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, therfore 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 comprehension 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' 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 the 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 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 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."

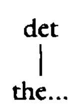
N → boy, girl, dog, cat, ice cream, candy, hotdogs
"The nouns in the mental dictionary include *boy, girl,* . . ."

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

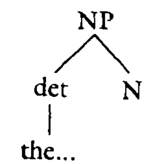
P → with, in, near
"The prepositions include *with, in, near.*"

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.)

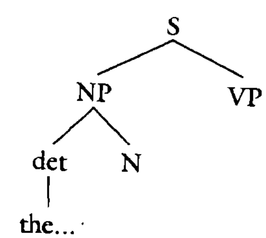


Determiners, like all words, have to be part of some larger phrase. The parser can figure out which phrase by checking to see which rule has "det" on its right-hand side. That rule is the one defining a noun phrase, NP. More tree can be grown:



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.

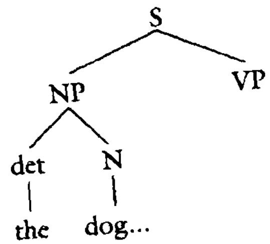
In the meantime, the tree continues to grow, for NP's cannot float around unattached. Having checked the right-hand sides of the rules for an NP symbol, the parser has several options. The freshly built NP could be part of a sentence, part of a verb phrase, or part of a prepositional phrase. The choice 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 the rules confirms the prediction: *dog* is part of the N rule.

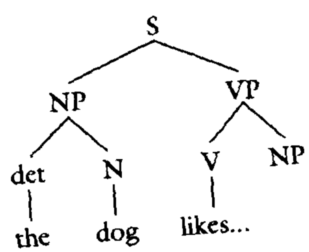
This allows *dog* to be integrated into the tree, completing the noun phrase:



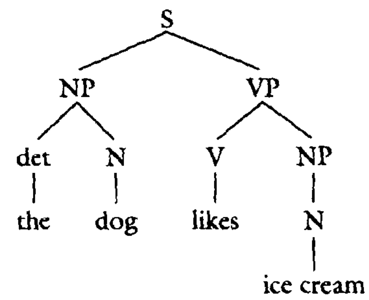
The parser no longer has to remember that there is an NP to be completed; all it has to keep in mind is the incomplete S.

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 the 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 short-sighted attempt to balance this year's budget at the same time that the governor's cronies and bureaucratic hacks are lining their pockets.

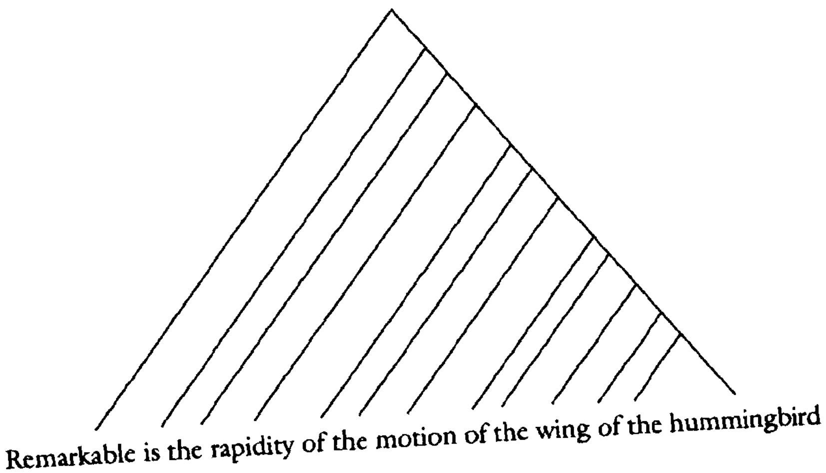
Many linguists believe that the reason that languages allow phrase movement, or choices among more-or-less synonymous constructions, is to ease the load on the listener's memory.

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

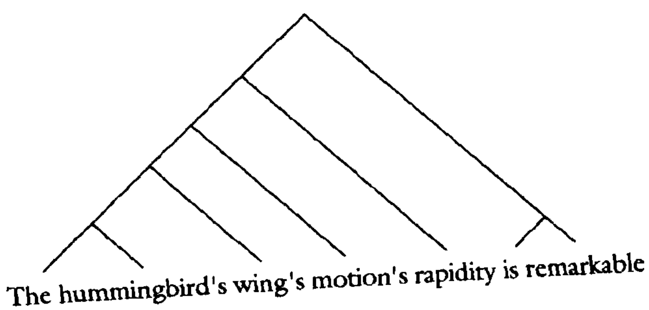
> Remarkable is the rapidity of the motion of the wing of the hummingbird.
>
> This is the cow with the crumpled horn that tossed the dog that worried the cat that killed the rat that ate the malt that lay in the house that Jack built.
>
> Then came the Holy One, blessed be He, and destroyed the angel 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:



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:



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.

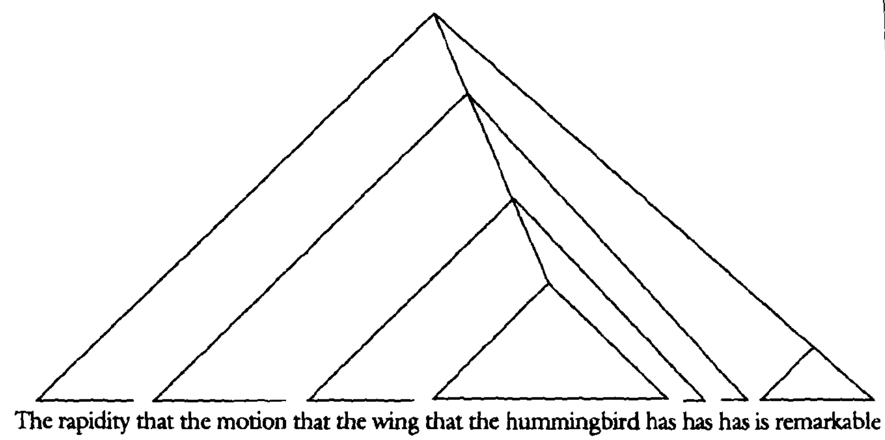
The clause ***that the motion has*** has been embedded in the noun phrase containing ***The rapidity***. The result is a bit stilted but easy to understand. One can also say

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:



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 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 a 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 consis-

tent 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

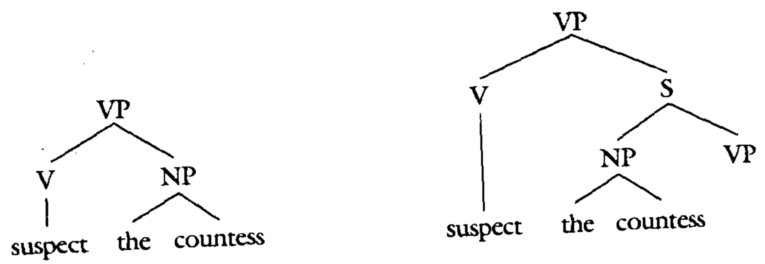
> Flip said that Squeaky will do the work yesterday.

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." (Anne 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 responsible 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: Vesco?
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 reelection 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 were 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 communciation 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 to one another, Go to, let us make brick, and burn them thoroughly. And they had brick for stone, and slime had they for mortar. 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)