ABSTRACT
        This review of Bever's psycholinguistics survey is
for the most part favorable. Commentary is centered on sections 1, 2,
4, and 6 of the report. The survey's first part is judged significant
in that Wundt's pioneering work in psycholinguistics is discussed.
The second section, on grammar as a psychological process, is found
obscure in its denunciation of the competence-performance
distinction. Evidence presented in section 4, on language perception
processes, is not found convincing enough to refute the theory of the
psychological reality of linguistically defined surface structure.
The failure to present relevant data is cited as another fault of the
report as well as forcing conclusions based on weak evidence. Section
6 presents three implications of the theory for research in automatic
language processing. The arguments, that psycholinguistic research is
not a constraint on computational linguistics, and that automatic
language processing has as a goal the development of man-made
communication, are held to be unsound. The argument that the
development of computational algorithms based on grammars generates
advances in linguistic theory is considered more valid.
(Author/AM)

LNLJ III   Paul G. Chapin

REVIEW OF

THOMAS G. BEVER, A SURVEY OF SOME RECENT

WORK IN PSYCHOLINGUISTICS *

    This is a stimulating and extremely valuable work, offering the author's
theoretical interpretation of the last few years' experimental research in
the psychology of language perception and its relation to research in automatic
language processing. No small part of its value resides in the fact that it
is written without technical terminology or a welter of statistical detail,
thus rendering it accessible to the reader with no special background in
experimental psychology, such as the present reviewer.

    The report is organized into six sections. The first is a brief review
of the history of the subject, with attention devoted both to relevant psy-
chologizing and to relevant linguisticizing. The most significant part of
this section is a fairly detailed exposition of the pioneering work of Wundt,
which will be new to most readers. Already in 1900 Wundt perceived that
a sentence is a psychological whole, not just a sequence of psychological
events, and that a distinction must be maintained between the external form
of sentences (word order, etc.) and their underlying logical structure. It
is a minor tragedy for the progress of our understanding of human psycho gy
that these insights were ignored for some six decades.

    The second section is the most obscure. It is entitled "The Study of
Grammar as a Psychological Process", and it argues that the classical
distinction made by generative grammarians between "competence" and
"performance" is an artificial one (and presumably also the Saussurean
langue-parole distinction, though this is not specifically claimed). This is
argued on the ground that the linguist is engaged in a psychological process
when he is doing linguistics--he is making decisions which can only be justified
by appeals to intuition about the domain of facts which he will describe and
theorize about. Analytical division of language into levels (phonological,
syntactic, etc.), judgments of grammaticality, the adoption of the sentence
as the basic unit of analysis, judgments of ambiguity, and judgments of
structural relations are offered as examples of such decisions. Therefore,
goes the argument, it is false to distinguish between idealized

or "pure" grammar--competence--and "contaminated" speech behavior--
performance--since psychological processes are engaged in abstracting
the former from the latter.

This argument rests on a basic confusion between the linguist as data
source and the linguist as data analyst. There are manifest advantages to
serving as one's own informant, and linguists frequently do so (there
are also manifest dangers, and any careful linguist will check his subtler
intuitions with several other native speakers before relying on them).
But in principle any native speaker of the language under study could
serve as informant, providing the intuitions of grammaticality, ambiguity,
structure, etc., which are the primary data of linguistics. That the
"mentalistic" character of these data does not destroy the possibility
of doing linguistics as an objective science has already been well argued
(e.g. by Katz 1964), and it is surely not Bever's intention to dispute
this. But with data in hand, the linguist doffs his informant's hat, dons
his scientist's hat, and proceeds to analyze the data and construct ex-
planatory models, a task possible only for someone with rather spe-
cialized training and experience. In this endeavor he is no better or
worse off than the theoretician working in any other domain of inquiry
with respect to the interference of his own psyche in the performance
of his work. If I understand Section two correctly, Bever is denying the
possibility of this transition. Yet the various examples he offers of
"intuitive" decisions linguists make point clearly to the distinction that
is being ignored. Bever concedes that a high degree of intersubjective
consistency can be achieved with respect to judgments of grammaticality,
ambiguity, and structural relations. On the other hand, he points to
the numerous different proposals that have been made, explicitly or im-
plicitly, as to the basic unit of linguistic analysis; he might well also
have discussed the numerous controversies regarding the number and
nature of linguistic levels. Yet it is just exactly the former class of
judgments which are those which can be made by any native speaker,
and constitute the data of linguistics, and the latter class which are made
not on the basis of some sort of linguistic intuition, but on the basis of
assumptions regarding what will constitute the most comprehensive and
revealing theoretical model; in short, scientific judgments.

The discussion in Section two is marred by the adducement of
several claims which are either ill-considered or irrelevant. I shall
discuss these briefly.

(1) "A grammar...describes a set of 'grammaticality' intuitions
which native speakers can train themselves to have." Whether intuitions
of grammaticality are acquired by training (apart, of course, from the
training involved in learning the language in the first place, if there is
any) is an empirical question, and one for which the prima facie evi-
dence seems to indicate a negative answer. It is certainly necessary
for an informant to understand the nature of the task (to distinguish it,
for example, from requests for information about prescriptively approved

usage); but once this is done, the ability to make highly sophisticated judgments about grammaticality seems to flow directly from native command of the language, with no special training required.

(2) "**By assumption** the natural domain of linguistic theory is the set of all _intuitively_ well-formed sentences: any systematic facts about speech which are outside that domain are by assumption part of the 'performance' of the language " (Italics his.) As mentioned above, one of the assumptions of generative grammatical theory is that the most revealing model of language can be constructed if the sentence is taken as the basic unit of analysis. However, it is preposterous to claim that any systematic fact about speech which is not about sentences is assumed to be part of performance. Very few phonological studies extend their domain of investigation beyond the level of the phonological word, but the authors of th se studies, insofar as they have accepted the distinction at all, have always assumed that they were describing competence. In syntax, there has been some study of the syntactic constraints on word derivation, and some study of "discourse analysis," i.e., of intersentential syntactic processes within a structured discourse. In neither case has it been assumed that these were studies of performance.

(3) "Linguists defended themselves from the onslaught of inexplicable psychological facts about speech by invoking the distinction between competence and performance." Again (with reference to the various proposals as to the basic unit of linguistic analysis): "There are...many arbitrary lines which have been drawn by linguists in order to partition off the kinds of facts about language which they felt prepared to describe." In his Foreword, the author refers to the work as "a highly personal review." Even so, I consider remarks like the foregoing about serious theoretical proposals tendentious.

(4) "...there is no guarantee that a linguistic grammar itself is either a _direct_ or _ideal_ representation of the linguistic structure." (Italics his.) This is obviously true, just as it is true that there is no guarantee that General Relativity is a direct or ideal representation of that aspect of the cosmos which it purports to describe, or that the double helix is a direct or ideal representation of the DNA molecule. The relevance of this to any argument whatsoever is hard to see.

(5) A number of examples of fuzzy data are offered, to show that decisions of grammaticality, ambiguity, and structure are not always clear-cut. Again, this is a problem common to all empirical fields, and can scarcely be taken as an indictment of any particular theoretical claim in linguistics.
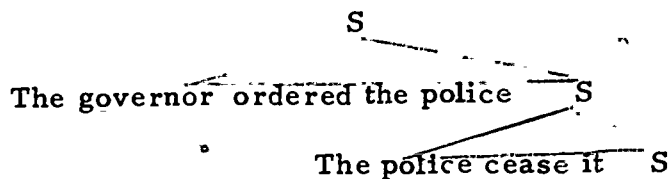
In sum, the competence-performance distinction is not holy writ, but the arguments advanced by Bever fail to show that it is any more pernicious in linguistics than similar idealizing assumptions in other fields, such as the ideal gas, the frictionless surface, or the economic man. Bever feels impelled to revoke the distinction in order to press the central claim of the work, that linguistically defined structures are operative in language perception, while linguistically defined processes (grammatical rules) are not (this claim will be discussed in some detail below). Otherwise, he feels the claim leads to a dilemma. However, the reasoning here is obscure. The claim is a priori perfectly plausible; whether it is true or not is of course an empirical question, and one on which the competence-performance distinction has no particular bearing. The "dilemma" that Bever sees arising from his claim is a spurious one, all right, but its spuriousness is grounded in the unnecessary assumption that psychological reality entails active deployment in perception, and not in the distinction between competence and performance.

The central part of the report is contained in Sections three and four, which are interpretive surveys of the experimental evidence regarding (principally short-term) memory of sentences perceived and processes of language perception, respectively. The thesis of Section three is that recent experimental evidence has tended to confirm the "coding hypothesis," which is essentially that sentences are remembered in terms of their underlying syntactic structures. The most striking experiment described was an investigation of the effectivenesss of various words in a sentence as prompts to the recall of the sentence. It was discovered that the word "police" was significantly more effective as a prompt for sentence (a) than for sentence (b):

(a) The governor ordered the police to cease drinking.
(b) The governor ordered the police to prevent drinking.

For the words "governor" and "ordered", there was no significant difference in their effectiveness as prompts for the two sentences. This finding correlates exactly with the difference in the presumed underlying structures of (a) and (b):

(a)   S

         The governor ordered the police   S

                          The police cease it   S

                                    The police are drinking

(b)

```
                                        S
        The governor ordered the police      S
                            The police prevent it    S
                                        Somebody is drinking
```
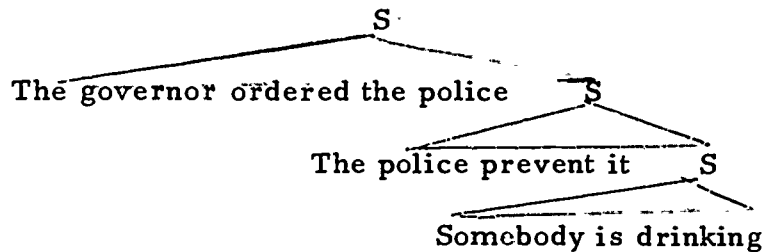
In (a), "police" occurs three times in the underlying structure, in (b) only twice, while "governor" and "ordered" have a single appearance in each case.

Section four develops the theory that language perception is governed by a set of "perceptual strategies," heuristic processes for immediate sentence analysis. These fall into two major categories, called by the author segmentation strategies and functional labelling strategies. The former are employed in reducing lengthy utterances to manageable, significant subparts; the latter operate to assign functional roles to the individual members (words) of the subparts so obtained. A variety of experimental evidence for the existence and the nature of these strategies is discussed.

One of the great successes of modern psycholinguistics has been the discovery of a new experimental technique for studying the psychological status of internal sentence structure. I refer to the famous "click" experiments, first described in (Fodor and Bever 1965). These were originally taken to demonstrate the psychological reality of linguistically defined surface structure of sentences. However, as more and more experiments have been performed with this technique, the pattern of results which has emerged has persuaded Bever that the data must be interpreted differently: that the psychologically real structural divisions in a sentence which attract the subjective positioning of clicks are in fact divisions between sentences of the underlying structure--clauses of the main sentence--and not surface phrase structure divisions. If true, this would indicate that in the process of language perception we immediately segment the utterance into subparts corresponding to the principal subparts of the underlying structure--no "surface parsing" step takes place.

An examination of the supporting evidence offered leaves this reviewer much less convinced of the validity of this claim than Bever appears to be. The original experiments were consistent with either position, as Bever acknowledges. The example he provides of the first sentences used is:

(c) Because it rained yesterday the picnic will be cancelled.

The subjective location of clicks was attracted to the point between "yesterday" and "the", which is both the major surface phrase structure division of (c) and the point of division between the subordinate clause and the main clause of (c). Bever then offers several examples culled from the early experimental materials of presumable underlying clauses which are embedded in the main sentence with successively less overt marking of their clausal status. It is important to note these examples exactly.

(d) The man who nobody likes is leaving soon.
(e) Nobody likes the man who is leaving soon.
(f) The reporters who were assigned to George drove to the airport.
(g) The reporters assigned to George drove to the airport.
(h) Only the metropolitan district of Hamburg was leveled by the war.

For examples (d) and (e), it is stated that "the points at the extremes of the embedded clauses are as effective in attracting the subjective location of clicks as they are in sentences with two entirely separate clauses." Example (f) is offered for contrast with (g), and it is observed that for (g) (as well as for (f), presumably) "clicks were displaced perceptually to the point following 'George'. " Finally, (h) is offered as an example of a case in which "one logical structure sentence was so completely embedded within another that there was absolutely no explicit clause boundary in the apparent structure"; the result was that "the point between 'Hamburg' and 'was' was found to be effective in attracting subjective click location even though this point is not a break between clauses in the surface phrase structure. " Bever concludes, "These data suggest that an early step in the perceptual organization of a string of words is to isolate those sequences in the surface order which correspond to underlying structure sentences."

This conclusion is startling, to say the least. In examples (f) through (h), the point to which the clicks were subjectively attracted is the primary surface phrase structure division of the sentence. In example (d), the embedded clause has two "extremes", one before "who" and one after "likes. " The latter coincides with the primary surface constituent break of the sentence, the former with one of the second-level breaks. In (e) the point before "who" is also at a second-level phrase structure break. Both of these second-level breaks, however, are the major breaks in their vicinity. Thus all of these data are totally compatible with the view that clicks are attracted to major breaks in surface constituent structure.

It would have been much more strongly indicative, though not conclusive, of the validity of Bever's claim if he had discovered that in sentence (g), say, the point before "assigned" had been as effective in attracting clicks as the point after "George."

There is a hint in the quote above with reference to example (h) and in several other places that Bever has some notion of "surface clause", never defined, which necessarily plays a predominant role in click attraction, if surface structure plays any role in click attraction at all. It is possible to give a precise definition to this notion (see for example Ross 1967b), and it is even possible that it is important in segmentation procedures; but Bever never offers us any reason to believe that the non-existence of surface clauses, however defined, in sentence (h) entails that click attraction to the point following "Hamburg" is caused by under-lying clausal structure rather than surface constituent structure.

Bever recognizes that the data presented so far are open to alternative interpretations, which he proceeds to list and argue against. The two alternatives suggested to the hypothesis that clicks are attracted to points corresponding to divisions between underlying structure sentences are (1) that attraction of clicks is toward "every surface phrase structure division, including those within clauses", or (2) that it is toward "those points at which an underlying structure sentence division concides with an explicit clause division in the surface structure." The principal experimental evidence adduced against these hypotheses is from an unpublished experiment by Bever, Kirk, and Lackner, the description of which I shall quote in full:

> Bever, Kirk, and Lackner used the same technique of click location in which they systematically varied the within-clause surface phrase structure of 25 sentences. They found no tendency for the number of errors into a break to be correlated with the relative depth of that break in the surface phrase structure....Bever, Kirk, and Lackner examined their results and found several instances in which a within-clause phrase structure break corresponded to a division between sentences in the [underlying constituent structure]. These breaks did attract the subjective location of clicks.

Most disappointingly, not a single example of the experimental sentences used is given, so that the reader can judge for himself the linguistic accuracy of the experimenters' judgments as to depth of

surface structure break or point of division between underlying structure sentences. There is more than one instance in the text where my own linguistic estimates differ from those tendered by the author; this fact reflects on the competence of neither of us, but it does leave me unconvinced by claims which depend on the analyses I find questionable. One example is apropos to this discussion. Another experiment taken as disconfirming of alternative hypothesis (2) was that the position after the main verb was significantly more effective in attracting clicks in sentence (j) than in sentence (k):

(j)   They desired the general to fight.
(k)   They defied the general to fight.

It is postulated that this difference depends on underlying structure differences between (j) and (k), "since there is no obvious difference in the surface structures of these sentences corresponding to the differences in the underlying structure." While it is a moot question exactly what constitutes an "obvious" difference in surface structure, linguistic discussion of these sentence types has generally postulated a surface structural difference between them of exactly the sort which would predict the experimental results described under either of the alternative hypotheses (see for example Rosenbaum 1967 [to which Bever explicitly refers], especially the surface structure analyses of the sentence "I hope for you to come on time," p. 86, and of the sentence "I reminded John to visit his mother," p. 89). Coincidentally, a sentence exactly analogous to (j) ("I desire you to win the prize") is offered by McCawley (1968, p. 72) as an ungrammatical example. One can only speculate on the effect on experimental subjects who share McCawley's dialect.
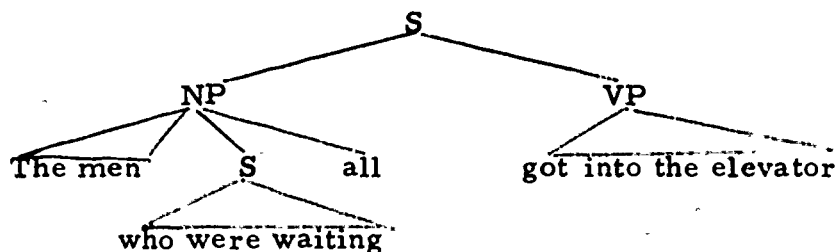
The other major experiment described in this section actually seems to confirm alternative hypothesis (2). The example sentences are:

(1) They watched the light turn green.
(m) They watched the light green car.
(n) They watched the light on the corner.

It was discovered that the point after "watched" was significantly more effective in attracting clicks in sentence (1) than in sentences (m) or (n). This difference correlates with the fact that sentence (1) contains an explicitly marked sentence embedded in the surface phrase structure, while sentences (m) and (n) do not, although according to generally accepted analyses, all three sentences contain an embedded sentence in the underlying structure. Since the presence of an "S" node in surface

9

I would postulate the following surface structures for these sentences (irrelevant details omitted):

(o)

```
                          S
              NP                      VP
    The men    S    all      got into the elevator
          who were waiting
```

(p)

```
                          S
            NP                    VP
    The beautiful fairy princess    belongs to    S
                                      whoever slays the golden dragon
```

If these analyses are correct, then in sentence (o), all has the boundary of an underlying (and explicit surface) clause preceding it, and a major surface constituent break following it; in sentence (p), belongs to has a clause boundary following and major constituent break preceding. Clicks should be placed objectively during all of (o) and belongs to of (p).

Since the embedded clauses in (o) and (p) are still explicitly marked as embedded sentences in the surface structure, attraction of the clicks to the clause boundaries will support only a weakened version of Bever's hypothesis. Thus an appropriate follow-up experiment, in the event that click attraction in(o) and (p) is to clause boundaries rather than to major surface constituent breaks, is to try the same experiment with sentences like (o') and (p'), derivatives of (o) and (p) but without embedded S nodes in the surface structure:

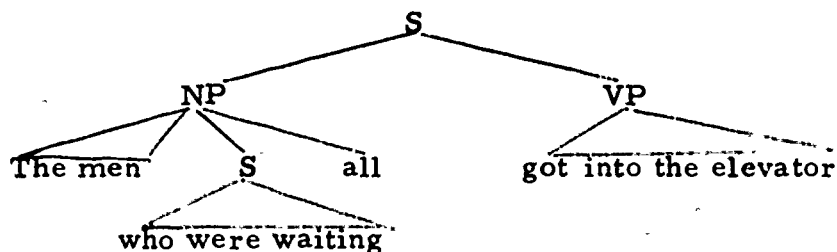(ó) The men waiting all got into the elevator.
(ṕ) The beautiful fairy princess belongs to the slayer of the golden dragon.

Attraction of clicks away from the major surface constituent break in such sentences would be strong evidence for the validity of Bever's hypothesis.
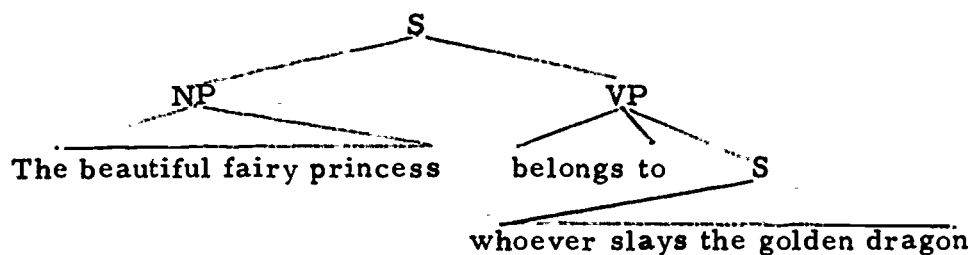
One control which I feel should be maintained in future click experiments is to place the click somewhere near the middle of the test sentence, without a preponderance of sentential material to one side or the other of the

I would postulate the following surface structures for these sentences (irrelevant details omitted):

(o)

(p)

If these analyses are correct, then in sentence (o), all has the boundary of an underlying (and explicit surface) clause preceding it, and a major surface constituent break following it; in sentence (p), belongs to has a clause boundary following and major constituent break preceding. Clicks should be placed objectively during all of (o) and belongs to of (p).

Since the embedded clauses in (o) and (p) are still explicitly marked as embedded sentences in the surface structure, attraction of the clicks to the clause boundaries will support only a weakened version of Bever's hypothesis. Thus an appropriate follow-up experiment, in the event that click attraction i n(o) and (p) is to clause boundaries rather than to major surface constituent breaks, is to try the same experiment with sentences like (o') and (p'), derivatives of (o) and (p) but without embedded S nodes in the surface structure:

(ó) The men waiting all got into the elevator.
(p') The beautiful fairy princess belongs to the slayer of the golden dragon.

Attraction of clicks away from the major surface constituent break in such sentences would be strong evidence for the validity of Bever's hypothesis.

One control which I feel should be maintained in future click experiments is to place the click somewhere near the middle of the test sentence, without a preponderance of sentential material to one side or the other of the

objective location of the click. Recent linguistic studies (see Ross 1967b) have begun to indicate that "heavy" noun phrases, i.e., noun phrases with lengthy embedded clauses, have some poorly understood linguistic properties of their own, simply due to their "weight". It seems likely that there are corresponding psychological properties, that a constituent which requires a relatively great amount of time to process will assume relatively great psychological importance, which may interfere with other properties of the sentence under study. It is for this reason that the VP of (o) is not simply "got in," nor the NP of (p) simply "She."

The other major category of perceptual strategies are the _functional labelling_ strategies. Given the segmentation of the perceived utterance generated by the segmentation strategies, the functional labelling strategies assign functional roles ("actor", "action", "object", "modifier") to each of the elements of a segment. Bever sees a hierarchy of such strategies.

The simplest functional labelling strategy to describe intuitively (and no doubt the most difficult to formalize) is the semantic plausibility strategy. If the selectional restrictions on the lexical items in a segment are such that only one functional labelling is possible, that labelling will be assigned immediately. The principal evidence for this is that passive sentences are perceptually more complex than their corresponding actives unless the semantic relations are unique, in which case there is no difference in complexity. That is, sentences (q) and (r) are equally simple, while (s) is more difficult than (t):

(q) The cookie was eaten by the dog.
(r) The dog ate the cookie.
(s) The horse was followed by the cow.
(t) The cow followed the horse.

When semantic plausibility does not determine functional labelling, as in (s) and (t), other interpretive strategies must be brought into play. Bever claims that the primary one of these is the following: "any Noun-Verb-Noun sequence within a potential underlying constituent structure unit in the surface structure corresponds to actor-action-object." Violation of the sequence expected by this strategy is claimed to account for the relatively greater perceptual complexity of (s) than (t), as well as for the greater complexity of (v) than (u):

(u) They are fixing benches.
(v) They are performing monkeys.

12

Most of the research in this area of psycholinguistics has been directed at confirming the hypothesis that perceptual complexity of sentences is directly correlated with their linguistically defined derivational complexity-- that the more transformations that apply in the derivation of a sentence, the more complex it will be to process perceptually. This hypothesis would explain the complexity differential in sentence pairs (s)-(t) and (u)-(v), given presently generally accepted linguistic analyses of these sentences. It would of course have to be modified to accommodate the finding that when the semantic relations are unique, the complexity differential does not arise, at least for the passive case (it is unclear why semantic plausibility strategies are not useful in sentence (v)). However, Bever lists three experiments, the results of which he asserts are inconsistent with the hypothesis and tend to confirm the proposed primary functional labelling strategy. As this point is of central importance, I shall consider each of the three experiments in some detail.

The first experiment was one of a number exploring the perceptual properties of double center-embedded sentences, which have been known for a long time to be of enormous perceptual complexity. It was discovered that subjects tend to interpret

(w) The man the girl the boy met believed laughed.

as

(x) The man, the girl, and the boy met, believed, and laughed.

Bever takes this as showing that the subjects have imposed a general "actor-action" schema on (w). While the result is consistent with this interpretation, it by no means forces it. All that is really shown by this experiment is that sentences with compound subjects and objects are perceptually simpler than center-embedded sentences, which is not surprising. Several hypotheses have been advanced as to the source of the complexity of center-embedding--any of them would predict the described results. Whether sentence (w) or (x) is syntactically more complex is an empty question at this point; at the present level of our knowledge of grammar, the sentences are linguistically incommensurable.

The second experiment also dealt with doubly center-embedded sentences. Subjects were confronted with the sentence

(y) The editor authors the newspaper hired like laughed.

Under the only possible interpretation of the entire sentence, "authors" must be taken as a noun phrase. However, even after training (repeating the experiment with a number of different but analogous sentences) subjects had a strong tendency to interpret "authors" as a transitive verb, with "the editor" as subject and "the newspaper" as direct object. Sentence (y) was

much more difficult to comprehend than

(z) The editor the authors newspapers hired liked laughed.

where the second "the" eliminates the possibility of interpreting "authors" as a verb. It is claimed that this shows the power of the proposed primary functional labelling strategy.

It is fairly clear that the difference in syntactic complexity between (y) and (z) is minimal, and that some perceptual strategy not correlated with linguistic complexity is at work. However, the primary functional labelling strategy proposed by Bever is not the only candidate. I would like to advance a quite different hypothesis: a primary goal in language processing is to terminate constituents of the highest possible level at the earliest possible point. This hypothesis would explain the difference between (y) and (z). It would also coincide with another extremely insightful notion of Bever's, mentioned somewhat later in the text. Bever suggests that there may be a perceptual explanation for the interesting generalization made by Vendler (Vendler 1961) that order of adjectives preceding nouns in English is determined by the "noun-likeness ' of the adjective. That is, an adjective which is more like a noun, such as a color name, will appear closer to the head noun of the noun phrase than one less noun-like, such as an adjective of size. Compare "little red house" with "red little house". Bever proposes as the relevant perceptual strategy the following: given a determiner, expect a noun. Then if a more "noun-like" adjective precedes one that is less noun-like, the listener is likely to interpret the more noun-like adjective as the head noun of the noun phrase, thus throwing his computation of the sentence out of phase. This is an attractive idea, but Bever is obliged to postulate a distinct perceptual strategy to accommodate it, while it flows as a direct consequence from the more general hypothesis advanced above.

Testing the proposed hypothesis should be relatively straightforward. It would predict, for example, that (aa) would be perceptually more complex than (bb):

(aa) Only the tallest loggers fell the giant redwoods.
(bb) Only the tallest loggers will fell the giant redwoods.

In (aa) interpretation of "fell" as an intransitive verb in the past tense permits spurious sentence closure immediately following, while in (bb) the "will" eliminates the possibility of that interpretation. Similarly, (cc) should be more complex than (dd):

(cc) Everyone expected to arrive before 7:00 did so.
(dd) Everyone expecting to arrive before 7:00 did so.

The point following "7:00" in (cc) can be interpreted as the end of a sentence, while in (dd) it cannot.

Negative results on these and similar experiments would disconfirm the hypothesis. Testing the hypothesis vis-à-vis Bever's functional labelling strategies, however, must wait until the latter are spelled out in more detail. That is, positive results on the experiments just proposed could be explained by appropriate enrichments of Bever's theory. A finding that (aa) is indeed more complex than (bb), for example, could be taken as showing that the functional labelling "actor-action" takes precedence over "actor-action-object." If (cc) is more complex than (dd), that could be because on the correct interpretation of (cc), an "object" ("everyone") precedes an "action" ("expected"). Until the theory of functional labelling strategies is more fully formulated, critical empirical tests of it cannot be devised.

The third experiment investigated immediate recall of sentences like

(ee) Quickly the waiter sent the order back.
(ff) The waiter quickly sent back the order.

It was discovered that in immediate recall, subjects tended to eject the "extraneous" material to the extremes of the sentence; 87% of the syntactic order errors were from sentences like (ff) to sentences like (ee), rather than the reverse. Bever takes this as indicating that "subjects have a strong tendency to reconstruct a sentence they just heard to maximally conform to a [noun-verb-noun] sequence."

This experiment is on its face the most impressive of the three. However, two serious questions are raised. First of all, its relevance in determining the relative perceptual complexity of sentences is not entirely clear. Typical experimental techniques for investigation of perceptual complexity are determination of response accuracy and response latency (delay in response time) in performing tasks such as paraphrasing perceived sentences or verifying pictures as appropriate or inappropriate to the situation described by the sentence. While these techniques are apparently not so reliable as the click experiments, the experimenter can at least be fairly confident that they measure the same variable or class of variables. Immediate recall experiments, on the other hand, require rather different sorts of performance on the part of the subject. It is possible, for example, that aural patterning plays a significant role. It

15

should not be assumed that results from these experiments are to be included in the same class of evidence as results arising from experiments of the other type.

Secondly, the results described appear to be contradictory to the results of another experiment described in support of the coding hypothesis. In that experiment, the test sentences were exemplified by

(gg) John eagerly called Mary.
(hh) Eagerly John called Mary.

The reported result was that in recall, significantly more errors were from sentence (hh) to sentence (gg) than vice-versa. Since the linguistic assumption was that sentence (gg) is closer to an underlying form than (hh), this was taken as supporting the coding hypothesis. But there is no apparent linguistic difference between sentences (ee) and (hh) or sentences (ff) and (gg) with regard to the status of the adverb. Thus if the latter experiment confirms the coding hypothesis, the former disconfirms it; on the other hand, if the former experiment supports the theory of func- tional labelling strategies, the latter militates against it.

There is of course another difference between sentences (ee) and (ff), namely the position of the particle "back." Linguists studying the verb-particle construction (e.g. Fraser 1965) have in general postulated that the verb and its particle are contiguous in underlying structure, and that a transformational rule of particle movement is responsible for their separation in sentences like (ee). If this analysis is correct, then a preponderance of errors in recall toward discontinuous verb-particle combinations would also tend to disconfirm the coding hypothesis. Insofar as this linguistic analysis is more convincing than Ross' analysis of auxiliaries (Ross 1967a), the third experiment is that much more damaging to the coding hypothesis than an earlier experiment described by Bever which showed that English sentences with auxiliaries of varying degrees of complexity are not significantly different in ease of retention (Ross argued that auxiliaries are actually main verbs of embedded sentences in the underlying structure; if this is true, that experiment would falsify the coding hypothesis as it is currently formulated).

There is one interpretation available of the third experiment which would preserve the coding hypothesis in something like its present form. A distinction was mentioned earlier between "heavy" and "light" noun phrases. One of the facts observed which have led to the establishment of this distinction is that particle movement is increasingly less likely around increasingly heavy noun phrases. Conversely, particle movement seems to become increasingly more likely around increasingly light noun phrases, to the point that it is quasi-obligatory around one- or two-word

noun phrases in colloquial speech. It may be (pace the "eagerly" example) that the same holds for adverb fronting--that it is much more likely around a light subject noun phrase than around a heavy one. If both of these suggestions are true, then a simple explanation of the experimental results would be that while subjects are storing the test sentence in its underlying form, they are upon regeneration of it applying transformations which are for them almost obligatory. It would be easy to test this hypothesis; simply repeat the experiment with test sentences like

The tallest red-haired waiter in the room quickly sent back the irate customer's order.

Quickly the tallest red-haired waiter in the room sent the irate customer's order back.

and observe whether there is any significant deviation from the error pattern observed for sentences (ee) and (ff). The relative disturbance of the "actor-action-object" sequence in the two instances is the same. It would still remain, of course, to explain the apparently contradictory results of the experiment in which the adverb was "eagerly".

To summarize this part of the discussion, Bever has described three different experiments which he claims "converge on one common explanation: Any [noun-verb-noun] sequence in the surface structure is assumed to correspond directly to actor-action-object in the [underlying structure]." I have tried to show that while at least two of the three experiments are quite significant, their significance is not necessarily that claimed.

Bever finally comes to grips with the question of how it is that we correctly interpret sentences in which the semantic relations are not unique and the primary functional labelling strategy fails (e.g. "John was kissed by Mary"). He asserts that "there is a heterogeneous set of strategies attached to specific lexical items, primarily verbs [but also including prepositions and other 'function words'--PGC]." Thus presumably the combination of some form of "be", a verb in the past participle, and a prepositional phrase introduced by "by" serves to inform us that a sentence is passive, and we invert the primary functional labelling strategy to interpret it. Several experiments are described to support this assertion.

It was discovered that doubly center-embedded sentences are significantly easier to comprehend if the subject noun phrases are separated by relative pronouns. That is, (11) was simpler than (mm):

(11) The boy who the man who the girl likes saw laughed.
(mm) The boy the man the girl likes saw laughed.

This was interpreted as showing that when the sequence NP_1-Relative pronoun -NP_2 is perceived, NP_1 is interpreted as the object and NP_2 as the subject of the action in some underlying clause, the presence of the relative pronoun being the key to the employment of this strategy.

A series of experiments investigated the complexity differential between sentences with main-verbs of differing lexical properties. The basic finding was that sentences containing main verbs which can grammatically be followed by either a noun phrase object or a sentential complement, such as see, are significantly more complex than otherwise identical sentences whose main verbs can only take a noun phrase object, such as hit. This leads Bever to suggest the following hypothesis: "The listener scans the individual lexical items in a sentence and attempts to project the deep structure of the sentence by combining all the potential [underlying constituent] structures which the lexical items have in common." This hypothesis correctly predicts the results of the preceding experiments, since a verb like see is compatible with more potential underlying structures than a verb like hit, thus requiring more time to compute their intersection with the potential structures compatible with the other lexical items in the sentence. However, the hypothesis is at least as important for what it denies as for what it asserts. To understand this, and to discuss the hypothesis in that perspective, it is necessary to survey the overall structure of the argument to this point.

The segmentation strategies, according to Bever, isolate sequences of the perceived sentence corresponding to potential underlying constituent structure sentences (clauses). The internal semantic content of these clauses is then computed by a hierarchy of strategies: if only one arrangement of the elements of the clause is semantically plausible, that is the interpretation assigned; otherwise the sequence actor-action-object is searched for; finally, all else failing, the intersection of the underlying structures compatible with the lexical elements of the clause is computed. What specifically does not happen, and this is the important point,- is any syntactic unravelling, any sort of "transformational inversion" by means of which underlying syntactic structures are computed from

surface syntactic structures. Not only are surface syntactic structures perceptually unimportant (the conclusion of the discussion of segmentation strategies), but there is no correlation between derivational complexity (roughly, the number of transformations applied in the derivation of a sentence) and perceptual complexity. Such correlations as have appeared to exist in the past are actually due to the fact that the derivationally more complex structures (passives and participial constructions of the "They are performing monkeys" class are the chief examples) violate the expectations of one of the functional labelling strategies, specifically the strategy which expects noun-verb-noun to correspond to actor-action-object.

I have discussed the steps of this argument individually, suggesting in several cases that the evidence offered is open to interpretations which are compatible with the view that derivational complexity at least some-times results in perceptual complexity, although Bever's arguments show clearly that it is impossible to maintain the one-to-one correlation of the earliest, simplistic formulations. That is, unless the experimental results are to be impeached, passive sentences are less complex when the semantic relations are unique, suggesting something like Bever's semantic plausibility strategy; some sort of strategy quite apart from straight syntactic computation is required to account for the greater complexity of comprehending sentences with apparent internal sentences, whether it is Bever's primary functional labelling strategy or the strategy I have hypothesized of fast constituent closure; and the internal syntactic properties of lexical items play a greater role than that envisioned in earlier models. However, granting the existence of all of these factors in the highly complex process of language perception is not tantamount to granting the assumption that derivational complexity is not a factor. Rather more direct sorts of evidence are required for a claim like this. Bever alludes to only two results of direct relevance.

The first result cited as the failure of an "attempt to demonstrate the perceptual reality of transformations" is a comparison of the experiments investigating errors in adverb placement in sentence recall which were discussed above, the results of which appeared mutually contradictory. Bever's comment is, "in [the experiment in which the adverb tended to be moved toward the beginning of the sentence] the same deep structure and transformational differences were included in the experimental mater-ials as in [the experiment in which the adverb tended to be moved toward the center of the sentence]. Yet the transformational differences had no perceptual effect." Insofar as this comment is comprehensible at all, it is subject to the remarks I have already made about those experiments, and I shall not consider it further.

The other result referred to is more interesting. An attempt was made to show that reduced passives are more complex than full passives, that is, that a sentence like (oo) is more complex than a sentence like (nn):

(nn) The corn pudding was eaten by Mildred.
(oo) The corn pudding was eaten.

The attempt failed. Under previously accepted linguistic analyses, this would constitute a clear counterexample to the claim that derivational complexity is a factor in perceptual complexity. The linguistic assumption has been that the underlying form of a sentence like (oo) is an active sentence with an indefinite subject, something like

(pp) Someone ate the corn pudding.

The derivation of (oo) would then proceed in two steps: passivization of (pp), and deletion of "by someone" from the passive version of (pp). Derivation of (nn), on the other hand, would only involve passivization, and no subsequent deletion. On this account, more transformations would be involved in the derivation of (oo) than of (nn), and on the hypothesis that transformational complexity entails perceptual complexity, (oo) should be perceptually more complex than (nn). However, recent studies of the English passive (Chomsky 1969, Hasegawa 1968) have called into question the accuracy of this linguistic analysis. In particular, it has been proposed that full English passive sentences are the result of the application of at least two separate rules, one of which places the underlying object noun phrase in the subject position, and another which transposes the underlying subject noun phrase to its position as object of the agentive by-phrase. While the studies referred to do not specifically consider reduced or agentless passives (those which are full sentences, that is), it is not an unreasonable extension of their arguments to suppose that only the former rule is involved in the derivation of such sentences. If this is the case, then a full passive would actually be transformationally more complex than a reduced one. It would be interesting to test whether the full passive is perceptually more complex, adjusting for whatever complexity is introduced by its extra length (the experimental results reported by Bever, recall, were only that the reduced passive was not more complex than the full passive).

As the foregoing discussion shows, psycholinguistic experimentation and hypothesizing which rests too heavily on a particular linguistic analysis may find itself vulnerable to linguistic attack; indeed, given the fluid state of the field of linguistics today, it almost surely will.

This presents the experimenter with a problem, since he does not wish to wait until all linguistic controversies are settled to begin his work. What he needs are experimental materials which will yield significant results regardless of the direction linguistic theo·y takes. In concluding my discussion of this central part of the work under review, I shall suggest some lines of research which should yield results which can be taken with fair confidence as confirming or disconfirming of the hypothesis that derivational complexity is a factor in perceptual complexity.

One experimental strategy is to investigate constructions in which two adjacent syntactic elements can appear with equal grammaticality and without change in cognitive meaning in either of their possible orders. One such example would appear to be the noun phrase and the adverb in sentences which have already been discussed such as

(gg) John eagerly called Mary.
(hh) Eagerly John called Mary.

Another example in English is a construction of participial modifier and noun as in

(qq) These were the results expected.
(rr) These were the expected results.

French has a number of adjectives which can either precede or follow the noun (grand, nouveau, etc.). Tests involving these adjectives in both pre-noun and post-noun position could be run on native speakers of French. In all of these cases, most transformational grammarians would agree that one of the forms involved one more movement transformation than the other, although they might disagree on which was which. Thus a finding that there is no significant perceptual difference between two sentences related in this way would be disconfirming of the hypothesis.

Positive results in these experiments, however, might be argued to be due to violations of functional labelling strategies. Thus other experiments are called for in which the order of elements is held constant and derivational complexity is increased, as well as experiments in which the derivational complexity is held constant and the order of elements scrambled. If the hypothesis that derivational complexity is a factor in perceptual complexity and the hypothesis that perception is guided by functional labelling strategies are contraries, these experiments should provide an empirical basis for a choice between them.

21

The simplest examples I can think of of sentences whose order of functional elements is <u>actor-action-object</u> but which are probably derivationally more complex than simple declarative sentences are the so-called cleft sentences, like

(ss) It was the boy from Boston that won the prize.

Experimentation with these sentences would require a control for the effects of their extra length. Even so, a finding of greater perceptual complexity for cleft sentences than for their corresponding simple declaratives would be somewhat equivocal, since on Bever's theory the subject would no doubt segment a cleft sentence into two substrings, corresponding to two sentences in the underlying constituent structure. Another class of putative examples includes

(tt) There will be three papers assigned by the professor.
(uu) There will be several children playing in the yard.

which are generally thought to be transformational derivatives of

(vv) Three papers will be assigned by the professor.
(ww) Several children will be playing in the yard.

and which retain the order of functional elements of their sources.

It is important that sentences of this class be presented in the future tense, since in the present and past they are ambiguous, being also derivable by relative clause reduction from sources like

(xx) There were several children who were playing in the yard.

With the problem clearly set, it is likely that serious attention to it by a number of trained grammarians will yield examples which are better than these.

The problem of finding sentences of the other sort, sentences in which the derivational complexity is held constant and the order of elements scrambled, is rather more difficult. The most convincing examples I know again involve the cleft sentence construction, with sentences differing as to whether the actor or the object is brought into the "topic" position:

(aaa) Mary was given a new hat by John.
(bbb) A new hat was given Mary by John.

There are two difficulties here, however. In the first place, there are
many dialects 'for which (bbb) is not grammatical. Secondly, the role
of indirect objects in sequential functional labelling strategies has yet
to be investigated.

A recent linguistic discovery by Grinder (1969) has, I believe, impor-
tant bearing on the matter under consideration. In an investigation of
Samoan, Grinder discovered that the Samoan equivalents of all of the
following sentences are grammatical:

(ccc) John dropped the yam and the coconut.
(ddd) It was the yam that John dropped and the coconut. (With no
        comma after dropped.)
(eee) What did John drop and the coconut?

There are excellent grounds for believing that in Samoan (ddd) and (eee) are
transformational derivatives of the structure underlying (ccc) (with appro-
priate modifications for the interrogative pronoun which need not concern
us here). However, there is an intermediate stage in the derivation; the
immediate derivatives of (ccc) are equivalent to

(fff) It was the yam that John dropped it and the coconut.
(ggg) What did John drop it and the coconut?

Then (ddd) is derived from (fff) and (eee) from (ggg) by (optional) deletion
of the it, an anaphoric pronoun whose antecedent is yam in (fff) and what
in (ggg).

The significance of these facts is in their relationship to a hypothesis
proposed by Ross (1967b). Ross observed that sentences like (ddd) and
(eee) are ungrammatical in English (and every other language with which
he was familiar) and suggested that this was not an isolated or accidental
grammatical fact, but rather a reflection of a universal constraint on
grammars to the effect that no conjunct may be moved out of a coordinate
sequence of conjuncts (with certain refinements irrelevant to the present
discussion). This constraint was one of several proposed by Ross in his
development ot a general theory of constraints on syntactic rules with
unbounded domains of application. In quite another context, Ross found
it necessary to establish a distinction between "chopping rules" and
"copying rules", the former being transformational rules which delete
some constituent or move some constituent without leaving a trace, while

the latter copy some constituent, the original from which the copy was made then being pronominalized. Ross observed that while chopping rules are subject to the proposed constraints, copying rules are not. Given this qualification, the rules deriving (fff) and (ggg) do not violate Ross' coordinate structure constraint, since the presence of the it shows them to be copying rules. However, another, independent rule deletes the it, the result being a surface structure identical to the surface structure which would have resulted if the Samoan rules in question had been chopping rules and thus in violation of Ross' constraint.

Now if Ross' constraints are indeed universal principles of grammar, or even first approximations to such principles, it is plausible to expect that they are not fortuitous linguistic abstractions, but have a psycholinguistic basis. Discovery of the psychological implications of and explanations for theoretical principles of linguistics will be an exciting and fruitful field of research for decades to come (the fifth section of the work under review offers a number of speculative but very promising suggestions of exactly this sort); but the point to note here is that it is not at all obvious how Ross' coordinate structure constraint could have any psychological value at all in a model of language perception such as that outlined by Bever, according to which the hearer goes from perceived utterance to semantic interpretation, with no intervening syntactic computation, given the facts of Samoan. That is, if the Samoan listener decodes sentences like (ddd) and (eee) by (in some sense) first restoring the deleted pronoun and then returning the noun phrase which had been removed from the position of the pronoun back to that position, he is only performing perceptual tasks which are required in any case in other languages; but if the Samoan listener is capable of interpreting sentences in which some conjunct has been moved out of a coordinate structure indefinitely far from the beginning, to the beginning, without going through the intermediate stage of restoring the deleted pronoun, then of what value is Ross' constraint, which would seem to relieve the hearer of this task? It was observed earlier that psychological reality does not necessarily entail active deployment in perception, and the same observation may apply here; nonetheless, I believe this is an instance where psycholinguists are obliged to pay careful attention to internal linguistic evidence.

The balance of the fourth section is devoted to some ramifications of the theory--learning of the perceptual strategies, the oscillations of attention during speech perception, and the perceptual effects of ambiguity. The most interesting discovery reported was that subjects find it significantly easier (as measured by latency of paraphrase response) to find an interpretation for an ambiguous sentence than for an unambiguous

one. That is, if two potential interpretations are available, the task of discovering one of them is simpler than the task of discovering a unique interpretation. This finding suggests parallelism in semantic processing.

The fifth section, as mentioned above, explores the possible perceptual foundations for some theoretical principles of linguistics. It is this section which contains the discussion referred to earlier of Vendler's generalization about adjective ordering in English and a possible perceptual explanation for it. Possible psychological foundations for certain constraints on pronoun-antecedent relationships are also discussed.

Finally in the sixth section Bever draws out the implications of the theory for research in automatic language processing. He quite correctly observes that if the theory is true, if human language perception proceeds by isolation of potential underlying constituent structure clauses and semantic interpretation of the elements of those clauses by the strategies discussed, then automatic syntactic analysis procedures such as those devised by Petrick (1965) or the MITRE group (Zwicky et al., 1965), which proceed by surface parsing and transformational inversion, are uninteresting as candidates for perceptual models. Now it is clear from the findings reported by Bever that a number of processes are involved in language perception which were in no way reflected in these procedures. In particular, much more attention must be paid in the future to lexical properties by those developing automatic syntactic analysis procedures which lay claim to being in some sense psychological models. However, in this review I have tried to show that on the crucial issues, the psychological importance of surface syntactic structures and the contribution of derivational complexity to perceptual complexity, the case is not yet closed. Further experimentation is called for along the lines I have suggested. If the results of those experiments confirm Bever's theory, then those who wish to model human language perception with a procedure which directly incorporates a grammar will be forced to rethink their positions.

It can be argued that workers in computational linguistics need not be constrained by the results of research in psycholinguistics. For example, Quillian writes (1969):

> We...believe that, given the present state of psychological theories, almost any program able to perform some task previously limited to humans will represent an advance in the psychological theory of that performance.

I fear I cannot agree with this position. The seventeenth and eighteenth centuries saw the construction of marvelous automata--machines shaped like humans which could walk, run, do acrobatics, and perform many other tasks previously limited to humans. Yet their construction added not a single paragraph to our understanding of human physiology. Mimicry does not imply understanding, as the existence of parrots testifies.

Another argument is that an appropriate goal of research in automatic language processing is the development of efficient man-machine communication in something like natural language, and that this can be viewed essentially as an engineering task. While it is not particularly fruitful to argue about goals, all experience to date has indicated that it is considerably simpler and more economical to train human beings to use the restricted sort of language required to communicate with a computer than to endow the computer with even a small fraction of the creativity and variety of the human linguistic capacity. In my opinion, it is likely to remain so.

Still another position is that the development of computational algorithms based on grammars generates advances in linguistic theory. This position has somewhat more validity than the others mentioned. I believe it is historically accurate to say that the genesis of two theoretically important linguistic notions, the notion of recoverable deletion and the notion of lexical items as complex symbols or bundles of features rather than atomic members of classes, was in such research. It is not unlikely that further research of this sort will yield other linguistically interesting results. This research then amounts to research into the formal properties of grammars, and is justifiable as such, with no necessity for claims of psychological validity.

If it is not out of place in a sober article such as this, I would like to suggest in closing that ultimately the greatest value of attempts at the simulation of human intelligence may be a humanistic one. As we devote our best efforts to the reproduction of intelligent human behavior by the most powerful general-purpose computing machinery at our disposal and continue to discover the inherent limitations on what that machinery can do, it may serve to check our boundless faith in technology, and provide the basis for a new appreciation of what an awesome thing a human being is. Perhaps fewer human beings will be maimed or destroyed by the latest advances in technology if those in charge of such things can be forced to ponder anew the ancient question:

What is man, that Thou art mindful of him?

REFERENCES

Chomsky, N. (1969). "Remarks on Nominalization", in Jacobs and Rosenbaum (eds.) Readings in English Transformational Grammar. Waltham, Mass: Ginn Blaisdell.

Fodor, J., and T. Bever (1965). "The Psychological Reality of Linguistic Segments", Journal of Verbal Learning and Verbal Behavior 4, 414-429.

Fraser, B. (1965). The Verb-Particle Construction in English. Unpublished Ph.D. dissertation, M.I.T.

Grinder, J. (1969). "Conjunct Splitting in Samoan", Linguistic Notes from La Jolla 2, 46-79 (Published by the Linguistics Department of the University of California, San Diego).

Hasegawa, K. (1968). "The Passive Construction in English", Language 44, 230-243.

Katz, J. (1964). "Mentalism in Linguistics", Language 40, 124-138.

McCawley, J. (1968). "Lexical Insertion in a Transformational Grammar Without Deep Structure", in Darden, Bailey, and Davison (eds.), Papers from the Fourth Regional Meeting, Chicago Linguistic Society, 71-80 (Published by the Department of Linguistics, University of Chicago).

Petrick, S. (1965). A Recognition Procedure for Transformational Grammars. Unpublished Ph.D. dissertation, M.I.T.

Quillian, R. (1969). "The Teachable Language Comprehender. A Simulation Program and Theory of Language", Communications of the Association for Computing Machinery 12, 459-476.

Rosenbaum, P. (1967). The Grammar of English Predicate Complement Constructions. Cambridge, Mass.: The M.I.T. Press.

Ross, J. (1967a). "Auxiliaries as Main Verbs". Mimeo, M.I.T.

Ross, J. (1967b). Constraints on Variables in Syntax. Unpublished Ph.D. dissertation, M.I.T.

Vendler, Z. (1961). "Order of Adjectives", report from the Transformations and Discourse Analysis Project, University of Pennsylvania.

Zwicky, A., J. Friedman, B. Hall, and D. Walker. "The MITRE Syntactic Analysis Procedure for Transformational Grammars", Proceedings of the Fall Joint Computer Conference, 1965, 317-326. Washington, D.C.: Spartan Books.

POSTSCRIPT

Since this review was written, I have been informed that the "unpublished" experiment by Bever, Kirk, and Lackner described on page   has been published as "The underlying structures of sentences are the primary units of immediate speech processing", in Perception and Psychophysics, 1969, Vol. 5(4), pp. 225-234 (this particular issue of Perception and Psychophysics was inaccessible to me at the time of writing).   The published version includes a complete list of the sentences used, with the structural analyses assumed, and a complete description of the experimental methodology.   Also, a revised version of Sections two and four of the Survey, with which the review was primarily concerned, will appear in published form in J. Hayes (ed.), Language and Cognition, Englewood Cliffs, New Jersey: Prentice-Hall, in press.   The revisions will deal in part with some of the issues raised in the review.   Thus some aspects of the review are obsolete, in the sense that they are irrelevant to continuing discussion, although I believe them to be fair with respect to the work under review. It is possible that a revised version of this review will be published in response to the articles named; as regards the present version, the reader is requested to respect the injunction on the inside front cover of the Notes. Comme s and criticisms addressed to me are, of course, welco.   and solicited.