

# 8

## Inferences in Comprehension

Herbert H. Clark

*Stanford University*

One evening as Margaret and Jeffrey were relaxing at home, the telephone rang and Margaret answered.

"It's George," she told Jeffrey after a short time. "He's having a party Saturday night and wants us to come. Are you interested in going?"

How should I answer? Jeffrey asked himself. From the way she asked the question, she can't be very anxious to go. Yes, I'd very much like to go, but it wouldn't be fair to drag her to something she doesn't like. I'd better give her a way out.

"It doesn't make much difference to me one way or the other," Jeffrey answered.

His answer is so neutral, he can't really want to go, thought Margaret. And if he doesn't want to go, I certainly wouldn't want to force him. He'd be hard to live with all day Sunday.

"Okay, then, I will make excuses for us," she said, and after a short interchange with George, she said goodbye, hung up the telephone, and sat down.

"George's party sounded as if it would be a lot of fun," she said after a moment. "I'm sorry you didn't want to go."

"*Me* not want to go? But it was *you* who said you didn't want to go. I was just giving you a way out."

The twists and turns of this conversation illustrate nicely how much of what we understand comes from what we infer. Jeffrey had quite legitimately understood Margaret as having said she didn't really want to go, even though the words she used — *Are you interested in going?* — by themselves hardly convey this message. Jeffrey understood what he thought she meant by drawing an inference. If Margaret had genuinely wanted to go, he thought, she would have said *I'd love to go — how about you?* or *We can make it, can't we?* or *It sounds great and Saturday night is free, isn't it?* — a question revealing enthusiasm while asking Jeffrey for his answer. What she did say, however, was *Are you interested*

*in going?* This was a noncommittal question that put the burden of the decision on Jeffrey. And when Margaret put the burden of the decision on him, Jeffrey knew, she wanted him to decide because she wouldn't go otherwise. As a result, Jeffrey decided that she meant she didn't want to go. Margaret's understanding of Jeffrey's answer took just as complicated a route. She saw that his answer, *It doesn't make much difference to me one way or the other*, was noncommittal, and for Jeffrey a noncommittal answer really meant "no." Hence, what he meant here was that he didn't want to go. Although Margaret made the mistake of asking so noncommittal a question in the first place, they were both right from then on to infer what they did. Margaret conventionally *did* mean she didn't want to go when she used a noncommittal question, and Jeffrey conventionally *did* mean he didn't want to go when he used a noncommittal answer.

#### Authorized Inferences

What is illustrated in this conversation is something I will call the *authorized* inference, an inference the speaker intended the listener to draw as an integral part of the message being conveyed. Jeffrey meant Margaret to understand that he didn't want to go to the party when he said *It doesn't make much difference to me one way or the other*. He would judge her as having misunderstood if she hadn't taken it that way. On the other hand, many inferences we draw in conversations are unauthorized, inferences the speaker did not necessarily mean to convey with a particular sentence. I will return to the distinction between authorized and unauthorized inferences later on, for it turns out to be crucial for theories of comprehension.

In the study of comprehension, it is important to discover how we draw authorized inferences as we listen to people talk. Comprehension is best thought of as problem solving. The problem to be solved is, What did the speaker mean? or more accurately, What did the speaker intend us to understand by what he said? We solve this problem, as Margaret and Jeffrey's conversation suggests, using three main ingredients:

1. The explicit content of the sentence.
2. The circumstances surrounding the utterance.
3. A tacit contract the speaker and listener have agreed upon as to how sentences are to be used.

Roughly speaking, we take what is actually said (1), register the relevant features of the present circumstances (2), implicitly consult the contract we have with the speaker about what such a sentence would mean under such circumstances (3), and from this deduce the intended meaning of the utterance. To illustrate, Margaret took what Jeffrey actually said (*It doesn't make much difference to me one way or the other*), registered the relevant circumstances (he is answering a question about going to a party; he and I are on intimate terms), and consulted their tacit contract about the use of such sentences (a noncommittal answer

between intimates means a qualified “no”). From the ingredients as she perceived them, she deduced that Jeffrey meant “No, I don’t really want to go to the party.”

This, of course, is only a rough characterization of how we understand what other people say. Before it can have any real substance, we must be able to specify (a) the tacit contracts people have with each other; (b) the inferences these contracts lead to; and (c) the way listeners actually use the contracts in drawing authorized inferences. In this chapter, I will take up a fundamental type of authorized inference and follow it through these three specifications with some care. In the first section, I will take up a tacit contract called the given–new contract, one that Haviland and I have recently discussed in a series of papers (Clark & Haviland, 1974, 1977; Haviland & Clark, 1974). In the second section, I will present a taxonomy – as yet incomplete – of the inferences people draw on the basis of this contract. In the final section, I will discuss some ways in which listeners draw the inferences they do. Now, although I will be concerned with one particular type of inference, its lessons are very broad indeed. Above all else it reaffirms how much of comprehension is problem solving of the most sophisticated kind, for it demands extensive use of general knowledge, subtle judgments about the circumstances present, and in the end just plain skill.

### THE GIVEN–NEW CONTRACT

What I will be concerned with is reference – to objects, events, and states of affairs. When someone says *The man over there is gullible*, he is using part of the sentence, namely, *the man over there*, to refer to a particular object, and he is leaving it to us to infer what object that is. We draw this inference by consulting the content of the sentence (*the man over there*), the circumstances surrounding the utterance (e.g., the direction of the speaker’s gaze), and a tacit contract about the use of noun phrases for reference. On the basis of all this, we decide that he is referring to Gerald, a man standing nearby. Reference of this kind, then, is an example par excellence of authorized inference. The definite noun phrase, however, is not the only linguistic device available for referring, and the referents themselves do not have to be objects. When someone says *What Maxine did was leave*, he is using *what Maxine did* to refer to an act Maxine carried out, and when he says *What Maxine was was brave*, he is using *what Maxine was* to refer to some state of affairs that holds for Maxine.

#### The Function of Given and New Information

All these instances of reference fall under what linguists have called given information as distinguished from new information (see Halliday, 1967, 1970; Chafe, 1970, 1974; Kuno, 1972, 1975; and with the terms *presupposition* and

*focus* in place of *given* and *new*, Akmajian, 1973; Chomsky, 1971; Jackendoff, 1972). In English, each assertion is said to convey given information and new information, a distinction that is obligatorily indicated in its syntax and intonation. The part of the sentence said to convey given information is conventionally required to carry information that the listener already knows or could know. The part of the sentence said to convey new information is conventionally required to convey information the listener doesn't already know but that the speaker would like to get across. In *What Maxine did was leave*, *Maxine did something* conveys given information — the listener is expected to know that Maxine did something — and *Maxine left* conveys new information — the listener is not expected to know that Maxine left. As this example makes clear, the distinction is not reflected in one set of words versus another, but in one set of underlying propositions versus another.

Why are we obliged to distinguish between given and new information? The obvious answer is that they serve an important function in communication. But for them to be useful, the speaker and listener must agree to use them in the conventional way. The speaker must try to construct his utterances so that the given information actually *does* convey information he believes the listener already knows or could know, and so that the new information actually *does* contain information he believes the listener doesn't already know. The listener, for his part, agrees to interpret each utterance on the assumption that the speaker is trying to do this. In short, the speaker and listener agree to hold to what Haviland and I have called the given–new contract, a tacit agreement between the speaker and listener about how given and new information are to be used in sentences. It is only by holding to this agreement that the speaker and listener can gain any advantage from the distinction.

For the listener, the main consequence of this tacit agreement is that he can make use of the given–new strategy. (For more detail, see Clark & Haviland, 1974, 1977; Haviland & Clark, 1974). He is assumed to absorb each assertion into memory in three steps:

Step 1: Identify the given and new information.

Step 2: Search memory for a proposition matching the given information and call it the “antecedent.”

Step 3: Add the new information to memory by replacing the given information by its antecedent.

To see how this strategy works, imagine that the listener is confronted with the sentence *It was Maxine who hit Max*. At Step 1, he divides the sentence into given and new thus:

- |           |            |
|-----------|------------|
| 1. Given: | X hit Max. |
| New:      | X = Maxine |

At Step 2, he realizes that he should already know about an event in which someone hit Max, so he searches memory for such an event. When he finds one,

say,  $E_{31}$  *hit Max* ("some entity labeled  $E_{31}$  hit Max"), he assumes that it must be the event the speaker was referring to with the given information, so he labels it thus:

2. Antecedent:  $E_{31}$  hit Max.

At Step 3, he notes that since  $X$  corresponds to  $E_{31}$ , he should replace the  $X$  in  $X = \textit{Maxine}$  by  $E_{31}$  and add the resulting proposition to memory thus:

3. Add:  $E_{31} = \textit{Maxine}$

By this time his memory has been updated by the new information in this sentence yet has not been cluttered by all the information in the sentence he already knew — the given information. More important, he has added the new information at just that place in memory where the speaker intended it to be added.

### Implicatures

This strategy works smoothly and precisely in many very simple cases. Consider Sequence (A):

(A) I saw someone hit Max. It was Maxine who hit him.

In order to give self-contained examples, I will assume that the only information the listener hearing the second sentence has in episodic memory is the information conveyed by the first sentence. Hence, when the listener applies the given—new strategy to the second sentence, the only place he can search for antecedents is in the information provided by the first sentence. In this instance, the listener will search for an antecedent to  $X$  *hit Max* (the given information of the second sentence); he will find a matching proposition directly expressed in the first sentence; he will call it the antecedent (say,  $E_{58}$  *hit Max*); and after changing the new information to  $E_{58} = \textit{Maxine}$ , he will add it to memory.

It is far more typical, however, for there to be no direct antecedent in memory. Consider Sequence (B):

(B) Max had a black eye. It was Maxine who hit him.

Once again, the listener will search memory for a proposition matching  $X$  *hit Max*, but in this instance he will not find one. *Max had a black eye* simply does not say, or even necessarily imply, that someone hit Max, and Step 2 in the strategy will fail. In instances like this, Haviland and I have assumed, the listener attempts to introduce an antecedent that he connects to information already in memory in a way he thinks the speaker must have intended it to be connected. This process is called *bridging*, and it results in the addition of a set of one or more propositions to memory, a set that is called an *implicature*. In Sequence (B), the listener would most likely add this implicature to memory:

(B') Max had a black eye because someone hit him.

Once he has done this, of course, he has a proposition that can serve as the antecedent, namely, *someone hit Max*, and he can add the new information to memory appropriately.

The implicature in (B'), I argue, is an integral part of the message the speaker of (B) is trying to get across; indeed, it is just as much a part of the message as the information (B) conveys explicitly. My reasoning here is this: If I were to say (B) to Susan Haviland, I would mean for the proposition *X hit Max* in the second sentence to refer to some particular event. More than that, I would not have used the sentence unless I was sure she could figure out precisely what event I was referring to. Of course, I know well that she does not yet realize directly that someone hit Max, and I know she knows I know it. Yet I am confident that if I pretended she knew that someone hit Max, she would be able to figure out just what event I was talking about. I am sure she would reason this way: "Ah, if Clark can pretend that I already know that someone hit Max, he must think that that should be obvious from what I already know. I already know that Max had a black eye, which could have had any number of causes. But given Clark's pretense, he must intend me to think that it was caused by someone hitting Max. Without this inference he could not expect me to identify the referent for *X hit Max*." In listening to ongoing speech, of course, we are not aware of reasoning like this, but that makes it no less plausible. Many implicatures are difficult to account for without such reasoning.

In its most general form, then, the tacit agreement the speaker and listener have with each other on the use of given and new information goes like this:

Given-New Contract: The speaker agrees to try to construct the given and new information of each utterance in context (a) so that the listener is able to compute from memory the unique antecedent that was intended for the given information, and (b) so that he will not already have the new information attached to the antecedent. (Clark & Haviland, 1977, p. 9)

This contract is critical to the listener. It allows him to be confident that the speaker had referents in mind in the first place, and it gives him a way to figure out what those referents are. If the speaker is cooperative and is adhering to the given-new contract, the listener should be able to identify the referents uniquely on the basis of what he can be confident the speaker thinks the listener knows. Elsewhere, Haviland and I have examined some consequences of this contract. Here I will take up the problem of implicatures.

### VARIETIES OF GIVEN-NEW IMPLICATURES

The implicatures the given-new contract leads us to draw as we listen to others talk take many forms. To give some idea of their variety, I will present a brief taxonomy of these implicatures as found in naturally occurring speech. As with any taxonomy, this one is hardly complete. No taxonomy can be truly complete

without a theory to explain it, and there does not appear to be any theory for this taxonomy at this time. All one can hope for is that this taxonomy will suggest the kind of theory to be worked out in the future. As before, I will illustrate each implicature with a two-sentence sequence in which the first sentence is meant to constitute all of the episodic information the listener has available to him as he listens to the second sentence. That is, the only place he can search for antecedents, or bridges to antecedents, is in the information conveyed in the first sentence. Many of the implicatures I will illustrate, however, could just as well have been inferred on the basis of nonlinguistic sources. The two-sentence sequences are simply the most convenient way of displaying the implicatures briefly and clearly.

### Direct Reference

One of the commonest types of implicature is direct reference – when, for example, a noun phrase refers directly to an object, event, or state just mentioned. This type of implicature is so simple and so well known that we may overlook the fact that there is an inference required here at all. Consider the following examples:

#### A. Identity:

1. I met a man yesterday. The man I met yesterday told me a story.
2. I ran two miles the other day. My two-mile run the other day did me good.
3. Her house was large. The largeness of her house surprised me.

#### B. Pronominalization:

4. I met a man yesterday. He told me a story.
5. I ran two miles the other day. It did me good.
6. Her house was large. That surprised me.

#### C. Epithets:

7. I met a man yesterday. The bastard stole all my money.
8. I ran two miles the other day. The whole stupid business bored me.
9. Her house was large. The immensity made me jealous.

#### D. Set membership:

10. I met two people yesterday. The woman told me a story.
11. I met two doctors yesterday. The taller one told me a story.
12. I swung three times. The first swing missed the ball by a mile.

The implicatures for these four categories are straightforward. For the identity in Example (1), the implicature is roughly this:

- (1') The antecedent for *the man I met yesterday* is the entity referred to by *a man*.

Obvious as this inference may be, the listener still has to draw it if he is to properly understand the second sentence in Example (1). It is conceivable that the speaker did not mean this antecedent to be the same entity as that referred to by *a man* in the first sentence, so the listener must make a leap — perhaps only a millimeter leap — in drawing the inference. I call this case identity because the referring expression *the man I met yesterday* contains all and only the information the listener knows about the entity being referred to. Examples (2) and (3) are analogous except that what is referred to in Example (2) is an event and in Example (3) a state.

For pronominalization, the principle is the same except that pronouns use only a subset of the properties that characterize the previously mentioned object, event, or state. In Example (4), the entity *he* refers to would be completely characterized as *the man I met yesterday*; *he* uses an abbreviated characterization that retains only the features “male” and “singular.” The same goes for Examples (5) and (6). In reality, there are many expressions with which one could refer to this man, and they range in progression from the full *the man I met yesterday* down to the sparse *he*: *the man I met yesterday*, *the man I met*, *the man who was met yesterday*, *the man who was met*, *the man*, and *he*. Some of the conceivable expressions in this series — *the person*, *the adult*, and *the met one* — do not work very well in examples like (1) and (4) for reasons I do not fully understand.

Epithets, on the other hand, add information about the referent, as in this implicature for Example (7):

- (7') The antecedent for *the bastard* is the entity referred to by *a man*; that entity is also a bastard.

Epithets turn out to be surprisingly restricted in productivity. Not just any added information will do. Replace *the bastard* in Example (7) by *the rancher*, *the robber*, or *the President*, and there is no longer an obvious implicature. We would normally take *the rancher*, *the robber*, and *the President* as referring to someone other than the man mentioned in the first sentence.

With set membership, reference is made to one or more members of a set. What the given information does is (a) identify the set and (b) provide a way of distinguishing the referent from the rest of the set. The implicature in Example (10) looks roughly like this:

- (10') One of the entities referred to by *two people* is a woman and the other is not; this woman is the antecedent to *the woman*.

In Example (10), the noun phrase *the woman* enables us to infer (a) that its referent belongs to the set referred to by *two people* and (b) that one of the two people is a woman and the other is not. We infer (b) because if both people were women, the speaker of Example (10) could not have expected us to be able to



figure out the referent uniquely. Examples (11) and (12) work along similar lines.

### Indirect Reference by Association

Very often what is referred to is not an object, state, or event mentioned previously, but only something indirectly associated with such an object, state, or event (see Chafe, 1972). The associated pieces of information are sometimes completely predictable from what has been mentioned, but often they are not. Here I will give only three levels of predictability, although in reality these levels very likely lie along a continuum:

#### E. Necessary parts:

13. I looked into the room. The ceiling was very high.
14. I hit a home run. The swing had been a good one.
15. I looked into the room. The size was overwhelming.

#### F. Probable parts:

16. I walked into the room. The windows looked out into a garden.
17. I went shopping yesterday. The walk did me good.
18. I left at 8 p.m. The darkness made me jumpy.

#### G. Inducible parts:

19. I walked into the room. The chandeliers sparkled brightly.
20. I went shopping yesterday. The climb did me good.
21. I left at 8 p.m. The haste was necessary given the circumstances.

The implicatures we are induced to draw here are once again obvious: In Example (13), the implicature would look something like this:

- (13') The room referred to by *the room* has one and only one ceiling; that ceiling is the antecedent of *the ceiling*.

Since every room has one and only one ceiling, the ceiling is a "necessary part" of the room mentioned in the first sentence in Example (13) and is therefore easy to refer to. Similarly, the swing is a necessary event for the home run in Example (14), and the size is a necessary property of the room in Example (15). In Example (16), the implicature would be as follows:

- (16') The room referred to by *the room* has more than one window; those windows are the antecedent of *the windows*.

What makes the windows only a "probable part" is the fact that not all rooms have windows, although they often do. In this example, then, the implicature adds an important bit of information, namely, that the room mentioned does have windows, and this is information that is not found anywhere else.

It is the classification "inducible parts" that shows how critical the given—new contract is for drawing these inferences. As we hear *the room* in Example (19), we could infer that the room has a floor, some walls, and a ceiling and that it might well have windows, furniture, and lights. But we would not infer, for example, that it has chandeliers. Yet this is precisely the inference that is forced by the second sentence in Example (19):

- (19') The room referred to by *the room* has chandeliers; they are the antecedent for *the chandeliers*.

Here, then, is a clear example in which the search for an antecedent has forced us to draw an inference we would not otherwise draw. By the mere act of referring, the speaker of Example (19) has induced us to build and store away in memory a piece of information that is not predictable from anything else.

#### Indirect Reference by Characterization

In many instances, what is referred to is an object that plays a role in an event or circumstance mentioned previously. For example, a murder is an event that requires one or more murdering agents, a murder weapon or instrument of some sort, and a victim. Once a speaker has mentioned a murder, he can refer to objects that play these roles as long as he characterizes the roles clearly. As with associated parts, roles can vary from complete predictability—all murders require a victim—to almost complete unpredictability. I will nevertheless give just two levels here:

##### H. Necessary roles:

20. There was a murder yesterday. The victim was a terrorist.
21. I went shopping yesterday. The time I started was 3 p.m.
22. I trucked my trunk to New York. The truck was quick.

##### I. Inducible roles:

23. John died yesterday. The murderer got away.
24. John was murdered yesterday. The knife lay nearby.
25. John went walking at noon. The park was beautiful.

The implicatures for "necessary roles" are just as we should expect. The implicature for Example (20), for example, is roughly as follows:

- (20') The event referred to by *a murder* had a murder victim; that person is the antecedent for *the victim*.

Examples (21) and (22) have similar implicatures. For the "inducible roles" we are forced to make broader leaps of inference. The implicature for Example (23) might look like this:

- (23') Some one person caused John to die yesterday; that person is the antecedent for *the murderer*.

In Example (24), we infer that the murder weapon was a knife, not a gun, a bomb, or poison. And in Example (25), we infer that John went walking in a park.

These two categories — necessary and inducible roles — cover a lot of ground, for in English, noun phrases can characterize the roles they refer to in many ways. Some have been illustrated already. The noun *victim* characterizes the role it refers to by itself, since *victim* means “person to whom something bad has happened.” But the same effect could have been achieved with a noun plus restrictive adjectives or relative clauses, as in *the one who was murdered*, *the one who died*, or *the dead man*. Similarly, *the murderer* in Example (23) is explicitly agentive, meaning “the one who did the murder,” although it could be replaced by *the one who did the murder* or something similar. Of course, the more fully the referent is characterized, the easier it is to infer its identity.

It is not always easy to separate “parts” from “roles.” In Example (24), the knife was classified as a “role” in the action of murdering, not as a “part.” An event, like a murder, may have another event, like stabbing, as a “part,” but it cannot have a concrete object as a “part.” Concrete objects play “roles” in events. Ultimately, this distinction may break down, for one could argue that stabbing, though an event, plays a role in a murder just as a knife does: The stabbing is the cause of the death. For now, it is convenient to retain the distinction. Concrete objects are “parts” of other concrete objects but play “roles” in events or states.

### Temporal Relations

In many sentences — perhaps most — the given information refers back to an event — call it A (for antecedent event) — and A occurred at a particular point or period in time. Now, if A has not been mentioned before, the listener must build some temporal relationship between A and an event that *has* been mentioned before; if he does not, he will not have A anchored in time. Just as we try to anchor each object referred to to some place we already know, so we try to anchor each event referred to to some time we already know. Yet the relationship between A and previously mentioned events need not be merely temporal. It may carry an additional notion of cause, reason, or consequence. There appear to be five major classes of temporal relations:

#### J. Reasons:

26. John fell. What he wanted to do was scare Mary.
27. John went to the party. The one he expected to meet was Mary.
28. John had a new suit on. It was Jane he hoped to impress.

#### K. Causes:

29. John fell. What he did was trip on a rock.
30. John went to the party. The one who invited him was Mary.
31. John had a new suit on. It was Jane who told him to wear it.

## L. Consequences:

- 32. John fell. What he did was break his arm.
- 33. John went to the party. The one he saw first was Mary.
- 34. John met Sally. What he did was tell her about Bill.

## M. Concurrences:

- 35. Max lives in New York. Moritz is crazy too.
- 36. Max lives in New York. Moritz isn't very sane either.
- 37. Alex went to a party last night. He's going to get drunk again tonight.

## N. Subsequences:

- 38. John looked left. John looked right.
- 39. John arrived at the party. He got himself a drink.
- 40. John met Sally. They talked for ten minutes.

The implicatures in Classes J through N are easy to illustrate. A reason is something that answers the question "What for?" and a cause is something that answers the question "How come?" The implicature induced by Example (26) gives a reason:

- (26') John fell for the reason that he wanted to do something; that something is the event being referred to by *what he wanted to do*.

On the other hand, the implicature induced by Example (29) gives a cause:

- (29') John fell because he did something; that something is the event being referred to by *what he did*.

For both reasons and causes, A (the event referred to) occurs *before* the event mentioned previously. In Example (26), John's wanting to scare Mary occurred before his fall, and in Example (29) John's tripping on a rock occurred before the fall. In contrast, consequences occur *after* the previously mentioned event. So the implicature induced by Example (32) is roughly as follows:

- (32') John did something because he fell; that something is the event being referred to by *what he did*.

Thus, Examples (29) and (32) lead to exactly the opposite implicatures: In Example (29), A (the antecedent event) caused John's falling, whereas in Example (32) John's falling caused A.

The class of implicatures called concurrences are usually induced with the help of an adverb like *too*, *either*, *again*, or *still*. The implicature for Example (35) is roughly as follows (see Lakoff, 1971):

- (35') Everyone who lives in New York is crazy; therefore, Max is crazy; this state is the state being referred to by the given information *someone other than Moritz is crazy*.

In Examples (35), (36), and (37), the listener is expected to draw the implicature that being in one state, or the occurrence of one event, necessarily

entails the existence of another state, or the occurrence of another event. In Example (37), the listener is expected to infer that Alex's going to the party last night necessarily entails his getting drunk at that party.

The last class of implicatures, the subsequences, are the easiest to describe but the hardest to explain. In Example (38), we would normally infer that John looked left *and then* John looked right. The second-mentioned event, we would infer, occurred immediately subsequent to the first-mentioned event. But how does this come about? The sentence *John looked right*, with the main stress on *right*, can be divided into given and new information in three distinct ways (see Chomsky, 1971; Clark & Haviland, 1977; Jackendoff, 1972):

- a. Given:     *X* happened.  
New:        *X* = John look right
- b. Given:     John did *X*.  
New:        *X* = look right
- c. Given:     John looked in *X* direction.  
New:        *X* = right

These three divisions correspond to the three questions that *John looked right* can legitimately answer: (a) *What happened?* (b) *What did John do?* and (c) *Which way did John look?* Imagine that the listener takes interpretation (a). In Example (38), since it is given that something happened, the listener must find the event *A* that is being referred to. No event mentioned in the first sentence is compatible with John's turning right, so the listener draws the implicature that *A* occurred after the event mentioned previously. He would have done the same if he had interpreted *John looked right* as in (b) or (c), too. So the implicature of (38) is roughly as follows:

- (38') Something happened after John looked left; that is the event being referred to by *something happened*, the given information of *John looked right*.

Subsequences are extraordinarily common. When someone describes a number of events, allotting one event to a sentence, we normally infer that these events occurred in the order in which they were mentioned. Indeed, children are able to draw this inference before they understand words like *before* and *after* (Clark, 1971), and adults make use of this inference in reconstructing sentences from memory (Clark & Clark, 1968; Smith & McMahon, 1970). Thus, subsequence is the major way by which we sequence events as we hear them. Subsequence can be thought of as having consequence as a special case. Consequence is subsequence plus causality: The first-mentioned event not only precedes the second-mentioned event but also causes it.

From these illustrations, it is clear that the placement of two sentences adjacent to each other can be taken to imply one of at least four relationships — reason, cause, consequence, or subsequence. These are illustrated in the

following four sequences:

- (C) John fell. He wanted to scare Mary.
- (D) John fell. He tripped on a rock.
- (E) John fell. He broke his arm.
- (F) John fell. He stood up.

To figure out just which of the four relationships the second sentence has to the first, the listener must look at information outside the sentences themselves. Probably the most useful information is the plausibility of the four relationships. In Sequence (E), it is implausible for John's breaking his arm to be a *reason* for his falling or for it to be a *cause* of his falling. It is quite plausible, however, that it was not only subsequent to his falling but also caused by his falling. If this is roughly correct, the route by which the listener draws the right implicature — the inference the speaker intended him to draw — is complicated. It goes something like this: The listener assumes that the speaker meant the second event to have some temporal relationship to the first. The relationship he meant could not have been an implausible one, for then the speaker could not be sure that the listener would identify it uniquely. Hence, the speaker must have meant the most obvious, the most plausible, relationship.

### DRAWING INFERENCES

Here, then, is a prime example of authorized inference. The speaker wants to get his message across. As part of that message he needs to refer to objects, events, and states the listener already knows. But he cannot refer to them directly — there is no way of pointing to objects, events, and states in the listener's mind directly. He has to rely on his tacit agreement with the listener, the given–new contract. By this agreement, he is confident that if he builds sentences in a particular way, the listener will infer the identity of the referents he means. The listener views the process in much the same way. He assumes that the speaker wants to refer to objects, events, and states and that the speaker is cooperatively relying on the given–new contract. Thus, he, the listener, can infer the identity of referents and draw other implicatures on the basis of this contract.

So far, I have briefly described the contract and classified the implicatures to which it leads. Here I turn to the process by which these implicatures might be drawn. Before I can do that, however, I must reexamine the notion of authorized versus unauthorized inferences.

#### Authorized and Unauthorized Inferences

Imagine that someone has just heard a passage describing how a washing machine works. For many investigators in psychology and artificial intelligence, this listener would be said to have understood the passage if he had come to know

how a washing machine works as characterized by the passage. For them, understanding a message is synonymous, or virtually synonymous, with understanding the situation the message describes. I want to argue that this view of comprehension is incomplete because it fails to distinguish between authorized and unauthorized inferences — between inferences the speaker meant the listener to draw as an integral part of the message and inferences the listener drew without the speaker's authorization. It fails to distinguish between the intended meaning of a message and the implications of a message, a distinction that we listeners keep straight as best we can. Some examples will make this distinction clearer.

Tina is sitting in an easy chair reading a book when Michael walks into the room, leaving the door open. Tina says to Michael, "It's getting cold in here." She means for Michael to understand this as a polite request to close the door, and he understands her to mean that. Michael has thereby drawn an authorized inference, one he believes Tina meant him to draw.

Basil is a defendant in a murder trial. In his testimony, he says, "I was home at 6 p.m." By this, he means for the jury to think he believes he was home at 6 p.m. But the jury, from other evidence, realizes that Basil has just lied — he could not have been home at 6 p.m. because he was seen 60 miles away at 6:15. The jury further infers that he must have lied about the blood on his tie, too, which he said had come from a steak he had for dinner. The jury finally concludes that Basil must be the murderer. Basil, of course, did not mean this to be part of his message at all. The jury has drawn unauthorized inferences from what Basil said.

Jurgen asks Gisela the question, "Is Gordon a psychologist?" and Gisela replies, "Is the Pope Catholic?" Although Gisela's answer is not a *direct* answer to Jurgen's question, she intends him to understand it to mean, "Definitely yes," and Jurgen understands it as such. The inference that the answer is "Definitely yes," then, is one that she authorized.

Ian has just arrived home after a late night at work, and his wife Maggie calls out, "Is that you, honey?" Now Maggie meant for Ian to take this as a mere question of concern, one to be answered by "Yes" or some such response. But Ian infers something more. He recalls that whenever she has called out like this before, she has worried all evening, watched television for consolation, and been cranky all the next day. He therefore infers that she worried and watched television that night and that she will be cranky the next day. Maggie clearly did not mean to convey this by what she said, and so Ian has drawn unauthorized inferences.

Ned and Jane have just seen a very bad movie, and Jane says to Ned, "Wasn't that terrific?" She means for Ned to take this as a sarcastic exclamation about the movie, and he understands it this way. He has drawn the authorized inference that she thought the movie was terrible.

Susan is in the financial district doing business when she decides to stop at Sam's Bar for a midafternoon cocktail. As she reaches the door, a bartender

steps in her way and says, "Sorry, the bar has just closed." He means her to understand this as a statement of fact — the bar has just closed — plus an apology that he cannot help it. She realizes that this is what he meant and walks away. But she draws further inferences. Since bars do not close until later, he was barring her deliberately. Financial districts are notorious for their male-only bars, so he must have been excluding her because she was a woman. Although he did not mean to insult her, she drew the unauthorized inference that he had. Susan could therefore distinguish between what the bartender meant her to understand and the other things she inferred.

From these examples, it is not hard to see what distinguishes authorized from unauthorized inferences. For each authorized inference, the speaker intended the listener to draw an inference, and the listener realized that the speaker intended him to do so. For each unauthorized inference, there was no such intention on the speaker's part nor any attribution of such an intention on the listener's part. For each authorized inference, the speaker relied on one or another speaker-listener agreement about how sentences are to be used. Tina relied on conventions about indirect requests, Gisela on conventions about indirect answers, and Jane on conventions about sarcasm. For the unauthorized inferences, there were no such agreements. The listeners in these instances drew inferences from the intended meaning together with other information, relying on logical requirements *outside* any conventions about the use of language.

The line between authorized and unauthorized inferences, of course, is not always as clean as these examples suggest. At times we may think an agreement is in force when it is not and therefore believe the inference we draw is authorized when it is not. Speakers can make mistakes too, as when they utter sentences that by normal agreements should be taken as meaning something they did not intend them to mean. Yet the distinction between authorized and unauthorized inferences is one that speakers try to keep track of. For example, Susan knew what the bartender meant her to believe when he said, "Sorry, the bar has just closed," even though she took it in quite a different light.

### Inferring the Identity of Referents

Earlier in the chapter, I briefly described the given-new strategy by which listeners are assumed to integrate the novel information of an assertion into memory. It has three steps:

- Step 1: Identify the given and new information.
- Step 2: Search memory for a proposition matching the given information and call it the antecedent.
- Step 3: Add the new information to memory by replacing the given information by its antecedent.

It is at Step 2 in this strategy that listeners infer the identity of referents, drawing all of the implicatures described in the earlier section on varieties of



implicature. Yet as it now stands, Step 2 is not very illuminating. It appears to involve an intricate process of problem solving in which authorized inferences and the given—new contract play central roles. Though there is little to go on other than the logical considerations discussed so far, one can describe a rough but plausible model for the process by which listeners draw these implicatures.

Like other kinds of problem solving, the process of inferring referents has a goal, a “data base” or store of information, a set of constraints or boundary conditions, and some fundamental mental operations. Briefly, the process requires the following four components:

*Goal:* Identify the referent of the given information.

*Data Base:* Information in previous sentences; real-world knowledge of objects, events, and states.

*Constraints:* Given—new contract

1. The referent matches the given information.
2. The referent is one the speaker believes the listener can figure out from what he knows.
3. The referent is unique.

*Mental operations:*

1. Build a candidate referent from the data base.
2. Check whether a candidate referent conforms to the constraints.

To proceed, the listener sets up the goal of identifying the referent, and attempts to find a way to achieve it. To do this he builds a candidate referent from the data base and checks whether it conforms to the constraints. If it does, he assumes that his goal has been reached and that he has identified the referent. If it does not, he tries another candidate referent, and so on. For some candidate referents he will be forced to add certain assumptions, and if this candidate is accepted, so are the assumptions. This way the listener arrives at the intended referent plus the other inferences he was supposed to draw.

To illustrate, consider a simple case of pronominalization, as in Example (4):

(4) I met a man yesterday. He told me a story.

In the second sentence it is given that there is a singular male entity, so the problem to be solved is “Identify the referent of *he*.” The listener first builds a candidate referent from the data base, say, “the person referred to by *I*,” and checks it against the constraints. As it happens, it fails the first test — “The referent matches the given information” — because *I* and *he* cannot be coreferential. The listener then builds another candidate referent, say, “the person referred to by *a man*,” and then checks the constraints, finds that they are all satisfied, and accepts the problem as solved: *He* refers to the person referred to by *a man*.

A sequence like Example (19) offers a much more challenging illustration:

(19) I walked into the room. The chandeliers sparkled brightly.

Here the listener's goal is, "Identify the referent for *the chandeliers*." As a first candidate referent, he might set up "the person referred to by *I*" and as a second, "the object referred to by *the room*." Both candidates fail and are rejected. With no other explicit candidates to try out, the listener is forced to begin making assumptions. He might note that although there is no obvious relationship between chandeliers and people (*I*) or between chandeliers and walking, there is one between chandeliers and rooms — rooms may have chandeliers in them. He would therefore make the assumption that the room actually did have chandeliers in it and set up these chandeliers as a candidate referent for *the chandeliers*. These assumed chandeliers pass the first test (they are chandeliers), the second test (they are ones the speaker could plausibly expect the listener to think of), and the third test (they are unique). Thus, the listener accepts the chandeliers he assumed to be in the room mentioned as the referent to *the chandeliers*.

It is easy to see how this process leads to implicatures in the other sequences provided earlier. Take Example (10):

(10) I met two people yesterday. The woman told me a story.

If the first and second tests are to be met, the referent for *the woman* has to be one of the two people mentioned in the first sentence and has to be a woman. More interesting, if the third test is to be met, the other person *cannot* be a woman — it must be a man, a boy, or a girl. If the other person were a woman, the referent would not be unique. This of course is precisely the inference we draw in hearing this pair of sentences.

Or take Example (26):

(26) John fell. What he did was trip on a rock.

If *what he did* in the second sentence referred to John's falling, the second sentence would be contradictory. If *what he did* referred to an event subsequent or consequent to John's falling, the second sentence would be implausible. But if *what he did* referred to a cause of John's falling, the second sentence would make good sense. Hence the listener assumes that something happened to cause John to fall and that that something was meant to be the referent of *what John did*. This, too, is just the implicature that was earlier attributed to Example (26).

An important consequence of this process is that implicatures are determinate. To see this, consider Example (35):

(35) Max lives in New York. Moritz is crazy too.

If the listener had the time and inclination, he could conceivably build an indefinitely long series of assumptions linking the second sentence to the first. For example, he might assume the following: Everyone who lives in New York breathes foul air; everyone who breathes foul air develops bronchial diseases; and bronchial diseases always drive one crazy. By this circuitous route, it would

follow that everyone who lives in New York is crazy and therefore Max is crazy. The last clause could serve as a referent to *someone besides Moritz is crazy*, the given information of the second sentence in Example (35). But the listener would never go to all this trouble. He would reject this chain of assumptions because the speaker could hardly have expected him to figure it out and to do so uniquely. Because of the second and third constraints, the listener will always pick the simplest assumption the speaker could plausibly expect him to make. In Example (35), the simplest assumption is the shortest chain: Everyone who lives in New York is crazy. Thus, in drawing implicatures the listener should always take the shortest route since this will make them determinate, unique.

Implicatures, or authorized inferences, differ from ordinary implications, or unauthorized inferences, in just this way. Implicatures are determinate, whereas implications may not be. On hearing *the room* in Example (19),

(19) I walked into the room. The chandeliers sparkled brightly.

we might begin drawing unauthorized inferences that there may be windows, doors, furniture, and people in the room. There is no principled stopping rule for such inferences — we could go as far as our imagination would allow. But when we hear *the chandeliers* and draw the authorized inference that it refers to chandeliers that are in the room, the process is determinate. In this instance there is a principled stopping rule, and it is this: Make the simplest assumption possible. The rule makes good sense. If implicatures were not determinate, the speaker could not consider them an integral part of what he wants to convey, for what he conveys must be determinate.

## CONCLUSIONS

Authorized inferences, or implicatures, I have argued, play a central role in communication. Speakers build utterances expecting their listeners to draw certain inferences, and listeners, in turn, comprehend utterances on the understanding that speakers mean them to draw such inferences. These inferences are therefore an integral part of what speakers try to convey. Indeed, people rarely consider them to be inferences at all — they are simply part of what the speaker said. Nevertheless, drawing inferences clearly takes skill and knowledge — children and eavesdroppers are often unsuccessful. It also consumes mental effort and time (Haviland & Clark, 1974).

In this chapter, I have taken up just one type of authorized inference — the identification of referents. I have argued three major points. First, listeners identify referents by use of an agreement they have with speakers called the given–new contract. Without that agreement they would have no basis for identifying referents — at least the ones the speaker intended. Second, in order to identify these referents listeners have to add certain bridging assumptions —

authorized inferences or implicatures — and these are limited in their variety. Third, listeners arrive at these implicatures through a process of problem solving in which they try out candidate referents, check them against constraints dictated by the given—new contract, and accept the candidate that satisfies them all.

Yet in investigating these inferences I have barely scratched the surface. Many questions remain to be answered and much work is left to be done. Is the taxonomy I have presented complete? How is it to be explained; that is, why these categories of implicature and not others? Do people identify referents by the process I have outlined? If so, how does it work in detail? These are important questions. If the approach I have suggested is roughly correct, their answers will affect the way we think about other aspects of comprehension as well. Whatever the process of identifying referents turns out to be, authorized inferences are here to stay.

#### ACKNOWLEDGMENTS

The preparation of this chapter was supported by Grant MH-20021 from the U.S. Public Health Service. It is based in part on a paper called "Bridging" that I presented at a workshop on "Theoretical Issues in Natural Language Processing," convened by R. Schank and B. L. Nash-Webber in Cambridge, Massachusetts, in June 1975. I thank Eve V. Clark for her suggestions in the preparation of this chapter and Susan E. Haviland for discussion over several years on the issues I have taken up here.

#### REFERENCES

- Akmajian, A. The role of focus in the interpretation of anaphoric expressions. In S. R. Anderson & P. Kiparsky (Eds.), *A Festschrift for Morris Halle*. New York: Holt, Rinehart, & Winston, 1973.
- Chafe, W. L. *Meaning and the structure of language*. Chicago: University of Chicago Press, 1970.
- Chafe, W. L. Discourse structure and human knowledge. In J. B. Carroll and R. O. Freedle (Eds.), *Language comprehension and the acquisition of knowledge*. Washington, D.C.: Winston, 1972.
- Chafe, W. L. Language and consciousness. *Language*, 1974, 50, 111-133.
- Chomsky, N. Deep structure, surface structure, and semantic interpretation. In L. A. Jakobovits & D. D. Steinberg (Eds.), *Semantics: An interdisciplinary reader in philosophy, psychology, linguistics, and anthropology*. Cambridge: Cambridge University Press, 1971.
- Clark, E. V. On the acquisition of the meaning of "before" and "after." *Journal of Verbal Learning and Verbal Behavior*, 1971, 10, 266-275.
- Clark, H. H., & Clark, E. V. Semantic distinctions and memory for complex sentences. *Quarterly Journal of Experimental Psychology*, 1968, 20, 129-138.
- Clark, H. H., & Haviland, S. E. Psychological processes as linguistic explanation. In D. Cohen (Ed.), *Explaining linguistic phenomena*. Washington: Hemisphere, 1974.

- Clark, H. H., & Haviland, S. E. Comprehension and the given-new contract. In R. Freedle (Ed.), *Discourse production and comprehension*. Norwood, N.J.: Ablex, 1977.
- Halliday, M. A. K. Notes on transitivity and theme in English: II. *Journal of Linguistics*, 1967, 3, 199-244.
- Halliday, M. A. K. Language structure and language function. In J. Lyons (Ed.), *New horizons in linguistics*. Baltimore: Penguin, 1970.
- Haviland, S. E., & Clark, H. H. What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior*, 1974, 13, 512-521.
- Jackendoff, R. S. *Semantic interpretation in generative grammar*. Cambridge, Mass.: MIT Press, 1972.
- Kuno, S. Functional sentence perspective: A case study from Japanese and English. *Linguistic Inquiry*, 1972, 3, 269-320.
- Kuno, S. Three perspectives in the functional approach to syntax. In R. E. Grossman, L. J. San, & T. J. Vance (Eds.), *Papers from the parasession on functionalism*. Chicago: Chicago Linguistic Society, 1975.
- Lakoff, G. The role of deduction in grammar. In C. J. Fillmore & D. T. Langendoen (Eds.), *Studies in linguistic semantics*. New York: Holt, Rinehart, & Winston, 1971.
- Smith, K. H., & McMahon, L. E. Understanding order information in sentences: Some recent work at Bell Laboratories. In G. Flores d'Arcais & W. J. M. Levelt (Ed.), *Advances in psycholinguistics*. Amsterdam: North Holland, 1970.

