What's what: talkers help listeners hear and understand by clarifying sentential relations*

VIRGINIA VALIAN
CUNY Graduate Center, New York
ROGER WALES
University of St. Andrews

Abstract

It was predicted that a talker would clarify the sentential relations of an utterance if a listener indicated difficulty in hearing and understanding. Subjects read syntactically clear and distorted sentences to a listener (experimenter) in an adjoining room. The experimenter often asked "What?" Subjects changed distorted versions to clear versions, while repeating clear versions essentially as first read. Other subjects were asked to make the sentences clear and simple to understand. The same basic results were obtained. Talkers thus seem to interpret a "What?" partly as a request for clearer sentential relations and respond accordingly. The results indicate that talkers have knowledge of underlying structure. Several alternate explanations can be rejected. A relative derivational theory of complexity is presented.

The present experiment systematically explores the talker's knowledge and use of syntactic structure within the context of an everyday speech situation. Every talker has experienced the phenomenon of saying something to someone who is having difficulty both hearing and understanding what has been said. This can occur, for example, in a noisy restaurant, or when the talker and listener are in different rooms. Listeners commonly signal their difficulty by asking "What?" The talker must then decide how, if at all, to

---

*The experiments were begun while the senior author was Research Associate and the junior author Visiting Professor at the Psychology Department, Massachusetts Institute of Technology, and supported in part by the Public Health Service, Grant No. HD 05168-04 to M.I.T.; we particularly thank M. Garrett for his help. The research was also supported by a grant from the City University of New York Faculty Research Award Program. R. Fienon and D. T. Langendoen gave linguistic advice; L. Goldstein helped with the apparatus.
change the original utterance. We call this a What? situation. The talker's change options are phonetic, syntactic, semantic, or any combination of the three. The syntactic options are the focus of the present experiment.

If it is true that clearly displayed syntactic relations within a sentence will aid a listener's semantic analysis (Fodor, Bever & Garrett, 1974), and if talkers in a What? situation want to maximize ease of comprehension on the part of listeners, then one can hypothesize uniform syntactic behavior by talkers when listeners query "What?" Talkers should clarify the sentential relations of their original utterance.

In many cases, clarification of sentential relations is equivalent to production of a sentence less transformationally removed than the original sentence from the deep structure representation. For example, in (1), the verb and its particle are separated by the noun phrase 'the number'. The fact that 'the

(1) Molly forgot to look the number up before leaving the house
(2) Molly forgot to look up the number before leaving the house

number' is the object of 'look up' is obscured. Placement of the particle next to the verb, as in (2), clarifies this syntactic relation and also results in a sentence which is less transformationally distorted than (1). If a talker had originally uttered (1), (s)he should utter (2) after being queried; if (s)he had originally uttered (2), (s)he should repeat the sentence essentially as first spoken. Behaviour of this sort would indicate that the talker's syntactic knowledge includes the relation between sentences which differ by whether an optional transformation has applied. This relation is not specified at the level of surface structure.

However, not all cases of clarification of sentential relations need be equivalent to removal of one or more transformations. For example, although (4) has clearer sentential relations than (3), (3) and (4) probably
derive from different deep structures, one of which contains more empty nodes than the other. If a queried talker changed (3) to (4), the only syntactic knowledge that could be imputed is that a sentence with an explicit subject presents clearer (because fuller) syntactic relations than a sentence without an explicit subject. (When an explicit subject is used, 'do' support is obligatory.)

Although there is evidence to support the claim that listeners are sensitive to underlying syntactic relations (Blumenthal & Boakes, 1967; Bever, Lackner & Kirk, 1969; Levelt, 1970) and that certain tasks are facilitated if listeners are presented with less distorted sentences (Hakes, 1972; Fodor &
Garrett, 1967; but see Bock & Brewer, 1974 for contrary results), there is no experimental evidence that talkers have implicit knowledge of deeper syntactic structure (though see Jarvella, 1972; Garrett, in press), nor that this knowledge could be employed when listeners signal a need for more explicit syntactic relations.

In Experiment I subjects read sentences exemplifying many different linguistic constructions. Half the sentences had relatively clearly displayed syntactic relations, half were relatively distorted; all were grammatical and acceptable. Subjects were told that the experimenter wanted to simulate a What? situation, which was briefly described. All experimental sentences were then queried. The prediction was that subjects would change distorted versions to clear ones and would repeat clear versions. In Experiment II the What? situation was not simulated, but subjects were asked explicitly to clarify and simplify what they had read in order to make it easier to understand. This change of instructions was used to determine if the behavior evoked in Experiment I could also be evoked by more self-conscious instruction. There was the same prediction as in Experiment I because it was hypothesized that talkers interpret a What? in part as a request for syntactic simplification and clarification.

In Experiment III subjects were given both forms of each sentence and asked to choose the simpler. This control condition tested whether an unspecified notion of simplicity would yield the same results as Experiments I and II; in these instructions there was also no mention of talkers or listeners. The prediction was that subjects would show a different pattern of results from Experiments I and II, because the task engages neither the subjects’ natural mode of responding nor delineates the relevant dimension of simplicity.

The formulation presented here can be seen as a relative, rather than absolute, derivational theory of complexity. An absolute derivational theory of complexity states that one sentence is more psychologically complex than another if there are more transformations in its derivational history (Miller, 1962; Mehler, 1967). The two sentences being compared are not required to have the same deep structure representation, nor is a distinction made between optional and obligatory transformations. A relative derivational theory of complexity, on the other hand, would require of the sentences to be compared that they be derived from the same deep structure, and would claim that the psychologically more complex sentence had one or more optional transformations in its history. As sketched, the theory does not specify what psychological complexity is, and is in any event a theory about listening. Further elaborations of the theory are needed to account for talkers’ behavior. One hypothesis is that talkers tacitly know that more
transformed sentences are harder for listeners to process; when listeners signal difficulty, talkers produce less transformed sentences to aid listeners' processing.

Experiment I

Method

Procedure and apparatus
Subjects were run individually. A subject was seated in a sound-free chamber, fitted with earphones and microphone and given a face-down stack of 177 cards, on each of which one sentence was typed. The subject was told that the experimenter wanted to simulate a situation of everyday life in which the subject would say something that a listener failed clearly to hear and understand, resulting in the listener saying “What?” The subject was told that the experimenter would be seated in the outer room, listening through earphones to the subject reading each sentence, and that there would be a varying level of noise present. As a result, the experimenter would often have to ask the subject “What?”, at which time the subject should try to act as (s)he would in that situation in real life, repeating the sentence verbatim or changing it in any way (s)he chose, whichever seemed most natural. The subject was also told that the experimenter was interested in the changes the subject might make.

The sequence of events was as follows:

   The experimenter said “OK”.
   The subject turned over a card, read it aloud, and turned it face down.

On 146 of the trials the experimenter asked “What?” in a natural, questioning intonation.

The subject repeated the sentence with or without changes.

The experimenter said “OK” 15 sec after (s)he said “What?”

The subject went on to the next card.

On the 31 occasions when the experimenter did not say “What?” (s)he said “OK” and the subject went directly to the next card. Subjects took a 5-minute break midway. They were asked at the end of the experiment if they thought they had responded naturally and all said they had, assuming they would have uttered the initial typed sentence.
Table 1  Examples of clear and distorted sentence versions for each linguistic construction type

<table>
<thead>
<tr>
<th>Clear</th>
<th>Distorted</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a Subject relative</td>
<td>The treasure that she found was valuable.</td>
</tr>
<tr>
<td>1b Object relative</td>
<td>Tony watered the plant that the florist had sold him.</td>
</tr>
<tr>
<td>2 Relative with copula</td>
<td>The people who were criticizing the politician were angry.</td>
</tr>
<tr>
<td>3 Object NP complement</td>
<td>Roger insisted that he had the right of way.</td>
</tr>
<tr>
<td>4 Subject NP complement - trans verb</td>
<td>It gratified Marcy that her thesis was a success.</td>
</tr>
<tr>
<td>5 Subject NP complement - intrans verb</td>
<td>It appears that William is going to Chicago.</td>
</tr>
<tr>
<td>6 Tag questions</td>
<td>Has the chef started our order yet, has he?</td>
</tr>
<tr>
<td>7 Manner adverbials</td>
<td>Ginny persuasively argued her case.</td>
</tr>
<tr>
<td>8 Deleted questions</td>
<td>Why don't you finish your homework now?</td>
</tr>
<tr>
<td>9 Permuted relatives</td>
<td>Somebody who loves me called me.</td>
</tr>
<tr>
<td>10 Verb plus particle</td>
<td>Jesse put on his shirt.</td>
</tr>
<tr>
<td>11 Todative</td>
<td>The salesman sold Jerry a watch.</td>
</tr>
<tr>
<td>12 Regular passive</td>
<td>The spy divulged the secret to Emma.</td>
</tr>
<tr>
<td>13 Double-agent passive</td>
<td>Tom took advantage of Lou (4 only)</td>
</tr>
<tr>
<td></td>
<td>Lou was taken advantage of by Tom.</td>
</tr>
</tbody>
</table>

Materials

Eight sentence pairs were created for each of 12 linguistic constructions so that the syntactic relations were clear in one version of a sentence and distorted in the matching version. A sentence was held to display clear sentential relations relative to its semantically- and lexically-equivalent mate under the following conditions. If both sentences were derived from the same deep structure, the sentence closer to the deep structure representation was the clear form (as in 1, 2, 3, 4, 5, 6, 9, 10, 11 and, arguably, 7 and 8). If the two sentences were derived from different deep structures, in the
clearer sentence the surface structure topic was the deep structure subject rather than deep structure object (as in 12 and 13), or the clearer sentence was derived from a deep structure with fewer empty nodes than the deep structure of the distorted one (as in 8). (The linguistic structure of each construction is discussed at greater length on pages 165–168.)

Table 1 lists each construction and gives an example of a clear and distorted version. For a 13th construction, double-agent passives, four clear forms and eight distorted forms were created. Double-agent passives are such that either of the two object noun phrases can be the subject (see Table 1). Thus, an active form has two corresponding passive forms. Data are reported for these thirteen constructions. Thirty-one filler sentences, which were not queried, were also constructed.

Sentences were not controlled for length. The clear versions ranged between three and twelve words, with a mean of 7.52 and standard deviation of 1.83. Seventy-five percent of all clear sentences were between six and nine words long. The distorted versions ranged between one and twelve words, with a mean of 6.92 and standard deviation of 1.83. Seventy-eight percent of all distorted sentences were between six and nine words long.

Data are not reported for an additional six constructions. For five constructions the sentence pairs differed along different sets of dimensions than the clear-distorted dimension, such as nouns vs. gerunds. The sixth construction consisted of four non-sentences which represented putative deep structure strings. A total of 319 sentences was constructed.

Design
Subjects were divided into two groups of 10 each (5 female, 5 male). One group read four clear and four distorted sentences from each of Constructions 1 - 13 and the other group read the eight complementary versions. For Construction 12 (double-agent passives) one group read two active and four passive sentences and the other group read the six complementary sentences. In two of the passives the direct object was surface subject and in the other two the prepositional object was surface subject. (The remaining 44 queried sentences were divided among the six additional constructions: eight for each of five constructions, four for the sixth construction.)

Each subject received a different random order of the 146 queried sentences. The 31 filler sentences occurred in one of two orders, under the constraint that no fewer than two and no more than eight experimental sentences intervened between each filler.

The design allowed computation of a 2-factor repeated measures analysis of variance with subjects repeated across construction, of which there were thirteen types, and across syntactic form, which was either clear or distorted.
The design also allowed for computation of a one-between one-within analysis, with sentences repeated across syntactic form and nested within construction.

Scoring

Subjects’ responses were divided into three categories: 1) same in critical respects as the original sentence; 2) different in critical respects from the original sentence; 3) unscorable. Substitutions of one lexical item for another or elimination of lexical items were ignored if a paraphrastic relation existed between the original and repeated sentence. A response was labeled unscorable if there was no paraphrastic relation (very rare) or if the subject chose a new sentence type that did not contain the syntactic construction being investigated. For example, if the subject removed the manner adverb from “Ginny persuasively argued her case”, the position of the adverb relative to the verb could not be assessed and the response was called unscorable. Unscorable responses accounted for 13% of the data.

The criteria according to which a repetition conformed in critical respects with the original varied depending on the construction. A brief summary of the criteria for each construction is presented below. In all cases the scoring procedure for the distorted versions was the inverse of the procedure for the clear versions. Thus, only the procedure for the clear versions is described.

(1) Relative. A same response required presence of the relative marker ‘that’, or division into two independent clauses. A different response required absence of the relative marker. An unscorable response occurred if the relative clause was changed to an adjective or if the relative clause was deleted. Twenty-nine percent of the responses were unscorable. The scoring is based on Smith (1964); Bever & Langendoen (1971) suggest that ‘that’ is introduced transformationally. If their analysis is correct, the relative marker is a case where clearer sentential relations are present in sentences transformationally more distant from the base.

(2) Relative and copula. A same response required the presence of the relative marker and copula, or presence of the relative marker and a tense change of the verb to past or present (instead of the progressive), or division into two separate clauses joined by a connective. A different response was scored if the marker and copula were absent, or if the marker, copula and verb were absent. An unscorable response occurred if the relative clause was deleted or if it was permuted to the end of the sentence and changed into an adverbial (e.g., “the children made a lot of noise by chewing gum”). Thirteen percent of the responses were unscorable.

(3) Object noun phrase complement. A same response required presence of the complementizer ‘that’ or some equivalent such as ‘like’; a different
response required absence of a complementizer. A response was unscorable if the matrix clause was deleted or if the subordinate clause was converted into a non-sentential noun phrase (Rosenbaum, 1967). Fourteen percent of the responses were unscorable.

(4) Subject noun phrase complement with transitive verb. A same response occurred if the complementizer 'that' was present, even if subjects changed the sentence to an object noun phrase complement either by passivizing the verb or by using an adjective instead of the psychological verb. A different response was scored if the complementizer was absent, again even if an object noun phrase complement was used. This criterion was based on the grounds that passivization retains the complement clause as subject of the sentence. Forty-three percent of the responses were passivizations or used a predicate adjective. An unscorable response occurred in the same conditions as for object noun phrase complements, or if the sentence was converted into a 'for-to' complement, or if the complement clause was changed into an adverbial by exchanging 'because', 'when', etc., for 'that'. Nineteen percent of the responses were unscorable.

(5) Subject noun phrase complement with intransitive verb. A same response was scored if the complementizer 'that' or 'like' was present. A different response was scored if a complementizer was absent. An unscorable response occurred in the same conditions as for subject noun phrase complements with transitive verbs. Fifty percent of the responses were unscorable.

(6) Yes–no (tag) questions. A same response required presence of the tag; a different response required absence of the tag (Katz & Postal, 1964). An unscorable response occurred if the repetition was not in the form of a question. One percent of the responses were unscorable.

(7) Manner adverbials. A same response was scored if the adverb was either directly before or directly after the verb. A different response was scored if the adverb was placed after the object noun phrase at the end of the sentence or if placed at the beginning of the sentence. An unscorable response occurred if the adverb was deleted. Eleven percent of the responses were unscorable.

(8) Deleted noun phrase-verb questions. A same response was scored if the (surface) subject noun phrase and its verb or copula were present; a different response was scored if either the noun phrase or verb was absent. An unscorable response occurred if a non-question form was used. Four percent of the responses were unscorable. Post-hoc linguistic analysis indicated that one sentence had been inappropriately included: the deleted elements of the other sentences could all be plausibly argued to be present on a designated list of deletable elements, but the deleted elements of the excluded sentence could not be. Therefore, this sentence was not included in the data analysis.
(9) **Permuted relatives.** A same response was scored if the relative clause was placed alongside the subject noun phrase (Ross, 1967). Compression of the relative to an adjective was also allowed. A different response was scored if the relative clause and subject noun phrase were separated by the main verb phrase, or if the main verb phrase was converted into the relative clause and the relative converted into the main verb phrase. An unscorable response occurred if the relative clause was eliminated, or if two separate sentences were created. Eight percent of the responses were unscorable.

(10) **Verb plus particle.** A same response was scored if the particle was placed directly after the verb (Chomsky, 1964). Passive constructions, such as “the clerk was bawled out by his supervisor” were allowed. A different response was scored if the particle was placed after the noun phrase. An unscorable response occurred if a verb which does not take a particle was substituted for the original verb, or if the object noun phrase was deleted so that the particle could appear in no position other than directly after the verb. Six percent of the responses were unscorable.

(11) **To-dative.** A same response was scored if the ‘to’ was present; a different response was scored if the ‘to’ was absent (Fillmore, 1965; Jackendoff & Culicover, 1971). An unscorable response occurred if the dative was eliminated. Eight percent of the responses were unscorable.

(12) **Double-agent passives.** A same response occurred if the sentence was repeated in the active voice; a different response was scored if the passive voice was used (Chomsky, 1957). An unscorable response occurred if a middle form (“Philip and Emily got out of touch”) was used. Four percent of the responses were unscorable.

(13) **Regular passives.** The same criteria were used as for (12). Four percent of the responses were unscorable.

A same response was arbitrarily scored with an 8, a different response with a 2.

**Subjects**
Subjects were linguistically naive paid volunteers with normal hearing. Ten subjects were eliminated from the experiment either halfway through or immediately after testing because of their failure to change more than 25% of their utterances; one was eliminated because he was a poor reader; one because she did not follow instructions. This left a total of 20 subjects whose data were analyzed.

**Results**
As Table 2 shows, there was an overall tendency to repeat the sentence essentially as read, but this tendency was stronger for the clear versions than
for distorted versions, as predicted. When changes were made in the clear versions, they were only 1½ times more likely to be a distorted version than to be an unscorable response. When changes were made in the distorted versions, however, they were three times more likely to be a clear version than to be an unscorable response.

Table 2 Percent response types to clear and distorted sentence versions collapsed across construction type

<table>
<thead>
<tr>
<th>Version</th>
<th>Response Type</th>
<th>stay</th>
<th>different</th>
<th>unscorable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment I</td>
<td>clear</td>
<td>69</td>
<td>18</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>distorted</td>
<td>42</td>
<td>45</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>55</td>
<td>32</td>
<td>13</td>
</tr>
<tr>
<td>Experiment II</td>
<td>clear</td>
<td>60</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>distorted</td>
<td>31</td>
<td>53</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>X</td>
<td>45</td>
<td>37</td>
<td>18</td>
</tr>
</tbody>
</table>

The major prediction that the ratio of same to different responses would be greater for the clear versions than for the different versions was strongly confirmed. Table 3 gives the mean response scores to clear and distorted versions for each construction type; the higher the score, the greater the proportion of stay responses. In one set of scores the scores are averaged across subjects, in the other set across sentences. With subjects as the repeated measure across construction and syntactic form, the effect of syntactic form was significant beyond the 0.001 level, \( F_1 (1, 19) = 87.02 \). With sentence items as the measure repeated across syntactic form and nested within construction, syntactic form was also highly significant, \( F_2 (1, 91) = 136.39, p < 0.001 \). These \( F \)s were used to compute \( F'_{\min} (1, 47) = 53.13, p < 0.001 \). (See Clark, 1973, for the formulae.)

There was a significant effect of construction type: the absolute score obtained by ignoring syntactic form and averaging the clear and distorted versions varied by construction, \( F_1 (12, 228) = 5.58; F_2 (12, 91) = 4.7; F'_{\min} (12, 241) = 2.55, p < 0.005 \). The interaction between construction type and syntactic form was also significant. That is, the ratio between the clear and distorted version scores varied by construction, \( F_1 (12, 228) = 18.59; F_2 (12, 91) = 14.62; F'_{\min} (12, 233) = 8.18, p < 0.001 \).
Table 3  
*Mean scores for clear and distorted sentence versions presented by construction*

<table>
<thead>
<tr>
<th>Construction*</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Experiment I: What?</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subjects Clear</td>
<td>6.27</td>
<td>5.65</td>
<td>6.13</td>
<td>6.95</td>
<td>7.33</td>
<td>5.35</td>
<td>5.65</td>
<td>7.90</td>
<td>7.37</td>
<td>7.40</td>
<td>6.10</td>
<td>8.00</td>
<td>8.00</td>
</tr>
<tr>
<td>Distorted</td>
<td>5.55</td>
<td>4.95</td>
<td>4.50</td>
<td>3.63</td>
<td>3.43</td>
<td>7.93</td>
<td>6.55</td>
<td>4.27</td>
<td>3.65</td>
<td>5.50</td>
<td>6.53</td>
<td>3.95</td>
<td>2.45</td>
</tr>
<tr>
<td>Sentences Clear</td>
<td>5.85</td>
<td>5.81</td>
<td>6.13</td>
<td>7.09</td>
<td>7.41</td>
<td>5.35</td>
<td>5.97</td>
<td>7.93</td>
<td>7.47</td>
<td>7.38</td>
<td>6.18</td>
<td>8.00</td>
<td>8.00</td>
</tr>
<tr>
<td>Distorted</td>
<td>5.55</td>
<td>4.90</td>
<td>4.45</td>
<td>3.60</td>
<td>2.96</td>
<td>7.93</td>
<td>6.54</td>
<td>4.39</td>
<td>3.79</td>
<td>5.49</td>
<td>6.48</td>
<td>3.91</td>
<td>2.47</td>
</tr>
<tr>
<td><strong>Experiment II: Simplify</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subjects Clear</td>
<td>6.53</td>
<td>5.60</td>
<td>7.13</td>
<td>6.83</td>
<td>7.65</td>
<td>4.53</td>
<td>4.85</td>
<td>7.80</td>
<td>7.05</td>
<td>7.47</td>
<td>7.47</td>
<td>7.85</td>
<td></td>
</tr>
<tr>
<td>Distorted</td>
<td>4.97</td>
<td>4.67</td>
<td>4.23</td>
<td>3.23</td>
<td>3.65</td>
<td>7.77</td>
<td>5.25</td>
<td>3.77</td>
<td>2.65</td>
<td>4.35</td>
<td>6.45</td>
<td>2.45</td>
<td>2.15</td>
</tr>
<tr>
<td>Sentences Clear</td>
<td>6.47</td>
<td>5.56</td>
<td>7.33</td>
<td>7.19</td>
<td>8.00</td>
<td>4.49</td>
<td>5.07</td>
<td>7.82</td>
<td>6.69</td>
<td>6.97</td>
<td>5.46</td>
<td>7.44</td>
<td>7.85</td>
</tr>
<tr>
<td>Distorted</td>
<td>5.37</td>
<td>5.17</td>
<td>4.11</td>
<td>2.97</td>
<td>3.06</td>
<td>7.77</td>
<td>5.33</td>
<td>3.83</td>
<td>2.62</td>
<td>4.28</td>
<td>6.26</td>
<td>2.46</td>
<td>2.16</td>
</tr>
</tbody>
</table>

*See Table 1 for name of each construction and examples.

As can be seen from Table 3, ten of the thirteen constructions showed the predicted effect of syntactic form, for both the subjects and items analyses. The effect was in the opposite direction for manner adverbials, to-datives and yes-no questions. By treating each construction as a one-factor repeated measures analysis, $F'_{min}$s were computed for those constructions which had significant $F_1$s and $F_2$s. An acceptable alpha level was set at 0.005 one-tail (0.05/10). Seven constructions showed a significant effect of syntactic form in the predicted direction: subject noun phrase complement with transitive verb, $F'_{min}$ (1,22) = 16.46, $p < 0.001$; subject noun phrase complement with intransitive verb, $F'_{min}$ (1,20) = 28.23, $p < 0.001$; deleted noun phrase-verb, $F'_{min}$ (1,15) = 24.14, $p < 0.001$; permuted relatives, $F'_{min}$ (1,14) = 29.66, $p < 0.001$; double-agent passive, $F'_{min}$ (1,24) = 434, $p < 0.001$; regular passive, $F'_{min}$ (1,26) = 51.95, $p < 0.001$; verb plus particle, $F'_{min}$ (1,18) = 9.26, $p < 0.005$. Yes-no questions showed a significant effect of syntactic form in the opposite direction, $F'_{min}$ (1,15) = 11.22, $p < 0.005$ (two-tailed). The remaining constructions showed no significant effect.

**Discussion**

The results strongly confirmed the prediction that talkers would clarify sentential relations when in a simulated What? situation. In some of the constructions, the clearer version was also closer to the deep structure representation of the sentence than was the distorted version. Thus, the
results can also be interpreted to suggest that the talker’s processing mechanism has access to deeper levels of representation than surface structure. One implication of this interpretation is that some transformations are psychologically real, and that a relative derivational theory of complexity is a viable theory. For these conclusions to stand, however, it is necessary 1) to assess the linguistic status of each construction and to establish the relation between the clear version and the deep structure representation, 2) to assess the significance of the unscorable responses, and 3) to consider alternative explanations.

In the case of constructions (1) - (5), the relatives and complements, it is generally agreed that the complementizer is optionally deletable, so that the deleted version is one transformation more removed from the deep structure representation than is the undeleted, clear version (but see Bever & Langendoen, 1971). The other relatively non-controversial constructions are 6) tag questions, 9) permuted relatives, and 10) verb plus particle, where the clear version is less transformationally distant from the deep structure than the distorted version is.

For these eight constructions, then, preference for the clear version is preference for a version closer to deep structure. In the present experiment, talkers showed a preference for the clear version in all but one of these constructions, 6) tag questions. One possible explanation for the extreme preference for the distorted version in this construction is that a tag question is not recognizable as a question until the tag is reached; by changing to a yes-no question, talkers emphasize the interrogative status of the sentence.* A related possibility is that subjects flesh out the question aspect and drop the statement aspect, since only one half is needed and the sentence as a whole is a question.** The shorter length of the distorted version cannot be the responsible factor, since in many of the other constructions the distorted versions are shorter than the clear versions but are not correspondingly preferred.

Another possible explanation is that the linguistic analysis used here (based on Katz & Postal, 1964) is incorrect, that yes-no and tag questions are not derived from the same disjunctive deep structure, such that an additional transformation deletes the second half of the disjunction. Rather, tag questions and negative yes-no questions (but not positive yes-no questions) could be derived from the same deep structure, with the tag formed by an additional optional copying rule, and all other rules in common between the two types (Akmajian & Heny, 1975). This would make the tag question

---

*We thank S. Cohen Leehy for this suggestion.
**We thank D. T. Langendoen for this suggestion.
more distant from the deep structure than the negative question. However, since half the yes-no questions in the present experiment were positive and half negative, this explanation would account for only a portion of the data.

The remaining constructions are discussed individually. 7) Manner adverbials were treated as if they are in immediate construction with the verb, rather than the entire verb phrase. Therefore, placing the adverb on either side of the verb was held to clarify its modifying relation to the verb and conform to its deep structure placement. However, if the adverb qualifies the entire verb phrase its deep structure position is unfixed, since it can be attached to the verb phrase either before the verb, directly after the verb, or at the end of the verb phrase. The only position ruled out is sentence initial. The very slight preference \( F_1 (1,19) = 1.67, p = 0.2; F_2 (1,7) = 0.45, \text{n.s.} \) in the present experiment for the sentence final position would seem to rule out the hypothesis that the adverb modifies only the verb, on the assumption that subjects are in fact clarifying sentential relations.

The preference for sentence final position is larger than it seems, since it contrasts with the combined total for 2 other positions, directly before and directly after the verb, both of which were scored as conforming to the clear version. When these positions are treated separately the preference for final position becomes apparent: 49% of all responses were in sentence final, 36% were before the verb, 4% were between the verb and subsequent object noun phrase. Indeed, this last position sounds quite awkward unless the object noun phrase is a prepositional phrase. One percent of the responses were in sentence initial position and 11% were unscorable. Thus, subjects do seem to prefer sentence final position. There is one linguistic treatment of manner adverbs which can account for the verb phrase final position. Chomsky (1965) locates manner adverbs at the end of the verb phrase, though without discussing other possible positions within the verb phrase.

8) Deleted noun phrase-verb questions can easily be argued to have more clearly marked sentential relations in the clear form, because the clear form includes the surface subject and verb or auxiliary. Any sentence which specifies this information marks sentential relations more clearly than a sentence which does not. But it is a matter of controversy how much deletion should be allowed in a grammar, with recent theory (Fiengo, 1974) eliminating as much deletion as possible, in order to constrain the weak generative capacity of the grammar. Thus, the clear and distorted sentences in this construction would not be transformationally related, but derived from different deep structures and semantically related. If deletion of a small list of designated elements were allowed (Chomsky, 1965), however, then
the sentences used in 8), all of which delete only pronouns, auxiliary verbs and the verb 'go', could be argued to be transformationally related, with the deleted versions more transformationally distant than the non-deleted versions with the additional assumption that a rule which would delete just these elements could be motivated. Since deletion is so controversial, it seems more conservative to account for subjects' responses only in terms of clarifying sentential relations and not in terms of a deep structure hypothesis.

11) To-datives have recently been argued (Langendoen, personal communication) to be generated in both forms in deep structure, rather than generated with a 'to' which is then optionally deleted. (In Burt, 1971, the 'to' is transformationally inserted; this alternative is objectionable because it builds up structure.) The very marginal preference in the present experiment for the distorted form would be most compatible with dual base generation. The preference was not significant, $F_1 (1,19) = 0.58$, $F_2 (1,7) = 0.297$, indicating that both forms were viewed as being equally clear.

12) Regular passive and 13) double-agent passive are both marked as passive in deep structure, so that changing the sentence from a passive to an active form is not conforming with one aspect of the deep structure representation. The switch to the active form does clarify sentential relations, however, by removing the discrepancy between the surface subject and the deep subject: in the active form the surface subject is also the deep subject and the surface object is also the deep object. Thus, the passive constructions are an example where clarifying sentential relations results, in one sense, in a form which is closer to deep structure, but in another sense is a choice for a different deep form altogether.

Summary of discussion by construction
For the three constructions where subjects preferred the version labeled as distorted, future linguistic analysis may suggest that the label was applied to the wrong version [6) tag questions] or that neither version is distorted relative to the other [7) manner adverbials and 11) to-datives]. In seven of the remaining eleven constructions which conformed to predictions, the clear version can reasonably be identified as the version which is closer to deep structure.

Alternate explanations
Memory difficulties can be eliminated as possible explanations: a separate group of eleven subjects was asked simply to repeat back each sentence after (s)he had read it. No subject made more than a total of 6 errors and all errors were small in scope. A more serious candidate objection, mentioned earlier, is that by excluding the unscorable responses from the analysis the
prediction is vacuously confirmed: responses at variance with the hypothesis are eliminated. However, the hypothesis does not demand that subjects maintain the same constructional type. It is a measure of the perceived awkwardness or redundancy of a construction that subjects eliminate it, but this perception and behavior by subjects is orthogonal to the hypothesis, which is only in effect if the construction is maintained. For example, subject noun phrase complements with intransitive verbs received the largest percentage of unscorable responses, 50%. In 78% of these responses subjects dropped the matrix clause “it happened that” or “it seemed that”, etc. The hypothesis being tested does not predict that subjects will maintain the matrix clause, but that if they do the complementizer ‘that’ will also be present.

Another objection might be that subjects’ responses to sentences of one version were contaminated by the presence of similar sentences in the other version. For example, it might be argued that subjects would not spontaneously have thought of manipulating the presence of complementizers had there not been sentence versions with and without the complementizer. There are cases, however, where the alternative form was almost never used, such as the two passive constructions and tag questions. Further, the occasional high percentage of unscorable responses indicates that subjects felt free to choose different constructions when the alternative form was not congenial. Finally, even if subjects’ behavior were contaminated as suggested, the objection does not explain why the clear version was preferred to the distorted version.

A final objection might be that subjects try to maximize redundancy in a What? situation and that in general the clear version was also the more redundant version, as well as the longer version. Constructions 1), 2), 6), 8) and 11) are all more redundant and longer in the clear version, but two of the three constructions which did not show the predicted effect are in this group: 6) tag questions and 11) to-datives. Thus, neither redundancy nor length can account for the results.

**Theoretical interpretation**

One important theoretical question is whether there are several different factors which can be involved in clarifying sentential relations or whether there is only one unifying factor. If there are several ways in which sentential relations can be clarified, one of which is to produce a sentence version which is closer to the deep structure representation, then the present results provide evidence that subjects have knowledge of deeper levels of syntactic representation than the surface level. If there is only one factor that is involved, the present results cannot be interpreted in this way, because there
exist examples where clarifying sentential relations is not the same as producing a transformationally less distorted sentence. Although it seems plausible that clarifying sentential relations should be multifactorial, this experiment cannot decide between the two possibilities.

A separate matter of interpretation concerns the task used in Experiment I. The instructions to subjects did not give any explicit directions about how to insure that their listener would hear and understand them, because the methodological goal was to simulate an actual What? situation as closely as possible. It was assumed that subjects would interpret the instructions and the subsequent queries as requests to make the sentence clearer and simpler to understand. The assumption was tested by attempting to replicate the findings of Experiment I by using different instructions, which explicitly asked the subject to make the sentence as clear and simple to understand as possible. A similar pattern of results would suggest that subjects do interpret the What? query in part as a request for syntactic clarification and simplification.

Experiment II

Method

Materials, apparatus, design
These were identical to Experiment I.

Procedure

Only the instructions and the experimenter's queries differed from Experiment I. Subjects were told that the experimenter was interested in the everyday situation that occurred when the listener asked for repetition because what was said was not as simple and clear as it could have been. Thus, the experimenter would often say "Again" to the subject, at which time the subject "should take a few seconds and think of how to say the sentence more simply and clearly". If the sentence was already clear and simple then the subject should repeat it as first read. Finally, the subject was reminded that the goal was to make the sentence as clear and simple to understand as possible.

Subjects

Subjects were twenty linguistically naive paid volunteers with normal hearing.
When talkers clarify sentential relations

Scoring
This was identical to Experiment I.

Results
As Table 2 shows, there was a smaller tendency than in Experiment I to repeat the sentence as read. Overall in Experiment I, 55% of the responses were stay responses, whereas in Experiment II, 45% were stay responses. When changes were made in the clear versions, they were equally likely to be an unscorable response as a distorted version, which is an exaggerated version of Experiment I results. When changes were made in the distorted versions, they were 3½ times more likely to be a clear version than an unscorable response. This is again an exaggeration of Experiment I results.

The major prediction that the main effects and interaction from Experiment I would be duplicated in Experiment II was confirmed. Table 3 gives the mean response scores to clear and distorted versions for each construction type, separately for sentences and items. The main effect for syntactic form was highly significant, $F_1(1,19) = 93.27; F_2(1,91) = 199.09; F'_{\min}(1,39) = 63.51, p < 0.001$.

The effect of construction type just missed significance, $F_1(12,228) = 3.75, p < 0.001; F_2(12,91) = 3.48, p < 0.001; F'_{\min}(12,252) = 1.8, p > 0.1$. The interaction between construction type and syntactic form was significant, $F_1(12,228) = 26.53; F_2(12,91) = 19.98, F'_{\min}(12,228) = 11.4, p < 0.001$. As can be seen from Table 3, the same ten constructions showed the predicted effect of syntactic form in both experiments, but there are higher scores for both the clear and distorted versions in Experiment I. Individual $F'_{\min}$s are not reported because the effects are so similar to those of Experiment I. The only difference of note was that an additional construction, 3) object noun phrase complement, was significant, $F'_{\min}(1,22) = 14.16, p < 0.005$.

The effect of experiment was tested both with subjects repeated across syntactic form and construction type and nested within experiment, and with items repeated across syntactic form and experiment and nested within construction. The effect of experiment was significant, $F_1(1,38) = 6.45; F_2(1,91) = 17.67; F'_{\min}(1,67) = 4.73, p < 0.05$. There were no significant interactions involving experiment. The main effects and interaction reported above for Experiments I and II were significant.

Discussion
The results confirmed the prediction that the effect of syntactic form would also be present under different instructions, suggesting 1) that subjects are
interpreting “What?” in part as a request for syntactic clarification and 2) that the syntactic knowledge is available when explicitly demanded as well as when implicitly requested. One consequence of the instructions which is not reflected in the pattern of results presented above was the change in subjects’ intonation patterns. In Experiment I subjects raised their voices and spoke more distinctly; that behavior was completely absent in Experiment II. This suggests that talkers interpret a “What?” as a request for clarity at all levels, but a request to simplify is interpreted only structurally and lexically.

The significant difference between the two instructional conditions is understandable in light of the emphasis placed in Experiment II on making the sentence as clear and simple to understand as possible, rather than on responding naturally, as was the case for Experiment I. The fact that no interactions involving experiment were significant confirms this interpretation.

Given the similarity between Experiments I and II, it might be objected that subjects would prefer the clear version in any task involving simplicity, reducing the interest of the results from Experiments I and II. A similar objection might be that subjects could be operating under a much cruder notion of simplicity than that proposed here, so that asking them just to choose the simpler of two sentences would produce the same pattern of results*. This was tested in Experiment III. It was expected that some, but not all, constructions would show the same effect of syntactic form as was found in Experiments I and II.

Experiment III

Method

Materials

The same materials, minus the filler sentences, were used from Experiment I. For construction 13), double-agent passives, four additional actives were constructed to contrast with the passives; they were constructed by changing the lexical items in the subject and prepositional noun phrases but otherwise repeating the sentence. The clear and distorted versions of each sentence were typed on index cards. Half the time the top sentence was the clear version; half the time it was the distorted version.

*This suggestion is due to J. Fodor.
Procedure and design
Each subject was given a stack of cards in the same random order and asked to indicate on a score sheet which of the two sentences on each card was simpler, by placing a T (for top) or B (for bottom). They were also given the option of placing an S, indicating that the two versions were the same, neither simpler than the other. No further instructions were given. The task required 10 to 15 minutes.

Scoring
Subjects' scores were computed by adding the number of times each version was preferred for each construction; a constant of 1 was added to each sum. Item scores were computed similarly. Thus, for each subject for each construction there were 2 numbers, representing the number of times (s)he chose the clear version (plus 1) and the number of times (s)he chose the distorted version (plus 1). S responses were eliminated from the analysis. The same was true for item scores.

Subjects
Subjects were 20 linguistically naive volunteers, some of whom were paid.

Results
Table 4 presents the mean preference frequencies for clear and distorted versions for each construction type, for both subjects and items. The overall

Table 4  Mean frequency preferences for clear and distorted sentence versions in Experiment III

<table>
<thead>
<tr>
<th>Construction</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clear</td>
<td>3.50</td>
<td>3.25</td>
<td>4.25</td>
<td>6.65</td>
<td>5.30</td>
<td>1.60</td>
<td>4.30</td>
<td>3.70</td>
<td>6.30</td>
<td>4.55</td>
<td>2.85</td>
<td>7.80</td>
<td>8.95</td>
</tr>
<tr>
<td>Distorted</td>
<td>5.30</td>
<td>5.65</td>
<td>4.50</td>
<td>2.65</td>
<td>3.75</td>
<td>7.25</td>
<td>3.40</td>
<td>4.10</td>
<td>1.65</td>
<td>3.00</td>
<td>5.55</td>
<td>2.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Sentences</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distorted</td>
<td>11.75</td>
<td>12.63</td>
<td>9.75</td>
<td>5.13</td>
<td>7.50</td>
<td>16.25</td>
<td>6.50</td>
<td>9.71</td>
<td>2.63</td>
<td>6.00</td>
<td>11.75</td>
<td>3.50</td>
<td>1.00</td>
</tr>
</tbody>
</table>

*See Table 1 for name of each construction and examples.

effect of syntactic form was significant for subjects, $F_1 (1,19) = 4.34; p < 0.05$, and items, $F_2 (1,91) = 19.01, p < 0.001$, but not when $F_{\text{min}}'$ was computed, $F_{\text{min}}' (1,28) = 3.53$, n.s. There was a significant effect of construction type, $F_1 (12,228) = 4.54; F_2 (12,91) = 11.4; F_{\text{min}}' (12,317)$
= 3.25, \( p < 0.001 \). There was also a significant interaction between form and construction, \( F_1(12,228) = 31.99; F_2(12,91) = 20.01; F_{\text{min}}'(12,205) = 12.31, \ p < 0.001 \).

Although the results cannot be compared directly with the results of Experiments I and II, because frequency rather than mean score is used, the directional differences can be compared. In 5 of the 13 constructions the clear/distorted ratio differs from Experiments I and II. For constructions 1) relative, 2) relative with copula, 3) object noun phrase complement and 8) deleted questions, the distorted version is preferred to the clear; this preference is at variance with that shown in Experiments I and II. For construction 6) manner adverbials, the clear version is preferred to the distorted version, again unlike Experiments I and II. Thus, the effect of syntactic form is considerably different in Experiment III.

Another measure of similarity is extent of correlation. In each of Experiments I, II and III, each construction was ranked according to the difference in score between the clear and distorted versions. The difference ranks were used to compute correlations between Experiments I and II and between Experiments I and III*. Although the Kendall rank correlation coefficient between Experiments I and II was highly significant (\( \tau_{Ss} = 0.90, z = 4.29, p < 0.001; \tau_{\text{items}} = 0.87, z = 4.14, p < 0.001 \)), the weaker correlation between Experiments I and III was also significant (\( \tau_{Ss} = 0.73, z = 3.48, p < 0.001; \tau_{\text{items}} = 0.67, z = 3.17, p < 0.001 \)). A Kendall partial rank correlation coefficient was calculated to determine whether the correlation between Experiments I and II could be due to effects of Experiment III. If this were the case, the new correlation coefficient should be quite small. Instead, the coefficient was large (\( \tau_{XY,ZSs} = 0.74, \tau_{XY,Z\text{items}} = 0.73 \)). Although the coefficient cannot be tested for significance (Siegel, 1956), the amount of remaining correlation is more than would be expected if the variables determining Experiment III results were the only source of commonality between Experiments I and II.

*We thank an anonymous reviewer for this suggestion.

Discussion

Although not significant by the \( F'_{\text{min}} \) statistic, there was nevertheless a small effect of syntactic form. Inspection of the individual constructions suggests, however, that this is due to very large effects for some constructions, rather than being a consistent phenomenon as in Experiments I and II. For example, in Experiments I and II only three constructions out of thirteen failed to show the predicted effect, whereas in Experiment III six
out of thirteen failed. More significant than the overall results, however, is the fact that the pattern of results apparent in Experiment I and mimicked exactly in Experiment II is quite different in Experiment III. The correlation results confirm this interpretation.

Subjects in Experiment III seemed to equate shorter length with greater simplicity. In the ten cases where one version was shorter than the other, subjects chose the shorter version as simpler eight times. This contrasts with the behavior of subjects in Experiments I and II who chose the shorter version four out of ten times.

Thus, it seems clear that the results of Experiments I and II are not obtainable under any and all conditions. Nor does a crude notion of simplicity explain the behavior of subjects in Experiments I and II.

Conclusions

Taken together, the three experiments suggest that talkers' behavior in a What? situation is syntactically uniform: talkers interpret the What? query as a request for clearer sentential relations and modify their speech accordingly. In most cases the clarification is equivalent to the production of a sentence less transformationally derived than the original, and the results therefore indirectly support a relative derivational theory of complexity. The new paradigm presented here seems successful in bringing out talkers' structural knowledge by employing a formalized version of a natural situation. The experiments demonstrate that the apparently diverse responses that subjects could make in such a situation are in fact systematically ordered. Finally, the experiments demonstrate that abstract linguistic generalizations play an important role in everyday speech.

References


Nous avons fait l'hypothèse qu'un locuteur clarifie les relations phrasistiques d'un énoncé si un auditeur indique une difficulté à entendre ou à comprendre celui-ci. Les sujets de l'expérience lisent des phrases déformées et des phrases syntaxiquement claires à l'expérimentateur, situé dans une pièce contiguë. L'expérimentateur pose souvent la question “Comment?” Les sujets changent les phrases déformées mais conservent la version initialement lue des phrases claires.

A d'autres sujets, on a demandé de rendre les phrases claires et simples à comprendre. Les résultats obtenus sont les mêmes. Les locuteurs semblent interpréter le “Comment?” comme une demande de clarification de relations phrasistiques. Ces résultats montrent que le locuteur connaît la structure sous-jacente des phrases. On peut rejeter d'autres explications de ce processus. On présente une “relative” théorie de la complexité des énoncés.