewed in this section.

The core grammars of English can be divided into two contrasting, and difficult to reconcile, approaches. One approach is to describe English grammar and clause combination using grammatical form (i.e. what it looks like) as a starting point, while others take grammatical function (i.e. what it does) to be the primary indicator of a clause type. There are two major English grammars that take a primacy of function approach: the corpus-based The Collins Cobuild English Grammar (Sinclair et al, 1991) and The Comprehensive Grammar of the English Language (Quirk et al, 1985).

### 2.1. Function-first Descriptions of English Clause Combination

The Collins Cobuild (Sinclair et al, 1991) makes the traditional distinction between two clause combination methods: coordination and subordination, with a corollary binary distinction between coordinate and subordinate conjunctions. Coordinating conjunctions are not considered a unique conjunction in most descriptive grammars, but in the Collins Cobuild (Sinclair et al., 1991) are rather extensive, including *and*, *but*, *nor*, *or*, *yet*, *then*, *so*, as well as the combinations, *and also*, *and yet*, *and then*, *and so*. The grammar describes one type of coordinate clause, and three types of subordinate clauses: adverbial, relative and reported clauses. Because these are functional classifications, clauses with different forms may nonetheless be the same type of clause. For example, all clauses can be finite or non-finite. One of the four subordinate clause types, identified as the 'reported clause', might have a grammatical form beginning with 'that' and be finite, as in 'she said *that she was eating her dinner*', or it may begin with a non-finite infinitive, 'she said *to eat your dinner*'. Both are reported clauses because they function as complement to a specific verb of saying or thinking and report the relevant statement or thought.

Similarly, relative clauses may be a non-finite form, 'the girl *reading the book*', or have a finite structure, such as 'the girl *who is reading the book*'. Relative clauses are further subcategorized into adjectival functions, those which post-modify a noun phrase as in the previous examples, or nominal, as in '*that the lecture continues through lunch* is a problem'. Adjectival relatives are either defining, as above, or non-defining when placed in parenthetical commas. Another central clause type is the adverbial clause. Eight types of adverbial clauses are described: 1. Time, 2. Condition, 3. Purpose, 4. Reason, 5. Result, 6. Concessive, 7. Place, 8. Manner. Typically adverbials are associated with a clausal form that begins after a comma and contains a subordinator or adverb, as in 'students finish quickly, *when/after/before they eat*' (adverbial of time). However, given the functional approach of the Collins Cobuild (Sinclair et al 1991) adverbial clauses can also be non-finite forms, as in 'students finish quickly *to get out of class early*' (adverbial of purpose).

The other major grammar of English with a function-first approach is The Comprehensive Grammar of the English Language (Quirk et al, 1985). This describes a four way functional classification of subordinate clauses, fulfilled by a range of forms. The types are nominal clauses, adverbial clauses, relative clauses, and comparative clauses. Three main forms are associated with the four functional types: finite clauses, non-finite clauses, and verbless clauses. Nominal Clauses share the definition given by the Collins Cobuild (Sinclair et al., 1991) above for 'relative nominal clauses'. They are essentially any clause able to function as an argument, e.g. a clause that is itself a subject, object or copula predicate in a sentence. However, to highlight how confusing different descriptions of English grammar can be, the label 'nominal relative clause' used by the Collins Cobuild (Sinclair et al, 1991) is also used in the Comprehensive Grammar (Quirk et al, 1985) but for something different. Here, the label is used for a fused relative clause, a specific subtype of the nominal clause (illustrated in example number 5 below). Nine subtypes of the nominal clause are described, five are finite, the subsequent three non-finite and the last verbless: 1. 'That' clause (*that the analysis was incomplete* didn't undermine it), 2. Wh- interrogative (I didn't understand *what the teacher meant*), 3. Yes-no interrogatives (she asked *whether I was hungry*), 4. Exclamatives (I



liked *how professional it was*), 5. Fused nominal relatives (*what I saw was so funny*), 6. To-infinitive (*to have another one couldn't hurt*), 7. ing participle (*having a coffee before work is essential*), 8. Bare infinitival (the cleaner saw the dog *make another mess*), 9. Verbless (*I finished the entire exam, my phone on in pocket*).

The second combined clause type in *The Comprehensive Grammar* (Quirk et al, 1985) is the adverbial clause, subcategorized into four further syntactic functions. The description is quite different from the description of adverbial clauses in the *Collins Cobuild* (Sinclair et al, 1991). Here, they function as either conjuncts for discourse continuation (the book was good, *since it was short*); subjuncts for viewpoint elaboration (the book was good, *as far as I am concerned*); adjuncts to denote circumstance (*I finished the work, while I was on my break*); or disjuncts for style, content and attitudinal comments on the main clause (the idea is nonsense, *however you look at it*). Comparative clauses are the third clause identified in the grammar and have two subtypes: equivalence, marked by correlative 'as...as', and non-equivalence marked by the comparative '-er' morpheme and 'than'. Relative clauses, the final of the four functional types of English combined clauses, are either restrictive or non-restrictive. However, it becomes perplexing, as the grammar in later chapters also describes four grammatical forms of post-modifying finite clauses: 1. Appositive (the book, *which I've never read*, is said to be a bestseller), 2. Nominal relatives (see previous), 3. Sentential (*I knew it would be boring, which is why I didn't go*), 4. Relative (*It's the job that I want*).

Clearly there is some confusion in the description of relative clauses in *The Comprehensive Grammar* (Quirk et al. 1985). It receives a double classification. In different chapters, the term 'relative clause' is described once as a function, then a specific form. As one of four central 'types' of English clause combination, the grammar defines relative clauses as a functional category of subordination as in the *Collins Cobuild* (Sinclair, 1991), fulfilled by constructions ranging from participles, infinitivals to finite constructions beginning with 'that/which'. Yet, a later description of post-modification identifies relative clauses as one of the four syntactic forms of post-modifying finites. Both cannot be true. A similar confusion in this grammar emerges in the description of a minor clause type called a 'comment clause', which is perhaps more discourse marker than clause, as in 'I tried to call you, *you know*'. The *Comprehensive Grammar* (Quirk et al., 1985) describes comment clauses as either to-infinitival, -ing, nominal relatives or adverbial clauses. The confusion here is that adverbial clauses were considered a function in the grammar, while the other clauses were described as forms. And, since all these forms can fulfill the adverbial function, either the entire class of adverbials can be comment clauses, making superfluous the mention a few forms separately, or what is meant is that comment clauses are adverbials but restricted to a few forms.

The point is not that there are not excellent descriptions of relative clauses and comment clauses in the grammar. *The Comprehensive Grammar* (Quirk et al. 1985) is one of the most valuable grammars written. However, perhaps those grammars which prioritize function, despite providing highly valuable descriptions of the English language, can be prone to a level of confusion between the analysis of structural form and syntactic function of combined clauses. The result is ambiguity as to what the term 'type' of English clause actually encompasses. It is probably for this reason that form is the most common approach in pedagogical grammars, more or less matching one form to one clause label so as allow learners to identify clause types by sight.

## 2.2 Form-first Descriptions of English Clause Combination

The function prioritizing *Cobuild Grammar* (Sinclair et al, 1991) was the first corpus based grammar; however, a form prioritizing corpus-based counterpart exists- the *Longman Grammar of Spoken and Written English* (Biber et al, 1999). This grammar describes eleven clause types based on form, the first eight finite forms and the final three non-finite: 1. Nominal clauses, 2. Adverbial clauses, 3. Relative clauses, 4. Comparative clauses, 5. Reporting Clauses, 6. Comment clause, 7. Question tags, 8. Declarative tags. 9. Infinitive clauses 10. -ing clauses, 11. -ed clauses. As form constitutes a clause type, a relative clause cannot be fulfilled by a range of constructions. The term 'relative clause' is a specific clause form that has a fully finite verbal element, is prototypically introduced by the anaphoric pronoun 'that/which', and is positioned in the combined clause structure post noun phrase. It is also either integrated into the NP constituent (restrictive) or syntactically parenthetical (non-restrictive). Participles or to-infinitival forms can post-modify an NP, but they are not relative clauses. They can simply fulfill the same post-modifying function. One might term them 'modifying clauses', but this is, like the term 'complement clause' often used for clausal arguments, a convenient shorthand term, not really a specific English clause.

Many of the forms of clause combination in this grammar were described as forms in the previous grammars reviewed, and defined more or less the same as in examples already given. For example, the description of the comparative clause matches that in *The Comprehensive Grammar* (Quirk et al, 1985), and the reporting clause in the *Longman Grammar* (Biber et al., 1999) is essentially the same as in the *Cobuild Grammar* (Sinclair et al., 1991) though restricted to the 'that' finite clause form. However, there are some combined clauses unique to the



Longman Grammar (Biber et al, 1999). These are question tags, as in ‘you finished the work, *didn’t you?*’, and declarative tags, ‘it seems bad, *it does*’. Adverbial clauses are given a formal definition, described as prototypically optional and marked by a subordinator, e.g. ‘I was hopeful *when I finished the interview*’. Adverbials have three functional subtypes: circumstance (I did all the homework, *except I didn’t do it very well*); stance (*if I had known*, I would never have gone); and linking (the essay needs revising, *that is section one at least*).

Finally, The Cambridge Grammar of the English Language (Huddleston et al, 2002) also starts with form and uses the notion of prototypical function in its description. Primary classification is grammatical form, with three syntactic functions recognized that the forms can fulfill: complement, modifier and adjunct. Clauses may prototypically fulfill one these functions, such as relative clauses tend to function as modifiers, but the Cambridge Grammar (Huddleston et al, 2002) is at pains to keep the notion of form and function as distinct features of English grammar. It identifies eleven combined clause forms; the first five finite, the second five non-finite, and the last without a verb: 1. Symmetric coordination, 2. Asymmetric coordination, 3. Content clause, 4. Comparative clause, 5. Relative clause, 6. To-infinitival, 7. Bare-infinitival, 8. Past participle, 9. Present participle, 10. Catenative clause, 11. Verbless.

There are few striking differences about this description. It identifies two coordinate clause forms: symmetric coordination, which has coordinated clausal constituents that are interchangeable, as in ‘she was watching TV *and I was reading*’; and asymmetric coordination with fixed constituents that if switched would deteriorate grammaticality, as in ‘he finished the book *and he closed it*’. The Cambridge Grammar (Huddleston et al., 2002) also does not recognize an ‘adverbial’ clause, either as a form or function. Those termed ‘adverbial clauses’ by other grammars are accounted for by other clause types. Functional definitions of adverbials are simply categorized as adjuncts, and formal definitions are considered finite content clauses complementing prepositions within an adjunct. Thus the ‘adverbial clause’ is not a distinctive clause in English grammar. Finally, a combined clause labeled the catenative clause is described as a form marked typically by an infinitive complementing the main clause verb, for example ‘she didn’t want *to have to do it again*’.

### 2.3. Hierarchy and English Clause Combination

The notion of a hierarchical cline representing the different levels of the grammatical integration in of combined clauses combination has previously been proposed by Hopper and Traugott (2002), Payne (2011) and Mathessien (2003). These studies have proposed ‘tighter’ and ‘looser’ clauses, with ‘tight’ meaning a clausal constituent has comparatively more integration into the clause with which it combines. Integration can be signalled by a range of features: fewer grammaticalized makers associated with the combination (e.g. no conjunctions or relative pronouns to mark subordination in ‘tighter’ forms); reduced verbal properties (e.g. non-finiteness); more association with the main clause argument (e.g. subject/object position); and more associated with grammatical than discourse structures (e.g. coordinate clauses more easily separate into independent sentences so are more discourse-like).

Payne (2011), using such considerations, presented a cline of combined clause integration, reproduced in Figure 1. Along this hierarchy he placed a description of English clause combination ranging from highly grammatically integrated structures such as compound verbs to those not at all grammatically integrated, constituting separate clauses.

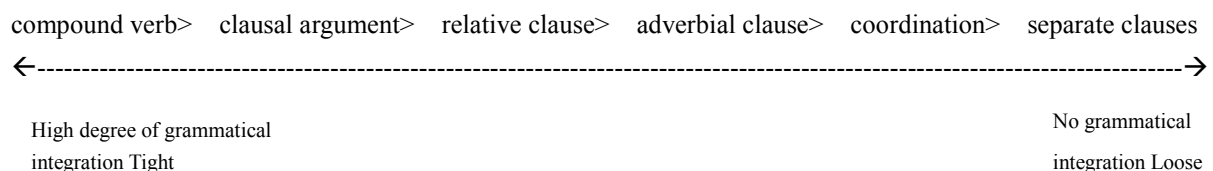


Figure 1: Hierarchy of Clause Combination (Payne, 2011)

The hierarchy in Payne (2011) derives from the following analysis. Compounds are considered the ‘tightest’ clause combination as two constituents verbal elements have been integrated; clausal arguments are essential to grammaticality so therefore second tightest; relative clauses are modifiers with an NP head that is typically an essential grammatical constituent, so third tightest; adverbial clauses are not essential to grammaticality, so less integrated; coordinate clauses are not subordinated to another constituent so even looser; and after coordination comes separate sentences.



Mathessien (2003) presented a slightly different hierarchy, as shown in Figure 2. This extends from syntactic clause combination to cohesion and coherence at the discourse level. For Mathessien (2003) the same underlying phenomenon is at work in grammatical relations as in connected discourse.

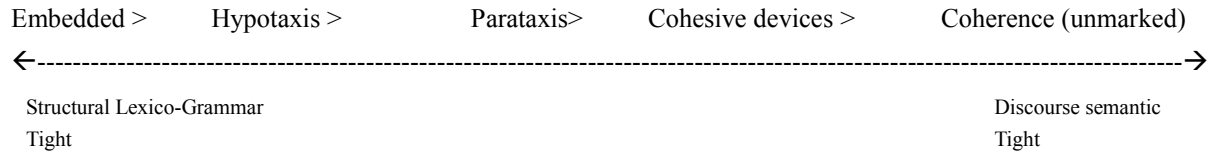


Figure 2: Hierarchy of Clause Combination (Mathessien, 2003)

Mathessien (2003) considered inter-clausal combination to range from tight syntactic embedding (e.g. infinitival clauses as complement to a main verb), to the looser relation of hypotaxis (e.g. a finite adverbial), to parataxis (e.g. coordination). Beyond these, discourse resources of English create relationships between clauses through cohesive devices, themselves a more explicit (ipso facto: tighter) form of coherence, which may be unmarked and based on pragmatic inference.

Hopper and Traugott (2002) also presented a hierarchy of clause combination, though somewhat lacking in detail as shown in Figure 3.

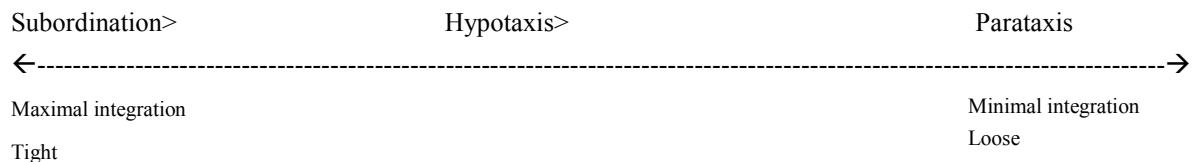


Figure 3: Hierarchy of clause combination (Hopper & Traugott, 2002):

For Hopper and Traugott (2002) and Mathessien (2003), parataxis is the syntactic independence of clauses which may (or may not) have pragmatic dependence. Hypotaxis is a more integrated clause, syntactically dependant within another clause's predicate but optional. Tighter still is subordination/embedding- clauses functioning as a constituent essential to grammaticality, e.g. verb arguments. Hopper and Traugott (2002) proposed their cline has an inverse relationship between syntactic tightness and overt linking devices, so the tighter the integration, the fewer markers signalling relations between the combined constituents. For example, tight clauses such as participle clauses do not have specific combining markers like relative pronouns or conjunctions found in the looser clauses. Similarly, tighter clauses lack tense/aspect marking. Brill (2011) indicated tense/aspect reduction suggest clauses depend for temporal continuity on the clause with which they integrate to contextualize the tense/aspect of the entire unit. Similarly, tighter clauses lack a grammatical subject position suggesting overarching continuity of the subject of the main clause<sup>1</sup>. A quantitative hypothesis derives from these considerations: first, the more integrated a clause, the more tense-aspect continuity it might tend to display between constituents. This would reflect why tense/aspect is unmarked in the tightest clauses. Second, the tighter a clause combination, the more often the same subject will be shared across constituents. This would reflect why the tightest forms lack a grammatical subject position.

The terms vary and are somewhat confounded, but all the proposals of a hierarchy reviewed acknowledge that clauses functioning as arguments/complements are more integrated than a modifying or adjunct/adverbial/hypotactic clause. This suggests that syntactic function has clear degrees of integration. The current study therefore proposes the following hierarchy of syntactic function in Figure 4, based on the three syntactic functions of Huddleston et al (2002):

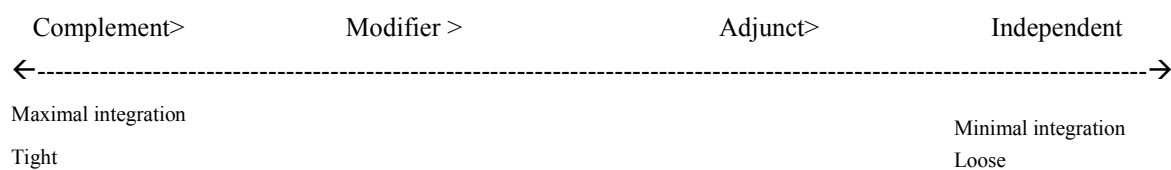


Figure 4: Hierarchy of syntactic function





### 3. Scope of the Study

The current study developed a hierarchy of English combined clauses through a synthesis and extension of the central grammars of English and an application of the notion of differing levels of integration. It chose form as the primary category to avoid form/function confusion. Once the hierarchy was established through descriptive grammatical analysis, a corpus of combined clause forms was analyzed to determine whether hierarchy in form had consequences for other features. Specifically, it was of interest as to whether those clauses which were more grammatically integrated patterned with the tighter syntactic functions, had quantitatively more inter-clausal tense-aspect continuity and more often shared the same subject across constituents.

### 4. Research Questions

Two research questions were addressed:

1. What is the hierarchy that exists amongst the combined clauses of English?
2. Does this hierarchy pattern quantitatively with particular syntactic functions, tense-aspect continuity and subject continuity in a way that reflects the hierarchy?

### 5. Method

#### 5.1. Data and Analysis

The research design consisted of a descriptive and a quantitative component. Research question one was addressed through a description of English combined clauses that synthesized and disambiguated those of the major grammars of English. Following the grammars of Huddleston et al. (2002) and Biber et al (1999), description was based on form rather than function. The combined clauses of English were distributed along a hierarchy from tight to loose based on an analysis of grammatical properties such as: the form of the verbal element, the presence/absence of grammaticalized markers, their prototypical argument roles, and the behaviour of the constituents with respect to movement or deletion. The description of the combined clause forms of English according to their hierarchical distribution, and the supporting analysis, is reported in section 6.1.

Research question two was addressed quantitatively. It built on the results of research question one by combining this description with a corpus analysis. To determine patterns of syntactic functions, tense-aspect continuity and subject continuity, a data set of 450 combined clauses was created. This represented 50 examples of each form along the established hierarchy. The combined clause data was taken from the Australian Corpus of English (ACE). The ACE is a corpus containing 17 genres and 500 samples of running text. It is modelled in design and balance after the Brown and LOB corpora sets.

#### 5.2 Procedure

The first 50 acceptable occurrences of each combined clause were taken from ACE sub-section A and K. Only those considered a canonical 'citation' form (i.e. typical forms one would expect in a descriptive grammar) were extracted. In the ACE data as in all discourse there exist numerous non-canonical forms, such as clauses with multiple embedding, subject dropping, combined clauses broken by long interpolations etc. The decision was made to pass over these clauses since the purpose of this study was not to study patterns of discourse but patterns in clear examples of specific grammatical forms.

The following principles guided extraction of clauses from the ACE. Clauses were taken from their surrounding discourse with any non-clausal preposed elements, but right-branching adjuncts were removed after 20 words. The principles are demonstrated in the following extracted coordinate clause:

#### Example 1: Extraction of combined clauses

*ACE 'Belles Letters' Corpus: <sample> G12*

#### Original sentence:

To kill time, Macquarie furl'd his past and corresponded with a range of individuals, some of them the darlings of the gods, others the play things of their sport.

#### Extracted combined clause:

To kill time, Macquarie furl'd his past and corresponded with a range of individuals

If, however, the initial clause of a combined form was itself in some larger sentence, perhaps subordinated in a larger stretch of discourse, then the subject head constituted the extraction point:

**Example 2: Extraction of combined clauses**

ACE 'Belles Letters' Corpus: <sample> G12

**Original sentence:**

In shaky health and out of sorts, Macquarie harboured grudge and grievance: that he may have damaged his standing in the eyes of the Commander-in-Chief harped on his mind, as did his unexpected and peremptory assignment to the East.

**Extracted combined clause:**

That he may have damaged his standing in the eyes of the Commander-in-Chief harped on his mind

Once 50 clauses for each clause type were extracted, all 450 (9x 50) combined clauses in the corpus was entered into SPSS v. 17, coded for three features: 1. Syntactic function (complement (C), modifier (M), or adjunct (A)); 2: Tense-aspect continuity (+TA) or discontinuity (-TA); 3: Same subject (EQ) or different subjects (SR) between the clause constituents in the combined form. Frequency statistics and proportions were calculated using SPSS to determine how these features quantitatively patterned according to combined clause type.

**6. Results**

*6.1. What is the Syntactic Hierarchy of Clause Combination in English?*

A description of English combined clauses emerged in which nine grammatical forms were identified as the central classes in English grammar. They are identified and distributed along a hierarchy of grammatical integration in figure 5.

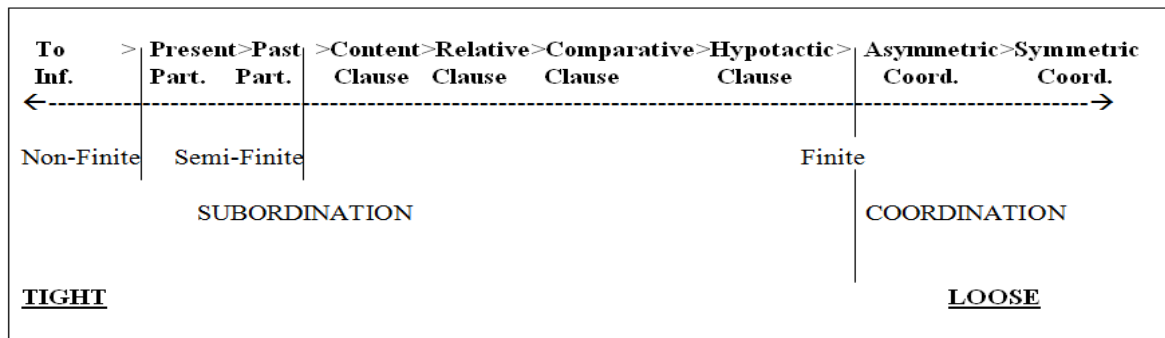


Figure 5: The Hierarchy of Combined Clauses in English

The result summarized in Figure 5 derived from the following analysis. English has three categories based on the finiteness of the verb: non-finite clauses were considered tightest as they had no independent tense-aspect properties in their form, while finite clauses were considered loosest being completely inflectionally independent, and an intermediate category was identified, following Payne (2011), of semi-finite in which clauses retained limited aspectual distinctions. From this three way tight>loose categorization, the infinitival clause was analysed as the tightest form, defined as an embedded non-finite clause with an infinitive verbal element that may or may not have a 'to' onset. This form was considered tightest due to being the only completely non-finite clause in English, and because the same form often merged into the verbal element of another to create the catenative construction. However, while the catenative construction speaks to the level of integration of the to-infinitival clause in the grammar, once a to-infinitival is merged with the verbal element to become catenative, it is perhaps less of a combined clause, contrary to Huddleston et al (2002), and more of combined verbal element nearer a serial verb or compound verb, similar to Payne (2011).

The present participle and past participle clauses were categorized as less integrated due to being semi-finite and manifesting aspectual distinctions. These aspectual distinctions indicate more independence from a main clause than a non-finite. Specifically, present participle clauses denoted simultaneity of the subordinate clause with the main clause, while a past participle denoted perfective completion (Givon, 2001). Hierarchically, the present participle was therefore considered a tighter form than the past participle by marking simultaneity in its verbal element with the main clause, whereas past participle combinations marked different temporal states between the constituents (Givon, 2001; Haiman, 1986).



Content clauses and relative clauses were identified as fully finite forms, both prototypically marked an onset 'that'. Superficially similar in form, they were disambiguated primarily by the fact that content clauses needed no grammatical antecedent, 'that' was not pronominal, and the content clause was embedded in a combined clause as an argument, hence its nomenclature as providing the 'content' of an argument typically in subject or object position. Relative clauses, however, required an antecedent noun phrase, 'that' was pronominal, and the clause could be removed from a sentence without making it ungrammatical. For these reasons, the relative clause was therefore a looser, less integrated, form than the content clause.

The comparative clause was defined as a fully finite clause marked by 'as...as' or an adjective inflection 'more/-er' correlated with a 'than' preposition. The form was difficult to place hierarchically. It did not function as an argument or core constituent like the content or relative, so seemed looser, yet sometimes a comparative clause could not be removed from its combination without weakening grammaticality, suggesting more formal integration than a relative. It was decided that it was tighter than optional hypotactic clauses (discussed next) as comparatives were sometimes essential, yet not as tight as content clauses, which were always essential. Finally they were positioned slightly less integrated than relatives which modified noun phrases in argument roles, while comparatives often completed adjectival subject complements, as in 'it was harder *than I thought*'.

The hypotactic clause was defined as a finite form, prototypically marked by a subordinator or adverb, and set apart from the main by a comma in writing or intonation unit in speech. It was usually optional, but always modified the verbal element of the main clause with adverbial semantics. The form, following the literature on 'adverbial complements', was sometimes essential to grammaticality in restricted contexts (Biber et al, 1999). The hypotactic clause had a wide range of onsets to mark it as a subordinate clause in relation to the main clause, subordinators and adverbs including: *after, then, for, when, as, because, in, so, therefore, however*. The clause type was labelled as hypotactic though often labelled an adverbial clause to distinguish it as a form not a function, avoiding the confusion covered in the literature review. The term adjunct was used for the functional equivalent of the adverbial clause.

Coordination was separated into two classes- symmetric for forms that had constituents able to be reversed while maintaining grammaticality, and asymmetric for forms that were fixed and ungrammatical if reversed. Coordinate forms were prototypically grammatically marked by a small set of coordinating conjunctions mainly, *and, but, so*. These could be elided for paratactic juxtaposition of clauses, termed appositive clauses in some grammars, but were not considered a unique clause here. Apposition seemed mostly a stylistic feature rather than a grammatical feature of English, and all appositive clauses responded to a symmetric or asymmetric description. Because asymmetric coordination exhibited grammatical restrictions, it was considered tighter than symmetric coordination. A coordination-subordination binary was kept for broad classification, however, considered 'fuzzy' categories given that asymmetric coordination indicated coordinated clauses can exhibit dependency between constituents.

The result of the analysis, as summarized in figure 5, indicated clear hierarchy amongst the combined clause forms of English, ranging from grammatical forms tightly integrated into each other to clauses with loose, more discourse-like, relationships. Interestingly, none were 'equally' integrated, as all had specific features that seemed legitimately to disambiguate their hierarchical position.

#### 6.2. What are the patterns of the hierarchy according to syntactic functions, tense-aspect and subject continuity?

As there were nine clause forms in the syntactic hierarchy, this resulted in 450 clauses being taken from the ACE (50 x 9) according to the procedures outlined in the methodology. The following presents the patterns of *syntactic function, tense-aspect and subject continuity* in the corpus data. Table 1 reports the frequency of syntactic functions according to clause type.

Starting with the highest frequency for a function, the results indicated coordinate clauses were categorically adjuncts, as they are always grammatically optional and never modify another head. The second highest frequency and proportion of adjuncts, 44 (88%), was found in hypotactic clauses, though these sometimes, 6 (12%), functioned as complements. The three loosest clauses therefore associated with the loosest syntactic function of adjunct. The next highest frequency occurred in comparative clauses, with 43 (86%) being complements. A few, 7 (14%), functioned as adjuncts. The fact they were never modifiers is an artefact. Comparatives cannot appear as modifiers and maintain their grammatical form, just as relative clauses cannot maintain their form and be complements since they must have an NP head and this head takes the complement role. A large frequency gap existed between comparative's adjunct frequency and the only slightly looser hypotactic clause.





Table 1. Clause Type and Syntactic Function

Clause Type	Syntactic function			Total
	Adjunct	Modifier	Complement	
to-infinitival	11	8	31	50
present participle	30	2	18	50
past participle	22	26	2	50
content clause	0	9	41	50
relative clause	13	37	0	50
comparative clause	7	0	43	50
hypotactic clause	44	0	6	50
asymmetric coordination	50	0	0	50
symmetric coordination	50	0	0	50
<b>Total</b>	<b>227</b>	<b>82</b>	<b>141</b>	<b>450</b>

Content clauses most commonly, 41 (82%), functioned as complements, fewer as NP modifiers, 9 (18%), in construction like ‘she made a complaint *that I’d spent her money*’. It is possible for content clauses to function as adjuncts but this did not occur in the data. The pattern for content clauses to associate with complements reflects why many grammars refer to it as a ‘complement’ clause. However, the current research believes this confounds form and function, and misleadingly implies the clause is categorically restricted to being a complement.

Past participles were distributed fairly evenly between adjuncts, 22 (44%), and modifiers, 26 (52%), rarely functioning as complements, 2 (4%). One might have expected an overwhelming majority to be modifiers given previous analyses, as in the *Collins Cobuild* (Sinclair et al., 1991) which somewhat equates them to relative clauses with obligatory ellipsis: ‘the man *who was seen yesterday*’ becomes ‘the man *seen yesterday*’. The different distribution patterns, however, support a distinction between these clauses.

In contrast to their past participle counterparts, present participles clauses were regularly complements, 18 (36%), but rarely modifiers, 2 (4%). They were most often adjuncts, 30 (60%). The infrequency of functioning as modifier is interesting as this is a regular textbook presentation of participle clauses (Swan, 2005). Rather than NP modification, for example ‘I sat next to a guy *wearing the baseball cap*’, the form seems more to appear as adjunct, as in ‘I had such a busy day, *running around all over town*’, or complement, ‘Can you stop *talking to the student next to you!*’.

Relative clauses, as expected, were overwhelmingly modifiers, 37 (74%). However the fact that 13 (26%) functioned as adjuncts reflects a significant number of non-restrictive and sentential relativisation in English. As mentioned, relative clauses could not function as a complement and retain grammatical form. Finally, to-infinitival clauses, the tightest form, were most often complements, 31 (62%), which is the tightest function. They also functioned as modifiers, 8 (16%), and adjuncts, 11 (22%), indicating they were the most evenly patterned of the clause forms across the syntactic functions. Therefore while there is a relationship between form and function in English strong enough to restrict function for some forms- e.g. relative clauses excluded as complements, comparatives excluded as modifiers, coordinated clauses restricted to adjuncts- form and function are nonetheless distinct properties of the grammar of English clause combination.

Frequencies of function presented a complex picture, but it would not be unfounded to suggest they reflected a kind of ‘natural syntax’ (Haiman, 1986) of English combined clauses in which tighter forms associate with tighter functions, and looser forms with looser functions. Table 1 however certainly does not present a lockstep pattern where the hierarchy of form matches quantitatively the hierarchy of function. After all, anomalies to such an interpretation include that semi-finites were more often adjuncts than they were complements, while comparatively looser content and comparative clauses favoured complements.

Table 2 reports the number of clauses which contained tense/aspect continuity between constituents. Tense/aspect continuity was considered not applicable to the non-finite and semi-finite clauses, which only had complete inflection in the main clause.



Table 2. Clause Type and Tense-Aspect Continuity

		tense/aspect discontinuity	tense/aspect continuity	N
Clause Type	to-infinitival	-	-	50
	present participle	-	-	50
	past participle	-	-	50
	content clause	33	17	50
	relative clause	27	23	50
	comparative clause	15	35	50
	hypotactic clause	21	29	50
	asymmetric coordination	18	32	50
	symmetric coordination	16	34	50
<b>Total</b>		<b>130</b>	<b>170</b>	<b>450</b>

None of the distributions strikingly favoured discontinuity or continuity between the tense-aspect of verbal elements in a form. The hypothesis that the tighter forms, because they were more integrated, would favour continuity, so that if one constituent was, for example, present simple the other would likely be also, was not supported. This hypothesis derived from the consideration that the tightest forms in English were non-finite and semi-finite, suggesting they relied more on the tense-aspect of main clause to temporally contextualize the whole combined form. In other words, forms lacking full inflection could be seen as having more integration with the main clause's tense-aspect. Consequently, as finite clauses moved along the hierarchy towards tight semi/non-finites, they might have higher frequencies of continuity reflecting this increasing integration. In fact, the data somewhat reflected the opposite tendency. It seems the tighter a combined clause of English, the less it tends to have tense/aspect continuity between constituents. The tighter content clauses and relative clauses, for example, had discontinuous tense/aspect more frequently than looser clauses like asymmetric and symmetric coordination.

Based on these results, either one might reject that tense-aspect relates to hierarchy, or offer a revised hypothesis. A revised hypothesis is that the lack of full inflection in highly integrated forms reflects that as clauses become tighter in English, tense/aspect relations between constituents become less relevant. Consequently, constituents tend to be less synchronized with each other. It is not that the main clause's tense/aspect absorbs their subordinate's tense/aspect that result in tighter clauses lacking full inflection, as originally proposed, but rather the functional accomplishment of the clause becomes less linked to tense/aspect as integration increases. In this revised formulation, one might expect the result found in this study. As one moves up the hierarchy of English combined clauses, the frequency of tense-aspect continuity would tend to decrease, terminating in maximal discontinuity, i.e. the combined clauses which lack full inflection.

Table 3 reports the number of clauses that maintained the same subject across combined clausal constituents, and the number which contained different subjects. The hypothesis was that the tighter the combined clause, the more often shared subjects would prevail reflecting their more integrated, more singular, construction. This proposal built on Payne (2011), in which the hierarchy of combined clauses ultimately terminated in a single subject clause with combined verbal elements (see figure 1).

Results indicated evidence for the hypothesis that combined clauses further along the hierarchy more frequently maintained subject continuity. The general trend shown was that same subjects declined in frequency as combined clause forms became looser, and increased as they became tighter. Putting aside asymmetric clauses for a moment, to-infinitival clauses, the tightest form, had the highest frequency with 34 out of 50 clauses having the same subjects. The next highest frequency occurred in the next tightest clause, present participles with 29 out of 50, and the third highest frequency in third tightest, past participles with 26 out of 50. It would be wrong to claim participle clauses preferred subject continuity, since proportions in past participles for example were comparable with 52% same subjects to 48% different subjects. However, the data does show they, and the other tighter clauses, more frequently maintained subject continuity in comparison to the looser clauses. Content clauses, relative clauses,



comparative clauses, hypotactic clauses, and symmetric coordination all had higher frequencies of different subjects between their constituents.

Table 3. Clause Type and Subject Continuity

		Same subjects	Different Subjects	N
Clause Type	to-infinitival	34	16	50
	present participle	29	21	50
	past participle	26	24	50
	content clause	14	36	50
	relative clause	10	40	50
	comparative clause	17	33	50
	hypotactic clause	19	31	50
	asymmetric coordination	35	15	50
	symmetric coordination	15	35	50
Total		199	251	450

The lowest frequency of subject continuity, 10 out of 50 clauses, occurred at the middle of the combined clause hierarchy in relative clauses. An information packaging constraint may explain the result. ‘Heavy’, i.e. longer, constituents tend to appear later in a clause structure, so usually in the predicate, while subjects tend to be ‘reduced’, i.e. short (Dubois, 2003). Relative clauses are ‘heavy’ and so, because of this information packaging constraint, prefer late positions in a clause, falling within predicates more often than pre-verbal subject NPs. Relative clauses would typically modify predicate NPs, which are infrequently co-referential with subject NPs in the same combined clause. This would lead to relative clauses infrequently having the same subjects as their main clause. This is likely the operating principle in why subject continuity is not quantitatively associated with the relative clause form.

The exceptional result was asymmetric coordination. As shown in table 3, it heavily preferred the same subject across clauses, with 35 out of 50 clauses exhibiting continuity. This was only one more continuous subject than the to-infinitivals, and thus not particularly different from the tightest form, but it was in stark contrast to all other looser clause forms. Asymmetric frequencies were the exact inverse of the form’s counterpart, symmetric coordination, which itself followed the general trend of the hierarchy with lower same subject frequencies. An asymmetric coordinate clause was defined as a form which could not have its clause constituents switched without deteriorating grammaticality. Since the corpus analysis revealed asymmetric clauses pattern with subject continuity, it seems that subject continuity is not conducive to interchanging constituents, despite both constituents being about the same thing/subject. The result suggests the discourse function of asymmetric coordination is for the second clausal constituent to provide an essential expansion or elaboration of the first clause. The second constituent in asymmetric coordination perhaps performs a task perhaps more like a predicate or hypotactic subordinate than its counterpart constituent in symmetric coordination. That is, asymmetric coordinate can indicate consecutive events, consequence, temporal sequence and so on with respect to a singular subject, becoming nonsense when the constituents are reversed: ‘she had a sleeping pill *and she slept well*’ but not \*‘she slept well *and she had sleeping pill.*’

**7. Discussion**

The results of this study have shown that both a description of the grammar and quantitative corpus patterns illustrate evidence for hierarchy across the range of English combined clauses. All combined clauses were shown to fit along a hierarchy from loose to tight, based on how integrated into each other their constituents were. The most significant grammars in English linguistics have different approaches to the analysis of combined clauses, however this study argued a form based approach best allows one to perceive combined clause hierarchy clearly in the grammar, and avoids form/function confounds shown to be problematic in some previous analyses. A hierarchy of function also exists, from complement to modifier to adjunct, but these three functions need to be distinguished from the notion of clause type. The current synthesis and extension of descriptions of English clause combination has indicated that hierarchy is an important, underemphasized, aspect of English clauses. Indeed, it is



often overlooked, for while some research has explored the notion, none of the major grammars of English have a discussion of hierarchy amongst the clause types.

Hierarchy amongst the combined clause forms was revealed by the corpus component of this research to have consequences for other aspects of English grammar and discourse. The effects of a clause's position on the formal hierarchy affected its patterns of syntactic function, subject continuity, and tense-aspect continuity. These patterns reflected the grammatical level of integration of the clause. Tighter clauses often exhibited a quantitative preference for maintaining the same subject across constituents, which suggests a more singular discourse unit. They also more often fulfilled the complement function. This reflects tight integration as complements are prototypically essential to the grammaticality of the sentence they occur in, while other functions are not. Finally, in the tighter clause forms, there was a decrease in the continuity of tense and aspect across constituents. It might be that this quantitative decrease in continuity reflected the formal hierarchy in that the tightest clauses of English are maximally discontinuous, having grammaticalized this discontinuity through marking full inflection on only one constituent.

This discussion of the corpus findings should not be taken as strong claims, however. There were patterns in the corpus component of the study that do not easily fit into an interpretation in which all the different clause types have their formal level of integration perfectly reflected in an isomorphic level of integration across other features of English. For example, some tight forms were frequently loose functions, such as present participles. Also, some loose forms frequently had subject continuity, such as asymmetric coordination. How hierarchy in form affects other aspects of English grammar and discourse needs to be pursued by future research that uses a larger data set and explores a wider range of variables and analytical considerations. Future research may be able to show further, stronger evidence that this hierarchy has a range of consequences across many parts of the grammar and discourse of the language. Indeed, it may lead to a revision and clarification of the hierarchy itself, not only an understanding of its qualities and patterns.

#### 8. Conclusion

This study has synthesized and extended the descriptions of clause combination in the 'bestsellers' of English grammar: The Longman Grammar of Spoken and Written English (Biber et al, 1999), The Cambridge Grammar of the English Language (Huddleston et al, 2002), The Collins Cobuild English Grammar (Sinclair et al, 1991) and The Comprehensive Grammar of the English Language (Quirk et al, 1985). The research produced a description of nine forms of combined clauses central to English grammar, and provided an analysis that distributed them along a hierarchy of grammatical integration. The forms were disambiguated from function, of which three were identified- complement, modifier and adjunct- which had their own hierarchical organization. A quantitative study of the clause forms using data from the ACE corpus indicated that formal hierarchy had consequences for how combined clauses patterned with syntactic functions, subject continuity and tense aspect continuity. The patterns broadly reflected that the higher on the hierarchy a form is, i.e. the tighter the integration of the combined clause, then the more it will tend to be tightly integrated in other grammatical and discourse features. An understanding of hierarchy and the levels of integration in the combined clauses of the English language has significant value for researchers, teachers and students whose goal is to better understand English grammar.

#### References

- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan., E. (1999). *The Longman grammar of spoken and written English*. Cambridge: Cambridge University Press.
- Bowern, C. (2008). Syntactic change and syntactic borrowing in generative grammar. In G. Ferraresi & M. Goldbach (eds.) (2009), *Principles of syntactic reconstruction*. Amsterdam: John Benjamins.
- Bril, I. (Ed.) (2011) *Clause-linking and Clause-hierarchy: Syntax and pragmatics*. Amsterdam: John Benjamins.
- Chomsky, N. (2007). Bilingualistic explorations: design, development, evolution. *International Journal of Philosophical Studies*, 15 (1), 1-21. <http://dx.doi.org/10.1080/09672550601143078>
- Cristofaro, S. (2005) *Subordination*. Oxford Univ. Press <http://dx.doi.org/10.1093/acprof:oso/9780199282005.001.0001>
- Du Bois, J. (2003). Argument structure: grammar in use. In J. Du Bois, L. Kumph, & W. Ashby (Eds) (2003) *Preferred argument structure: grammar as architecture for function*. Amsterdam: John Benjamins.
- Givón, T. (1995). *Functionalism and grammar*. Philadelphia: John Benjamins.



- Givón, T. (2001). *Syntax. Volume I*. Amsterdam/Philadelphia: John Benjamins.
- Haiman, J. (1985). *Natural syntax: Iconicity and erosion*. Cambridge: Cambridge University Press.
- Hopper, P., & Traugott, E. (2002). *Grammaticalization*. Cambridge, UK: Cambridge University Press
- Huddleston, R., & Pullum, G. (2002). *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
- Matthiessen, C (2002). Combining clauses into clause complexes: a multi-faceted view. In Joan Bybee & Michael Noonan (eds.), *Complex sentences in grammar and discourse: essays in honor of Sandra A. Thompson*. Amsterdam: Benjamins.
- Payne, T. (2011). *Understanding English grammar: a linguistic introduction*. New York: Cambridge University Press.
- Peters, P. (1986) *Australian Corpus of English (ACE)*. Sydney: Macquarie University, Department of Linguistics.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. Harlow: Longman.
- Sinclair, J (Ed) (1991). *Collins COBUILD English Grammar*. London & Glasgow: Collins.
- Sjolie, D., (2006). Phrase and clause grammar tactics for the ESL/ELL writing classroom. *The English Journal*, 95 (5): 35-40. <http://dx.doi.org/10.2307/30046586>
- Swan, M. (2005) *Practical English usage, 3<sup>rd</sup> edition*. Oxford: Oxford University Press.

---

## **Appendix A: Clause Combination in the Central English Grammars**

---

### Appendix A.1

Huddleston, et al (2002) 'The Cambridge Grammar of the English Language'

---

There are 10 clause types that can combine:

#### **FORMS:**

Independent

1. Coordinate

Finite

2. Content clause

3. Comparative Clause

4. Relative Clause

Non-finite

5. to-infinitival

6. bare-infinitival

7. past participle

8. present participle

9. Catenative chains

No-TAM

10. verbless

Coordination has 2 types which can be either marked or unmarked:

1. Symmetric

2. Asymmetric





The function of adverbial/adjunct is not restricted to a form.

Content clauses have 4 sub-types:

1. Declarative
2. Closed interrogative (wh)
3. Open Interrogative (if, whether)
4. Exclamative

Relative clauses have 3 subcategories:

1. Wh relative
2. non-Wh (that or bare)
3. Fused relatives

Relative clause types are either syntactically:

1. integrated
2. supplementary

Comparative clauses are complements to prepositions 'than' and 'as/like'. They exhibit 'obligatory reduction'.

---

Appendix A.2

Sinclair et al. (1991) 'Collins Cobuild English Grammar'

---

There are 4 clause types that can combine:

### **FUNCTIONS**

1. Coordinate
2. Adverbial clauses
3. Relative clauses
4. Reported clauses (essentially that clauses, sometimes to-inf)

Subordinate clauses, **3** types, can be finite or non-finite.

Basic distinction between **2** combination methods: coordination and subordination/ coordinate conjunctions and subordinate conjunctions.

There are **8** kinds of adverbial clauses, marked by particular subordinating conjunctions:

1. Time
2. Condition
3. Purpose
4. Reason
5. Result
6. Concessive
7. Place
8. Manner



Relative clauses: there are **2** types

1. Adjectival
2. Nominal (wh- clauses, fused relatives)

Adjectival relatives can be:

1. Defining
2. Non-defining

Coordination marked by several coordinators- and, but, nor, or, yet, then, so. And also combinations- and yet, and then, and so.

---

Appendix A.3

Quirk et al. (1985) 'A Comprehensive Grammar of the English Language'

---

There are 5 clause types that can combine:

**FUNCTIONS:**

Independent

1. Coordinate
2. Nominal Clause
3. Adverbial Clause
4. Relative Clause
5. Comparative Clause

These functions have **3** structural types

1. finite
2. non-finite
3. verbless

**FORMS:**

Nominal Clauses- subcategorize into **9** forms:

Finite

1. That- clause (like content clauses)
2. Wh- interrogative
3. Yes-no/alternative interrogatives (whether/if...or)
4. Exclamatives
5. Nominal relatives (fused)

Non-finite

6. To-infinitive
7. -ing participle
8. Bare infinitival
9. Verbless



Adverbial Clauses- subcategorize into 4 syntactic functions:

1. Conjuncts (discourse continuer)
2. Subjuncts (viewpoint marking: as far as BOB's concerned,...)
3. Adjuncts (circumstantial, ... while I was on vacation)
4. Disjuncts (style, content, attitudinal comment on the matrix)

Comparative clauses- 2 types

1. Equivalence (as..as)
2. Non-equivalence (-er, enough...too, more than)

Relative Clauses-

2 syntactic function types of relative post-modification:

1. Restrictive
2. Non-restrictive

4 syntactic types of 'post-modifying finite clauses':

1. Relative
2. Appositive
3. Nominal Relative (fused)
4. Sentential (clause is antecedent)

---

Appendix A.4

Biber et al (1999) 'The Longman Grammar of Spoken and Written English'

---

There are **11** forms of dependent clauses (plus independent coordination):

**FORMS:**

Finite:

1. Nominal clauses
2. Adverbial clauses
3. Relative clauses
4. Comparative clauses
5. Reporting Clauses
6. Comment clause
7. Question tags
8. Declarative tags

Non-Finite:

9. Infinitive
10. -ing clauses
11. -ed clauses



There are two types of nominal clauses:

1. Wh- interrogatives
2. Nominal relative clauses

Adverbial clauses:

Usually optional. Subordinator present. 3 types (further subtypes semantically)

1. Circumstance
2. Stance
3. Linking

Relative clauses have two types:

1. Restrictive
2. Non-restrictive

Comparative clauses have two types

1. as...as.
2. than- clauses

Reporting clauses have two functions:

1. Direct speech
2. Direct thought

Comment clauses: begin with phrases such as 'you know. I think, it seems' etc.

---

### Notes

<sup>i</sup> While some clauses lack a grammatical subject they can have subjects. For example, in 'The exam revealed John *to be behind the rest of the class*', the subject of the non-finite to-infinitival clause is 'John' but it is not grammatically coded into that clause in the way that 'the exam' is the grammatical subject of the main clause.