Linguistically, sentences contain both Given information (what the listener is expected to know already) and New information (what the listener is not expected to know already). According to a proposed Given-New Strategy, the listener, in comprehending a sentence, first searches memory for antecedent information that matches the sentence's Given information; he then revises memory by attaching the New information to that antecedent. To provide evidence for this strategy, we presented subjects with pairs of sentences, where the first (the context sentence) provided a context for the second (the target sentence). The subjects were required to press a button when they felt they understood the target sentences. Consistent with the proposed strategy, Experiment I showed that a target sentence with a definite noun phrase presupposing existence took less time to comprehend when its Given information had a direct antecedent in the context sentence than when it did not. Experiment II ruled out a repetition explanation for Experiment I, and Experiment III demonstrated the same phenomenon for target sentences containing the adverbs still, again, too, and either.
ACQUIRING NEW INFORMATION

It deals with just such integration of information. It rests on the unsurprising assumption that language is primarily used for imparting new information. The speaker’s purpose is to provide new information to his audience, and the listener’s is to extract the new information and integrate it with old information already in memory. Indeed, communication is a cooperative effort between the speaker and listener. The speaker syntactically identifies what he thinks his audience already knows, what we will call the *Given information*, and he similarly identifies what he thinks his audience does not already know, what we will call the *New information*. The listener, for his part, assumes that the speaker is cooperative and has good judgment about what he, the listener, does and does not know. The listener’s strategy, therefore, is to identify the syntactically marked Given and New information, treat the Given information as an address to information already in memory, and then integrate the New information into memory at that point (see also Clark, 1973b).

The syntactic devices the speaker has available for the marking of Given and New information are numerous (Halliday, 1967). The sentence in (1), for example, has roughly the following Given–New structure.

(1) The jokes Horace tells are awful.
Given: Horace tells jokes.
New: Those jokes are awful.

As this analysis indicates, (1) conveys two pieces of information, that Horace tells jokes and that they are awful. But (1) distinguishes syntactically between these two pieces of information. It presupposes the first and asserts the second. The speaker who chooses this sentence, therefore, assumes that his audience already knows that Horace tells jokes. The new information he wants to get across is that the jokes are awful. Restrictive relative clauses in an initial part of the sentence (here, *that Horace tells*) are a common device for conveying Given information. Sentences (2), (3), and (4) have similar analyses:

(2) Elizabeth tells awful jokes too.
Given: Someone else tells awful jokes.
New: Elizabeth tells awful jokes.

(3) It was Einstein who searched in vain for the unified field.
Given: Someone searched in vain for the unified field.
New: That someone was Einstein.

(4) The morning star is the evening star.
Given: There is a morning star (just one).
New: There is an evening star (just one).

Sentence (2) uses the adverbial *too* to distinguish Given and New information; (3) makes use of the so-called cleft construction in which the New information is introduced by *it is*; and (4) uses the definite article *the* to presuppose existence. These are only a few of the possible devices for distinguishing Given from New.

The listener’s success with the Given–New Strategy depends critically on whether the Given information, as so marked by the speaker, actually does match information already in memory. To borrow a term normally associated with pronominalization, the Given information must have an *Antecedent* in memory. If there is no Antecedent, the listener must construct one by elaborating information he already has, or he must construct one from scratch. It is only when the listener finds (or constructs) the Antecedent in memory that he can attach the New information to it, thereby integrating the New information with what he already knows. To illustrate, assume that the listener already knows that Horace tells jokes. This knowledge is coded, along with all sorts of information from both linguistic and other sources, in a complex of interrelated, primitive propositions similar in nature to those posited by Clark and Chase (1972), Kintsch (1972), Rumelhart, Lindsay, and Norman (1972), and others. When told
The jokes Horace tells are awful, the listener will extract the Given information (Horace tells jokes), search his memory structure for the matching Antecedent (Horace tells jokes), and on finding it, attach the New information (the jokes are awful) to this Antecedent. The result is a revised memory structure now containing the New information.

This model leads to several important predictions about the time it takes a listener, or reader, to understand sentences in experimentally manipulated contexts. The basic notion is that the listener will not feel he has fully comprehended a sentence until he has integrated its novel information in memory. So, if the Given information matches something in context, that is, if it has an Antecedent, the listener should be very quick to claim he comprehends the sentence. But if it does not, if there is no suitable Antecedent for the Given information, the listener should be very slow to claim he comprehends the sentence. In such cases the listener will search long and hard and, upon failure, will try another strategy. He may find that the Given information is semantically similar to something in context, so he can construct an Antecedent based on an inferential bridge; or he may find he has to set up a new, separate Antecedent. In any case, these additional steps will take time.

To illustrate such a prediction, imagine that a subject is presented with a sentence like (5) and is asked to indicate when he thinks he fully understands the sentence.

(5) The alligator was his favorite present.

This sentence, of course, presupposes the existence of a particular alligator (see sentence (4)). Our prediction is that the subject will be relatively fast in understanding (5) in such a sequence as (6),

(6) Ed was given an alligator for his birthday. The alligator was his favorite present.

but relatively slow for such a sequence as (7).

(7) Ed was given lots of things for his birthday. The alligator was his favorite present.

The context sentence in (6) makes the existence of a specific alligator clear, and so an alligator can serve as Antecedent to the Given information in the target sentence (5). The context sentence in (7), on the other hand, although still appropriate to the situation, does not mention that there is a specific alligator about. With no direct Antecedent for the Given information in the target sentence, the connection between the two sentences requires an extra inferential step, something like, “Ah, one of those ‘things’ must have been an alligator,” and this should require time.

**EXPERIMENT I**

This prediction was tested in Experiment I, in which subjects were presented pairs of context and target sentences, where the target sentence always contained a definite noun phrase (as in (5)). Half the time the context sentence preceding the target sentence explicitly mentioned the existence of the referent of the noun phrase (as in (6)), and half the time it did not (as in (7)). The context and target sentences of each pair were presented sequentially, and comprehension time was measured by having the subjects press a button when they felt they understood the target sentence. According to the Given-New Strategy, comprehension time of the target sentence should be faster when preceded by a context sentence of the first type (to be called Direct Antecedent pairs) than when preceded by one of the second type (to be called Indirect Antecedent pairs).

**Method**

Experiment I made use of 68 context-target pairs, 34 Direct Antecedent pairs and 34 Indirect Antecedent pairs. Each Direct Antecedent pair, such as

(8) We got some beer out of the trunk. The beer was warm.
was matched with an Indirect Antecedent pair containing the same target sentence and a slightly changed context sentence, such as:

(9) We checked the picnic supplies. The beer was warm.

These context-target pairs were made up to be as natural as possible, and they covered a large number of different topics. The 68 pairs were then divided up into two lists of 34 pairs. Each list contained half Direct and half Indirect Antecedent pairs, with the added constraint that if a Direct Antecedent pair occurred on one list, its matched Indirect Antecedent pair occurred on the other. Half the subjects received one list, and half the other. It was therefore possible to compare the comprehension time of the same target sentence (for example, The beer was warm) in both a Direct Antecedent and Indirect Antecedent context.

The context and target sentences were presented to the subject sequentially in a modified tachistoscope which had two 13 × 7 cm presentation fields, one 1 cm above the other. The context sentence appeared in the upper presentation field and the target sentence in the lower one. When the subject pressed a (black) button, the context sentence appeared in the upper field. Another press of the same button then caused the context sentence to disappear, the target sentence to appear in the lower field, and a clock to start. A press of a second (red) button removed the target sentence from view and stopped the clock. The sentences, typed in elite type, were viewed at a distance of 51 cm and were centered in their respective fields.

Each subject was given 57 such trials, 23 practice trials followed by 34 test trials. He was told that he would be reading pairs of sentences. He was to be sure to read and pay attention to the first sentence in each pair since it would be related to the second. But, he was told, it was the second sentence that we were interested in, and he was to be sure to press the red button as soon as he understood what it meant. Although the instruction, "Press the red button as soon as you understand what the second sentence means," was left deliberately vague (for example, subjects were not specifically told to understand the target sentence in the context of the first) no subject had any difficulty following it.

In this and the two subsequent experiments, the subjects were Stanford University students either paid $2.00 for their services or given credit towards a course requirement in introductory psychology. There were 16 subjects, seven women and nine men, in Experiment I.

Results and Discussion

As predicted, comprehension time of the target sentences was faster in the Direct Antecedent pairs than in the Indirect Antecedent pairs, 835 msec to 1016 msec, a difference of 181 msec. This difference was significant in an analysis of variance that treated both target sentences and subjects as random effects, \( \min F'(1,36) = 6.47, p < .025 \) (see Clark, 1973a). There were no reliable differences between the two lists nor were there any interactions.

Although the results of Experiment I conformed to our predictions, there is one obvious alternative explanation of these results. All of the Direct Antecedent pairs contained the repetition of a noun, while none of the Indirect Antecedent pairs did. Our results might therefore have arisen because of the simple facilitating effect of repetition (Smith, Chase, & Smith, 1973). To test for this possibility we performed Experiment II, in which both Direct and Indirect Antecedent pairs involved repetition.

EXPERIMENT II

In discourse, after a noun has been mentioned once, it is usually necessary to definitize it when it is mentioned a second time. So if we speak of an alligator once, we will thereafter refer to it as the alligator. There are, however, important exceptions to this general rule (see Chafe, 1972), as illustrated in (10).
(10) Ed wanted an alligator for his birthday.

Whereas in most sentences the phrase *an alligator* posits the existence of the beast, the phrase in (10) does not. The fact that Ed wants one does not mean that the speaker had a particular one in mind (except in a rather rare reading). For this reason, the target sentence in (11)

(11) Ed wanted an alligator for his birthday. The alligator was his favorite present.

has nothing for *the alligator* to match up to; there is no Antecedent for *the alligator* in memory. So even though the word *alligator* was mentioned in the context sentence, the connection will require an extra inferential step, something like, “Oh, he must have actually been given an alligator,” and therefore (11) should take a relatively long time to comprehend.

In Experiment II, therefore, we altered each Indirect Antecedent pair used in Experiment I along the lines of (11). If the difference between Direct and Indirect Antecedent pairs in Experiment I is attributable to repetition, that difference should disappear. If, on the other hand, it is attributable to the Given-New Strategy, it should remain.

Method

Experiment II was identical to Experiment I except for two features. Each context sentence in the 34 Indirect Antecedent pairs was altered, along the lines of (11), so that each contained a repetition of the critical noun of the target sentence, but did not posit the existence of the object referred to by the critical noun. The Indirect Antecedent pair replacing (9), for example, was:

(12) Andrew was especially fond of beer.
    The beer was warm.

There were 10 subjects, seven women and three men.

Results

Again as predicted, comprehension time for target sentences was faster for the Direct Antecedent pairs than for the Indirect Antecedent pairs, 1031 to 1168 msec, a difference of 137 msec, \( \text{min} F'(1,23) = 15.7, p < .001 \). There were no reliable differences between the two lists of sentences, nor were there any interactions. These results, therefore argue that mere repetition of the critical noun is not enough to account for the results of Experiment I.

EXPERIMENT III

Experiments I and II together clearly support the Given-New Strategy. Experiment III was designed to test the model vis-à-vis other kinds of presupposition and to see what effect negation has on the process. We chose to examine the types of presuppositions required by certain adverbs, namely, *still, either, again,* and *too,* as in the following examples.

(13) Vivien still knows her Ovid.
(14) Stephen can't dance the rhumba either.
(15) Lance is out hunting mushrooms again.
(16) Annie slew a gnu too.

Sentence (13) presupposes that Vivien used to know her Ovid; this is Given information. The New information is that she knows it now. In (14) it is Given that someone else cannot dance the rhumba, at least in one reading; what is New is that Stephen cannot. The Given information in (15) is that Lance has hunted mushrooms before; the New information is that he is doing it now. Sentence (16) is the positive counterpart of (14). Its Given information is that someone else slew a gnu; its New information is that Annie did.

As in the first two experiments, sentences such as (13) through (16) were paired up with context sentences to form either Direct Antecedent pairs, as in (17), or Indirect Antecedent pairs, as in (18).
(17) Last Christmas Eugene became absolutely smashed. This Christmas he got very drunk again.

(18) Last Christmas Eugene went to a lot of parties. This Christmas he got very drunk again.

In (17) the Given information of the target sentence, that he had got drunk before, is directly expressed in the context sentence (though here in slightly different words). This provides a Direct Antecedent. In (18) the Given information is not directly expressed in the context sentence, although it is consistent with information provided by the context sentence. So the context sentence provides only an Indirect Antecedent.

In this experiment, however, there was a third context of interest, namely one that provided a proper Antecedent, but did so in a negative form. Consider the context-target pair in (19):

(19) Last Christmas Eugene couldn't stay sober. This Christmas he got very drunk again.

The Given information of the target sentence, that he had got drunk before, is expressed by the context sentence, but only negatively. Whereas the target says got drunk, the context says couldn't stay sober. While these two expressions are synonymous, one must deduce that not sober implies drunk, or vice versa, and this operation should take an appreciable amount of time (see Clark & Chase, 1972). Such context-target pairs, to be called Negative Antecedent pairs, should therefore take longer than Direct Antecedent pairs, though not necessarily any longer than Indirect Antecedent pairs, which also require extra deductions.

Method

Experiment III used the same procedure as Experiment I, but used newly composed context-target pairs. There was a total of 144 pairs, 36 each for the adverbs still, either, again, and too. Of these, 48 (12 for each adverb) were Indirect Antecedent pairs, and 48 were Negative Antecedent pairs. For each Direct Antecedent pair there was an Indirect Antecedent pair and a Negative Antecedent pair with the same target sentence; the only differences among the three pairs lay in their context sentences, as illustrated in (17), (18), and (19) above. These 144 pairs were then divided into three lists, each of which contained equal numbers of the four adverbs and equal numbers of the three types of pairs. The lists were constructed so that the three pairs with the same target sentence always appeared on different lists. In actuality, there were also two kinds of Direct Antecedent pairs. Some repeated the critical lexical item from context to target sentence, as in (20).

(20) Fred was right about the new rules. He was right about the old ones too.

Some used synonymous words in the context and target sentences, as in (21).

(21) Fred was correct about the new rules. He was right about the old ones too.

These pairs were included to provide an additional check on the repetition explanation tested in Experiment II.

In Experiment III there were 27 subjects, 18 men and 9 women. Each list was presented to nine subjects, though the same 24 practice pairs were presented to all 27 subjects.

Results

The mean comprehension times, shown in Table 1, fully support the predictions based on the Given–New Strategy. The target sentences were comprehended most quickly in the Direct Antecedent condition (1023 msec) and were comprehended about equally slowly in the Indirect Antecedent (1097 msec) and Negative Antecedent (1088 msec) conditions. In an analysis of variance treating both subjects and target sentences as random effects, this
difference was highly reliable, $\text{min} F'(2,118) = 5.55, p < .005$. The Direct Antecedent sentences were fastest for each of the four adverbs taken separately; the Indirect and Negative Antecedent sentences were approximately equal for all the adverbs but 

\textit{either}, where the Indirect Antecedent condition took about 100 msec longer than the Negative Antecedent condition. Nevertheless, the interaction between the antecedent conditions and the adverbs was not reliable, $F' < 1$. The differences among the four adverbs were only marginally reliable, $\text{min} F'(3,52) = 2.55, p < .10$, and $\text{max} F'(3,52) = 2.79, p < .05$, and could be attributable to slight differences in average sentence length across adverbs.

The results also fail to support an explanation of these differences based on simple word repetition, as tested in Experiment II. Among the 48 Direct Antecedent pairs, there were 25 pairs with a surface repetition, counting \textit{find–found}, \textit{kept–keep}, \textit{lose–lost} as repetitions, and 23 without. Comprehension time for the target sentences averaged 1033 msec for the pairs with repetition and 1014 msec for those without, a difference of 19 msec in the wrong direction, $t(47) = 1.72$, n.s.

\section*{GENERAL DISCUSSION}

As we have argued, sentences are normally comprehended within some larger framework, represented internally by a complex network of interrelated propositions. Employing the Given–New Strategy, the listener takes in a sentence, breaks it into its syntactically defined Given and New information, and then attempts to add the New information to memory. This he does by treating the Given information as a pointer or address to some Antecedent already in memory. That is, he searches memory for a matching Antecedent to this Given information, and on finding it, attaches the New information to the Antecedent. If he cannot find a matching Antecedent, then he must (a) build some sort of bridging structure, (b) treat all information in the sentence as new and begin construction of a new separate structure, or (c) attempt to recompute what is Given and what is New in the sentence.

This model leads to important predictions about the comprehension of pairs of sentences in a laboratory setting. The subject, we assume, will treat the first sentence of such a pair as completely new information and set up a separate structure in memory. But when processing the second, he will try to comprehend it with respect to the information contained in the first. So he will take the Given information of the second and try to find a matching Antecedent in the information provided by the first. If the Antecedent is available directly, then he will comprehend the second (target) sentence quickly; if not, he will require more time, since he will have to construct a bridging structure. Consistent with the predictions, Experiment I showed that comprehension of a sentence containing a definite noun phrase, say, \textit{the beer}, took less time when the context sentence had explicitly posited the existence of some beer than when it had not. Experiments II and III showed that this result could not simply be a result of repetition of the critical noun. And Experiment III demonstrated the same phenomenon for the rather different kind of presuppositions associated with the adverbs \textit{still, either, again, and too}. In particular, it showed that when the target sentence had

\begin{table}
\centering
\begin{tabular}{lcccc}
\hline
& Still & Either & Again & Too & Means \\
\hline
Direct antecedent & 1031 & 1102 & 984 & 976 & 1023 \\
Indirect antecedent & 1058 & 1244 & 1040 & 1047 & 1097 \\
Negative antecedent & 1076 & 1141 & 1063 & 1065 & 1088 \\
\hline
\end{tabular}
\caption{Mean Comprehension Time (msec) for Target Sentences in Direct, Indirect, and Negative Antecedent Pairs Containing Still, Either, Again, or Too in Experiment III}
\end{table}
Given information that required a match with its negative (as with wrong in the target sentence and not right in the context), the target still took longer to comprehend, apparently because of the inference required for the match.

In testing the Given-New Strategy here, we have relied on a subjective criterion of comprehension, how long it takes the subject to feel he has understood a sentence. But as several psychologists have argued (for example, Dooling, 1972; Mistler-Lachman, 1972; Schwartz, Sparkman, & Deese, 1970; Wang, 1970a, b), there may be several different levels of comprehension. At a shallow level we may be able to judge how comprehensible a sentence is without actually comprehending it fully; at a deeper level we may be able to comprehend a sentence without relating it to context; and at a still deeper level we comprehend sentences with respect to context. Obviously, the interest here is in the third and deepest level. The present evidence is consistent with the assumption that most subjects were indeed using this criterion for their judgments on most sentences. If they had used either of the first two criteria, there should have been no reliable differences between conditions, since all conditions contained the identical target sentences. The differences actually found must therefore be attributed to the subject comprehending the sentences at the deepest level. Nevertheless, there may have been some instances of "incomplete" contextual comprehension, and so the absolute differences between conditions cannot be taken very seriously.

The Given-New Strategy is an attempt to formalize part of what people do in integrating new information with what they already know. It therefore ought to apply to previous studies where the main interest was in how subjects went about integrating new information. Consider the study by Bransford and Franks (1971) in which subjects were presented with a number of sentences derived from sets of four propositions (for example, The rock rolled down the mountain; The hut was tiny; The rock crushed the hut and The hut was at the edge of the woods). The presented sentences contained either one, two, or three of these four propositions in various combinations, but subjects were never presented with a sentence containing all four. Later, however, when subjects were given a recognition task, they were found to claim, with the highest confidence, that the sentence containing all four propositions (for example, The rock which rolled down the mountain crushed the tiny hut at the edge of the woods) was old, even though they had never seen it. According to Bransford and Franks, the subjects had built up information structures which eventually contained the same information as the four-proposition distractor sentences, and this is what led them to make their judgments. Rumelhart, Lindsay, and Norman (1971) illustrated this process more explicitly with the use of graph structure notation. Although these are plausible accounts for why the subjects erred on the four-proposition sentences, they nevertheless say little about how the subjects went about building up this information structure.

According to the present view, one important aspect of the Bransford and Franks sentences was that they contained definite noun phrases. As we have argued, the makes specific existential and uniqueness presuppositions and induces the subjects to believe that there is only one hut, rock, mountain, and so on. By the Given-New Strategy, the subject is therefore encouraged to treat each combination sentence as relevant to all others and to attach all instances of the same noun phrase to a common Antecedent. In the Bransford and Franks study, this strategy could have been especially important since four different sets of four propositions were mixed together on the same list. But if Bransford and Franks had replaced every definite article by an indefinite article, the result might have been somewhat different. The subjects should not have been quite so
compelled to build up a single structure for every set of four propositions, and they should have found it a little easier to keep each instance of a noun phrase separate. This is not to say that Bransford and Franks' results were due entirely to the fact that the sentences they used contained definite noun phrases, but rather that definiteness may have been an important contributing factor.

To consider another example, Bransford and Johnson (1973) found that subjects rated the following paragraph as rather incomprehensible:

The procedure is actually quite simple. First you arrange things into different groups depending on their makeup. Of course, one pile may be sufficient depending on how much there is to do. If you have to go somewhere else due to lack of facilities that is the next step, otherwise you are pretty well set. It is important not to overdo any particular endeavor. That is, it is better to do too few things at once than too many....

In addition, subjects found this paragraph very difficult to recall. But when told beforehand that the paragraph was about washing clothes, other subjects found the paragraph quite comprehensible and much easier to remember. Bransford and Johnson argued, rightly, that the topic "washing clothes" helps the subjects to build up some sort of semantic product, but they did not specify how this might be done.

Without the topic "washing clothes," we would argue, someone reading this paragraph simply has no way of constructing the intended Antecedents of each sentence. For example, the procedure in the first sentence presupposes that the subject already knows about some particular procedure. Without the topic, he does not, but with the topic, he does. The second sentence requires the subject to know what things are to be arranged; the third sentence requires him to know what one pile is sufficient for and what there is to do. Each of these involves Given information and requires an Antecedent in memory for proper integration. Without the topic, there is no way of computing the intended Antecedents, but with the topic there is, though some will be indirect. The Given-New Strategy provides a relatively natural account for why subjects only found the paragraph comprehensible when they were given the topic beforehand.

The Given-New Strategy also provides us with a clue about the role of redundancy in language. At first it seems rather odd that sentences should contain so much information that both the listener and speaker already accept as true. Language is a medium for communication, and an efficient speaker ought never to say things his listener already knows. But in light of the Given-New Strategy this redundant information actually becomes a prerequisite for communication. As Given information, it serves as an address directing the listener to where New information should be stored. And as such, of course, it is not really redundant at all. It simply serves a different purpose.

REFERENCES


Wang, M. D. The role of syntactic complexity as a determiner of comprehensibility. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 398-404. (b)

(Received January 14, 1974)