CHAPTER I

The Complexity of Linguistic Structure

1.1 A sociological problem

Those of us who make it our business to study language often find ourselves in the curious position of trying to persuade the world at large that we are engaged in a technically demanding enterprise. Mathematicians are not expected to be able to relate their work to others: “Oh, I never could do math!” And although biologists and neuroscientists may be expected to explain the goals of their research in a very general way, the formidable chemical and physiological details that constitute the real substance of their work are freely granted to be beyond the understanding of non-specialists.

But language seems to be a different story. When we begin describing the sorts of issue we deal with, a typical response is “Oh, yes, I know how hard language is: I tried to learn Russian once!” When we try to explain that, no, that’s not the half of it, we rapidly lose our audience’s attention. The reaction is understandable: who but a linguist wants to hear at a cocktail party or barbecue about current thinking on long-distance extraction or the role of extrametricality in stress assignment?

Language and biology provide an interesting contrast in this respect. People expect to be baffled or bored by the biochemical details of, say, cell metabolism, so they don’t ask about them. What interests people about biology is natural history—strange facts about animal behavior and so forth. But they recognize and respect the fact that most biologists don’t study that. Similarly, what interests people about language is its “natural history”: the etymology of words, where language came from, and why kids talk so badly these days. The difference is that they don’t recognize that there is more to language than this, so they are unpleasantly disappointed when the linguist doesn’t share their fascination.
It may well be that individuals who are attracted into linguistics have a certain talent for metalinguistic reflection—a delight in constructing ungrammatical sentences, finding curious ambiguities and implicatures, hearing and imitating accents, and the like—and that professional training as a linguist only amplifies this proclivity. It would then be no surprise that linguists' sense of what is interesting in language is different from that of our friends in biology, economics, and dentistry. It is just that we linguists have made the mistake of assuming everyone else is like us. We are sort of in the position of an avid opera-goer who hasn't quite caught on that he is in the company of a bunch of rock-and-rollers.

In itself this mutual misperception would be merely a harmless source of social annoyance. But it can have a more pernicious side. A good example is the debate in the middle 1990s over the proposal by the School Board of Oakland, California that the dialect of the African-American community ("Ebonics") be employed as an integral part of class instruction (see Perry and Delpit 1998). The proposal was based on the well-documented structural integrity of this dialect as a language from the point of view of linguistics, and on linguistic research suggesting that instruction in a vernacular can be a valuable scaffolding to support acquisition of literacy in the standard language (Labov 1972; Rickford 1999). However, scientific documentation was of no concern to the general public. During the height of the debate, it was not uncommon to find letters to the editor in newspapers and even scientific journals to the effect that whatever the linguists in their ivory towers might think, Ebonics is a barbaric perversion of English, and it is nonsense to encourage its use.1

This example is revealing in two ways. First, it underlines the importance of language to social identity, and the way linguistic issues come to be conflated with broader social attitudes. Unlike cell metabolism, language is something that people have a personal stake in. Second, it demonstrates the way that people feel entitled to enter the conversation in the absence of expertise, and even to demean the experts in the interest of making a point. The behavior is reminiscent of creationists reacting to evolutionary theory, or oil corporations reacting to evidence of global warming.

Of course the experts aren't always right, and often there are dueling experts arrayed on opposite sides of a politically charged issue. But one ought to feel obliged at least to consider some of the facts and arguments that the experts have to offer before rejecting them.

Even when broader social issues are not at stake, one finds a curious readiness on the part of the lay public to speculate about the nature of language, and to be satisfied with the answers they arrive at. Unfortunately, such attitudes sometimes extend to influential practitioners of disciplines impinging on linguistics such as philosophy, neuroscience, and computer science. The effect is to divide disciplines from each other when they should be providing mutual support.

The goal of this chapter, therefore, is to establish a baseline of what a theory of linguistic structure must be responsible for. For linguists, the chapter should serve as a reminder of the scope of the enterprise and as an orientation into the outlook of the present study. As for those in neighboring disciplines, I certainly don't expect them to follow every detail of the structure to be described. But I hope the chapter will help them see beyond "natural history" aspects of language and to recognize the subtler issues at stake in linguistic theory, so that in succeeding chapters I can bring these issues to bear on psychological and biological concerns.

To be sure, there are many other aspects to language besides raw structure, such as its role in social identity and its power in poetry. But it is in the study of structure that I think linguistic theory finds its deepest and most characteristic concerns, and it is structure that will be the principal focus here.

1.2 The structure of a simple sentence

A good way to get into the complexity of linguistic structure is through a rather full analysis of a very simple sentence of English:

(t) The little star's beside a big star.

The structure of this sentence is given in Fig. 1.1 (next page). Fig. 1.1 is organized into four domains or levels, each of which has a characteristic formal structure. I have given a pretty fair articulation of phonological (sound) structure and syntactic structure, though many particulars are still omitted. The gross articulation of semantic/conceptual structure (alias meaning) is given here, but much of the detail is unknown. I have also included a very sketchy articulation of a level I have called spatial structure, the level at which this sentence can be compared with the perception of the world.

Fig. 1.1 serves as a useful baseline because nearly every bit of structure encoded in it represents a broad consensus among linguists—even if there is disagreement
about exactly how various aspects are to be systematically formalized. In particular, some variant of the phonological and syntactic structure will be found in every introductory textbook of linguistics.

The format of Fig. 1.1 is, however, a bit idiosyncratic in the way it divides the structure into levels. I have chosen this format in anticipation of the theory of linguistic structure to be developed in Part II: each level of structure is to be thought of as the product of an independent combinatorial ("generative") system.

But this is getting ahead of the story. For the moment let us take a tour of Fig. 1.1.

1.3 Phonological structure

The phonological structure consists of four subcomponents or tiers. Down the middle is the *segmental structure*, the string of discrete speech sounds or phonemes, notated here in the phonetic alphabet (representing more or less standard American pronunciation). However, the speech sounds are actually composite, made up of a matrix of *distinctive features*. Fig. 1.2 zooms in on the segmental structure of the word *star*, now broken into distinctive features. These features define a similarity space among speech sounds. Here is a classic example that illustrates their usefulness. The English regular plural suffix has three pronunciations: 's' as in *cats*, 'z' as in *dogs*, and 'az' ('uhz') as in *horses*. The choice of which to use is determined by the final sound of the word that the suffix is attached to—in particular the sound's distinctive features. The [-voiced] sound 's' is used with words that end with a [-voiced] sound ('t', 'p', 'k', 'f', etc.); the [+voiced] sound 'z' is used with words that end with a [+voiced] sound ('d', 'b', 'g', 'm', etc. plus vowels); and 'az' is used with words that end with the sounds 's', 'z', 's' ('sh'), 'Z' ('zh'), 'c' ('ch'), or 't', all of which have feature compositions.

![Fig. 1.2. Detail of segmental structure of the word star](image-url)
close to those of 's' and 'z'. That is, the generalizations concerning the exact phonetic form of the plural suffix fall out very naturally from the similarity space defined by the distinctive features. Hundreds of phenomena of this sort have been studied by phonologists. These features play a role in child language acquisition, in historical change, and in speech errors, as well as in the description of many dozens of languages. It is a scientific question what the right set of features is—one that is gradually being settled through phonological and phonetic research.

However, phonological structure is more than just a sequence of phonemes: it is organized into a number of semi-independent "levels," labeled along the right-hand side of the phonology in Fig. 1.1. One is the grouping of speech sounds into syllables, indicated by o. Now it could have been that syllables were just unstructured strings of phonemes, like this:

![Image of a tree diagram with a branch labeled 's' and two leaves labeled 'b' and 'g'.]

(2)

But in fact there are hierarchical distinctions inside the syllable. A syllable has to have one segment that functions as a Nucleus—the sonorous core around which the syllable is built. This is designated by N in Fig. 1.1. The nucleus is usually a vowel, but consonants with the distinctive feature [sonorant] can also serve as syllabic nuclei. One of these is "l," which serves as nucleus in the second syllable of little, as seen in Fig. 1.1.

The rest of the syllable's structure is optional. The nucleus and any following material (called the Coda) are grouped as the Rime (the part of the syllable that remains the same in rhymes). The material before the nucleus is grouped as the Onset (the part that remains the same in alliteration). These are indicated in Fig. 1.1 by R and O respectively. Notice also in Fig. 1.1 that the segment D in little (a "flapped" d or t) is ambisyllabic: it serves both as coda of one syllable and as onset of the next.

Above the syllabic structure is the tier of prosodic structure, which has two components. The brackets indicate the organization of the syllables into intonational phrases; pauses in pronouncing the sentence can occur only between bracketed units. Within the brackets are the xs of the metrical grid, which indicates the relative stress of syllables. Syllables with no xs above them are unstressed; more xs indicate more stress, so that the word big receives the main stress of the sentence.

Looking now below the phonological string, we find the tier of morphophonological structure—the grouping of the speech stream into words (indicated by Wd). Notice that the words the and a do not have the symbol Wd below them. Rather, they are treated phonologically as clitics—phonological fragments that attach to adjacent words to form a larger Wd constituent. Finally, notice that the sound 'z' by itself also is a clitic, notated orthographically in sentence (1) by 's.'

1.4 Syntactic structure

Consider next the syntactic structure. This is a tree diagram of the familiar sort. The largest constituent, the sentence (S), divides into a noun phrase (NP) (which serves as the subject) and a verb phrase (VP) (which serves as the predicate); the NP divides into a Determiner, a modifying adjectival phrase (AP), and a head noun (N), which carries the features 3rd person count singular. (If this were a French or German sentence, the Det and A would also have number, and all constituents of NP would have grammatical gender as well.) The VP divides into a head verb (V) and a prepositional phrase (PP), the PP divides into a preposition and its NP object, and the NP divides like the subject NP. Attached to the V is an Inflection which includes present tense plus the features 3rd person singular, which agree with the subject.

The way I have notated this tree differs from standard convention in two respects. First, it is customary to put the words of the sentence at the bottom of the tree, as in (3).

![Image of a tree diagram with a branch labeled 'NP' and two branches labeled 'Det' and 'AP' with 'N', 'the', 'A', and 'star' as leaves.]

(3)

I have omitted the words for reasons to be discussed more fully in Chapter 5. The basic reason is that things like the and star are actually pieces of phonology, not syntax. The only aspects of words that play a role in syntax are the part of speech (Det, N, etc.) and syntactic features such as 3rd person singular and present tense.

The other way this notation differs from tradition is that some of the connections in the tree are notated with double lines. These are connections between phrases and their heads. The idea is that phrases like NP, VP, AP, and
PP are "projections" of their heads—N, V, A, and P respectively. For example, there couldn't be a prepositional phrase with a noun as head or a verb phrase with a preposition as head. So each phrase is to be thought of as a structural skeleton, indicated by the double lines, supplemented by elaborations, indicated by single lines. This is not a standard notation, but it makes more graphic an insight about syntactic structure that goes back at least to Zellig Harris in the 1940s and that was introduced into modern theory by Chomsky (1970) under the rubric of "X-bar" theory (see also Jackendoff 1977). 3

In Fig. 1.1, I have continued the double line from V up beyond VP to S, in effect treating the sentence as a syntactic projection of the verb. An important variant, originating in Chomsky (1986), is shown in (4). Here the sentence is the projection of Inflection (Tense).

![Diagram](image)

Again, as in phonology, there is ongoing debate about this seemingly simple choice; in section 1.6 we will return to some repercussions. Nevertheless, everyone agrees that overall there is constituent structure of this sort, and that constituents have heads.

A further aspect of X-bar theory has not been noted here: syntactic categories such as Noun and Verb, like phonological segments, are often treated as feature complexes rather than primitives. The most broadly accepted set (dating from Chomsky 1970) analyzes Noun as [+N, −V] and Verb as [−N, +V]; Adjective is [+N, +V] and Preposition is [−N, −V].

3 In most approaches to syntax, this distinction between single and double lines can be regarded as merely a visual aid, of no theoretical significance. However, some versions of syntactic theory take the distinction between the head of a phrase and its other constituents to be marked in the structure. For instance, in Head-Driven Phrase Structure Grammar (HPSG, Pollard and Sag 1994), the head falls under the HEAD-DTR ("head–daughter") attribute, and other constituents fall under the COMP–DTR ("complement–daughters") attribute. Under this approach the double line has a theoretical interpretation, as it stands for a different kind of connection between the head and the phrase. For the moment I leave open the decision of how to regard the double line notation. It is, however, illustrative of the sort of subtle but significant questions that can arise in syntactic theory.

1.5 Semantic/conceptual and spatial structure

If phonology and syntax are fairly well settled at least at this relatively gross level, there is considerably less agreement about the proper formulation of semantic/conceptual structure, even with respect to whether there is such a thing. Fig. 1.1 unabashedly presents one aspect of this structure in my own version of the notation. Chapter 12 will call this aspect the descriptive tier; it corresponds roughly to the information that might be encoded in a predicate logic. Chapter 12 will also motivate other tiers of semantic/conceptual structure, related to this tier in roughly the way the tiers of phonology are related. One is the referential tier, which corresponds roughly to the aspects of meaning added when moving from a predicate logic to a quantification logic. Another is the tier of information structure, the division of the content of the sentence into foreground and background (topic/focus and presupposition).

The structure given in Fig. 1.1 is a labeled bracketing, in which each pair of brackets surrounds a conceptual constituent. The label on a constituent designates it as belonging to a major conceptual type such as Situation, Event, State, Object, Place, or Property.

Two kinds of relation among conceptual constituents appear in this structure. The first is function–argument structure, notated as in (5).

![Bracket](image)

Here F is a function that maps a constituent of type Y and a constituent of type Z into a constituent of type X. (5) shows a two-place function. Be is such a function in Fig. 1.1. There are also one-place functions (such as beside in Fig. 1.1) and possibly three-place functions. The second kind of relation is modification, notated as in (6).

![Bracket](image)

(6) is a constituent of type X, in which the inner constituent, of type Y, specifies a further characteristic of the outer constituent. An example of this in Fig. 1.1 is the modification of the first Object constituent by the Property LITTLE.

Using this notation, the conceptual structure in Fig. 1.1 says that there is a Situation in the present, consisting of a State. This State is one of an Object being located in a Place; the function BE maps the Object and the Place into this State.

Now look at the Object that is the first argument of BE. It has three pieces of structure. The first designates the Object as of the category STAR (which presumably has more internal articulation, not noted here: see Chapter 11). The second piece is a marker DEF ('definite'), which indicates roughly that the identity of
the object in question can be fixed by either the previous discourse or the context (if there were two little stars around, one couldn’t say the little star). The third piece is a modifying constituent of the type Property, which designates the object as having the characteristic little.

Little has further internal structure, which again I have not formalized: basically it says that the overall size of the object in question is smaller than a pragmatically determined norm. This norm in turn may be chosen from (a) the average size of members of the category in question (here, stars), (b) the average size of stars in the contextual environment (here, only two), (c) the average size of all comparable objects in the contextual environment (as in The star is little, the circle is big), and perhaps others (Bierwisch and Lang 1989).

Turning to the rest of the semantic/conceptual structure in Fig. 1.1: the other Object, a big star, works the same way as the little star. A big star, however, serves as the argument of a function beside, which maps the Object into a region or Place—the region in which the first Object is located by the function be. Again, I can give some detail of beside: the region beside an object X is exterior to X, proximal to X, and in a horizontal direction from X. Hence in (7a) the Y is near but not beside the X. In addition, no other object can come between X and an object next to it; hence in (7b) the Y is also near but not beside the X.

(7) a. X
    Y

    b. X Z Y

I emphasize again that the notation for semantic/conceptual structure in Fig. 1.1 is by no means universally accepted. However, all the distinctions that have been mentioned here arise sooner or later in every theory of meaning (and what is sooner and what later, and by how much, is one of the major divides among theories).

If the details of semantic/conceptual structure are sketchy and open to dispute, those of spatial structure are hardly even touched upon. One can think of spatial structure variously as an image of the scene that the sentence describes, a schema that must be compared against the world in order to verify the sentence (a “mental model” in Johnson-Laird’s (1983) sense), the physical (or non-propositional) structure of the model in which the truth conditions of semantic/conceptual structure are applied, or perhaps other construals. (I will be more precise in Chapters 9 and 11.)

What is clear is that any such image requires two star-shaped objects (or object-schemas) in it. More interesting is that the features of beside must appear in some way also in this configuration, so that “beside-ness” can be verified in a visually presented array. I have notated the region “beside a big star” very crudely in Fig. 1.1 by means of a dotted line that represents the approximate boundary of the relevant region.

1.6 Connecting the levels

The structure of our sentence cannot just be the collection of these structures. It is necessary also to encode the relationships among them—how the parts of each structure are connected to parts of the others.

I have notated correspondences between units of phonological structure and syntactic structure with pre-subscripts. For example, the phonological clitic the carries the pre-subscript 0, which places it in correspondence with the initial Determiner in syntax. Similarly, correspondences between units of syntax and units of semantic/conceptual structure are notated with post-subscripts. For instance, the initial Det in syntax is coindexed with the feature DEF in semantic/conceptual structure. As with the structures themselves, many aspects of the correspondences noted in Fig. 1.1 are somewhat sketchy and imprecise; we will return to some of the details later. Still, their general outline should be clear.4

An important feature of these correspondences—whatever their details—is that for the most part they do not obtain between primitive elements of any of the levels: they are rather relations between composite units. The primitive units of phonological structure such as distinctive features and syllables are completely invisible to syntax and meaning. What this means is that the speech sounds are themselves meaningless. Only the assembly of a number of speech sounds into a word or clitic has a connection to syntax and/or meaning. Similarly, the word’s syntactic category such as Noun or Verb (itself a composite of syntactic features) is invisible to phonology. In other words, a word’s part of speech gives no hint as to how to pronounce it.

Various other primitive units of syntax have no direct connection to phonological structure. They are therefore by themselves unpronounceable. Consider for instance 3rd person. It must combine with the features singular and the element present tense before there is a unit that (sometimes) has a pronunciation as the verbal affix -s, as in the verb spms. In this particular sentence, however, even 3rd person singular present tense has no independent pronunciation. Rather, it is bundled up with the verb to form a unit that is often pronounced is, but that in this case is contracted into z (spelled’s), a clitic attached to the previous word.

Not every aspect of syntax corresponds to something in meaning either. For

4 In particular, alert readers may wonder how the pre-subscripts a and b connect the phonology to the syntax, and why the VP in syntax does not connect to anything in phonology. Please wait until Chapter 5.
example, the features 3rd person and singular on the verb are purely syntactic agreement features that have no particular effect in semantic/conceptual structure (in English at least).

Another thing to notice about these correspondences is that the units that are connected between phonology and syntax are not always the same units that are connected between syntax and conceptual structure. For example, the inflected verb (the upper V in Fig. 1.1) is connected to the phonology (subscript f), where it appears as the clitic ‘z’; but the bare verb and the inflection are connected separately to semantics (subscripts 6 and 7), where they are separate elements.

Generally speaking, the mapping between phonology and syntax preserves linear order, while the mapping between syntax and meaning tends to preserve the relative embedding of arguments and modifiers. In particular, the head of a syntactic phrase tends to map into the outermost function of the corresponding conceptual constituent. For instance, the preposition beside, the head of the PP, maps into the function BESIDE that governs the Place-constituent in Fig. 1.1.

In turn, some but not all parts of semantic/conceptual structure correspond to spatial structure—in Fig. 1.1, the two Object-constituents and the Place. Other parts of conceptual structure are harder to represent directly in any spatial format. For instance, little and big raise the problem of how to notate relative size in spatial structure; definiteness (def) raises the problem of how to notate uniqueness. My impression is that these explicit pieces of conceptual structure encode distinctions that are only implicit in spatial structure—so it is hard to see how to notate the relationship with a simple co-subscribing.

One aspect of this correspondence merits special attention. As noted a moment ago, the little clitic ‘z’ in phonology is of course the contracted verb is, which expresses the verb be in the 3rd person singular present tense, a smallish part of the syntactic tree. In turn, the verb be corresponds to the next-to-largest function in semantic/conceptual structure. The largest function in semantic/conceptual structure is present tense. But this appears as a feature of the verb be in syntax, and is not even an independent element in the phonology. So, by virtue of this two-step correspondence, elements of relative insignificance in phonology can correspond to major organizing features of meaning. (A similar situation arises in vision, where tiny features of a boundary can dramatically affect the three-dimensional interpretation of an array.)

This behavior on the part of Tense is a good illustration of the kinds of tensions that arise in syntactic theory. Tense has been noted in Fig. 1.1 as a feature on the verb, making it easy to match to phonology. But as a result it is more difficult to match to meaning, because it is necessary to say, exceptionally, that this feature attached to a verb (rather than the verb itself) maps to the function of largest scope in semantic/conceptual structure. We can improve the mapping to semantic/conceptual structure by adopting the alternative syntactic structure shown in (4), with Inflection as the head of the sentence: now the superordinate syntactic head maps to the largest-scope semantic function. But this creates a problem in mapping to the phonology, since the clitic ‘z’ now must match two separate pieces of syntax at once, the Verb and the Tense. So changing the syntactic analysis to simplify one mapping makes the other mapping more complex.

A third possibility is to keep both correspondences simple by localizing the complexity in the syntactic component itself. This has been the standard approach in generative grammar. The idea is that the syntactic structure of our sentence contains two different trees. The form given in Fig. 1.1 is the “surface structure,” which interfaces easily with the phonology; and the form in (4) is the “underlying (or deep) structure,” which interfaces easily with meaning. Then, internal to syntax, these two forms are related by a transformation that combines the underlying Inflection and Verb into the single unit found in surface structure. This approach was the major innovation in Chomsky’s Syntactic Structures (1957) and has been a staple of syntactic analysis ever since.

Whichever of these three ways to deal with Tense proves correct, the point is that there is a mismatch between phonology and meaning, which has to be encoded somewhere in the mapping among the levels of structure. If this mismatch is eliminated at one point in the system, it pops up elsewhere. Much dispute in modern syntax has been over these sorts of mismatch and how to deal with them. (I don’t think most linguists have viewed it this way, though.) We will encounter such mismatches pervasively in the course of Parts II and III.

1.7 Anaphora and unbounded dependencies

For the sake of completeness, let me step away from our little sentence for a moment, to mention briefly two syntactic phenomena that have been the focus of a great deal of research in linguistic theory and that will turn up now and again in the present study: anaphora and unbounded dependencies.

The set of constraints on the use of anaphoric elements such as pronouns and reflexives has come to be called Binding Theory (not to be confused with the neuroscientist’s notion of binding, to be taken up in Chapter 3). Some standard examples of reflexives appear in (8). (The notation * before a sentence indicates that it is judged ungrammatical.)
In particular, the conditions crucially involve linguistic structure, and not just linear order. Thus it is no surprise that these conditions have been a constant preoccupation of linguistic research. The main lines of dispute are whether the linguistic structure involved in conditions on anaphora is syntactic structure alone (Chomsky 1981; Lasnik 1989), or whether semantic/conceptual structure plays a role as well or even instead (six independent approaches among many appear in Jackendoff 1972; Fauconnier 1983; Kuno 1987; Levinson 1987; Van Hoek 1993; and Culicover and Jackendoff 1995).

For a different sort of phenomenon, consider the examples in (11). The italicized elements are understood as having a role appropriate to the position marked by t. For instance, in (11a), *which movie* is understood as the object of the verb *saw*.

(11) a. *Which movie* does Susan imagine that Sarah saw t last night?  
    [wh-direct question]

b. John was wondering *who* Sarah decided she would go to the movies with t on Sunday.  
    [Indirect question]

c. I didn’t like the movie *which* you said that everyone was talking about t the other day.  
    [Relative clause]

d. You may take *whichever sandwich* you find t on the table over there.  
    [Free relative]

e. *That movie*, I wouldn’t recommend that anyone consider taking their kids to t.  
    [Topicalization]

It is significant that the understood position can be within a subordinate clause, in fact deeply embedded within multiple subordinate clauses, as in (11e). For this reason, the relation between the italicized constituent and the understood position is called a long-distance dependency.

The analysis of such constructions within mainstream generative grammar is that the italicized constituent is actually in its understood position in underlying (deep) structure, and that it moves to the fronted position in the course of a syntactic derivation. The movement leaves behind an “unpronounced pronoun” called a trace, which is indicated by t in (11). However, other generative frameworks, especially Head-Driven Phrase Structure Grammar (Pollard and Sag 1994), have proposed analyses in which there is no movement, but instead the grammar directly establishes an anaphora-like relation between the italicized constituent and the trace (or a formal equivalent of the trace).

These constructions pose an interesting problem in that there are strong constraints on the structural position that an “extracted” constituent can occupy in relation to its trace. For example, an “extracted” constituent cannot be outside of a conjoined construction (12a), a relative clause (12b), an indirect question

(8) a. Joe adores himself.  

b. Joe thinks that Fred adores himself.  

 c. *Joe thinks that you adore himself.*  

Example (8a) shows that a reflexive pronoun in object position can co-refer with (or be bound by) an NP in subject position. (8b) shows that it must co-refer with the subject of the same clause, not just with any subject. (8c) shows that if the reflexive cannot be bound by the subject of its own clause—here because you and *himself* cannot co-refer—its use is ungrammatical (in English; certain other languages such as Japanese work differently).

Turning to ordinary pronouns, consider the examples in (9).

(9) a. Joe adores him.  
  [him ≠ Joe]  

b. Joe thinks that Fred adores him.  
  [him = Joe or third party]  

c. Joe thinks that you adore him.  
  [him = Joe or third party]  

d. He thinks that Joe adores Fred.  
  [he ≠ Joe or Fred]  

e. If you tickle Joe, he laughs.  
  [he can = Joe]  

f. If you tickle him, Joe laughs.  
  [him can = Joe]  

Example (9a) shows that a simple pronoun in object position, unlike the reflexive in (8a), must not co-refer with the subject of its clause. (9b, c) show that it can; however, co-refer with an NP in another clause—again unlike a reflexive. (9d) shows that a pronoun cannot co-refer with an NP in a subordinate clause on its right. However, (9e) shows that a pronoun can co-refer with an NP in a subordinate clause on its left. In (9f) the pronoun is, atypically, to the left of its antecedent; but this case is saved by the fact that the pronoun is in a subordinate clause.

Another sort of anaphoric element is the expression *do so*, which stands for a VP rather than an NP. Its relation to its antecedent resembles that of pronouns. Compare (10a–d) to (9c–f).

(10) a. Fred impressed the boss without trying to do so.  
  [do so = impress the boss]  

b. Fred did so without trying to impress the boss.  
  [did so ≠ impress the boss]  

c. Without TRYING to impress the boss, Fred did so.  
  [did so = impress the boss]  

d. Without trying to do so, Fred impressed the boss.  
  [do so = impress the boss]  

These examples make it clear that it is a complex matter to state the exact conditions under which an anaphoric element can co-refer with an antecedent.
(12c), or a noun complement (12d) that the trace is inside of. The examples in (12) all involve direct wh-questions, but the same thing happens with all the constructions in (10).

(12) a. *What did Beth eat peanut butter and if for dinner?
   b. *Who does Sam know a girl who is in love with it?
   c. *Who does Betty know which professor flunked it?
   d. *What food were you never aware of the hypothesis that you shouldn't eat it?

As with anaphora, it has been a preoccupation of linguistic research for three decades (starting with Ross 1967) to characterize precisely the environments from which “extraction” is possible. Again, one of the issues is whether the criteria are completely syntactic or partially semantic as well (Ertesvik-Shir and Lappin 1979; Deane 1991; Kluender 1992; Van Valin 1994; Culicover and Jackendoff 1997). But the overall outlines of the phenomenon are clear.

*   *   *   *   *

Any adequate theory of language must begin with the fact that even the simplest sentences contain at least this rich a structure. Although I don't feel comfortable making moral statements, I will make one nevertheless. In my opinion, if one wishes to join the conversation about the nature of language, one must recognize and acknowledge this complexity. One need not have an account of all of it, but one may not willfully ignore it and still expect to be allowed in the game. This is the minimum that scientific responsibility demands.

Having recognized all this complexity, the obvious question is: What can we make of it? That is what the rest of this book is about.

---

CHAPTER 2

Language as a Mental Phenomenon

2.1 What do we mean by “mental”?

The remarkable first chapter of Noam Chomsky's *Aspects of the Theory of Syntax* (1965) sets in place an agenda for generative linguistic theory, much of which has survived intact for over thirty-five years. The present chapter and the next two will be devoted to evaluating and rearticulating this agenda, and to replying to some of the more common and longstanding criticisms of the approach.

We follow *Aspects* by starting with the issue of the status of linguistic description. The standard techniques of linguistic research lead us to some posited structure, say Fig. 1.1, for the sentence *The little star's beside a big star*. How is such a structure to be understood? The fundamental claim of *Aspects* is that this structure is more than just a useful description for the purposes of linguists. It is meant to be “psychologically real”: it is to be treated as a model of something in the mind of a speaker of English who says or hears this sentence. What does this claim mean?

Often the answer is put in these terms: Fig. 1.1 is a model of a mental representation of the sentence. Unfortunately, I have to plunge right in and attempt to wean readers away from this terminology, which I think has led to unnecessary and prolonged misunderstanding. The problem is that the term "representation" suggests that it represents something—and for something to represent something else, it must represent it to someone. But we don't want to say that Fig. 1.1 "represents the sentence to the language user": that would suggest somehow that the language user has conscious access to all the structure in the figure, or could have it with sufficient introspective effort. Nor do we want to say that the figure represents the sentence to some entity within the language user's unconscious mind: that would conjure up the notorious homunculus,