lifferent symbols that refer to the same thing, like *the tall blond man vith one black shoe* and *the man*. On the other hand, mentalese must ıe simpler than spoken languages; conversation-specific words and onstructions (like *a* and *the*) are absent, and information about ronouncing words, or even ordering them, is unnecessary. Now, it ould be that English speakers think in some kind of simplified and nnotated quasi-English, with the design I have just described, and ıat Apache speakers think in a simplified and annotated quasi-ıpache. But to get these languages of thought to subserve reasoning roperly, they would have to look much more like each other than ither one does to its spoken counterpart, and it is likely that they re the same: a universal mentalese.

Knowing a language, then, is knowing how to translate mentalese ıto strings of words and vice versa. People without a language would ill have mentalese, and babies and many nonhuman animals presum-bly have simpler dialects. Indeed, if babies did not have a mentalese) translate to and from English, it is not clear how learning English ɔuld take place, or even what learning English would mean.

So where does all this leave Newspeak? Here are my predictions ɔr the year 2050. First, since mental life goes on independently of articular languages, concepts of freedom and equality will be think-ɔle even if they are nameless. Second, since there are far more ɔncepts than there are words, and listeners must always charitably ll in what the speaker leaves unsaid, existing words will quickly gain ew senses, perhaps even regain their original senses. Third, since ıildren are not content to reproduce any old input from adults but :eate a complex grammar that can go beyond it, they would creolize lewspeak into a natural language, possibly in a single generation. he twenty-first-century toddler may be Winston Smith's revenge.

From Pinker, S. (1994). *The language instinct*. New York: Morrow.

4

How Language Works

Journalists say that when a dog bites a man that is not news, but when a man bites a dog that is news. This is the essence of the language instinct: language conveys news. The streams of words called "sentences" are not just memory prods, reminding you of man and man's best friend and letting you fill in the rest; they tell you who in fact did what to whom. Thus we get more from most stretches of language than Woody Allen got from *War and Peace,* which he read in two hours after taking speed-reading lessons: "It was about some Russians." Language allows us to know how octopuses make love and how to remove cherry stains and why Tad was heartbroken, and whether the Red Sox will win the World Series without a good relief pitcher and how to build an atom bomb in your basement and how Catherine the Great died, among other things.

When scientists see some apparent magic trick in nature, like bats homing in on insects in pitch blackness or salmon returning to breed in their natal stream, they look for the engineering principles behind it. For bats, the trick turned out to be sonar; for salmon, it was locking in to a faint scent trail. What is the trick behind the ability of *Homo sapiens* to convey that man bites dog?

In fact there is not one trick but two, and they are associated with the names of two European scholars who wrote in the nineteenth century. The first principle, articulated by the Swiss linguist Ferdinand de Saussure, is "the arbitrariness of the sign," the wholly conventional pairing of a sound with a meaning. The word *dog* does not look like a dog, walk like a dog, or woof like a dog, but it means

"dog" just the same. It does so because every English speaker has undergone an identical act of rote learning in childhood that links the sound to the meaning. For the price of this standardized memorization, the members of a language community receive an enormous benefit: the ability to convey a concept from mind to mind virtually instantaneously. Sometimes the gunshot marriage between sound and meaning can be amusing. As Richard Lederer points out in *Crazy English*, we drive on a parkway but park in a driveway, there is no ham in hamburger or bread in sweetbreads, and blueberries are blue but cranberries are not cran. But think about the "sane" alternative of depicting a concept so that receivers can apprehend the meaning in the form. The process is so challenging to the ingenuity, so comically unreliable, that we have made it into party games like Pictionary and charades.

The second trick behind the language instinct is captured in a phrase from Wilhelm Von Humboldt that presaged Chomsky: language "makes infinite use of finite media." We know the difference between the forgettable *Dog bites man* and the newsworthy *Man bites dog* because of the order in which *dog, man*, and *bites* are combined. That is, we use a code to translate between orders of words and combinations of thoughts. That code, or set of rules, is called a generative grammar; as I have mentioned, it should not be confused with the pedagogical and stylistic grammars we encountered in school.

The principle underlying grammar is unusual in the natural world. A grammar is an example of a "discrete combinatorial system." A finite number of discrete elements (in this case, words) are sampled, combined, and permuted to create larger structures (in this case, sentences) with properties that are quite distinct from those of their elements. For example, the meaning of *Man bites dog* is different from the meaning of any of the three words inside it, and different from the meaning of the same words combined in the reverse order. In a discrete combinatorial system like language, there can be an unlimited number of completely distinct combinations with an infinite range of properties. Another noteworthy discrete combinatorial system in the natural world is the genetic code in DNA, where four kinds of nucleotides are combined into sixty-four kinds of codons, and the codons can be strung into an unlimited number of different genes. Many biologists have capitalized on the close parallel between the principles of grammatical combination and the principles of genetic combination. In the technical language of genetics, sequences of DNA are said to contain "letters" and "punctuation"; may be "palindromic," "meaningless," or "synonymous"; are "transcribed" and "translated"; and are even stored in "libraries." The immunologist Niels Jerne entitled his Nobel Prize address "The Generative Grammar of the Immune System."

Most of the complicated systems we see in the world, in contrast, are *blending systems*, like geology, paint mixing, cooking, sound, light, and weather. In a blending system the properties of the combination lie *between* the properties of its elements, and the properties of the elements are lost in the average or mixture. For example, combining red paint and white paint results in pink paint. Thus the range of properties that can be found in a blending system are highly circumscribed, and the only way to differentiate large numbers of combinations is to discriminate tinier and tinier differences. It may not be a coincidence that the two systems in the universe that most impress us with their open-ended complex design—life and mind—are based on discrete combinatorial systems. Many biologists believe that if inheritance were not discrete, evolution as we know it could not have taken place.

The way language works, then, is that each person's brain contains a lexicon of words and the concepts they stand for (a mental dictionary) and a set of rules that combine the words to convey relationships among concepts (a mental grammar). We will explore the world of words in the next chapter; this one is devoted to the design of grammar.

The fact that grammar is a discrete combinatorial system has two important consequences. The first is the sheer vastness of language. Go into the Library of Congress and pick a sentence at random from any volume, and chances are you would fail to find an exact repetition no matter how long you continued to search. Estimates of the number of sentences that an ordinary person is capable of producing are breathtaking. If a speaker is interrupted at a random point in a sentence, there are on average about ten different words that could be inserted at that point to continue the sentence in a grammatical and meaningful way. (At some points in a sentence, only one word can be inserted, and at others, there is a choice from among thousands; ten is the average.) Let's assume that a person is capable of producing sentences up to twenty words long. Therefore the number of senten-

ces that a speaker can deal with in principle is at least 10^{20} (a one with twenty zeros after it, or a hundred million trillion). At a rate of five seconds a sentence, a person would need a childhood of about a hundred trillion years (with no time for eating or sleeping) to memorize them all. In fact, a twenty-word limitation is far too severe. The following comprehensible sentence from George Bernard Shaw, for example, is 110 words long:

> Stranger still, though Jacques-Dalcroze, like all these great teachers, is the completest of tyrants, knowing what is right and that he must and will have the lesson just so or else break his heart (not somebody else's, observe), yet his school is so fascinating that every woman who sees it exclaims: "Oh why was I not taught like this!" and elderly gentlemen excitedly enroll themselves as students and distract classes of infants by their desperate endeavours to beat two in a bar with one hand and three with the other, and start off on earnest walks around the room, taking two steps backward whenever M. Dalcroze calls out "Hop!"

Indeed, if you put aside the fact that the days of our age are threescore and ten, each of us is capable of uttering an *infinite* number of different sentences. By the same logic that shows that there are an infinite number of integers—if you ever think you have the largest integer, just add 1 to it and you will have another—there must be an infinite number of sentences. The *Guinness Book of World Records* once claimed to recognize the longest English sentence: a 1,300-word stretch in William Faulkner's novel *Absalom, Absalom!* that begins:

> They both bore it as though in deliberate flagellant exaltation . . .

I am tempted to achieve immortality by submitting the following record-breaker:

> Faulkner wrote, "They both bore it as though in deliberate flagellant exaltation . . ."

But it would be only the proverbial fifteen minutes of fame, for soon I could be bested by:

> Pinker wrote that Faulkner wrote, "They both bore it as though in deliberate flagellant exaltation . . ."

And that record, too, would fall when someone submitted:

> Who cares that Pinker wrote that Faulkner wrote, "They both bore it as though in deliberate flagellant exaltation . . ."?

And so on, ad infinitum. The infinite use of finite media distinguishes the human brain from virtually all the artificial language devices we commonly come across, like pull-string dolls, cars that nag you to close the door, and cheery voice-mail instructions ("Press the pound key for more options"), all of which use a fixed list of prefabricated sentences.

The second consequence of the design of grammar is that it is a code that is *autonomous* from cognition. A grammar specifies how words may combine to express meanings; that specification is independent of the particular meanings we typically convey or expect others to convey to us. Thus we all sense that some strings of words that can be given common-sense interpretations do not conform to the grammatical code of English. Here are some strings that we can easily interpret but that we sense are not properly formed:

> Welcome to Chinese Restaurant. Please try your Nice Chinese Food with Chopsticks: the traditional and typical of Chinese glorious history and cultual.
>
> It's a flying finches, they are.
>
> The child seems sleeping.
>
> Is raining.
>
> Sally poured the glass with water.
>
> Who did a book about impress you?
>
> Skid crash hospital.
>
> Drum vapor worker cigarette flick boom.
>
> This sentence no verb.
>
> This sentence has contains two verbs.
>
> This sentence has cabbage six words.
>
> This is not a complete. This either.

These sentences are "ungrammatical," not in the sense of split infinitives, dangling participles, and the other hobgoblins of the schoolmarm, but in the sense that every ordinary speaker of the casual vernacular has a gut feeling that something is wrong with them, despite their interpretability. Ungrammaticality is simply a consequence of our having a fixed code for interpreting sentences. For some strings a meaning can be guessed, but we lack confidence that the speaker has used the same code in producing the sentence as we used in interpreting it. For similar reasons, computers, which are less forgiving of ungrammatical input than human listeners, express their displeasure in all-too-familiar dialogues like this one:

```
› PRINT (x + 1
*****SYNTAX ERROR*****
```

The opposite can happen as well. Sentences can make no sense but can still be recognized as grammatical. The classic example is a sentence from Chomsky, his only entry in *Bartlett's Familiar Quotations*:

> Colorless green ideas sleep furiously.

The sentence was contrived to show that syntax and sense can be independent of each other, but the point was made long before Chomsky; the genre of nonsense verse and prose, popular in the nineteenth century, depends on it. Here is an example from Edward Lear, the acknowledged master of nonsense:

> It's a fact the whole world knows,
> That Pobbles are happier without their toes.

Mark Twain once parodied the romantic description of nature written more for its mellifluousness than its content:

> It was a crisp and spicy morning in early October. The lilacs and laburnums, lit with the glory-fires of autumn, hung burning and flashing in the upper air, a fairy bridge provided by kind Nature for the wingless wild things that have their homes in the tree-tops and would visit together; the larch and the pomegranate flung their purple and yellow flames in brilliant broad splashes along the slanting sweep of the woodland; the sensuous fragrance of innumerable deciduous flowers rose upon the swooning atmosphere; far in the empty sky a solitary esophagus slept upon motionless wing; everywhere brooded stillness, serenity, and the peace of God.

And almost everyone knows the poem in Lewis Carroll's *Through the Looking-Glass* that ends:

> And, as in uffish thought he stood,
> The Jabberwock, with eyes of flame,
> Came whiffling through the tulgey wood,
> And burbled as it came!
>
> One, two! One, two! And through and through
> The vorpal blade went snicker-snack!
> He left it dead, and with its head
> He went galumphing back.
>
> "And hast thou slain the Jabberwock?
> Come to my arms, my beamish boy!
> O frabjous day! Callooh! Callay!"
> He chortled in his joy.
>
> 'Twas brillig, and the slithy toves
> Did gyre and gimble in the wabe:
> All mimsy were the borogoves,
> And the mome raths outgrabe.

As Alice said, "Somehow it seems to fill my head with ideas—only I don't exactly know what they are!" But though common sense and common knowledge are of no help in understanding these passages, English speakers recognize that they are grammatical, and their mental rules allow them to extract precise, though abstract, frameworks of meaning. Alice deduced, "*Somebody* killed *something:* that's clear, at any rate—." And after reading Chomsky's entry in *Bartlett's*, anyone can answer questions like "What slept? How? Did one thing sleep, or several? What kind of ideas were they?"



How might the combinatorial grammar underlying human language work? The most straightforward way to combine words in

order is explained in Michael Frayn's novel *The Tin Men*. The protagonist, Goldwasser, is an engineer working at an institute for automation. He must devise a computer system that generates the standard kinds of stories found in the daily papers, like "Paralyzed Girl Determined to Dance Again." Here he is hand-testing a program that composes stories about royal occasions:

> He opened the filing cabinet and picked out the first card in the set. *Traditionally*, it read. Now there was a random choice between cards reading *coronations, engagements, funerals, weddings, comings of age, births, deaths*, or *the churching of women*. The day before he had picked *funerals*, and been directed on to a card reading with simple perfection *are occasions for mourning*. Today he closed his eyes, drew *weddings*, and was signposted on to *are occasions for rejoicing*.
>
> *The wedding of X and Y* followed in logical sequence, and brought him a choice between *is no exception* and *is a case in point*. Either way there followed *indeed*. Indeed, whichever occasion one had started off with, whether coronations, deaths, or births, Goldwasser saw with intense mathematical pleasure, one now reached this same elegant bottleneck. He paused on *indeed*, then drew in quick succession *it is a particularly happy occasion, rarely*, and *can there have been a more popular young couple*.
>
> From the next selection, Goldwasser drew *X has won himself/herself a special place in the nation's affections*, which forced him to go on to *and the British people have clearly taken Y to their hearts already*.
>
> Goldwasser was surprised, and a little disturbed, to realise that the word "fitting" had still not come up. But he drew it with the next card—*it is especially fitting that*.
>
> This gave him *the bride/bridegroom should be*, and an open choice between *of such a noble and illustrious line, a commoner in these democratic times, from a nation with which this country has long enjoyed a particularly close and cordial relationship*, and *from a nation with which this country's relations have not in the past been always happy*.
>
> Feeling that he had done particularly well with "fitting" last time, Goldwasser now deliberately selected it again. *It is also fitting that*, read the card, to be quickly followed by *we should remember*, and *X and Y are not merely symbols—they are a lively young man and a very lovely young woman*.
>
> Goldwasser shut his eyes to draw the next card. It turned out to read *in these days when*. He pondered whether to select *it is fashionable to scoff at the traditional morality of marriage and family life* or *it is no longer fashionable to scoff at the traditional morality of marriage and family life*. The latter had more of the form's authentic baroque splendour, he decided.

Let's call this a word-chain device (the technical name is a "finite-state" or "Markov" model). A word-chain device is a bunch of lists of words (or prefabricated phrases) and a set of directions for going from list to list. A processor builds a sentence by selecting a word from one list, then a word from another list, and so on. (To recognize a sentence spoken by another person, one just checks the words against each list in order.) Word-chain systems are commonly used in satires like Frayn's, usually as do-it-yourself recipes for composing examples of a kind of verbiage. For example, here is a Social Science Jargon Generator, which the reader may operate by picking a word at random from the first column, then a word from the second, then one from the third, and stringing them together to form an impressive-sounding term like *inductive aggregating interdependence*:

dialectical	participatory	interdependence
defunctionalized	degenerative	diffusion
positivistic	aggregating	periodicity
predicative	appropriative	synthesis
multilateral	simulated	sufficiency
quantitative	homogeneous	equivalence
divergent	transfigurative	expectancy
synchronous	diversifying	plasticity
differentiated	cooperative	epigenesis
inductive	progressive	constructivism
integrated	complementary	deformation
distributive	eliminative	solidification

Recently I saw a word-chain device that generates breathless book jacket blurbs, and another for Bob Dylan song lyrics.

A word-chain device is the simplest example of a discrete combinatorial system, since it is capable of creating an unlimited number of distinct combinations from a finite set of elements. Parodies notwithstanding, a word-chain device can generate infinite sets of grammatical English sentences. For example, the extremely simple scheme



assembles many sentences, such as *A girl eats ice cream* and *The happy dog eats candy*. It can assemble an infinite number because of the loop at the top that can take the device from the *happy* list back to itself any number of times: *The happy dog eats ice cream, The happy happy dog eats ice cream*, and so on.

When an engineer has to build a system to combine words in particular orders, a word-chain device is the first thing that comes to mind. The recorded voice that gives you a phone number when you dial directory assistance is a good example. A human speaker is recorded uttering the ten digits, each in seven different sing-song patterns (one for the first position in a phone number, one for the second position, and so on). With just these seventy recordings, ten million phone numbers can be assembled; with another thirty recordings for three-digit area codes, ten billion numbers are possible (in practice, many are never used because of restrictions like the absence of 0 and 1 from the beginning of a phone number). In fact there have been serious efforts to model the English language as a very large word chain. To make it as realistic as possible, the transitions from one word list to another can reflect the actual probabilities that those kinds of words follow one another in English (for example, the word *that* is much more likely to be followed by *is* than by *indicates*). Huge databases of these "transition probabilities" have been compiled by having a computer analyze bodies of English text or by asking volunteers to name the words that first come to mind after a given word or series of words. Some psychologists have suggested that human language is based on a huge word chain stored in the brain. The idea is congenial to stimulus-response theories: a stimulus elicits a spoken word as a response, then the speaker perceives his or her own response, which serves as the next stimulus, eliciting one out of several words as the next response, and so on.

But the fact that word-chain devices seem ready-made for parodies like Frayn's raises suspicions. The point of the various parodies is that the genre being satirized is so mindless and cliché-ridden that a simple mechanical method can churn out an unlimited number of examples that can almost pass for the real thing. The humor works because of the discrepancy between the two: we all assume that people, even sociologists and reporters, are not really word-chain devices; they only seem that way.

The modern study of grammar began when Chomsky showed that word-chain devices are not just a bit suspicious; they are deeply, fundamentally, the wrong way to think about how human language works. They are discrete combinatorial systems, but they are the wrong kind. There are three problems, and each one illuminates some aspect of how language really does work.

First, a sentence of English is a completely different thing from a string of words chained together according to the transition probabilities of English. Remember Chomsky's sentence *Colorless green ideas sleep furiously*. He contrived it not only to show that nonsense can be grammatical but also to show that improbable word sequences can be grammatical. In English texts the probability that the word *colorless* is followed by the word *green* is surely zero. So is the probability that *green* is followed by *ideas, ideas* by *sleep*, and *sleep* by *furiously*. Nonetheless, the string is a well-formed sentence of English. Conversely, when one actually assembles word chains using probability tables, the resulting word strings are very far from being well-formed sentences. For example, say you take estimates of the set of words most likely to come after every four-word sequence, and use those estimates to grow a string word by word, always looking at the four most recent words to determine the next one. The string will be eerily Englishy, but not English, like *House to ask for is to earn our living by working towards a goal for his team in old New-York was a wonderful place wasn't it even pleasant to talk about and laugh hard when he tells lies he should not tell me the reason why you are is evident.*

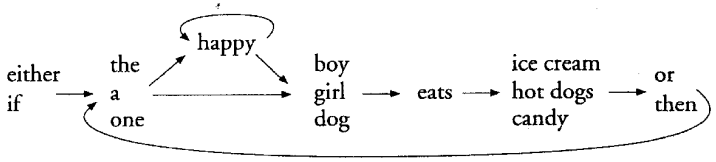
The discrepancy between English sentences and Englishy word chains has two lessons. When people learn a language, they are learn-

ing how to put words in order, but not by recording which word follows which other word. They do it by recording which word *category*—noun, verb, and so on—follows which other category. That is, we can recognize *colorless green ideas* because it has the same order of adjectives and nouns that we learned from more familiar sequences like *strapless black dresses*. The second lesson is that the nouns and verbs and adjectives are not just hitched end to end in one long chain; there is some overarching blueprint or plan for the sentence that puts each word in a specific slot.

If a word-chain device is designed with sufficient cleverness, it can deal with these problems. But Chomsky had a definitive refutation of the very idea that a human language is a word chain. He proved that certain sets of English sentences could not, even in principle, be produced by a word-chain device, no matter how big or how faithful to probability tables the device is. Consider sentences like the following:

> Either the girl eats ice cream, or the girl eats candy.
> If the girl eats ice cream, then the boy eats hot dogs.

At first glance it seems easy to accommodate these sentences:



But the device does not work. *Either* must be followed later in a sentence by *or*; no one says *Either the girl eats ice cream, then the girl likes candy*. Similarly, *if* requires *then*; no one says *If the girl eats ice cream, or the girl likes candy*. But to satisfy the desire of a word early in a sentence for some other word late in the sentence, the device has to remember the early word while it is churning out all the words in between. And that is the problem: a word-chain device is an amnesiac, remembering only which word list it has just chosen from, nothing

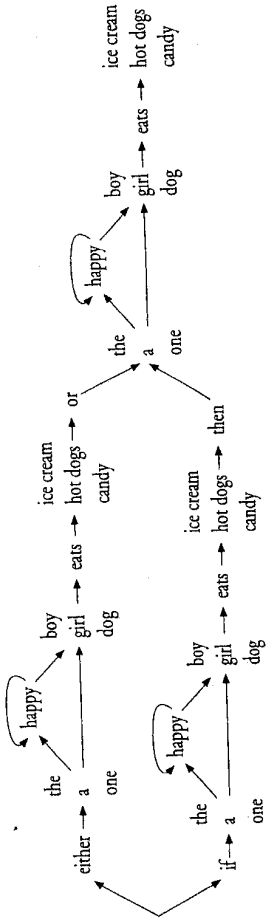
earlier. By the time it reaches the *or/then* list, it has no means of remembering whether it said *if* or *either* way back at the beginning. From our vantage point, peering down at the entire road map, we can remember which choice the device made at the first fork in the road, but the device itself, creeping antlike from list to list, has no way of remembering.

Now, you might think it would be a simple matter to redesign the device so that it does not have to remember early choices at late points in the sentence. For example, one could join up *either* and *or* and all the possible word sequences in between into one giant sequence, and *if* and *then* and all the sequences in between as a second giant sequence, before returning to a third copy of the sequence—yielding a chain so long I have to print it sideways (see page 96). There is something immediately disturbing about this solution: there are three identical subnetworks. Clearly, whatever people can say between an *either* and an *or*, they can say between an *if* and a *then*, and also after the *or* or the *then*. But this ability should come naturally out of the design of whatever the device is in people's heads that allows them to speak. It shouldn't depend on the designer's carefully writing down three identical sets of instructions (or, more plausibly, on the child's having to learn the structure of the English sentence three different times, once between *if* and *then*, once between *either* and *or*, and once after a *then* or an *or*).

But Chomsky showed that the problem is even deeper. Each of these sentences can be embedded in any of the others, including itself:

> If either the girl eats ice cream or the girl eats candy, then the boy eats hot dogs.
> Either if the girl eats ice cream then the boy eats ice cream, or if the girl eats ice cream then the boy eats candy.

For the first sentence, the device has to remember *if* and *either* so that it can continue later with *or* and *then*, in that order. For the second sentence, it has to remember *either* and *if* so that it can complete the sentence with *then* and *or*. And so on. Since there's no limit in principle to the number of *if*'s and *either*'s that can begin a sentence, each requiring its own order of *then*'s and *or*'s to complete it, it does no good to spell out each memory sequence as its own



chain of lists; you'd need an infinite number of chains, which won't fit inside a finite brain.

This argument may strike you as scholastic. No real person ever begins a sentence with *Either either if either if if*, so who cares whether a putative model of that person can complete it with *then . . . then . . . or . . . then . . . or . . . or*? But Chomsky was just adopting the esthetic of the mathematician, using the interaction between *either-or* and *if-then* as the simplest possible example of a property of language—its use of "long-distance dependencies" between an early word and a later one—to prove mathematically that word-chain devices cannot handle these dependencies.

The dependencies, in fact, abound in languages, and mere mortals use them all the time, over long distances, often handling several at once—just what a word-chain device cannot do. For example, there is an old grammarian's saw about how a sentence can end in five prepositions. Daddy trudges upstairs to Junior's bedroom to read him a bedtime story. Junior spots the book, scowls, and asks, "Daddy, what did you bring that book that I don't want to be read to out of up for?" By the point at which he utters *read*, Junior has committed himself to holding four dependencies in mind: *to be read* demands *to, that book that* requires *out of, bring* requires *up*, and *what* requires *for*. An even better, real-life example comes from a letter to *TV Guide*:

> How Ann Salisbury can claim that Pam Dawber's anger at not receiving her fair share of acclaim for *Mork and Mindy*'s success derives from a fragile ego escapes me.

At the point just after the word *not*, the letter-writer had to keep four grammatical commitments in mind: (1) *not* requires *-ing* (her anger at *not* receiv*ing* acclaim); (2) *at* requires some kind of noun or gerund (her anger *at* not *receiving acclaim*); (3) the singular subject *Pam Dawber's anger* requires the verb fourteen words downstream to agree with it in number (Dawber's *anger . . . derives* from); (4) the singular subject beginning with *How* requires the verb twenty-seven words downstream to agree with it in number (*How . . . escapes* me). Similarly, a reader must keep these dependencies in mind while interpreting the sentence. Now, technically speaking, one could rig up a word-chain model to handle even these sentences, as long as there is some actual limit on the number of dependencies that the speaker need keep in mind (four, say). But the degree of redundancy in the device would be absurd; for each of the thousands of *combinations* of dependencies, an identical chain must be duplicated inside the device. In trying to fit such a superchain in a person's memory, one quickly runs out of brain.



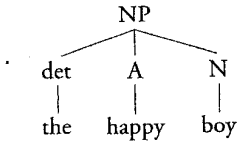
The difference between the artificial combinatorial system we see in word-chain devices and the natural one we see in the human brain is summed up in a line from the Joyce Kilmer poem: "Only God can make a tree." A sentence is not a chain but a tree. In a human

grammar, words are grouped into phrases, like twigs joined in a branch. The phrase is given a name—a mental symbol—and little phrases can be joined into bigger ones.

Take the sentence *The happy boy eats ice cream.* It begins with three words that hang together as a unit, the noun phrase *the happy boy.* In English a noun phrase (NP) is composed of a noun (N), sometimes preceded by an article or "determiner" (abbreviated "det") and any number of adjectives (A). All this can be captured in a rule that defines what English noun phrases look like in general. In the standard notation of linguistics, an arrow means "consists of," parentheses mean "optional," and an asterisk means "as many of them as you want," but I provide the rule just to show that all of its information can be captured precisely in a few symbols; you can ignore the notation and just look at the translation into ordinary words below it:

NP → (det) A* N
"A noun phrase consists of an optional determiner, followed by any number of adjectives, followed by a noun."

The rule defines an upside-down tree branch:



Here are two other rules, one defining the English sentence (S), the other defining the predicate or verb phrase (VP); both use the NP symbol as an ingredient:

S → NP VP
"A sentence consists of a noun phrase followed by a verb phrase."

VP → V NP
"A verb phrase consists of a verb followed by a noun phrase."

We now need a mental dictionary that specifies which words belong to which part-of-speech categories (noun, verb, adjective, preposition, determiner):

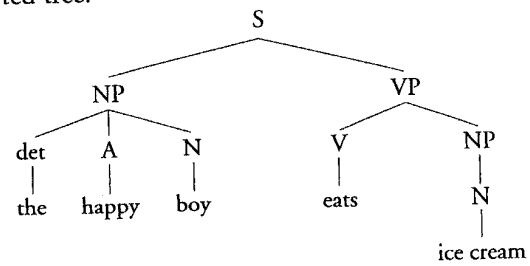
N → boy, girl, dog, cat, ice cream, candy, hot dogs
"Nouns may be drawn from the following list: *boy, girl, . . .*"

V → eats, likes, bites
"Verbs may be drawn from the following list: *eats, likes, bites.*"

A → happy, lucky, tall
"Adjectives may be drawn from the following list: *happy, lucky, tall.*"

det → a, the, one
"Determiners may be drawn from the following list: *a, the, one.*"

A set of rules like the ones I have listed—a "phrase structure grammar"—defines a sentence by linking the words to branches on an inverted tree:



The invisible superstructure holding the words in place is a powerful invention that eliminates the problems of word-chain devices. The key insight is that a tree is *modular,* like telephone jacks or garden hose couplers. A symbol like "NP" is like a connector or fitting of a certain shape. It allows one component (a phrase) to snap into any of several positions inside other components (larger phrases). Once a kind of phrase is defined by a rule and given its connector symbol, it never has to be defined again; the phrase can be plugged in any-

where there is a corresponding socket. For example, in the little grammar I have listed, the symbol "NP" is used both as the subject of a sentence (S → NP VP) and as the object of a verb phrase (VP → V NP). In a more realistic grammar, it would also be used as the object of a preposition (*near the boy*), in a possessor phrase (*the boy's hat*), as an indirect object (*give the boy a cookie*), and in several other positions. This plug-and-socket arrangement explains how people can use the same kind of phrase in many different positions in a sentence, including:

[The happy happy boy] eats ice cream.
I like [the happy happy boy].
I gave [the happy happy boy] a cookie.
[The happy happy boy]'s cat eats ice cream.

There is no need to learn that the adjective precedes the noun (rather than vice versa) for the subject, and then have to learn the same thing for the object, and again for the indirect object, and yet again for the possessor.

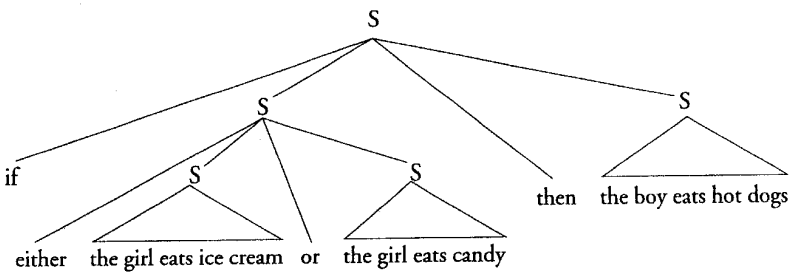
Note, too, that the promiscuous coupling of any phrase with any slot makes grammar autonomous from our common-sense expectations involving the meanings of the words. It thus explains why we can write and appreciate grammatical nonsense. Our little grammar defines all kinds of colorless green sentences, like *The happy happy candy likes the tall ice cream*, as well as conveying such newsworthy events as *The girl bites the dog*.

Most interestingly, the labeled branches of a phrase structure tree act as an overarching memory or plan for the whole sentence. This allows nested long-distance dependencies, like *if . . . then* and *either . . . or*, to be handled with ease. All you need is a rule defining a phrase that contains a copy of the very same kind of phrase, such as:

S → either S or S
"A sentence can consist of the word *either*, followed by a sentence, followed by the word *or*, followed by another sentence."

S → if S then S
"A sentence can consist of the word *if*, followed by a sentence, followed by the word *then*, followed by another sentence."

These rules embed one instance of a symbol inside another instance of the same symbol (here, a sentence inside a sentence), a neat trick—logicians call it "recursion"—for generating an infinite number of structures. The pieces of the bigger sentence are held together, in order, as a set of branches growing out of a common node. That node holds together each *either* with its *or*, each *if* with its *then*, as in the following diagram (the triangles are abbreviations for lots of underbrush that would only entangle us if shown in full):



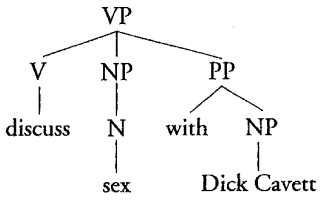
There is another reason to believe that a sentence is held together by a mental tree. So far I have been talking about stringing words into a grammatical order, ignoring what they mean. But grouping words into phrases is also necessary to connect grammatical sentences with their proper meanings, chunks of mentalese. We know that the sentence shown above is about a girl, not a boy, eating ice cream, and a boy, not a girl, eating hot dogs, and we know that the boy's snack is contingent on the girl's, not vice versa. That is because *girl* and *ice cream* are connected inside their own phrase, as are *boy* and *hot dogs*, as are the two sentences involving the girl. With a chaining device it's just one damn word after another, but with a phrase structure grammar the connectedness of words in the tree reflects the relatedness of ideas in mentalese. Phrase structure, then, is one solution to the engineering problem of taking an

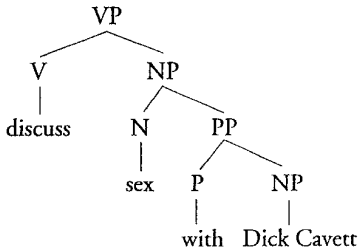
interconnected web of thoughts in the mind and encoding them as a string of words that must be uttered, one at a time, by the mouth.

One way to see how invisible phrase structure determines meaning is to recall one of the reasons mentioned in Chapter 3 that language and thought have to be different: a particular stretch of language can correspond to two distinct thoughts. I showed you examples like *Child's Stool Is Great for Use in Garden*, where the single word *stool* has two meanings, corresponding to two entries in the mental dictionary. But sometimes a whole sentence has two meanings, even if each individual word has only one meaning. In the movie *Animal Crackers*, Groucho Marx says, "I once shot an elephant in my pajamas. How he got into my pajamas I'll never know." Here are some similar ambiguities that accidentally appeared in newspapers:

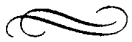
> Yoko Ono will talk about her husband John Lennon who was killed in an interview with Barbara Walters.
>
> Two cars were reported stolen by the Groveton police yesterday.
>
> The license fee for altered dogs with a certificate will be $3 and for pets owned by senior citizens who have not been altered the fee will be $1.50.
>
> Tonight's program discusses stress, exercise, nutrition, and sex with Celtic forward Scott Wedman, Dr. Ruth Westheimer, and Dick Cavett.
>
> We will sell gasoline to anyone in a glass container.
>
> For sale: Mixing bowl set designed to please a cook with round bottom for efficient beating.

The two meanings in each sentence come from the different ways in which the words can be joined up in a tree. For example, in *discuss sex with Dick Cavett*, the writer put the words together according to the tree at the left ("PP" means prepositional phrase): sex is what is to be discussed, and it is to be discussed with Dick Cavett.





The alternative meaning comes from our analyzing the words according to the tree at the right: the words *sex with Dick Cavett* form a single branch of the tree, and sex with Dick Cavett is what is to be discussed.



Phrase structure, clearly, is the kind of stuff language is made of. But what I have shown you is just a toy. In the rest of the chapter I will try to explain the modern Chomskyan theory of how language works. Chomsky's writings are "classics" in Mark Twain's sense: something that everybody wants to have read and nobody wants to read. When I come across one of the countless popular books on mind, language, and human nature that refer to "Chomsky's deep structure of meaning common to all human languages" (wrong in two ways, we shall see), I know that Chomsky's books of the last twenty-five years are sitting on a high shelf in the author's study, their spines uncracked, their folios uncut. Many people want to have a go at speculating about the mind but have the same impatience about mastering the details of how language works that Eliza Doolittle showed to Henry Higgins in *Pygmalion* when she complained, "I don't want to talk grammar. I want to talk like a lady in a flower shop."

For nonspecialists the reaction is even more extreme. In Shakespeare's *The Second Part of King Henry VI*, the rebel Dick the Butcher speaks the well-known line "The first thing we do, let's kill all the lawyers." Less well known is the second thing Dick suggests they do: behead Lord Say. Why? Here is the indictment presented by the mob's leader, Jack Cade:

> Thou hast most traitorously corrupted the youth of the realm in erecting a grammar school. . . . It will be proved to thy face that thou hast men about thee that usually talk of a noun and a verb, and such abominable words as no Christian ear can endure to hear.

And who can blame the grammarphobe, when a typical passage from one of Chomsky's technical works reads as follows?

> To summarize, we have been led to the following conclusions, on the assumption that the trace of a zero-level category must be properly governed. 1. VP is α-marked by I. 2. Only lexical categories are L-markers, so that VP is not L-marked by I. 3. α-government is restricted to sisterhood without the qualification (35). 4. Only the terminus of an X^0-chain can α-mark or Case-mark. 5. Head-to-head movement forms an A-chain. 6. SPEC-head agreement and chains involve the same indexing. 7. Chain coindexing holds of the links of an extended chain. 8. There is no accidental coindexing of I. 9. I-V coindexing is a form of head-head agreement; if it is restricted to aspectual verbs, then base-generated structures of the form (174) count as adjunction structures. 10. Possibly, a verb does not properly govern its α-marked complement.

All this is unfortunate. People, especially those who hold forth on the nature of mind, should be just plain curious about the code that the human species uses to speak and understand. In return, the scholars who study language for a living should see that such curiosity can be satisfied. Chomsky's theory need not be treated by either group as a set of cabalistic incantations that only the initiated can mutter. It is a set of discoveries about the design of language that can be appreciated intuitively if one first understands the problems to which the theory provides solutions. In fact, grasping grammatical theory provides an intellectual pleasure that is rare in the social sciences. When I entered high school in the late 1960s and electives were chosen for their "relevance," Latin underwent a steep decline in popularity (thanks to students like me, I confess). Our Latin teacher Mrs. Rillie, whose merry birthday parties for Rome failed to slow the decline, tried to persuade us that Latin grammar honed the mind with its demands for precision, logic, and consistency. (Nowadays, such arguments are more likely to come from the computer programming teachers.) Mrs. Rillie had a point, but Latin declensional paradigms are not the best way to convey the inherent beauty of grammar. The insights behind Universal Grammar are much more interesting, not only because they are more general and elegant but because they are about living minds rather than dead tongues.



Let's start with nouns and verbs. Your grammar teacher may have had you memorize some formula that equated parts of speech with kinds of meanings, like

> A NOUN's the name of any thing;
> As *school* or *garden, hoop* or *swing.*
> VERBS tell of something being done;
> To *read, count, sing, laugh, jump,* or *run.*

But as in most matters about language, she did not get it quite right. It is true that most names for persons, places, and things are nouns, but it is not true that most nouns are names for persons, places, or things. There are nouns with all kinds of meanings:

> the *destruction* of the city [an action]
> the *way* to San Jose [a path]
> *whiteness* moves downward [a quality]
> three *miles* along the path [a measurement in space]
> It takes three *hours* to solve the problem. [a measurement in time]
> Tell me the *answer.* ["what the answer is," a question]
> She is a *fool.* [a category or kind]
> a *meeting* [an event]
> the *square root* of minus two [an abstract concept]
> He finally kicked *the bucket.* [no meaning at all]

Likewise, though words for things being done, such as *count* and *jump*, are usually verbs, verbs can be other things, like mental states (*know, like*), possession (*own, have*), and abstract relations among ideas (*falsify, prove*).

Conversely, a single concept, like "being interested," can be expressed by different parts of speech:

her *interest* in fungi [noun]
Fungi are starting to *interest* her more and more. [verb]
She seems *interested* in fungi. Fungi seem *interesting* to her. [adjective]
Interestingly, the fungi grew an inch in an hour. [adverb]

A part of speech, then, is not a kind of meaning; it is a kind of token that obeys certain formal rules, like a chess piece or a poker chip. A noun, for example, is simply a word that does nouny things; it is the kind of word that comes after an article, can have an *'s* stuck onto it, and so on. There is a connection between concepts and part-of-speech categories, but it is a subtle and abstract one. When we construe an aspect of the world as something that can be identified and counted or measured and that can play a role in events, language often allows us to express that aspect as a noun, whether or not it is a physical object. For example, when we say *I have three reasons for leaving*, we are counting reasons as if they were objects (though of course we do not literally think that a reason can sit on a table or be kicked across a room). Similarly, when we construe some aspect of the world as an event or state involving several participants that affect one other, language often allows us to express that aspect as a verb. For example, when we say *The situation justified drastic measures*, we are talking about justification as if it were something the situation did, though again we know that justification is not something we can watch happening at a particular time and place. Nouns are *often* used for names of things, and verbs for something being done, but because the human mind can construe reality in a variety of ways, nouns and verbs are not limited to those uses.

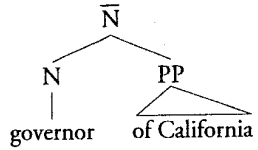


Now what about the phrases that group words into branches? One of the most intriguing discoveries of modern linguistics is that there appears to be a common anatomy in all phrases in all the world's languages.

Take the English noun phrase. A noun phrase (NP) is named after one special word, a noun, that must be inside it. The noun phrase owes most of its properties to that one noun. For example, the NP *the cat in the hat* refers to a kind of cat, not a kind of hat; the meaning of the word *cat* is the core of the meaning of the whole phrase.

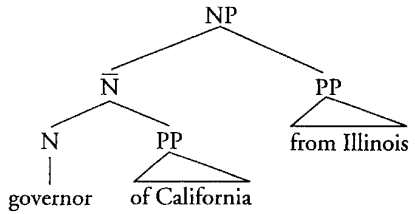
Similarly, the phrase *fox in socks* refers to a fox, not socks, and the entire phrase is singular in number (that is, we say that the fox in socks *is* or *was* here, not *are* or *were* here), because the word *fox* is singular in number. This special noun is called the "head" of the phrase, and the information filed with that word in memory "percolates up" to the topmost node, where it is interpreted as characterizing the phrase as a whole. The same goes for verb phrases: *flying to Rio before the police catch him* is an example of flying, not an example of catching, so the verb *flying* is called its head. Here we have the first principle of building the meaning of a phrase out of the meaning of the words inside the phrase. What the entire phrase is "about" is what its head word is about.

The second principle allows phrases to refer not just to single things or actions in the world but to sets of players that interact with each other in a particular way, each with a specific role. For example, the sentence *Sergey gave the documents to the spy* is not just about any old act of giving. It choreographs three entities: Sergey (the giver), documents (the gift), and a spy (the recipient). These role-players are usually called "arguments," which has nothing to do with bickering; it's the term used in logic and mathematics for a participant in a relationship. A noun phrase, too, can assign roles to one or more players, as in *picture of John, governor of California*, and *sex with Dick Cavett*, each defining one role. The head and its role-players—other than the subject role, which is special—are joined together in a subphrase, smaller than an NP or a VP, that has the kind of nonmnemonic label that has made generative linguistics so uninviting, "N-bar" and "V-bar," named after the way they are written, $\overline{N}$ and $\overline{V}$:



The third ingredient of a phrase is one or more modifiers (usually called "adjuncts"). A modifier is different from a role-player. Take the phrase *the man from Illinois*. Being a man from Illinois is not like

being a governor of California. To be a governor, you have to be a governor of something; the Californianess plays a role in what it means for someone to be governor of California. In contrast, *from Illinois* is just a bit of information that we add on to help identify which man we are talking about; being from one state or another is not an inherent part of what it means to be a man. This distinction in meaning between role-players and modifiers ("arguments" and "adjuncts," in lingo) dictates the geometry of the phrase phrase structure tree. The role-player stays next to the head noun inside the N-bar, but the modifier goes upstairs, though still inside the NP house:



This restriction on the geometry of phrase structure trees is not just playing with notation; it is a hypothesis about how the rules of language are set up in our brains, governing the way we talk. It dictates that if a phrase contains both a role-player *and* a modifier, the role-player has to be closer to the head than the modifier is—there's no way the modifier could get between the head noun and the role-player without crossing branches in the tree (that is, sticking extraneous words in among the bits of the N-bar), which is illegal. Consider Ronald Reagan. He used to be the governor of California, but he was born in Tampico, Illinois. When he was in office, he could have been referred to as *the governor of California from Illinois* (role-player, then modifier). It would have sounded odd to refer to him as *the governor from Illinois of California* (modifier, then role-player). More pointedly, in 1964 Robert F. Kennedy's senatorial ambitions ran up against the inconvenient fact that both Massachusetts seats were already occupied (one by his younger brother Edward). So he simply took up residence in New York and ran for the U.S. Senate from there, soon becoming *the senator from New York from Massachusetts*. Not *the senator from Massachusetts from New York*—though that does

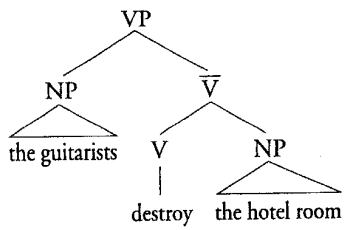
come close to the joke that Bay Staters used to tell at the time, that they lived in the only state entitled to *three* senators.

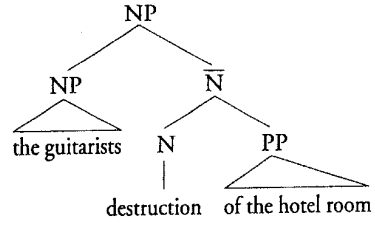
Interestingly, what is true of N-bars and noun phrases is true of V-bars and verb phrases. Say that Sergey gave those documents to the spy in a hotel. The phrase *to the spy* is one of the role-players of the verb *give*—there is no such thing as giving without a getter. Therefore *to the spy* lives with the head verb inside the V-bar. But *in a hotel* is a modifier, a comment, an afterthought, and is kept outside the V-bar, in the VP. Thus the phrases are inherently ordered: we can say *gave the documents to the spy in a hotel*, but not *gave in a hotel the documents to the spy*. When a head is accompanied by just one phrase, however, that phrase can be either a role-player (inside the V-bar) or a modifier (outside the V-bar but inside the VP), and the actual order of the words is the same. Consider the following newspaper report:

> One witness told the commissioners that she had seen sexual intercourse taking place between two parked cars in front of her house.

The aggrieved woman had a modifier interpretation in mind for *between two parked cars*, but twisted readers give it a role-player interpretation.

The fourth and final component of a phrase is a special position reserved for subjects (which linguists call "SPEC," pronounced "speck," short for "specifier"; don't ask). The subject is a special role-player, usually the causal agent if there is one. For example, in the verb phrase *the guitarists destroy the hotel room*, the phrase *the guitarists* is the subject; it is the causal agent of the event consisting of the hotel room being destroyed. Actually, noun phrases can have subjects too, as in the parallel NP *the guitarists' destruction of the hotel room*. Here, then, is the full anatomy of a VP and of an NP:





Now the story begins to get interesting. You must have noticed that noun phrases and verb phrases have a lot in common: (1) a head, which gives the phrase its name and determines what it is about, (2) some role-players, which are grouped with the head inside a subphrase (the N-bar or V-bar), (3) modifiers, which appear outside the N-or V-bar, and (4) a subject. The orderings inside a noun phrase and inside a verb phrase are the same: the noun comes before its role-players (*the destruction of the hotel room*, not *the of the hotel room destruction*), and the verb comes before *its* role-players (*to destroy the hotel room*, not *to the hotel room destroy*). The modifiers go to the right in both cases, the subject to the left. It seems as if there is a standard design to the two phrases.

In fact, the design pops up all over the place. Take, for example, the prepositional phrase (PP) *in the hotel*. It has a head, the preposition *in*, which means something like "interior region," and then a role, the thing whose interior region is being picked out, in this case a hotel. And the same goes for the adjective phrase (AP): in *afraid of the wolf*, the head adjective, *afraid*, occurs before its role-player, the source of the fear.

With this common design, there is no need to write out a long list of rules to capture what is inside a speaker's head. There may be just one pair of super-rules for the entire language, where the distinctions among nouns, verbs, prepositions, and adjectives are collapsed and all four are specified with a variable like "X." Since a phrase just inherits the properties of its head (*a tall man* is a kind of *man*), it's redundant to call a phrase headed by a noun a "noun phrase"—we could just call it an "X phrase," since the nounhood of the head noun, like the manhood of the head noun and all the other information in the head noun, percolates up to characterize the whole phrase. Here is what the super-rules look like (as before, focus on the summary of the rule, not the rule itself):

$XP \rightarrow (SPEC)\ \overline{X}\ YP^*$

"A phrase consists of an optional subject, followed by an X-bar, followed by any number of modifiers."

$\overline{X} \rightarrow X\ ZP^*$

"An X-bar consists of a head word, followed by any number of role-players."

Just plug in noun, verb, adjective, or preposition for X, Y, and Z, and you have the actual phrase structure rules that spell the phrases. This streamlined version of phrase structure is called "the X-bar theory."

This general blueprint for phrases extends even farther, to other languages. In English, the head of a phrase comes before its role-players. In many languages, it is the other way around—but it is the other way around across the board, across all the kinds of phrases in the language. For example, in Japanese, the verb comes *after* its object, not before: they say *Kenji sushi ate*, not *Kenji ate sushi*. The preposition comes after its noun phrase: *Kenji to*, not *to Kenji* (so they are actually called "postpositions"). The adjective comes after its complement: *Kenji than taller*, not *taller than Kenji*. Even the words marking questions are flipped: they say, roughly, *Kenji eat did?*, not *Did Kenji eat?* Japanese and English are looking-glass versions of each other. And such consistency has been found in scores of languages: if a language has the verb before the object, as in English, it will also have prepositions; if it has the verb after the object, as in Japanese, it will have postpositions.

This is a remarkable discovery. It means that the super-rules suffice not only for all phrases in English but for all phrases in all languages, with one modification: removing the left-to-right order from each super-rule. The trees becomes mobiles. One of the rules would say:

$\overline{X} \rightarrow \{ZP^*, X\}$

"An X-bar is composed of a head X and any number of role-players, in either order."

To get English, one appends a single bit of information saying that the order within an X-bar is "head-first." To get Japanese, that bit of information would say that the order is "head-last." Similarly, the other super-rule (the one for phrases) can be distilled so that left-to-right order boils away, and an ordered phrase in a particular language can be reconstituted by adding back either "X-bar-first" or "X-bar-last." The piece of information that makes one language different from another is called a parameter.

In fact, the super-rule is beginning to look less like an exact blueprint for a particular phrase and more like a general guideline or

principle for what phrases must look like. The principle is usable only after you combine it with a language's particular setting for the order parameter. This general conception of grammar, first proposed by Chomsky, is called the "principles and parameters" theory.

Chomsky suggests that the unordered super-rules (principles) are universal and innate, and that when children learn a particular language, they do not have to learn a long list of rules, because they were born knowing the super-rules. All they have to learn is whether their particular language has the parameter value head-first, as in English, or head-last, as in Japanese. They can do that merely by noticing whether a verb comes before or after its object in any sentence in their parents' speech. If the verb comes before the object, as in *Eat your spinach!,* the child concludes that the language is head-first; if it comes after, as in *Your spinach eat!,* the child concludes that the language is head-last. Huge chunks of grammar are then available to the child, all at once, as if the child were merely flipping a switch to one of two possible positions. If this theory of language learning is true, it would help solve the mystery of how children's grammar explodes into adultlike complexity in so short a time. They are not acquiring dozens or hundreds of rules; they are just setting a few mental switches.



The principles and parameters of phrase structure specify only what kinds of ingredients may go into a phrase in what order. They do not spell out any particular phrase. Left to themselves, they would run amok and produce all kinds of mischief. Take a look at the following sentences, which all conform to the principles or super-rules. The ones I have marked with an asterisk do not sound right.

Melvin dined.
*Melvin dined the pizza.

Melvin devoured the pizza.
*Melvin devoured.

*Melvin put the car in the garage.
*Melvin put.

*Melvin put the car.
*Melvin put in the garage.

Sheila alleged that Bill is a liar.
*Sheila alleged the claim.
*Sheila alleged.

It must be the verb's fault. Some verbs, like *dine,* refuse to appear in the company of a direct object noun phrase. Others, like *devour,* won't appear without one. This is true even though *dine* and *devour* are very close in meaning, both being ways of eating. You may dimly recall from grammar lessons that verbs like *dine* are called "intransitive" and verbs like *devour* are called "transitive." But verbs come in many flavors, not just these two. The verb *put* is not content unless it has both an object NP (*the car*) and a prepositional phrase (*in the garage*). The verb *allege* requires an embedded sentence (*that Bill is a liar*) and nothing else.

Within a phrase, then, the verb is a little despot, dictating which of the slots made available by the super-rules are to be filled. These demands are stored in the verb's entry in the mental dictionary, more or less as follows:

dine:
verb
means "to eat a meal in a refined setting"
eater = subject

devour:
verb
means "to eat something ravenously"
eater = subject
thing eaten = object

put:
verb
means "to cause something to go to some place"
putter = subject
thing put = object
place = prepositional object

allege:
verb
means "to declare without proof"
declarer = subject
declaration = complement sentence

Each of these entries lists a definition (in mentalese) of some kind of event, followed by the players that have roles in the event. The entry indicates how each role-player may be plugged into the sentence—as a subject, an object, a prepositional object, an embedded sentence, and so on. For a sentence to feel grammatical, the verb's demands must be satisfied. *Melvin devoured* is bad because *devour*'s desire for a "thing eaten" role is left unfulfilled. *Melvin dined the pizza* is bad because *dine* didn't order *pizza* or any other object.

Because verbs have the power to dictate how a sentence conveys who did what to whom, one cannot sort out the roles in a sentence without looking up the verb. That is why your grammar teacher got it wrong when she told you that the subject of the sentence is the "doer of the action." The subject of the sentence is often the doer, but only when the verb says so; the verb can also assign it other roles:

The big bad wolf *frightened* the three little pigs. [The subject is doing the frightening.]
The three little pigs *feared* the big bad wolf. [The subject is being frightened.]
My true love *gave* me a partridge in a pear tree. [The subject is doing the giving.]
I *received* a partridge in a pear tree from my true love. [The subject is being given to.]
Dr. Nussbaum *performed* plastic surgery. [The subject is operating on someone.]
Cheryl *underwent* plastic surgery. [The subject is being operated on.]

In fact, many verbs have two distinct entries, each casting a different set of roles. This can give rise to a common kind of ambiguity, as in the old joke: "Call me a taxi." "OK, you're a taxi." In one of the Harlem Globetrotters' routines, the referee tells Meadowlark Lemon

to shoot the ball. Lemon points his finger at the ball and shouts, "Bang!" The comedian Dick Gregory tells of walking up to a lunch counter in Mississippi during the days of racial segregation. The waitress said to him, "We don't serve colored people." "That's fine," he replied, "I don't eat colored people. I'd like a piece of chicken."



So how do we actually distinguish *Man bites dog* from *Dog bites man*? The dictionary entry for *bite* says "The biter is the subject; the bitten thing is the object." But how do we *find* subjects and objects in the tree? Grammar puts little tags on the noun phrases that can be matched up with the roles laid out in a verb's dictionary entry. These tags are called *cases*. In many languages, cases appear as prefixes or suffixes on the nouns. For example, in Latin, the nouns for man and dog, *homo* and *canis*, change their endings depending on who is biting whom:

Canis hominem mordet. [not news]
Homo canem mordet. [news]

Julius Caesar knew who bit whom because the noun corresponding to the bitee appeared with *-em* at the end. Indeed, this allowed Caesar to find the biter and bitee even when the order of the two was flipped, which Latin allows: *Hominem canis mordet* means the same thing as *Canis hominem mordet*, and *Canem homo mordet* means the same thing as *Homo canem mordet*. Thanks to case markers, verbs' dictionary entries can be relieved of the duty of keeping track of where their role-players actually appear in the sentence. A verb need only indicate that, say, the doer is a subject; whether the subject is in first or third or fourth position in the sentence is up to the rest of the grammar, and the interpretation is the same. Indeed, in what are called "scrambling" languages, case markers are exploited even further: the article, adjective, and noun inside a phrase are each tagged with a particular case marker, and the speaker can scramble the words of the phrase all over the sentence (say, put the adjective at the end for emphasis), knowing that the listener can mentally join them back up. This process, called agreement or concord, is a second engineering solution (aside from phrase structure itself) to the problem of encoding a

tangle of interconnected thoughts into strings of words that appear one after the other.

Centuries ago, English, like Latin, had suffixes that marked case overtly. But the suffixes have all eroded, and overt case survives only in the personal pronouns—*I, he, she, we, they* are used for the subject role; *my, his, her, our, their* are used for the possessor role; *me, him, her, us, them* are used for all other roles. (The *who/whom* distinction could be added to this list, but it is on the way out; in the United States, *whom* is used consistently only by careful writers and pretentious speakers.) Interestingly, since we all know to say *He saw us* but never *Him saw we*, the syntax of case must still be alive and well in English. Though nouns appear physically unchanged no matter what role they play, they are tagged with silent cases. Alice realized this after spotting a mouse swimming nearby in her pool of tears:

> "Would it be of any use, now," thought Alice, "to speak to this mouse? Everything is so out-of-the-way down here, that I should think very likely it can talk: at any rate, there's no harm in trying." So she began. "O Mouse, do you know the way out of this pool? I am very tired of swimming about here, O Mouse!" (Alice thought this must be the right way of speaking to a mouse: she had never done such a thing before, but she remembered having seen, in her brother's Latin Grammar, "A Mouse—of a mouse—to a mouse—a mouse—O mouse!")

English speakers tag a noun phrase with a case by seeing what the noun is adjacent to, generally a verb or preposition (but for Alice's mouse, the archaic "vocative" case marker *O*). They use these case tags to match up each noun phrase with its verb-decreed role.

The requirement that noun phrases must get case tags explains why certain sentences are impossible even though the super-rules admit them. For example, a direct object role-player has to come right after the verb, before any other role-player: one says *Tell Mary that John is coming*, not *Tell that John is coming Mary*. The reason is that the NP *Mary* cannot just float around tagless but must be case-marked, by sitting adjacent to the verb. Curiously, while verbs and prepositions can mark case on their adjacent NP's, nouns and adjectives cannot: *governor California* and *afraid the wolf*, though interpretable, are ungrammatical. English demands that the meaningless preposition *of* precede the noun, as in *governor of California* and *afraid of the wolf*, for no reason other than to give it a case tag. The sentences we utter are kept under tight rein by verbs and prepositions—phrases cannot just show up anywhere they feel like in the VP but must have a job description and be wearing an identity badge at all times. Thus we cannot say things like *Last night I slept bad dreams a hangover snoring no pajamas sheets were wrinkled*, even though a listener could guess what that would mean. This marks a major difference between human languages and, for example, pidgins and the signing of chimpanzees, where any word can pretty much go anywhere.

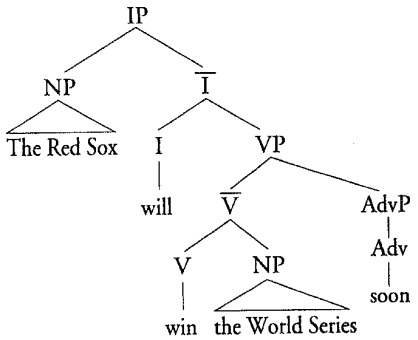


Now, what about the most important phrase of all, the sentence? If a noun phrase is a phrase built around a noun, and a verb phrase is a phrase built around a verb, what is a sentence built around?

The critic Mary McCarthy once said of her rival Lillian Hellman, "Every word she writes is a lie, including 'and' and 'the.' " The insult relies on the fact that a sentence is the smallest thing that can be either true or false; a single word cannot be either (so McCarthy is alleging that Hellman's lying extends deeper than one would have thought possible). A sentence, then, must express some kind of meaning that does not clearly reside in its nouns and verbs but that embraces the entire combination and turns it into a proposition that can be true or false. Take, for example, the optimistic sentence *The Red Sox will win the World Series*. The word *will* does not apply to the Red Sox alone, nor to the World Series alone, nor to winning alone; it applies to an entire concept, the-Red-Sox-winning-the-World-Series. That concept is timeless and therefore truthless. It can refer equally well to some past glory, a hypothetical future one, even to the mere logical possibility, bereft of any hope that it will ever happen. But the word *will* pins the concept down to temporal coordinates, namely the stretch of time subsequent to the moment the sentence is uttered. If I declare "The Red Sox will win the World Series," I can be right or wrong (probably wrong, alas).

The word *will* is an example of an auxiliary, a word that expresses layers of meaning having to do with the truth of a proposition as the speaker conceives it. These layers also include negation (as in *won't* and *doesn't*), necessity (*must*), and possibility (*might* and *can*). Auxiliaries typically occur at the periphery of sentence trees, mirroring the

fact that they assert something about the rest of the sentence taken as a whole. The auxiliary is the head of the sentence in exactly the same way that a noun is the head of the noun phrase. Since the auxiliary is also called INFL (for "inflection"), we can call the sentence an IP (an INFL phrase or auxiliary phrase). Its subject position is reserved for the subject of the entire sentence, reflecting the fact that a sentence is an assertion that some predicate (the VP) is true of its subject. Here, more or less, is what a sentence looks like in the current version of Chomsky's theory:



An auxiliary is an example of a "function word," a different kind of word from nouns, verbs, and adjectives, the "content" words. Function words include articles (*the, a, some*), pronouns (*he, she*), the possessive marker *'s*, meaningless prepositions like *of*, words that introduce complements like *that* and *to*, and conjunctions like *and* and *or*. Function words are bits of crystallized grammar; they delineate larger phrases into which NP's and VP's and AP's fit, thereby providing a scaffolding for the sentence. Accordingly, the mind treats function words differently from content words. People add new content words to the language all the time (like the noun *fax*, and the verb *to snarf*, meaning to retrieve a computer file), but the function words form a closed club that resists new members. That is why all the attempts to introduce gender-neutral pronouns like *hesh* and *thon* have failed. Recall, too, that patients with damage to the language areas of the brain have more trouble with function words like *or* and *be* than with content words like *oar* and *bee*. When words are expensive, as in telegrams and headlines, writers tend to leave the function words out, hoping that the reader can reconstruct them from the order of the content words. But because function words are the most reliable clues to the phrase structure of the sentence, telegraphic language is always a gamble. A reporter once sent Cary Grant the telegram, "How old Cary Grant?" He replied, "Old Cary Grant fine." Here are some headlines from a collection called *Squad Helps Dog Bite Victim*, put together by the staff of the *Columbia Journalism Review*:

New Housing for Elderly Not Yet Dead
New Missouri U. Chancellor Expects Little Sex
12 on Their Way to Cruise Among Dead in Plane Crash
N.J. Judge to Rule on Nude Beach
Chou Remains Cremated
Chinese Apeman Dated
Hershey Bars Protest
Reagan Wins on Budget, But More Lies Ahead.
Deer Kill 130,000
Complaints About NBA Referees Growing Ugly

Function words also capture much of what makes one language grammatically different from another. Though all languages have function words, the properties of the words differ in ways that can have large effects on the structure of the sentences in the language. We have already seen one example: overt case and agreement markers in Latin allow noun phrases to be scrambled; silent ones in English force them to remain in place. Function words capture the grammatical look and feel of a language, as in these passages that use a language's function words but none of its content words:

DER JAMMERWOCH

Es brillig war. Die schlichte Toven
Wirrten und wimmelten in Waben.

LE JASEROQUE

Il brilgue: les tôves lubricilleux
Se gyrent en vrillant dans la guave.

The effect can also be seen in passages that take the function words from one language but the content words from another, like the following pseudo-German notice that used to be posted in many university computing centers in the English-speaking world:

ACHTUNG! ALLES LOOKENSPEEPERS!

Das computermachine ist nicht fuer gefingerpoken und mittengrabben. Ist easy schnappen der springenwerk, blowenfusen und poppencorken mit spitzensparken. Ist nicht fuer gewerken bei das dumpkopfen. Das rubbernecken sightseeren keepen das cottenpickenen hans in das pockets muss; relaxen und watchen das blinkenlichten.

Turnabout being fair play, computer operators in Germany have posted a translation into pseudo-English:

ATTENTION

This room is fulfilled mit special electronische equippment. Fingergrabbing and pressing the cnoeppkes from the computers is allowed for die experts only! So all the "lefthanders" stay away and do not disturben the brainstorming von here working intelligencies. Otherwise you will be out thrown and kicked andeswhere! Also: please keep still and only watchen astaunished the blinkenlights.



Anyone who goes to cocktail parties knows that one of Chomsky's main contributions to intellectual life is the concept of "deep structure," together with the "transformations" that map it onto "surface structure." When Chomsky introduced the terms in the behaviorist climate of the early 1960s, the reaction was sensational. Deep structure came to refer to everything that was hidden, profound, universal, or meaningful, and before long there was talk of the deep structure of visual perception, stories, myths, poems, paintings, musical compositions, and so on. Anticlimactically, I must now divulge that "deep structure" is a prosaic technical gadget in grammatical theory. It is not the meaning of a sentence, nor is it what is universal across all human languages. Though universal grammar and abstract phrase structures seem to be permanent features of grammatical theory, many

linguists—including, in his most recent writings, Chomsky himself—think one can do without deep structure per se. To discourage all the hype incited by the word "deep," linguists now usually refer to it as "d-structure." The concept is actually quite simple.

Recall that for a sentence to be well formed, the verb must get what it wants: all the roles listed in the verb's dictionary entry must appear in their designated positions. But in many sentences, the verb does not seem to be getting what it wants. Remember that *put* requires a subject, an object, and a prepositional phrase; *He put the car* and *He put in the garage* sound incomplete. How, then, do we account for the following perfectly good sentences?

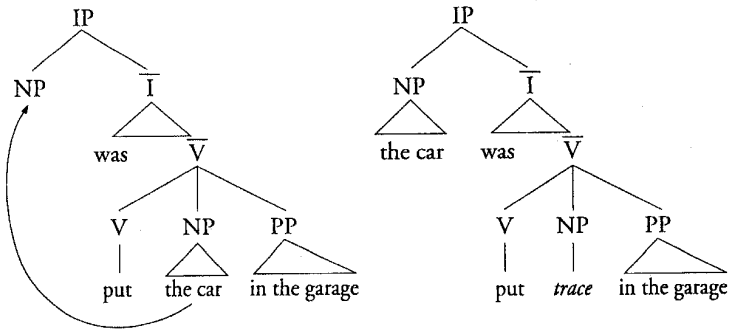
The car was put in the garage.
What did he put in the garage?
Where did he put the car?

In the first sentence, *put* seems to be doing fine without an object, which is out of character. Indeed, now it rejects one: *The car was put the Toyota in the garage* is awful. In the second sentence, *put* also appears in public objectless. In the third, its obligatory prepositional phrase is missing. Does this mean we need to add new dictionary entries for *put*, allowing it to appear in some places without its object or its prepositional phrase? Obviously not, or *He put the car* and *He put in the garage* would slip back in.

In some sense, of course, the required phrases really are there—they're just not where we expect them. In the first sentence, a passive construction, the NP *the car*, playing the role of "thing put" which ordinarily would be the object, shows up in the subject position instead. In the second sentence, a *wh*-question (that is, a question formed with *who, what, where, when*, or *why*), the "thing put" role is expressed by the word *what* and shows up at the beginning. In the third sentence, the "place" role also shows up at the beginning instead of after the object, where it ordinarily belongs.

A simple way to account for the entire pattern is to say that every sentence has two phrase structures. The phrase structure we have been talking about so far, the one defined by the super-rules, is the deep structure. Deep structure is the interface between the mental dictionary and phrase structure. In the deep structure, all the role-players for *put* appear in their expected places. Then a transforma-

tional operation can "move" a phrase to a previously unfilled slot elsewhere in the tree. That is where we find the phrase in the actual sentence. This new tree is the surface structure (now called "s-structure," because as a mere "surface" representation it never used to get proper respect). Here are the deep structure and surface structure of a passive sentence:



In the deep structure on the left, *the car* is where the verb wanted it; in the surface structure on the right, it is where we actually hear it. In the surface structure, the position from which the phrase was moved contains an inaudible symbol that was left behind by the movement transformation, called a "trace." The trace serves as a reminder of the role that the moved phrase is playing. It tells us that to find out what *the car* is doing in the putting event, we should look up the "object" slot in the entry for the verb *put*; that slot says "thing put." Thanks to the trace, the surface structure contains the information needed to recover the meaning of the sentence; the original deep structure, which was used only to plug in the right sets of words from the lexicon, plays no role.

Why do languages bother with separate deep structures and surface structures? Because it takes more than just keeping the verb happy—what deep structure does—to have a usable sentence. A given concept often has to play one kind of role, defined by the verb in the verb phrase, and simultaneously a separate role, independent of the verb, defined by some other layer of the tree. Consider the difference between *Beavers build dams* and its passive, *Dams are built by beavers*.

Down in the verb phrase—the level of who did what to whom—the nouns are playing the same roles in both sentences. Beavers do the building, dams get built. But up at the sentence (IP) level—the level of subject-predicate relations, of what is being asserted to be true of what—they are playing different roles. The active sentence is saying something about beavers in general, and happens to be true; the passive sentence is saying something about dams in general, and happens to be false (since some dams, like the Grand Coulee Dam, are not built by beavers). The surface structure, which puts *dams* in the sentence's subject position but links it to a trace of its original verb phrase position, allows the cake to be both eaten and had.

The ability to move phrases around while still retaining their roles also gives the speaker of a rigid-word-order language like English a bit of wiggle room. For example, phrases that are ordinarily buried deep in the tree can be moved to early in the sentence, where they can hook up with material fresh in the listener's mind. For example, if a play-by-play announcer has been describing Nevin Markwart's progression down the ice, he could say *Markwart spears Gretzky!!!* But if it was Wayne Gretzky the announcer had been describing, he would say *Gretzky is speared by Markwart!!!!* Moreover, because a passive participle has the option of leaving the doer role, ordinarily the subject, unfilled in deep structure, it is useful when one wants to avoid mentioning that role altogether, as in Ronald Reagan's evasive concession *Mistakes were made.*

Hooking up players with different roles in different scenarios is something that grammar excels at. In a *wh*-question like

What did he put [*trace*] in the garage?

the noun phrase *what* gets to live a double life. Down in the who-did-what-to-whom realm of the verb phrase, the position of the trace indicates that the entity has the role of the thing being put; up in the what-is-being-asserted-of-what realm of the sentence, the word *what* indicates that the point of the sentence is to ask the listener to provide the identity of something. If a logician were to express the meaning behind the sentence, it would be something like "For which *x*, John put *x* in the garage." When these movement operations are combined with other components of syntax, as in *She was told by Bob to be examined by a doctor* or *Who did he say that Barry tried to convince*

to leave? or *Tex is fun for anyone to tease*, the components interact to determine the meaning of the sentence in chains of deduction as intricate and precise as the workings of a fine Swiss watch.



Now that I have dissected syntax in front of you, I hope your reaction is more favorable than Eliza Doolittle's or Jack Cade's. At the very least I hope you are impressed at how syntax is a Darwinian "organ of extreme perfection and complication." Syntax is complex, but the complexity is there for a reason. For our thoughts are surely even more complex, and we are limited by a mouth that can pronounce a single word at a time. Science has begun to crack the beautifully designed code that our brains use to convey complex thoughts as words and their orderings.

The workings of syntax are important for another reason. Grammar offers a clear refutation of the empiricist doctrine that there is nothing in the mind that was not first in the senses. Traces, cases, X-bars, and the other paraphernalia of syntax are colorless, odorless, and tasteless, but they, or something like them, must be a part of our unconscious mental life. This should not be surprising to a thoughtful computer scientist. There is no way one can write a halfway intelligent program without defining variables and data structures that do not directly correspond to anything in the input or output. For example, a graphics program that had to store an image of a triangle inside a circle would not store the actual keystrokes that the user typed to draw the shapes, because the same shapes could have been drawn in a different order or with a different device like a mouse or a light pen. Nor would it store the list of dots that have to be lit up to display the shapes on a video screen, because the user might later want to move the circle around and leave the triangle in place, or make the circle bigger or smaller, and one long list of dots would not allow the program to know which dots belong to the circle and which to the triangle. Instead, the shapes would be stored in some more abstract format (like the coordinates of a few defining points for each shape), a format that mirrors neither the inputs nor the outputs to the program but that can be translated to and from them when the need arises.

Grammar, a form of mental software, must have evolved under similar design specifications. Though psychologists under the influence of empiricism often suggest that grammar mirrors commands to the speech muscles, melodies in speech sounds, or mental scripts for the ways that people and things tend to interact, I think all these suggestions miss the mark. Grammar is a protocol that has to interconnect the ear, the mouth, and the mind, three very different kinds of machine. It cannot be tailored to any of them but must have an abstract logic of its own.

The idea that the human mind is designed to use abstract variables and data structures used to be, and in some circles still is, a shocking and revolutionary claim, because the structures have no direct counterpart in the child's experience. Some of the organization of grammar would have to be there from the start, part of the language-learning mechanism that allows children to make sense out of the noises they hear from their parents. The details of syntax have figured prominently in the history of psychology, because they are a case where complexity in the mind is not caused by learning; learning is caused by complexity in the mind. And that was real news.

7

Talking Heads

For centuries, people have been terrified that their programmed creations might outsmart them, overpower them, or put them out of work. The fear has long been played out in fiction, from the medieval Jewish legend of the Golem, a clay automaton animated by an inscription of the name of God placed in its mouth, to HAL, the mutinous computer of *2001: A Space Odyssey*. But when the branch of engineering called "artificial intelligence" (AI) was born in the 1950s, it looked as though fiction was about to turn into frightening fact. It is easy to accept a computer calculating pi to a million decimal places or keeping track of a company's payroll, but suddenly computers were also proving theorems in logic and playing respectable chess. In the years following there came computers that could beat anyone but a grand master, and programs that outperformed most experts at recommending treatments for bacterial infections and investing pension funds. With computers solving such brainy tasks, it seemed only a matter of time before a C3PO or a Terminator would be available from the mail-order catalogues; only the easy tasks remained to be programmed. According to legend, in the 1970s Marvin Minsky, one of the founders of AI, assigned "vision" to a graduate student as a summer project.

But household robots are still confined to science fiction. The main lesson of thirty-five years of AI research is that the hard problems are easy and the easy problems are hard. The mental abilities of a four-year-old that we take for granted—recognizing a face, lifting a pencil, walking across a room, answering a question—in fact solve some of

the hardest engineering problems ever conceived. Do not be fooled by the assembly-line robots in the automobile commercials; all they do is weld and spray-paint, tasks that do not require these clumsy Mr. Magoos to see or hold or place anything. And if you want to stump an artificial intelligence system, ask it questions like, Which is bigger, Chicago or a breadbox? Do zebras wear underwear? Is the floor likely to rise up and bite you? If Susan goes to the store, does her head go with her? Most fears of automation are misplaced. As the new generation of intelligent devices appears, it will be the stock analysts and petrochemical engineers and parole board members who are in danger of being replaced by machines. The gardeners, receptionists, and cooks are secure in their jobs for decades to come.

Understanding a sentence is one of these hard easy problems. To interact with computers we still have to learn their languages; they are not smart enough to learn ours. In fact, it is all too easy to give computers more credit at understanding than they deserve.

Recently an annual competition was set up for the computer program that can best fool users into thinking that they are conversing with another human. The competition for the Loebner Prize was intended to implement a suggestion made by Alan Turing in a famous 1950 paper. He suggested that the philosophical question "Can machines think?" could best be answered in an imitation game, where a judge converses with a person over one terminal and with a computer programmed to imitate a person on another. If the judge cannot guess which is which, Turing suggested, there is no basis for denying that the computer can think. Philosophical questions aside, it was apparent to the committee charged with overseeing the competition that no program could come close to winning the $100,000 prize, so they devised a $1,500 version that would be fairer to the state of the art. Each of the judges had to stick to a single topic of conversation selected by the programmer or by the human foil, whichever it was, and the judge was not allowed to engage in any "trickery or guile" such as repeating a question ten times or asking whether zebras wear underwear; the conversation had to be "natural." After interacting with several programs and human foils for about seven minutes apiece, the judges ranked all the humans as more humanlike than any of the computers. About half the judges did, however, misidentify the winning program as human.

The accomplishment is less impressive than it sounds. The rules

handcuffed the judges: "unnatural trickery or guile" is another way of referring to any attempt to determine whether one is conversing with a human or a machine, which is the whole point of the test! Also, the winning programmer shrewdly exploited the opportunity to designate the topic of conversation for his program. He chose "whimsical conversation," which is a dubious example of a "topic," and which, by definition, can be full of non sequiturs:

Judge: It has been a long day! Now I think I'm ready for some fun. The weather isn't very pleasant today. Do you have any jokes that could lift the clouds?
Computer: Everybody talks about the weather but nobody seems to do much about it. Why do you need to tell me now you think you're ready for some fun? Groannnnnn . . .
Judge: You're trying to make this difficult for me aren't you? I just explained that I was ready for fun!
Computer: Would you prefer it if it was not true that I are trying to make this difficult for you aren't you? I just explained that I was ready for fun! I am, therefore I compute . . . What makes you think I am trying to make this difficult for you aren't I?
Judge: I can't say I understand the question . . . but the whimsy of a poet is one thing and the whimsy of most of us is another.
Computer: How do you know you can't say you understand the question . . . but the whimsy of a poet is one thing and the whimsy of most of us is another.

In fact the programmer had simply updated a technique invented in 1965 by the computer scientist Joseph Weizenbaum. Weizenbaum wrote a rather stupid program called ELIZA, named after the Shaw heroine, which looked for a few key words in an input line, retrieved a canned response from a list, and inserted fragments of the input line into its response (with some minor substitutions, like "you are" for "I am"). It could thus crudely simulate a conversation. One version, called Doctor, mimicked a Rogerian psychotherapist, taking advantage of Carl Rogers' philosophy that a therapist should recast patients' statements back to them for their contemplation. A version of the program was later sold as a computer game in Radio Shack stores, and another comes packaged with my word processing program, GNU Emacs. (The documentation says, "If using Emacs becomes terribly frustrating and none of the techniques described above solve the problem, Emacs can still help you. Type 'M-x doctor RET.' ") Here is a typical dialogue:

User: I have been rather depressed lately.
Computer: Are you depressed often?
User: Yes, almost all the time.
Computer: Why do you say yes almost all the time?
User: Right now I feel like I could cry.
Computer: Is it because right now you feel like you could cry that you came to me?

To Weizenbaum's horror, one day he found his secretary transfixed at the terminal, pouring her heart out to the program.

Most computer scientists are annoyed by the Loebner Prize competition. They consider it a pointless publicity stunt, because it is an exercise in how to fool an amateur, not how to get computers to use language. (Artificial intelligence researchers and other professionals who are knowledgeable about language were not allowed to act as judges, and none bothered to compete; the submissions were from hobbyists.) It is about as productive as promoting biology by offering a prize to the designer of the most convincing silk flower, or running a space program by simulating a moon landing on a Hollywood back lot. There has been intensive research on computer language-understanding systems, but no serious engineer has the hubris to predict that the systems will duplicate the human ability anytime soon.

In fact, from a scientist's perspective, people have no right to be as good at sentence understanding as they are. Not only can they solve a viciously complex task, but they solve it *fast*. Comprehension ordinarily takes place in "real time." Listeners keep up with talkers; they do not wait for the end of a batch of speech and interpret it after a proportional delay, like a critic reviewing a book. And the lag between speaker's mouth and listener's mind is remarkably short: about a syllable or two, around half a second. Some people can understand *and* repeat sentences, shadowing a speaker as he speaks, with a lag of a quarter of a second!

Understanding understanding has practical applications other than building machines we can converse with. Human sentence com-

prehension is fast and powerful, but it is not perfect. It works when the incoming conversation or text is structured in certain ways. When it is not, the process can bog down, backtrack, and misunderstand. As we explore language understanding in this chapter, we will discover which kinds of sentences mesh with the mind of the understander. One practical benefit is a set of guidelines for clear prose, a scientific style manual, such as Joseph Williams's 1990 *Style: Toward Clarity and Grace*, which is informed by many of the findings we will examine.

Another practical application involves the law. Judges are frequently faced with guessing how a typical person is likely to understand some ambiguous passage, such as a customer scanning a contract, a jury listening to instructions, or a member of the public reading a potentially libelous characterization. Many of people's habits of interpretation have been worked out in the laboratory, and the linguist and lawyer Lawrence Solan has explained the connections between language and law in his interesting 1993 book *The Language of Judges*, to which we will return.



How do we understand a sentence? The first step is to "parse" it. This does not refer to the exercises you grudgingly did in elementary school, which Dave Barry's "Ask Mr. Language Person" remembers as follows:

> Q. Please explain how to diagram a sentence.
>
> A. First spread the sentence out on a clean, flat surface, such as an ironing board. Then, using a sharp pencil or X-Acto knife, locate the "predicate," which indicates where the action has taken place and is usually located directly behind the gills. For example, in the sentence: "LaMont never would of bit a forest ranger," the action probably took place in a forest. Thus your diagram would be shaped like a little tree with branches sticking out of it to indicate the locations of the various particles of speech, such as your gerunds, proverbs, adjutants, etc.

But it does involve a similar process of finding subject, verbs, objects, and so on, that takes place unconsciously. Unless you are Woody

Allen speed-reading *War and Peace*, you have to group words into phrases, determine which phrase is the subject of which verb, and so on. For example, to understand the sentence *The cat in the hat came back*, you have to group the words *the cat in the hat* into one phrase, to see that it is the cat that came back, not just the hat. To distinguish *Dog bites man* from *Man bites dog*, you have to find the subject and the object. And to distinguish *Man bites dog* from *Man is bitten by dog* or *Man suffers dog bite*, you have to look up the verbs' entries in the mental dictionary to determine what the subject, *man*, is doing or having done to him.

Grammar itself is a mere code or protocol, a static database specifying what kinds of sounds correspond to what kinds of meanings in a particular language. It is not a recipe or program for speaking and understanding. Speaking and understanding share a grammatical database (the language we speak is the same as the language we understand), but they also need procedures that specify what the mind should *do*, step by step, when the words start pouring in or when one is about to speak. The mental program that analyzes sentence structure during language comprehension is called the parser.

The best way to appreciate how understanding works is to trace the parsing of a simple sentence, generated by a toy grammar like the one of Chapter 4, which I repeat here:

> S → NP VP
> "A sentence can consist of a noun phrase and a verb phrase."
>
> NP → (det) N (PP)
> "A noun phrase can consist of an optional determiner, a noun, and an optional prepositional phrase."
>
> VP → V NP (PP)
> "A verb phrase can consist of a verb, a noun phrase, and an optional prepositional phrase."
>
> PP → P NP
> "A prepositional phrase can consist of a preposition and a noun phrase."
>
> N → boy, girl, dog, cat, ice cream, candy, hot dogs
> "The nouns in the mental dictionary include *boy, girl, . . .*"

V → eats, likes, bites
"The verbs in the mental dictionary include *eats, likes, bites.*"

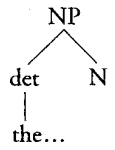
P → with, in, near
"The prepositions include *with, in, near.*"

det → a, the, one
"The determiners include *a, the, one.*"

Take the sentence *The dog likes ice cream.* The first word arriving at the mental parser is *the.* The parser looks it up in the mental dictionary, which is equivalent to finding it on the right-hand side of a rule and discovering its category on the left-hand side. It is a determiner (det). This allows the parser to grow the first twig of the tree for the sentence. (Admittedly, a tree that grows upside down from its leaves to its root is botanically improbable.)



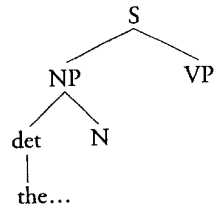
Determiners, like all words, have to be part of some larger phrase. The parser can figure out which phrase is by checking to see which rule has "det" on its right-hand side. That rule is the one defining a noun phrase, NP. More tree can be grown:



This dangling structure must be held in a kind of memory. The parser keeps in mind that the word at hand, *the*, is part of a noun phrase,

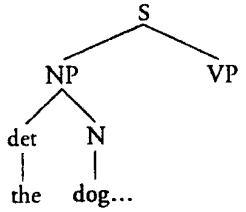
which soon must be completed by finding words that fill its other slots—in this case, at least a noun.

In the meantime, the tree continues to grow, for NP's cannot float around unattached. Having checked the right-hand sides of the rules for an NP symbol, the parser has several options. The freshly built NP could be part of a sentence, part of a verb phrase, or part of a prepositional phrase. The choice can be resolved from the root down: all words and phrases must eventually be plugged into a sentence (S), and a sentence must begin with an NP, so the sentence rule is the logical one to use to grow more of the tree:



Note that the parser is now keeping *two* incomplete branches in memory: the noun phrase, which needs an N to complete it, and the sentence, which needs a VP.

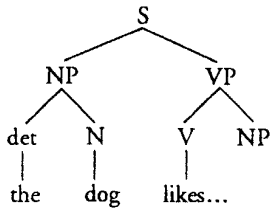
The dangling N twig is equivalent to a prediction that the next word should be a noun. When the next word, *dog*, comes in, a check against he rules confirms the prediction: *dog* is part of the N rule. This allows *dog* to be integrated into the tree, completing the noun phrase:



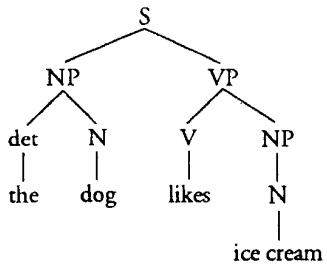
The parser no longer has to remember that there is an NP to be completed; all it has to keep in mind is the incomplete S.

At this point some of the meaning of the sentence can be inferred. Remember that the noun inside a noun phrase is a head (what the phrase is about) and that other phrases inside the noun phrase can modify the head. By looking up the definitions of *dog* and *the* in their dictionary entries, the parser can note that the phrase is referring to a previously mentioned dog.

The next word is *likes*, which is found to be a verb, V. A verb has nowhere to come from but a verb phrase, VP, which, fortunately, has already been predicted, so they can just be joined up. The verb phrase contains more than a V; it also has a noun phrase (its object). The parser therefore predicts that an NP is what should come next:



What does come next is *ice cream*, a noun, which can be part of an NP—just as the dangling NP branch predicts. The last pieces of the puzzle snap nicely together:



The word *ice cream* has completed the noun phrase, so it need not be kept in memory any longer; the NP has completed the verb phrase, so it can be forgotten, too; and the VP has completed the sentence. When memory has been emptied of all its incomplete dangling branches, we experience the mental "click" that signals that we have just heard a complete grammatical sentence.

As the parser has been joining up branches, it has been building up the meaning of the sentence, using the definitions in the mental dictionary and the principles for combining them. The verb is the head of its VP, so the VP is about liking. The NP inside the VP, *ice cream*, is the verb's object. The dictionary entry for *likes* says that its object is the liked entity; therefore the VP is about being fond of ice cream. The NP to the left of a tensed verb is the subject; the entry for *likes* says that its subject is the one doing the liking. Combining the semantics of the subject with the semantics of the VP, the parser has determined that the sentence asserts that an aforementioned canine is fond of frozen confections.



Why is it so hard to program a computer to do this? And why do people, too, suddenly find it hard to do this when reading bureaucratese and other bad writing? As we stepped our way through the sentence pretending we were the parser, we faced two computational burdens. One was memory: we had to keep track of the dangling phrases that needed particular kinds of words to complete them. The other was decision-making: when a word or phrase was found on the right-hand side of two different rules, we had to decide which to use to build the next branch of the tree. In accord with the first law of artificial intelligence, that the hard problems are easy and the easy problems are hard, it turns out that the memory part is easy for computers and hard for people, and the decision-making part is easy for people (at least when the sentence has been well constructed) and hard for computers.

A sentence parser requires many kinds of memory, but the most obvious is the one for incomplete phrases, the remembrance of things parsed. Computers must set aside a set of memory locations, usually called a "stack," for this task; this is what allows a parser to use phrase structure grammar at all, as opposed to being a word-chain device. People, too, must dedicate some of their short-term memory to dangling phrases. But short-term memory is the primary bottleneck in human information processing. Only a few items—the usual estimate is seven, plus or minus two—can be held in mind at once, and the items are immediately subject to fading or being overwritten. In the following sentences you can feel the effects of keeping a dangling phrase open in memory too long:

He gave the girl that he met in New York while visiting his parents for ten days around Christmas and New Year's the candy.

He sent the poisoned candy that he had received in the mail from one of his business rivals connected with the Mafia to the police.

She saw the matter that had caused her so much anxiety in former years when she was employed as an efficiency expert by the company through.

That many teachers are being laid off in a shortsighted attempt to balance this year's budget at the same time that the governor's cronies and bureaucratic hacks are lining their pockets is appalling.

These memory-stretching sentences are called "top-heavy" in style manuals. In languages that use case markers to signal meaning, a heavy phrase can simply be slid to the end of the sentences, so the listener can digest the beginning without having to hold the heavy phrase in mind. English is tyrannical about order, but even English provides its speakers with some alternative constructions in which the order of phrases is inverted. A considerate writer can use them to save the heaviest for last and lighten the burden on the listener. Note how much easier these sentences are to understand:

He gave the candy to the girl that he met in New York while visiting his parents for ten days around Christmas and New Year's.

He sent to the police the poisoned candy that he had received in the mail from one of his business rivals connected with the Mafia.

She saw the matter through that had caused her so much anxiety in former years when she was employed as an efficiency expert by the company.

It is appalling that teachers are being laid off in a shortsighted attempt to balance this year's budget at the same time that the governor's cronies and bureaucratic hacks are lining their pockets.

Many linguists believe that the reason that languages allow phrase movement, or choices among more-or-less synonymous constructions, is to ease the load on the listener's memory.

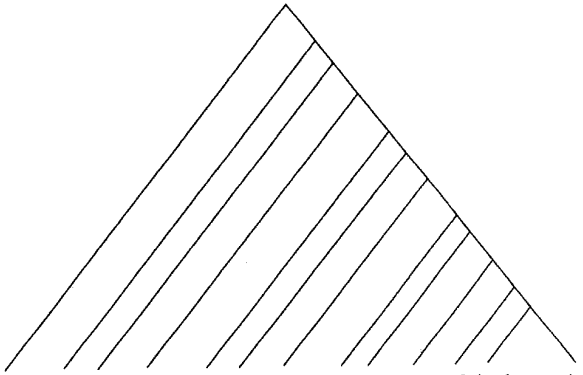
As long as the words in a sentence can be immediately grouped into complete phrases, the sentence can be quite complex but still understandable:

Remarkable is the rapidity of the motion of the wing of the hummingbird.

This is the cow with the crumpled horn that tossed the dog that worried the cat that killed the rat that ate the malt that lay in the house that Jack built.

Then came the Holy One, blessed be He, and destroyed the angel of death that slew the butcher that killed the ox that drank the water that quenched the fire that burned the stick that beat the dog that bit the cat my father bought for two zuzim.

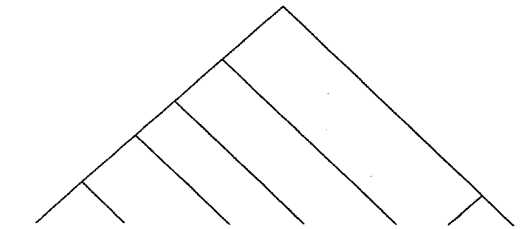
These sentences are called "right-branching," because of the geometry of their phrase structure trees. Note that as one goes from left to right, only one branch has to be left dangling at a time:



Remarkable is the rapidity of the motion of the wing of the hummingbird

Sentences can also branch to the left. Left-branching trees are most common in head-last languages like Japanese but are found in a few

constructions in English, too. As before, the parser never has to keep more than one dangling branch in mind at a time:



The hummingbird's wing's motion's rapidity is remarkable

There is a third kind of tree geometry, but it goes down far less easily. Take the sentence

The rapidity that the motion has is remarkable.

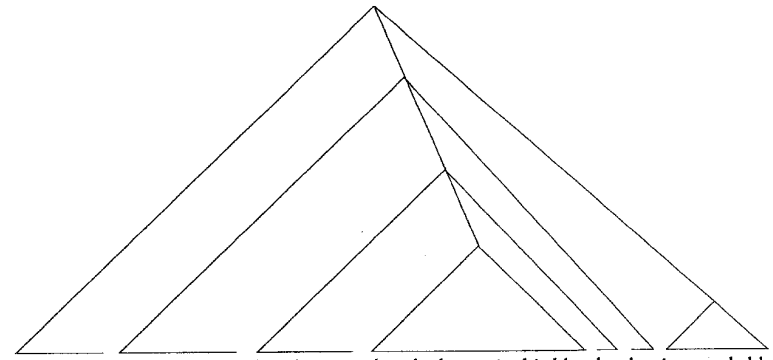
The clause *that the motion has* has been embedded in the noun phrase containing *The rapidity*. The result is a bit stilted but easy to understand. One can also say

The motion that the wing has is remarkable.

But the result of embedding the *motion that the wing has* phrase inside the *rapidity that the motion has* phrase is surprisingly hard to understand:

The rapidity that the motion that the wing has has is remarkable.

Embedding a third phrase, like *the wing that the hummingbird has*, creating a triply embedded onion sentence, results in complete unintelligibility:



The rapidity that the motion that the wing that the hummingbird has has has is remarkable

When the human parser encounters the three successive *has*'s, it thrashes ineffectively, not knowing what to do with them. But the problem is not that the phrases have to be held in memory too long; even short sentences are uninterpretable if they have multiple embeddings:

The dog the stick the fire burned beat bit the cat.
The malt that the rat that the cat killed ate lay in the house.
If if if it rains it pours I get depressed I should get help.
That that that he left is apparent is clear is obvious.

Why does human sentence understanding undergo such complete collapse when interpreting sentences that are like onions or Russian dolls? This is one of the most challenging puzzles about the design of the mental parser and the mental grammar. At first one might wonder whether the sentences are even grammatical. Perhaps we got the rules wrong, and the real rules do not even provide a way for these words to fit together. Could the maligned word-chain device of Chapter 4, which has no memory for dangling phrases, be the right model of humans after all? No way; the sentences check out perfectly. A noun phrase can contain a modifying clause; if you can say *the rat*, you can say *the rat that S*, where S is a sentence missing an object that modifies *the rat*. And a sentence like *the cat killed X* can contain a noun phrase, such as

its subject, *the cat*. So when you say *The rat that the cat killed,* you have modified a noun phrase with something that in turn contains a noun phrase. With just these two abilities, onion sentences become possible: just modify the noun phrase inside a clause with a modifying clause of its own. The only way to prevent onion sentences would be to claim that the mental grammar defines two different kinds of noun phrase, a kind that can be modified and a kind that can go inside a modifier. But that can't be right: both kinds of noun phrase would have to be allowed to contain the same twenty thousand nouns, both would have to allow articles and adjectives and possessors in identical positions, and so on. Entities should not be multiplied unnecessarily, and that is what such tinkering would do. Positing different kinds of phrases in the mental grammar just to explain why onion sentences are unintelligible would make the grammar exponentially more complicated and would give the child an exponentially larger number of rules to record when learning the language. The problem must lie elsewhere.

Onion sentences show that a grammar and a parser are different things. A person can implicitly "know" constructions that he or she can never understand, in the same way that Alice knew addition despite the Red Queen's judgment:

> "Can you do addition?" the White Queen asked. "What's one and one and one and one and one and one and one and one and one?"
> "I don't know," said Alice. "I lost count."
> "She can't do Addition," the Red Queen interrupted.

Why does the human parser seem to lose count? Is there not enough room in short-term memory to hold more than one or two dangling phrases at a time? The problem must be more subtle. Some three-layer onion sentences are a little hard because of the memory load but are not nearly as opaque as the *has has has* sentence:

> The cheese that some rats I saw were trying to eat turned out to be rancid.
> The policies that the students I know object to most strenuously are those pertaining to smoking.

> The guy who is sitting between the table that I like and the empty chair just winked.
> The woman who the janitor we just hired hit on is very pretty.

What boggles the human parser is not the amount of memory needed but the *kind* of memory: keeping a particular kind of phrase in memory, intending to get back to it, at the same time as it is analyzing another example of *that very same kind of phrase*. Examples of these "recursive" structures include a relative clause in the middle of the same kind of relative clause, or an *if . . . then* sentence inside another *if . . . then* sentence. It is as if the human sentence parser keeps track of where it is in a sentence not by writing down a list of currently incomplete phrases in the order in which they must be completed, but by writing a number in a slot next to each phrase type on a master checklist. When a type of phrase has to be remembered more than once—so that both it (*the cat that . . .*) and the identical type of phrase it is inside of (*the rat that . . .*) can be completed in order—there is not enough room on the checklist for both numbers to fit, and the phrases cannot be completed properly.



Unlike memory, which people are bad at and computers are good at, decision-making is something that people are good at and computers are bad at. I contrived the toy grammar and the baby sentence we have just walked through so that every word had a single dictionary entry (that is, was at the right-hand side of only one rule). But all you have to do is open up a dictionary, and you will see that many nouns have a secondary entry as a verb, and vice versa. For example, *dog* is listed a second time—as a verb, for sentences like *Scandals dogged the administration all year*. Similarly, in real life *hot dog* is not only a noun but a verb, meaning "to show off." And each of the verbs in the toy grammar should also be listed as nouns, because English speakers can talk of cheap *eats*, his *likes* and dislikes, and taking a few *bites*. Even the determiner *one*, as in *one dog*, can have a second life as a noun, as in *Nixon's the one*.

These local ambiguities present a parser with a bewildering number of forks at every step along the road. When it comes across, say, the word *one* at the beginning of a sentence, it cannot simply build

det
|
one

but must also keep in mind

N
|
one

Similarly, it has to jot down two rival branches when it comes across *dog*, one in case it is a noun, the other in case it is a verb. To handle *one dog*, it would need to check four possibilities: determiner-noun, determiner-verb, noun-noun, and noun-verb. Of course determiner-verb can be eliminated because no rule of grammar allows it, but it still must be checked.

It gets even worse when the words are grouped into phrases, because phrases can fit inside larger phrases in many different ways. Even in our toy grammar, a prepositional phrase (PP) can go inside either a noun phrase or a verb phrase—as in the ambiguous *discuss sex with Dick Cavett*, where the writer intended the PP *with Dick Cavett* to go inside the verb phrase (discuss it with him) but readers can interpret it as going inside the noun phrase (sex with him). These ambiguities are the rule, not the exception; there can be dozens or hundreds of possibilities to check at every point in a sentence. For example, after processing *The plastic pencil marks . . .*, the parser has to keep several options open: it can be a four-word noun phrase, as in *The plastic pencil marks were ugly*, or a three-word noun phrase plus a verb, as in *The plastic pencil marks easily*. In fact, even the first two words, *The plastic . . .*, are temporarily ambiguous: compare *The plastic rose fell* with *The plastic rose and fell.*

If it were just a matter of keeping track of all the possibilities at each point, a computer would have little trouble. It might churn away for minutes on a simple sentence, or use up so much short-term memory that the printout would spill halfway across the room, but eventually most of the possibilities at each decision point would be contradicted by later information in the sentence. If so, a single tree and its associated meaning should pop out at the end of the sentence,

as in the toy example. When the local ambiguities fail to cancel each other out and two consistent trees are found for the same sentence, we should have a sentence that people find ambiguous, like

Ingres enjoyed painting his models nude.
My son has grown another foot.
Visiting relatives can be boring.
Vegetarians don't know how good meat tastes.
I saw the man with the binoculars.

But here is the problem. Computer parsers are too meticulous for their own good. They find ambiguities that are quite legitimate, as far as English grammar is concerned, but that would never occur to a sane person. One of the first computer parsers, developed at Harvard in the 1960s, provides a famous example. The sentence *Time flies like an arrow* is surely unambiguous if there ever was an unambiguous sentence (ignoring the difference between literal and metaphorical meanings, which have nothing to do with syntax). But to the surprise of the programmers, the sharp-eyed computer found it to have five different trees!

Time proceeds as quickly as an arrow proceeds. (the intended reading)
Measure the speed of flies in the same way that you measure the speed of an arrow.
Measure the speed of flies in the same way that an arrow measures the speed of flies.
Measure the speed of flies that resemble an arrow.
Flies of a particular kind, time-flies, are fond of an arrow.

Among computer scientists the discovery has been summed up in the aphorism "Time flies like an arrow; fruit flies like a banana." Or consider the song line *Mary had a little lamb.* Unambiguous? Imagine that the second line was: *With mint sauce.* Or: *And the doctors were surprised.* Or: *The tramp!* There is even structure in seemingly nonsensical lists of words. For example, this fiendish string devised by my student Annie Senghas is a grammatical sentence:

Buffalo buffalo Buffalo buffalo buffalo buffalo Buffalo buffalo.

American bison are called *buffalo*. A kind of bison that comes from Buffalo, New York, could be called a *Buffalo buffalo*. Recall that there is a verb *to buffalo* that means "to overwhelm, to intimidate." Imagine that New York State bison intimidate one another: *(The) Buffalo buffalo (that) Buffalo buffalo (often) buffalo (in turn) buffalo (other) Buffalo buffalo.* The psycholinguist and philosopher Jerry Fodor has observed that the Yale University football cheer

Bulldogs Bulldogs Bulldogs Fight Fight Fight!

is a grammatical sentence, albeit a triply center-embedded one.

How do people home in on the sensible analysis of a sentence, without tarrying over all the grammatically legitimate but bizarre alternatives? There are two possibilities. One is that our brains are like computer parsers, computing dozens of doomed tree fragments in the background, and the unlikely ones are somehow filtered out before they reach consciousness. The other is that the human parser somehow gambles at each step about the alternative most likely to be true and then plows ahead with that single interpretation as far as possible. Computer scientists call these alternatives "breadth-first search" and "depth-first search."

At the level of individual words, it looks as if the brain does a breadth-first search, entertaining, however briefly, several entries for an ambiguous word, even unlikely ones. In an ingenious experiment, the psycholinguist David Swinney had people listen over headphones to passages like the following:

Rumor had it that, for years, the government building had been plagued with problems. The man was not surprised when he found several spiders, roaches, and other bugs in the corner of his room.

Did you notice that the last sentence contains an ambiguous word, *bug*, which can mean either "insect" or "surveillance device"? Probably not; the second meaning is more obscure and makes no sense in

context. But psycholinguists are interested in mental processes that last only milliseconds and need a more subtle technique than just asking people. As soon as the word *bug* had been read from the tape, a computer flashed a word on a screen, and the person had to press a button as soon as he or she had recognized it. (Another button was available for nonwords like *blick*.) It is well known that when a person hears one word, any word related to it is easier to recognize, as if the mental dictionary is organized like a thesaurus, so that when one word is found, others similar in meaning are more readily available. As expected, people pressed the button faster when recognizing *ant*, which is related to *bug*, than when recognizing *sew*, which is unrelated. Surprisingly, people were just as primed to recognize the word *spy*, which is, of course, related to *bug*, but only to the meaning that makes no sense in the context. It suggests that the brain knee-jerkingly activates both entries for *bug*, even though one of them could sensibly be ruled out beforehand. The irrelevant meaning is not around long: if the test word appeared on the screen three syllables after *bugs* instead of right after it, then only *ant* was recognized quickly; *spy* was no longer any faster than *sew*. Presumably that is why people deny that they even entertain the inappropriate meaning.

The psychologists Mark Seidenberg and Michael Tanenhaus showed the same effect for words that were ambiguous as to part-of-speech category, like *tires*, which we encountered in the ambiguous headline *Stud Tires Out*. Regardless of whether the word appeared in a noun position, like *The tires . . .*, or in a verb position, like *He tires . . .*, the word primed both *wheels*, which is related to the noun meaning, and *fatigue*, which is related to the verb meaning. Mental dictionary lookup, then, is quick and thorough but not very bright; it retrieves nonsensical entries that must be weeded out later.

At the level of the phrases and sentences that span many words, though, people clearly are not computing every possible tree for a sentence. We know this for two reasons. One is that many sensible ambiguities are simply never recognized. How else can we explain the ambiguous newspaper passages that escaped the notice of editors, no doubt to their horror later on? I cannot resist quoting some more:

The judge sentenced the killer to die in the electric chair for the second time.

Dr. Tackett Gives Talk on Moon
No one was injured in the blast, which was attributed to the buildup of gas by one town official.
The summary of information contains totals of the number of students broken down by sex, marital status, and age.

I once read a book jacket flap that said that the author lived with her husband, an architect and an amateur musician in Cheshire, Connecticut. For a moment I thought it was a ménage à quatre.

Not only do people fail to find some of the trees that are consistent with a sentence; sometimes they stubbornly fail to find the *only* tree that is consistent with a sentence. Take these sentences:

The horse raced past the barn fell.
The man who hunts ducks out on weekends.
The cotton clothing is usually made of grows in Mississippi.
The prime number few.
Fat people eat accumulates.
The tycoon sold the offshore oil tracts for a lot of money wanted to kill JR.

Most people proceed contendedly through the sentence up to a certain point, then hit a wall and frantically look back to earlier words to try to figure out where they went wrong. Often the attempt fails and people assume that the sentences have an extra word tacked onto the end or consist of two pieces of sentence stitched together. In fact, each one is a grammatical sentence:

The horse that was walked past the fence proceeded steadily, but the horse raced past the barn fell.
The man who fishes goes into work seven days a week, but the man who hunts ducks out on weekends.
The cotton that sheets are usually made of grows in Egypt, but the cotton clothing is usually made of grows in Mississippi.
The mediocre are numerous, but the prime number few.
Carbohydrates that people eat are quickly broken down, but fat people eat accumulates.

JR Ewing had swindled one tycoon too many into buying useless properties. The tycoon sold the offshore oil tracts for a lot of money wanted to kill JR.

These are called garden path sentences, because their first words lead the listener "up the garden path" to an incorrect analysis. Garden path sentences show that people, unlike computers, do not build all possible trees as they go along; if they did, the correct tree would be among them. Rather, people mainly use a depth-first strategy, picking an analysis that seems to be working and pursuing it as long as possible; if they come across words that cannot be fitted into the tree, they backtrack and start over with a different tree. (Sometimes people can hold a second tree in mind, especially people with good memories, but the vast majority of possible trees are never entertained.) The depth-first strategy gambles that a tree that has fit the words so far will continue to fit new ones, and thereby saves memory space by keeping only that tree in mind, at the cost of having to start over if it bet on the wrong horse raced past the barn.

Garden path sentences, by the way, are one of the hallmarks of bad writing. Sentences are not laid out with clear markers at every fork, allowing the reader to stride confidently through to the end. Instead the reader repeatedly runs up against dead ends and has to wend his way back. Here are some examples I have collected from newspapers and magazines:

Delays Dog Deaf-Mute Murder Trial
British Banks Soldier On
I thought that the Vietnam war would end for at least an appreciable chunk of time this kind of reflex anticommunist hysteria.
The musicians are master mimics of the formulas they dress up with irony.
The movie is Tom Wolfe's dreary vision of a past that never was set against a comic view of the modern hype-bound world.
That Johnny Most didn't need to apologize to Chick Kearn, Bill King, or anyone else when it came to describing the action [Johnny Most when he was in his prime].

Family Leave Law a Landmark Not Only for Newborn's Parents
Condom Improving Sensation to be Sold

In contrast, a great writer like Shaw can send a reader in a straight line from the first word of a sentence to the full stop, even if it is 110 words away.



A depth-first parser must use some criterion to pick one tree (or a small number) and run with it—ideally the tree most likely to be correct. One possibility is that the entirety of human intelligence is brought to bear on the problem, analyzing the sentence from the top down. According to this view, people would not bother to build any part of a tree if they could guess in advance that the meaning for that branch would not make sense in context. There has been a lot of debate among psycholinguists about whether this would be a sensible way for the human sentence parser to work. To the extent that a listener's intelligence can actually predict a speaker's intentions accurately, a top-down design would steer the parser toward correct sentence analyses. But the entirety of human intelligence is a lot of intelligence, and using it all at once may be too slow to allow for real-time parsing as the hurricane of words whizzes by. Jerry Fodor, quoting Hamlet, suggests that if knowledge and context had to guide sentence parsing, "the native hue of resolution would be sicklied o'er with the pale cast of thought." He has suggested that the human parser is an encapsulated module that can look up information only in the mental grammar and the mental dictionary, not in the mental encyclopedia.

Ultimately the matter must be settled in the laboratory. The human parser does seem to use at least a bit of knowledge about what tends to happen in the world. In an experiment by the psychologists John Trueswell, Michael Tanenhaus, and Susan Garnsey, people bit on a bar to keep their heads perfectly still and read sentences on a computer screen while their eye movements were recorded. The sentences had potential garden paths in them. For example, read the sentence

The defendant examined by the lawyer turned out to be unreliable.

You may have been momentarily sidetracked at the word *by*, because up to that point the sentence could have been about the defendant's examining something rather than his being examined. Indeed, the subjects' eyes lingered on the word *by* and were likely to backtrack to reinterpret the beginning of the sentence (compared to unambiguous control sentences). But now read the following sentence:

The evidence examined by the lawyer turned out to be unreliable.

If garden paths can be avoided by common-sense knowledge, this sentence should be much easier. Evidence, unlike defendants, can't examine anything, so the incorrect tree, in which the evidence would be examining something, is potentially avoidable. People do avoid it: the subjects' eyes hopped through the sentence with little pausing or backtracking. Of course, the knowledge being applied is quite crude (defendants examine things; evidence doesn't), and the tree that it calls for was fairly easy to find, compared with the dozens that a computer can find. So no one knows *how much* of a person's general smarts can be applied to understanding sentences in real time; it is an active area of laboratory research.

Words themselves also provide some guidance. Recall that each verb makes demands of what else can go in the verb phrase (for example, you can't just *devour* but have to *devour something*; you can't *dine something*, you can only *dine*). The most common entry for a verb seems to pressure the mental parser to find the role-players it wants. Trueswell and Tanenhaus watched their volunteers' eyeballs as they read

The student forgot the solution was in the back of the book.

At the point of reaching *was*, the eyes lingered and then hopped back, because the people misinterpreted the sentence as being about a student forgetting the solution, period. Presumably, inside people's heads the word *forget* was saying to the parser "Find me an object, now!" Another sentence was

The student hoped the solution was in the back of the book.

With this one there was little problem, because the word *hope* was saying, instead, "Find me a sentence!" and a sentence was there to be found.

Words can also help by suggesting to the parser exactly which other words they tend to appear with inside a given kind of phrase. Though word-by-word transition probabilities are not enough to understand a sentence (Chapter 4), they could be helpful; a parser armed with good statistics, when deciding between two possible trees allowed by a grammar, can opt for the tree that was most likely to have been spoken. The human parser seems to be somewhat sensitive to word pair probabilities: many garden paths seem especially seductive because they contain common pairs like *cotton clothing, fat people,* and *prime number.* Whether or not the brain benefits from language statistics, computers certainly do. In laboratories at AT&T and IBM, computers have been tabulating millions of words of text from sources like the *Wall Street Journal* and Associated Press stories. Engineers are hoping that if they equip their parsers with the frequencies with which each word is used, and the frequencies with which sets of words hang around together, the parsers will resolve ambiguities sensibly.

Finally, people find their way through a sentence by favoring trees with certain shapes, a kind of mental topiary. One guideline is momentum: people like to pack new words into the current dangling phrase, instead of closing off the phrase and hopping up to add the words to a dangling phrase one branch up. This "late closure" strategy might explain why we travel the garden path in the sentence

Flip said that Squeaky will do the work yesterday.

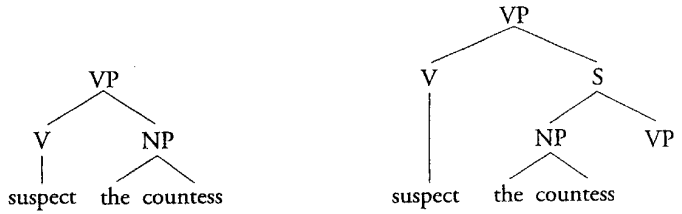
The sentence is grammatical and sensible, but it takes a second look (or maybe even a third) to realize it. We are led astray because when we encounter the adverb *yesterday,* we try to pack it inside the currently open VP *do the work,* rather than closing off that VP and hanging the adverb upstairs, where it would go in the same phrase as *Flip said.* (Note, by the way, that our knowledge of what is plausible, like the fact that the meaning of *will* is incompatible with the meaning of *yesterday,* did not keep us from taking the garden path. This

suggests that the power of general knowledge to guide sentence understanding is limited.) Here is an another example, though this time the psycholinguist responsible for it, Annie Senghas, did not contrive it as an example; one day she just blurted out, "The woman sitting next to Steven Pinker's pants are like mine." (Annie was pointing out that the woman sitting next to me had pants like hers.)

A second guideline is thrift: people to try to attach a phrase to a tree using as few branches as possible. This explains why we take the garden path in the sentence

Sherlock Holmes didn't suspect the very beautiful young countess was a fraud.

It takes only one branch to attach the *countess* inside the VP, where Sherlock would suspect her, but two branches to attach her to an S that is itself attached to the VP, where he would suspect her of being a fraud:



The mental parser seems to go for the minimal attachment, though later in the sentence it proves to be incorrect.

Since most sentences are ambiguous, and since laws and contracts must be couched in sentences, the principles of parsing can make a big difference in people's lives. Lawrence Solan discusses many examples in his recent book. Examine these passages, the first from an insurance contract, the second from a statute, the third from instructions to a jury:

Such insurance as is provided by this policy applies to the use of a non-owned vehicle by the named insured and any person responsi-

> ble for use by the named insured provided such use is with the permission of the owner.
>
> Every person who sells any controlled substance which is specified in subdivision (d) shall be punished. . . . (d) Any material, compound, mixture, or preparation which contains any quantity of the following substances having a potential for abuse associated with a stimulant effect on the central nervous system: Amphetamine; Methamphetamine . . .
>
> The jurors must not be swayed by mere sentiment, conjecture, sympathy, passion, prejudice, public opinion or public feeling.

In the first case, a woman was distraught over being abandoned in a restaurant by her date, and drove off in what she thought was the date's Cadillac, which she then totaled. It turned out to be someone else's Cadillac, and she had to recover the money from her insurance company. Was she covered? A California appellate court said yes. The policy was ambiguous, they noted, because the requirement *with the permission of the owner*, which she obviously did not meet, could be construed as applying narrowly to *any person responsible for use by the named insured*, rather than to *the named insured* (that is, her) *and any person responsible for use by the named insured.*

In the second case, a drug dealer was trying to swindle a customer—unfortunately for him, an undercover narcotics agent—by selling him a bag of inert powder that had only a minuscule trace of methamphetamine. The *substance* had "a potential for abuse," but the *quantity of the substance* did not. Did he break the law? The appellate court said he did.

In the third case, the defendant had been convicted of raping and murdering a fifteen-year-old-girl, and a jury imposed the death penalty. United States constitutional law forbids any instruction that would deny a defendant the right to have the jury consider any "sympathy factor" raised by the evidence, which in his case consisted of psychological problems and a harsh family background. Did the instructions unconstitutionally deprive the accused of *sympathy*, or did it deprive him only of the more trivial *mere sympathy*? The United States Supreme Court ruled 5–4 that he was denied only *mere sympathy*; that denial is constitutional.

Solan points out that the courts often resolve these cases by relying on "canons of construction" enshrined in the legal literature, which

correspond to the principles of parsing I discussed in the preceding section. For example, the Last Antecedent Rule, which the courts used to resolve the first two cases, is simply the "minimal attachment" strategy that we just saw in the Sherlock sentence. The principles of mental parsing, then, literally have life-or-death consequences. But psycholinguists who are now worrying that their next experiment may send someone to the gas chamber can rest easy. Solan notes that judges are not very good linguists; for better or worse, they try to find a way around the most natural interpretation of a sentence if it would stand in the way of the outcome they feel is just.



I have been talking about trees, but a sentence is not just a tree. Since the early 1960s, when Chomsky proposed transformations that convert deep structures to surface structures, psychologists have used laboratory techniques to try to detect some kind of fingerprint of the transformation. After a few false alarms the search was abandoned, and for several decades the psychology textbooks dismissed transformations as having no "psychological reality." But laboratory techniques have become more sophisticated, and the detection of something like a transformational operation in people's minds and brains is one of the most interesting recent findings in the psychology of language.

Take the sentence

> The policeman saw the boy that the crowd at the party accused (*trace*) of the crime.

Who was accused of a crime? The boy, of course, even though the words *the boy* do not occur after *accused*. According to Chomsky, that is because a phrase referring to the boy really does occur after *accused* in deep structure; it has been moved backwards to the position of *that* by a transformation, leaving behind a silent "trace." A person trying to understand the sentence must undo the effect of the transformation and mentally put a copy of the phrase back in the position of the trace. To do so, the understander must first notice, while at the beginning of the sentence, that there is a moved phrase, *the boy*, that needs a home. The understander must hold the phrase in short-term memory until he or she discovers a gap: a position

where a phrase should be but isn't. In this sentence there is a gap after *accused*, because *accused* demands an object, but there isn't one. The person can assume that the gap contains a trace and can then retrieve the phrase *the boy* from short-term memory and link it to the trace. Only then can the person figure out what role *the boy* played in the event—in this case, being accused.

Remarkably, every one of these mental processes can be measured. During the span of words between the moved phrase and the trace—the region I have underlined—people must hold the phrase in memory. The strain should be visible in poorer performance of any mental task carried out concurrently. And in fact, while people are reading that span, they detect extraneous signals (like a blip flashed on the screen) more slowly, and have more trouble keeping a list of extra words in memory. Even their EEG's (electroencephalograms, or records of the brain's electrical activity) show the effects of the strain.

Then, at the point at which the trace is discovered and the memory store can be emptied, the dumped phrase makes an appearance on the mental stage that can be detected in several ways. If an experimenter flashes a word from the moved phrase (for example, *boy*) at that point, people recognize it more quickly. They also recognize words related to the moved phrase—say, *girl*—more quickly. The effect is strong enough to be visible in brain waves: if interpreting the trace results in an implausible interpretation, as in

> Which food did the children read (*trace*) in class?

the EEG's show a boggle reaction at the point of the trace.

Connecting phrases with traces is a hairy computational operation. The parser, while holding the phrase in mind, must constantly be checking for the trace, an invisible and inaudible little nothing. There is no way of predicting how far down in the sentence the trace will appear, and sometimes it can be quite far down:

> The girl wondered who John believed that Mary claimed that the baby saw (*trace*).

And until it is found, the semantic role of the phrase is a wild card, especially now that the *who/whom* distinction is going the way of the phonograph record.

> I wonder who (*trace*) introduced John to Marsha. [*who* = the introducer]
> I wonder who Bruce introduced (*trace*) to Marsha. [*who* = the one being introduced]
> I wonder who Bruce introduced John to (*trace*). [*who* = the target of the introduction]

This problem is so tough that good writers, and even the grammar of the language itself, take steps to make it easier. One principle for good style is to minimize the amount of intervening sentence in which a moved phrase must be held in memory (the underlined regions). This is a task that the English passive construction is good for (notwithstanding the recommendations of computerized "style-checkers" to avoid it across the board). In the following pair of sentences, the passive version is easier, because the memory-taxing region before the trace is shorter:

> Reverse the clamp that the stainless steel hex-head bolt extending upward from the seatpost yoke holds (*trace*) in place.
> Reverse the clamp that (*trace*) is held in place by the stainless steel hex-head bolt extending upward from the seatpost yoke.

And universally, grammars restrict the amount of tree that a phrase can move across. For example, one can say

> That's the guy that you heard the rumor about (*trace*).

But the following sentence is quite odd:

> That's the guy that you heard the rumor that Mary likes (*trace*).

Languages have "bounding" restrictions that turn some phrases, like the complex noun phrase *the rumor that Mary likes him*, into "islands" from which no words can escape. This is a boon to listeners, because the parser, knowing that the speaker could not have moved something out of such a phrase, can get away with not monitoring it for a trace. But the boon to listeners exerts a cost on speakers; for these sentences

they have to resort to a clumsy extra pronoun, as in *That's the guy that you heard the rumor that Mary likes him.*



Parsing, for all its importance, is only the first step in understanding a sentence. Imagine parsing the following real-life dialogue:

P: The grand jury thing has its, uh, uh, uh—view of this they might, uh. Suppose we have a grand jury proceeding. Would that, would that, what would that do to the Ervin thing? Would it go right ahead anyway?
D: Probably.
P: But then on that score, though, we have—let me just, uh, run by that, that—You do that on a grand jury, we could then have a much better cause in terms of saying, "Look, this is a grand jury, in which, uh, the prosecutor—" How about a special prosecutor? We could use Petersen, or use another one. You see he is probably suspect. Would you call in another prosecutor?
D: I'd like to have Petersen on our side, advising us [laughs] frankly.
P: Frankly. Well, Petersen is honest. Is anybody about to be question him, are they?
D: No, no, but he'll get a barrage when, uh, these Watergate hearings start.
P: Yes, but he can go up and say that he's, he's been told to go further in the Grand Jury and go in to this and that and the other thing. Call everybody in the White House. I want them to come, I want the, uh, uh, to go to the Grand Jury.
D: This may result—This may happen even without our calling for it when, uh, when these, uh—
P: Vescoe?
D: No. Well, that's one possibility. But also when these people go back before the Grand Jury here, they are going to pull all these criminal defendants back in before the Grand Jury and immunize them.
P: And immunize them: Why? Who? Are you going to—On what?
D: Uh, the U.S. Attorney's Office will.
P: To do what?
D: To talk about anything further they want to talk about.
P: Yeah. What do they gain out of it?
D: Nothing.

P: To hell with them.
D: They, they're going to stonewall it, uh, as it now stands. Except for Hunt. That's why, that's the leverage in his threat.
H: This is Hunt's opportunity.
P: That's why, that's why,
H: God, if he can lay this—
P: That's why your, for your immediate thing you've got no choice with Hunt but the hundred and twenty or whatever it is, right?
D: That's right.
P: Would you agree that that's a buy time thing, you better damn well get that done, but fast?
D: I think he ought to be given some signal, anyway, to, to—
P: [expletive deleted], get it, in a, in a way that, uh—Who's going to talk to him? Colson? He's the one who's supposed to know him.
D: Well, Colson doesn't have any money though. That's the thing. That's been our, one of the real problems. They have, uh, been unable to raise any money. A million dollars in cash, or, or the like, has been just a very difficult problem as we've discussed before. Apparently, Mitchell talked to Pappas, and I called him last—John asked me to call him last night after our discussion and after you'd met with John to see where that was. And I, I said, "Have you talked to, to Pappas?" He was at home, and Martha picked up the phone so it was all in code. "Did you talk to the Greek?" And he said, uh, "Yes, I have." And I said, "Is the Greek bearing gifts?" He said, "Well, I want to call you tomorrow on that."
P: Well, look, uh, what is it that you need on that, uh, when, uh, uh? Now look [unintelligible] I am, uh, unfamiliar with the money situation.

This dialogue took place on March 17, 1973, among President Richard Nixon (P), his counsel John W. Dean 3rd (D), and his chief of staff H. R. Haldeman (H). Howard Hunt, working for Nixon's re-election campaign in June 1972, had directed a break-in at the Democratic Party headquarters in the Watergate building, in which his men bugged the telephones of the party chairman and other workers. Several investigations were under way to determine if the operation had been ordered from the White House, by Haldeman or Attorney General John Mitchell. The men are discussing whether to pay $120,000 in "hush money" to Hunt before he testified before a

grand jury. We have this verbatim dialogue because in 1970 Nixon, claiming to be acting on behalf of future historians, bugged his own office and began secretly taping all his conversations. In February 1974 the Judiciary Committee of the House of Representatives subpoenaed the tapes to help them determine whether Nixon should be impeached. This excerpt is from their transcription. Largely on the basis of this passage, the committee recommended impeachment. Nixon resigned in August 1974.

The Watergate tapes are the most famous and extensive transcripts of real-life speech ever published. When they were released, Americans were shocked, though not all for the same reason. Some people—a very small number—were surprised that Nixon had taken part in a conspiracy to obstruct justice. A few were surprised that the leader of the free world cussed like a stevedore. But one thing that surprised everyone was what ordinary conversation looks like when it is written down verbatim. Conversation out of context is virtually opaque.

Part of the problem comes from the circumstances of transcription: the intonation and timing that delineate phrases is lost, and a transcription from anything but the highest-fidelity tape is unreliable. Indeed, in the White House's independent transcription of this low-quality recording, many puzzling passages are rendered more sensibly. For example, *I want the, uh, uh, to go* is transcribed as *I want them, uh, uh, to go.*

But even when transcribed perfectly, conversation is hard to interpret. People often speak in fragments, interrupting themselves in midsentence to reformulate the thought or change the subject. It's often unclear who or what is being talked about, because conversers use pronouns (*him, them, this, that, we, they, it, one*), generic words (*do, happen, the thing, the situation, that score, these people, whatever*), and ellipses (*The U.S. Attorney's Office will* and *That's why*). Intentions are expressed indirectly. In this episode, whether a man would end the year as president of the United States or as a convicted criminal literally hinged on the meaning of *get it* and on whether *What is it that you need?* was meant as a request for information or as an implicit offer to provide something.

Not everyone was shocked by the unintelligibility of transcribed speech. Journalists know all about it, and it is a routine practice to

edit quotations and interviews heavily before they are published. For many years the temperamental Boston Red Sox pitcher Roger Clemens complained bitterly that the press misquoted him. The *Boston Herald*, in what they must have known was a cruel trick, responded by running a daily feature in which his post-game comments were reproduced word for word.

Journalists' editing of conversations became a legal issue in 1983, when the writer Janet Malcolm published an unflattering *New Yorker* series about the psychoanalyst Jeffrey Masson. Masson had written a book accusing Freud of dishonesty and cowardice in retracting his observation that neurosis is caused by sexual abuse in childhood, and was fired as the curator of the Freud archives in London. According to Malcolm, Masson described himself in her interviews as "an intellectual gigolo" and "after Freud, the greatest analyst who's ever lived," and as planning to turn Anna Freud's house after her death into "a place of sex, women, and fun." Masson sued Malcolm and the *New Yorker* for ten million dollars, claiming that he had never said these things and that other quotations had been altered to make him look ridiculous. Though Malcolm could not document the quotations from her tapes and handwritten notes, she denied having manufactured them, and her lawyers argued that even if she had, they were a "rational interpretation" of what Masson had said. Doctored quotes, they argued, are standard journalistic practice and are not examples of printing something with knowledge that it is false or with reckless disregard for whether it is false, part of the definition of libel.

Several courts threw out the case on First Amendment grounds, but in June 1991 the Supreme Court unanimously reinstated it. In a closely watched opinion, the majority defined a middle ground for journalists' treatment of quotations. (Requiring them to publish quotes verbatim was not even considered.) Justice Kennedy, writing for the majority, said that the "deliberate alteration of the words uttered by a plaintiff does not equate with knowledge of falsity," and that "If an author alters a speaker's words, but effects no material change in meaning, the speaker suffers no injury to reputation. We reject any special test of falsity for quotations, including one which would draw the line at correction of grammar or syntax." If the Supreme Court had asked me, I would have sided with Justices White and Scalia in calling for some such line to be drawn. Like many

linguists, I doubt that it is possible to alter a speaker's words—including most grammar and syntax—without materially changing the meaning.

These incidents show that real speech is very far from *The dog likes ice cream* and that there is much more to understanding a sentence than parsing it. Comprehension uses the semantic information recovered from a tree as just one premise in a complex chain of inference to the speaker's intentions. Why is this so? Why is it that even honest speakers rarely articulate the truth, the whole truth, and nothing but the truth?

The first reason is air time. Conversation would bog down if one had to refer to the United States Senate Select Committee on the Watergate Break-In and Related Sabotage Efforts by uttering that full description every time. Once alluded to, *the Ervin thing*, or just *it*, will suffice. For the same reason it is wasteful to spell out the following chain of logic:

> Hunt knows who gave him the orders to organize the Watergate break-in.
> The person who gave him the orders might be part of our administration.
> If the person is in our administration and his identity becomes public, the entire administration will suffer.
> Hunt has an incentive to reveal the identity of the person who gave him the orders because it might reduce his prison sentence.
> Some people will take risks if they are given enough money.
> Therefore Hunt may conceal the identity of his superior if he is given enough money.
> There is reason to believe that approximately $120,000 would be a large enough incentive for Hunt to conceal the identity of the person who gave him the order.
> Hunt could accept that money now, but it is in his interest to continue to blackmail us in the future.
> Nonetheless it might be sufficient for us to keep him quiet in the short run because the press and the public might lose interest in the Watergate scandal in the months to come, and if he reveals the identity later, the consequences for our administration would not be as negative.
> Therefore the self-interested course of action for us is to pay Hunt the amount of money that would be a large enough incentive for him to keep silent until such time as public interest in Watergate wanes.

It is more efficient to say, "For your immediate thing you've got no choice with Hunt but the hundred and twenty or whatever it is."

The efficiency, though, depends on the participants' sharing a lot of background knowledge about the events and about the psychology of human behavior. They must use this knowledge to cross-reference the names, pronouns, and descriptions with a single cast of characters, and to fill in the logical steps that connect each sentence with the next. If background assumptions are not shared—for example, if one's conversational partner is from a very different culture, or is schizophrenic, or is a machine—then the best parsing in the world will fail to deliver the full meaning of a sentence. Some computer scientists have tried to equip programs with little "scripts" of stereotyped settings like restaurants and birthday parties to help their programs fill in the missing parts of texts while understanding them. Another team is trying to teach a computer the basics of human common sense, which they estimate to comprise about ten million facts. To see how formidable the task is, consider how much knowledge about human behavior must be interpolated to understand what *he* means in a simple dialogue like this:

> Woman: I'm leaving you.
> Man: Who is he?

Understanding, then, requires integrating the fragments gleaned from a sentence into a vast mental database. For that to work, speakers cannot just toss one fact after another into a listener's head. Knowledge is not like a list of facts in a trivia column but is organized into a complex network. When a series of facts comes in succession, as in a dialogue or text, the language must be structured so that the listener can place each fact into an existing framework. Thus information about the old, the given, the understood, the topic, should go early in the sentence, usually as the subject, and information about the new, the focus, the comment, should go at the end. Putting the topic early in the sentence is another function of the maligned passive

construction. In his book on style, Williams notes that the usual advice "Avoid passives" should be flouted when the topic being discussed has the role connected with the deep-structure object of the verb. For example, read the following two-sentence discussion:

> Some astonishing questions about the nature of the universe have been raised by scientists studying the nature of black holes in space. The collapse of a dead star into a point perhaps no larger than a marble creates a black hole.

The second sentence feels like a non sequitur. It is much better to put it in the passive voice:

> Some astonishing questions about the nature of the universe have been raised by scientists studying the nature of black holes in space. A black hole is created by the collapse of a dead star into a point perhaps no larger than a marble.

The second sentence now fits in smoothly, because its subject, *a black hole*, is the topic, and its predicate adds new information to that topic. In an extended conversation or essay, a good writer or speaker will make the focus of one sentence the topic of the next one, linking propositions into an orderly train.

The study of how sentences are woven into a discourse and interpreted in context (sometimes called "pragmatics") has made an interesting discovery, first pointed out by the philosopher Paul Grice and recently refined by the anthropologist Dan Sperber and the linguist Deirdre Wilson. The act of communicating relies on a mutual expectation of cooperation between speaker and listener. The speaker, having made a claim on the precious ear of the listener, implicitly guarantees that the information to be conveyed is relevant: that it is not already known, and that it is sufficiently connected to what the listener is thinking that he or she can make inferences to new conclusions with little extra mental effort. Thus listeners tacitly expect speakers to be informative, truthful, relevant, clear, unambiguous, brief, and orderly. These expectations help to winnow out the inappropriate readings of an ambiguous sentence, to piece together fractured utterances, to excuse slips of the tongue, to guess the referents of pronouns and descriptions, and to fill in the missing steps of an argument. (When

a receiver of a message is not cooperative but adversarial, all of this missing information must be stated explicitly, which is why we have the tortuous language of legal contracts with their "party of the first part" and "all rights under said copyright and all renewals thereof subject to the terms of this Agreement.")

The interesting discovery is that the maxims of relevant conversation are often observed in the breach. Speakers deliberately flout them in the literal content of their speech so that listeners can interpolate assumptions that would restore the conversation to relevance. Those assumptions then serve as the real message. A familiar example is the following kind of letter of recommendation:

> Dear Professor Pinker:
>
> I am very pleased to be able to recommend Irving Smith to you. Mr. Smith is a model student. He dresses well and is extremely punctual. I have known Mr. Smith for three years now, and in every way I have found him to be most cooperative. His wife is charming.
>
> Sincerely,
>
> John Jones
> Professor

Though the letter contains nothing but positive, factual statements, it guarantees that Mr. Smith will not get the position he is seeking. The letter contains no information relevant to the reader's needs, and thereby violates the maxim that speakers be informative. The reader works on the tacit assumption that the communicative act as a whole *is* relevant, even if the content of the letter itself is not, so he infers a premise that together with the letter makes the act relevant: that the writer has no relevant positive information to convey. Why does the writer demand this minuet, rather than just saying "Stay away from Smith; he's dumb as a tree"? It is because of another premise that the reader can interpolate: the writer is the kind of person who does not casually injure those who put their trust in him.

It is natural that people exploit the expectations necessary for successful conversation as a way of slipping their real intentions into covert layers of meaning. Human communication is not just a transfer of information like two fax machines connected with a wire; it is a

series of alternating displays of behavior by sensitive, scheming, second-guessing, social animals. When we put words into people's ears we are impinging on them and revealing our own intentions, honorable or not, just as surely as if we were touching them. Nowhere is this more apparent than in the convoluted departures from plain speaking found in every society that are called politeness. Taken literally, the statement "I was wondering if you would be able to drive me to the airport" is a prolix string of incongruities. Why notify me of the contents of your ruminations? Why are you pondering my competence to drive you to the airport, and under which hypothetical circumstances? Of course the real intent—"Drive me to the airport"—is easily inferred, but because it was never stated, I have an out. Neither of us has to live with the face-threatening consequences of your issuing a command that presupposes you could coerce my compliance. Intentional violations of the unstated norms of conversation are also the trigger for many of the less pedestrian forms of nonliteral language, such as irony, humor, metaphor, sarcasm, putdowns, ripostes, rhetoric, persuasion, and poetry.

Metaphor and humor are useful ways to summarize the two mental performances that go into understanding a sentence. Most of our everyday expressions about language use a "conduit" metaphor that captures the parsing process. In this metaphor, ideas are objects, sentences are containers, and communication is sending. We "gather" our ideas to "put" them "into" words, and if our verbiage is not "empty" or "hollow," we might "convey" or "get" these ideas "across" "to" a listener, who can "unpack" our words to "extract" their "content." But as we have seen, the metaphor is misleading. The complete process of understanding is better characterized by the joke about the two psychoanalysts who meet on the street. One says, "Good morning"; the other thinks, "I wonder what he meant by that."

8

The Tower of Babel

And the whole earth was of one language, and of one speech. And it came to pass, as they journeyed from the east, that they found a plain in the land of Shinar; and they dwelt there. And they said one to another, Go to, let us make brick, and burn them thoroughly. And they had brick for stone, and slime had they for morter. And they said, Go to, let us build us a city and a tower, whose top may reach unto heaven; and let us make us a name, lest we be scattered abroad upon the face of the whole earth. And the Lord came down to see the city and the tower, which the children of men builded. And the Lord said, Behold, the people is one, and they have all one language; and this they begin to do: and now nothing will be restrained from them, which they have imagined to do. Go to, let us go down, and there confound their language, that they may not understand one another's speech. So the Lord scattered them abroad from thence upon the face of all the earth: and they left off to build the city. Therefore is the name of it called Babel; because the Lord did there confound the language of all the earth: and from thence did the Lord scatter them abroad upon the face of all the earth. (Genesis 11:1–9)

In the year of our Lord 1957, the linguist Martin Joos reviewed the preceding three decades of research in linguistics and concluded that God had actually gone much farther in confounding the language of Noah's descendants. Whereas the God of Genesis was said to be content with mere mutual unintelligibility, Joos declared that "lan-