• the arbitrariness of the sign
• discrete combinatorial system (examples: grammars, DNA)
• A generative grammar consists of a finite collection of discrete elements and a finite number of rules for combining them, yielding an infinite output. The property that allows for an infinite output is called “recursion” (discussion of which appears at the bottom of p.100).
• Structure is distinct from meaning. That is, a sentence can be well-formed according to the rules of syntax, but nonetheless may be nonsensical or meaningless to some degree. On the other hand, it is possible to convey information using ungrammatical sentences (where “ungrammatical” means “not among the well-formed sentences given by the grammar”).
• Page 90 begins a discussion of a machine for generating strings of words, a word-chain device. So a string like Casey often sings is given the following analysis: Casey – often – sings where the relationship between the subject and verb is not formally expressed. The only relations expressible are those between “Casey” and “often” and between “often” and “sings”. In a model such as this, it is easy to see that there is no way to account for the subject-verb agreement between “Casey” and “sings”.
• Why is the model of the word-chain insufficient to account for the properties of natural language? Pinker gives several other examples.
• The word-chain model describes the surface string in terms of what words are next to what other words. As such, it is not powerful enough to describe relationships that hold between nonadjacent words or phrases (“long-distance dependencies”). So we cannot understand the whole structure or meaning of a sentence simply by looking at the linear order of the words in the surface string, since we also need to know something about the relationships between words that are not adjacent. This is the kind of information that the hierarchical structure of the sentence can give us.
• Words are not simply strung along one after the other. Instead, there are complex hierarchical relations that hold among them. Our sentence Casey often sings has the following hierarchical structure:

Here we see that often and sings are collected together under one node, corresponding to the verb phrase. The other underlined node represents the subject, which here consists of only one word. It is easy to see that the relation of subject-verb agreement can be expressed in the tree structure using the structural notion of “sisterhood”.
• Note that subject-verb agreement is now “local”. In the word-chain model, the subject and the verb were nonadjacent. In the hierarchical tree structure, they are adjacent at the level of the underlined nodes.
• On p.98, Pinker begins his discussion of phrase structure grammars, the tree structures they yield, and how these structures are to be interpreted. The X’ (“X-bar”) structures introduced on p.107 are a more
highly articulated version of the simplified structures that are used to introduce the discussion of trees here.

- What we think of as “parts of speech” are nothing more than grammatical category labels. They tell us how that word may be used in the sentence, what positions in the tree it may appear in, and what relation it may have to other parts of speech. These labels do not have any direct relation to a word’s meaning. Similarly, it is misleading to think of the “subject” and “object” as code words for the “doer” and “receiver” of an action. In the syntax, subject and object are merely positions in a tree.

- Words are combined to make phrases, and phrases are combined to make sentences. Every phrase has a “head”, which corresponds to a single part of speech. The head of a NP is N, the head of a VP is V, and so on. Likewise, every head is associated with a phrase.

- In *Casey often sings*, the phrase “Casey” is an NP which has as its head the noun “Casey”, the phrase “often sings” is a VP which has the head “sings”, and “often” is an AdvP headed by the adverb “often”.

- The X’ level is the level of structure between the head (X) and full phrase (XP) levels. Why do we need an additional level of structure? Consider the phrase *drink milk*. Here, *milk* is a role-player with respect to *drink* (that is, “milk” is an argument of “drink”). We represent this fact by showing the argument as a sister of the verb, and use the V´ level to accomplish this.

```
  VP
   \ /
    V V'
      NP
        drink milk
```

Because we want “object of” to be a very local relationship, we attach the verb and its NP object as closely as possible. The V´ level is of course “closer” to V than VP is. A modifier would not be attached as the sister of the verb, as this position is reserved for arguments.

- In English, the head comes first and then the object, but it could just as well be the other way around. The notion of “sisterhood” says nothing about the order in which the elements under a given node appear. We write them in the order we do simply because that is the way they fall out in English. In English we say “drink milk”, and in Japanese it is “milk drink”. In both cases the structure consists of a head V and an argument NP that combine to form V´. The only difference is in the order in which they appear underneath the V´.

- Hierarchical structure is as a universal property of human language, and the specific order in which heads and their arguments appear is determined on a language-particular basis. That is, the child already knows that heads their arguments combine to form part of a phrase that looks like the one drawn above, but the child will have to figure out whether the phrases of its language are head-final (like Japanese) or head-initial (like English).

- Word order and case-marking serve identical functions: that of determining relations among words in the surface string. They serve as clues to the hierarchical structure that must be inferred from the surface string.

- The verb determines what arguments may appear along with it.

- The head of the sentence is a verbal element, namely, inflection. A sentence is of category IP, for “inflectional phrase”. To this point the discussion in the book has centered on nouns, verbs, prepositions – those words which contribute meaning to the sentence. These are known as content
words. Function words, on the other hand, give information about grammatical relations. Both content words and function words serve as heads and project phrases.

- Finally, the notions of “deep” and “surface” structure are introduced (“d-structure” and “s-structure”). We know that the passive *John was arrested by Bill* is related to the active *Bill arrested John*. We can relate them formally by showing that they are two different surface realizations of the same d-structure.