

Chapter 2

Clause Combination in English

2.1 Introduction

As mentioned in the previous chapter, most language use in English or otherwise, does not consist of isolated propositions. Human communication is not a sequence of independent, unrelated clauses made up of simple subjects followed by simple predicates. This, of course, has long been noted, despite a certain bias remaining in linguistics for analyses to be focused on what happens internally within a clause rather than across clauses and larger spans of discourse (Halliday 1965; Morris and Hirst 1991). The pertinent point, however, is that since human communication is almost always a sequence of coherently connected propositions, clause combination is naturally a central feature of language. This is acknowledged in the Adaptive Approach to Grammar (Givón 2002), but even in schools of linguistics fundamentally opposed to the tenets of functional-cognitive linguistics, it has been argued that clausal recursion is the only unique feature of human communication and therefore the core component of an innate language faculty (Chomsky 2007). This itself is quite a strong claim, rejected by other linguists (Everett 2012), and is not really held to in the current book, but it does demonstrate the importance of clause combination in contemporary linguistics regardless of the school to which one belongs.¹

In applied linguistics, it is well established that the knowledge of the forms and functions of different combined clauses is one of the more complex aspects of understanding English grammar for ESL students (Sjolie 2006). English clause structure tends to become more complicated in subordination; for example, the SVO order of the language can switch to VSO in subordinate clauses such as in *I would have said something, had I been there*. Furthermore, in second language research, it has been found that the acquisition of subject relative clauses tends to precede other relativization strategies, likely because the structure is analogous to

¹A version of this chapter was first published in Green (2012).

an SVO main clause (O'Grady et al. 2003). Applied work directed at mapping the educational development of native speakers has similarly shown that the knowledge of the patterns of clause combination is important for developing communicative competence appropriate for different genres and registers (Myhill 2009). As Biber et al. (2004) demonstrates the patterns of clause grammar vary systematically across different written and spoken genres and registers, and even within specific sub-categories of a genre, so both native and non-native speaking students need to learn this.

Given that the clause system is such an important object of study, it must also be an important enterprise to describe the properties of clause combination in English grammar as completely and accurately as possible. Yet, many grammars of English, whether written for native speakers or second language learners, lack any special emphasis on the relationship between clauses in the system, and what makes them a unified system set of more similar or more dissimilar constructions. Further complicating matters is what was noted in the previous chapter, namely that linguists and grammarians do not share a widely accepted description of the English clause system. This chapter reviews in more depth some of these disagreements, which exist on things as fundamental as the number of clause types, what the clauses should be called, how they should be defined, and even the existence of such commonly accepted clause types such as the English adverbial clause.

The chapter aims to highlight some of the difficulties in describing the clause system that have prevented widespread acceptance of any single model of the English clause system. To do so, it will compare and contrast the clause systems presented in some of the most widely referenced grammars of the English language currently in use. These include: *The Longman Grammar of Spoken and Written English* (Biber et al. 1999), *The Cambridge Grammar of the English Language* (Huddleston and Pullum 2002), *The Collins Cobuild English Grammar* (Sinclair 1990) and *The Comprehensive Grammar of the English Language* (Quirk et al. 1985). These grammars can be considered some of the most important descriptions of clause combination in English language study in that they are the most used and most cited in English linguistics; they are the 'best sellers' and the weighty reference tomes in institutions and offices everywhere. Once these grammars have been described and compared, the chapter develops a synthesized model of the clause system and organizes this system as a hierarchy of constructions. The notion of a clause hierarchy is largely absent, or at least underspecified, in many of the major reference grammars of English, and so the chapter (and indeed the book in general) hopes to contribute to its acceptance as fundamental property of the clause system. The clause system is organized much more according to a hierarchy of different levels of clause integration than can be meaningfully captured by only the binary categories of subordination and coordination (Payne 2011; Givón 2001).

To begin, the above reference grammars can be divided into two contrasting, and difficult to reconcile, approaches to the English clause system. One approach to describing what a speaker knows about their language is to describe English clause grammar using grammatical *form* (i.e. what it looks like) as a starting point, while the other is to take grammatical *function* (i.e. what it does) to be the primary

indicator of a clause category. This variation in the starting point of the analysis is partly why a widely accepted model of the English clause system has eluded grammarians, linguists and students. Two of the major English grammars that start their description of English grammar through functional categorization are the corpus-based *Collins Cobuild English Grammar* (Sinclair 1990) and *The Comprehensive Grammar of the English Language* (Quirk et al. 1985). An overview of these function-first descriptions of English clause combination is now given.

2.2 Function-first Descriptions of the English Clause System

2.2.1 *The Collins Cobuild English Grammar*

The Collins Cobuild English Grammar (Sinclair 1990) identifies itself as a “Grammar of Functions”, which it defines as “the patterns of language and the things you can do with them” (p. iv). It is one of the first corpus-based grammars in which entirely authentic data is used, believing that there “is no justification for inventing examples” (p. vii). In its analysis of the clause system of English, it begins with the traditional distinction between two clause combination methods: coordination and subordination, with a corollary binary distinction between coordinate and subordinate conjunctions. Coordination is defined as the combination of independent clauses, while subordination is defined as a clause that is dependent on another. Coordinating conjunctions, unlike most other descriptive grammars, are a rather extensive set in the *Collins Cobuild* (Sinclair 1990). They include *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 in this model of grammar. For example, all the above clauses can be either finite or non-finite and retain their primary class category. An adverbial clause of purpose may be either ‘I read contemporary fiction *to relax after lectures*’ or ‘I often read *Venus in Furs so that I can relax after lectures*’. The Collins Cobuild does note that there are form-function correspondences, and that one function tends to be fulfilled by one form; for example, noun groups tend to be subjects and topics. However, with clause combination, this one-to-one correspondence does not seem to be maintained.

Let us go through the three types of subordinate clauses individually. One of the three subordinate clauses is identified as the ‘reported clause’, which can 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.

A second type of clause is the relative clause, which in the *Collins Cobuild* (Sinclair 1990) 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: 'I suspect that linguists, *who I'm sure mean well*, have many interpersonal problems'.

The final subordinate clause type in this grammar is the adverbial clause. Eight subtypes of adverbial clauses are described: (1) Time, (2) Condition, (3) Purpose, (4) Reason, (5) Result, (6) Concessive, (7) Place, (8) Manner. These are defined semantically, though the grammar notes that typically the adverbial clause is formally associated with a clause that begins after a comma and contains a subordinator or adverb. The subtypes of adverbial clauses are also typically associated with a set of connectors that signal their function in relation to the main clause, such as in 'students finish exams quickly, *when/after/before they eat*' (adverbial of time). Although this particular example is finite in its verbal element and has a grammatical subject position, given the general functional approach of the *Collins Cobuild* (Sinclair 1990) an adverbial clause can also be any non-finite form as long as it fits into the semantic classifications above. For example, 'students finish exams quickly *to get out of class early*' is an adverbial of purpose, despite being a non-finite infinitival form. It is also worth noting that since the types of adverbial clauses described by the *Collins Cobuild* are functional-semantic types, some may have exactly the same form, yet have a different classification. For example, 'the student studied hard *so he received a passing grade*' may be an adverbial clause of purpose in one context of use with the communicative intent being on the reason why he was studying hard, while 'the student studied hard *so he received a passing grade*' would be an adverbial clause of result if the communicative intent was on the outcome.

The clause grammar of English, as described in the *Collins Cobuild* and reviewed above, might be summarized in the following manner:

Collins Cobuild English Grammar (Sinclair 1990)

There are four central clause types. These FUNCTIONAL categories can be finite or non-finite:

1. Coordinate
2. Adverbial clauses
3. Relative clauses
4. Reported clauses

These four clause functions are divided into two combination methods: coordination and subordination. When finite (and in some cases of non-finite subordination), these clauses are marked by explicit coordinate and subordinate conjunctions:

1. Coordination is marked by several coordinators: and, but, nor, or, yet, then, so. There are also multi-word coordinators: and yet, and then, and so
2. Subordination is marked by several subordinators: when, while, that, which, so, because. There are also multi-word subordinators: so that, after which, in order to, whenever.

Subordinate clauses have three types, either finite or non-finite: Adverbial, Relative and Reported.

1. Adverbial: There are eight kinds:
 1. Time, 2. Condition, 3. Purpose, 4. Reason, 5. Result, 6. Concessive, 7. Place , 8. Manner
2. Relative clauses: There are two kinds, each with two subcategories:
 1. Adjectival:
 - 1.1. Defining
 - 1.2. Non-defining
 2. Nominal:
 - 2.1. (Fused) relatives
 - 2.2. *Wh*-clauses
3. Reported clauses: there are two kinds which are specified according to form and used prototypically as indirect speech or as predicative verbal complements:
 1. That clauses
 2. To-infinitival clauses

In summary, what we see in the *Collins Cobuild* (Sinclair 1990) is a model of English grammar that uses the functions of English clauses as the central organizing principle of the system. The grammar makes no claim that this is the organizing principle when cognitively processing and producing English clauses in discourse; however, there is an implication from the way they have organized their description of the clause system that in the native English speaker's mind, function is the higher order category in which a range of forms can exist. The benefit of such a description is that it explains rather neatly why a noun can be easily modified by a finite clause just as well as by a non-finite clause in, for example, 'the book *that was being read*' and 'the book *being read*'. These are interchangeable in the system because they are the same clause type, namely a relative clause.

2.2.2 A Comprehensive Grammar of the English Language

The other major grammar of English with a function-first approach is the seminal *A Comprehensive Grammar of the English Language* (Quirk et al. 1985). This

grammar also begins with a binary subordination and coordination distinction in the clause system (though see the next chapter for their discussion of gradience in the set of conjunctions). It makes a four-way functional classification of subordinate clauses, which are fulfilled by a range of different forms. The clause types, in terms of their higher order functional categories, are nominal clauses, adverbial clauses, relative clauses, and comparative clauses. Three main forms are associated with these four functional types: finite clauses, non-finite clauses, and verbless clauses. As with the *Collins Cobuild* (Sinclair 1990), it is useful to examine the clauses one by one in the following section and then offer an overall summary of their description of the clause system.

The first principal clause category in the *Comprehensive Grammar's* (Quirk et al. 1985) description of English grammar is the nominal clause, which is essentially any clause able to function as an argument of a verb, e.g. a clause that is itself a subject, an object or a copula predicate in a sentence such as '*That we do not know how a child acquires language* is a sad fact of linguistic science'. However, in the *Comprehensive Grammar*, nine subtypes of the nominal clause are described; quite a few more than in the *Collins Cobuild* above, which has only two. In the following list of these subtypes, the first five nominal clauses are finite, the subsequent three are non-finite and the last is 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*).

One can see that a key challenge in understanding the English clause system has already emerged. This is that different descriptions of English grammar use either the same label for different things, or different labels for the same thing (see Pullum 2009, for an extensive discussion of this very same problem in relation to determiners vs. determinatives). The label 'nominal clause', as we have seen, is used both in the *Collins Cobuild* (Sinclair 1990) and in the *Comprehensive Grammar* (Quirk et al. 1985). However, they are not referring to the same clause. In the *Collins Cobuild* (Sinclair 1990) the nominal clause is a subtype in the category of relative clauses. Furthermore, the definition of a nominal clause that they offer covers only senses 3, 4 and 5 of what the *Comprehensive Grammar* (Quirk et al. 1985) defines under the term nominal clause. The *Comprehensive Grammar's* (Quirk et al. 1985) definition additionally encompasses what the *Collins Cobuild* (Sinclair 1990) place in an entirely separate clause category, namely the reported clause and senses 1, 2, 6 and 7. Additionally, a verbless nominal clause is included in the *Comprehensive Grammar* (Quirk et al. 1985), as well as an *-ing* participle nominal clause. If one were simply glancing at a range of reference grammars to remind themselves of what 'reported clause' was, say for a lesson they were preparing on English grammar, it is easy to see how one's understanding of the English clause system could quickly become confused by these different models of

the clause system. Two very well-respected reference grammars use the same (meta)language for different analyses.

The second combined clause type in the *Comprehensive Grammar* (Quirk et al. 1985) is the adverbial clause, again a functional category. This clause is further subcategorized into four syntactic functions. The description is quite different from the description of adverbial clauses in the *Collins Cobuild* (Sinclair 1990). In the *Comprehensive Grammar* (Quirk et al. 1985), an adverbial clause can function as a conjunct for discourse continuation, as in ‘the book was good, *since it was short*’; subjuncts which are used for viewpoint elaboration, as in ‘the book was good, *as far as I am concerned*’; adjuncts which denote circumstance, as in ‘I finished the work, *while I was on my break*; or disjuncts for style, content and attitudinal comments on the main clause, as in ‘the theory is nonsense, *however you look at it*’. The adverbial clause subcategories are not restricted to particular finite or non-finite forms; however, it does seem that some restrictions are inherent in the model. For example, it seems likely that subjuncts are almost always finite clauses, but there may be counter examples that the reader can think of.

Comparative clauses are the third clause identified in the *Comprehensive Grammar* (Quirk et al. 1985) and have two subtypes: equivalence, marked by correlative *as...as*: ‘She is *as good a linguist as he is*’; and non-equivalence marked by the comparative *-er* morpheme and *than*: ‘She is a better linguist *than he is*’. This again contrasts with the *Collins Cobuild*, where although such patterns are discussed (Sinclair 1990, p. 99), they are never clearly identified as a central combined clause type of English. Thus, whether or not the *Collins Cobuild* (Sinclair 1995) intended to include a comparative clause in their model of the clause system similar to that described by the *Comprehensive Grammar* (Quirk et al. 1985) is an open question.

Relative clauses, the final of the four functional types of English clauses identified in the *Comprehensive Grammar* (Quirk et al. 1985), are subcategorized as either restrictive or non-restrictive. The definitions are the same as in the *Collins Cobuild*’s (Sinclair 1990) description of defining and non-defining relative clauses given above. Relative clauses, being a functional category of subordination as with the *Collins Cobuild* (Sinclair 1990), can be filled by constructions ranging from non-finite participles and infinitivals to finite constructions beginning with *that/which/who*. However, at this point, another challenge emerges in trying to pin down the English clause system. This is the blurred line between form and function in definitions, which becomes evident when relative clauses are considered in the *Comprehensive Grammar* (Quirk et al. 1985) as a whole. For example, in the later chapters of this reference grammar, there are descriptions of four grammatical forms of post-modifying finite clauses for noun phrases (Quirk et al. 1985, p. 1244). These are: (1) Appositive (the book, *which I’ve never read*, is said to be great), (2) Nominal relatives (see above examples), (3) Sentential (I knew it would be boring, *which is why I didn’t go*), and (4) Relative (It’s the job *that I want*). The problem is mostly with number 4, since what has happened is that term relative clause had been defined twice in the *Comprehensive Grammar* (Quirk et al. 1985) and in one chapter it is given a function definition and in another defined according to form. As one of the four

central types of English clause combination, a relative clause is a functional category according to Quirk et al. (1985, p. 1047) that can be a finite or non-finite clause that modifies a noun. Yet, as one of the four post-modifying clauses of nouns, it is defined specifically as the finite form (4) above.

This seems to happen elsewhere in the grammar also. In the description of a minor clause type called a ‘comment clause’ (which is often considered a discourse marker rather than a 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* clauses, nominal relatives or adverbial clauses. The problem here is that adverbial clauses were defined principally as a function in the grammar, while these other clauses were all described as forms. Since all these forms can fulfil the adverbial function, either the entire class of adverbials can be comment clauses, making superfluous the mention of a few forms separately, or what is meant must be that comment clauses are adverbials but restricted to a few forms within that class, namely the finite adverbials. It is not clear.

The point is not that the descriptions of relative clauses and comment clauses in the *Comprehensive Grammar* (Quirk et al. 1985) are invalid. This is one of the most detailed grammars of the language ever written, and rightly famous. The point is, however, that grammars face the challenge of disambiguating form and function, and when this is not consistent, it can damage the internal validity of the model of the clause system being developed (Aarts 2007). By allowing a wide range of forms into a class, if definitions are not kept straight, the function-first approach to the clause results in ambiguity as to what the categories are in the English clause system. It is probably (partly at least) for this reason a focus on form is the most common approach in pedagogical grammars; it allows students to more or less match one form to one clause label, so learners are able to identify the English clause types by sight.

Overall, the clause system of English, according to the *Comprehensive Grammar*, might be sketched out as follows:

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

There are five central clause types. These are FUNCTIONAL categories that can be finite or non-finite:

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

These functions have three structural types

1. finite
2. non-finite
3. verbless

1. Nominal Clauses—subcategorize into nine forms across the three structural types:

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

Verbless

2. Adverbial Clauses—subcategorize into four syntactic functions:

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

3. Comparative clauses—subcategorize into two types

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

4. Relative Clauses—subcategorize into two syntactic functions:

1. Restrictive
2. Non-restrictive

But, there are four syntactic forms of ‘post-modifying finite clauses’:

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

What we see in the *Comprehensive Grammar* (Quirk et al. 1985) is a similar approach to the representation of the English clause system to that found in the *Collins Cobuild* (Sinclair 1990). They place the forms of the clause system within higher order functional categories. Despite this shared orientation, the two grammars differ significantly in their model of the language as the summaries of their descriptions indicate. Let us now look at the other approach to the clause system which inverts the function-first approach, but have had just as significant an impact on the study of English grammar as the *Comprehensive Grammar* (Quirk et al. 1985) and the *Collins Cobuild* (Sinclair 1990).

2.3 Form-first Descriptions of the English Clause System

2.3.1 *The Longman Grammar of Spoken and Written English*

The function prioritizing *Collins Cobuild Grammar* (Sinclair 1990) was the first widely used grammar based entirely on corpus data. However, a form prioritizing corpus counterpart was later developed, known as the *Longman Grammar of Spoken and Written English* (Biber et al. 1999). This corpus based grammar goes a little further than the previous one in that it embeds in the description of the English clause system the statistical variation that occurs according to genre, register, spoken and written language. The focus here will be on summarizing the clause grammar model that is described; however, the reader is encouraged to refer to the grammar itself for the important discussion of frequency variation that affects the clause system.

This grammar describes thirteen clause types (Biber et al. 1999, pp. 192–210), which are defined according to their structural form. They are: (1) Nominal clauses, (2) Adverbial clauses, (3) Relative clauses, (4) Comparative clauses, (5) Reporting Clauses, (6) Comment clauses, (7) Question tags, (8) Declarative tags (9) Infinitive clauses (10) *-ing* clauses (11) *-ed* clauses, (12) Verbless (13) Supplementive clauses. The first eight of these are finite forms and the next four are non-finite forms, while the final Supplementive clause is problematic and one we shall come back to in a moment. As form constitutes the primary categorization of the English clause system in this model, a relative clause, for example, cannot be fulfilled by a range of different constructions. The term ‘relative clause’ in the *Longman Grammar* (Biber et al. 1999) is a specific clause structure that must have a fully finite verbal element, is prototypically introduced by an anaphoric pronoun such as *that/which/who*, and is prototypically positioned after 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 in the *Longman Grammar* (Biber et al. 1999) description of the clause system. They can simply fulfil the same post-modifying function.

Many of the thirteen forms of clause combination in this grammar were described in the grammars summarized in the previous sections, and defined more or less the same as in the examples given there. As such, we need not go through each of them again. There are a few differences that need to be mentioned, however. The description of the reporting clause in the *Longman Grammar* (Biber et al. 1999) is essentially the same as in the *Cobuild Grammar* (Sinclair 1990), though the *Longman Grammar* restricts its definition to the *that* finite clause form only. Further, there are some clauses unique to the *Longman Grammar* (Biber et al. 1999) when compared to the other grammars that have so far been described. These are question tags, as in ‘you finished the work, *didn’t you?*’, and declarative tags, ‘it seems bad, *it does*’.

One of the most systematic differences in the *Longman Grammar* (Biber et al. 1999) from the previous grammars is that the names of many clause types are the same but now refer to individual forms rather than refer to individual functions (and a range of forms). For example, adverbial clauses, a term we have encountered in the previous grammars, are given a formal definition which restricts them to a finite clause marked by a subordinator and (typically) are syntactically optional in relation to the main clause, as in ‘I was hopeful *when I finished the interview*’. Again in an inverse fashion to the function-first grammars, it is the forms that are sub-categorized for function. That is, adverbial clauses have three functional sub-types according to their semantics: 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*).

There is perhaps one minor difficulty in the *Longman Grammar* (Biber et al. 1999) that is reminiscent to the one we looked at in the *Comprehensive Grammar* (Quirk et al. 1985). This is the clause type given the label the Supplementive clause (Biber et al. 1999, p. 201). This clause is defined as any clause that has no clear grammatical integration into the main clause, essentially any adjunct or parenthetical clause such as ‘he wore the jacket, *looking like a fool*, but was never embarrassed’ or ‘*Though read rapidly and not closely*, she understood the book well’. The issue that one might point out here is that it is somewhat inconsistent with the approach of the *Longman Grammar* (Biber et al. 1999) to define all clause types as structural forms, since this Supplementive clause category is a functional class.

The clause grammar of English as described by the *Longman Grammar* might be summarized as follows:

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

1. There are thirteen forms of combined clauses, with a binary division between dependent clauses and coordination. These are FORMAL categories:

Finite:

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

Non-Finite:

10. Infinitive
11. -ing clauses
12. -ed clauses
13. Verbless

2. There are two types of nominal clauses:

1. *Wh*-interrogatives
2. Nominal relative clauses

3. There are three types of adverbial clauses (with further semantic subtypes consistent with all of those described earlier in the *Collins Cobuild*). The three types of adverbial clauses are functional subcategories that can have different forms

1. Circumstance
2. Stance
3. Linking

4. There are two types of Relative clauses:

1. Restrictive
2. Non-restrictive

5. There are two types of Comparative clauses, with specific forms:

1. *as...as*
2. *than* clauses

6. There are two types of Reporting clauses. These are functional subcategories that can have different forms:

1. Direct speech
2. Direct thought

7. Clauses can be Supplementative (a functional distinction)

What the *Longman Grammar* (Biber et al. 1999) therefore presents is a large reference grammar that like the *Collins Cobuild* (Sinclair 1990) shares the goal of describing the clause system using only corpus data and authentic examples. Yet, these two corpus grammars have inverted approaches to the English clause system. Even though they identify many of the same properties of the English clause system, the hierarchical organization is different and the definitions associated with clauses have largely been switched, with formal categories becoming functional categories and functional categories becoming formal categories.

2.3.2 *The Cambridge Grammar of the English Language*

The final English grammar that will be considered is the *The Cambridge Grammar of the English Language* (Huddleston and Pullum 2002). It, like the *Longman Grammar* (Biber et al. 1999), starts with form as the primary classifying principle of the English clause system, and subcategorizes the clause forms for their possible functions. According to this grammar, the set of clause forms can fulfil up to three syntactic functions, which are: complement, modifier and adjunct. A complement is

any clause that is grammatically necessary to a matrix clause, as in the argument completing clausal constituent unit such as ‘here is *what you asked for*’. A modifier is any clause embedded within another constituent that modifies its head and that would not lead to ungrammaticality if removed, as in ‘the meal *that I ordered* was splendid’. Finally, an adjunct is any other optional clause, as in ‘graduate school is a great way to quickly build debt *while you are young*’. These three functions are not properties of the combined clause system per se, but rather are the core functional categories of English grammar generally, and apply to the parts of speech and phrasal constituents also. Furthermore, although clauses may prototypically fulfil a particular function, such as relative clauses as modifiers, the *Cambridge Grammar* (Huddleston and Pullum 2002) very consistently keeps the notion of form and function distinct in its description of the English clause system and quite strongly proposes that that nothing but syntactic and morphological properties are relevant to defining a clause in English grammar.

The *Cambridge Grammar* (Huddleston and Pullum 2002) identifies eleven combined clause forms that can participate in the three functions: (1) Symmetric coordination, (2) Asymmetric coordination, (3) Content clause, (4) Comparative clause, (5) Relative clause, (6) To-infinitivals, (7) Bare infinitivals, (8) Past participle clauses, (9) Present participle clauses, (10) Catenatives, (11) Verbless clauses. These forms are further divided into three subcategories according to the finiteness of the verb: the first five are finite, the second five are non-finite, and the last is verbless. Many of the clause forms are defined the same as the forms described in the *Longman Grammar* (Biber et al. 1999). For example, the past and present participle clauses are identical to the *-ing* and *-ed* clauses in the earlier grammar. Other clauses are also similar, with some changes in names or the parameters of the definition. For example, the content clause is the *Cambridge Grammar*’s (Huddleston and Pullum 2002) name for the reporting clauses and comment clauses of the *Longman Grammar* (Biber et al. 1999). However, the definition is slightly extended to include clauses such as ‘the suggestion *that he become a linguist* was laughed at’. The *Cambridge Grammar* (Huddleston and Pullum 2002) demonstrates rather convincingly that this is the same clause form that appears in ‘the suggestion is *that he become a linguist*’, with an alteration of function between modifier and complement.

There are few striking differences about this description of clause combination in comparison to those previously discussed. One of the most notable is that the *Cambridge Grammar* (Huddleston and Pullum 2002) identifies two coordinate clause forms: symmetric coordination and asymmetric coordination. Symmetric coordination has coordinated clausal constituents that are interchangeable, as in ‘she was watching TV *and I was reading*’. Asymmetric coordination has fixed constituents, which if switched would make a sentence either less grammatical or ungrammatical, as in ‘he kissed me goodnight *and I left for home*’ or ‘he finished the book *and he closed it*’. While one might object that this distinction is largely semantic and discourse-pragmatic, the argument of the *Cambridge Grammar* seems to be that they respond differently to syntactic tests and are therefore different kinds of syntactic structures.

Perhaps one of the most notable features of the English clause system according to the *Cambridge Grammar* (Huddleston and Pullum 2002) is that it does not recognize an ‘adverbial’ clause in the language, neither as a form nor as a function. This is quite remarkable given the ubiquity of the adverbial clause category in the study of English. Those clauses termed adverbial clauses in the other grammars, and taught in school, are accounted for in the *Cambridge Grammar* (Huddleston and Pullum 2002) by the strict division between form and function and the reanalysis of all adverbial clause subordinators as prepositions. The grammar first takes the step of arguing that words such as ‘because’ or ‘since’ are always a preposition even when followed by a clause. Therefore, there is no longer an extensive class of subordinators in the English language and as a consequence, the authors argue that they see no principled form-based category distinction between clauses that complement prepositions and the content clause category. Thus, the entire class of adverbial clauses, as extensively described in the *Longman Grammar* (Biber et al. 1999) are merged into the content clause category and treated as content clauses with the function of complement to a preposition and adjunct in relation to their matrix clause. In sum, there is no adverbial clause in the English clause system according to the *Cambridge Grammar* (Huddleston and Pullum 2002).

Another clause type unique to this grammar is the catenative clause. This is described as a form marked typically by an infinitive functioning as complement to the matrix clause verb, for example ‘she didn’t want *to have to do it again*’. It is not so much that catenatives themselves are unique to the *Cambridge Grammar* (Huddleston and Pullum 2002) as they are often described as part of a verb sequence. What is unique is that they are treated as examples of clause combination. Indeed, the *Cambridge Grammar* (Huddleston and Pullum 2002) classifies most verb combinations as bi-clausal combinations, even auxiliary and main verb combinations. What this means in practice is that structures traditionally thought of as a single verbal element and thus a single clause, such as the present perfect ‘the bear *has eaten* the honey’ are reanalysed as one finite clause followed by a separate non-finite complement clause ‘the bear has *eaten the honey*’.

The clause system of English, according to the *Cambridge Grammar*, is as follows:

The Cambridge Grammar of the English Language (Huddleston and Pullum 2002)

There are ten clause types that can combine. These are formal classes:

1. Coordinate
2. Content clause
3. Comparative Clause
4. Relative Clause
5. to-infinitival
6. bare infinitival
7. past participle
8. present participle

9. Catenative chains

No-TAM marker

10. verbless

Coordinate clauses have two types:

1. Symmetric
2. Asymmetric

Content clauses have four types:

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

Relative clauses have three types:

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

Relative clauses are either syntactically:

1. integrated
2. supplementary

Comparative clauses have no subtypes but so have different connectors:

1. than
2. as
3. like

What the preceding section of this chapter has shown in general is that all the major reference grammars English, despite the wealth of detail and the expertise of the linguists involved, do not share an agreed model of the English clause system. One does not want to overstate the problems, since certainly a good deal of common ground exists across the grammars, but there are definitely some problem areas. For example, some grammars classify the central clause types of the language according to their functions, while others use the form (clause syntax and morphology) as the primary organizing principle of the system. Further, the grammars do not agree on the number of clause types in English, the labels that should be attached to them, their definitions or which subtypes of clauses belong in which of the categories. It is the goal of the next section to offer a synthesis of the foregoing descriptions of grammar, in order to propose a model of the English clause system that captures how it is unified by an underlying, but very real, aspect of the system: namely, the clause hierarchy.

2.4 The Concept of Clause Hierarchy

What I would like to argue for the remainder of this chapter is that a very important feature of the English clause system has been largely overlooked in most books on English grammar, including the major reference grammars just discussed. What they do not explicitly articulate is that the clause system is organized as a clause hierarchy and that this is part of what a speaker knows about their language. The notion of hierarchy, or a cline along which clauses distribute according to their different levels of grammatical integration, should have a prominent place in the analysis of the clause system. Once it has been described, it will be demonstrated that this organizing principle can be detected in several patterns of the English language that support the view clause hierarchy has some cognitive representation in the system of English grammar.

Of course, clause hierarchy is not a new idea despite not being the focus of study in the major reference grammars of English just reviewed. Particularly in the functional-typological schools of linguistics, this notion of clause hierarchy has been elaborated on by Hopper and Traugott (2003), Payne (2011), Matthiessen (2002), Aarts (2007) and Givón (1979, 2001). These researchers have proposed that clause systems in many languages can be described as a set of ‘tighter’ or ‘looser’ clauses. A ‘tight’ clause means that a clausal constituent has, in comparison to a ‘loose’ clause, more dependence on the clause with which it combines, typically a main clause. Other models of hierarchy exist also, which distribute clauses along a cline of verbal to nominal propositions, another way of saying that some clauses are more noun-like things in the propositions they package, while others are more verb-like processes (Aarts 2007). While much of this work has been cross-linguistic, Givón (1993) has worked on the idea in relation to the English clause system, and the following chapters draw on his ideas while updating it with a synthesis of the major reference grammars of English outlined above, most of which have been published after his original work.

The clause hierarchy of a language can be determined by the analysis of the dependence and integration of each member of the clause system in relation to their patterns with superordinate matrix clauses, e.g. the main clause. Integration can be signalled by features such as fewer grammaticalized makers associated with the combination (e.g. no conjunctions or relative pronouns to mark subordination); reduced verbal properties (e.g. non-finiteness); ability to function as a main clause argument (e.g. subject/object role); and/or more or less similarity to structures that occur at discourse level (e.g. coordinate clauses are more similar to independent sentences than are infinitival clauses). For example, Payne (2011), using these principles as tests of hierarchical position, developed the cline of combined clause integration represented in Fig. 2.1. His hierarchy extends from discourse, defined as separate sentences, through to compound verbs where clause combination becomes more or less a single verbal element.

The hierarchy that Payne (2011) proposes derives from the following analysis. Compound verbs, though uncommon in English but found, he argues, in patterns

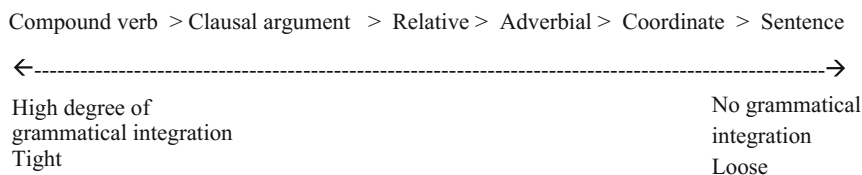


Fig. 2.1 Hierarchy of clause combination (Payne 2011, p. 329)

such as ‘please *go get* the book for me’ are considered the ‘tightest’ clause combination as they have two verbal elements placed adjacently in a verb phrase, one of which lacks full finiteness. Clausal arguments are essential to grammaticality, in that they function as arguments of the verb in the main clause, either as a subject clause, a complement clause or an object. Since they have more independent verb phrases and can have independent subject slots, they not as integrated as compound verbs. Relative clauses, Payne (2011) suggests, are prototypically modifiers with an NP head that is itself usually an essential grammatical constituent. Given that modifiers are not needed for the grammaticality of the main clause, they are third tightest. Adverbial clauses are not essential to grammaticality, so even less integrated; while coordinate clauses are not subordinated to another constituent at all so at the end of the cline, after which propositions have few syntactic relations and exist in separate sentences or intonation units with terminal contours. Payne’s (2011) hierarchy of clause combination sheds some light on how the clauses of English relate to each other as a unified set, taking the grammatical analyses discussed in the previous sections a little further. However, it does reflect some of the form/function problems discussed previously. For example, Payne’s (2011) hierarchy is sometimes using function and sometimes formal categories of clause combination in order to position the English clauses along his cline of integration.

Matthiessen (2002) presents a slightly different model of clause hierarchy, shown in Fig. 2.2. This extends from syntactic clause combination to cohesion and coherence at the discourse and text level. For Matthiessen (2002), the same underlying phenomenon is at work when a speaker combines propositions at the lexicosemantic level as when they combine propositions at discourse level (further discussion of the relationship between the grammar of clause combination and the signalling of discourse level relations is had in Chap. 6).

Matthiessen (2002) considers clause combination to range from tight syntactic embedding (e.g. infinitival clauses as complement to a main verb), to the looser

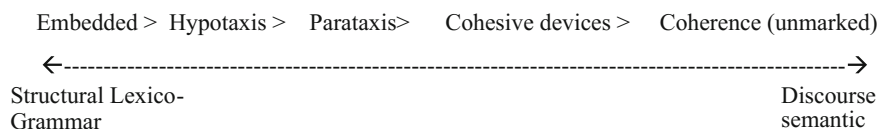


Fig. 2.2 Hierarchy of clause combination (Matthiessen 2002, p. 36)

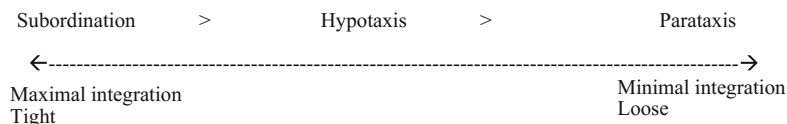


Fig. 2.3 Hierarchy of clause combination (Hopper and Traugott 2003)

relations of hypotaxis (e.g. a finite adverbial), to parataxis (e.g. coordination). Beyond these, the discourse resources of English create relationships between clauses via the use of cohesive devices, themselves a more explicit (ipso facto: tighter) form of general propositional coherence that may be achieved through inference and pragmatics. Matthiessen's (2002) hierarchy is a function based distribution of the English clause system's different levels of integration.

Hopper and Traugott (2002), in their seminal book on grammaticalization, also offer a hierarchy of clause combination, shown in Fig. 2.3. It is somewhat lacking in detail, as their aim is to capture typological generalizations across languages and they are not focussed on the English clause system.

This model shares much with Matthiessen (2002). For Hopper and Traugott (2003), parataxis is the syntactic independence of clauses which may have pragmatic dependence. Hypotaxis is a more integrated clause that is syntactically dependant within another clause's predicate but nevertheless optional in the sense that it can be removed without affecting the grammaticality of the main clause. Tighter still is subordination, which like Matthiessen's (2002) embedding, covers all clauses which function as a constituent essential to grammaticality, e.g. verbal arguments.

What is interesting about Hopper and Traugott (2003) is that they also propose 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 as found in the looser clauses. Similarly, tighter clauses lack tense/aspect marking. This tense/aspect reduction causes combined clauses to depend for temporal continuity on the clause with which they integrate (on this point, see also Bril 2010). Tighter clauses also lack a grammatical subject position suggesting overarching continuity of the participants in the main clause. Rather, grammatical subjects may be substituted for by logical subjects, as in 'The exam revealed John *to be behind the rest of the class*', which has the subject of the non-finite infinitival clause logically co-referent with the object of the main clause, but the relationship is not one of grammatical concord where number and person are reflected in the verbal element of the subordinate clause.

2.4.1 A Hierarchy for the English Clause System

Drawing upon the models of hierarchy previously discussed and the clause systems described in the major reference grammars, this section offers a synthesis that distributes the core categories of the English clause system along a hierarchy of integration. The focus here will be on describing the common core of the system (Biber et al. 1999); that is, the principal categories and their relationship to one another. Analysis of the numerous sub-types that exist within each category will largely be left until the next chapter. To begin, it seems wise to follow Huddleston and Pullum (2002) and separate form and function in our description of the English clause system, so as to do our best not to encounter any of the definitional problems highlighted previously in the review of the major reference grammars and models of clause hierarchy. Let us start with the development of a hierarchy of functional integration. Even though the terms vary and are somewhat confounded, all the models of hierarchy just reviewed acknowledged that clauses functioning as arguments/complements are more integrated than are modifying clauses, and these are more integrated than adjunct, hypotactic and paratactic clauses. This suggests that syntactic function has its own degrees of integration, which can be modelled apart from form (though this is not a claim that they are psycholinguistically independent as discussed in the next chapter). The functional category names by Huddleston and Pullum (2002) seem more than sufficient with the addition of an independent category to account for parataxis/juxtaposition of clauses) and can be distributed along a hierarchy such as that in Fig. 2.4.

The model in Fig. 2.4 can be taken to mean clauses that function in a matrix clause as a complement of the main verb are maximally integrated into that clause. Such a function is at the highest hierarchical node in a clause, a primary constituent. Modifiers are less integrated, as they typically play no essential grammatical role in their matrix clause and are embedded in a phrasal constituent within the matrix clause. Adjuncts are even less integrated as they can be removed without affecting grammaticality. This is similar to modifiers, but adjuncts exist outside the matrix clause syntax—i.e. they do not function within any constituents that fulfil a verb valency role in the main clause. If a clause appears in a more integrated function, it is a more integrated constituent into the clause with which it combines. However, the case can be made that any functional hierarchy is secondary to a formal hierarchy. This is because in language the form that a clause takes is highly correlated to its typical function, due to functional pressures on the form during historical

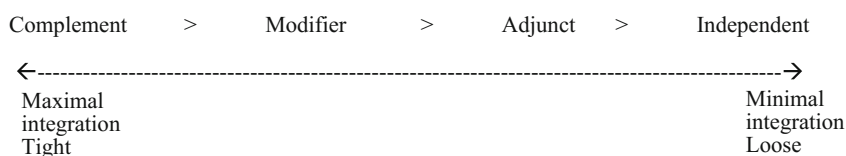


Fig. 2.4 Hierarchy of syntactic function

development (Givón 2001; Hopper and Traugott 2003). Function is a more abstract category, which as Huddleston and Pullum (2002) point out is not a property of the clause system per se but applies across the grammar in general. In other words, from a cognitive perspective one might propose that function has shaped form over time and therefore a clause hierarchy is a formal system that reflects function rather than a functional system that reflects form.

Let us therefore examine the different forms that constitute the common core of English clauses, and model them as the clause hierarchy of English. It is suggested that the reference grammars of English reviewed previously in the chapter can be synthesized into a description of English clause combination that has nine types. Relatively familiar category labels are given in (1) with prototypical examples taken from BNC (Spoken), Christine version (Sampson et al. 2000).

(1) *Core category types of English clause system*

- (a) Coordinate (symmetric): He hated Leeds *and* everyone hates Leeds [T10_3780]
- (b) Coordinate (asymmetric): I must tread gingerly this week *or* things will backfire [T12_7053]
- (c) Adverbial: I'll show you, *when you come out* [T11_7746]
- (d) Comparative: It is probably more *than the car is worth* [T20_5958]
- (e) Relative: I've got another girl *that does all my buying* [T22_0363]
- (f) Content: I think *that is right* [T22_7053]
- (g) Past participle: You don't want the ground *splattered with horrible things* [T17_7869]
- (h) Present participle: there was a problem *buying the house* [T22_5604]
- (i) Infinitival: I've got enough for tomorrow *to get me there* [T16_5058].

The clauses in (1) are to be taken as formally defined categories in the English clause system. That is, a form-first approach like the *Longman Grammar* (Biber et al. 1999) and *Cambridge Grammar* (Huddleston and Pullum 2002) rather than the *Collins Cobuild* (Sinclair 1990) and the *Comprehensive Grammar* (Quirk et al. 1985). For example, the term relative clause is retained for a category defined (prototypically) as a finite form of noun modification that contains a relative pronoun within an NP. The forms can be distributed along a hierarchy of grammatical integration, as summarized in Fig. 2.5.

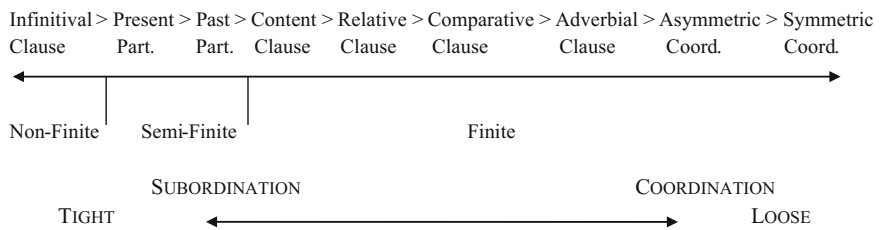


Fig. 2.5 The hierarchy of combined clauses in English

The position of each clause on the hierarchy of forms represented in Fig. 2.5 has the following principles behind it. One starts with the observation that English has three categories based on the finiteness of the verb: non-finite, semi-finite and finite. Non-finite clauses are considered tightest as they have no independent tense-aspect properties in their form, while finite clauses are considered loosest as they are completely inflectionally independent. An intermediate category is identified, following Payne (2011) and Givón (2001), of semi-finite in which clauses retain limited aspectual distinctions. Given this three way tight to loose categorization based on tense and aspect, the infinitival clause is analysed as the tightest form, defined as a non-finite clause with an infinitive verbal element that may or may not have a *to* onset. This form can be considered tightest due to being the only completely non-finite clause in English. The scope of the tense-aspect-mood of the verbal element in the matrix clause to which it attaches, determines the interpretation of these properties in the infinitival clause. Furthermore, its position at the top of the hierarchy can be supported by its other patterns. For example, a common construction in which the form appears is where the infinitive merges into the verbal element of another to create the pattern as in ‘I want *to go to the theatre*’, known as a catenative chain. Unlike most other clauses in the system, this is very hard to disambiguate as either bi-clausal or mono-clausal.

The present participle and past participle clauses are categorized as less integrated than the infinitival clause on the hierarchy due to their morphological features of being semi-finite. Retaining their aspectual distinctions indicates more independence from the main clause verbal element than with a fully non-finite clause. Furthermore, the aspect of the present participle clause in English denotes simultaneity of the subordinate clause with the main clause proposition, while a past participle denotes perfective completion and a sequence of events (Givón 2001). Langacker (2008) phrases it as a profiling distinction: the present participle profiles an event or state that has not been completed within a larger discourse event, while the past participle clause has a more independent profile. Hierarchically then, the present participle exhibits more integration than the past participle, which combines different states between propositions (Givón 2001; Haiman 1985).

Content clauses and relative clauses are both fully finite forms, both prototypically marked by a subordinator onset, and indeed often this onset has the same form of *that*. While therefore they are somewhat superficially similar in form in some circumstances, they are disambiguated on the clause hierarchy by a few key properties that give them different levels of integration. For example, content clauses need no grammatical antecedent and *that*, when it does appear, is not pronominal nor does it have a structural role in the subordinate clause such as subject or object. Also, the content clause embeds in a combined clause as an argument, and is semantically essential to a combined clause unit: hence its nomenclature as providing the ‘content’ of an argument and its other common name in many grammars as a ‘complement clause’. Thus, this form has as one of its features a correlation with the most integrated function in Fig. 2.4. Relative clauses, in contrast to content clauses, require an antecedent noun phrase and *that* is pronominal (even in less prototypical category members such as fused relative

clauses). Finally, the relative clause could be removed from a sentence without making it ungrammatical, so its prototypical place on the functional hierarchy of integration is lower than the content clause, in the modifier function of Fig. 2.4.

The comparative clause is defined as a fully finite clause marked by an adjective inflection *more/-er* followed by a *than* preposition, or the correlative pattern *as...as*. The form is difficult to place hierarchically. As has long been known, it is a complex category in English that seems to be both a single class but a heterogeneous range of forms (Huddleston 1967). For example, scholars often include in the comparative clause category clauses introduced by ‘so’, ‘so that’, and ‘the same as’. At its core however, the clause is finite, introduced by a clause connector, and not typically offset intonationally from its matrix clause. Since the clause does not function as an argument or core constituent like the content or relative clause, it is perhaps looser than these, and yet the comparative clause is surely higher in the clause system hierarchy than adverbial clauses. For example, the comparative clause is not often easy to remove, nor a background proposition, e.g. *she ran faster than all her colleagues did*. Rather, it is often central to the communicative intent. It also exhibits reduction phenomena such as sluicing and verb substitution where the main clause verb is carried over to the subordinate clause. Relatives tend to be constituents in noun phrases which are themselves in argument roles, while comparatives tend to complete adjectival subject complements, as in ‘it was harder *than I thought*’. Thus, the comparative clause is positioned between relative clauses and adverbial clauses.

The adverbial clause in the model described in Fig. 2.5 is defined as a finite form, with a grammatical subject, and typically marked by a subordinator. It is often set apart from the main clause by a comma in writing or intonational reset in speech. It is usually optional, i.e. patterns with the adjunct in function, but the form is multifunctional and can sometimes be essential to grammaticality albeit in restricted contexts as an adverbial complement (Biber et al. 1999). Regardless, it always modifies a verbal element in the main clause with semantics of time, place, consequence, condition, etc., and so has a wide range of onsets to mark it as a subordinate clause in relation to the main clause, including: *after, then, for, when, as, because, in, so, therefore, however*. Given the discussion previously in the chapter, one might ask why the adverbial clause category has been retained as part of the clause system in contrast to Huddleston and Pullum (2002) who argue that it is identical to the content clause. The reasons why this author is not fully persuaded that the category does not exist are as follows. A content clause has a semantically empty connector, prototypically ‘that’ (ellipsis is beside the point for the argument), whereas the adverbial clause has semantically meaningful connectors (Winter 1977). Furthermore, if one accepts the argument that adverbial clauses are content clauses only in a different function, then ‘that’ should be able to occur in them since this is part of the structure of the content clause. Yet, this is disallowed as demonstrated in (2):

(2) *Content and Adverbial clauses*

She said *that I was there*.

She had this idea *I was there*.

She had this idea, *because I was there*.

*She had this idea, *because/since/when/after that I was there*.

The content clause subordinator cannot co-occur when it functions as complement to a preposition in an adjunct to the matrix clause. If this connector problem is disregarded as unimportant to disambiguating the clause categories, then the argument is that they are the same clause because they are both fully finite, SVO clauses with obligatory grammatical subjects. But from this it follows that they are also in no formal way different from a canonical main clause which have these properties also, the only difference being they are functional dependents to another clause.

Coordination, following Huddleston and Pullum (2002), is separated in the hierarchy into two classes: symmetric for forms that have constituents able to be reversed while maintaining grammaticality, and asymmetric for forms that are fixed. Both are fully finite clauses, with grammatical subjects and are typically SVO. They exhibit no syntactic embedding in the clause with which they combine, though often at the interpropositional level they exhibit discourse level embedding and provide background information to the prior clause. Thus, a coordination–subordination distinction is retained, albeit with ‘fuzzy’ category boundaries. Coordinate clauses are prototypically grammatically marked by a core set of coordinating conjunctions mainly, *and*, *or*, *but*, though following the *Collins Cobuild* (Sinclair 1990) and the *Comprehensive Grammar* (Quirk et al. 1985) they may have stacked conjunctions such as ‘and then’, ‘but yet’, ‘yet also’ and retain their class category. Usually, stacked conjunctions signal asymmetric coordination as the clausal constituents cannot be reversed. Because asymmetric coordination exhibits certain grammatical restrictions on its clausal constituents, and correlates with conjunctions that signal dependency relations, e.g. ‘and then’, it can be considered higher on the hierarchy than symmetric coordination. However, this is not a strong commitment. Another possible analysis is that there is only one clause category of coordination in the system, in which asymmetric coordination is a subtype and/or radial category member (see Chap. 3).

The clause system of English therefore exhibits a range of formal, and functional, properties that signal differential levels of integration and provide principled reasons for modelling the system as the clause hierarchy in Fig. 2.5. The clause hierarchy unifies the clause system of English in such a way that describes how the properties that constitute each clause vary in relation to others in the system along a cline of integration into another superordinate clause. This is an observation beyond describing the different uses that a clause can be put to in discourse, which other grammars cover very well. The remainder of this chapter is a preliminary exploration of evidence that the clause hierarchy has demonstrable effects in the English language.

2.4.2 *Some Patterns of the Clause Hierarchy*

If hierarchy is a not just an artefact of the linguistic analysis above but a real property of English grammar, i.e. part of what speakers know about their language, one would expect it to have some detectable effects in the language. It is the aim of much of the rest of the book to investigate and describe the effects of clause hierarchy in the language. The rest of this chapter will offer a first approach, and consider if there is evidence in corpus data that syntactic functions, tense-aspect continuity and subject continuity in combined clauses meaningfully reflect the position on the clause hierarchy proposed in Fig. 2.5. The more integrated a clause, the more one might expect it to have a more integrated function and share the tense, aspect and participants of the clause with which it combines. To test this idea, a data set of 450 combined clauses was created. This represented 50 examples of each clause type along the hierarchy. The combined clause data was taken from the Australian Corpus of English (ACE). The ACE is modelled in design and balance after the Brown and LOB corpora, though collected decades later during 1985–89. The corpus contains 17 genres and 500 samples of running text.

The first 50 occurrences of each combined clause in the clause hierarchy were taken from ACE sub-sections A and K. Only those approximating a canonical ‘citation’ form (i.e. typical forms one would expect in a descriptive grammar) were extracted, i.e. the common core approach (Biber et al. 1999). In the corpus data, as in all discourse, there exist numerous non-canonical forms, such as clauses with multiple embedding, subject dropping, combined clauses broken up by long stylistic interpolations, etc. The decision was made to pass over these clauses in selecting clauses for analysis given the restricted scope of this study, though non-prototypical category members are explored in following chapters.

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 at phrasal or clausal boundaries that extended beyond 20 words. The principles are demonstrated in the following extracted present participle clause:

Extraction of combined clauses

ACE ‘Belles Letters’ Corpus:

- (a) Original sentence:

Under the influence of alcohol, Johnston behaved obnoxiously to his brother discussing him with Thomas Keneally and others as an example of Australian working class culture, listing his virtues and his ‘ocker’ weaknesses as if Jack was not even in the room

- (b) Extracted combined clause (present participle):

Under the influence of alcohol, Johnston behaved obnoxiously to his brother discussing him with Thomas Keneally and others.

In some cases however the target clause of a combination was itself a constituent in some larger sentence, perhaps subordinated in a larger stretch of discourse. In such cases, the subject head constituted the extraction point:

Extraction of combined clauses

ACE 'Belles Letters' Corpus:

(a) 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.

(b) Extracted combined clause (content 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 along the clause hierarchy were extracted, all 450 (9×50) combined clauses in the corpus were entered into SPSS and coded for:

1. Syntactic function [complement (C), modifier (M), or adjunct (A)]
2. Same subject (EQ) or different subjects (SR) between the clause constituents in the combined form.
3. Tense-Aspect (\pm TA) continuity between the clause constituents in the combined form.

The following presents the quantitative patterns of syntactic function, tense-aspect and subject continuity in the combined clause data. Table 2.1 reports the frequency of syntactic functions according to clause type along the clause hierarchy.

Starting with the highest frequency for a function, the results indicate coordinate clauses were categorically adjuncts, an artefact since they are always grammatically optional and are non-headed constructions. The second highest frequency and

Table 2.1 Clause type and syntactic function

		Syntactic function			Total
		Adjunct	Modifier	Complement	
Clause Type	To-infinitival	9	12	29	50
	Present participle	32	7	11	50
	Past participle	22	27	1	50
	Content clause	1	9	40	50
	Relative clause	13	37	0	50
	Comparative clause	44	0	6	50
	Adverbial clause	45	0	5	50
	Asymmetric coordination	50	0	0	50
	Symmetric coordination	50	0	0	50
Total		231	92	127	450

proportion of adjuncts, 45 (90 %), was found in adverbial clauses, though 5 (10 %) of these functioned as adverbial complements. Comparative clauses were difficult to code, much as they were difficult to situate on the clause hierarchy. Huddleston and Pullum (2002) describe them as functioning as complements to prepositions, but given the focus here was the clause, they were coded according to their function in the superordinate clause: e.g. they were considered adjuncts in a clause combination such as ‘Many more birds had been sold in retail outlets *than had been bred in captivity*’, and complements in patterns such as ‘the faces aren’t *as pretty as they once were*’. Under this analysis, 44 (81 %) in the data were adjuncts and 6 (12 %) were complements. Generally speaking, the clauses that are at the loosest end of the clause hierarchy also patterned with the least integrated function of being an adjunct.

Content clauses most commonly, 40 (80 %), functioned as complements, with fewer as NP modifiers in constructions like ‘she made a complaint *that I’d spent her money*’. It is marginally possible for content clauses to function as adjuncts but this only occurred once in the data in the borderline case between content and adverbial clause in, ‘now *that I think about it*, my feet are a lot better’. Though at the middle of the formal hierarchy, the content clause tends to associate with the tightest function of verb complementation; nevertheless, the common term in many grammars to refer to this as a ‘complement’ clause because they are typically predicates of perception-cognition-utterance verbs implies the clause form is categorically restricted to being a verbal complement.

Past participle clauses were distributed fairly evenly between adjuncts, 22 (44 %), and modifiers, 27 (54 %), rarely functioning as complements, 1 (2 %). One might have expected a majority to be modifiers given a common emphasis on their use as relative clause substitutes: e.g. ‘the man *seen yesterday*’. But, it appears that the modifier use is not the overwhelmingly dominant functional pattern of the form. Present participle clauses were more regularly complements than the past participle clauses at 11 (22 %), and more rarely modifiers, 7 (14 %). They were most often adjuncts, 32 (64 %). Again, the infrequency of functioning as a modifier is interesting as this is a common presentation of participle clauses. Rather than NP modification, the participle clause patterned mostly in the data as adjuncts in constructions such as ‘Mr Shandran is arguing for an extended stay of execution *pending a plea for the Malaysian Supreme Court*.’

Relative clauses, as expected, were overwhelmingly modifiers, 37 (74 %). However, the fact that 13 (26 %) functioned as adjuncts reflects a number of non-restrictive and sentential relatives in the data. There were no fused relatives functioning as complements in the data. Finally, to-infinitival clauses, the tightest form, were most often complements, 29 (58 %); thus in the tightest function. They also consistently functioned as modifiers, 12 (24 %), and adjuncts, 9 (18 %), indicating they were the most evenly patterned of the clause forms across the syntactic functions.

Frequencies of function present a complex picture, particularly towards the more integrated end of the clause hierarchy, but it would not be completely unfounded to suggest the clause system reflects some evidence of a kind of ‘natural syntax’ (Haiman 1985), whereby the tighter forms tend to associate with tighter functions,

Table 2.2 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
	Adverbial clause	21	29	50
	Asymmetric coordination	18	32	50
	Symmetric coordination	16	34	50
Total		130	170	450

and looser forms with looser functions. For the loosest clauses on the hierarchy, patterning with the loosest adjunct function was quite clear. However, it was not the case that clauses higher in the hierarchy, i.e. the infinitival and participle clauses, clearly patterned with the most integrated function.

Table 2.2 reports the number of clauses which maintained 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.

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 tend to be also, seems unsupportable. This hypothesis was based on the consideration that the tightest forms in English were non-finite and semi-finite, suggesting they relied more on the tense-aspect of the main clause to temporally contextualize the whole combined form. Consequently, as finite clauses moved along the hierarchy towards tight semi/non-finites, they might have higher frequencies of tense-aspect continuity reflecting this trend towards increasing temporal integration. The data suggest the opposite tendency. It seems in Table 2.2 that the tighter a combined clause of English, the less it tends to have tense/aspect continuity between constituents. Content clauses and relative clauses, for example, had discontinuous tense/aspect more frequently than looser clauses like asymmetric and symmetric coordination.

Table 2.3 reports the frequency of clauses that maintained the same subject across combined clausal constituents, and the number which contained different subjects (sometimes referred to as switch reference). Included here are both the grammatical subjects appearing in the finite clauses and logical subjects in the semi-finite and non-finite clauses (e.g. *for him to go on looking*). The hypothesis

Table 2.3 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
	Adverbial clause	19	31	50
	Asymmetric coordination	35	15	50
	Symmetric coordination	15	35	50
Total		199	251	450

was that the tighter the combined clauses, the more often they would exhibit subject continuity reflecting they were a more integrated combination.

Results in Table 2.3 indicate some evidence for the hypothesis that combined clauses further along the hierarchy more frequently maintained subject continuity. The general trend shown in Table 2.3 is that shared subjects declined in frequency as combined clause forms became looser, and increased as they became tighter. Putting aside asymmetric coordinate clauses for the moment, to-infinitival clauses, the tightest form, had the highest frequency of 34 out of 50 clauses with 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. With past participles, however, it should also be noted that distributions were comparable with 52 % same subjects and 48 % different subjects. However, the data does show that as a group the non-finite and semi-finite clauses much more frequently maintained subject continuity in comparison to the looser clauses. Content clauses, relative clauses, comparative clauses, adverbial clauses, and symmetric coordination all had higher frequencies of different subjects between their constituents.

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 (Givón 2001) so usually in the predicate of clauses, while subjects in a matrix clause tend to be ‘reduced’, i.e. short (Du Bois 2003). Relative clauses are heavy constituents and therefore prefer late positions in a clause, tending to fall within predicates more often than pre-verbal constituents. Rather than subject NPs, relative clauses therefore typically modify predicate NPs and these are infrequently co-referential with subject NPs in the same clause. Consequently, this would lead to relative clauses infrequently having the same subjects as the main clause.

One of the most interesting distributions for subject continuity is for the coordinate clause category. Asymmetric coordination heavily preferred the same subject across clauses, with 35 out of 50 clauses exhibiting subject continuity. This was only one more continuous subject than the to-infinitival clauses and thus not particularly different from the tightest form. It is however in stark contrast to all other

clauses low on the hierarchy, in particular the other form of coordination. Asymmetric frequencies were the exact inverse of symmetric coordination, which itself followed the general trend of the hierarchy to exhibit less subject continuity as clauses became less integrated. Recall that following Huddleston and Pullum (2002), an asymmetric coordinate clause was defined as a form which could not have its clausal constituents switched without deteriorating grammaticality. Since the data here indicate asymmetric clauses patterned with subject continuity, it seems that subject continuity is not conducive to interchanging constituents, despite both constituents being about the same thing/subject. This suggests the discourse function of asymmetric coordination is for the second clausal constituent to provide an expansion or elaboration of the first clause, perhaps more like a predicate or adverbial subordinate clause than its counterpart symmetric coordination, which is closer to an independent sentence.

2.5 Summary

This chapter has shown that the clause system of English can be described and understood as a clause hierarchy. Well beyond the binary categories of coordination and subordination, English clauses exist along a grammatical cline progressing from looser to tighter integration. The clauses each have a distinct level of integration in relation to one another, which is reflected in the grammatical properties of the clauses. This chapter has also reviewed several of the major reference grammars of English, noting their different approaches to describing the clause system and areas where they disagree, sometimes quite substantially. The clause hierarchy that has been developed is a synthesized and extended description of clause combination as described in the: *The Longman Grammar of Spoken and Written English* (Biber et al. 1999), *The Cambridge Grammar of the English Language* (Huddleston and Pullum 2002), *The Collins Cobuild English Grammar* (Sinclair 1990) and *The Comprehensive Grammar of the English Language* (Quirk et al. 1985). A preliminary investigation of some corpus data in this chapter indicated that hierarchy amongst the combined clauses has consequences for other aspects of English grammar and discourse, which suggests that the clause hierarchy is a cognitively salient feature of the English clause system. The following chapters will explore the patterns of the clause hierarchy more completely.

References

- Aarts, Bas. 2007. *Syntactic gradience: The nature of grammatical indeterminacy*. Oxford: Oxford University Press.
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad, and Edward Finegan. 1999. *Longman grammar of spoken and written English*. London: Longman.

- Biber, Douglas, Susan Conrad, and Viviana Cortes. 2004. If you look at...: Lexical bundles in university teaching and textbooks. *Applied linguistics* 25(3): 371–405.
- Bril, Isabelle (ed.). 2010. *Clause linking and clause hierarchy*. Amsterdam: John Benjamins.
- Chomsky, Noam. 2007. Biolinguistic explorations: Design, development, evolution. *International Journal of Philosophical Studies* 15(1): 1–21.
- Du Bois, John. 2003. Argument structure. In *Preferred argument structure: Grammar as architecture for function*, eds. John Du Bois, Lorriane Kumpf and William Ashby, 11–60. Amsterdam: Benjamins.
- Everett, Daniel. 2012. What does Pirahã grammar have to teach us about human language and the mind? *Wiley Interdisciplinary Reviews: Cognitive Science* 3(6): 555–563.
- Givón, Talmy. 1979. *On understanding grammar*. New York: Academic Press.
- Givón, Talmy. 1993. *English grammar: A function based introduction*. Amsterdam: Benjamins.
- Givón, Talmy. 2001. *Syntax: an introduction*, vol. 1. Amsterdam: Benjamins.
- Givón, Talmy. 2002. *Bio-linguistics: the Santa Barbara lectures*. Amsterdam: Benjamins.
- Green, Clarence. 2012. An analysis of hierarchy in English clause combination. *International Journal of Applied Linguistics and English Literature* 1(2): 1–17.
- Haiman, John. 1985. *Iconicity in syntax*. Amsterdam: Benjamins.
- Halliday, M.A.K. 1965. Speech and situation. *English in Education* 2: 14–17.
- Hopper, Paul, and Elizabeth Traugott. 2003. *Grammaticalization*. Cambridge: Cambridge University Press.
- Huddleston, Rodney. 1967. More on the English comparative. *Journal of Linguistics* 3(1): 91–102.
- Huddleston, Rodney, and Geoffrey Pullum. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
- Langacker, Ronald. 2008. *Cognitive grammar: A basic introduction*. Oxford: Oxford University Press.
- Matthiessen, Christensen. 2002. Combining clauses into clause complexes: A multi-faceted view. In *Complex sentences in grammar and discourse: Essays in honor of Sandra A. Thompson*, eds. Joan Bybee and Michael Noonan, 235–319. Amsterdam: Benjamins.
- Morris, Jane, and Grahame Hirst. 1991. Constructions: A new theoretical approach to language. *Computational Linguistics* 17(1): 21–48.
- Myhill, Debra. 2009. Becoming a designer: Trajectories of linguistic development. In *The Sage handbook of writing development*, ed. Roger Beard, Debra Myhill, Jeni Riley, and Martin Nystrand, 402–415. London: Sage Publications.
- O'Grady, William, Miseon Lee, and Miho Choo. 2003. A subject-object asymmetry in the acquisition of relative clauses in Korean as a second language. *Studies in Second Language Acquisition* 25(3): 433–448.
- Payne, Thomas. 2011. *Understanding English grammar: A linguistic introduction*. Cambridge: Cambridge University Press.
- Pullum, Geoffrey. 2009. Lexical categorization in English dictionaries and traditional grammars. *Zeitschrift für Anglistik und Amerikanistik* 57(3): 255–273.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartik. 1985. *A comprehensive grammar of the English language*. London: Longman.
- Sampson, Geoffrey, Anna Rahman, and Alan Morris. 2000. *The Christine corpus (Release 2)*. Sussex: University of Sussex.
- Sinclair, John. 1990. *Collins COBUILD English grammar*. London: Collins.
- Sjolie, Dennis. 2006. Phrase and clause grammar tactics for the ESL/ELL writing classroom. *English Journal* 1: 35–40.
- Winter, Eugene. 1977. A clause relational approach to English texts: A study of some predicative lexical items in written discourse. *International Science* 6: 1–92.

Patterns and Development in the English Clause
System

A Corpus-Based Grammatical Overview

Green, C.

2017, IX, 199 p. 36 illus., Hardcover

ISBN: 978-981-10-2880-9