Chapter 8

Comprehending Sentence Structure

Janet Dean Fodor

8.1 From Word String to Sentence Meaning

8.1.1 Deducing Structure

When we hear or read a sentence, we are aware more or less instantaneously of what it means. Our minds compute the meaning somehow, on the basis of the words that comprise the sentence. But the words alone are not enough. The sentence meanings we establish are so precise that they could not be arrived at by just combining word meanings haphazardly. A haphazard word combiner could misunderstand (1) as meaning that all bears love; it could interpret (2) as meaning that pigs fly and rabbits can't.

(1) Love bears all.

(2) Pigs and rabbits can't fly.

But people (except when something is badly wrong) do not comprehend sentences in so vague a way. We combine word meanings according to a precise recipe that is provided by the syntax of the language. A sentence is more than a string of words. It is a highly structured object in which the words are organized into phrases and clauses in accord with general principles of syntactic patterning. For instance, the string of words in (2) does not have the structure shown by the brackets in (3a). Its structure is (3b), and this is why it means what it does.

(3) a. [Pigs [and rabbits can't] fly].

b. [Pigs and rabbits] [can't fly].

What is interesting is that this syntactic structure that drives sentence comprehension is not manifest in the stimulus. It is there but not overtly displayed. It is not reliably marked in the phonetic form of a spoken sentence or in the orthographic form of a written one. The prosody of the sentence (the melody and timing with which it is spoken; see chapter 9) does usually reflect some aspects of the syntactic structure, but not all; in written language, structural information is missing entirely except as indicated by an occasional comma. Thus, if perceivers are to use syntactic structure in comprehending sentences, it seems that they must first deduce the syntactic structure.

In order to do so, they must be applying their knowledge of the structural principles of the language. Different natural languages (English, Spanish, Japanese, and so forth) exhibit slightly (not entirely) different sentence patterns, which shape their sentences and dictate how word meanings are to be integrated. Linguists study languages in order to discover what these patterns are. But the speakers and perceivers of a language evidently have this information in their heads. We can't introspect about it, we have no conscious access to it (which is why linguists have to infer it laboriously by observing sentences just as entomologists observe insects); but we acquire this knowledge as infants learning the language, and we draw on it, unconsciously and automatically, to calculate sentence structure every time we speak or understand.

For example, people who know English know that the subject of sentence (1) is the abstract noun *love* and that its verb is *bears*, even though *love* can be a verb in other contexts and *bears* can be a noun. The fact that *love* is the subject here and not the verb follows from the facts of English grammar. In English (though not in Welsh or Maori) a sentence cannot begin with a verb unless it is an auxiliary verb (like *may* in *May Charlotte leave early?*) or an imperative verb (as in *Watch your back!*) or in a special "topicalized" construction such as *Leave you I must but forget you I never will.* Sentence (1) does not fit any of these patterns. The beginning of it may look like an imperative (*Love bears, don't fear them!*) but at least in modern English the *all* at the end of (1) scotches this possibility. Therefore the first word of (1) is not a verb. It is not an adjective or a preposition or anything else, so it must be a noun. This fact then entails that the next word, *bears*, functions as a verb here, not as a noun. This is because in English (unlike Korean) the subject cannot be followed directly by the object (or other noun phrase); the verb must come between them. (There can be two nouns in a row in English if they form a compound noun, like *honey bees*; but *love bears* does not work as a compound noun in (1).)

Perceivers thus PROJECT structure onto a string of words, deducing it from their mentally stored grammar of the language. This is an extraordinary feat that underlies even the most ordinary uses of language. Psycholinguists are interested in finding out HOW people put their unconscious grammatical knowledge to work. What exactly are these mental deductions, which so rapidly and reliably deliver sentence meanings to our

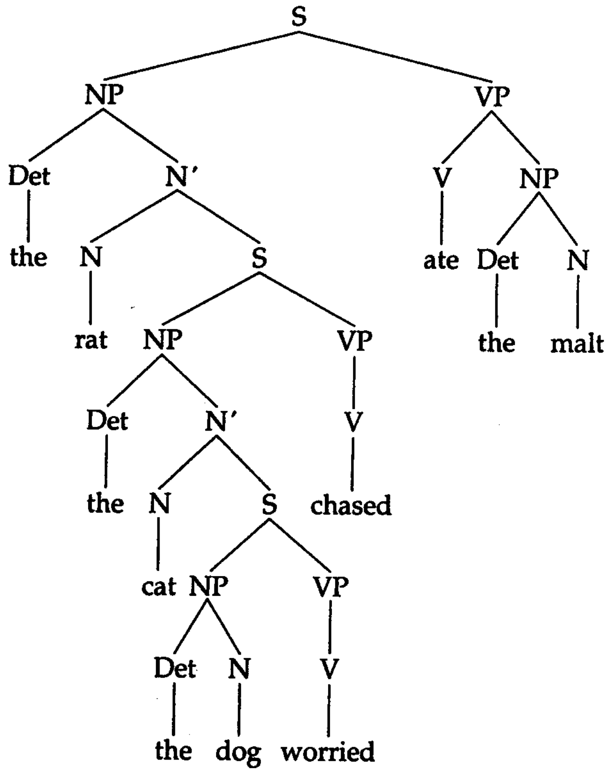
conscious minds? They are not directly observable, so we need to be resourceful in finding methods for investigating them. One strategy that can be helpful is to look at extreme cases, where a great many properties of the sentence have to be projected from very little perceptible evidence. These would be sentences where the ratio of "invisible" structure to overt words is very high. An example is (4).

(4) The rat the cat the dog worried chased ate the malt.

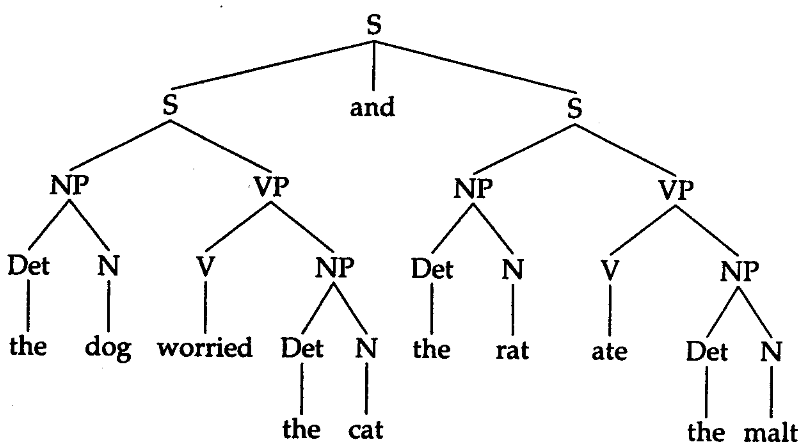
This is what is called a doubly center-embedded relative clause construction; it has one relative clause in the middle of another relative clause in the middle of a main clause. It was noticed some years ago by Yngve (1960) that sentences like this, though well formed in accord with the syntactic rules of English, are extremely difficult to structure and understand. Parts of (4) are easy enough to parse. The phrase *the cat the dog worried* is clear, and if we call this cat *Socks*, the sentence would be *The rat Socks chased ate the malt*, which is also comprehensible. But apparently there is just too much structure in (4) as a whole for our mental comprehension routines to cope with. Miller and Chomsky (1963) and, more recently, Frazier (1985) have observed that doubly center-embedded relative clause constructions have a very dense syntactic structure. A tree diagram of the structure of (4) is shown in (5), and you can see that it has many nonterminal (higher) nodes relative to its terminal (word-level) nodes. This is especially true at the beginning of the sentence, where the first six words require three clause structures to be built. Compare this with the structure shown in (6) for a sentence which has the same number of words but has fewer nodes more evenly distributed, and which is perfectly easy to understand.[1]

1. In this chapter, traditional tree diagrams (or structural bracketings) and category labels will be used to represent sentence structures. S = sentence or clause; NP = noun phrase; VP = verb phrase; Det = determiner (article). Readers should note that newer conventions (e.g., clauses as CP or IP) are in use in many recent works in linguistics.

(5)



(6)



The difficulty of (4) (and other sentences with multiple center-embeddings) shows that our mental sentence comprehension routines, though remarkably efficient most of the time, do have their limits.

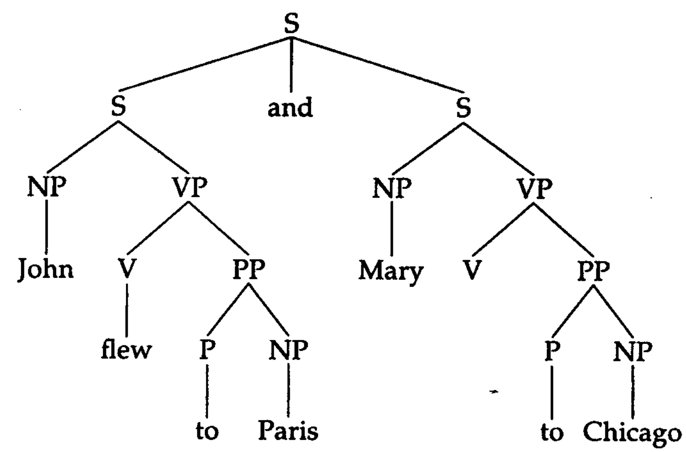
8.1.2 Empty Categories

Another way in which a sentence can have a high load of structure to be projected by the perceiver is if some of its constituents do not overtly appear in the word string. These non-overt constituents are what linguists call *empty categories.* They are *categories* in the syntactician's sense; that is, they are noun phrases or verbs or relative pronouns, and so forth. They are *empty* in the sense of lacking any phonological (or orthographic) realization. Thus, an empty category is a piece of the sentence structure, but it is not pronounced (or written) by the sentence producer, so it is not audible (or visible) to the sentence perceiver. The perceiver must deduce both its existence and its properties. An example is the "missing" verb *flew* in the second clause of sentence (7).

(7) John flew to Paris, and Mary to Chicago.

Mary is a noun phrase (NP) and *to Chicago* is a prepositional phrase (PP). A clause cannot normally consist of just an NP followed by a PP; it must have a verb. It seems reasonable to suppose, then, that the structure of (7) is (8), where there is a verb in both clauses in accord with general structural principles, but where the second verb is phonologically empty.

(8)



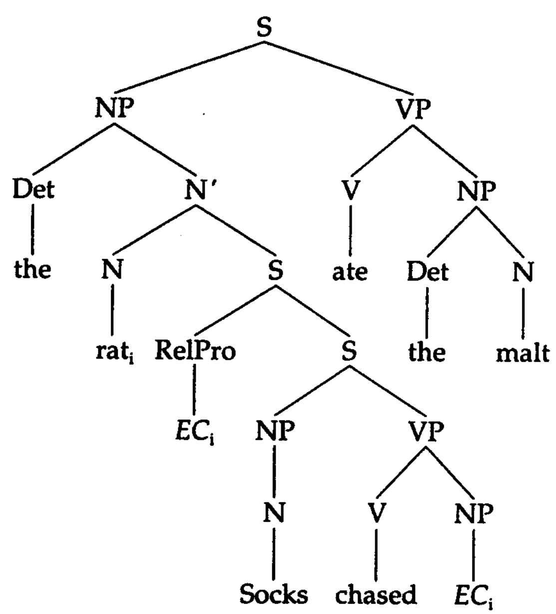
The empty verb has a quite specific meaning. The second clause of (7) clearly means that Mary FLEW to Chicago, not that she drove to Chicago, or that she wrote to Chicago, and so forth. On the other hand, if the first clause of (7) had been *John drove to Paris,* then the second clause would have meant that Mary DROVE to Chicago, not that she flew there. It is always the verb in the first clause that identifies the empty verb in the second clause. Thus, the grammar of English does not allow just any verb in any context to be empty; there are strict principles governing where empty categories (henceforth ECs) can appear in sentences and how they can be interpreted. If there were not, it would be impossible for perceivers to reconstruct ECs as they compute the sentence structure.

Sentence (4) above also contains some ECs, though they are not shown in (5). (If they were, the contrast in complexity between (5) and (6) would be even clearer.) A relative clause in English, as in many languages, has a "gap" in it where a noun phrase would normally appear. Consider example (9), which is the simplified version of (4) with just one relative clause.

(9) The rat Socks chased ate the malt.

The relative clause *Socks chased* modifies the noun *rat*; *rat* is the head noun of the whole complex NP *the rat Socks chased*. The relative clause means that Socks chased the rat, but the word *rat* doesn't actually appear as the object of the verb *chased*. The object of *chased* is missing. Since this is a verb that normally MUST have an object, we may assume that it does have an object in the relative clause but that the object is phonologically empty. Thus, the structure of (9) is as shown in (10).

(10)



The empty object NP in (10) is coindexed with the head noun *rat*; this is to indicate the fact that the EC is interpreted as referring to the rat. Note also that another EC is shown in (10), in the pre-clause position where a relative pronoun (*who, whom, which*) often appears. In a sentence like *The rat which Socks chased ate the malt* the relative pronoun is *which*, and it provides a link between the head noun and the EC in the relative clause. Sentence (9) is exactly similar except that its relative pronoun also happens to be an EC. Thus, the empty relative pronoun mediates the relation between *rat* and the empty object of *chased*. The fact that sentence (9) is easy to understand shows that this double linkage of ECs does not overload the human sentence processing routines.

Linguists have argued that the structure (10) is derived from an underlying structure in which the relative pronoun follows *chased*, in the "normal" position for the object of a verb. The underlying structure is transformed by moving the relative pronoun from the object position to the pre-clause position adjacent to the head noun that the clause modified. Whenever an element moves out of its underlying position, an EC coindexed with it is created in that position. An EC that is thus "left behind" by a movement operation is called a *trace* of the element that moved. Thus the EC after *chased* in (10) is the trace of the relative pronoun which moved leftward. This kind of trace is usually referred to as a WH-trace, since it results from the movement of a WH-phrase—an expression such as *who* or *which* or *with whom* or *how many of the elephants*, and so forth.

WH-phrases (which occur in many other languages though they do not typically start with *wh* except in English) appear in questions as well as in relative clauses. The question in (11) was formed by movement of the WH-phrase *which of the elephants* to the beginning of the question, leaving a WH-trace in its underlying position after the verb *tickling*.

(11) [Which of the elephants]$_i$ was Bertram tickling *WH-trace*$_i$?

Note that, in a question, an auxiliary verb (the verb *was* in (11)) moves to the left of the subject NP; the usual order would be *Bertram was*, but here it is *was Bertram*. This is Subject-Auxiliary Inversion, discussed by Lasnik in chapter 10 of this volume. Though it is an important aspect of the syntax of questions in English, it is not germane to the concerns of the present chapter; we will sidestep it by focusing on relative clauses and embedded questions, where it does not apply. An embedded question is shown in (12).

(12) The ringmaster asked [which of the elephants]$_i$ Bertram was tickling *WH-trace*$_i$.

The main clause of (12) is declarative, but it has a question embedded within it as the object of the verb *asked*. The WH-phrase *which of the elephants* has moved to the beginning of the question clause. In all cases of WH-movement, the semantic role of the removed phrase is determined by its underlying position. In (12), for instance, the WH-phrase originated in a position following the verb *tickling*, and the meaning is that the elephant was the object of the tickling action. The surface position of the WH-phrase itself cannot signal its semantic role, because the WH-phrase is always at the front of its clause regardless of the meaning. To comprehend a sentence that contains a WH-phrase, a perceiver therefore needs to know what its UNDERLYING position was. The underlying position is marked in the surface structure by a trace, but that is not very helpful; since a trace is an EC, it is inaudible and invisible to the perceiver. Let us consider how the perceiver (more precisely, the unconscious sentence processing routines in the perceiver's mind/brain) could set about locating the crucial EC position.

Consider first some simple strategies that WON'T work. The trace is not always at the end of the sentence, as is shown by (13). It is not always immediately following a verb, as is shown by (14), where it is the object of the preposition *with*. It is not always an object, as shown by (15), where it is the subject of the verb *were*. Example (16) shows that the trace may immediately follow its antecedent WH-phrase, and example (17) shows that it may be separated from the antecedent phrase by several intervening clauses.

(13) You can always tell [which books]$_i$ Walter read *WH-trace*$_i$ in the bathtub.

(14) I wonder [which of his books]$_i$ Walter lit the fire with *WH-trace*$_i$.

(15) Do you recall [which books]$_i$ Walter proclaimed *WH-trace*$_i$ were unreadable?

(16) Walter never did find out [which books]$_i$ *WH-trace*$_i$ were on the reading list.

(17) It is remarkable [how many books]$_i$ Walter tried to bribe his roommate to inform the instructor that he had every intention of reading *WH-trace*$_i$ soon.

It seems the only way for the comprehension routines to locate a trace is to find a position in the sentence that "needs" a phrase of the kind that has been moved. In examples (13)–(17), an NP has moved, so somewhere in the sentence there must be a "gap" that is "NP-shaped"; that is, a position where an NP would normally occur. In example (18) the adverb *where* has moved, so the gap later in the sentences must be "adverb-shaped." In (19) the PP *to whom* has moved, so there is a gap later that is suited to a PP.

(18) The waiter asked where$_i$ we would like to sit *WH-trace*$_i$.

(19) Marsha is the person [to whom]$_i$ I am most indebted *WH-trace*$_i$ for my recent success on Broadway.

In each case the comprehension mechanism must be on the look-out for a gap of just the right type to fit the WH-phrase "filler" at the beginning of the clause. If it is lucky, the sentence contains just one gap of the appropriate category, and it is immediately recognizable AS a gap. If so, the processing mechanism can build the correct structure, with an EC in the right position, coindexed to the WH-phrase filler. Then, when semantic interpretation processes occur, the WH-phrase will be interpreted as having the sematic role normally associated with a phrase in the position that the EC is in.

However, in some cases the processing mechanism may not be able to tell, at least on the basis of neighboring words, whether there is a gap in some position or not. In (13) the verb *read* is missing an NP object, so that is where trace must be. But how could the processor establish that? It is not the case that *read* ALWAYS has an object; it can function as an intransitive verb in examples like (20) and (21).

(20) Walter would never admit that he read in the bathtub.

(21) You can always tell [which books]$_i$ Walter read about *WH-trace*$_i$ in the *New York Review*.

In (20) there is no WH-phrase, no movement, and so no trace. In (21) there is WH-movement, but the trace is in a different position, following the preposition *about* rather than the verb *read*. Thus, the fact of the matter is that in (13) there MUST be a trace after *read*, but just looking at the position after *read* does not SHOW that there must be. Its presence there must be inferred. What entails that the trace is after *read* in (13) is that there is no other place in this sentence where it could be. When all impossible positions have been excluded, the only possible one must be the right one. In sentence (21) there are TWO possible positions for an empty NP: one after *read* and one after *about*. Only one of them can be the real trace site. And it is the position after *about* that wins, because it has the greater need: an NP object for *about* is obligatory, whereas *read* can do without one. (Note that English also contains a verbal particle *about* that does not need an object NP, as in *The children ran about all afternoon;* but the preposition *about* that occurs in *read about*, as in (21), MUST have an object.)

These examples illustrate the fact that locating an EC can involve a global inference over a whole sentence, finding and comparing candidate positions. But a sentence is not globally available to the perceiver; it is received one word at a time. This is obviously so in speech; and even in reading, the words are usually identified one after the other. Furthermore, it seems clear that normally we do not wait until we have heard or read an entire sentence before comprehending the beginning of it. This means that the processing routines will often be faced with a decision to make before they have enough information about the sentence to be able to make it. The first eight words of sentences (13) and (21) are identical. At the word *read* the question arises: Is there a trace next? If the processor guesses yes, it will be right about (13) but wrong about (21). If it guesses no, it will be right about (21) but wrong about (13). If it does not guess at all, it will fall behind in interpreting the sentence and may never recover. In some cases the information that resolves the issue arrives only much later. In (22), for example, there is a doubtful trace site after *was reading*, whose status is not resolved for another seven or eight words. In (22a) it is eventually shown to be the true trace position by the fact that the sentence ends without any other possible trace position; in (22b) it is shown NOT to be a real trace position by the fact that there is an undeniable trace after *about* seven words later.

(22) This is the book that Walter was reading
  a. *WH-trace*$_i$ to his friends and fellow students on Friday.
  b. to his friends and fellow students about *WH-trace*$_i$ on Friday.

Thus, we see that the inference from grammatical constraints to sentence structure is often quite intricate, and the facts that should feed it are not always there when needed. How do the processing routines cope?

8.1.3 Ambiguity

The uncertainty that complicates the perceiver's task of detecting ECs is just one instance of a very common problem for the processing routines: ambiguity. There are fully ambiguous sentences such as Chomsky's example (Chomsky 1965) shown in (23).

(23) Flying planes can be dangerous.

In cases of full ambiguity, the linguistic facts do not resolve the meaning; the perceiver must decide on some other basis (topic of conversation, plausibility, knowledge about the speaker) which of the two meanings the speaker intended. Full ambiguity sometimes arises with ECs, as in (24) where the WH-phrase is the PP *to whom* whose trace might be in the clause with *say* (to whom did she say it?) or in the clause with *mailed* (to whom did she mail them?). Nothing in the word string shows which analysis is intended.

(24) To whom did Eloise say she had mailed three postcards?

More common than full ambiguity is temporary ambiguity. That is, for a processor receiving words over time, it may be that some early part of the word sequence is ambiguous but the ambiguity is then resolved by words that follow. The examples (13), (21), and (22) discussed above are temporarily ambiguous with respect to the trace position. Chomsky's example (23) can be turned into a case of temporary ambiguity if we change the verb so that it disambiguates one or the other of the two meanings. In (25), both sentences begin with the words *flying planes* whose structure is temporarily ambiguous (*flying* could be an adjective or a verb), but its structure is subsequently disambiguated by the singular or plural predicate.

(25) a. Flying planes is dangerous.
  b. Flying planes are dangerous.

In (25a) the disambiguator is the singular verb *is*, which requires the meaning that it is dangerous to fly planes; in (25b) the disambiguator is the plural verb *are*, which requires the meaning that planes which fly are dangerous.

Full ambiguity is an obvious threat to successful communication, but even temporary ambiguity can be troublesome for a system that is working at close to full capacity. Research on sentence comprehension has uncovered many varieties of temporary ambiguity. A handful of examples are shown in (26), with the disambiguating word underlined in each. (In some cases, such as f., what disambiguates is the fact that the sentence comes to an end without any more words.) Some of these examples involve ambiguities of trace position, and some involve other sources of ambiguity. In some cases the ambiguity is easy to spot, and in others it is very difficult.

(26) a. The cotton clothing is made of <u>comes</u> from Mississippi.
  b. Sally found the answer to the physics problem <u>wasn't</u> in the book.
  c. The package dropped from the airplane <u>reached</u> the ground safely.
  d. The commander of the army's bootlaces <u>are</u> broken.
  e. They told the boy that the girl shouted at <u>in</u> the playground to go home.
  f. Eloise put the book that she'd been reading all afternoon in the library.
  g. Have the soldiers given their medals <u>by</u> their sweethearts.
  h. He put the candy in his mouth on the table.

Where there is ambiguity, the sentence processing mechanism lacks guidance as to what structure to build. However, experimental data and perceivers' judgments on sentences like those in (26) suggest that the processor does not just grind to a halt when it encounters an ambiguity; rather, it makes a guess. The characteristic sign of a guessing system is that sometimes it wins and sometimes it loses. If the sentence happens to end in a way that fits the guess, processing will be easy—in fact, just as easy as if there had been no ambiguity at all. But if the sentence happens to end in a way that fits the other structure—the one the processor did not guess—then there will be trouble later on; at the disambiguation point the structural analysis of the sentence will be impossible to continue, and the processor will have to back up and try the other analysis instead. In psycholinguistic parlance this situation is called a "garden path": The processor makes a mistake and proceeds blithely on, not realizing there is any problem until later, when things take a sudden turn for the worse. The examples in (26) are all garden path sentences; that is, they all end in the unexpected direction, and the processing routines exhibit some distress (though in varying degrees) on encountering the disambiguating word. Consider (26a) (from Marcus 1980). The first six words are temporarily ambiguous. If followed by *expensive handwoven fabric from India*, the sentence is easy to process; there is no garden path. The sentence is about some cotton clothing and tells us that it is made of expensive stuff. In (26a) the same six words are followed by *comes from Mississippi*, and the sentence is extremely difficult to process—so much so that it may appear at first to be ungrammatical. The sentence is about cotton from which clothing is made, and tells us where it comes from. The fact that the first way of ending the sentence is easier to process than the second is our evidence that the processor makes a guess about the structure of the subject NP at the beginning, before it encounters the disambiguating information later on. It guesses that the structure is *[the cotton clothing]*, rather than *[the cotton]* plus a relative clause. Interestingly, this is true for virtually all perceivers, so it is not a matter of individual experience but reflects some basic fact about the way the human brain works.

One of the projects of psycholinguistic research is to map out the structural guesses that the sentence processor makes, by establishing which sentence completions are easy and which are difficult for all sorts of temporary ambiguity. From this we can hope to infer what kind of a machine this processor is. The basis for the inference is the plausible assumption that when there are no external restrictions, a mechanism will do what comes most naturally to it. By this logic, sentence (26a) could be very helpful in ruling out certain hypotheses about the design of the processor: it would rule out any kind of machine for which the analysis *[the cotton]* + relative clause would be easier to spot, or easier to build, than the analysis *[the cotton clothing]*. This empirical program has been under way for some years. Results show that the human sentence processor's guesses are far from random; they exhibit very consistent general tendencies. With regard to phrasal structure, what the human processor likes best is simple but compact structures, which have no more tree branches than are necessary, and the minimal tree-distance (walking up one branch and down another) between any pair of adjacent words.

With regard to ECs too, all the evidence suggests that in case of ambiguity the human sentence processor does not stop and wait until more information arrives. It makes a guess in order to be able to carry on parsing the sentence, and its guesses are not random. It appears to err systematically in the direction of overeagerness, anticipating ECs before they occur. Sometimes the remainder of the sentence confirms that guess, but sometimes it does not.

8.1.4 Anticipating ECs

Looking back at the examples in (22), test your own judgment about which is easier to parse. For most people the structure for (22a) is computed smoothly, while the analysis of (22b) hiccups at the *about on* sequence that shows the early gap to be wrong. It seems, then, that in processing both sentences, the processor notices the early gap position and likes it; it guesses that this is the real gap for the EC. By good fortune this turns out several words later to be correct for (22a). But in (22b), where a later gap position is correct, this guess causes a garden path.

A number of experimental results support this idea that the human sentence processor is inclined to be over-hasty in postulating traces. For instance, Frazier and Clifton (1989) tested sentences as illustrated in (27).[2] Like the examples above, these have an ambiguous trace site that is disambiguated in two different ways by the words that follow.

(27) a. What$_i$ did the cautious old man whisper *WH-trace*$_i$ to his fiancée during the movie last night?

b. What$_i$ did the cautious old man whisper to his fiancée about *WH-trace*$_i$ during the movie last night?

The correct trace positions are marked in (27). The doubtful trace position is after the verb *whisper*. The verb *whisper* (like the verb *read* in earlier

2. These are just examples of the sentences tested. In this experiment, and in all the others discussed in this chapter, many sentences of the same general type are tested, in order that statistical tests can be made, to distinguish chance performance from the phenomena of interest. Many other aspects of the linguistic materials and experimental procedure must also be carefully controlled. Details will not be discussed here but can be found in the original experimental reports in the articles referenced.

examples) sometimes has an object (*I whispered a message to my friend*) and sometimes has no object (*I whispered to my friend*). So when *whisper* appears without an overt NP following it, this might be because it has an EC as its object, or because it has no object at all. In (27a) it has an EC object; in (27b) it is intransitive, with no object at all. The problem for the processor is that word by word the two sentences in (27) are identical all the way up to *fiancée*; thus, when it encounters the word *whisper*, it has absolutely no way of knowing whether there is a trace after it or not. Frazier and Clifton's experiment was designed to tell us what the processor does in such a circumstance where the word string is not informative. If the processor guesses that *whisper* has a trace as its object, it should find the last part of sentence (27a) easier to cope with than the last part of sentence (27b). If it guesses that *whisper* has no object at all, it should find the rest of (27b) easier to parse than the rest of (27a). If it does nothing but just sits and waits for more information, it should find the two endings more or less equally easy.

The sentences were presented to subjects visually, on a CRT screen, one chunk (a few words) at a time, in what is called a "self-paced reading" paradigm; that is, the subject pushes a button after reading each segment of the sentence, to call up the next one. The subject can spend as long as is necessary to read each segment, but the amount of time taken is recorded by a computer and is used by the experimenter as a measure of how difficult that part of the sentence is to process. (The reading time is influenced, of course, by many properties of the words and the sentence structure, but this is no problem as long as other factors are held constant and only the factor of interest in the experiment is varied.) In Frazier and Clifton's experiment, subjects were faster at reading sequences like *to his fiancée during the movie* in (27a) than sequences like *to his fiancée about during the movie* in (27b). This is not very surprising since the former sequences are one word shorter; but the advantage for the (a) type was present even when the reading times were adjusted for the length difference. Thus, the experimental results provided confirmation of the intuitive judgment that early trace sentences like (27a) are easier to understand than are late trace sentences like (27b). It seems that if the processor has a filler on its hands, it is inclined to jump at the first gap it finds. Frazier and Clifton call this the Active Filler Hypothesis: "When a filler has been identified, rank the option of assigning it to a gap above all other options."

To sum up so far: We have seen that the human sentence processing system is capable of making mistakes about ECs. Because they are not signaled explicitly in the input, it can be difficult sometimes to distinguish between an EC and a nothing at all. The processor is at the mercy of each sentence to provide some cues from which it can deduce the difference between an EC and a nothing. If a sentence does not provide the information, or provides it too late, the processor has to resort to guessing. And, inevitably, roughly half the time it will guess wrong. It may recover itself later so that ultimately the sentence is understood correctly, but time and effort are wasted on the garden path; thus, processing is less than perfectly efficient. However, to be fair to this mental processing system that we are all equipped with, it should be noted that ANY system—except one with the leisure to sit back and wait for the needed information to arrive—would be bound to make mistakes on temporarily ambiguous sentences. Even a computer, with its vast memory and computational resources, has significant trouble with ambiguity. What is impressive about the human language processor is how FEW mistakes it makes in general. When relevant information IS available, we see over and over again that the processor makes extremely efficient use of it.

8.1.5 Using Linguistic Information

As an informal demonstration of how the processor takes advantage of information available, consider its performance on sentence (28).

(28) What$_i$ are boxes$_j$ easy to store *EC*$_j$ in *WH-trace*$_i$?

This is a very easy sentence to understand; yet, it contains two ECs only a word apart, each of which has to be assigned an antecedent. The indices in (28) show the correct antecedent–gap pairings. The WH-word *what* has moved from its underlying position as the object of the preposition *in*. There is also an understood object of the verb *store*, associated with the subject NP *boxes*. (Linguists analyze this association as involving WH-movement also, where what moves is a phonologically empty element, somewhat as in the case of the empty relative pronoun in (9)/(10) above. But the details are not important here.) Notice how clear it is that the question in (28) is about storing boxes in something, not about storing something in boxes: a good answer would be "closets," not "pencils." There seems to be no temptation at all to construe (28) the wrong way round, about storing things in boxes. This means that the processor has not only found the right fillers and gaps in (28) but has also matched them up the right way. In principle, either filler could go with either gap, but there is a constraint in the grammar that insists on one pairing and prohibits the other. This constraint requires two filler-gap relations in a sentence to be nested one within the other, not intersecting. Thus indices i j j i as in (28) are acceptable, but i j i j would not be. This Nested Dependency Constraint is the reason why (28) is NOT ambiguous but only refers to storing boxes in something else. Thus, this is a case where the word string and the grammar do determine, between them, a uniquely correct analysis. And the processor evidently knows it. It has no trouble working out the implications of the constraint for the analysis of this sentence, or in building a structure with two ECs at very close quarters. No wonder, then, that we see no sign of strain in the processor in its dealings with normal questions and relative clauses that have only one gap apiece.

Our informal assessment, then, is that as long as there is no ambiguity, sentence processing seems to be completely smooth and efficient. Experimental data support this impression. Whatever useful information a sentence provides, or the grammar of the language provides, the processing mechanism takes advantage of. Consider a very basic point. Since a WH-trace is created when a WH-phrase moves, there cannot be a WH-gap in a sentence if there isn't any WH-filler. This is potentially useful information; it could save the processor the trouble of going on a gap-hunt in sentences with no filler phrase. Does the processing device take account of it? If it did not, then regardless of whether a filler was present or not, it would process the rest of a sentence in exactly the same way. But, on the contrary, Crain and Fodor (1985) and Stowe (1986) have shown that the processor treats sentences with and without filler phrases differently. Crain and Fodor tested sentences like (29) and (30).

(29) Could the little child have forced us to sing those French songs for Cheryl last Christmas?

(30) Who$_i$ could the little child have forced us to sing those French songs for *WH-trace*$_i$ last Christmas?

In (29) there is no filler; in (30) there is a filler *who* at the beginning. Where is the trace associated with *who*? It must be in a position suitable for an NP. The correct NP position is after the preposition *for*, but as the processor works its way through the sentence, it will encounter various other potential NP positions along the way, and it needs to check each one in case it should turn out to be the right one. Consider the position after the verb *forced* in (30). What appears at that position is the pronoun *us*, and a pronoun constitutes an NP. The fact that this overt NP occurs at this point means that this is NOT the gap position. But for all that the processor knows, it MIGHT have been. This is clear from sentence (31), which is identical to (30) as far as the verb *forced*, but does have a gap right after it.

(31) Who$_i$ could the little child have forced *WH-trace*$_i$ to sing those French songs for Cheryl last Christmas?

As we know from the Active Filler Hypothesis, when the processor is in need of a gap it does not wait to see how things will turn out, but actively seeks out likely sites. The processor might thus ANTICIPATE a gap after *forced* in (30) and then be a little put out when it discovers *us* there, realizing that its hunt for the gap is not over after all. This phenomenon

has been called the "filled gap effect." We would expect it to occur in (30). What we want to know is whether it ALSO occurs in (29), where there is no filler. If it does, this would mean that the processor is not taking full advantage of the grammar of the language, to reduce its workload on-line. But if there is NO filled gap effect in (29), this would show that the processor is smart enough to hypothesize a gap only when there is a filler. This is just what the experimental data reveal.

Subjects were presented with sentences like (29) and (30) in a self-paced reading task, as described earlier for the sentences in (27) except that each "chunk" was just one word; that is, as each word was read, the subject pushed the button to call up the next one. What is of interest here is the reading time for an overt NP such as the word *us* in (30), where a filled gap effect is expected, as compared with that same NP in a sentence like (29), where there is no filler. In fact, the reading time data showed that an overt NP after the verb is significantly more difficult to process in sentences like (30) than in sentences like (29). We can conclude that the processor seeks a gap only when it has a filler to go with it. In general, it seems that the human processing routines are highly sensitive to the grammar. By referring to the grammar, they can postulate ECs just where it is reasonable to do so, not unrestrainedly. A processor that concocted ECs without cause, simply at whim, could see ECs everywhere, between any two words in any sentence. A processor that didn't bother itself with ECs at all would misunderstand many sentences. But the sentence processor in the human brain apparently avoids both kinds of mistake, by paying close attention to the grammatical constraints on the language.

Being sensitive to grammatical constraints presumably exacts a cost in complexity of the computations the processor must perform. But it also confers real benefits in accuracy of comprehension, and all the more so in sentences where the correct structure is not transparently signaled by the overt word sequence. Stowe (1986) investigated the processor's sensitivity to other constraints on the grammatical transformation that creates WH-traces. WH-movement can move a phrase across many clauses in a sentence, as we saw in example (17) above. There is, in fact, no restriction at all on how far a WH-phrase may move between its original position in the underlying structure and its final position in the surface form of the sentence.[3] Nevertheless, a WH-phrase cannot move freely. It can move only "leftward," that is, closer to the beginning of the sentence. It can move

3. There is linguistic evidence which suggests that a phrase doesn't ever move across unbounded stretches of a sentence, but moves in short steps from one clause to the next one. (See Haegeman 1991, chapters 6 and 7, for more information.) This is called *successive cyclic movement*. It affects the linguistic description of the "island constraints" discussed below, but is not important to our present concerns.

only upward, from a more deeply embedded sentence position to a less deeply embedded one. And there are certain kinds of constructions within a sentence that a WH-phrase cannot move out of in any direction. Ross (1967) coined the term "island" for such constructions. The WH-phrase is marooned on the island; it can move within it but cannot escape from it. A subject phrase constitutes an island, as does an adverbial clause, and any complex noun phrase. In theoretical linguistics the important question is what it is about these constructions that makes them islands, when other constructions (for example, object complement clauses) are not islands. In psycholinguistics, our interest focuses on whether the processor KNOWS these island constraints and respects them in its search for traces.

Stowe tested sentences like (32), where the phrase that is bracketed is a subject island; *the silly story about Greg's older brother* is the subject of the verb *was* in the subordinate clause. Therefore, no WH-element is allowed to cross the borders of this phrase. This means there can be no trace INSIDE this phrase whose antecedent is OUTSIDE it.

(32) The teacher asked what$_i$ [the silly story about Greg's older brother] was supposed to mean *WH-trace*$_i$.

In fact, sentence (32) does NOT violate the island constraint, since the WH-element *what* did not move out of the subject island; it started out as the object of the verb *mean* and moved to the top of the question clause introduced by *asked*. In doing so, it passed right over the subject island and was unaffected by it. But consider (32) from the processor's point of view. It proceeds through the words of the first clause, then encounters the filler *what*, and so it initiates a gap hunt. The fact that the subject phrase is an island means that it would be pointless for the processor to look for a gap inside the subject phrase; if the gap were there, the sentence wouldn't be well formed. But does the processing mechanism know that? And does it know it in the heat of the moment, on-line, or does it only catch up with the linguistic constraints later, when the sentence is over and it has had time to get its breath back? Stowe's experiment used the "filled gap effect" to show that the human sentence processor does exploit the island constraint on-line to limit the scope of its gap search. Let's consider the logic behind the design of this experiment.

The island in (32) must not contain the trace associated with the word *what* outside the island; so the preposition *about* inside the island, which needs an NP object, must have an OVERT object. And indeed it does: *about* is followed by the the overt NP *Greg's older brother*. But a processor that did not know (or did not care) about island constraints might anticipate a gap after *about* and then show signs of distress when it encounters *Greg's older brother* there instead. In fact, Stowe did not find any filled gap effect in sentences like (32). She measured reading time for the word *Greg's* in (32) (that is, the first word after *about*, which shows that there is no gap there), and compared it with reading time for the same word in sentence (33). This sentence is exactly like (32) except that it has no WH-filler, so the processor should know that it could not contain a WH-trace.

(33) The teacher asked if [the silly story about Greg's older brother] was supposed to mean anything.

The data showed no difference between (32) and (33). This contrasts with (30) and (29) above, where the findings did show a filled gap effect. Consider what can be inferred from this. We interpreted the filled gap effect in (30) as indication that when there is a filler, the processor contemplates the idea of a gap in any legitimate gap position. So now: the ABSENCE of a filled gap effect in (32) indicates that the processor is NOT prepared to contemplate the idea of a gap in a position that is illegitimate because it is inside an island.

Further experimental support for the precision with which the processor obeys the dictates of the grammar comes from the way in which the processor matches up fillers with gaps. Informally, we noted earlier that sentence (28) is easy to process even though it contains two fillers and two gaps. The fillers arrive first; therefore, when it finds the first gap, the processor has to choose which of the fillers to associate with it. As far as we can tell, it does so unerringly, with no temptation to pick the wrong one. Does the processor ever ignore grammatical constraints and consider an illegal antecedent for an EC? The experimental evidence suggests that it does not.

Consider a WH-trace in a relative clause construction. There is only one right answer as to where its filler is. The filler must be the the relative pronoun (overt or empty) that moved up from its underlying position to a position adjacent to the head noun that the relative clause modifies. This is the only kind of antecedent–trace relation that the grammar allows. For example, in the construction *the rat that Socks chased around the mat*, the trace is after *chased*, and the only legitimate antecedent for it is the relative pronoun *that* at the front of the relative clause. The antecedent cannot be any other NP inside the relative, nor any NP outside the relative. If we want to make a good test of how well the human sentence processor makes use of this grammatical information, we need to construct examples that contain a variety of other tempting NPs, to see whether the processor falls for them. Swinney et al. (1988) tested sentences like (34).

(34) The policeman saw the boy$_i$ [that$_i$ the crowd at the party accused *WH-trace*$_i$ of the crime].

The bracketed phrase in (34) is a relative clause modifying the head noun *boy*. The relative pronoun *that* is coindexed with this head noun, and the trace is coindexed with the relative pronoun; as a result, the trace denotes the boy. Therefore the meaning of the clause is that the crowd at the party accused the boy. It cannot mean that the policeman was accused, or that the crowd was accused, because the trace cannot have *policeman* or *crowd* as its antecedent. But does the processor realize this? Swinney et al. used an experimental paradigm known as *cross-modal priming* to find out. They reasoned that when the processor finds a trace and assigns an antecedent to it, the meaning of the antecedent would be momentarily mentally activated. Thus, if sentence (34) is interpreted correctly, the meaning of *boy* should be active when the word *boy* is processed; then its activation would decline, but it would become activated again when the trace after *accused* is identified. All we need, then, is a method of detecting the mental activation of meaning, at specific points in a sentence.

Swinney et al. went fishing for the meaning of *boy* at the trace position, using a related word as bait. The sentence was spoken; and just as the word *accused* ended and the word *of* began, the word *girl* appeared briefly on a CRT screen. The subjects in the experiment had to respond to the word *girl* by reading it aloud; the time between when it appeared and when a subject began to pronounce it was recorded. Why test with the word *girl*? Because it is a close associate of the word *boy*, and words that are associated tend to "prime" each other. So, if the meaning of *boy* is active at the trace position, responses to *girl* should be facilitated at that position. That is, responses to *girl* should be faster there than responses to *girl* in other contexts, and faster than responses to other words that are unrelated to *boy* (though otherwise comparable to *girl* in length and inherent difficulty). Note that because the sentence is spoken and the word is presented visually, the word can be simultaneous with the relevant part of the sentence, without the two stimuli physically masking each other. This mix of auditory and visual presentation is the cross-modal aspect of cross-modal priming; it is important because it makes it possible to examine sentence processing right at the moment of interest, without having to stop or slow down the sentence in order to do so.

Using this method, Swinney et al. were able to determine how the sentence processing device interprets a trace on-line, that is, which of the NPs in the sentence it assigns as its antecedent. They tested not just the correct antecedent (*boy* in the case of (34)) but also incorrect antecedents (*policeman* and *crowd* in (34)). The results showed that the correct antecedent was activated at the trace position, but incorrect antecedents were not, indicating that in interpreting the trace the processor had respected the grammatical constraints on the language. Thus, once again, it appears that the human sentence processing routines make reliable use of grammatical facts and are speedy enough to be able to do so at the instant at which they become relevant to the sentence at hand.

Consider one final example of the impressive fine-tuning of the processing routines. In another experiment using the same cross-modal priming paradigm, Swinney et al. (1988) tested sentences like (35).

(35) The boxer$_i$ visited the doctor that the swimmer at the competition had advised him$_i$ to see about the injury.

Imagine that the processor has received just the beginning of this sentence so far, up to the verb *advised* as shown in (36). For all that the processor can tell at this point there might be a WH-trace after *advised*. If there were, its antecedent would be the relative pronoun that is associated with the head noun *doctor*, as in the example shown in (37).

(36) The boxer visited the doctor that the swimmer at the competition had advised....

(37) The boxer visited the doctor$_i$ that the swimmer at the competition had advised *WH-trace*$_i$ to see an expert about the injury.

The Active Filler Strategy predicts that the processor will briefly anticipate a trace after *advised*, just until the next word arrives and shows that there is none. Since the trace would be coindexed with *doctor*, and the processor is good at getting its coindexings right, we could expect priming of an associated word such as *nurse* if it were visually presented right after the word *advised* in (35). But then, immediately after that, the pronoun *him* is heard, so the processor discovers that the object of *advised* is actually an overt pronoun, not a trace. And a pronoun is governed by different grammatical constraints than a trace is. Like a trace, a pronoun picks up its reference from a prior phrase. But for a pronoun the grammar stipulates that the antecedent cannot be the subject of the same clause, and it cannot be a relative pronoun. In (35), therefore, the antecedent of *him* cannot be *swimmer* and cannot be *doctor*. That leaves only *boxer* (or, less probably, the pronoun could refer to someone not mentioned in the sentence at all). Thus, once the processor encounters the word *him* in (35), there should be priming of a target word associated with *boxer*, but not of words associated with *doctor* or *swimmer*. The results of the experiment were almost as predicted. Immediately after the verb *advised*, only the meaning of *doctor* was activated. This suggests that the processor guessed (wrongly) that there was a trace at that position, and assigned it an antecedent, without waiting to see what actually came next. Immediately after the pronoun *him*, just a fraction of a second later, there was activation of both *doctor* and *boxer*, but not of *swimmer*. Activation of *doctor* there has no good linguistic explanation, but it is reasonable to suppose that this was leftover activation from the preceding position. The sudden onset of activation of *boxer* at this point suggests that the processor rapidly consulted its linguistic knowledge and computed that *boxer* is the right antecedent for the pronoun (though it was not for the trace). This is a particularly striking result because it shows that the parser's success on other sentences, such as (34), is not due merely to some simple-minded strategy such as "if you need an antecedent, find a relative pronoun." It seems that the processor knows very well that different kinds of dependent elements have different antecedents.

8.2 Are Empty Categories Real?

8.2.1 Linguistic Explanations

The conclusion we have arrived at is that the human sentence processing device is—except when ambiguity trips it up—very quick, very well informed, and very accurate. Because this is so, we can hope to use experimental results on the processing of ECs to answer some linguistic questions about the status and formal properties of ECs. They have been the focus of a great deal of theoretical attention in recent years, and a wealth of facts and interesting generalizations have been established. But there are still some unanswered questions and some conflicting theoretical proposals. Indeed, the very existence of ECs has been challenged. Since they are no more visible to linguists than to other language users, the claim that they exist rests on certain theoretical assumptions about how the OBSERVABLE facts are to be captured. Recall that the two essential properties of an EC are (a) that it has no phonological form and (b) that it is a constituent of a sentence structure with exactly the same syntactic status as any overt constituent. The justification for assuming, despite point (a), that something falls under point (b) is that once ECs are postulated, they are seen to obey the very same general principles that apply to overt constituents. Therefore, some properties of questions, relative clauses, and other constructions containing ECs do not have to be stipulated separately; they follow in large part from known properties of other constructions. This kind of consolidation of facts in two superficially different domains is what warms linguists' hearts.

The most striking similarities are between ECs and overt pronouns and reflexives. (Note: In linguistic discussions, reflexives may also be called *anaphors*.) Pronouns and reflexives share with ECs the property of being dependent for their content (their reference in the world) on their association with an antecedent phrase. The pronoun *her* in (38) is dependent for its content on the antecedent *Sarah*; the reflexive *themselves* in (39) is dependent on the antecedent *the children*.

(38) Sarah$_i$ bowed low when the audience applauded her$_i$.

(39) The new au pair hopes that the children$_i$ won't tire themselves$_i$ out.

Included in the grammar of English (or in Universal Grammar, true of all human languages) are three *binding principles*, which regulate the relationship between a dependent element and its antecedent. And these principles have been argued to apply to ECs just as to non-empty pronouns and reflexives. Principle A applies to reflexives; Principle B applies to pronouns; and Principle C applies to full referential noun phrases (such as *Benjamin Franklin* or *this bathmat* or *the fastest kangaroo in Australia*). The evidence suggests that each of these principles also applies to one kind of EC.

Let's consider Principle C first. It says that a referential NP cannot be dependent for its reference on any other NP. More specifically, a referential NP may not corefer with any NP that stands in a potential antecedent position relative to it. We may think of a potential antecedent as a subject or object that is higher in the structure. (This is very approximate; for details, see Lasnik, 1989, 1990.) This is why in (40), the NP *Thomas* cannot refer to the same person as the higher subject NP *he*. (It can in certain special circumstances, such as if Thomas is unaware that he is Thomas, but we will set aside those cases.)

(40) He$_i$ thinks you deliberately insulted Thomas$_j$.

Chomsky (1976) observed that if we were to assume that Principle C also applies to WH-trace, that would provide an explanation for an otherwise puzzling restriction on coreference in some WH-constructions. Consider the meaning of example (41).

(41) Please tell us who$_i$ he$_j$ thinks you insulted *WH-trace*$_i$.

As one reads the pronoun *he* in (41), it is clear that it refers to someone who has not yet been mentioned in the sentence. And in particular, it is clear that *he* does not corefer with *who*. There's a very clear implication in (41) that one person thinks a different person was insulted. (Contrast this with *Tell me who thinks you insulted him* where the *who* and the *him* CAN refer to the same person.) This requirement of noncoreference in (41) does not follow from any other fact about WH-constructions that we have mentioned so far. But Principle C can explain it. Principle C prevents coreference between *he* and the WH-trace in (41). The structural relation between *he* and the WH-trace is just like the relation between *he* and *Thomas* in (40). The NP *he* stands in a potential antecedent position relative to the trace in (41) and relative to the referential noun phrase *Thomas* in (40). So if WH-trace is like a referential NP in falling under Principle C, coreference will be outlawed in both examples and for the same reason. There is one final step in explaining (41), which is that, as usual, the

WH-trace must corefer with its WH-antecedent, in this case *who*. Since *he* cannot corefer with the WH-trace, it follows that *he* cannot corefer with *who*, either. (Note: It might look as if Principle C would prevent the necessary coreference between the WH-phrase and its trace; in fact it does not, because the WH-phrase is not in a subject or object position but moved into the pre-clause position.)

Another kind of trace that linguists have postulated is known as NP-trace, and it seems to obey Principle A, the binding principle that applies to reflexives. NP-trace is the by-product of a transformational movement called NP-movement, which differs from WH-movement in certain ways: it applies to NPs only; it applies to any NP, not just to WH-phrases; and it moves the NP into subject position. (For more technical details, see Haegeman, chapters 3 and 6.) NP-trace occurs in passive sentences like (42), and in "raising" constructions like (43).

(42) The pie$_i$ was eaten *NP-trace*$_i$ by the guests.

(43) The pie$_i$ seems to me *NP-trace*$_i$ to have disappeared.

NP-trace marks the underlying position of the moved constituent and determines its semantic role in the sentence. Thus, the pie is understood as the object of the eating in (42), and as the subject of the disappearing in (43). In both examples the phrase *the pie* moved from its underlying position into the subject position of the sentence. But NP-movement is much more restricted than WH-movement. NP-movement must be quite local. An example such as (44) is ungrammatical because the moved NP is too far away from its underlying position. (An asterisk indicates an ungrammatical sentence.)

(44) *The pie$_i$ was believed that the guests ate *NP-trace*$_i$.

This fact about NP-movement would be explained if NP-trace is like an overt reflexive and falls under Principle A. Principle A requires a reflexive to have a local (nearby) antecedent. (For the exact definition of "local," see Lasnik 1989, chapter 1). Principle A allows the reflexive *themselves* in (39), where its antecedent is *the children* in the same clause, but disallows it in (45), where *the children* is further away (and the nearer NP *the new au pair* is singular, not plural).

(39) The new au pair hopes $_S$[that the children$_i$ won't tire themselves$_i$ out].

(45) The children$_i$ hope $_S$[that the new au pair won't tire themselves$_i$ out].

Comparing the ungrammatical relationship in (44) and the ungrammatical relationship in (45), we see that they are similar: In both cases the subject of a higher clause is coindexed with the object of a lower clause. So Principle A can account for both examples in the same way. NP-trace and overt reflexives must have local antecedents, and in (44) and (45) the antecedents are not local enough. Here is another case where a fact about movement constructions receives a natural explanation as long as we make the assumption that when a phrase moves, it leaves a silent constituent behind. Without a trace in (44), there would be nothing for Principle A to apply to.

This is a clever form of argument and makes a strong case for the claim that sentences contain elements that cannot be heard or seen. It is not completely decisive evidence, however, as long as there are OTHER satisfactory explanations for the same range of facts. And, indeed, other linguists have proposed alternative explanations, which make no mention of ECs at all, or even of movement transformations. For instance, it has been argued that a passive sentence like (42) is generated just as it appears, with a subject and a passive verb and an optional PP. It is not derived by rearranging the subject and object. Instead, it is assigned a meaning in which the subject of the sentence denotes the patient, and the PP denotes the agent, so that (42) means just the same as the active sentence *The guests ate the pie*. The fact that a passive construction must be local can be explained without assuming that any NP moves from one place to another in the sentence structure, creating a trace. It follows simply from the fact that ALL relationships between a verb and its arguments (its subject and objects) are local. Example (44) is ungrammatical because *the pie* is supposed to be understood as the patient of the eating event, but it is too far from the verb *ate* for that to be so. Thus, this non-movement theory can account for the facts we attributed to NP-movement and Principle A a moment ago. It needs a little more elaboration for examples like (43), but we need not work through that here. (See Bresnan 1978 for the classic presentation, and chapter 1 of Bresnan 1982 for more details.) For our purposes what is important is that the linguistic facts lend themselves to more than one analysis, and it is not easy on purely linguistic grounds to determine which one is correct. Both cover the facts tolerably well, though in quite different ways. Because of this indeterminacy, it would be very satisfactory for the EC hypothesis if psycholinguistic research could offer independent confirmation of the existence of ECs, in just those positions in which linguists have postulated them. And it has been argued, by MacDonald (1989) and by McElree and Bever (1989), that experiments showing that traces activate their antecedents, such as we discussed in the previous section, provide strong support for the EC hypothesis. How could there be antecedent activation by traces if there were no traces?

Is it true that experimental evidence of antecedent activation settles the issue of ECs? Probably not, unfortunately, in any deep sense, though the experimental results may shed some welcome light nonetheless. To be decisive, the experimental evidence would have to reveal the presence of SYNTACTIC ENTITIES at the postulated trace positions. But the antecedent activation paradigms that psycholinguists rely on at present only reveal SEMANTIC INTERPRETATION. They show that perceivers are UNDERSTANDING the sentences correctly, that they are establishing the right relationships between verbs and their arguments, even though some of those arguments are not in their "canonical" positions but are elsewhere in the sentence. But this fact is compatible with any linguistic analysis that accurately associates form and meaning for sentences, whether the analysis involves ECs or not. Therefore, these experiments, however successful they are on their own terms, necessarily fall short of proving that ECs exist. They do not definitively establish that what we have been CALLING "antecedent activation" effects really ARE antecedent reactivation effects; that is, that they are the result of perceivers mentally computing a syntactic structure in which there is an EC that is coindexed with an antecedent. Perhaps one day we will find ways of achieving this more interesting goal. In the meantime, though, we may be able to make some progress toward it if we can establish some of the finer parameters of the comprehension process.

The experiment by MacDonald addressed NP-trace. Experimental data on NP-trace are particularly welcome, since NP-trace is more vulnerable to skepticism than WH-trace is. The linguistic arguments for ECs work best for WH-trace, but do not give much of a grip on NP-trace. Let's consider just two kinds of argument. First, there are facts which suggest that though an EC has no phonological content, it may nevertheless have phonological effects. It can get in the way of other phonological processes, such as the merging of *want* and *to* into *wanna*, in colloquial speech. In (46b) the two words have coalesced. This is known as *wanna*-contraction.

(46) a. Does Esther want to sing?

b. Does Esther wanna sing?

c. Who does Esther wanna sing "Hello, Dolly" with?
(She wants to sing it with Martin.)

d. Who$_i$ does Esther want to sing "Hello, Dolly" with *WH-trace*$_i$?

e. *Who does Esther wanna sing "Hello, Dolly"?
(She wants Martin to sing it.)

f. Who$_i$ does Esther want *WH-trace*$_i$ to sing "Hello, Dolly"?

g. There hafta be two student representatives on every committee.

h. There$_i$ have *NP-trace*$_i$ to be two student representatives on every committee.

Contraction is also possible in the question (46c), whose structure is shown in (46d). But for most speakers *wanna* is not acceptable in (46e). The structure of (46e) is shown in (46f), and an interesting difference from the other examples is apparent. In (46e,f) the *want* and the *to* are not adjacent; they are separated by a WH-trace. This offers a good explanation for why contraction is blocked here; phonological merger can only occur when words are in close contact. But of course, this explanation is available only if WH-trace is assumed to exist as a genuine entity in linguistic structure. For NP-trace there is no such phonological evidence. Relevant examples are not plentiful, but what little evidence there is shows that words on either side of NP-trace CAN coalesce in colloquial speech. An example is shown in (46g). The *hafta* in (46g) is a reduction of *have* and *to*. Contraction occurs here despite the fact that an NP-trace intervenes. The structure of (46g) is shown in (46h); it is a raising construction in which the subject NP (the 'dummy' NP *there*) has moved up from the lower clause, leaving an NP-trace in between *have* and *to*. (Note: We have a terminological tangle here. Anyone who doubts the reality of NP-trace wouldn't want to describe (46g,h) in these terms anyway. But for convenience in discussing the experiments let's continue to talk AS IF sentences contain WH-trace and NP-trace even while we raise the question of whether or not they do.) Some interesting explanations have been proposed to account for why phonological contraction processes care about WH-trace but completely ignore NP-trace. But whatever the reason for it, this fact robs us of a potentially valuable form of evidence for NP-trace. For all that (46) shows, NP-trace might not exist.

Second: a WH-phrase often wears on its sleeve the fact that it started life in the position now occupied by its trace, but an NP-trace does not. For those English speakers (increasingly few) who make a *who* / *whom* distinction, a WH-pronoun is nominative or accusative in accord with its UNDERLYING position before movement. In (47a) it is nominative *who* because it started out as SUBJECT of *is clever* in the lower clause; in (47b) it is accusative *whom* because it started out as OBJECT of *consult* in the lower clause.

(47) a. I know who$_i$ Josephine thinks *WH-trace*$_i$ is clever.

b. I know whom$_i$ Josephine ought to consult *WH-trace*$_i$.

By contrast, the form of the moved NP in a passive or raising construction is always in accord with its SURFACE position. For instance, a pronoun subject of a passive sentence can be nominative *he* as in (48a), but it cannot be accusative *him* as in (48b), even though a pronoun in the underlying position would have been accusative, as shown by (48c).

(48) a. He$_i$ was greeted *NP-trace*$_i$ by the guests.

b. *Him$_i$ was greeted *NP-trace*$_i$ by the guests.

c. The guests greeted him.

Thus, the positive evidence for NP-movement (and NP-trace) is considerably less compelling than the positive evidence for WH-movement (and WH-trace). Furthermore, the possibility of generating the sentence directly, without moving phrases around, is made easier for "NP-movement" constructions by the fact that the supposedly moved NP occupies a perfectly normal subject position, rather than the special before-the-clause position that a WH-phrase appears in, and also by the fact that NP-movement does not span great distances as WH-movement does. In view of this, there are three hypotheses worth considering about ECs: (1) that there are none; (2) that WH-trace exists, but not NP-trace; (3) that both WH-trace and NP-trace are real. Let us see what the experimental data can tell us.

8.2.2 Experimental Evidence

MacDonald (1989) tested "short" passive sentences (passives without an agent *by*-phrase) such as the second sentence in (49).

(49) The terrorists wanted to disrupt the ceremonies.
[The new mayor at the center podium]$_i$ was shot *NP-trace*$_i$.

MacDonald was looking for reactivation, at the NP-trace position, of its antecedent. In (49) the antecedent is the NP *The new mayor at the center podium*. For this purpose she used an experimental paradigm that had already been used to demonstrate antecedent reactivation for overt pronouns. Thus, if the passive sentences gave a similar result, that would be consistent with the hypothesis that a passive sentence contains an element similar to a pronoun, though not directly observable. The task MacDonald used is known as visual probe recognition (VPR). A sentence is presented on a screen; then a single word is presented which the subject must judge as having been, or not been, in that sentence. Many other sentences were presented in which the correct answer was "no," but in all the sentences of interest the correct answer was "yes." The probe word for the passive sentences was always the head noun of the antecedent NP. So for (49), the probe was *mayor*. The first sentence in (49) is there to provide some context and to suggest a plausible agent for the action in the passive sentence that follows. The logic behind this experimental design is much

the same as for the cross-modal priming paradigm discussed earlier. The idea is that when the processor encounters an EC and assigns it its antecedent, the antecedent is mentally activated and so performance on it will be facilitated. Thus, "yes" responses to the probe word *mayor* after (49) should be faster than normal. How fast is normal? To establish that, MacDonald tested similar sentences that had an adjective in place of the passive verb, such as *furious* in (50).

(50) The terrorists wanted to disrupt the ceremonies.
The new mayor at the center podium was furious.

Linguistic theories are all agreed that an adjective is not accompanied by an NP-trace, so (50) contains nothing that should trigger reactivation of the subject noun. The interesting comparison is between the response time for deciding whether *mayor* occurred in (49), where the hypothesis says there is an NP-trace, and the response time for deciding whether *mayor* occurred in (50), where there is no NP-trace. Any difference between them (after the difficulty of particular words is factored out) must be due to the facilitation of *mayor* in (49) due to its reactivation by the NP-trace.

As MacDonald had predicted, probe recognition was faster for the passive verb sentences than for the adjective sentences. McElree and Bever (1989) obtained similar results, for raising as well as passive constructions, using a similar visual probe recognition task. The evidence from VPR thus seems to offer solid support for the existence of NP-trace. Curiously, though, NP-trace has not shown such clear effects in the cross-modal priming (CMP) paradigm that provided evidence of antecedent reactivation for WH-trace in constructions like (34) and (35). Osterhout and Swinney (1993) tested passive sentences with CMP and did not observe significant reactivation until one second AFTER the NP-trace position. This contrasts with results for WH-trace, where the effect is standardly observed when the test word occurs right at the trace position. A second is a long time in sentence processing, so this discrepancy suggests that there is a serious delay in finding an NP-trace or in assigning it its antecedent. So here we have something of a mystery. WH-trace makes a strong showing in CMP. NP-trace does well in VPR but shows up only very weakly in CMP. (For no good reason, WH-trace has not yet been clearly tested with VPR.) Why might this be so? And, especially, what could be the explanation for the apparently contradictory NP-trace results? Do the VPR data overestimate the NP-trace phenomenon, or do the CMP data underestimate it?

This is a research issue still under study, but it seems likeliest that the CMP data provide the more realistic assessment. The CMP task is "on-line." Subjects respond during the processing of the sentence while they are computing the syntactic structure and meaning. By definition an EC is a syntactic entity. And this means that if we are going to see signs of its existence, the best time to observe them should be DURING processing. By the time processing is over, what perceivers retain of a sentence is mostly just its meaning; syntactic details are rapidly lost. (Test yourself: What can you recall of the meaning of the sentence prior to this one? How much of it can you repeat verbatim? That's much harder.) This fact about timing is important to our attempt to make sense of the apparently conflicting experimental results. The VPR task occurs after the sentence is over, and it is basically a memory task. The question "Was this word in the sentence?" taps a representation of the sentence that is the end-result of sentence processing and is now laid down in short-term memory. So, even if an EC had been computed along the way, it would have served its purpose by then and would have been discarded, once its effect on the meaning of the sentence had been established. Because of this, the VPR result for passive and raising sentences tells us only that the subject phrase is more strongly represented in memory when the predicate is a passive verb than when it is an adjective. And though that is evidence of SOME difference between passive verbs and adjectives, it might not be evidence of NP-trace.

Let us take stock here. Arguably, CMP is a better experimental paradigm than VPR for revealing the presence of ECs. But NP-trace does not show up robustly in CMP, especially by comparison with WH-trace, which gives immediate significant effects. This looks very much like evidence for hypothesis (2): that WH-trace exists but NP-trace does not. This is not an unreasonable inference to draw, at least as an interim conclusion pending more experiments in future. But if this is to be the verdict, we must find some other explanation for the memory enhancement that occurred in the experiments on passive and raising constructions. That must have some other source, if it is not due to the presence of NP-trace. If we cannot imagine any other explanation, perhaps we should think twice about giving up the idea of NP-trace.

In fact, a plausible account suggests itself. Passive sentences occur quite infrequently; certainly they are less common than sentences with adjectival predicates. In the processing of a sentence, then, a passive verb would be less expected than an adjective. And there is reason to believe that passive sentences are more effortful to process than active sentences are. (The evidence for this is murkier than one might expect. See Forster and Olbrei 1973 for a summary.) How could this make passive sentences EASIER than adjectival sentences in the VPR experiments? It sounds as if it would predict exactly the opposite. But this is not so, because there is known to be an INVERSE relation between VPR performance on a word, and its expectability and ease of processing in the sentence. Earlier research (not all of it on language) had shown that the greater the effort that goes into processing a stimulus, the more distinctive its subsequent memory representation is. Cairns, Cowart, and Jablon (1981) applied this finding to sentences like those in (51).

(51) a. Kathy wanted a snapshot of my baby, but she unfortunately forgot her camera today.

b. Kathy finally arrived at my baby shower, but she unfortunately forgot her camera today.

In (51a) the word *camera* is more predictable than it is in (51b), because of the reference to a snapshot in the first clause of the sentence. On the basis of the prior research, we would expect that in the context in which it is less predictable, the word *camera* should be harder to process but easier to recall. Cairns, Cowart, and Jablon confirmed this for their test sentences by comparing three tasks: two measures of processing difficulty, and VPR as a measure of short-term memory. In one task, subjects listened to the sentence over headphones and pushed a button when they had understood the sentence well enough to be able to answer a comprehension question about it. Time of understanding was higher for sentences like (51b) with unpredictable words than for sentences like (51a). In another task, subjects listened to the sentences and pushed a button as soon as they heard a word beginning with a specified phoneme (speech sound), such as *t*. This is called *phoneme monitoring*. In (51a) and (51b), subjects told to monitor for *t* would push the button on hearing the beginning of *today*. The heavier the processing load at the point in the sentence at which the phoneme occurs, the slower the subject should be at detecting the phoneme; thus, slow monitoring reveals difficult sentence processing. Note that the *t* of *today* immediately follows the word *camera*, so the phoneme monitoring time should reflect how easy or difficult it is to integrate *camera* into the sentence. The results confirmed that sentences like (51b) were harder to process than sentences like (51a). However, when the same sentences were tested in VPR, with the word *camera* as the probe word for (51a,b), performance was significantly better for sentences like (51b) than for sentences like (51a).

Let us call this the *depth-of-processing effect*. It links three things: the less predictable an item is, the greater the effort of processing it, and the easier it is to recall. This provides a good explanation for why passive sentences fare better than adjectival sentences in VPR experiments. Because passive sentences are less common than adjectival sentences, they receive more attention during processing, so they are easier to probe in short-term memory. To pin this down more tightly, we would like to know if the passive sentences in McDonald's experiment were harder to process than the adjectival sentences were. In fact, no difference between them was observed. However, the only measure available is whole sentence reading time, and possibly this is not sensitive enough to show the difference we are looking for. A different way of checking up on the depth-of-processing explanation would be to confirm that a passive verb is indeed less expected by the processing routines than an adjective is. Wasow et al. (1994) tested the beginnings of MacDonald's sentences, up to the verb *be* (*was* in example (49)), in a sentence completion task. The beginning of the sentence appeared on a screen; the subject immediately read it aloud and continued speaking, completing the sentence in any way that came to mind. Fewer passive predicates were produced than adjectives, by a highly significant margin. If this production test is a good indicator of what perceivers are anticipating as they are reading a sentence, we can conclude that the passive verbs in MacDonald's experiment would have been less expected than the adjectives and hence more memorable. It seems quite likely, then, that the VPR facilitation for passive sentences is really a depth-of-processing effect, a matter of memory enhancement that has nothing to do with NP-trace or ECs at all.

To sum up: At very least, the experimental data seem to be telling us that NP-trace does not have the same status as WH-trace. And though caution is in order, since firm facts are still scarce, the data seem most compatible with hypothesis (2): that WH-trace appears in mental representations of sentence structure, but NP-trace does not. This is not the outcome that was anticipated by McElree and Bever or by MacDonald when they devised their experiments on NP-trace, but it is in keeping with their conception that sentence processing facts can illuminate linguistic issues on which standard linguistic methodology does not definitively pronounce (or has not yet).

Suppose this conclusion holds up as further evidence is gathered in future. How extensive would its impact be on current conceptions of linguistic structure? It could turn attention toward theories of language such as Lexical Functional Grammar and Phrase Structure Grammar, which are not committed to NP-trace and which account for the same language facts by different descriptive means, such as a lexicon that contains passive verbs on a par with active ones. Alternatively, it may be that even within Government and Binding theory, which makes the most extensive use of ECs, it is possible to reassess the need for NP-trace, or at least to reassess its supposed similarity to overt anaphors. In recent years it has been established that ECs (presupposing their existence) are subject to an additional constraint, the *Empty Category Principle* (Chomsky 1981). The ECP applies ONLY to phonologically empty elements; it does not apply to overt lexical pronouns and reflexives. Though motivated primarily by other facts, it turns out that the ECP can account for properties of passive and raising constructions that were previously attributed to the binding principles. This changes the internal dynamics within the theory in ways that

have not yet been fully explored in psycholinguistics, and it may offer some interesting reconciliation of the apparent misfit between theory and experimental results. However, even if the ECP breaks the analogy between NP-trace and reflexives, it does not lessen the theory's commitment to NP-trace. Far from it—the ECP can apply to passives only if there is an EC in passives for it to apply to. And, in fact, the direction this theory is taking in general is toward more ECs rather than fewer. Current analyses assume many more movement operations than previously, such as movement of the verb and the subject in the derivation of every sentence; and every movement leaves a trace. Perhaps it might be argued, by way of defense against the experimental conclusions, that a single NP-trace in a passive sentence would naturally be undetectable amid this throng of other ECs. Such possibilities underline the need for psycholinguistics to keep on polishing up its premises and its methods, so that we can look more and more closely at how sentences are mentally represented. With new experimental techniques, or a new understanding of how theory and experiments should mesh, NP-trace might be banished once and for all, or it might bounce back stronger than ever. This is a lively field of research, so no one would want to bet now on which of these turns it will take over the next few years; it will be interesting to see.

Suggestions for Further Reading

Discussion of empty categories makes up a large part of the literature on syntactic theory for the last ten years and more, particularly within Government and Binding theory. To learn more about the linguistic motivation for ECs, and to see them at work in the syntactic derivations of sentences, it would be advisable to begin with textbook presentations such as Haegeman 1991 or Freidin 1992. These texts contain references to the original books and research articles on which they are based, such as Chomsky 1981, 1986a, 1986b, and they offer advice about how best to move on to reading this more technical literature.

Several linguistic theories do not employ movement transformations and assume either WH-trace only or no ECs at all. For Lexical Functional Grammar, the classic presentation is the paper by Bresnan 1978; a more recent discussion is Kaplan and Zaenen 1989; the most comprehensive collection of related papers is in the book edited by Bresnan 1982. For Head-driven Phrase Structure Grammar (HPSG) and its predecessor Generalized Phrase Structure Grammar (GPSG), a classic presentation is Gazdar 1981; a recent discussion is Sag and Fodor 1994; the most comprehensive presentation is in two volumes by Pollard and Sag 1987, 1994. For references on Categorial Grammar and its modern extensions see chapter 9 by Steedman in the present volume. Textbook level presentations of these and related theoretical positions can be found in Blake 1990, Horrocks 1987, and Sells 1985.

Experimental research on ECs is largely summarized in Foder 1989, which sets it in the context of earlier work on the same kinds of sentences before ECs were "invented." Fodor 1993 updates and extends this survey. Most of the material in the present chapter is drawn from these two papers. The most influential experimental studies have been those by MacDonald 1989, by McElree and Bever 1989, by Swinney, Nicol, and colleagues (reported in Nicol and Swinney 1989), and by Tanenhaus and colleagues (summarized in Tanenhaus, Garnsey, and Boland 1990 and in Boland et al. in press). An interesting recent debate starts with a challenge to empty categories by Pickering and Barry 1991 and continues with articles by Gorrell 1993 and by Gibson and Hickok 1993, and a reply by Pickering 1993. For evidence on the psychological status of ECs that derives from studies of aphasic subjects, see the references in question 8.10 below. A good source of discussion and references on garden-path sentences of many kinds is Gibson (in press).

Problems

8.1 Draw a tree diagram showing the structure of sentence (i). How many ECs does it contain (assuming both WH-trace and NP-trace exist)? What is the deep structure position of the phrase *which pie*?

(i) I have forgotten which pie Bertram said seemed to him to have been nibbled by mice.

8.2 State what category of constituent (NP, N, V, and so on) is "missing" in each of the following sentences:

(i) Pigs can play croquet but rabbits can't.
(ii) Bert was snappish to his mother but he wasn't to his father.
(iii) Tebaldi's recording of Tosca is terrific, but Melba's of Manon is more moving.
(iv) The cow's in the corn but I don't care.
(v) Enjoy!

8.3 For any or all examples (26b–h) in section 8.1.3:

(i) State where the temporary ambiguity begins.
(ii) Describe the two possible analyses of the word string.
(iii) Make up a sentence that starts the same but ends in a way compatible with the other analysis.
(iv) Decide which of the two analyses is easier to compute. (After noting your own intuitions of processing difficulty, try the sentences on some friends to see how uniform judgments are.)
(v) For each analysis, try to make up a sentence which has that analysis but is not even temporarily unambiguous.

Questions for Further Thought

8.1 For any language other than English that you know (or know someone who knows), find as many examples as you can of sentences containing ECs. What are the categories (NP, PP, VP, and so on) of these ECs? Are they traces of movement? Or are they deletions of repeated material (as in problem 8.2, (i)–(iv))? Or are they interpreted by reference to the discourse (as in (v) of problem 8.2)? Does English permit ECs in similar contexts? If not, do the ECs you have found have properties that would make them useful in testing psycholinguistic hypotheses not testable in English? (For example: Is NP-trace in passives more anticipatable on-line than it is in English? Does any EC appear in the same context as an overt pronoun, so that the two could be precisely compared in an experiment?)

8.2 The phonological argument for the existence of WH-trace, based on *wanna*-contraction, has been challenged by Schachter (1984). He observed that when a putative WH-trace is in subject position in a tensed clause, contraction is possible across it. In *Who do you think's clever?* the WH-trace is between *think* and *is*, but nevertheless the *is* can contract to *'s*. Read Schachter's article, and, if you wish, refer also to pp. 374–394 of Fodor (1993) and pp. 435–443 of Barss (1993). Does *think's*—contraction prove that WH-trace does not occur in subject position? that it doesn't exist at all? Is there any way to reconcile *think's*-contraction with the facts of *wanna*-contraction?

8.3 If there is a WH-trace in subject position in sentences like *Who do you think is clever?* (see question 8.2), it should reactivate its antecedent during sentence processing, just as WH-trace in object position has been shown to do. Construct a set of experimental sentences that could be used in a cross-modal lexical decision experiment to determine whether this is so. Design your sentences in matched pairs so that a comparison can be made between cases where there is verb contraction across the putative trace and cases where there is no contraction. Would you expect these to show the same amount of antecedent activation?

8.4 Pickering and Barry (1991) combine facts about center-embedding and facts about WH-constructions, to argue that WH-trace does not exist. (This is much more dramatic than claiming that experimental data DO NOT CONFIRM that WH-trace exists.) Pickering and Barry argue that IF there were WH-traces, a sentence like (i) should be as difficult to process as doubly center-embedded sentences like (4) in section 8.1.1.

(i) John found the saucer [on which]$_i$ Mary put the cup [into which]$_j$ I poured the tea WH-trace$_j$ WH-trace$_i$.

(4) The rat the cat the dog worried chased ate the malt.

But this is clearly not so. Sentence (i) is long but it is relatively easy to comprehend, while sentence (4), though shorter, is more difficult. Read the article by Pickering and Barry, and, if you wish, refer to the responses to it by Gorrell (1993) and by Gibson and Hickok (1993). Is this sentence processing argument against WH-trace valid? If not, why not? If so, can linguistic theory capture the facts of WH-constructions without WH-trace?

8.5 We have seen that the processor copes very efficiently (ambiguities aside) with ECs. So now it is not so clear, after all, that what pushes the processor beyond its limits is a heavy ratio of structural nodes to overt words (see section 8.1.1). So now it is not clear why the doubly center-embedded sentence (4) is so strikingly difficult to process. Can you explain this? Many ideas have been proposed over the years. Read Gibson (in press, chapter 3, and, if you wish, also chapter 4); look up some of the earlier works that Gibson cites. Is there any one explanation of the center-embedding effect that is convincing?

An interesting clue is that ungrammatical sentences with three NPs but only two VPs, such as (i), tend to be perceived (incorrectly) as more grammatical than sentences with three NPs and three VPs, such as (4).

(i) *The rat the cat the dog chased ate the malt.

(4) The rat the cat the dog worried chased ate the malt.

Does any account of the center-embedding effect explain this fact? If so, does it predict WHICH of the three NPs in (i) lacks a VP in (i)?

8.6 Visual probe recognition (VPR) experiments show facilitation for the head noun of the antecedent of an NP-trace (the noun *mayor* in (49) in section 8.2.2). It has been assumed that this is because the antecedent is reactivated at the trace position. If this explanation is correct, then a word that is in the sentence but NOT in the antecedent phrase should NOT be facilitated. The trace explanation would be undermined by across-the-board facilitation for the whole sentence. MacDonald (1989) checked for facilitation of a non-antecedent word (a word denoting the understood agent of the passive, such as *terrorists* in example (49)). She found a very slight effect that might have been due to facilitation but was not statistically significant. This is difficult to evaluate; perhaps facilitation occurred but was attenuated because the test word occurred far back in a previous sentence. Make up some sentences to put this issue to a more stringent test. This is more of a challenge than it may sound. One way to do it would be to compare responses to the same word when it occurs (a) in the antecedent phrase in a passive sentence and (b) in a non-antecedent constituent of a passive sentence. Ideally, the contexts in which the target word appears should otherwise be as similar as possible, the sentences should be equally plausible, and so forth. And the target word should be the same distance from the end of the sentence iny each case so that neither is nearer to the probe task than the other is. (Can all these criteria be satisfied at once?)

8.7 Evidence for the psychological reality of ECs may be provided by the pattern of deficits in aphasic speech, where brain damage has impaired the normal functioning of the grammar and/or the language processing systems (see chapter 13 of this volume). Read Caplan and Hildebrandt (1988) and Grodzinsky et al. (1993), and assess the evidence they provide for WH-trace, NP-trace, and other ECs. Read Zurif et al. (1993) to see how experimental methods outlined in this chapter can be adapted to the study of aphasic speech.

References

Barss, A. (1993). Transparency and visibility: Sentence processing and the grammar of anaphora. In G. Altmann and R. Shillcock, eds., *Cognitive models of speech processing: The second sperlonga meeting.* Hillsdale, NJ: Lawrence Erlbaum.

Blake, B. (1990). *Relational grammar.* London: Routledge.

Boland, J. E., M. K. Tanenhaus, S. M. Garnsey, and G. N. Carlson (in press). Verb argument structure in parsing and interpretation: Evidence from WH-questions. *Journal of Memory and Language.*

Bresnan, J. (1978). A realistic transformational grammar. In M. Halle, J. Bresnan, and G. A. Miller, eds., *Linguistic theory and psychological reality.* Cambridge, MA: MIT Press.

Bresnan, J., ed. (1982). *The mental representation of grammatical relations.* Cambridge, MA: MIT Press.

Cairns, H. S., W. Cowart, and A. D. Jablon (1981). Effects of prior context upon the integration of lexical information during sentence processing. *Journal of Verbal Learning and Verbal Behavior* 20, 445–453.

Caplan, D., and N. Hildebrandt (1988). Disorders affecting comprehension of syntactic form: Preliminary results and their implications for theories of syntax and parsing. *Canadian Journal of Linguistics* 33, 477–505.

Chomsky, N. (1965). *Aspects of the theory of syntax.* Cambridge, MA: MIT Press.

Chomsky, N. (1976). Conditions on rules of grammar. *Linguistic Analysis* 2, 303–351. Reprinted in N. Chomsky, *Essays on form and interpretation.* Amsterdam: North Holland.

Chomsky, N. (1981). *Lectures on government and binding.* Dordrecht, Netherlands: Foris.

Chomsky, N. (1986a). *Knowledge of language: Its nature, origin and use.* New York: Praeger.

Chomsky, N. (1986b). *Barriers.* Cambridge, MA: MIT Press.

Crain, S., and J. D. Fodor (1985). How can grammars help parsers? In D. Dowty, L. Karttunen, and A. Zwicky, eds., *Natural language parsing: Psychological, computational, and theoretical perspectives.* Cambridge: Cambridge University Press.

Fodor, J. D. (1989). Empty categories in sentence processing. *Language and Cognitive Processes* 4, 155–209.

Fodor, J. D. (1993). Processing empty categories: A question of visibility. In G. Altmann and R. Shillcock, eds., *Cognitive models of speech processing: The second sperlonga meeting.* Hillsdale, NJ: Erlbaum.

Forster, K. I., and I. Olbrei (1973). Semantic heuristics and syntactic analysis. *Cognition* 2, 3, 319–347.

Frazier, L. (1985). Syntactic complexity. In D. Dowty, L. Karttunen, and A. Zwicky, eds., *Natural language parsing: Psychological, computational, and theoretical perspectives.* Cambridge: Cambridge University Press.

Frazier, L., and C. Clifton (1989). Successive cyclicity in the grammar and the parser. *Language and Cognitive Processes* 4, 93–126.

Freidin, R. (1992). *Foundations of generative syntax.* Cambridge, MA: MIT Press.

Gazdar, G. (1981). Unbounded dependencies and coordinate structure. *Linguistic Inquiry* 12, 155–184.

Gibson, E. A. F. (in press). *Memory limitations and processing breakdown.* Cambridge, MA: MIT Press.

Gibson, E., and G. Hickok (1993). Sentence processing with empty categories. *Language and Cognitive Processes* 8, 147–161.

Gorrell, P. (1993). Evaluating the direct association hypothesis: A reply to Pickering and Barry (1991). *Language and Cognitive Processes* 8, 129–146.

Grodzinsky, Y., K. Wexler, Y. C. Chien, S. Marakovits, and J. Solomon (1993). The breakdown of binding relations. *Brain and Language* 45, 396–422.

Haegeman, L. (1991). *Introduction to government and binding theory.* Oxford: Blackwell.

Horrocks, G. C. (1987). *Generative grammar.* New York: Longman.

Kaplan, R. M., and A. Zaenen (1989). Long-distance dependencies, constituent structure, and functional uncertainty. In M. R. Baltin and A. S. Kroch, eds., *Alternative conceptions of phrase structure.* Chicago: University of Chicago Press.

Lasnik, H. (1989). *Essays on anaphora.* Dordrecht, Netherlands: Kluwer.

Lasnik, H. (1990). Syntax. In D. N. Osherson and H. Lasnik, eds., *Language: An invitation to cognitive science.* Vol. 1. 1st ed. Cambridge, MA: MIT Press.

MacDonald, M. C. (1989). Priming effects from gaps to antecedents. *Language and Cognitive Processes* 4, 1–72.

Marcus, Mitchell P. (1980). *A theory of syntactic recognition for natural language.* Cambridge, MA: MIT Press.

McElree, B., and T. G. Bever (1989). The psychological reality of linguistically defined gaps. *Journal of Psycholinguistic Research* 18, 21–35.

Miller, G. A., and N. Chomsky (1963). Finitary models of language users. In R. D. Luce, R. Bush, and E. Galanter, eds., *Handbook of mathematical psychology.* Vol. 2. New York: Wiley.

Nicol, J., and D. Swinney (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research* 18, 5–19.

Osterhout, L., and D. Swinney (1993). On the temporal course of gap-filling during comprehension of verbal passives. *Journal of Psycholinguistic Research* 22, 273–286.

Pickering, M. (1993). Direct association and sentence processing: A reply to Gorrell and to Gibson and Hickok. *Language and Cognitive Processes* 8, 163–196.

Pickering, M., and G. Barry (1991). Sentence processing without empty categories. *Language and Cognitive Processes* 6, 229–259.

Pollard, C., and I. A. Sag (1987). *Information-based syntax and semantics.* Vol. 1, *Fundamentals.* CSLI Lecture Notes no. 13. Stanford: Center for the Study of Language and Information (distributed by the University of Chicago Press).

Pollard, C., and I. A. Sag (1994). *Head-driven phrase structure grammar.* Chicago: University of Chicago Press.

Sag, I. A., and J. D. Fodor (1994). Extraction without traces. In R. Aranovich, W. Byrne, S. Preuss, and M. Senturia, eds., *Proceedings of the West Coast Conference on Formal Linguistics* 13, 365–384. Published for Stanford Linguistics Association by CSLI Publications, Stanford, CA.

Schachter, P. (1984). Auxiliary reduction: An argument for GPSG. *Linguistic Inquiry* 15, 514–523.

Sells, P. (1985). *Lectures on contemporary syntactic theories.* Stanford, CA: CSLI.

Stowe, L. A. (1986). Evidence for on-line gap location. *Language and Cognitive Processes* 1, 227–245.

Swinney, D., M. Ford, U. Frauenfelder, and J. Bresnan (1988). On the temporal course of gap-filling and antecedent assignment during sentence comprehension. In B. Grosz, R. Kaplan, M. Macken, and I. Sag, eds., *Language structure and processing*. Stanford: Center for the Study of Language and Information.

Tanenhaus, M. K., S. M. Garnsey, and J. Boland (1990). Combinatory lexical information and language comprehension. In G. T. M. Altmann, ed., *Cognitive models of speech processing: Psycholinguistic and computational perspectives*. Cambridge, MA: MIT Press.

Wasow, T., K. Henniss, L. Re, and J. D. Fodor (1994). Thematic roles and expectations in processing passive sentences. Unpublished manuscript, Stanford University and Graduate Center, City University of New York.

Yngve, V. (1960). A model and an hypothesis for language structure. *Proceedings of the American Philosophical Society* 104, 444–466.

Zurif, E., D. Swinney, P. Prather, J. Solomon, and C. Bushell (1993). An on-line analysis of syntactic proceeding in Broca's and Wernicke's Aphasia. *Brain and Language* 45, 448–464.