

**NOTICE WARNING CONCERNING COPYRIGHT RESTRICTIONS:**  
The copyright law of the United States (title 17, U.S. Code) governs the making of photocopies or other reproductions of copyrighted material. Any copying of this document without permission of its author may be prohibited by law.

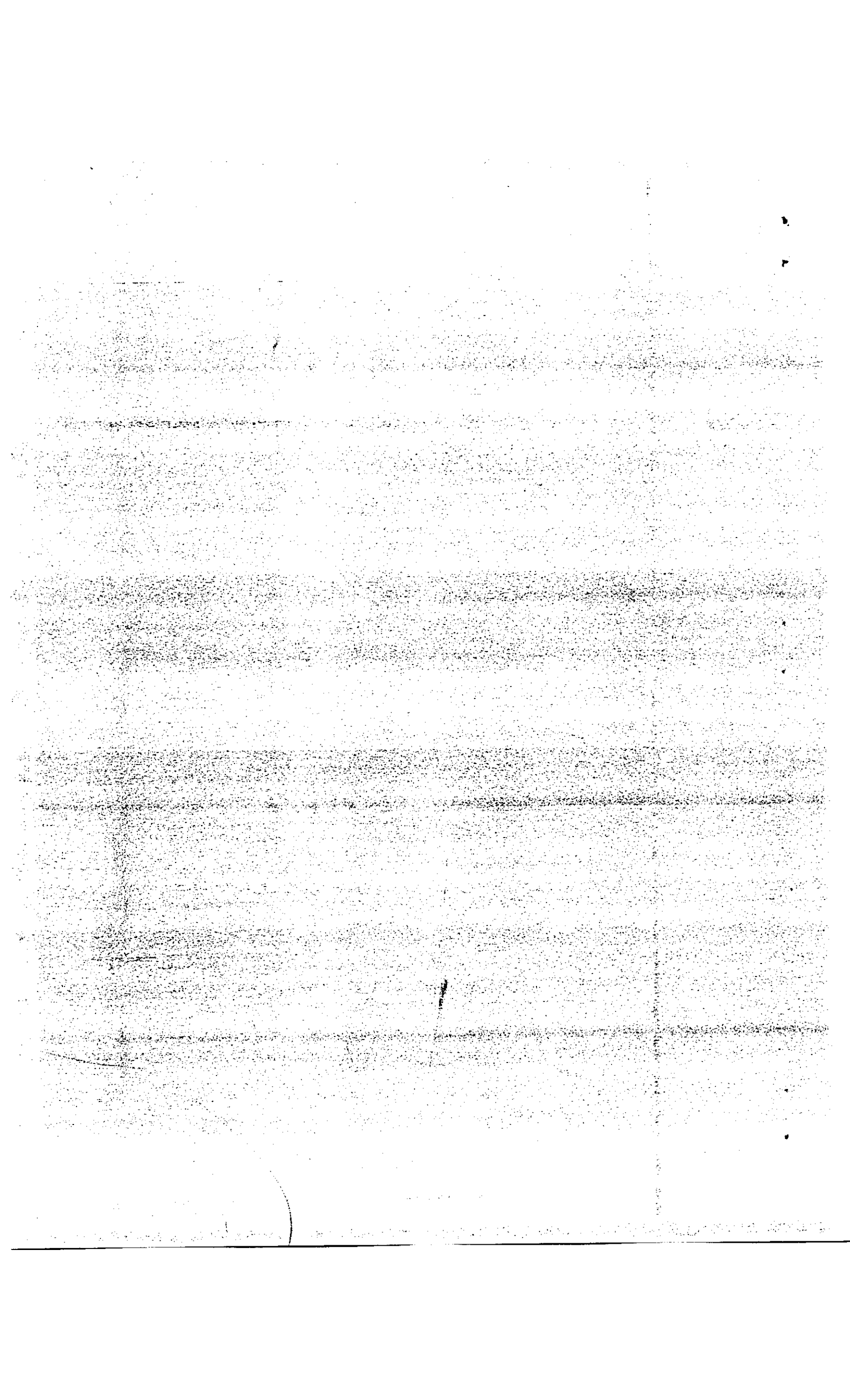
from: J. R. Hayes, ed., Cognition  
and the Development of Language;  
New York: John Wiley and Sons.  
In press.

ON TWO HYPOTHESES CONCERNING PSYCHOLINGUISTICS

W. C. Watt

March 1969

This work was supported by the Advanced Research Projects Agency  
of the Office of the Secretary of Defense under Contract No.  
F 44620-67-C-0058.



\* This work was supported by the Advanced Research Projects Agency of the Office of the Secretary of Defense under Contract No. F 44620-67-C-0058. I acknowledge with pleasure the helpful comments of H. H. and E. V. Clark.

1 The 'Correlation Hypothesis' (or Hypotheses) must not be confused with the 'Correspondence Hypothesis', Hayes's term (Introduction to this volume) for the hypothesis which holds that "...the derivational history of the sentence corresponds step by step to the sequence of psychological processes which are executed when a person processes the sentence" (ibid., p. 00).

The second hypothesis, the 'Strong Inclusion Hypothesis', has had a much more peculiar history, since it appeared in early writings as a tacit assumption, and then yielded in later writings to the corresponding null assumption, without ever (to my knowledge) having been exposed to light, examined, and accepted or rejected. As a result, the downfall of the Strong Inclusion Hypothesis (which I do not think many, apprised of the nature of the Hypothesis, would now dispute) has come about without occasioning a general awareness of the consequences of this downfall. And this is the reason for our here examining the Hypothesis in some detail, and the corresponding null hypothesis as well; we will thus be able to determine what corollaries, if any, had the Strong Inclusion Hypothesis as their sole supporter.

It is probably superfluous, but I should like to note at the outset that although I will try to mount as strong a case as possible for each of my arguments, on the other hand presently-available evidence on almost any psycholinguistic point is so scanty as to blunt any claim that this or that hypothesis has truly been disconfirmed. (If enough future evidence is to be concentrated on a given point, research directed at gathering that evidence must be provoked by formulating hypotheses which seem worth rejecting or upholding.)

The paper proceeds in three sections. Sections 1 and 2 examine, independently, the two hypotheses; Section 3 concludes.

1. The Correlation Hypothesis.

1.0. The 'Correlation Hypothesis' as I will term it concerns the nature of the relation between the grammar constructed for e.g. English in accordance with principles of linguistics, and the mental grammar of the ('idealized') speaker of English. As far as modern linguistics is concerned<sup>2</sup> the contemporary approximative notion of the nature of the linguistic grammar (hereafter, LG) was born simultaneously with the notion that speakers do have mental grammars (MG) in some sense similar to that LG. Indeed, LG and ('idealized') MG were tied very closely together, for the essential business of the LG was seen as accommodating in an ('idealized') MG the linguistic information available to the ('idealized') native speaker<sup>3</sup>. This view of the LG/MG relation has been encapsulated in a version of the Correlation Hypothesis: the LG differs from the MG only in that the LG represents "axiomatically" (Fodor and Garrett, 1966, p. 139) what the MG represents in some other way. LG and MG generate the same set of sentences and assign those sentences the same structural descriptions and derivational paths; the sentences transformationally related (derived from equivalent or identical Deep Structures) in the LG are similarly related in the MG; and in general LG and MG differ only representationally<sup>4</sup>.

The MG in question could only be a Competence Grammar (CG), as contrasted with a means of making and parsing sentences (or worse yet, utterances). Further clarifying this contrast, we see that between our MG and what we say there must be interposed a performative mechanism, whose 'use' illustrates one sort of performative factor, whose 'misuse' betrays another sort. Having access in some way to the MG, this performance

<sup>2</sup> For coverage of some notions ancestral to this idea, see Chomsky (1966).

<sup>3</sup> I am using the annoyingly-parenthesized term "('idealized')" to indicate that it need not be held that the MG thus qualified is, in fact, the possession of any one English-speaker.

Individual MG differ from each other in idiosyncratic ways corresponding to idiosyncratically-affective facets of the individual language-acquisition process. Trivially, no two speakers have exactly the same lexicon--same words with exactly the same definitions--(Harris, 1954, p. 150); less trivially, speakers differ on minor syntactic points. Unless the LG is specifically geared to capture just the MG of some particular speaker (the grammarian's), then it will probably capture the MG of no speaker. Moreover, if speakers vary much in their MG then the LG cannot represent a composite speaker without losing some of the very 'simplicity' that contributes to its adequacy as a (putative) representation of the MG.

This point is related to, but quite different from, two other points due to Chomsky. First, it is evident that the speaker's MG and more generally his linguistic competence must be considerably abstracted from his actual performance, since the latter's set of "utterances" only intersects with the set of sentences of competence: there must be (infinitely many) sentences too complex or too stilted or just too long ever to be used--no performance will ever contain them--; and then actual discourse consists less of "sentences" as such than of "...interrupted fragments, false starts, lapses, slurring, and other phenomena..." (Chomsky, 1962, p. 531).  
The second point is that the LG is an "idealized" MG (Chomsky, 1967, p. 398), in that the LG expresses in compact and formal notation what the MG doubtless expresses in some other (and far more abstract) way. Contrasting my point with the two just outlined, I was claiming that the performance of speakers belies, and the LG idealizes, not a set of identical MG, but rather a set of slightly-differing MG. The point is obvious enough, and will not be referred to again.

<sup>4</sup> "...every speaker of a language has mastered and internalized a generative grammar that expresses his knowledge of his language." (Chomsky, 1965, p. 8.) "...a reasonable model of language use will incorporate, as a basic component, the generative grammar that expresses the speaker-hearer's knowledge of the language..." (ibid., p. 9). "A general linguistic theory...must therefore be regarded as a specific hypothesis, of an essentially rationalist cast, as to the nature of mental structures and processes"(ibid., p. 53). "The generative grammar represents the information concerning sentence structure that is available, in principle, to one who has acquired the language. It indicates how, ideally--leaving out any limitations of memory, distractions, etc.--he would understand a sentence...." (1963, pp. 326f.). "...the technical term 'competence' refers to the ability of the idealized speaker-hearer to associate sounds and meanings strictly in accordance with the rules of his language. The grammar of a language, as a model for idealized competence, establishes a certain relation between sound and meaning--between phonetic and semantic representations. ... To discover this grammar is the primary goal of the linguistic investigation of a particular language." (1967, p. 398).

Chomsky's conception of the nature of the speaker's MG must be sharply distinguished from two other conceptions that abut on it. First, he does assert that speakers have an MG, not just the products generated in accordance with the corresponding LG: they have a grammar, not a set of "grammatical" sentences. They neither have a mere inventory of words and phrases with their grammatical (combinatory) properties--this notion, which may be ascribed to de Saussure, Chomsky rejects in (1963, p. 328)--nor do they have some inventory of sentence-blanks into which appropriate words can be put, a notion which may be ascribed to Bolinger and which Chomsky rejects in (1964, p. 54, n.). That is, Chomsky means just what he says when he refers to "...the rules of the grammar represented in the brain..." (1963, p. 330). But on the other hand he specifically enjoins against the interpretation that, therefore, the MG's model of generation is a model of production (or, reversed somehow, a model of reception). For further remarks on this, see below.



mechanism composes sentences whose analyses accord with those given them by the MG, and imposes on input sentences the analyses they would have had if generated by the MG. In so doing, the performance mechanism obeys what we might call 'positive performative factors', in that knowing how to speak English involves knowing how to select and utter, from the vast synonymous paradigms generally offered by the language, a sentence that can be controlled in speaking and understood by the auditor; and a sentence, moreover, that through sequencing its main elements focusses attention in the way desired<sup>5</sup>. The mechanism is also constrained to reveal what we might call 'negative performative factors', in that its (mental) productions, in the course of being realized as sound, are often distorted by memory lapses and other performative difficulties<sup>6</sup>.

The performance mechanism must produce sentences (apart from the 'negative performative factors') which will be as if they had been produced by invoking the rules of the MG in their prescribed order; but their composition need not have consisted of actually invoking those rules in that way.<sup>7</sup> To take a simple example, in the LG and in the CG, Pronominalization must follow Passivization, in order that an abstract structure of the form "John killed John" may be converted into the grammatical "John was killed by himself" rather than into the ungrammatical "\*Himself was killed by John". But the accepted model of the over-all linguistic capacity does not insist that speakers, in composing such sentences, actually first ready an intermediate structure through Passivization, and only then invoke Pronominalization. (This would be the fallacy Propter hoc, ergo post hoc.)

In sum, taking as our text the grammatical sentence "That that that that Byron detested Hunt distressed Shelley saddened Clare will perturb

<sup>5</sup> See for example Chomsky (1965, pp. 11; 221, n. 32; 224-5, n. 9) for further references to these aspects of performance; cf. Firbas (1959; 1964) and Halliday (esp. 1967).

<sup>6</sup> As see Chomsky (1965, pp. 4, 10-15).

<sup>7</sup> Present-day linguistic terminology has bred endless misunderstanding of this issue, in that "generate", "rewrite as", and so on, are naturally and persistently misconstrued as describing production rules, the result being that the generative grammar is regarded as, in fact, a set of instructions which a speaker follows in order to compose sentences. This has happened despite the fact that from the very beginning Chomsky emphasized that this was a misconstrual: e.g., "A grammar does not tell us how to synthesize a specific utterance; it does not tell us how to analyze a particular given utterance. In fact, these two tasks...are both outside the scope of grammars..." (1957, p. 48). He has returned to this theme many times (e.g. 1964, p. 10; 1965, pp. 3-9 and elsewhere), restating it with increasing asperity. Most recently (1967, p. 399) he put it in these words: "...although we may describe the grammar G as a system of processes and rules that apply in a certain order to relate sound and meaning, we are not entitled to take this as a description of the successive acts of a performance model...--in fact, it would be quite absurd to do so. ... The grammatical rules that generate phonetic representations of signals with their semantic interpretations do not constitute a model for the production of sentences, although any such model must incorporate the system of grammatical rules."

Even so, the mistake has been made many times, even by Katz, if I read him correctly. Thus the latter scholar at one time felt at least that in composing a sentence a speaker necessarily first produces a full syntactic structure, and then "...utilizes the phonological component of his linguistic description to produce a phonetic shape for it" (1964, p. 132; reiterated on the following page, and see especially the footnote on that page). Vestiges of this notion are still active.

Note that, while the accepted notion of performance does not insist that the MG's transformational rules be invoked one-by-one in sequence, neither does it insist that they not be, as see just below in (1.1.1.).

<sup>5</sup> See for example Chomsky (1965, pp. 11; 221, n. 32; 224-5, n. 9) for further references to these aspects of performance; cf. Firbas (1959; 1964) and Halliday (esp. 1967).

<sup>6</sup> As see Chomsky (1965, pp. 4, 10-15).

<sup>7</sup> Present-day linguistic terminology has bred endless misunderstanding of this issue, in that "generate", "rewrite as", and so on, are naturally and persistently misconstrued as describing production rules, the result being that the generative grammar is regarded as, in fact, a set of instructions which a speaker follows in order to compose sentences. This has happened despite the fact that from the very beginning Chomsky emphasized that this was a misconstrual: e.g., "A grammar does not tell us how to synthesize a specific utterance; it does not tell us how to analyze a particular given utterance. In fact, these two tasks...are both outside the scope of grammars..." (1957, p. 48). He has returned to this theme many times (e.g. 1964, p. 10; 1965, pp. 3-9 and elsewhere), restating it with increasing asperity. Most recently (1967, p. 399) he put it in these words: "...although we may describe the grammar G as a system of processes and rules that apply in a certain order to relate sound and meaning, we are not entitled to take this as a description of the successive acts of a performance model...--in fact, it would be quite absurd to do so. ... The grammatical rules that generate phonetic representations of signals with their semantic interpretations do not constitute a model for the production of sentences, although any such model must incorporate the system of grammatical rules."

Even so, the mistake has been made many times, even by Katz, if I read him correctly. Thus the latter scholar at one time felt at least that in composing a sentence a speaker necessarily first produces a full syntactic structure, and then "...utilizes the phonological component of his linguistic description to produce a phonetic shape for it" (1964, p. 132; reiterated on the following page, and see especially the footnote on that page). Vestiges of this notion are still active.

Note that, while the accepted notion of performance does not insist that the MG's transformational rules be invoked one-by-one in sequence, neither does it insist that they not be, as see just below in (1.1.1.).

<sup>5</sup> See for example Chomsky (1965, pp. 11; 221, n. 32; 224-5, n. 9) for further references to these aspects of performance; cf. Firbas (1959; 1964) and Halliday (esp. 1967).

<sup>6</sup> As see Chomsky (1965, pp. 4, 10-15).

<sup>7</sup> Present-day linguistic terminology has bred endless misunderstanding of this issue, in that "generate", "rewrite as", and so on, are naturally and persistently misconstrued as describing production rules, the result being that the generative grammar is regarded as, in fact, a set of instructions which a speaker follows in order to compose sentences. This has happened despite the fact that from the very beginning Chomsky emphasized that this was a misconstrual: e.g., "A grammar does not tell us how to synthesize a specific utterance; it does not tell us how to analyze a particular given utterance. In fact, these two tasks...are both outside the scope of grammars..." (1957, p. 48). He has returned to this theme many times (e.g. 1964, p. 10; 1965, pp. 3-9 and elsewhere), restating it with increasing asperity. Most recently (1967, p. 399) he put it in these words: "...although we may describe the grammar G as a system of processes and rules that apply in a certain order to relate sound and meaning, we are not entitled to take this as a description of the successive acts of a performance model...--in fact, it would be quite absurd to do so. ... The grammatical rules that generate phonetic representations of signals with their semantic interpretations do not constitute a model for the production of sentences, although any such model must incorporate the system of grammatical rules."

Even so, the mistake has been made many times, even by Katz, if I read him correctly. Thus the latter scholar at one time felt at least that in composing a sentence a speaker necessarily first produces a full syntactic structure, and then "...utilizes the phonological component of his linguistic description to produce a phonetic shape for it" (1964, p. 132; reiterated on the following page, and see especially the footnote on that page). Vestiges of this notion are still active.

Note that, while the accepted notion of performance does not insist that the MG's transformational rules be invoked one-by-one in sequence, neither does it insist that they not be, as see just below in (1.1.1.).

Dr. Psoriasis would have amused Byron", we see that the CG accounts for the correct surface analysis of this sentence and for its Deep Structure, transformational derivation, and meaning; while the performative mechanism accounts for how the sentence was actually composed, in so doing showing which 'positive performative factors' were disobeyed and which 'negative' ones were thereby run afoul of, hence in effect accounting for the fact that a paraphrase of this sentence is likely to be easier both to utter and to comprehend: e.g., "It would have amused Byron that Dr. Psoriasis will be perturbed at Clare's being saddened by Shelley's distress at Byron's detestation of Hunt".

We have sketched, then, a version of the Correlation Hypothesis that we will call  $CH_{CG}$ , a version based on the presumption that the basic human MG is a Competence Grammar, served by and in service to a performance mechanism, which grammar is the grammar that the LG is an (attempted) account of. This version may be expressed as:

$$MG \equiv CG = LG$$

where "=" means "is isomorphic to", "is essentially equivalent to", "is no more than representationally different from", "is idealized by", "is axiomatized by", or something of the sort.

We have explicitly acknowledged that the MG need not be identical to the performance mechanism, and we cannot demand that the composition of sentences consist of invoking the MG's rules one by one. However we have already noted, in n. 4, that a "reasonable model of language use", with Chomsky, "will incorporate, as a basic component, the generative grammar that expresses the speaker-hearer's knowledge of the language"<sup>8</sup>. This

<sup>8</sup> A later comment (Chomsky, 1967, p. 399): "...it is important to distinguish clearly between the function and properties of the perceptual model PM [the input aspect of the performance model] and the competence model G that it incorporates. ...PM makes use of much information beyond [italics mine, wcv] the intrinsic sound-meaning association determined by the grammar G...".

must mean that, in some way, the performance mechanism refers to the MG in the process of composing and parsing sentences. (If it did not, then the MG would have no use at all.) This in turn means that, while it is a mistake to identify the generations of the MG with the productions of the performance mechanism, on the other hand it would be a mistake of equal magnitude to completely dissociate the two. And, in fact, if the MG is actually "incorporated" into the "model of language use", then the relation must be rather close.<sup>9</sup> We can imagine a very abstract relation between the two components which, still, is a very close relation. For example, suppose that the set of generations of the MG were finite and that the full generation of each sentence were available on a mental 5x8 index-card; and that the performance mechanism had "only" to find the right card to connect incoming sound to interpretation or intended meaning to outgoing sound. This notion completely divorces the MG's rules from the rules of performance, since the latter rules consist of ways of checking 5x8 cards, in whatever manner is most efficient, from bottom to top or top to bottom or from the middle in both directions. Even so, a sentence which the MG generates with a very complex derivation would (to continue the analogy) occupy more of the 5x8 card than would a simple derivation; and ceteris paribus should take more performative time and effort to recover. Thus, again ceteris paribus, "generative MG complexity" would bear a direct relation, even in this maximally abstract 'index-card' notion, to "performative complexity". Then if everything else were equal we could experimentally observe performative complexity, establish whether our account of MG complexity corresponded properly, and adjust the MG derivations accordingly. We would have to make adjustments, if we believed our

<sup>9</sup> "The psychological plausibility of a transformational model of the language user would be strengthened, of course, if it could be shown that our performance on tasks requiring an appreciation of the structure of transformed sentences is some function of the nature, number, and complexity of the grammatical transformations involved" (Miller and Chomsky, 1963, p. 481).



experimental results, since the alternative would be to exile the MG to an absurdly peripheral status in which its existence would be defined to be beyond even the most indirect proof: exactly as with the soul.

As it happens, everything else is not equal (we will shortly devote a few pages to this topic), and so the proposed measure of related complexity is somewhat snarled. But, as is obvious, this does not free us from the burden of explicating that relation; we must ultimately be able to show what factors complicate the relation, and to show this in such a way that, when everything else is made equal (by being taken into account), then the relation will indeed hold: performatively complex sentences will be complex in the MG. If the relation still does not hold when we have taken everything into account that is at all plausible, then we must again think of changing the MG to fit the facts, if this can be done.

Our purpose in the rest of this Section on the Correlation Hypotheses can now be summarized: we will show that there are many discrepancies between reasonably-hypothesized performative complexities and MG (= current LG) complexities; we will examine a variety of possible complicating and distortive factors that must be taken into account by any attempt to explain these discrepancies without changing the MG; we will conclude that these factors are not enough, or are not germane; we will consider what alterations to the MG are thus implied; we will conclude that, apparently, these alterations cannot be made without changing the MG into something other than a CG, that is to say other than a grammar "=" to the LG; and we will then suggest that the MG is, by virtue of these alterations, an "Abstract Performative Grammar", whose nature we will try, very tentatively, to sketch.

1.1. On the CH<sub>CG</sub>

1.1.1. If we compare Chomsky's comments of nn.7 and 9, above, we see that while the general thesis that production mirrors generation is labeled "absurd", nonetheless the issue of whether or not producing sentences involves invoking the (generative) sequence of transformations, is left open. There is no paradox here. The notion that in producing a sentence one converts the sentence's MG-generation directly into a production algorithm, producing the sentence by beginning with 'S' and ending with a phoneme-to-sound conversion, may be naturally divided into separate notions concerning separate parts of the derivation. One notion holds that a speaker produces a sentence's Deep Structure by starting with 'S', expanding 'S' to 'NP + Predicate', and expanding each node in turn until the categorial component of the Base is exited from and the lexical items are added, resulting in a specification of the sentence's full Deep Structure. A second notion, distinct from the first, holds that one then takes that Deep Structure and produces from it a surface structure by rewriting the Deep Structure through the successive application of appropriate transformations purloined from the MG. These two notions are distinct in the sense that the first can be rejected on grounds that leave the second unscathed. Within the conventional sense of generative grammar, syntax is generative and semantics only interpretive; and before lexical substitution takes place the Deep Structure's meaning is not fully specified, in that any lexical item, of whatever meaning, can be substituted for a given node unless the substitution (taking context into account) is unlawful. Thus the first notion insists, utterly contrary to sense, that one must initially "...select sentence type, then determine subcategories, etc.,

finally, at the last stage, deciding what he is going to talk about..." (Chomsky, 1965, p. 197, n. 8). Given the currently-accepted notion of grammar (which is not in dispute here)<sup>10</sup> the first notion is thus utterly ridiculous, as Chomsky makes manifest. But the second notion is not subject to the same criticism, since it only maintains that a speaker uses the rules of the MG to alter a structure whose meaning is already established into another structure with the same meaning. The second notion is not at all contrary to sense, and it was obviously worth subjecting to experimentation. Since mental activities are not accessible to direct observation, however, it was tested only in the weaker "5x8" version sketched above: the experiments were designed to show, not whether performance mirrored competence, but whether performative complexity mirrored competence complexity. This is the weaker version in the sense that if it were upheld the stronger version would remain still in doubt; but (of more immediate importance) if it were disconfirmed, then the stronger version would be seriously threatened.

The weaker version easily withstood early experimentation. These results were first summarized, cautiously, in Miller and Chomsky (1963, pp. 481-483); some of them, with later ones, have more recently been summarized by Fodor and Garrett (1966, pp. 143-148). The latter exposition may be consulted for details; here it must suffice to say that these experiments indeed seemed "impressive successes", as Fodor and Garrett put it (ibid., p. 143); they showed, or seemed to show, for example, that Active sentences, which in the CG are more simply derived than Passives, take performatively less time to produce (Miller, 1962; Miller and McKean, 1964), are performatively easier to remember (Mehler, 1963), and require

<sup>10</sup> It is, of course, in dispute elsewhere; representative recent papers in marked opposition to major aspects of the Aspects model(s) are Ross, 1967; Lakoff, 1968; McCawley, 1968; and Fillmore, 1968. These works, and others by these and other like-minded authors cited in them, do not present a unified front; but they are similar in that they essentially propose that behind the Deep Structures of Aspects there lies much more abstract complex structure than could be accommodated at all convincingly in an Aspects-like Base component. These Abyssal Structures clearly need transformations just to reach the level of Chomsky's Deep Structure. I have used one of the two Aspects models because it provides a firm and moderately well-known basis in linguistics for the variety of psycholinguistic comments on which this paper is partly grounded, and because these comments are not in general falsified by any of the new models currently under debate. (The comments would have to be greatly altered, but not so as to be replaced by their contraries.) Lastly, it seems possible that the psycholinguistic reality of the 'abyssal structures' is different from that of the 'deep structures'. The deep structures to a large extent have forms influenced by the notion of 'kernel', or 'elementary sentence': the canonical and maximally simple sentential surface (Harris, 1957, pp. 334-336; 1965, pp. 364-367 et passim; Chomsky, 1957, pp. 45f, 61-84; 1965, pp. 17f.). But the abyssal structures scarcely betray any such influence. If there is any psycholinguistic reality to the 'kernel sentence', as has sometimes been suggested (Chomsky, 1955, p. 23; 1957, p. 104, n. 11; Miller and Chomsky, 1963, p. 483; Chomsky, 1965, p. 18), then/ the 'abyssal structures' might indeed have a psycholinguistic reality, even under the CH<sub>CG</sub>, that is distinctively more abstract in some way. (A more abstract psycholinguistic reality could be attained, outside the MG proper, in the 'archival linguistic competence' proposed below in [1.2.2.4].) Some of the examples cited below, as affecting the plausibility of CH<sub>CG</sub>, are drawn from the literature opposed to Aspects; but in these cases it seems to me that the analyses proposed are incorporable into an Aspects-like grammar.

less memory-storage (Savin and Perchonock, 1965). Results like these were obtained for some (few) other sentence-types; and in general results were consistently favorable to the thesis that performative complexity mirrors the generative complexity of the CG (that is, of the current version of the LG).

But, Fodor and Garrett claim, other experimental results undermine these early successes; results which show, for sentence 'A' more complexly-derived in the CG than sentence 'B', either that 'A' and 'B' are performatively equally-complex, or that 'A' is performatively the less complex of the two. However, some of the results they cite are afflicted with the evanescence that seems to characterize so much work in this field<sup>11</sup>. In fact, only one of their cited pieces of evidence seems valid; it is that the sentence "John phoned the girl up" is more complex in the LG than the sentence "John phoned up the girl", though performatively they are of equal complexity.

That is, their counter-evidence consists of the one example, plus the fact that "John runs faster than Bill does" and "John runs faster than Bill" are performatively of equal complexity, which could be explained by the facts that while the former is less complexly-derived in the CG (as see the discussion of 'Deletion' below), on the other hand the latter is shorter.<sup>12</sup>

However, though they had no solid data, I think Fodor and Garrett were right in concluding that the evidence favoring a close relation between generative and performative complexity was weak, and in inferring, from the threat to this relation, a threat to the stronger notion that performance mirrors generation rule-by-rule. From these conclusions they drew the further inference that the relation between the CG and the performance mechanism must henceforward be assumed to be "more abstract", a term which

- <sup>11</sup> They cite a result due to Mehler (ibid.) and to Miller and McKean (ibid.): more complex auxiliaries are no higher in performative complexity than are simple ones. But this result is now put into doubt by some work by Clark and Stafford (in press). They also cite an unpublished result due to Fodor, Jenkins, and Saporta: 'John runs faster than Bill runs' is harder to recognize (process) than 'John runs faster than Bill' or 'John runs faster than Bill does'. But Fodor and Garrett provide for this result the partial explanation (p. 150) that English-speakers find the first sentence strange, hence hesitate over it.
- <sup>12</sup> They also (p. 150) cite Slobin (1963) as authority for the 'fact' that truncated Passives ("John was found") take less time to process than do full Passives (no example). However, I have it from Slobin (personal communication, October 2, 1968) that neither in the cited reference nor anywhere else has he demonstrated this 'fact', or tried to do so. Fodor and Garrett in a subsequent paper (1967, p. 290) cite this 'fact' as an assumption: or rather they (quite correctly) state its contrary as "counter-intuitive".

they leave quite vague but which seems to entail that the performative recognition routine include (or have access to) all the rules of the CG, but must additionally include special performative rules for imposing surface structure on raw in-coming sentences.<sup>13</sup> (As many readers will recognize, this proposal is strongly reminiscent of one due to Miller and Chomsky [1963, pp. 476-480].)

In a second attack on the problems of complexity-relations (1967) Fodor and Garrett report the results of an experiment and suggest two ways in which those results could be accounted for. Identifying the notion that performance complexity must mirror the transformational complexity of the CG as the "Derivational Theory of Complexity", or DTC, they return to their example of the truncated Passives and point out that according to DTC if sentence X has a more complex transformational derivation in the CG than sentence Y, then DTC must predict X to be performatively more complex even though X's extra complexity consists of the fact that X is derived from Y by deleting part of Y. DTC must predict "The boy was hit by someone" to be less complex (p. 290) than the shorter "The boy was hit" which derives (with Fodor and Garrett) from the former sentence by transformationally deleting "by someone".<sup>14</sup> This DTC prediction is counter to sense and probably incorrect; since DTC is discredited, when it makes a correct prediction it probably does so by accident. Thus their experimental result that (e.g.) "The man the dog bit died" is more complex than "The man whom

- <sup>13</sup> They specifically suggest (p. 140) that the CG/performance relation is more "abstract" than the 'Analysis-by-Synthesis' model would imply (this is scarcely surprising). Some of their other comments are sensible enough (e.g., the discussion on p. 141 of the necessity for the recognition routine to recognize NP of derived phrase-structures); but (since the recognizer seems to incorporate or have full access to the CG) they do not seem to demonstrate a "particularly abstract relation between the grammar and the recognizer" (ibid.) such as they claim characterizes the human language faculty.
- <sup>14</sup> These examples will be returned to below.



the dog bit died" from which it derives by deleting "whom"--this result, predicted accidentally by DTC, is in want of a more satisfactory explanation.

The first explanation to be found in their pages is a simple and persuasive one. They conjecture (pp. 290f.) that relative pronouns ease parsing, so that their absence increases performative complexity. (It is their absence, not the Deletion that brings about that absence, that results in the complexity-increase.) As they say, the presence of the relative pronoun "...is evidence of the application of an embedding transformation, and that transformation can apply only where certain grammatical relations hold between the noun phrases in the sentence." That is, if I understand what they are saying (their subsequent remarks are confusing), a function of relatives is to provide a strong clue to the derived phrase-structure of an incoming sentence, enabling the auditor to impose a unique "labeled bracketing" from which, in turn, he can retrieve (however he does this) the correct Deep Structure and from that an understanding. The labeled bracketing, for instance, takes a raw incoming string (relatives italicized) like "The tiger which the lion that the gorilla chased killed was ferocious" and turns it into the derived (surface) phrase-structure, with its characteristic polytomy, that is shown simplified in Figure 1. That derived structure is evidence of the action of the transformations that derived it from the deeper structure, simplified, of Figure 2; when the structure of Figure 2 has been gained, the sentence has been parsed deeply enough to receive a semantic interpretation.

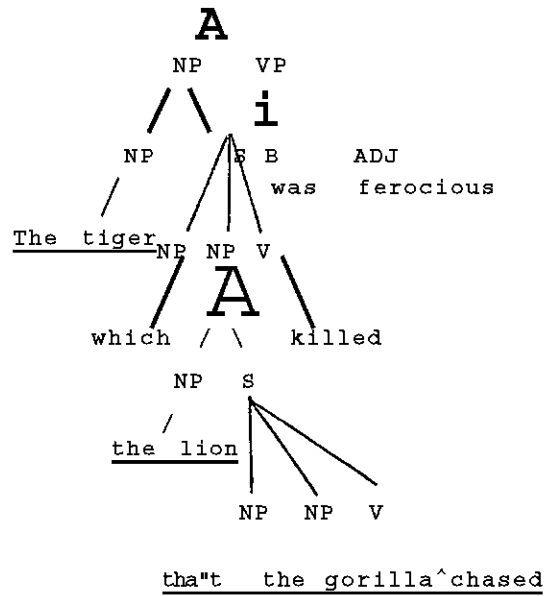


Figure 1

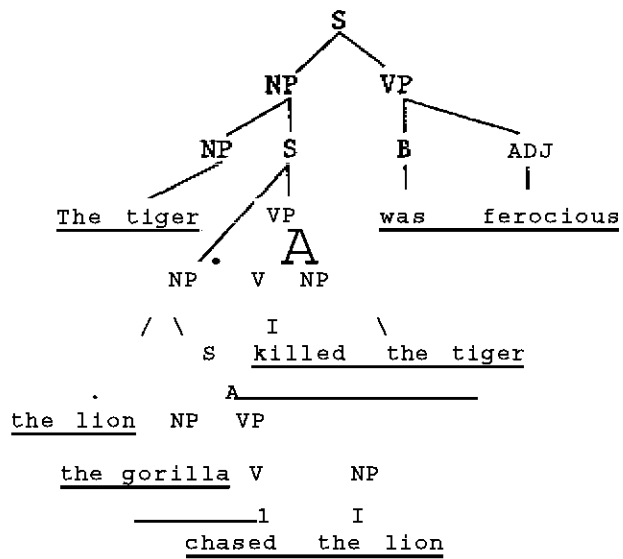


Figure 2

They do not go at all deeply into how the relatives actually help here, but presumably they help in this way: "The tiger which the lion..." as an opening sequence predicts that the Predicate for the Subject "The tiger" will be postponed until the end of a clause modifying "The tiger", which clause can tentatively be assigned a Deep Structure including a Verb Phrase in which "the tiger" will appear as an Object. The Subject of that Verb Phrase may or may not be "the lion". When "...that the gorilla..." is added there is added the prediction that the Predicate for the Subject "the lion" will be postponed until the end of a clause modifying "the lion", which clause will have a Deep Structure including a Verb Phrase (whose Subject may or may not be "the gorilla") in which "the lion" will appear as an Object. The tentatively-assigned Deep Structures are, simplified, as in Figure 3. As we

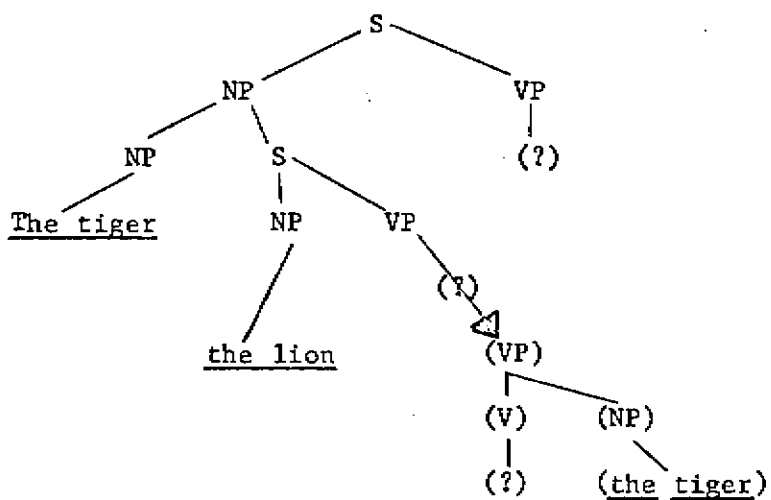
(Put Figure 3 about here)

see, a fair amount of Deep Structure is assigned. Moreover, a further tentative assignment would make each  $[ \rightarrow ]$  VP identical with the VP  $[ \rightarrow ]$  node already known to dominate it (so that, e.g. for the second example in Figure 3, the anticipated Deep Structure would be that of Figure 4.) We

(Put Figure 4 about here)

might expect probability--that is, frequency--to play a part here; indeed, without it, some unique assignments would be out of the question until virtually the whole of the sentence had been processed; notice that "The tiger which the lion that the gorilla..." could also, very rarely, serve to open e.g. "The tiger which the lion that the gorilla strangled ate, I was glad to see the last of"; and this sentence has little to do with the tentatively assigned Deep Structure of Figure 3b. That is, a very early unique tentative

(a) "The tiger which the lion...."



(b) "The tiger which the lion that the gorilla...."

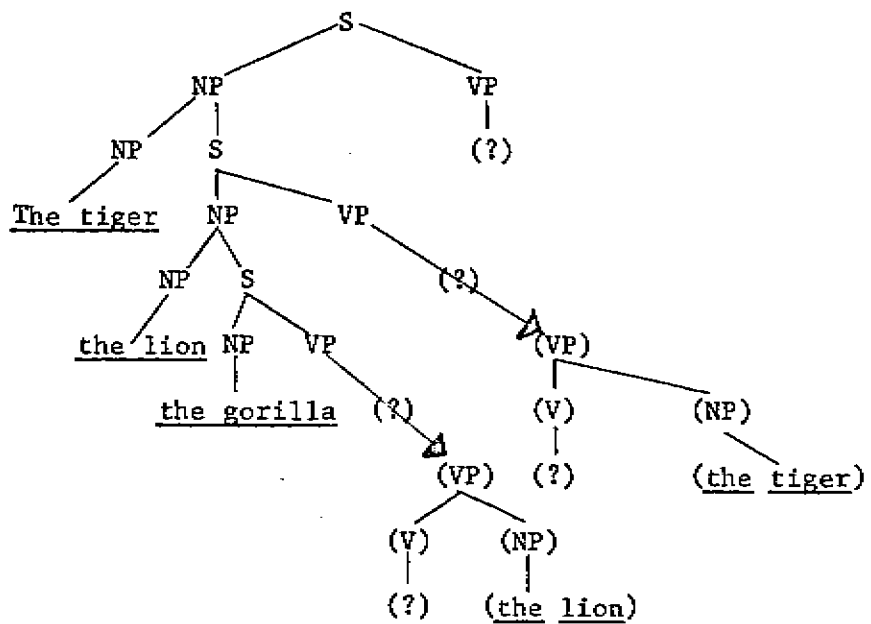


Figure 3

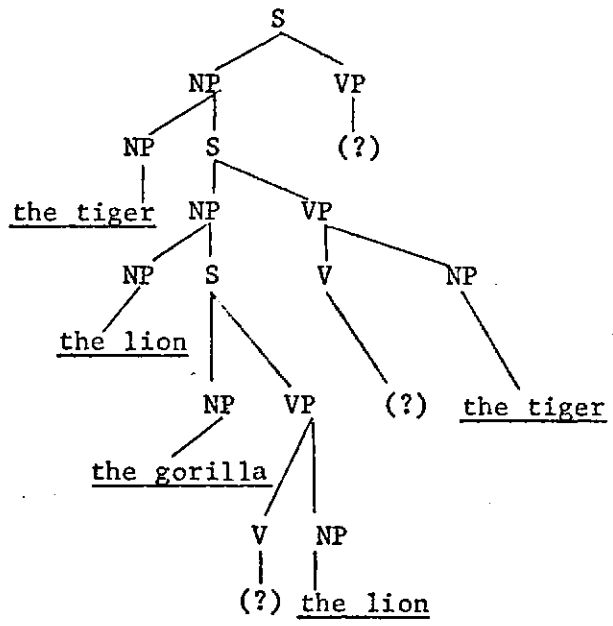


Figure 4

assignment is often unwarranted except psycholinguistically. We might then suppose that the rare and thus mis-predicted sentence-type is parsed after rescinding a prior unique tentative assignment.

In any case, it is clear I think that presence of the relatives greatly reduces the number of high-probability alternatives; certainly removal of the relatives increases that number. From "The tiger the lion the gorilla..." no part of the ultimate correct structure can be tentatively assigned except the (presumptive) Subjecthood of "The tiger"; the string of NP could just as well open a sentence like "The tiger the lion the gorilla and the pangolin make better pets than the cobra or the Gaboon viper" (see Fodor and Garrett's sentence 11, p. 290.) Thus correct Deep Structure cannot tentatively be assigned until later in the in-coming string. Whence, presumably, the greater complexity inflicted by the absence of the relatives: for it seems quite reasonable to suppose that the longer the auditor must go without retrieving correct Deep Structure, the more unprocessed sentence he will have to hold in memory (having perhaps 'cleared the register' after a misassignment now vitiated), and the more effort will be required of him. That is, psycholinguistic parsing complexity increases with the amount of Deep Structure whose correct assignment is postponed; with the length of sentence over which the postponement must be carried; and with the complexity of misassignments whose rescission returns the processor to an earlier point in the sentence.

This, at least, is (roughly sketched) a theory that has some credibility; its formulation, above, was directly inspired by remarks of Fodor and Garrett in the paper cited. I will term this theory the 'Theory of Cumulative Assignments', or TCA; to summarize, it maintains that as we process a sentence (whether or not concurrently with our hearing/reading of it) we tentatively assign to it a surface-structure consistent with as much initial string as

it takes to uniquely determine such an assignment (which must, then, be fragmentary until the end of the sentence is reached, though that fragmentariness diminishes); this surface-assignment tacitly implies a deep structure from which that surface is transformationally derived; and--as to 'complexity'--the more quickly the correct surface (hence deep) structure can be assigned (the fewer constituents it takes to support a unique assignment) the more easily-parsed a sentence is: performatively, the less complex. (If Fodor and Garrett should refuse to champion TCA, I would do it.)

Notice, incidentally, that TCA predicts (rightly or wrongly) that certain sentences-types are performatively more complex than we might have guessed from their LG complexity alone. Suppose a sentence were such as to open with a sequence that systematically spawned roughly equiprobable assignments, so that no definite assignment could be made until later in the string. For example, the sequence "Floyd Thursby was shot by..." could open any of these four sentences: "Floyd Thursby was shot by a streetlight", "Floyd Thursby was shot by midnight", "Floyd Thursby was shot by misadventure", and "Floyd Thursby was shot by a vixen". If these four sentence-endings were equally probable, then the auditor would have no (noncontextual) reason to settle on any one of the four kinds of "by..." phrase that are possible here. It would follow directly that, though none of these sentences is linguistically ambiguous, the 'performative ambiguity' just described would increase the psycholinguistic complexity of each. Thus to the usual explanations of the complexity of the Passive we would be forced to add that of 'performative ambiguity', an aspect of TCA.

Fodor and Garrett's second explanation for their experimental results --which however they do not distinguish from their first--is essentially an unnecessary appendage to TCA. This explanation--their only explicit one--holds that in processing a sentence we render special homage to the sentence's verbs, making "direct inductions" of a set of "base configurations" on the basis of how each verb is classified in the lexicon "...according to the base structure configurations it may enter" (p.295). Let us term this variant the 'Theory of Verb-Dominated Direct Inductions', or TVDDI. As we see, TVDDI differs from TCA on two main issues: (1) TVDDI assigns special processing importance to verbs; (2) TVDDI insists, at every point 'v' in the sentence being analyzed, where 'v' is any verb, that the processing mechanism consult its lexicon, list all of the possible "base configurations" that that verb could fit into, and then return to the sentence under analysis, determining which of the possible "configurations" fits the already-analyzed surface of the sentence, and also, presumably, matching each listed possible configuration against new pieces of sentence, as these are analyzed, if those pieces could be commanded by the verb, as by being in the verb's complement. Thus TVDDI, to refer to our Figure 4, above, would have us get to the sentence's first verb, "chased", consult the lexicon, list all possible "configurations" that "chase" can have, compare each of these to the "configuration" actually assigned to that verb by the structure defined thus far, and then choose the "configuration(s)" (and attendant predictions) that match. (Whereas the alternative, a natural part of TCA, would have us, rather, consult the MG's lexicon to ascertain simply whether or not "chase" can occur in the



"configuration" already determined by the previously-analyzed structure. TCA could, in addition, have the processing mechanism simply leave blank any part of the sentence not yet analyzed, rather than, as with TVDDI, necessarily listing possibilities for such parts if those parts fall into potential complements for an already-reached verb. Thus, under TCA, if a verb (e.g., "chase") were revealed by the lexicon to have a "configuration" that met the requirements of the assigned structure, then the processing mechanism, once a "matching configuration" were found, would be oblivious to all other possibilities; while under TVDDI all possibilities would first be "listed", a match being attempted only then.)

TCA is to my mind the more plausible explanation of the processing of "chase" as integrated into the deep structure of Figure 4. There, the V is determined to be one that can take NP as a (direct or indirect) object; thus the verb "know", had it been met instead of "chase", would have been acceptable also, since "know" has a possible "configuration" that matches the requirements of the deep structure assigned thus far. Notice that in TCA there is a (partially-completed) deep structure: the processor is not constrained, as under TVDDI, to compare possible deep structure "configurations" only with derivable surfaces.^"\*

But on the other hand TVDDI is not utterly implausible as a model of the processing of extremely complex sentences: e.g., of those made very confusing by loss of several relative pronouns. For in these sentences the amount of structure assigned before the verb is reached may, owing to the complexity, be very slight, so that when a verb is reached the processor might possibly, if only in desperation, consult its lexicon, list possibilities, and then (lacking any deep-structure hypothesis) compare these possibilities against the surface to see whether that surface could be derived from one of them. That is, TVDDI may represent the strategy of the processing mechanism faute de mieux: failing TCA. TVDDI would thus be resorted to only when control over the sentence-processing had already been lost, and when the respondent was clutching at straws.

This much more limited role for Fodor and Garrett's TVDDI is, in fact, the most that could be supported by their own evidence. Their postulation of TVDDI was principally based (p. 295) on the fact that where a sentence had a verb whose complement elsewhere in English could be other

15 "...the program [that speakers] use to recover the grammatical structure of sentences...must consult a lexicon which classifies the verb in the sentence according to the base structure configurations it may enter. Second, it must run through each such deep structure configuration, asking whether the surface material in the sentence can be analyzed as a transformed version of that deep structure" [italics mine, wcw] (ibid., p. 295).

than the sort of complement--direct-object NP--it had in the sentence at hand, then among the errors that respondents made occasionally there would be an error in which the verb was, in fact, given the wrong sort of complement. (Fodor and Garrett remark [ibid.] that only in such cases do errors of this sort crop up, which fact, since they could crop up in other of their experimental sentences only at the cost of producing gross ungrammaticality, is not too surprising.) But this fact does not at all imply, as Fodor and Garrett think it does, that TVDDI is a basic processing maneuver used by English-speakers: it only implies that when TVDDI is used, it sometimes engenders mistakes. For if TVDDI were used consistently, as a basic strategy, sentences containing "versatile" verbs (capable of more than one complement-'configuration') should always be rather more complex than otherwise-similar sentences containing less versatile verbs. Such sentences are predicted even by TCA to be slightly more complex--two or three 'configurations' must be searched and compared with the sentence's accumulated Deep Structure--; but TVDDI must predict them to be much more complex: the several 'configurations' must be searched, then listed, then compared with a piece of surface-structure; and this, if it is a typical such piece of surface, will have undergone great distortion in the course of being transformed from the canonical--that is, Deep--form in which any lexical 'configuration' is given. (In Fodor and Garrett's experimental sentences the transformation-induced distortions were severe.) This complexity prediction must hold, if TVDDI is to be upheld as a basic processing strategy, even when presence of relatives makes processing easier. But on Fodor and Garrett's evidence this is simply not so: their sentences 1 and 9, the only ones containing

"versatile" verbs, are by their own criteria adjudged to be on the average less complex than sentence 2 (whose verb is not "versatile"), where all three sentences have their relatives.<sup>16</sup> That the "versatile" verbs have not been shown to cause greater complexity when the relatives are present means that TVDDI has not been shown to be a basic processing maneuver.

But if TVDDI did take over when TCA failed, serving as an abnormal processing maneuver for the difficult de-related sentences, then sentences 1 and 9 should show a dramatic decrease in complexity when, having had their relatives restored, they are amenable to TCA methods instead of the more demanding TVDDI ones. But this seems not to be true either (p. 292).<sup>17</sup> Thus, while TVDDI could be at most an abnormal processing tactic, it does not seem, on the admittedly slender evidence we have, even to be that.

But if TVDDI is not used, even in desperation, how are we to explain the errors in which a "versatile" verb was fitted with a complement it could only have acquired from the general MG lexicon? I do not really have an alternative to TVDDI to offer, but from Fodor and Garrett's evidence I think we can glean some indication of where to look for one.

As they remark, three of their experimental sentences were such that each of the sentence's three nouns could be matched sensibly with but one of the sentence's three verbs, if only because prior unique matching of the other two verbs had eliminated all save one possibility for the third. Thus one could completely lose track of the surface and still recapture the only rational Deep Structure, using what we might call 'puzzle-solving' methods. Such methods should work about as well, I should

<sup>16</sup> The "versatile-verb" sentences are minutely more complex than sentence 2 when all three are denuded of their relatives; this may or may not have some significance elsewhere, but in any case it cannot affect the outcome of the present argument, as see the remainder of this paragraph. (Sentence 1 is by a negligible margin the most complex sentence of all, when without relatives; but since the only other "versatile-verb" sentence, 9, is not next-most-complex, we can scarcely attribute sentence 1's position to its containing a "versatile" verb. In fact, sentence 1's position appears to be of no independent significance at all.)

<sup>17</sup> It may be pertinent to recall Slobin's finding (1966) that Passive sentences whose surface Actor-Object pair could not be mistaken for deep-Subject and deep-Object were easier to process than the "reversible" Passives susceptible to this error. ("The cat is being chased by the dog" is reversible because "The cat is chasing the dog" is sensible, hence a likely error; "The flowers are being watered by the girl" is irreversible because "The flowers are watering the girl" is not sensible [hence an unlikely error].) Possibly the irreversible Passives need not be detransformed, in any sense, to the Active-like Deep Structure in order to reveal their deep-Subject and deep-Object: the Deep Structure can be seized 'logically', as it were.

Turner and Rommetveit found (1967, 1967a) that "reversible" pictures (containing an Actor-Victim pair that would have been wellformed in another picture as Victim-Actor, e.g. two humans) did not affect performance in quite this way, so that Slobin's result seems to apply specifically to linguistic irreversibility.

think, if the respondents were to be presented with a mere list of the sentence's chief words--preferably in order--and told to make a sensible sentence out of them. In any case these three sentences were, without their relatives, the least 'complex' of the lot. (They were the least 'complex' when their relatives had been restored, too--indicating that TCA may receive help from unexpected quarters.) But such a 'puzzle-solving' method might have been tried also with the "versatile"-verb sentences, though (predictably?) with less success: once control over the surface were lost, respondents might well, on a partial recollection of the surface's chief words only, construct a Deep Structure, hence a new surface, where most of the old words turned up but where the relations among them had suffered a change.<sup>18</sup> Certainly a respondent would have to have recourse to his lexicon to do this--as in the three sentences containing unique assignments, for that matter--though apparently in much less exigent a way than would be predicted by TVDDI. But with this speculative remark we must close this discussion.

We entered upon this lengthy discussion of TCA, and therefore contrastively of TVDDI, because TCA seems to be a very serious and far-ranging hypothesis which might explain many discrepancies between the comparative sentential complexities predicted by the "Derivational Theory of Complexity", based firmly on  $CH_{CG}$ , and the performative complexities actually observed in experiment. That is, TCA cannot be ignored when matching complexities; it may be of help in explaining discrepancies and so in saving much of the Derivational Theory. Certainly the relation between the Linguistic Grammar and speakers' performances has proved, not surprisingly<sup>19</sup>, to be other than the simple relation predicted by the Derivational Theory;

<sup>18</sup> Unhappily, both of their sentences containing "versatile" verbs were such that the "versatile" verb was followed by another verb<sub>2</sub> which was such that verb<sub>2</sub> could occur in some other sentence as the "versatile" verb's complement, so that the sequence (versatile verb) + (other verb) bid fair to be error-producing. Thus, in "The pen which the author whom the editor liked used was new", a plausible mistake is "like to use"; while in "The man whom the girl that my friend married knew died", a possible mistake is "knew to have died". (Or, ignoring or lacking "whom", "knew had died".) The same broad method used in desperation to construct a correct Deep Structure for the three 'unique-assignment' sentences, might if tried on these two 'versatile verb' sentences have permitted construction of a Deep Structure on the same rough principles: but a Deep Structure, in this case, that was the wrong one.

Certainly any tendency to give e.g. "like" the wrong complement-- because all the lexical entries for "like" were being rifled--would have been if anything encouraged by the presence to the right of "like" of the possible complement "use". If the claimed tendency (essential to TVDDI) were actually to exist, it might be observed even when no possible complement was present to offer encouragement. This could be tested: subjects could be given a sentence like the one in question but where "used" was replaced by "owned"; since "\*liked to own" is impossible, any error in that direction would have to be in obedience to a (strong) tendency to give "like" the wrong complement. Unfortunately, as needs hardly be said, failure of the tendency to produce the named result would be of next to no significance, since the result would be ungrammatical, a consequence that speakers shy away from.

<sup>19</sup> As see, again, Miller and Chomsky, 1963, pp. 476-480; cf. n. 9, above.



and certainly TCA is a step in the right direction.

However, it is the wrong step.

1.1.2. We are coming presently to the question of why it is the wrong step, but before doing that we must ask whether DTC itself, which still has some life in it, cannot be improved by trying to refine somewhat its gross "count of transformations" as the way of assessing transformational complexity. Certainly such a refinement is proposed unequivocally by Miller and Chomsky's reference<sup>20</sup> to the "number and complexity" of the transformations used to derive a sentence. Since each transformation is a sequence of (one or more) "elementary operations", an obvious way of improving the calculation of transformational complexity would be to stop counting transformations and, instead, count elementary

<sup>20</sup> Cited in n.9 above.

operations. These are of <sup>^</sup>f our kinds: deletion, adjunction, substitution,  
21

and permutation. Each of these can be counted separately, and if justified, weighted; certainly permutation is the most complex of these operations and should presumably be weighted more heavily. Indeed, as to the permutations, further refinements can easily be envisaged. Thus it is certainly reasonable to suppose that an operation that permutes a large string-segment over a long distance is psycholinguistically harder to handle than one that permutes a brief string-segment over a short distance. Still further possible refinements come to mind: for instance, one might suppose that DTC would be improved if its calculations of transformational complexity took into account the relative complexities of the structural descriptions mandatorily matched by the transformations in order for them to apply.

In short, DTC could probably be bettered in a number of ways, and so be better fitted to accommodate observed disparities between CG-complexity and apparent psycholinguistic complexity. Certainly, at the very least, in asserting DTC to be incapable of coping with this fact or that we must try to show that no DTC, including all envisaged improvements (such as those sketched just above), could cope with the facts adduced.

1.1.3. We have been examining problems in the way of accepting **C&QG**, that version of the Correlation Hypothesis in which the Mental Grammar is a Competence Grammar and is therefore "=" to a Linguistic Grammar like the one now under construction by linguists. We have observed that one can affirm belief in CH,, without accepting the notion that one uses the CG to produce sentences by invoking its rules in sequence, somehow reversing this procedure to parse sentences (and somehow piecing out the CG, for

<sup>21</sup> Chomsky has proposed (1965, p. 144) that permutation can be eliminated from this set, being replaced by combinations of the other elementary operations. We will continue to use the term "permutation" here, but the term can be interpreted in accordance with Chomsky's proposal, and it may well be that the best way of gaging the complexity due to a given permutation would be to sum the complexity due to its component operations.

parsing, with an ability to impose surface-structure on raw in-coming strings). But we noted that refusal to accept this version of  $CH_{CG}$  does not entirely free predictions of sentential complexity based on the LG from according with the psycholinguistic sentential complexity indicated by performance: we gave the "5x8 index-card" model as a version of  $CH_{CG}$  which still made the two complexities identical. Then we quoted Fodor and Garrett to the effect that there were discrepancies between the two complexities. Though we disputed the evidence on which they decided this point, we suggested that nonetheless their position was probably well-taken, in that it seemed plausible that evidence for such discrepancies would be forthcoming. Anticipating such evidence (we have presented hardly any as yet) we examined in some detail a proposal, suggested by some remarks by Fodor and Garrett, which introduces into complexity-calculations a principle, here called TCA, which looks as if it might explain some of the expected discrepancies. (In Fodor and Garrett it explained phenomena which were not discrepant.) We also mentioned another such proposal, and we cited some of the obvious ways in which a straightforward 'DTC' sort of measurement might be refined. In summary, we will have forearmed ourselves with these ways of dealing with the expected discrepancies, should we meet them:

(1) Refining the way in which psycholinguistic complexity is calculated as a function of linguistic complexity, by improving the Derivational Theory of Complexity;

(2) Introducing, as needed, purely performative determinants of psycholinguistic complexity: e.g., TCA.

At this point in our exposition the  $CH_{CG}$  is still perfectly viable, in its weaker version; TCA was meant to 'save'  $CH_{CG}$ , and it may appear to have done so. But it has not; and we have devoted the preceding sections to ways of mitigating discrepancies in order chiefly to prepare the way for an exhibit of discrepancies which no envisaged 'refinements' appear able to account for. We have explored at some depth the means of defending  $CH_{CG}$ , but we have not explored at any depth the really telling evidence against  $CH_{CG}$ .

This now becomes our purpose. We have reserved this discussion until now, a bit out of logical order perhaps, so that the evidence against  $CH_{CG}$  can be used immediately as evidence in favor of what to my mind is the most obvious counter-hypothesis. We know of course that, as has been mentioned above, all evidence on any psycholinguistic point is excessively scanty; but now we will present what evidence we can, concentrating on a single case (the truncated Passives) that has the status, in the literature on this troubled topic, of a minor classic.

1.2. On the CH<sub>APG</sub>

In this section we consider a new version of the 'Correlation Hypothesis' which states that the (idealized) 'Mental Grammar' is not a 'Competence Grammar' at all, but rather something of mostly-unknown properties that I will term an 'Abstract Performative Grammar', or APG. I will maintain that the APG is far, however, from exhausting the (idealized) linguistic capacity, in that the APG is as it were sustained by a Competence abstract knowledge about the APG. The CH<sub>APG</sub>, stated roughly, hypothesizes two components, then: the APG itself, and a necessary 'archival' Competence faculty.

The hypothesis will be defended on two rather different grounds: (1) the APG hypothesis appears to explain predictable performative discrepancies which any performative explanation based on a conventional CG--however strained that basis--cannot explain or cannot explain as well; and (2) the APG hypothesis appears to accord better with the learning-sequence of certain syntactic devices, when we take into consideration the fact that any CG hypothesis must seemingly predict, as late stages in the acquisition of those devices, certain events which are wanting in plausibility.

1.2.1. We first take up some complexity-discrepancies which do not appear to be reconcilable with any account of performance that is tied, in any reasonable and consistent way, to a CG.

1.2.1.1. The first and best-known such case is touched on by Fodor and Garrett (1967). They remark (p. 290) that truncated passives like "The boy was hit" are "transformationally derived" from "The boy was hit by someone", and so are predicted by the DTC to be more complex, a judgment

that Fodor and Garrett correctly label as "counter-intuitive". Fodor and Garrett, however, though it at first seems that this example is introduced to illustrate the sort of problem they are attacking, do not return to it again in their paper, and indeed their suggested supplement to DTC - TVDDI (or even TCA) - does not seem to promise much in this area. Of course this fact does not of itself mean that every CG-based performative explanation must deliver a "counter-intuitive" judgment in such cases: we cannot conclude this until possible revisions of the primitive DTC have also been taken into account. And this we will do.

First of all, however, we must clear up Fodor and Garrett's (and many others') exposition of these truncated Passives, for it is seriously in error on precisely the issue at hand: the correspondence between full Passives and their truncates. This becomes plain when we ask what complex symbol is associated with the word "someone", and when we ask what complex symbol is substituted for the "Δ" dominating the 'Agent' which is to be deleted to form the truncate. If that complex symbol is such that "someone" is the associated word, then "...by someone" could underlie a truncated Passive like "John was hit"; but if that complex symbol is not one that "someone" is associated with in the lexicon, then e.g. "John was hit" must not be underlain by "...by someone". To consider this matter in its more basic aspect, the question is whether or not the deep Subject which is to become a "...by Agent" phrase and is then to be deleted can have the complex symbol of "someone". That is, can its complex symbol contain such factors as (+object), (+animate), (+animal), and (+human).



The Passive 'Agent' could be "someone", and still be deleted, only if "someone" were recoverable after the deletion.<sup>22</sup> This entails, first, that from e.g. "The boy was hit..." we be able to recover "...by Agent", that is "...by NP"; and we can, because the grammar identifies the truncation as a truncation: as having resulted, precisely, from deletion of "...by NP". But it also entails that we be able to recover, as factors in the complex symbol dominated by that 'agent' NP, all of the factors of the complex symbol of "someone"--for this is the only sense in which we can say that "someone" is recoverable. In particular, we must be able to recover the fact that the 'agent' NP, as the deep Subject of the underlying Base, headed a complex symbol containing (+human). This would be possible only if the verb "hit" had to have, as its deep Subject, an NP whose complex symbol contained (+human), for it is from just the fact that the verb is "hit" that (+human) must be recovered.<sup>23</sup> If "hit" could have had a deep Subject whose complex symbol did not include (+human)-- which included (-human)--then just from the fact that the verb is "hit" we could not recover the fact that the deep Subject was (+human). But of course "hit" can have a (-human) deep-Subject/surface-Agent, as in "The aging diva was hit by an egg."<sup>24</sup> So then "The aging diva was hit" cannot derive from an underlying structure whose "...by Agent" phrase included (+human) in the complex symbol of the source (the deep-Subject) of the 'Agent' NP.

But the so-called "full Passive" with "...by someone"--e.g., "The aging diva was hit by someone"--demands an underlying structure whose "...by Agent" phrase contains, precisely, a complex symbol that includes (+human). From this it follows immediately that "The diva was hit" and "The diva was hit by someone" do not have the same underlying structure.

<sup>22</sup> Only "recoverable" deletions are allowed because (Chomsky, 1965, esp. pp. 144-147) otherwise deletion, a transformation, could introduce meaning-changes (in this case, meaning-losses); but since it is the Deep Structure and not the transformations with which the semantic component is associated, these changes would go unnoticed by that component, so that, for a sentence allowing such a deletion, the semantic component would be ascribing a meaning to the sentence which the surface, containing the lacunae introduced by deletion, could not sustain.

<sup>23</sup> It could not be recovered from the mere fact of truncation or from, say, the nature of the deep Object (= surface Subject, in the Passive).

<sup>24</sup> This example argues, in addition, that "hit" can in its Deep Structure have an (-animate) Subject: e.g., in the Deep Structure representable as "An egg hit the aging diva." Counterarguments to this thesis can be mounted, but in general it seems to me that they fail. English appears to invest with a special character any (-animate) object that is, or appears to be, moving under its own power; and the most fundamental aspect of this character is the ability of such objects to occur as real underlying Subjects of certain verbs. This might be stated in the grammar by inclusion of a factor (animated), having the positive value (+animated) when an (-animate) noun is to serve as deep Subject.

They are not paraphrastic. "The boy was hit" cannot derive, by "truncation" from "The boy was hit by someone"; in fact the two sentences are not directly comparable, either linguistically or--more to the point--psycholinguistically.

In view of the psycholinguistic importance ascribed to the truncated Passives, the facts just cited impel us to undertake a brief examination of the issue of recoverable deletion: in particular, of what it is that is deleted. As Chomsky has said (1964, p. 42) each truncated Passive must derive from an underlying structure containing (in the "by..." phrase) an "unspecified Noun Phrase"; and in fact our task largely resolves into that of determining the nature of this Noun Phrase's degree of specificity. At one time (ibid., p. 41) Chomsky asserted that such an unspecified Noun Phrase was to be realized by its appropriate 'designated element', either by some maximally unspecific word like "it", "someone", or "something",--realizable on the surface--or else, in the absence of any appropriate surface word in the language, by a mere "abstract 'dummy element'". We have already observed that, for e.g. "hit" (and many other verbs), "someone" (and therefore "something") is overly specific; and "it" is too specific for the same reason that "something" is: it is not recoverable. So the only choice open to the 'by Agent'-phrase for "hit", there being no candidate for "appropriate" unspecific word, is the "abstract dummy", "Δ". Any such "Δ", since it has no way of being realized on the surface by a word, must be obligatorily deleted (1965, p. 222, n. 1).

In the one of the two models of grammar that Chomsky alternately entertains in Aspects <see esp. pp. 120-123) we find that every lexical category is rewritten, as the last act of the base component's categorial subcomponent, as "A"; but that most of these 'A' elements will be rewritten by the lexical subcomponent as a "lexical entry" of the form (D,C), where "...D is a phonological matrix and C a complex symbol" (1965, p. 122). The observations of (1964) and (1965) can now be fitted together: if the "A" of the Passive's "...by A" phrase has not been rewritten as a lexical entry—roughly, as a word—then "A" is the only terminal and it automatically triggers deletion of the "...by A" phrase. Of course failure to rewrite "A" has to have happened while the Passive was still a Deep Structure, as yet unpassivized, with "A" occupying its Deep Structure position of Subject. In sum, then, if a passivizable sentence in its Deep Structure has a vacuous unrewritten "A" Subject, and is Passivized, then the resulting "

25

by A" phrase is obligatorily deleted. (This synoptic view can be seen explicitly: 1965, pp. 128f., 137.)

Now, one could impose either of two conditions on the connection between deletion and recoverability. The one would hold that nothing can be deleted that cannot be recovered; the other would hold that everything that can be recovered, must have been deleted. We might call these the 'weak' and 'strong' conditions, respectively. Clearly the use of "A" just cited meets the 'weak' condition: we can always recover "A". But it utterly fails to meet the 'strong' condition, since from the typical truncated Passive we can recover linguistic information much more specific than is conveyed by mere "A". Even from "John was hit" we can recover "...by (-abstract) (-subject)". From "John was divorced" we can recover

25 It would seem a natural next step to specify that if a Passivizable sentence has an unrewritten " $\Delta$ " deep Subject then Passivization is obligatory and subsequent "...by  $\Delta$ " deletion is also obligatory. The alternative, letting Passivization be optional and then having any unPassivized " $\Delta$ "-Subject sentences be aborted, seems awkward. However, Chomsky does not take this further step.

"...by (-abstract) (+object) (+animate) (+animal) (+human)" and presumably (-male). Even singularity and plurality are sometimes recoverable. Thus there are in English verbs which by lexical definition describe transitive actions that can be performed upon one only once and by only one person, and these verbs therefore, when with singular deep Objects, take only singular deep Subjects. Five of them are: "sire", "father", "beget", "bear" (= give birth to), and "deflower". Thus from "Gladys was deflowered" we can recover nearly everything about the missing Agent but his name. Similarly, there are verbs which by lexical definition require plural Subjects (and do not accept mass Subjects, apparently); two of these are "bracket" and (when synonymous with that verb) "flank". From "Make sure each adverse comment is bracketed [by  $\Delta$ ]" we can recover the fact that the missing " $\Delta$ " is (+plural). Even more specific recoveries may be possible; thus, from "George was gored" we might plausibly claim to be able to recover, among other factors, (+animal), (+large), and (+horned) [I am indebted for this example to H. H. Clark], and also of course, barring cuckolds, (-human). In short, so much specific linguistic information is recoverable from truncated Passives that " $\Delta$ " will fail the strong condition of recoverability much of the time.

I submit that only the 'strong' condition on recoverability is intuitively satisfying, and that only the 'strong' condition is consistent with the many other ways in which the LG reflects the competence of the idealized speaker: for if we can recover this linguistic information we must be recovering it in the Deep Structure (where else could it reside?), and so it must be in the Deep Structure to be recovered.

As we see, neither "someone" nor " $\Delta$ " can (abstractly) be the 'Agent' that is generally deleted to form truncated Passives, since "someone" is too specific (not all of its factors are actually recoverable, as from "hit"), while " $\Delta$ " is not specific enough (as for "deflower", "flank", "divorce", etc.). "Someone" fails the weak recoverability condition; " $\Delta$ ", the strong. The needed corrective, quite clearly, is to incorporate the 'strong' condition in the grammar, but without violating the 'weak' condition. It now remains to ascertain how this may be done.

We continue, as above, with the Aspects model in which the lexical subcomponent is called upon to substitute, for each of the " $\Delta$ " symbols that the categorial subcomponent generates as its terminal string, a (D,C) "lexical entry" consisting of a word in abstract phonological specification (D) together with (C), the word's meaning (and fine syntactic function) in the form of a complex symbol of factors (e.g., (human)) headed by values (let us assume + or -). In this scheme the only source of the semantic factors of present interest to us is this substitution maneuver which automatically carries with it the abstract phonological specification of a word; yet, as we have seen, we will sometimes want to associate some factors (C) with some " $\Delta$ " symbols without at the same time substituting a word (D): there is no D.

We want to substitute for " $\Delta$ " exactly the factors that can be recovered after " $\Delta$ " has obligatorily been deleted. No fewer factors, and no more. Thus for the "...by  $\Delta$ " phrase of "hit" we will want (+object) but not (+animate) nor (-animate). For the phrase of "divorce" we will want (+human) but not (+plural) nor (-plural), and in addition we will want (+male) if the superficial Subject (deep-Object) is (-male), but (-male) if the Subject is (+male): generally, ( $\alpha$ male) for 'Agent' " $\Delta$ " when the surface-Subject

is ( $\alpha$ male), with  $\alpha$  ranging over + and -. For the phrase of "deflower", which demands a (+human) deep-Object, we will want (+human) and (+male); and furthermore, if the deep-Object (surface-Subject) is singular, then we will want (-plural). For "flank" we will want, whatever the surface-Subject, (+plural). For "gore" we will want, among other factors, (+animal) and (+hornèd).<sup>26</sup>

Now, most of these recoverable factors will be recovered, quite obviously, just on the strength of the verb: it is known what the restrictions on the verb's deep Subject are, and this is the way we (or the grammar) will recover the missing factors. Moreover even when the nature of the deep-Object (surface-Subject) contributes to determining what can be recovered, as when the 'Agent' of the Passive of "deflower" must be (-plural) if the surface-Subject is (-plural)--even this restriction can be incorporated into a restriction on the verb. In fact, it must be so incorporated, since it is a restriction on "deflower" that the deep-Subject be (-plural) if the deep-Object is.

The fact that recoverable 'Agents' are recovered from restrictions on the verb directs us toward an efficient way of providing in the grammar for preservation of the 'strong' recoverability condition. Since we are attempting to have the grammar generate deep-Subjects whose complex symbols are totally anticipated by the lexical restrictions on what can occur as deep-Subject of the verb, and since therefore the complex symbol of such a deep-Subject is effectively provided by the lexicon's contextual statement of the verb, it is entirely natural to derive such a Subject's complex symbol from the contextual statement of the verb.



<sup>26</sup> It might seem that to use "gore" correctly we would now require that the grammar include an encyclopedic knowledge of which animals are and are not horned. But this is entirely false. We insist only that the grammar tell us, if someone says "I was gored by a canary", that that person thinks canaries weigh a ton and have horns.

The verb remains selectionally restricted according to what deep-Subjects and deep-Objects it can take (Chomsky, 1965, esp. pp. 95-99, 165), and so the usual procedure will be for the NP to gain their complex symbols before the verbs do. Optionally, however, the " $\Delta$ " dominated by NP need not gain a complex symbol from the lexicon; it will remain devoid of factors after the generation has passed through the NP-substitution rules. (So far, there is nothing novel.) But, when the verb gains its complex symbol in conformity with the contextual constraints specified in the lexicon, I propose that if the deep-Subject is "blank" (has no complex symbol), then all of the factors specified in the verb's contextual restriction for deep-Subject be copied out under [NP,S] to form a complex symbol dominated by the verb's deep-Subject. (We ignore for the remainder of this paper the possibility of similarly treating the deep-Object.) That is, if the deep-Subject is "blank", it automatically acquires all and only the factors specified for deep-Subject in the lexicon's selectional restriction on the verb. Thus any verb being introduced into a sentence with a "blank" deep-Subject (a) must have a contextual statement that accords with the non-blank categories--e.g. here, the deep-Object--and (b) must impose its own requirements on any "blank" categories.

As we noted, only the deep-Subjects that the process of lexical substitution left "blank" can derive a (thin) complex symbol from "copying" in this way. The usual deep-Subject (e.g., "the professor", "the wart-hog", "the King of Zembla") is far more specified (has many more factors) than is required by the verb, hence is far more specified than could be determined by (or recovered from) the verb. For all such " $\Delta$ " the more orthodox method of substituting a complex symbol will prevail: for our present purpose,

that method can consist entirely of substituting for the " $\Delta$ " of [NP,S] any (D,C) that the lexicon makes available for such a substitution--that is, almost any lexical item listed as a Noun.<sup>27</sup>

After introduction of the verb's complex symbol, with copying into the deep-Subject if called for, the sentence will have for its deep-Subject either a complex symbol drawn conventionally from the lexicon, or else a complex symbol copied from the lexical entry for the verb. If a verb makes a dual requirement of its deep-Subject and deep-Object ("divorce" insists that they have different values for (male); "deflower" insists that the deep-Subject have (-plural) if the deep-Object does), then this requirement is reflected exactly in the factors that are drawn from the verb's contextual statement for insertion into the deep-Subject's complex symbol. Suppose that the deep-Object has a complex symbol that includes (-plural) and (-male): then "deflower" can be introduced as the verb, but only, according to its lexical context-statement, if also the deep-Subject is (-plural) and (+male). If the deep-Subject already has a complex symbol, it must not include (+plural) or (-male). If the deep-Subject is "blank", then (-plural) and (+male) are copied from the verb's contextual statement into the deep-Subject's complex symbol. Thus, when all of these operations are over, "divorce" will have either a conventional deep-Subject--e.g. "Agatha" or "the heartless scoundrel"--or else it will have just a "thin" complex symbol containing, among a few other factors, e.g. (-male) and (-plural).

<sup>27</sup> Among these lexical substitutions will be the (D,C) for "someone". Since e.g. "hit" does not specify that its deep-Subject be (+human), no (+human) deep-Subject (such as "someone") can be conferred on "hit" through "copying;" thus, unless for "hit" the deep-Subject can acquire "someone" from the lexicon, it can never acquire "someone" at all. We will shortly return to this point from another angle.

Now, we could insist on treating these "thin" C's just as, before, we treated the "A" that had no C at all: that is, we could insist that all such C be deleted, after Passivization. But some of the deep-Subjects resulting from "copying" will have complex symbols C very similar to some C in the lexicon to which a lexical item D is joined. The "copied" deep-Subject of "deflower", if the deep-Object is singular^ will have very nearly the complex symbol C of the D "someone" provided by the lexicon. The two C's will differ only in that the "copied" C of the singular deep-Subject of "deflower" will be specified (-hnale), while the C given in the lexicon for "someone" will (obviously) not have this specification. Let us see how these two C's might be matched. "Someone" is no more specific than (+human) (-plural), being indiscriminate respecting (male)ness; it should consequently appear in the lexicon with a nonvalued factor: (male) rather than either (-hnale) or (-male). Since the (-hnale)ness of the copied deep-Subject of "deflower" is by definition recoverable, obviously it does not matter what D we adjoin to the "copied" C, so long as that D does not carry with it new, hence irrecoverable, factors. Or new values on factors. That is, to any "copied" C we can adjoin any lexical D whose lexical C: (a) contains no factor not included in the copied C; (b) contains every factor included in the copied C; and (c) contains no value on any factor— if it designates a value for that factor at all—that is different from the value on that factor in the "copied" C. Returning to the D "someone", it can now be substituted for the singular C copied into the deep-Subject from the verb "deflower".

Some such operation must be possible, for if it were not, the grammar would claim that in the sentence "Moll has been deflowered

by someone" we could not absolutely recover the fact that her deflowerer was (+male), and this would be, in fact, manifestly false. Some such solution is imperative, so there is strong motivation for the modification of the accepted version of the grammar (e.g., Chomsky, 1965, p. 165, Rule 21-ii) needed to accomplish such a solution.

Three final points.

(1) "Someone" can be withdrawn from the lexicon to fit a copied C; but also of course it can be withdrawn as a (D,C) pair, in a conventional lexical substitution, to instantiate a deep-Subject " $\Delta$ ". The severe selectional constraints on the verb which permit copying of "someone" must be imposed on the deep-Subject in any case: if the deep-Subject has a C gained in the lexicalization pass, then the verb's contextual restrictions must countenance that C (or else that verb cannot be introduced); if the deep-Subject is "blank" after the lexicalization pass, then the verb's contextual restrictions "copy" a C into the deep-Subject. The contextual restrictions in question are really a dual constraint on the deep-Subject and its verb, imposed by the deep-Subject if that is first to be introduced, imposed by the verb if it is first. This means that whenever "someone" can be gained from "copying" it could also have been gained from direct lexical substitution (a verb that forbade its deep-Subject to be "someone" could scarcely have specified such a deep-Subject through "copying"). Thus, in permitting the D "someone" to be adjoined to a copied C, we have introduced an undoubted redundancy into grammar.

The redundancy is not difficult to remove, and the manner of its removal has ancillary advantages; but to go deeper into this topic would

require more space than we have here and so for the present, having pointed the redundancy out, we will ignore it.<sup>28</sup>

(2) To provide in the most natural way for the generation of both "Gladys was deflowered by someone" and its paraphrase "Gladys was deflowered", we have allowed the deep-Subject's finding a D in the lexicon to be an option; if e.g. "someone" is substituted then the full Passive will result, but if no D is substituted then the unlexicalized " $\Delta$ " and the thin complex symbol it heads will, most simply, be deleted, and the paraphrastic truncate will result. Notice then that it is not quite right to say, even for a truncate like this, that the word "someone" can be recovered. More accurately, a C can be recovered with which some C in the lexicon does not "disagree", and to the latter the D "someone" or the D "somebody" is joined.

(3) For any C drawn conventionally from the lexicon, a D is provided by the lexicon; indeed, conventionally, the pair (D,C) is withdrawn together. But for some of the C that result from "copying", no D exists; in fact the copied C to which a D can be joined is the exception rather than the rule. Any C containing (+human) but omitting (-plural) cannot acquire the D "someone", which requires (-plural); any such C will not find a D at all. (It is just such a C that "divorce" has for its "copied" deep-Subject.) For such cases, where the C has no way of breaking surface, we have two choices. The first and simplest one is to insist, as in (2) just above, that such a D-less C be obligatorily deleted (with the result, e.g., "Gladys was divorced"). The second choice is to introduce a way of deriving from such a C, where possible, a phrase like "...by one or more persons" or the like. The C would have to undergo one or more segmentalization transformations (Postal, 1966); but this would be possible, and might well be

eliminating

<sup>28</sup> The redundancy in question cannot be removed, as we have noted before, just by/ "someone" from the lexicon, since there are many cases in which "someone" occurs in a sentence where it cannot have been generated through "copying"-- e.g., in "The actor Farine d'Avoine was divorced by someone last year". Nor can the redundancy be removed just by eliminating "copying" as a source for "someone", since to do so would make deletion of a "copied" C obligatory even when that C matched the lexicon's C for "someone", where this obligatory rule would, altogether artificially, be introduced just to remove the redundancy (and where the relation between the paraphrases "Moll was deflowered" and "Moll was deflowered by someone" would be made unnaturally remote).

This and other questions are taken up, with a consideration of some residual problems, in Watt, forthcoming.



desirable. We cannot take this subject further here. But we will assume that "one or more persons", if not actually derived from the C of an appropriate "copied" deep-Subject, would in the grammar be identified as synonymous with it.

We have taken this discussion far enough now, skipping only minor details, to return to our main point. But first let us summarize the series of steps involved in the generation of sentences exhibiting recoverable deletion: (1) the categorial subcomponent terminates, leaving each lexical category (Noun, Verb, and so on) represented by a " $\Delta$ " symbol; (2) for every " $\Delta$ " Noun symbol the lexicon will, optionally in the case of deep-Subjects, substitute a (D,C) pair of complex symbol cum phonological matrix; (c) for every " $\Delta$ " Verb symbol the lexicon substitutes a (D,C) pair conformable to the verb-context and also (d) copies, if the deep-Subject has no C, into the vacant C of the deep-Subject all of the factors that the lexicon stipulates for the verb's deep-Subject in that sentence; (e) for such a "copied" C there may now, optionally, be substituted any D whose C in the lexicon "agrees" with that C; (f) any D thus substituted will be realized on the surface; (g) if no D is substituted, unless segmentalization is introduced as another option, then deletion is obligatory (and this is the only source of the truncates' recoverable deletions).

We now return to the main issue of determining what full Passives are directly related to what truncates. We see that, formally, a full Passive corresponds to a truncate if the Agent of the full Passive contains in its complex symbol all and only the factors, with no discordant values, specified in the lexicon for inclusion in the deep Subject of the Agent's verb.

Informally, the Agent is "someone" only for the rare "singular deep-Subject if singular deep-Object" verbs, like the five cited above. If there were similar verbs for (-human) or (-animate) Subject-Object pairs, then the truncate's corresponding full Passive would have a "...by something" Agent phrase. For any verb whose deep Subject must be (+human) but which is indeterminate as to plurality<sup>29</sup> the surface overt Agent is, minimally, "person or persons" or the like. For many verbs (e.g., "hit") the minimal surface Agent is, alas, "one or more persons or things" or the like. There are, of course, the sentences like "Dwayne was hit (struck, kicked, flunked, exonerated,...) by someone", but such sentences are not directly related to any truncate, since from "Dwayne was hit (struck, kicked, flunked, exonerated,...)" one cannot recover (+human) because for "hit" and "strike" a (-human) rock could have served; for "kick" a (-human) ostrich could have served; and for "flunk" and "exonerate" either a (+plural) group ("the professors"/"jurymen") or a (-count) group ("the Department"/"Army") could have served. For a common verb picked at random, the chance that that verb occurs in a directly-related truncate/full-Passive pair of sentences is rather remote.

However, such verbs are not that difficult to find, by any means (as see n. 29), and certainly there are quite enough for meaningful psycholinguistic testing. But that testing has been somewhat complicated by the findings summarized just above. Heretofore the problem of testing the comparative psycholinguistic complexity of full Passives and truncates has been construed as that of comparing each truncate with a full Passive consisting of the truncate augmented by "...by someone". In this way

---

<sup>29</sup> E.g. "divorce", "marry", "assassinate", "dub", "knight", "cauterize", "inoculate", "trepan", "catechize", "shrive", "harangue", "indite", "abdicate", and countless other equally common verbs.

the great imponderable of all psycholinguistic investigation--the comparative influence on sentential complexity of meaning-differences--has been entirely avoided, since the assumption was that the truncates and their "...by someone" counterparts were paraphrastic. But now, as we see, there are very few verbs (e.g., "deflower") for which this assumption was warranted. On the other hand, many verbs (like those in n. 29) take only (human) Agents, and for these the truncate at least has the counterpart in "...by one or more persons" or the like. The difficulty is that while it is not hard to stretch out a truncate so that it occupies the same amount of recording tape as that truncate plus "...by someone", on the other hand "...by one or more persons" is more difficult to compensate for, if the stretched truncate is not to sound peculiar. (Of course the proportion by which the latter Agent is oversized could be reduced by using much longer truncates than has been customary.) In addition, the slight unnaturalness of the phrase, for many, might slow responses and so skew results. But it is clear how the revised experiments might proceed.<sup>30</sup>

Having come back to our main topic, we now ask if it seems likely that the truncates will be found through experiment to be in fact, as predicted by DTC, more complex than the full Passives from which in essence they derive. This does not seem likely. What does seem likely, on the contrary, is that experiment will at last confirm what has so often been declared, namely, that truncates are more complex in the LG but simpler in the MG, even when the length-difference has been compensated for.<sup>31</sup>

Unless, of course, mitigations for these predicted data can be found, such as TCA (or even TVDDI). Since we must consider some other cases much like the truncated Passives--cases which should also be mitigated if the truncated Passives are--we defer for the moment a consideration of whether

That is, it is clear how the experiments might proceed rigorously in the context of present knowledge. However it is quite possible, I think, that the complexity experiments could be considerably simplified, without being vitiated by the factors sketched above in this subsection, if one or two experiments were first conducted on aspects of language-use related to truncation.

We have shown that e.g. "John was divorced" is directly related, not to (1) "John was divorced by someone", nor to (2) "John was divorced by some people", but only to (3) "John was divorced by one or more persons" or the equivalent. Does this mean that "John was divorced" is ambiguous with respect to plurality of the Agent? If it were, and were a paraphrase of (3), then (3) would have to be ambiguous too (Hiz, 1964). But this would be ridiculous: we say a sentence is ambiguous when it could have more than one meaning, not when (a) it overtly specifies a choice of meanings and specifies that no choice between those meanings can be made, as in "...by one or more persons", nor when (b), as in the truncate, it tacitly (contextually) specifies the choice of meanings and remains silent on which choice is correct. It is clearly preferable to describe both (a) and (b) as "indeterminate" rather than as "ambiguous".

About "ambiguity" we have a little psycholinguistic data; about "indeterminateness" we have none. The 'prior experiment' mentioned above, then, involves testing whether or not indeterminateness affects linguistic performance in the same way that ambiguity does. The facts as to ambiguity appear to be that ambiguity slows processing when two (or more) interpretations are about equally likely (MacKay, 1966), but does not do so when only one is likely (Foss, Bever, and Silver, 1968). If indeterminateness proved to be like ambiguity in this respect then the testing of truncates whose indeterminateness is narrow and known could be greatly simplified. At the very least many truncates could be limited to but one 'likely' interpretation by being read in a prejudicial context; for example, "John was divorced" could be given in a context in which the auditor would be heavily disposed to think that only one Agent did the divorcing. Since "John was divorced" is indeterminate only with respect to plurality of the (necessarily (+human)) Agent, this truncate could then be tested for complexity against the simpler "...by someone" full Passive, the latter having been made the truncate's local paraphrase (in essentially the

sense of Harris, 1965, p. 388, n. 35; and see p. 390, n. 41). Furthermore it is probable that, psycholinguistically, sentences containing verbs like "divorce" could be made still easier to experiment with. For example it seems likely that the sentence "John has just been divorced" would invariably be interpreted as having a singular Agent, even though it is (remotely) possible that John was just divorced by three women simultaneously. Such a sentence, if 'performatively univocal' as we might term it, could then be tested without bothering to prejudice its prospective auditors.

However, it should be kept in mind that such simplifications of the psycholinguist's task have no linguistic warrant whatever.

31 This essentially assumes that prior experiment will establish what contribution length alone makes to performance-time, testing numbers of different string-lengths of similar or identical immediate constituent-structure: e.g., "William Henry Harrison" vs. "John Tyler"; e.g. "My adversaries are very persistent" vs. "My adversaries are very very persistent" (Watt, 1968); and so on.

We also assume that care will have been taken to guard against complications due to the constraint imposed by context or by general expectancy jointly on the Deep Structure and the surface which makes e.g. the truncate "John has just been divorced" derivable from a Deep Structure additionally specified so as to coincide with the Deep Structure that in the grammar underlies "John has just been divorced by someone." I foresee no real difficulty here--the effect is that of narrowing the versatility of "divorce", insisting that its Deep Subject be (-plural)--but then unforeseen difficulties have been the bane of psycholinguistic experimentation in this area.

or not the anticipated discrepancy between LG and MG can, still, be attributed to some 'mitigation' or other.

1.2.1.2. There are other apparent discrepancies between LG and MG complexities; the number of cases is easy to multiply, and none is very surprising.

(a) One observed discrepancy is due to Fodor and Garrett. They found (1967, pp. 293-4) essentially that the comparative complexity predicted by DTC, on the LG, for sentences like "The tired soldier fired the shot", was not confirmed by an experiment they ran; that is, sentences containing prenominal adjectives (e.g., "tired") did not exhibit the comparative complexity they should have to be derived, as in the LG, from a structure representable as "The soldier, which soldier was tired, fired the shot".<sup>32</sup>

(b) "John hammered the nail" and "Charlotte Corday knifed Marat" derive from sources representable as "John used a hammer to act upon the nail" and "Charlotte Corday used a knife to act upon Marat" (for essentially this view, see Lakoff, 1968); but it is implausible that the first pair of sentences should be psycholinguistically more complex than the second pair.

(c) "Mary grows flowers" and "The heroes cracked the glass" derive from sources on the order of "Mary causes flowers to grow" and "The heroes caused the glass to crack" (Chomsky, 1965, p. 189), but the first pair of sentences is certainly, counter to prediction, simpler.

(d) "Dee is hard to please" derives from a source on the order of "For anyone to please Dee is hard", with the first sentence, counter to sense, predicted to be the more complex.

<sup>32</sup> The experiment compared the complexity introduced by inserting two adjectives (3 transformations @) with that due, as before, to two deletions (with Fodor and Garrett, 1 transformation @). Contrary to DTC, the adjectived sentences proved less complex than the deleted ones, according to Fodor and Garrett. One must own, however, that their evidence does not justify great confidence in their claim. If their Table 5, p. 293, pertains just to the adjectived sentences, as it seems to on internal evidence (Fodor and Garrett say otherwise, *ibid.*), then the adjectived sentences were proven less complex than the deleted-relatives sentences of their Table 1 (p. 292), but not less complex than the same sentences of Table 2 (*ibid.*). The latter sentences were identical to the former ones except that they were read "expressively", whatever that may have come to.



(e) "Dee is eager to please" derives from a source very like "Dee is eager to please", and should therefore be demonstrably less complex than the more distantly-derived "Dee is hard to please"; again, a prediction counter to expectation.

(f) "There's a dragon in the street" derives from an ultimate source on the order of "A dragon is in the street" via an intermediate string of the form "In the street there is a dragon" (Watt, 1967); the first sentence is further from the ultimate source than is the third sentence, and so it is predicted, again counter to sense, to be more complex.

(g) "I read the book while in England" probably derives from a source something like "I read the book while I was in England" (Chomsky, 1965, p. 219), with the shorter sentence therefore predicted to be the more complex.

(h) "Brutus killed Caesar" has a source that we can represent as "Brutus caused Caesar to be (become) dead" (Lakoff, 1965, IX, pp. 9-12), but if we realize that source on the surface, in the 'representative' form just given, sense tells us that "Brutus killed Caesar" is simpler.

(i) "Guy Grand ate the pickled eel with revulsion" perhaps has an intermediate source on the order of "Guy Grand had revulsion while he ate the pickled eel" (Chomsky, *ibid.*), and in any case it has a deeper source like "Guy Grand was revolted {while eating the pickled eel" / by eating the pickled eel"}; and this form, since it is clearly a Passive, must derive in turn from something on the order of {"The pickled eel / Eating the pickled eel"} revolted Guy Grand {"while he ate it" / "}. and the second (more likely) of these structures has a still deeper source on the order of "Guy Grand's eating the pickled eel revolted him" or, more deeply, "(Guy Grand ate the pickled eel) [and it] revolted him".<sup>33</sup> This

33 As always, these illustrative sentences are used only to suggest their respective structures; the deeper the structure the more strained is the attempt to provide an illustrative (surface) sentence. In the last illustration I have inserted "[and it]" to make a viable surface.

means that under DTC we must expect these sentences to decrease in complexity roughly in the order of their citation here; whereas sense suggests that experiment will show exactly the opposite to be true.

And so on.<sup>34</sup>

Taken as a group, the sentences that are more complex in the LG but discrepantly of less performative difficulty have one thing in common: they resemble the truncates in that they are shorter than the underlying structure from which, more directly, derive their longer and performatively more complex counterparts.<sup>35</sup> This serves to remind us that the performative discrepancies here predicted are not in the least surprising, since sense would tell us that, barring distinct complicating factors (e.g., those explained by TCA), shorter sentences should be simpler to deal with than their longer paraphrases. In fact, the abundance of cases where the 'more distantly derived' sentence is shorter reminds us that DTC was clearly wrong from the beginning, intuitively, insofar as cases like these are concerned. DTC was designed to cover such cases as Active/Passive, where the sentence predicted to be more complex, because more distantly derived, was also, providentially as it turns out, the longer. Indeed, when DTC was first being promulgated it was more generally the case that transformation typically lengthened structure. More recent analyses have if anything reversed that state of affairs (and see p. 10, above). In short, the cases cited, saving (f) and part of (i), may be systematic exceptions to DTC; and if so they are the sort of systematic exception that unseats a theory altogether. For if only 'more transformations' were involved in increasing performative complexity, the crucial test of such a claim would be whether or not clear cases could be cited where the shorter sentence was harder to process than its longer, but LG-simpler, paraphrase.

<sup>34</sup> There is also the result, reported by Fodor and Garrett (1966, p. 150), that "John phoned the girl up", while "derived from" "John phoned up the girl", is no more complex. This result is impossible to evaluate, however, since "phone" is (at least in my speech) much more common than "phone up", so that the "up" at the end might have come as a surprise, after the sentence had seemed to be finished, or else just struck auditors as strange. It would be crucial to the demonstration of this discrepancy to test some verbs which occur more naturally with e.g. "up" and verbs which occur only with e.g. "up" (or "down", "out", and so on). (For example, "single out", "eke out", "jot down", "divvy up".) Lacking such evidence, we must decline to accept the discrepancy as a proven generalization, though it is certainly plausible enough.

<sup>35</sup> Of course case (f) and part of (i) differ from the others in this regard.

1.2.1.3 We now ask whether or not this apparent set of exceptions can be explained in such a way as to leave unjeopardized.

(1) We first take up the case of the full Passives and the truncates, according with the account of these forms sketched above in (1.2.1.1.). We assume that the truncates will be shown empirically to be psycholinguistically less complex, as a rule, than the corresponding full Passives, even when the differences in string-length have been taken into account. That is, we assume that the truncates will as a rule take less time to parse or compose. Now we ask if this discrepancy appears to be explainable by the contemplated revisions in the DTC. The answer is in the negative. If the truncate derives by deleting a "copied" deep-Subject, which the full Passive does not delete, then no revised DTC can account for the discrepancy. If the full Passives in question derive their overt Agents via segmentalization, as with "by one or more persons" or the like, then for these Passives (but only these) DTC offers a possible explanation, since segmentalization might be more complex an operation than deletion. Thus DTC might explain some discrepantly complex full Passives, but it cannot explain all. Moreover, no other plausible revision of the treatment of the truncates looks to be more amenable to explanation in terms of a revised DTC; for example, we might propose to derive truncates more distinctively than is done in the account given above, having them result, not from "copying" plus deletion, but from a failure to "copy". But this would afford little improvement in any DTC measurements. The truncate represents a deletion, not just of the deep-Subject's "copied" C, but also of the deep-Subject's NP node (formally, the head of the C); and it was never proposed to derive the NP by "copying", so that a mere "failure to copy" would not deprive

the deep-Subject of its NP. But suppose that the deep-Subject were so deprived: that is to say, suppose there were no deep-Subject at all, in Deep Structures destined for truncation. Such an omission would require introduction of a special Passive rule (one requiring no deep-Subject) to generate the truncations; not a desirable move. Furthermore, even if it were proposed that truncates be derived just by failing to "copy" a C into the deep-Subject, the DTC measurements would still not be much improved: though it is a moderately complex operation considered in isolation, "copying" does not add much complexity, since it is only part of the already-complex operation (a transformation) of introducing the verb (as see above). In short, no way comes to mind of revising DTC so as to force DTC to predict with general accuracy the comparative performative complexities of full Passives and truncates.

TCA will not help either, since every full Passive that fully corresponds to a truncate (of n words), includes that truncate as its first n words: any TCA-predicted delays in assigning surface and deep structure must affect such sentences equally. Passing to questions of relative probability, we see that it is certainly true that truncates are more common than full Passives by a factor of four to one (Svartvik, 1966, p. 141); but this factor is beyond question less than that by which e.g. "The tiger...was ferocious" is more probable than "The tiger...I was glad to see the last of" (see 1.1.1., above). Thus, though the full Passives are clearly the less probable, with completion "surprising" to that mild extent, the margin of likelihood is scarcely so great as to promise much in mitigation of an overturned DTC prediction.

In sum, none of the mitigations we have looked at promise to palliate the predicted discrepancy between the truncates' greater LG complexity and lesser performative complexity.<sup>36</sup>

(2) We continue with a brief inspection of the other discrepancies mentioned above:

The "Mary grows flowers/  $\left\{ \begin{array}{l} \text{causes} \\ \text{makes} \end{array} \right\}$  flowers  $\left\{ \begin{array}{l} \text{to} \end{array} \right\}$  grow" case (1.2.1.2.c) is another discrepancy not easily explained. No revision of DTC promises much, because the longer sentence is very close to the Deep Structure, while the shorter sentence is rather clearly a distorted (hence more distant) version of that Structure. TCA does not offer much either; for example, from the first two words ("Mary causes..." or "Mary grows..."), more correct Deep Structure is obtained for the "causes" sentence than for the "grows" sentence, because the "grows" sentence could, up to "Mary grows", be an intransitive sentence like "Mary grows an inch with every doughnut." Neither "shorter" nor "longer" sentence appears to continue, at any point, in so improbable a way as to cause rescission of a prior Deep Structure assignment made on a probabilistic basis. In sum, the discrepancy stands.

The "Dee is hard to please" case appears equally unyielding. Taking all the sentences of (1.2.1.2.d), together with the more complex paraphrase "It is hard [for anyone] to please Dee", it is obvious that TCA essentially predicts the sentences to be about equal by its standards. DTC, however modified, must gage "Dee is hard to please" and "It is hard to please Dee" as of highest derivational complexity (yet we anticipate these to be performatively no more complex than the others). The transitional probabilities are unpredictable in this case, but it does not seem at all likely that the factor that complicates the performance of "To please Dee is hard [for anyone]" is the unlikelihood of "Dee", "hard", "anyone", or any other

<sup>36</sup> It is almost certainly true that, while the truncates are only four times more common than full Passives in general, their likelihood is greater still when they are compared with just their paraphrastic full Passives (e.g., "Moll Flanders was deflowered by someone last year"). But it is hard to believe this to be relevant. The unlikelihood of the paraphrastic full Passives is surely not so extreme as to cause them to be resisted because bizarre; and it cannot be maintained very plausibly that, while completing a truncate with a lexical NP is not very "surprising", nonetheless completion with e.g. "...by someone" is so "surprising" as to mitigate DTC.

Verbs seem to vary in the extent to which their naturalness in a Passive is increased if an overt "by Agent" phrase is given. ("There has never been a time when France was not governed" sounds more awkward than "Nixon was elected to the Presidency in 1968"). If it is more "surprising" for a verb that needs no overt "by Agent" to have one than for a verb that does need such a phrase, it should not be hard to lend empirical verification to this fact.



element in this sentence. This case, too, cannot be dismissed.

The "Dee is eager to please" sentences of (1.2.1.2.e) were predicted by the simple DTC to be less complex than the "Dee is hard to please" sentences just examined. TCA finds no distinctions here; DTC cannot be revised in any obvious way to remove the discrepancy; again the probabilities seem irrelevant; and so again the case stands.

The "There's a dragon in the street" sentences of (1.2.1.2.f) seem just as impervious. TCA appears irrelevant; DTC appears beyond the needed modification; and probability seems to be as irrelevant as TCA. Again, the discrepancy stands.

The several other discrepancies noted in (1.2.1.2.)--just over half of the total--are subject to more uncertainty of judgment than those just considered, and the few statements that one can now make about their mitigation might better be postponed. Case (1.2.1.2.a) is hard to evaluate because it seems that "The tired soldier fired the shot" might be derivable from a structure like "The (soldier was tired) soldier fired the shot" rather than from a structure like "The soldier who was tired fired the shot"; this change would alter the DTC prediction rather violently, since the permutation that (presumably) contributed most heavily to the predicted complexity would now have been removed. The cases of (1.2.1.2.b), "John hammered the nail/used a hammer to act upon the nail", seem particularly hard to compare because of their inequality as to string-length; for the time being we pass these by, then. Case (1.2.1.2.g) involves a simple deletion ("I read the book while [I was] in England", but deletions of this sort seem so minor that the undoubted DTC prediction seems sub-critical. The cases of (1.2.1.2.h) are more interesting. Certainly DTC

cannot be revised in any obvious way so as to predict "Brutus caused Caesar to be(come) dead" to be performatively more complex than "Brutus killed Caesar"; but TCA appears to predict the shorter sentence to be simpler, a possible mitigation. As to probability, it does seem completely unlikely that any sentence-fragment "Brutus caused Caesar to..." should actually be completed with "...be dead", or that "Brutus made Caesar..." should be completed with "...dead". So reception of these sentences might be slowed because of this 'surprise' element. But before pronouncing this a case of mitigation on 'surprise' grounds we ought to test the comparative complexity of other sentence-pairs that are like "kill"/"dead" except that their 'surprisingness' is reversed: e.g., "Brutus angered Caesar"/"Brutus made Caesar angry". Lastly, the complex case (1.2.1.2.i) seems to me to require further linguistic analysis before we can base any DTC prediction on a firm LG account of these sentences. But for the time being it does not at all seem likely that "Guy Grand's eating the pickled eel revolted him" or "To eat the pickled eel revolted Guy Grand" will be predicted by DTC, on any LG account, to be more complex than "Guy Grand ate the pickled eel with revulsion"; and if not, the predicted discrepancy will stand.

(3) Considering just the processing aspect of performance, we have several fairly clear cases where LG complexity and performative complexity seem unalterably at odds. If we now turn briefly to the compositional aspect of performance, the number of discrepancies if anything increases. TCA can no longer mitigate the discrepancies, since the composer of a sentence can hardly run into the problem of assigning Deep Structure to his output: the output has its origin in that Deep Structure. Nor should probabilistic factors play a role, since the speaker of a sentence is

unlikely to run into performative difficulties caused by his surprise at how the sentence turns out. Yet the same DTC-predicted discrepancies, with the same difficulties of sufficiently altering the DTC, obtain for the composition of sentences as obtain for their reception. Thus, for example, DTC (revised or unrevised) predicts that truncates should take more time to compose than full Passives, stringwise length aside. If, as anticipated, this prediction fails, then the receptive discrepancy will be matched by a compositional or productive discrepancy. In like manner each performative discrepancy considered above from its receptive aspect has a compositional aspect; and, as we see, it is even harder to propose mitigations for compositional discrepancies than it is for receptive ones.

(4) In sum, the discrepancies we noted earlier (now expanded to have both receptive and compositional facets) do not all have explanations such as to leave  $CH_{CG}$  unthreatened: indeed, roughly half of these discrepancies remain, after a number of proposed 'mitigations' have been tried, as discrepant as ever. Certainly it would be exceedingly unwise to claim that no other mitigations can plausibly be put forward;<sup>37</sup> but our attempt, and failure, to find such mitigations effectively shifts the onus probandi to the opposing position, that of the defenders of  $CH_{CG}$ . And this comment, by reminding us that the  $CH_{CG}$  has now a rival hypothesis, returns us to our main theme.

1.2.2 If there are discrepancies between the complexity-predictions of the MG and those observable in the behavior of the users of that MG, and if no mitigating explanations can be found for those discrepancies, then we must change our model of the MG. We must design a new MG in which the performatively-simpler truncates are more simply derived than the performatively-complex

<sup>37</sup> For example, it might quite plausibly be contended that purely semantic complexity in the Deep (or Abyssal) Structure might affect performative complexity. (But we have sidestepped this possibility, in all cases save 1.2.1.2.a and e, by comparing paraphrases only.) Or again, it might be suggested that performative complexity might be partly a function of Miller and Chomsky's "N(Q)" measurement (1963, p. 480) of sentential simplicity. (But this measurement gages only constituent structure, whether deep or derived; in gaging the complexity of a derived constituent structure it ignores completely the manner--simple or complex--in which that structure was derived, except insofar as that derivation conforms to the rough overall judgment that transformation decreases the number of branches per node and so increases N(Q).) Or yet again, there is the proposal due to Harris (1968, pp. 186f.) that sentences increase in performative difficulty upon their greatly diverging from the simple high-frequency formats (such as SAAD); "...in He went home, I think one is less sure what is the subject". (Most of the discrepancies noted here are not subject to this mitigation, it seems; and note that while "There's a dragon in the street" is quite divergent from the SAAD, it is itself of a high-frequency format, and so is not predicted to be more complex by this criterion.) Certainly Harris' proposal seems more than reasonable: and it does not, as cf. the "There's a dragon" case, reduce to yet another statement about divergences from Deep Structure or about transitional probabilities.

full Passives; in which the performatively-simpler "Mary grows flowers" has the simpler derivation; and "Dee is hard to please"; and "There's a dragon in the street"; and all the other cases we may expect future experimentation to disclose. In general: it does not appear that the **MG** actually derives these sentences (even in the "**5x8**" sense) in the way we have thought, and so we must consider changing the **MG** so that these sentences receive new and simpler derivations. Afterward, if the result of our changes could also be an **LG**, as customarily constrained by meta-theoretic considerations, then we can change the **LG** too and thus restore the equation **MG**  $\leftarrow$  **LG** (with the usual freedom of interpretation of "=").

(Incidentally, the demand just made does not quite correspond to the demand that all "brief" paraphrases, being the performatively simple ones, be entered as simple in our new **MG**, the more complex sentences either being derived from them or, perhaps, having independent derivations. For it is not always the case that the briefer paraphrase promises to be the performatively simpler one, as we have seen in 1.2.1.2., just above. There is nothing particularly distinctive, linguistically speaking, about the set of sentences [taken in its entirety] that are to be given simpler derivations in the **MG**.)

Pursuing this idea, we see that in the new **MG** the truncated Passive must, rather than being derived from deletion of a "copied" "by Agent" phrase, be related to the full Passive in quite another way. If the truncate were to be derived by a Passive-like transformation from a Deep Structure like the Active, then that transformation could not convert the deep-Subject into a "by Agent" phrase, since that phrase would only have to undergo deletion, exactly the discrepancy-producing operation we want

to avoid. In short, to introduce into the new MG a truncate whose derivation is of the simplicity apparently required, we must essentially introduce the truncate as a new variety of Deep Structure. But we have no mandate to make the full Passive a new variety of Deep Structure. Then we are speaking of two Deep Structures for transitive-verb sentences: one like the Active, and one like the truncate. And we are speaking of three derived surface structures for transitive-verb sentences: one the Active (from the first-named Deep Structure), one the truncate (from the second-named), and one the full-Passive (derived in the usual way from the first-named).<sup>38</sup>

1.2.2.1 Setting aside for the moment any reservations we might have about this step, let us determine what is to be done to accommodate the other discrepancies in the revised MG. Clearly, in general, the performatively-simpler sentences ("Mary grows flowers", "Dee is hard to please", perhaps "It is hard to please Dee", and "There's a dragon in the street") are to be entered as new Deep Structures, or very close to Deep Structures, with their performatively-complex counterparts ("Mary makes flowers grow", "[For anyone] to please Dee is hard", and "In the street there's a dragon" or "\*A dragon is in the street") either removed from the set of Deep Structures or, possibly, left there in the expectation that, thus reduced, the discrepancy between the MG prediction and actual performance might be explained on other grounds.

Assuredly, the resulting MG looks rather little like the LG we are accustomed to. It seems natural enough, in fact, to ask at this point whether we should not debate jettisoning the notion of the MG altogether, substituting for it some much looser abstract linguistic faculty: a set

---

<sup>38</sup> Postulating that in the MG the Deep Structure for the truncate is quite dissociated from that for the full-Passive implies that while the Passive tends to be detransformed in memory to the underlying Active-like Deep Structure, the truncate should not exhibit the same tendency, inasmuch as it is not, in the MG, underlain by an Active-like Deep Structure at all. And, in fact, Slobin has reported experimental results (1968) that lend strong empirical support to just this conclusion.

Note that, by strict DTC, a truncate is now predicted to have (ceteris paribus) the same performative complexity as its paraphrastic Active, if any (e.g., "Moll was deflowered" and "Somebody deflowered Moll"). This prediction, while not implausible, deserves careful checking, especially in view of Gough's finding (1966, pp.494-496) that truncates and non-paraphrastic Actives are not equal in performative complexity. In Gough's experiment, the Actives paraphrased the truncates except for having fulfilled Subjects (e.g., "the girl"); this fact of course makes the complexities uncomparable, and as Gough himself points out (p.495) the experiment was biased in favor of the Active in any case: his Subjects were asked to verify whether or not presented sentences described subsequently-presented pictures, and the pictures were ones to describe which "...a speaker of English normally uses the active voice...". That is, the pictures displayed overt actor and overt acted-on; so perforce did the Actives; but of course the truncates did not.

of sentences (with recursive rules to extend that set indefinitely), or some way of generating (determining) sentences that is so disorganized as not to constitute a "grammar" at all, in the usual sense of that partially-defined term. That is, why not take to their logical extreme the alterations proposed here, moving all sentences' derived constituent structures into the Deep Structure, and so having no explicit way at all of coupling sentences together as having the same subcategorial and co-occurrence constraints. The full Passive, for example, would (just like the truncate) be added directly to the set of Deep Structures. However, this extreme move has no psycholinguistic justification at all; quite the contrary. The set of anticipated discrepancies is far from encouraging it and the performative complexity of the full Passive militates against it; while on the other hand there is at least enough tendency for people to remember complexly-derived transforms as simpler ones (Mehler, 1963; Clark and Clark, 1968; but cf. Turner and Rommetveit, 1968) that a general detransformation toward the Deep Structure is almost inescapable. Moreover there are other facts about the linguistic behavior of individuals that lead one to believe that people do have an MG, a complex grammar in which such unobvious facts as the constituency of the Verb Phrase in the Deep Structure are "psychologically real" beyond the possibility of doubt. Lakoff and Ross (1966) have made some observations about the Benefactive Object in English that illustrate this point. As they say, the Benefactive Object ("Baron Geauxbois bought a new Cord for Tom") must not be inside the Verb Phrase (VP) in English, since if it were, by a reasoning they defend, it would behave like the Direct Object ("Baron Geauxbois bought the jury") or the Indirect Object ("Baron Geauxbois gave a new Cord to Tom") which are



inside the VP. For instance it would resemble them in becoming the surface-subject of a Passive ("The jury was bought by Baron Geauxbois" and "Tom was given the new Cord by Baron Geauxbois"). But it does not resemble them, since "\*Tom was bought a new Cord by Baron Geauxbois", the equivalent Passive on the Benefactive Object, is ungrammatical. And, since the Benefactive Object is not therefore in the VP, it can (like other elements not in the VP) be referred to in a "do so" construction: hence "Baron Geauxbois bought a new Cord for Tom, and will do so for Barbara next Tuesday" is grammatical, while the same construction using a Direct or Indirect Object--these are inside the VP--would not be grammatical: "\*Baron Geauxbois gave a new Cord to Tom, and will do so to Barbara next Tuesday", and "\*Baron Geauxbois bought the jury, and will do so [? to] the judge next Tuesday". The two facts--no Passive, yes "do so"; yes Passive, no "do so"--are coupled, in Lakoff and Ross's treatment, and are explicitly predicted by whether or not the Object--Benefactive or Direct/Indirect--is inside the VP. It might still, I suppose, be argued that these facts would make as much sense in some other and much looser account--one in which the formation of Passives and "do so" sentences was oblivious to the constituency of the VP--but any such notion is completely overthrown, to my mind, by the following facts. In my own speech the Passive on the Benefactive Object which Lakoff and Ross condemned--e.g. "Tom was bought a new Cord by Baron Geauxbois"--is perfectly grammatical. There can be no doubt on this score, since I find completely grammatical the Benefactive-Object Passives of a number of verbs: "obtain", "procure", "steal", and others. Thus, according to Lakoff and Ross's criterion,

in my MG the Benefactive Object is inside the VP. Then, again according to Lakoff and Ross, I should find ungrammatical the (to them, grammatical) "do so" sentences like "Baron Geauxbois bought a new Cord for Tom, and will do so for Barbara next Tuesday". Now, if I should find this sentence ungrammatical this would be a striking fact, since in every other case of which I am aware if an American finds a sentence-type grammatical then I do too: mine is generally an "inclusive" dialect. I do not find the Benefactive Object "do so" sentence grammatical. (In fact, I even have trouble understanding it and am slightly incredulous that everyone else is not similarly puzzled.) Thus the claim that, for my MG, the Benefactive Object is in the Deep-VP, turns out to have a rather immediate verification, one which strengthens in turn the claim that the MG of the speakers with whom Lakoff and Ross were concerned is such that the Benefactive Object is outside the VP. Then, if so unobvious a fact as the constituency of the deep Verb Phrase can be demonstrated to have psychological reality in the MG, one is not at all disposed to assume that transformations lack this reality, in some sense. Notice that on the basis of VP-constituency in the Deep Structure one can predict facts about surface-structure elements which are not in the surface VP at all-- e.g., the surface-Subject of the Benefactive-Object Passive--though they are elements withdrawn from the Deep VP by, precisely, transformation. Thus, in conclusion, there is good evidence that the MG is in form very like the LG, with its complex interplay of rules and elements, and there is no evidence to the contrary. There is only the evidence, developed above in this paper, that the contents of the MG seem not to be exactly that of the LG to which we are accustomed.

1.2.2.2. It is clear that the kind of MG we are moving toward must be judged by criteria of economy different from those by which the LG is now judged. This kind of MG--let us now begin to call it the "Abstract Performative Grammar", or APG--is quite uneconomical on the points where it differs from the LG (or CG "=" to the LG). In developing the truncates, for example, without regard to their extremely close linguistic relation to the full Passives, the APG will fail to exploit that close relation in the obvious and natural manner: by giving the two types of sentence a common derivation up to the point (deletion) where they must part ways. It is a commonplace of linguistic discussion that an irrefragable argument for inclusion of transformations in the grammar is that only transformations permit capture of this basic economy, for only transformations can operate upon essentially one underlying structure (like the Active), properly alter it into two other structures (the abstract surface Active and the abstract common Passive/truncate), and then alter the common Passive/truncate into Passive and truncate. Thus the severe restrictions that determine Actives containing (roughly speaking) transitive verbs and their objects, and that determine Passives containing surface-Subjects and, optionally, a surface-by+Agent--these restrictions can be made one set of restrictions essentially on one underlying structure, the one underlying Active, Passive, and truncate. (Thus to say that a verb is transitive in an Active sentence with an object is tantamount to saying that the same verb can occur Passivized in a normal Passive sentence; information about the Active and the Passive is reduced to information about the set (Active, Passive)). This has been the view since the outset of transformational work (e.g. Chomsky, 1957, p. 43) and it has been maintained up to the present day. That is, despite the

uncertainties about the precise nature of evaluative measures and about the precise form of the grammars to which these measures are to be applied, even so there has been the persistent notion that a grammar that made correct generalizations (that collapsed duplicative rules into fewer and non-duplicative rules) should turn out to be more highly-valued than a grammar that failed to make such generalizations.<sup>39</sup>

An APG does not include all of the generalizations that the LG includes, and so it is definitely less highly-valued by the criterion referred to in p.39 just above. In particular, the APG is not economical in its treatment of the set of sentences (Passive/truncate), and in general we can see that there are sets of sentences (x, y) such that there are possible "generalizations" about them that are realized in the LG but not in the APG. Thus we may say that the LG and APG differ most basically in that insofar as the sentences x and the sentences y are taken as a single set, and insofar as the set (x, y) is that to which the evaluation measure is applied, to that extent the LG will consistently be more highly-valued than the APG.

However, the APG was clearly introduced to effect another and quite different sort of economy. For we have concluded that if truncates are performatively less complexly-derived than the LG "predicts", then there is warrant, provided no mitigations for this discrepancy can be found, for changing the MG so as to simplify the derivation of the truncates; and thus, tacitly, we have taken the position that the MG must realize certain economies of derivation not realized in the LG (CG). Taking the Passives and truncates as our (x, y) example, the LG minimizes the number of rules by which the set (x, y) is derived; but the APG minimizes the number of

"We have a generalization when a set of rules about distinct items can be replaced by a single rule (or, more generally, partially identical rules) about the whole set [*italics mine, wcw*], or when it can be shown that a 'natural class' of items undergoes a certain process or set of similar processes. ... The problem is to devise a procedure that will assign a numerical measure of valuation to a grammar in terms of the degree of linguistically significant generalization that this grammar achieves. The obvious numerical measure to be applied to a grammar is length, in terms of number of symbols [that is, symbol-tokens, wcw]." (Chomsky, 1965, p. 42.) Chomsky then proceeds to remark that if length is to be the criterion there must be adopted consistent notations and rule-formats so that the criterion can be correctly applied.

rules by which the set (x) is derived and, separately, minimizes the number of rules by which the set (y) is derived. Implicitly so far, but now explicitly, we have been saying that the APG is at least sometimes to be measured by an "abstract performative" evaluation measure in which a set of rules is more highly-valued if it individually derives each paradigm of sentences with the fewest number of rules. Thus in the APG the truncates receive a generative treatment that is essentially the treatment they would receive in the LG if there were no full Passives in English (and if the residual generalizable similarities between truncate and transitive Active--e.g. identity of their verb-sets--were ignored). To take another example, "Mary grows flowers" receives a treatment in the APG equivalent to the treatment that that sentence would receive in the LG if none of the sentences like "Flowers grow" or "Mary makes (something happen)" were included in English.

How far the APG is to be evaluated by the proposed "abstract performative" evaluation measure is still an open question, since there is no a priori way of prescribing the extent to which the APG should obey "abstract performative" rather than "competence" or "linguistic" simplicity criteria. Certainly at this point there seems to be no reason to insist that the APG derive the full Passive by the minimum number of steps, since the full Passive (taken as a monolithic set) seems to be performatively complex. However, conversion to an APG highly-valued only on the "abstract performative" criterion is not impossible.

By the same token, the actual form of the rules specific to the APG will not be stated here beyond the kind of more-or-less obvious descriptions exemplified in discussions above. Presumably the truncate will have a direct APG derivation

much like N--is--Adjective sentences; and "Mary grows flowers" will, with a complex underlying lexical entry for "grows", have an APG derivation like "He buys cars". "There's a dragon in the street" might in the APG have a directly-derived pleonastic "there-is" (there + be + Tense) adjunct. These investigations are scarcely begun, and attempting to state such rules with any finality would be entirely premature.

1.2.2.3. Two of the fundamental facts to be learned about a language in the course of its acquisition are which sentences of superficial similarity must in fact be distinguished as different, and which sentences of superficial difference must be generalized as being basically similar. Thus the learner of English must realize that "Dee is hard to please" and "Dee is eager to please", despite their superficial similarity, must be viewed as being quite different at a more basic level (the level we call Deep Structure), as see (1.2.1.2.d) and (1.2.1.2.e), above. Thus also the learner of English must realize that "The King of Zembla rewarded the winning athletes" and "The winning athletes were rewarded by the King of Zembla" are, despite their dissimilarity on the surface, basically the same at some deeper level (again, that of Deep Structure). To quote Chomsky on differentiation, "...the grammar of English, as a characterization of competence, must, for descriptive adequacy, assign different deep structures [to sentences superficially similar but fundamentally different]. The grammar that each speaker has internalized does distinguish these deep structures..." (1967, p. 433). And on generalization, "...a certain empirical claim is made, implicitly, concerning natural language. It is implied that a person learning a language will attempt to formulate generalizations [for sentences

superficially different but fundamentally similar or identical] that can easily be expressed.... in terms of the notations available in [the theory of grammar], and that he will select grammars containing these generalizations over other grammars..." (1965, p. 45).

Evidence that speakers differentiate and generalize is not hard to find. That the speaker "knows" there to be a profound difference between the superficially-similar sentences "I persuaded the doctor to examine John" and "I expected the doctor to examine John" is shown by the fact that all speakers find the difference manifest when the embedded sentence (The-doctor-examine-John) is Passivized, since the resultant Passives have distinctively different relations to their Active counterparts just cited. "I persuaded John to be examined by the doctor" does not have the same truth value as the corresponding Active (one of the two sentences can be false, the other true). But "I expected John to be examined by the doctor" does have the same truth value as the corresponding Active (if one is true the other is; if not, not) (Chomsky, 1967, pp. 432f.). The latter fact about paraphrasticity is, in turn, one kind of evidence favoring generalization. Other evidence in favor of generalization is that Speakers also commonly "know" that for every sentence of type a (e.g., Active) there can be constructed a sentence of type b (e.g., Passive) with the same meaning or (as with Affirmative/Negative) a constant difference or increment of meaning. (Informants do not have trouble turning Actives into Passives.)

As to the above noted differentiations, no psycholinguistic evidence that I know of encourages us to think that they are not incorporated directly into the MG (now identified as an APG), in some way. On the other hand, we have evidence that seems to indicate



that some generalizations (e.g., that truncates correspond to full Passives) are not incorporated in this way. Yet surely the manifest nature of the generalization must be (unconsciously) realized by everyone who learns English. How then can we possibly explain the seeming failure to form the appropriate rule in the MG?

The sentences of performative simplicity that we propose to enter in the APG as derived (like the Active) "directly" from Deep Structures have in common another attribute in addition to their comparative simplicity: they are early-learned. For most of the cases considered here this must remain for now just a plausible assumption: we assume, then, that a child acquiring English will acquire "Mary grows flowers" before "Mary makes flowers grow"; "There's a dragon in the street" before "In the street there's a dragon"; "John hammered the nail" before "John used the hammer to act upon (or hit) the nail"; and so on.<sup>40</sup> But one of these assumptions has some support, since it is known (Harwood, 1959, pp. 248f.; Brown and Hanlon, this volume, p.000) that children at least use truncated Passives before full ones. This fact alone does not entail that children use truncated Passives before they are able to use full ones, since it is a well-known result (Fraser, Bellugi, and Brown, 1963) that comprehension precedes production and therefore that the abilities of the MG can outstrip those of the compositional performative mechanism (McNeill, 1966, pp. 76-82). Indeed, one of Fraser et al.'s results is specifically that children comprehend full Passives before they produce them; but these comprehensions, it seems, might have been achieved just by comprehending the truncate portion of the full Passives concerned. Pending further evidence, we find, as stated, "some support" for our conjecture that MG's command truncates before full Passives.

<sup>40</sup> There may well be exceptions to this pattern, however (the assumption may be risky in any case). For example, children are often observed making sentences like (an actual example) "He made her dead", meaning "He killed her"; but this does not at all entail that the child cannot make sentences with causative" verbs ("kill", = "make [be] dead"; "grow", = "make grow"). This would have to be determined empirically, of course, but it might be only that the child initially has trouble with "causative" verbs which have a form different ("dead"/"kill") from that of the "caused" predicate ("be dead"; "grow"). This at least seems to be an avenue of investigation worth pursuing. (It is perhaps worth remarking that the children of Harwood's study (*ibid.*), whose age ranged from 4.11 to 5.8, had both kinds of construction. "Make" plus a verb or adjective ("You make me sick", cited p. 239) was common, and so was e.g. "break" in the "causative" sense of "make...break". The "causative" verbs (incorporating "make") seem to have been more common than the construction "make + verb" (*ibid.*, pp. 246f.); but this of course does not demonstrate anteriority of the former construction.)

It is also worth remarking that empirical failure of this generalization would not jeopardize the APG hypothesis itself, but only the hypothesis, here being proposed, concerning the explanation for the apparent failure of the MG to embody all possible generalizations. If it should turn out that some of the performatively-simple sentences here in question are learned late rather than early, a more complex explanation for the failure to generalize would have to be put forward. That step may perhaps await the empirical disconfirmation.

What shall we suppose the child's truncate-formation rule to be at this early pre-Passive stage? Shall we assume that he has what we might call a "Pre-Passive Passivization" rule which acts like the true Passivization rule except that it attacks only Actives that are to be turned into truncates? This is a possible assumption, perhaps. But in formulating such a transformation within his MG a child must (unconsciously, of course) be formalizing (and maximally exploiting) the generalization that for every Active of a certain form there is a paraphrastic truncate<sup>41</sup>. But such Actives scarcely exist: only those, in fact, with a verb like "deflower" or "beget" and a very indefinite Subject (like "someone"). That is, truncates do not correspond to Active sentences in anything like a simple way, except in cases which the child can hardly have come into contact with. Then, since truncates have neither surface-forms like the Active nor an obvious relation to the Active's Deep Structure, there is no good reason to suppose that a child who makes truncates (and Actives, of course), but no full Passives, has made the generalization needed to permit him to construct those truncates by recourse to a "Pre-Passive Passivization" rule<sup>42</sup>.

It seems more than reasonable, in fact, to think that the child is constructing truncates ("Those cookies were baked") by the same rules he uses to construct simple Predicate-Adjective sentences ("Those cookies were good"). The surfaces are very similar, and participles like "baked" are quasi-adjectival<sup>43</sup>. But if this is in fact the way the child makes truncates his truncate-formation rule is as simple and direct as is his rule for forming the Predicate-Adjective sentences: that is, both come directly from a Deep Structure on the order of "Subject-NP + (Be-verb + Tense) + Predicator", where a Predicator is either a Participle ("baked") or an adjective ("good")<sup>44</sup>. That is, the child has a much simpler way of making truncates than the adult

41 He might, just conceivably, be formulating the sort of "meaning-change" generalization such as may underlie the awareness of the similarity between e.g. Affirmative and Negative, which sort of similarity is realized in the LG as a systematic similarity of Deep Structure. That is, he might be forming the generalization that the truncates, while not synonymous with the Actives, can be viewed as preserving all of the Active except what is irrecoverable from the Active's deep-Subject, so that each truncate would be seen as incorporating a diminution of the semantic content of any one of a set of Actives. (For example, the truncate "The Hindenburg was sabotaged" corresponds "by diminution" to the infinite set of Actives including e.g. "A maniac sabotaged the Hindenburg," "(n) maniacs sabotaged the Hindenburg" (where n is any positive integer), and so on; but it does not correspond to e.g. "\*The Julian calendar sabotaged the Hindenburg." The relationship does not seem a simple one, though for a possible further simplifying factor see n.42, just below.

42 The paucity of Actives corresponding paraphrastically to truncates is exactly matched, of course, by a paucity of matching full Passives in the speech of the older child who uses full Passives. The older child's generalization 'Truncate = (full Passive minus by- $\Delta$ )'--a necessary step to achieving the LG account--could not be one formed on the basis of overwhelming linguistic evidence as such. But notice that a generalization to the truncate from the corresponding full-Passive is aided-- as the generalization from the corresponding Active is not--by the very close correspondence of surface-form. And note further that in language-use ("parole") the missing Agent is often recoverable from deictic context or from preceding linguistic material; and this, in a way that is not at all obvious at present, must play a role in promoting the correct LG generalization. However, as we have noted, it seems likely that the generalization is never made at all in the APG.

43 Some past participles are more adjectival than others: for example, some ("interested," "tired," "annoyed") can be modified by "very," while others ("baked") cannot. But generally they can all be compared ("more interested"; "more tired" or even "tireder"; "more baked"); modified by "quite"; and so on. The subject has been studied elsewhere, and need not be dwelt on here.

This comment on the APG is manifestly derived from the LG analysis, due to Postal and Lakoff (Lakoff, 1965, Appendix, A, passim) in which Verbs and Adjectives (and not just Past-Participles and Adjectives) are essentially the same "VERB" grammatical category, differing (ibid., 0, p. 15) only by one syntactic feature: conventional verbs are (+V) and (-Adjectival), "whereas conventional adjectives are (+V) and (+Adjectival). Notice that if the ruling grammatical category were 'predicator' rather than 'verb', so that (+V) would not automatically head this category's complex symbol, then verbs could be (-i-Predicator) (+Verb) (-Adjective); adjectives could be (+Predicator) (-Verb) (4-Adjective), and past participles could be (+Predicator) (+Verb) (^Adjective). That is—without necessarily advocating this description for the LG--the close relation between LG and derivative APG descriptions is easily demonstrated.

would if the adult's MG were "=" to the LG.

We see then that to demand that the child later acquire the LG's way of making truncates is to insist that he abandon a simple way of making truncates in favor of a more complex way. It is to insist that he forego the simplicity of his (putative) original derivation of the truncates in order to achieve a simplifying generalization holding for truncates and full Passives together; it is to insist that he give up a simple way of deriving the set of sentences  $\underline{x}$  because, when he acquires the set  $\underline{y}$ , the set  $(x,y)$  can be more simply derived if he uses the necessarily more complex derivation of  $\underline{y}$  to derive  $\underline{x}$  also, amplified (as by providing for recoverable deletion) where needed. It is to insist, in short, that CG or LG or "competence" simplicity is superordinate to APG or "abstract performative" simplicity, and to insist that the efficiency with which individual sentences are composed and parsed is subordinate to the efficiency with which sets of sentences are stored.

In sum, the presumptive abandonment of the early derivation of the truncates in favor of the LG derivation is not an event of linguistic ontogenesis that inevitably accords with intuition; in fact, especially in view of the probable performative simplicity of the truncates in the adult MG, the event is somewhat lacking in plausibility. And, as we see, defenders of the CG hypothesis are not spared the burden of proof, for the CG hypothesis must include an explanation of why  $(x,y)$  generalization is paramount, just as the APG hypothesis must include an explanation, presumably along the lines sketched here, of why  $(x,y)$  generalization is sometimes subordinated <sup>45</sup>.

CG and APG criteria of economy are competitive by nature, no matter how we refine their definition. They need not always come

<sup>45</sup> For a persuasive account of some CG generalizations, see n. 46, just below.

into conflict; the set  $(x,y)$  is often such that a single derivation is possible (obeying CG simplicity) that is still very simple by the APG criterion: e.g., when  $x$  is the set of transitive-verb Actives and  $y$  the set of intransitive-verb Actives; or perhaps when  $x$  is the set of simple Affirmative sentences and  $y$  the set of corresponding Negatives.<sup>46</sup> But sometimes, we have supposed, the two criteria do conflict, as in the cases considered above. In the case of the truncates, it appears that the APG criterion is the stronger. But we are not obliged to assume that the APG criterion is always the stronger; it would be rather more reasonable to anticipate that determination of which is stronger is at least partly contingent on the degree of complexity that an  $(x,y)$  generalization would introduce into the derivation of  $x$  alone. Investigating this possibility must assuredly be an early task in any exploration of the APG hypothesis<sup>47</sup>.

In conclusion, we see that the chief difference between CG and APG boils down to this: the CG puts a premium on over-all economy and so makes all significant generalizations; the APG puts a premium on economy of derivation of individual sentential paradigms, and so balks at incorporating some of these generalizations.

1.2.2.4. Having determined what must apparently be altered in the MG--having determined that the MG is apparently an APG---we now ask whether or not the LG could be revised in turn so that, once again, the formula  $MG = LG$  could be restored (as see (1.0.), above). To continue with the example of the truncates, could the APG "direct" analysis of that set of sentences replace their conventional treatment in the current LG? We see at once that if the APG analysis of the truncates were to replace the present LG derivation (from the full Passives via by- $\Delta$  deletion) then the LG would fail to make the obviously-correct  $(x,y)$  generalization (necessary to preserve high-valuedness



<sup>46</sup> McNeill, in his penetrating study of linguistic development in the child (1966), proposed a specific reason for the child's development of a Negative transformation (that is, one that forms Negative surfaces from Negative Deep Structures). "The pressure...to devise transformation rules may come from the cognitive clutter that results from not having them" he remarks (pp.61f.). "In Period 1 [a stage of acquisition] the child had to remember only two rules for the placement of the negative ['no' or 'not']. In Period 2, he had to remember five. By Period 3, if it were not for the transformations, he would have had to remember six or seven. It is possible, then, that the load on the child's memory by Period 2 was so great that the...transformational rules we observe in Period 3 were precipitated." And he continues, "The child needs to process sentences in short intervals of time; presumably it takes less time and a child tends to forget less when the placement of the negative is done by transformational rules rather than by independent...[Deep Structure] rules."

Essentially the transformation that McNeill proposes replaces a burgeoning number of special phrase-structure rules for positioning of the negative. The new rule consists of a single instruction which takes the negative particle from the beginning of the sentence and adjoins it to the auxiliary ("does", &c). Thus the different ways of making different Negative sentences are collapsed into one, so that the child's need "to process sentences in short intervals of time" is plausibly served by this simplification.

But if we grant that the child already knows how to make correct truncates, modeled on 'Noun - is - adjective' sentences, then no analogous economy can be achieved by substituting a new transformational account of the truncates for the old "direct" account. He has no burgeoning set of special-purpose rules to be gotten rid of: he has only the freedom expressed by the node 'Predicator' in the rule, needed for predicate-adjective sentences in any case, 'S → NP + be + Predicator plus other rules (Question; Negation...) also needed anyway. Thus the child, by McNeill's criterion, has nothing to gain from a new (and more complex) derivation of the truncate. McNeill's explanation for the addition of transformations to the MG, which to me is quite persuasive for the Negatives, seems to have no bearing on the truncates, and so this best explanation of CG-generalization seems to have a systematic set of exceptions: those we have been claiming to be ungeneralized in the APC. The two explanations--McNeill's of cases favoring generalization and hence "competence" economy, mine of cases opposing

generalization because favoring "abstract performative" economy--are far from mutually exclusive and evidence that supports one criterion does not necessarily refute the other.

<sup>47</sup> It is also possible that other factors play a role in determining, for a given case, whether CG or APG criterion shall prevail: for example, length of time between acquisition of x and acquisition of y (= length of time for the derivation of x to harden); relative frequency of x and of y; nearness of x and/or y to some z; and so on.

by a "competence" criterion of simplicity) about the truncates and full Passives, unless the LG also gave up its present account of the full Passives and adopted an (x,y) account in which both truncates and full Passives derived "directly" for a Passive-like Deep Structure. But then the LG would only have to abandon the obviously-correct (x,y) generalization about the full Passives and the Actives<sup>48</sup>. In short, there is no escaping

the conclusion---manifest in the preceding discussions, in fact---that if the MG is to become an APG then the MG can no longer be "=" to the LG, because the LG could not itself become an APG (without changing a fundamental characteristic---high-valuedness by a "competence" simplicity criterion--- that it has been assumed from the beginning of generative transformational studies to possess).

That is,

$$MG \equiv APG \neq LG.$$

But then the LG must embody many generalizations that the MG does not; yet the LG is supposed to be a representation of the human linguistic faculty. We are therefore constrained either to renounce this representative function of the LG, or else identify where in the human linguistic faculty the LG's now-distinctive generalizations, absent from the APG, are after all represented. There can be no question of renouncing the representative function of the LG, since there is no reason whatsoever to claim that speakers have not, tacitly, made the generalizations at issue: it is impossible to believe that tacit realization of the full-Passive/truncate relationship can fail utterly to be made. Then we are obliged to infer that the LG is perfectly correct in making its non-APG generalizations; and so we must seek to identify how these generalizations are represented in the linguistic

<sup>48</sup> This is a simplification, since some truncate-like sentences do have a derivation like the one the APG seems to provide for the true truncates: NP--be+Tense--Predicator. Thus "The glass is broken" or "The glass has been broken for some time" seem to contain, not a participle derived from a verb via Passivization (with absence of the deep-Subject arranged by deletion), but rather just an adjective: "broken". The meaning of such sentences seems to be on the order of: "The glass (has a crack) (is not whole)" and "The glass has (had a crack) (not been whole) for some time", respectively. Thus, when the LG analyses a "pseudo-truncate" in this fashion, since no Passives are involved, the LG's (truncates, full-Passives) generalization is unaffected. (But, it seems, in the APG both pseudo-truncates and true truncates have the 'adjective' ('predicator') analysis, with the consequence that the LG, though able to adopt some of the APG's analyses, cannot adopt all.)

Obviously the sentence "The glass broke" might now be resolved into "The glass became broken", thus appearing as an inchoative sentence subject to the analysis of Lakoff (1965, IV, pp. 4-14). The deceptively-simple Active "John broke the glass" would now be seen to have a deeper source like "John made the glass (be, become) broken", revealing itself to be a causative sentence in Lakoff's sense (ibid., pp. 14-18). Thus introduction of the adjective "broken" suggests an LG analysis for the true truncate along these lines: "The glass was broken" ← "The glass was broken by Δ" ← "Δ broke the glass" ← "Δ made the glass (be, become) broken". But this analysis is perfectly captured by the (x, y) generalization just indicated, and so the LG easily changes over to this analysis of these truncates without losing any part of its "Competence" status.

Of course if the LG should adopt this analysis then such a sentence as "The glass was broken" would be ambiguous, meaning either "The glass wasn't whole" or else "The glass was broken by Δ". But this seems perfectly in order. Notice however that many superficially-similar sentences are not ambiguous in this way: "The house was broken into" absolutely requires an Agent, a deleted "Δ". With some verb/adjective pairs (e.g., "divorce"/"divorced") there will be both ambiguous sentences ("John was divorced") and unambiguous sentences ("John was being divorced at this time last Monday"). The latter, again, can only be a truncated Passive; "divorced" when a pure adjective is a "stative"

(no activity) adjective like "tall" rather than a "non-stative" adjective like "noisy" or "foolish" (Lakoff, *ibid.*, Appendix, A, pp. 9f.), and a "stative" adjective cannot occur in such a "non-stative" (activity) environment.

We have gone into this matter at this length (while at the same time oversimplifying and omitting many details) because it directly affects our psycholinguistic measurements of the performative difficulty of the putative truncates. For now we see that it would be possible to conclude that true truncates were performatively simple when in fact it was only pseudo-truncates like "The glass is broken" that were being tested. Or the sentences being tested might be ambiguous (might be sets of non-paraphrastic homonymous strings), as "The glass was broken" seems to be, in which case (a) the "prejudiced" reading would, quite covertly, be the "adjectival" one, thus again yielding a misleadingly "simple" performance; or (b) the sentence would be (tacitly) recognized as ambiguous, in which case either "truncate" or "pseudo-truncate" reading would be complicated by this covert ambiguity (see n. 30, above). In testing speakers' performances on the truncates, then, these factors must be carefully excluded. Once the risks are identified, however, this is easy to do. Most sentences can quickly be typed as "truncate", "pseudo-truncate", or "both" (ambiguous): e.g., respectively, "That window was broken deliberately"; "That window has been broken for weeks"; and "That window was broken on Friday".

There appears to be no way, regarding these sentence-types, of reconciling APG and LG; if the LG incorporates the APG's analysis of the truncates it does so only for the superficially similar pseudo-truncates. As I hope is obvious, there is no way of merging truncate and pseudo-truncate by claiming that e.g. "The glass was broken" is both purely adjectival and also has, as a complement to the adjective "broken", a deleted Passive "by A" phrase. (For a different proof that the related "The glass broke" could not have such a source, see Lakoff, *ibid.*, IV, p. 17.) This claim would entail—pretending illustratively but counterfactually that "A" can in this instance be realized as "someone"—that "The glass was broken" would derive from "The glass was broken by someone", which like any Passive would have an Active-like source like "Someone broke the glass", which, we have said, in turn has the source "Someone made the glass be broken". But if "broken" always has a deleted "by A" phrase, then the last sentence must in its turn come from "Someone made the glass be broken by someone", which comes from "Someone made someone make the glass be broken", which comes from "Someone made someone make the glass be broken by someone", and so on in an infinite regress.

faculty if they are exiled from the MG where formerly they were lodged.

It would be premature to attempt any definitive answer to this question, clearly; but it is at least obvious, under  $CH_{APG}$ , that the linguistic faculty must contain, besides the MG (APG), knowledge about the MG. The linguistic faculty must include the knowledge that Actives and full Passives, though they are independently-derived in the MG, are closely-related under the paraphrase bond. This then is the Archival Competence Faculty anticipated in (1.2.), above<sup>49</sup>. Thus the linguistic content of the CG is now redistributed over two components: the MG, now an APG; and the Archival faculty just mentioned. And the content of these two components taken together is, quite properly, represented "axiomatically" in the LG.

This should not really be a surprising result. That the generalizations of the LG should be uniformly represented in the linguistic faculty could not be an a priori assumption. The facts that the LG is a grammar, and that the linguistic faculty appears to include a grammar, never strictly entailed that the two grammars were essentially identical: this might have turned out to be the case; but it seems that it cannot. The MG seems not to bear the direct relationship to the LG that it has generally been thought to: but neither the existence of the MG nor the form and content of the LG has been in any way jeopardized. This point is perhaps worth emphasizing. It means that the fundamental notions of linguistics, concerning the LG, have been in no way threatened by the APG hypothesis. It also means, of course, that the nature of the MG can no longer be inferred directly from the nature of the LG: in many cases (in all cases, at first) it must be discovered independently; but this in turn means that one can try to find out what is actually in the MG without being hampered by the assumption that everything in the LG is [in

<sup>49</sup> Many readers will at this juncture be reminded of Zellig Harris' model of transformational grammar and of his way of commenting on that model (1957; 1965; 1968). The reminder is apt. It is very probable, to my mind, that the Archival Competence Faculty just postulated will in many respects have the "equilibrium" character (Harris, in lecture, about 1962) of this model, in which sentences are not derived from abstract (non-sentential) underlying structures, but are rather related (psycholinguistically, "tacitly known to be related") to other sentences. I would insist that the sentences thus archivally related be assigned their ("derived") constituent structures, and that these structures be related; and I cannot agree with Harris that two sentences bound by this relationship must, for any n-tuple of words occurring in both together, be identically-ordered (identically-sequenced with all other like sentences) on a scale of grammaticality (1957, pp. 288f.; 1965, p. 368, and n.) Some Actives improve with Passivization, but most do not. But in general it seems to me that an Harrisian or "equilibrium" Archival Competence Faculty may well be what is indicated.

For readers unacquainted with Harris' concept of transformational grammar, insofar as it contrasts with Chomsky's, Grunig's account (1965-1966) may serve as a useful introductory comparison.

some isomorphic way] also in the MG.

I have referred to the new sort of MG as an "Abstract Performative Grammar" and have used the term "abstract performative" criterion of simplicity. The term was adopted for the obvious reason that an important concern of the MG seems to be to conserve on the derivations of individual sentence-paradigms, an issue which does seem to be basically performative. But the APG has been clearly distinguished from the mechanism actually used to parse sentences: that is, it is not to be confused with the mechanism that makes use of or refers to the rules of the MG in the course of acts of linguistic behavior as such. (We opened this discussion by specifying that the MG we were going to concern ourselves with would be, in the most abstract sense possible, the MG whose analyses were imposed on outgoing and incoming sentences, by whatever mechanism and in whatever way; we emphasized that this was the MG that must be "=" to the LG if any MG is; and it is this MG that we have now hypothesized not to be "=" to the LG.) The APG is no less "abstract" in this sense than the CG was: it too is a passive faculty, it too could just as well be assumed (but for the non-finiteness of its output) to exist in the maximally abstract "5x8" version cited facetiously above. Hence the "Abstract Performative Grammar" is both "Performative" and "Abstract". It is also envisaged to be, in every respect but that of obeying a "Competence" or LG simplicity criterion for sentential generalizations, a "Grammar". Thus, I submit, the APG is fitly so named.

If the APG hypothesis were to be definitively upheld it might be natural to distinguish two kinds of "Competence": "Abstract Performative" Competence (not an oxymoron), and "Archival" Competence; but this thorny issue were better postponed.



1.3. We began by specifying as the object of our interest the abstract MG whose analyses (derivations) were the ones that the performance mechanism, in whatever way, imposed on sentences being composed and on sentences being parsed. We readily accepted the notion that the rules of this MG are probably not used one-after-the-other to compose or in reverse order to parse; indeed, we postulated a somewhat facetious model in which sentential derivations were printed out on 5x8 cards and in which parsing and composing consisted of finding the right card. We concluded that even this model strongly supported the notion that MG complexity would be reflected in performance (behavior) complexity---unless distortions intervened---because the more complex sentences would occupy more of the 5x8 card and so take longer to access. These preliminaries defined the nature of our problem: to ascertain whether or not available and reasonably-predicted evidence supported the hypothesis ( $CH_{CG}$ ) that this abstract MG was in fact the LG. The problem first resolved into that of seeing what performative factors (e.g., TCA) might intervene between MG and performance in such a way as to distort the correlation, as by making sentences that were relatively simple in the MG, relatively complex to parse or compose. Having discussed what seemed to be the most plausible of the possible significant distorting factors, we then listed a number of (predicted) discrepancies between MG and performative complexities. And then we asked whether or not these discrepancies seemed all to be "mitigated" by the distorting factors already discussed. We concluded that they were not all mitigated. This led to our conclusion that therefore  $CH_{CG}$  was wrong, that the MG was not a CG " $\approx$ " to the LG, and that an alternative hypothesis should be advanced, conformable to evidence of performance and evidence of children's learning-sequences, about the nature of the MG. We stated a new hypothesis,

that the MG was in fact an Abstract Performative Grammar. This version of the Correlation Hypothesis, CH<sub>APG</sub>, was then further discussed, a very tentative sketch of the APG was given, and it was suggested that the linguistic faculty consisted of (at least) two components, the APG and an Archival Competence Faculty. The latter was briefly outlined.

I have agreed with Fodor and Garrett (1966, 1967) that there is little promise in the notion that the complexity of the sentential analyses of the LG will be directly reflected in the complexity of the human performance of those sentential analyses. From this point on however our opinions radically diverge. Where they persist in the assumption that the MG is a CG is an LG, and further assume that the performance-discrepancies are to be accounted for by an "abstract" relationship (1967, p. 296) between the CG and performance, I on the other hand have rejected this assumption, because no plausible performance factor actually seemed capable of accounting for all discrepancies and because in any case the assumption had no a priori warrant; and I proposed the new hypothesis just summarized. Thus the predicted discrepancies between LG and performance are explained in completely different ways: in Fodor and Garrett's account the discrepancies occur between MG (= LG) and performance, whereas in my account, CH<sub>APG</sub>, the basic source of these discrepancies occurs between LG and MG (= APG). As I made clear, however, none of the possible performance distortions of comparative complexities---e.g., TCA or even TVDDI, conceivably---can be ignored; indeed, some discrepancies will be wholly or partially mitigated by taking such distortions into account.

Since the LG/performance discrepancies are by CH<sub>APG</sub> asserted to occur chiefly between LG and MG, the new hypothesis must, to be complete, include an account of how the APG can be one thing and the LG something else, where both are correct. The main outlines of that account have already been

indicated: The LG must include the linguistic content (e.g., the (x,y) generalizations) of both MG (= APG) and the Archival Competence. But the details of this proposal---hence its confirmation---very much remain to be worked out.

The contention between  $CH_{CG}$  and  $CH_{APG}$ , as I hope is clear, is no mere terminological issue. The two hypotheses propound quite different notions about the nature of the MG, about the nature of the relation between MG and LG, and about an important aspect of the ontogeny of the rules being integrated into the MG. These, surely, are the most serious issues confronting present-day psycholinguistics.

2. The Strong Inclusion Hypothesis.

2.0. The question to be raised below is, briefly, this: To what degree is the observer entitled to assume, of a "wellformed English sentence" emitted by a child, that that sentence is assigned by the child's mental grammar the structure (the derivation) assigned the same sentence by the adult mental grammar? I resume use of the term 'mental grammar' (MG) to avoid argument over whether or not this device is a CG or an APG; but I note in passing that, under the CG hypothesis, the question at hand has the form: To what degree is the observer entitled to assume that a child's 'wellformed English sentence' is assigned by the child's CG the structure (derivation) assigned the same sentence by the LG, the grammar of English constructed by linguists? Because our question is quite divorced from both CG or APG hypotheses (or so I will maintain here), I will generally develop this section as if the CG were the MG at issue, chiefly because this tactic permits appeal to familiar examples from the linguistic literature. Where it seems advisable I will insert tentative comments on the possible effect of a substitution of APG for CG.

Ignoring the APG for the moment, then, and thus feigning belief in the proposition that the (idealized) adult MG is in some sense (perhaps the "axiomatic" sense) the Linguistic Grammar, we broaden our basic question and restate it in this way: Is the child's (idealized) MG 'strongly included' in the adult's (idealized) MG?

To many the meaning of 'strong inclusion' will be plain from the term and the context; but an informal definition will render that meaning more explicit. We first quote, for a similar set of properties, an informal definition due to Chomsky and Miller (1963, p. 297): "Two grammars will

be called weakly equivalent if they generate the same language [set of strings - wcw]; they will be called strongly equivalent if they generate the same set of structural descriptions." Each such structural description is of a string: thus 'strong equivalence' is also expressible as holding for two grammars if they generate the same set of strings-with-structural-description. Let us, following Katz and Postal (1964, pp. 24-26), single out the notion 'string with one structural description'--that is, 'unambiguously derived string'--and let us term anything answering to that description "a sentence". (Thus our "sentence" is identical to Katz and Postal's "sentoid" [ibid.] .) So every "sentence" is unambiguous. (And what is called an "ambiguous sentence" in conventional terminology would for us be a "set of homonymous sentences".) So, finally, two grammars may (at first) be called weakly equivalent if they generate the same set of strings, strongly equivalent if they generate the same set of sentences. A language such as English, or such as the English subsets controlled by children, can be regarded as a set of strings (sentential surfaces) or as a set of sentences (strings with structural derivations). Since it is natural to apply Chomsky and Miller's terms to languages as well as to their grammars, we derive the statement: Two languages will be termed weakly equivalent if they contain the same strings (are stringwise identical), strongly equivalent if they contain the same sentences .

We adopt the abbreviation 'p-inclusion' for 'proper inclusion'. We now state the meaning of the term "strong inclusion": One language will be said to be weakly p-included in another language if all its strings are p-included, strongly p-included if all its sentences are p-included.

The strong and weak p-inclusion of grammars, as distinct from languages, is defined analogously.

The 'Strong Inclusion Hypothesis' (hereafter, SIH) can be simply described as holding: If a language is weakly p-included in another language, then it is also strongly p-included. Thus, according to the SIH, the set of wellformed strings generated by a child (whether or not the child's MG also generates some informed strings) is strongly p-included in the language generated by an adult's MG, hence (unless the APG hypothesis is adopted) is strongly p-included in the set of sentences constituting the English language. Or, in particular, since any one string constitutes a (one-member) set of the kind in question, every wellformed string produced by a child must be assumed to be assigned by that child's MG the structural description, or derivation, assigned that string by the adult's MG and hence, by the CG hypothesis, by the grammar of English.

2.1. As we see immediately, these notions of "strong" and "weak" p-inclusion are overly simplified, for there are several ways in which a set of sentences can be p-included in English in one way but not p-included in another, A sentence may exhibit:

- (i) Stringwise wellformedness. The sentence is weakly p-included in English.
- (ii) Surface-structural wellformedness. A sentence that is weakly p-included in English may be:
  - (a) A parroting;
  - (b) A quasi-holophrase;^ or
  - (c) A sentence with an orthodox constituent structure, which

<sup>50</sup> We define and discuss these terms in n.52, below.

may be:

(c') an incorrect surface-structure, or

(c'') the correct surface-structure, in which case the sentence is strongly, as well as weakly, p-included.

(iii) Deep-structural wellformedness. A sentence may have:

(d) No deep-structure--e.g., a parroting--; or

(e) A deep-structure, which may be:

(e') incorrect or

(e'') correct (the sentence is "deeply p-included").

(iv) Derivational wellformedness. If the sentence is of both surface-structural and deep-structural wellformedness--if it is both 'strongly' and 'deeply' p-included--then it will have transformationally derived its correct surface from its correct base through a transformational path which is:

(f') an aberrant path, or

(f'') the correct one: in which case we will say that the sentence is maximally p-included.

(NB that maximal p-inclusion entails strong and deep p-inclusion.)

The Strong Inclusion Hypothesis, stated too simply above, now breaks down into these Hypotheses:

SIH<sub>1</sub>: If a language (set of sentences) L<sub>1</sub> is stringwise (weakly) p-included in another language L<sub>2</sub> then L<sub>1</sub> is surface-structurally p-included in L<sub>2</sub> (L<sub>1</sub> is strongly p-included);



SIH<sub>2</sub>: If a language L<sub>1</sub> is surface-structurally p-included in L<sub>2</sub> then L<sub>1</sub> is deep-structurally ("deeply") p-included in L<sub>2</sub>;

SIH<sub>3</sub>: If a language L<sub>1</sub> is surface-structurally and deeply p-included in L<sub>2</sub> then L<sub>1</sub> is maximally p-included.

SIH<sub>4</sub>: If a language L<sub>1</sub> is weakly (stringwise) p-included in L<sub>2</sub> then L<sub>1</sub> is maximally (surface-and deep-structurally and derivationally) p-included.<sup>51</sup>

Clearly the SIH version of greatest utility, were it only valid, is SIH<sub>4</sub>, since SIH<sub>4</sub> hypothesizes that any sentence that has the appearance of being in every sense a correct English sentence, is one in every sense. IF SIH<sub>4</sub> were upheld then whenever a child emitted a stringwise-wellformed sentence he would necessarily have emitted a completely-wellformed English sentence; whereas if SIH<sub>4</sub> fell, but, say, the weaker SIH<sub>1</sub> were upheld, then all that could be decided, of a stringwise-wellformed utterance, would be that that utterance had the correct surface-structure.

2.2. The four Inclusion Hypotheses need only be stated to be put into doubt, as it seems to me, for it is perfectly clear that all of them can be disconfirmed by counterexamples.

2.2.1. On formal grounds:

(i) SIH<sub>1</sub> is disconfirmed because there are sentences stringwise-identical but surface-structurally dissimilar, so a sentence may be stringwise identical with some English string (hence weakly p-included in English), but not strongly identical (hence not strongly p-included). A "parroting" or a quasi-holo-phrase is such a string;<sup>52</sup> elsewhere, weak identity without strong identity produces one kind of ambiguity: e.g., "They are flying planes", whose surface structure is either something like "NP--are+flying--planes", or else something like "NP\_\_\_\_are\_\_\_\_flying + planes."

<sup>51</sup> For completeness, we ought to list such further inclusion hypotheses as

SIH<sub>5</sub>: If a language L<sub>1</sub> is stringwise (weakly) p-included in another language L<sub>2</sub> then L<sub>1</sub> is deeply p-included; and so on. But the range of plausible inclusion hypotheses seems to be covered by SIH<sub>1</sub>--SIH<sub>4</sub>, and so the others, though formally on a par with those, will be ignored here.

<sup>52</sup> A "parrotting" is a sentence emitted by a child in direct imitation of a sentence just heard; such an emission, which need signify nothing whatever about the child's MG (compare the adult's imitation or burlesque of a phrase in an unknown foreign language, as in e.g. "Have you seen Slava Domnulu's new opera Acest om nu stie nimica?") has long been recognized by students of developmental psycholinguistics as an obvious counter-example to any hypothesis like the one under view. (However, McNeill asserts [1966, pp. 68f.], on the basis of a finding of Ervin-Tripp [Ervin, 1964], that children may not imitate forms their MG's do not already generate, in which case of course "parrottings" would not fail of strong p-inclusion. For a contrary view cf. Turner and Rommetveit, 1967<sub>a</sub>, pp.654ff.)

A "quasi-holophrase" is a sentence or sentence-part that contains as an unanalyzed lump a string that in the English LG is broken up into a sequence of distinguished segments (Hiž, 1961, pp. 44f.)--essentially, of morphemes. A quasi-holophrase contains at least one distinguished segment, however, and so fails to be a complete holophrase: this seems to be true in the case cited by Brown and Hanlon (p.000): the "you" segment of "D'you...", as Brown and Hanlon point out, is certainly a distinguished segment.

For another instance, the Subject I call "Language Acquisition Device #6" made extensive use at the age of 3,1 of the form "Would you mind..." (doing, getting, putting on, &c). All of his sentences containing "would you mind..." were stringwise quite grammatical, and they were used under, and only under, the right circumstances. But his understanding of this phrase was incomplete. If in signifying compliance with his request one said "No", meaning "No, I wouldn't mind...", then he showed disappointment, for his grammar demanded "Yes" as the affirmative reply. Superficially, his MG could have contained either of two mistaken notions: (1) the word "mind" was misunderstood as its antonym, thus replacing an antonymous V able to occur, with the intended meaning, in the phrase "would you \_\_\_\_\_"

doing..."; or (2) the entire phrase "would you mind" was underparsed, with "mind", in all probability, going completely unanalyzed. Now, (1) insists that LAD #6 substitute "mind" for one of these verbs: "favor", "sanction", "indorse", "countenance", or the like; none of these V are in common use in LAD #6's household, and I think the likelihood of his having the antonym of "mind" as an inferred (or universal/innate) semantic concept, is remote. Then (2) is the more reasonable explanation, and "would you mind..." was a quasi-holophrase.

(The upshot of numerous disappointing responses to his use of the locution was that LAD #6 despaired of his respondents' obduracy, and at 3;3 relinquished the expression.)

To avoid possible confusion, it should be pointed out that the term "holophrase" has been used in two completely different (but related) senses. Commonly (e.g., here, and Lenneberg, 1953) the term is applied to cases where a string that could be analyzed (is analyzed in the LG) is apparently used as if a single morpheme: examples are "all right", "everybody", "scarecrow". The second use appears in McNeill (ibid.), who applies the term (pp. 63f.) to cases where a single word is used to stand for an entire phrase or sentence, as where the utterance "Milk" can mean, for a one-year-old, " 'I want some milk', 'The milk is on the floor', 'Don't give me any more milk; I want Pablum', etc."

(ii)  $SIH_2$  is disconfirmed because there are sentences surface-structurally identical but basally (Deep-Structurally) dissimilar: e.g., the two sentences of the form "The cops must stop drinking by midnight", which have surface structures something like "NP--must--stop--drinking--Adverb" but which have these two Deep Structures, approximately: "NP--must--stop--themselves--from--drinking--Adverb" or "NP--must--stop--(unspecified-human-NP) [e.g., "everybody"]--from--drinking--Adverb".

(iii)  $SIH_3$  is disconfirmed because it is certainly possible for a sentence to be structurally and deeply p-included in English, but yet not derivationally p-included: it would have the right surface and the right Deep structure but would have derived the latter from the former in the wrong way. Suppose that both in the LG of English and in some aberrant weakly p-included APG, the Deep Structure of the truncated Passive were something like:

$$[S [NP \Delta]_{NP} [VP^V NP]_{VP}]_S$$

And suppose that the surface in both LG and weakly p-included APG is NP---be+tense---Participle. Deep and surface structures are identical, but LG and APG can still differ by having different transformational paths. E.g., in this APG, Deletion + Special-Passivization: a Vacuous-Deep-Subject-Deletion transformation, which transforms the above Deep Structure to:

$$[S [VP^V NP]_{VP}]_S$$

---plus an obligatory Passivization transformation that would attack any structure of this form. As against, in the LG, Passivization + Deletion: a transformation that Passivizes normally (yielding a conventional "Gladys was deflowered by  $\Delta$ " sort of structure) and a transformation that then

deletes the "by+A" phrase.

In sum, on the basis considered, there is no warrant for any of the Strong Inclusion Hypotheses.

2.2.2. However, it might still be maintained that there is a specifically natural-language basis for one or more of the SIH: it might be maintained that one or another SIH applies whenever both including and included sentence-sets are "Languages" in the same sense. Of course the defender of any such thesis would be obliged to admit immediately that the notion "natural language" is utterly illdefined, necessarily so for the present; perhaps he could provide a rough idea of what he meant by stipulating that a "natural language" L must have a grammar G that is obviously "economical" in the "Competence" or LG sense, but even so this would be a necessary but not sufficient condition, since many sentence-sets, while "economically" describable, are too small to be a natural language: or they might lack attributes of which we are entirely ignorant. The only way of being sure that it is a "natural language" that is at issue is to provide the notion with sort of ostensive definition (Kotarbinska, 1960; Grice, 1968, p. 2.40) and then to restrict oneself to the appropriate ostensions (examples). So let

53

us take as our including that illdefined sentence-set "English"; and as our weakly p-included let us take what appears to be a dialect of English: that dialect in which transitive-verb sentences can be Passivized ("That '37 Hudson Terraplane was bought by a born sucker") but in which Passivizations of locative verbs ("That chair was sat in b^ Dr. Psoriasis as he chuckled over the misdeeds of the locative verbs") are either disallowed completely or else, no less exigently, consistently labeled as much less natural. English, we will assume, is the broader dialect in which

<sup>53</sup> That English is a language is not so self-explanatory an assumption as it might seem. It is easy to find statements to the effect that a natural language can include phenomena not amenable to the sort of generalization found generally in the grammar of that language (Harris, 1951, pp. 346f.; Chomsky, 1962, p. 543, n. 28; Valdman, 1968, p. 125). Even such large paradigms as the Benefactive-Object Passive might fall under this heading. Nevertheless, the existence of these exceptional phenomena does not affect our definition of the notion "natural language", since that definition is ostensive. (It must be noted, incidentally, that this ostensive definition is not a conventional one, in that, though we can point to many parts of the illustrative object--that is, English sentences--we cannot ever actually point to the object itself, the English language, except in the peculiar sense of pointing to a device (an MG, ideally) capable of enumerating all the parts of the object. But this appeal to the notion of ostensive definition seems strained to so slight an extent that, in view of its possible clarificatory function, I have let it remain.)

both Passivizations are altogether grammatical.

We assume that the grammar  $G$  of  $L_1$  and the grammar  $G$  of  $L_2$  are both highly-valued by (intuitively-understood) LG criteria of simplicity: this will mean no more, in actuality, than that neither  $G(L_1)$  nor  $G(L_2)$  will fail to realize obvious generalizations or to achieve obvious reductions in the number and/or complexity of rules.

Rather than trying to disconfirm one-by-one a "natural-language" defense of each of the SIH, let us see whether we cannot show, on "natural-language" grounds, that at least one of the SIH must be false.

We assume, in the absence of any discernible reason to the contrary, that  $L_1$  is weakly and structurally p-included in  $L_2$ , so that the Passivized transitives and the unpassivized locatives ("Dr. Psoriasis sat in that chair...") have the same respective surface-structures in  $L_1$  and  $L_2$ . Now we ask if from that fact and the fact of the status of  $L_1$  and  $L_2$  as "natural languages" we can infer (a) that they have the same Deep Structure, and (b) that their surfaces are derived via the same transformations. We make the elementary observation that an optimal  $G(L_2)$  will derive Passivized transitives and Passivized locatives via the same Passivizing transformation: that transformation, in  $G(L_2)$ , will be general enough to attack either sort of verb. But in  $G(L_1)$ , which cannot Passivize locatives, the application of the Passivizing transformation must be more restricted. Either in  $G(L_1)$  the transitives and locatives have the same Deep Structures as they do in  $G(L_2)$ , but the Passivizing transformation in  $G(L_1)$  refers specifically to the verb's transitiveness [a reference absent from  $G(L_2)$ ], or else in  $G(L_1)$  the Passivizing transformation has the same form as it does in  $G(L_2)$  but it is prevented from attacking the locatives because in  $G(L_1)$  the

locatives have a non-Passivizable Deep Structure [hence a Deep Structure different from the one they have in  $G(L_2)$ ]. Then  $G(L_1)$  either has a Passivizing transformation different from that of  $G(L_2)$  or has a locative Deep Structure different from that of  $G(L_2)$ ; and so either  $SIH_3$  or  $SIH_2$  has been disconfirmed (on "natural language" grounds), and in any case  $SIH_4$  has been disconfirmed.<sup>54</sup>

Since both  $SIH_4$  and either  $SIH_2$  or  $SIH_3$  are shown to have no "natural language" basis, clearly there is in general no natural language basis for the set of SIH and one is not encouraged to expect that a natural-language basis for one of the other SIH (say,  $SIH_1$ ) will be forthcoming.

2.3. We have disconfirmed all of the SIH on general counterexamples, and have disconfirmed some (and by extrapolation, all) of the SIH on specifically "natural-language" counterexamples. We now briefly take up the appropriate contrary hypotheses, discounting just the strongest of these -- 'If a natural language  $L_1$  is weakly p-included in a natural language  $L_2$  then  $L_1$  is NOT maximally p-included in  $L_2$ ' -- and letting the rest fall by inference. We will do this

by showing that an  $L_1$  can be a natural language (someone's dialect) and yet be maximally p-included in an  $L_2$  (another dialect). This in fact can be granted immediately on trivial examples: for example, English less one word ("inconcinuous") or some one aphorism ("The more the merrier") is, clearly, a natural language that could be the  $L_1$  of some speaker; and clearly this  $L_1$  could be maximally p-included in English. But English itself, standardly speaking, can be maximally p-included in a dialect that includes all of English plus some peculiar outgrowth of its own. General



<sup>54</sup> A further word on the locative verbs. It was maintained by Chomsky (1965, pp. 104-106, partly revising an earlier treatment) that verbs like these, though "intransitive", when co-occurring with a V-Complement (but not with a VP-Complement), may be Passivized by the ordinary Passive transformation, because in Chomsky's 1965 formulation the Passivization transformation specifies, not that the verb be transitive with a Direct Object in its Complement, but only that the verb be followed, in its Complement, both by an NP and by a Manner Adverbial instantiated as 'by+passive'. This rule, which has V-Locative passivization as a desirable by-product, was created to account "...automatically", in Chomsky's words, "for the restriction of passivization to verbs that take Manner Adverbials freely". The underlying assumption was, apparently, that the "by NP" phrase of the Passive is a Manner Adverbial, so that if a V could take such an Adverbial in the Passive, it ought to be able to take other such Adverbials in the Active.

Lakoff, however, has shown (1965, Appendix, F, pp. 1-3) that there is a class of verbs--the 'Stative' verbs like "know", "believe", "see", and "hear"--that do not occur "freely" with ordinary Manner Adverbials but that, nevertheless, freely passivize, as in "That Viewvathit is Luxembourg's greatest composer is believed by all Andorrans". We note that also the 'Locative' verbs do not occur freely with Manner Adverbials. In fact, "sleep in" seems to me to occur no more freely with such Adverbials than does "sleep" itself, which was for a long time used by Chomsky, in the familiar "Colorless green ideas sleep furiously", to illustrate (among other like facts that such verbs do not co-occur with such Adverbials. Nor can either Lakoff's observation on the 'Statives' nor mine on the 'Locatives' be palliated by Chomsky's statement (ibid., p. 218, n. 28) that "...the generalization that relates Manner Adverbials to passivization...[is not] invalidated by the fact that certain items must be listed, in the lexicon, as conflicting with this generalization...". For in both cases it appears that the verbs in question are, as a class, unable to occur "freely" with Manner Adverbials. Thus to list in the lexicon, individually, each such verb as being "deficient" in this regard, would be (exactly in the sense endorsed by Chomsky) to miss a generalization.

In addition, Lakoff has also shown (ibid., passim) that Adverbs do not occur as such in the Deep Structure at all, typically, hence cannot be available in the Deep Structure in the way Chomsky described them (though they might still be correctly sited through prior transformation to cue Passivization). However, this fact is almost irrelevant in view of the facts presented above and their inescapable consequence: if the Passive's "by NP" phrase is indeed a Manner Adverbial, it is of an entirely different type from the ordinary adverbs like "furiously", "mechanically", "briskly", or the like. So cueing Passivization by the presence of a Manner Adverbial is not a good idea in any case.

This leaves the account of V-Locative Passivization still open. Lakoff and Ross (1966, p. 7) have suggested that perhaps e.g. "remain in" is a transitive verb, taking NP as its Object. However, a Passive sentence with "remain in" had earlier been labeled by Lakoff (ibid., p. 13) as ungrammatical, so that it is not clear that this Lakoff-Ross suggestion is meant to provide a means of V-Locative Passivization.

American English has sentences like (1) "Your transmission needs fixing" and like (2) "Your transmission needs to be fixed", but none like (3) "Your transmission needs fixed". But (3) is common in the Pittsburgh and general Western Pennsylvania dialect, and some Pittsburgh speakers (of an "inclusive" dialect) have all three sentence-types, seemingly with equal grammaticality. Their G differs from English, presumably, not in surrogating any English rule or structure with another rule or structure, but only in supplementing the English G with an optional rule for deleting "to be" from (2); hence the  $L_1$  English is maximally p-included in the  $L_2$  "inclusive" Pittsburgh dialect, and the contrary hypotheses are overturned.

2.4. As we have seen with the quasi-holophrases and perhaps the "parrot-ings", it has always been clear that in at least these few aberrant cases a child's superficially-wellformed sentence might, covertly, be not wellformed at all: that a sentence might be weakly p-included in English but not p-included in any deeper way. (This limited observation holds equally, of course, for LG and APG alike.) But, as we have also seen, the general case for the deeper p-inclusions has no a priori warrant whatever. This fact seems only twice to have been brought out into the open, but even so there is manifest in the psycholinguistic literature a growing (if tacit) awareness of the dubiety of the deeper inclusion hypotheses as they apply to children's language. Thus, for example, Brown and Fraser (1964) specifically defend a strong inclusion hypothesis (p. 71); but Brown and Hanlon (this volume, p.000) take pains to show that the (LG) analyses that they assign children's stringwise-wellformed sentences are defensible on the basis of a demonstrated consistency with the rest of

the children's language. Thus the present discussion does not appear in a climate of wholesale (but tacit) acceptance of the inclusion hypotheses (and, as we mentioned at the outset, it is not our main purpose to overturn them, though clarifying the degree to which they are in doubt is a necessary step toward our goal of exhibiting the consequences of the hypotheses' downfall).

As was said just above, the inclusion hypotheses (to generalize them, vaguely, as a set) have been specifically discounted twice in the literature, though without being specifically formulated: I refer to the rejections of Chomsky and McNeill. Chomsky, for his part, has made comments (1965, p. 202, n. 19; 1967a, pp. 86f.) to the effect that a child's superficially-wellformed sentences, in the early stages of acquisition, need not be underlain by the same Deep Structures as underlie them in the speech of an adult: that is, to particularize, Chomsky has in effect denied  $SIH_2$ , hence  $SIH_4$  as well. McNeill (1966, pp. 55f.) reaffirms the rejection of Hypotheses 2 and 4, pointing out that such a child-sentence as "I don't see you", for example, probably does not have the transformational or Deep Structure history that it has in the LG (or in the adult MG). (For future ease of reference, let us tag this the "Chomsky-McNeill Null Hypothesis".) McNeill also (in effect) states another null hypothesis which counters the  $SIH$  from another (and, on the surface, contradictory) direction: he holds that at the earlier stages a child's sentence (whether or not superficially wellformed) is generally a pronunciation of the Deep Structure: the result of applying the phonological rules directly to the sentence's Deep Structure with no intervention of transformations and so no (distinct) surface structure at all (ibid., pp. 54-65). If McNeill's Null Hypothesis should prove valid then

at this early stage when such a Deep-Structure pronunciation resulted in a stringwise-wellformed sentence,  $SIH_2$  alone would be upheld, though vacuously, accompanied by the fall of  $SIH_3$ ,  $SIH_4$ , and (presumably) even  $SIH_1$ . Since in fact McNeill's Null Hypothesis seems quite plausible, having been ably defended by its author, and since by allowing Deep Structures an ontogenetic development we can make the two null hypotheses mutually compatible, it seems that the literature already contains statements interpretable as holding that at one stage or other a child's speech disconfirms every one of the inclusion hypotheses, without exception.

2.5. If however the strong inclusion hypotheses fail in the case where all of a child's utterances (when normalized of course to MG-sentences) are stringwise wellformed, how much greater and more certain must be their failure when only some of a child's utterances have wellformed surfaces. For in such a case the whole of the child's "Language" cannot possibly be p-included (even weakly) in English; all "natural-language" basis for an individual sentence's deeper inclusion is therefore forfeit, and the formal bases of the inclusion hypotheses can now be dismissed out of hand.<sup>55</sup>

2.6. Above, we have seen that the set of stronger inclusion hypotheses fails on both general [2.2.1.] and natural-language [2.2.2.] grounds, so that in a young child's speech a superficially wellformed sentence may be

I assume that no one could adhere to the notion that a child's partial set of stringwise-wellformed sentences forms a Language in its own right, hence a set of sentences subject to whatever tenuous support the "natural-language" basis can lend to the inclusion hypotheses. To assume this one would have to assume that the child's language in fact consists of two sublanguages, and , containing the wellformed and illformed sentences respectively, and that the child's MG in fact consists of two subgrammars,  $G(-fL^{\wedge})$  and  $G(-L^{\wedge})$  respectively. But this notion is, clearly, preposterous. If a child says "I fixed up it" it is not because he has a special grammar,  $G(*L^{\wedge})$ , devoted to generations of this sort, but because he has a defective  $G(L^{\wedge})$ , one which incorrectly fails to distinguish those NP that can follow Verb-tflemora (e.g. "fixed+up") from those—e.g. "it"—that cannot. (If it should be maintained that a child in his very earliest stages generates each sentence by a rule idiosyncratic to that one sentence, thus in a way having as many (one-rule) "grammars" as he can generate sentences, we could still avoid the quibble over whether or not such devices are indeed "grammars" by pointing out that we are not concerned with such cases in the present instance, since at so early a stage the child will scarcely emit any stringwise-wellformed sentences at all.)

assigned by the child's MG, covertly, a deeper analysis completely at variance with that assigned it by the English LG and/or the adult MG. But we have also seen that even in the speech of a child who controls rather little of English it cannot be assumed that every sentence will be covertly misanalyzed, since though the stronger inclusion hypotheses fail so also [2.3.] do the corresponding contrary hypotheses. We have noted [2.4.] that failure of stronger p-inclusion has long been acknowledged in the case of the distinctively childish quasi-holophrases and, though now doubtfully, in the case of "parroting"; and we have noted that, perhaps partly due to what I have termed the Chomsky-McNeill and McNeill null hypotheses, faith in the stronger p-inclusion hypotheses is now on the wane. Finally, we noted [2.5.] that if some covert misanalyses could be expected even when all of a child's sentences were superficially wellformed (weakly p-included), this could be expected all the more, of a superficially-wellformed sentence, when the child's set of generated weakly p-included sentences comprised only a subset of the set of his generated sentences.

We have thus covered one way in which a child's generations may be misleading; but there is, of course, another side of the coin: certainly the young child generates sentences which are not even weakly p-included in English but which are nonetheless maximally p-included in the child's own MG. (And, to generalize and 'idealize', in the MG of the English-speaking child at the stage in question.) We have already noted one such case: for when the child generates "I fixed up it" he generates a sentence which is not weakly p-included in English but which is, by definition, maximally p-included in the child's MG; and any sentence generated by the child's MG yet not (at some level) identically generated by the English LG,

falls in a similar class. In fact, the case where a child's generation is weakly p-included in English but not p-included in English at some deeper level (thus violating one or another of the stronger inclusion hypotheses) is only a special case, as we see, of a more general phenomenon found in the speech of the younger child: generations that are maximally p-included in the child's MG but not maximally p-included in the English LG.<sup>56</sup> Thus while generations from the child's MG that are stringwise-wellformed may be misleading in one way (inducing the unwary to assume them wellformed on all levels), on the other hand the child's generations that are not stringwise-wellformed are misleading in a quite different way, since their utter lack of superficial wellformedness with respect to English disguises a total wellformedness with respect to the child's own MG.

Naturally this sort of deceptiveness is (in the context of the discussion thus far) quite without importance, since "wellformedness with respect to the child's own MG" can be simply assumed, just because the sentence was in fact generated by the child: that is, it is maximally p-included in the child's MG by definition. A veneer of LG stringwise-wellformedness may be a trifle misleading because from it one might infer LG maximal wellformedness; but if a sentence is stringwise illformed then it is also surface-structurally and transformationally illformed and it is probably also deep-structurally illformed: its overt illformedness cannot hide a thoroughgoing covert wellformedness, and since its maximal p-inclusion in the child's MG is "guaranteed" anyway, (as see just above), its LG-illformedness cannot deceive us on this score either. Nothing about a sentence can lead us falsely to infer that it was generated by the child's MG, in



<sup>56</sup> Our use of the phrase "maximally p-included in the child's MG" must of course include the vacuous cases when, for example, a "maximally p-included" sentence has no Deep Structure (as distinct from surface structure) at all.

the present context, since it either was or wasn't; superficial illformedness is always underlain by at least partial deeper illformedness; and so the most deceptive cases are still those we considered first, those where superficial LG-wellformedness in child-productions covers their deeper LG-illformedness.

We want now to change our angle of attack slightly, and to do this with greater clarity we want to condense to its rudiments that part of the preceding discussion that concerns p-inclusion in LG and in the child's MG. We see that a given string, surface-structure, deep-structure, or transformational path can be wellformed in the English LG or the Child's MG-- let us tag these conditions "E" and "C" respectively--or illformed in one or the other: "\*E" or "\*C", let us say. Joining these conditions, we say that a string (or surface-structure, etc.) may be wellformed in both LG and Child's MG, in one or the other, or in neither: that is, may have the set of properties (E, C); (E,\*C) or (\*E, C); or (\*E, \*C). The sentence-sets thus defined are represented diagrammatically in Figure 5.

(Put Figure 5 about here)

One's reading of the diagram is much simplified, of course, if one considers string, surface- and deep-surface, and transformational path all together, ignoring the fact that a given sentence may be, say, (\*E, C) as a string but (E, C) as a deep-structure. That is, we can concentrate on maximal p-inclusion alone, and then "a sentence"--at every level--would be uniquely assigned to one of the four sentence-sets of Figure 5. In this case, where we assume that covert illformedness is detected and tagged as \*E and where emission by a child assures the tag C, assignment of "a sentence" to one of the four sets is quite mechanical.

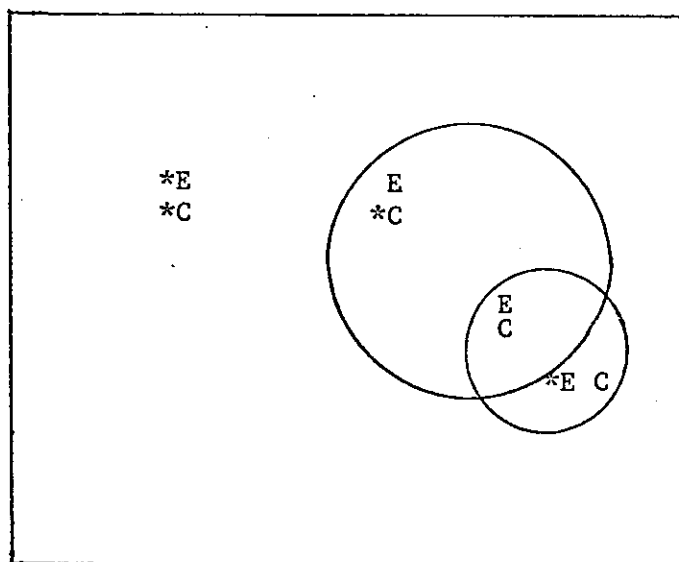


Figure 5

2.7. It is hardly necessary to add, however, that in actual practice the assignment of a child's sentence to either (E, C) or (\*E, C) is not mechanical at all; any such assignment, as we have noted, presupposes that the sentence has been generated by (and so is maximally p-included in) his MG, and of course this presupposition has no a priori warrant whatsoever. A child's having produced a sentence is no assurance that his MG generates it, since between what his MG generates and what he produced the child's performative mechanism may have intervened in such a way as to distort the generation, producing an utterance not identical with the underlying MG-generation and so misrepresentative of it. In fact, that such distortions are common occurrences is well recognized among developmental psycholinguists. It is well known that the young child's ability to comprehend surpasses his ability to produce (his grasp exceeds his reach), and the most natural way of explaining this fact is to assume that his MG generates all that he can comprehend, but that in trying to access that MG for composition his performative limitations intervene and limit his output (Fraser, Bellugi, and Brown, 1963; McNeill, 1966, pp. 76-82; and see section [1.2.2.3.], above). One example of the consequence of such limitations is, possibly, the so-called "telegraphic speech" to be observed in the speech of young children: as Chomsky has conjectured (1967a, p.88) it may be that "telegraphic" sentences (they omit material somewhat in the manner of a telegram) might owe their deviant surface form partly to the intervention of a defective performative mechanism which (chiefly because of specific memory limitations) is unable to pass through to the surface the full complexity of what the mechanism can compose by accessing the MG, with the result that the performative mechanism acts as a "filtering

device" that "operates on deep structures in some non-normal fashion" so as to produce the "telegraphic" result.<sup>57</sup>

Even if there were no plausibility to this argument at all, however, one would manifestly have no reason to assume that the speech-productions of young children are completely undistorted replicas of compositions based on the child's MG. Such, certainly, is not the case with adults, whose speech is full of errors and distortions of all kinds (as see n. 3, above); there is scarcely a sentence in an adult's casual speech that is not refracted by performative factors, down to and including cases where, plainly, the speaker has completely lost track of where in the sentence he is, and either trails off or veers in the wrong direction.<sup>58</sup> It is entirely reasonable to expect that children also have MG whose generations, when accessed by the compositional performative device and then performed as speech, become garbled: and probably in the earlier stages, as plausibly in the case of "telegraphic speech", garbled in ways peculiar to the speech of children. What this means is that some of the sentences that a child produces are not generated by his MG at all: they are only refractions of such generations, and only the (recovered) generations themselves are properly labeled C; the refractions are properly left outside the (C, \*C) set of categories altogether.

Thus, where in preceding sections we observed that a child-produced sentence might have covert irregularities, being therefore covertly (\*E) rather than (E), now we see that a child-produced sentence may covertly be (\*C) rather than (C).

Moreover, this potential of a covert \*C status obtains both for \*E and for E items: both for outputs that are not English and for outputs

McNeill however argues at one point, though a bit ambiguously, that "telegraphic speech" results, not from the interposition of a defective performative mechanism, but entirely from defects in the accessed MG (ibid., pp. 18f.). It would, of course, be quite possible for a defective MG and a defective performance mechanism to work in consort, yielding jointly deformities which neither would yield alone,

58

The very common "whom...is" ("whom" as Subject) sentences are of this sort: e.g. [*italics mine*], "Can a 40-year-old double divorcee find love and happiness married to a 22-year-old boy whom everyone thinks is interested in her daughter?"—from an inquiry in Playboy, April 1969, (Jj), No. 4, p. 56.

that are. For we have no real reason to suppose that the interposition of performative error results inevitably in the production of non-English- of \*E sentences--since it is in theory perfectly possible for performative error to produce the appearance of a full-fledged English sentence. And this is possible, moreover, whether the MG's generation was (E, C) or (\*E, C), since performative error could in theory either distort one well-formed English sentence into the outward form of another--as in the first case--or, as in the second case, distort an illformed (\*E, C) sentence into the outward form of some wellformed sentence--whether (E, C) or, quite deceptively, (E, \*C). Neither possibility is in the least far-fetched. Certainly cases of the performative distortion of one sentence into another sentence (from the MG-generated sentence into a sentence other than the one intended) can actually be observed in the productions of adults. Thus we find this sentence addressed to "Dear Abby": "I would like your opinion, which I respect highly."<sup>59</sup> It is of course clear that the opinion the writer would like, and the opinion the writer already respects highly, cannot be the same opinion: the writer has constructed what we may call a "portmanteau sentence" by mistakenly merging two different (and differently-indexed) occurrences of the word "opinion."<sup>60</sup> But the result of the confusion is a sentence, nonetheless; or at least the appearance of one. (What the Deep Structure of this product might be is an open question.)<sup>61</sup> But then if adults can hit upon the semblance of a sentence through error, so can children. And, as to the second case, where an illformed Deep Structure can be distorted into a wellformed sentence: this surely can happen also.<sup>62</sup> Thus on both counts it is to be expected that children, even more than adults, produce a set of sentences having in part only the most deceptive

<sup>59</sup> The Pittsburgh Press, October 10, 1968, p.40. The comma is mine.

<sup>60</sup> Note also that, to be correct, the first "opinion" should be a count noun, the second a mass noun.

<sup>61</sup> We might suppose, however, that the Deep-Structure is quite orthodox, bearing the correct complex symbols for both nouns associated with the D "opinion", and distinctively indexing them. But the rest of this distorted composition--the nature of its transformations and of its surface structure--is completely opaque.

<sup>62</sup> For example, a mistake near the surface might correct a deeper error of the sort that gives a plural verb to a singular noun: this kind of "corrective" error would seem reasonable to expect on the basis of the similar errors in the other direction, e.g. the error that from a correct deeper structure derives the incorrect "They think having a million dollars make them qualified for public office".



kind of relationship to the generations of the MG. And since these problems can arise unpredictably with any child-emitted sentence, they must be assumed possible with every such sentence; and so every sentence, if one is to be sure of its analysis and assignment, must be scrutinized rather closely.

We will take a brief look at such scrutiny in a moment; but first we should summarize all of the foregoing discussion as succinctly as possible. Pretending (for clarity's sake only) that the primary problem in treating a child's sentence is its proper categorization, we present such a summary in the decision-diagram of Figure 6.

(Put Figure 6 about here)

It is, I think, obvious that the diagram of Figure 6 has two main divisions: that below the conclusion "C AT ALL LEVELS", in which all of the decisions about E and \*E at all levels are perfectly straightforward; and that above the cited conclusion, where the decision as to C, \*C, or "no assignment" is made. The latter division rather clearly constitutes a Gordian knot which can either be cut--by making the simplifying assumption that of course what a child says is generated by his MG--or else unraveled, through testing the sentence under examination to ascertain, as best one can at least, whether or not a judgment as to MG-generation can be made. As we see, the most difficult and delicate decisions, almost undoubtedly, are those made at the outset of the diagram in answering the question: "Generated by the child's MG?". Though this statement does not blunt the manifest difficulty of deciding e.g. whether or not the sentence, if stringwise and surface-structurally wellformed, is also deep-structurally and transformationally wellformed, hence maximally p-included. A question, again, better settled non-arbitrarily.

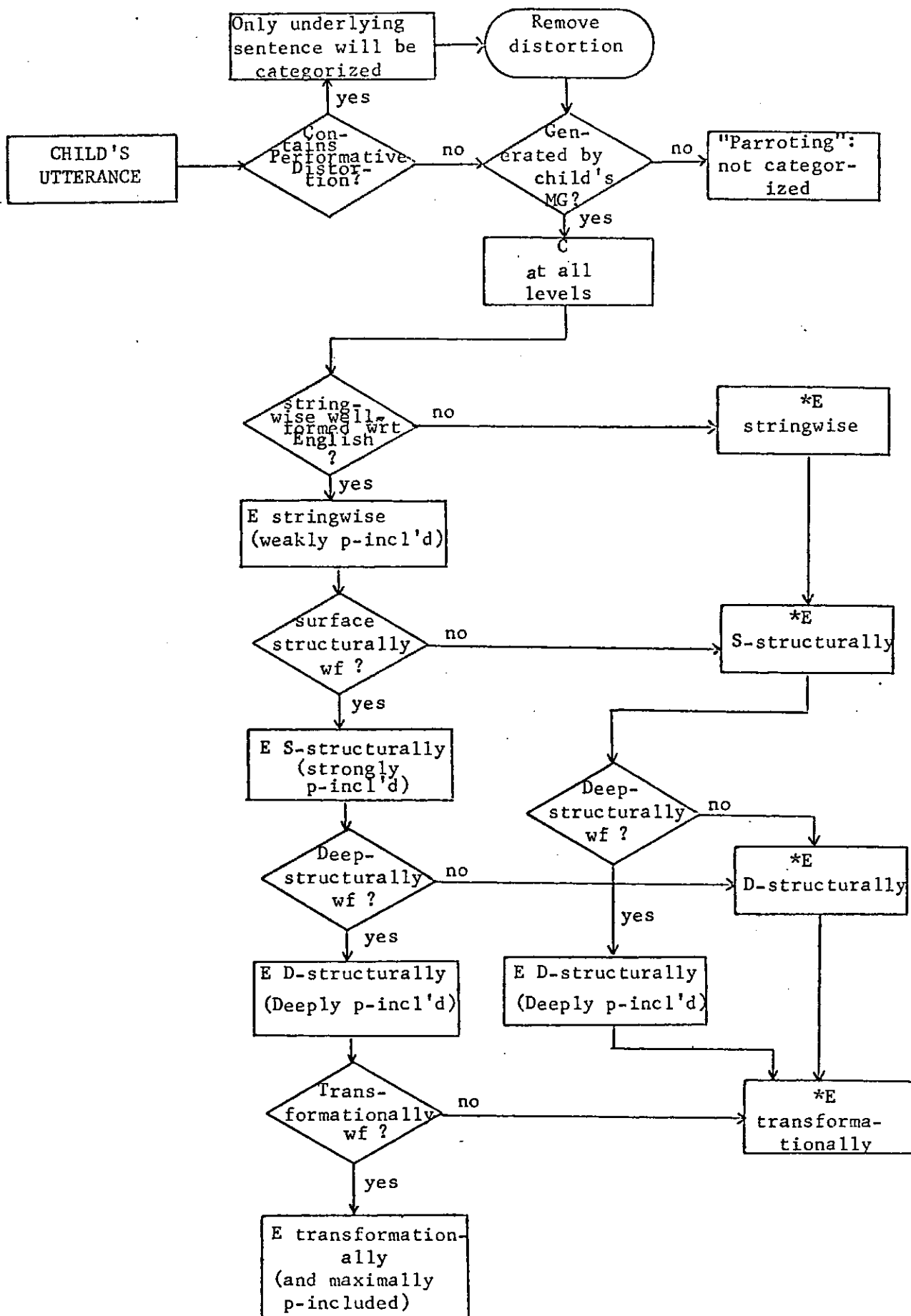


Figure 6. Decision-Diagram for Child-Produced Sentences.

2.8. We have gotten far enough to have outlined some of the main problems involved in analyzing and categorizing child-uttered sentences and to have seen, in brief, that there is great need for methods here: heuristics for aiding the determination of C or \*C and of the wellformedness of sentential levels hidden from the ear. The developmental psycholinguist is willy-nilly a "field-worker" in the old-fashioned sense of working systematically through a corpus of data and returning to his subjects to elicit further data. The introspection that is helpful or even essential (Postal, 1966a, pp. 92f.) for much of modern generative work is almost totally lacking when an adult analyzes the language of the child, since the adult no longer commands the language that the child commands and his knowledge of sentencehood and of meaning and connection of meaning--generally, his intuitions--are not serviceable to him; he is not a fluent speaker of the language under analysis.<sup>63</sup> And so the greatest single present need in this field is for a Methods in Psycholinguistics. The need for such a manual is all the greater in view of the fact that young children (whose language is most distant from our intuitions) are unusually recalcitrant interviewees. At the risk of repeating it once too often we may cite here the engaging dialog reported by Brown and Bellugi (1964, p. 135):

Interviewer: Adam, which is right, 'two shoes' or 'two shoe'?

Adam: Pop goes the weasel!

It will not however surprise the reader to find that while we have come far enough to delineate this problem, we have no panaceas at hand. There is no manual and we cannot compile one now. Still, there are a few useful heuristics available, and I should like in drawing

<sup>63</sup> There are perhaps limited exceptions to this generalization. For example, McNeill (1966, pp. 37f.) reports that adults presented with two (paraphrastic) illformed childish sentences can judge which is further from English (hence, by implication, earlier-learned) in about four out of five cases.

toward a close to invite attention to one or two of these in particular. They are to be found in Brown and Hanlon (this volume, pp. 000-000); these scholars have been unusually punctilious in giving their reasons for settling on specific analyses of child-produced sentences. (1) One Brown-Hanlon heuristic was used to determine whether or not a child who said "Me and Diandros are working" had an MG in which "me and Diandros" was (a) the superficial Subject and (b) plural in number. As they note, both (a) and (b) are supported by the "we" of the related produced sentence, "Me and Diandros are working, aren't we?"; the use of "we" seems to refer to "me and Diandros" as one NP; the NP of the tag e.g. "...aren't we" commonly refers directly to the superficial Subject of the antecedent Declarative ("Me and Diandros are working...") to which the tag is attached; and "we" is of course plural and inclusive of the speaker. (2) A similar heuristic was used to determine what the MG under study had to say about the Verb Phrase and Auxiliary: since "...are working..." was tagged with "aren't..." and "...made..." was tagged with "didn't...", Brown and Hanlon rightly infer that the MG that generates these correct tags probably embodies knowledge approximately to the effect that "n't" is attached to the first segment of the Auxiliary, which is "are" in the first example and "did" in the second; and they further infer that the MG must know that the "do + tense" auxiliary corresponds to a " $\emptyset$ " auxiliary accompanying "make + tense" in the sentence e.g. "I made a mistake, didn't I?"

These are, to reiterate, heuristics and not what Chomsky has called "discovery procedures" (1957, pp. 50-56): they are not litmus-paper tests to be applied mechanically. Thus, for example, the first heuristic will will not always work because the "we" of an "aren't we?" tag need not

refer to the superficial subject of the antecedant Declarative or indeed to a single NP in the underlying Deep Structure, as cf. "I'm working and Diandros is working, aren't we?" And the second heuristic involves appeal to a simplified notion of how tags relate to their antecedants--a notion which may be correct for fledgling MG's but which has not yet been shown to be so--for it is not the case that the form of the tag need have any simple relationship to the surface-form of the antecedant, as witness e.g. "Let's go to the movies, shall we?". But I think Brown and Hanlon would readily agree that these exceptions do not invalidate their heuristics: they only show areas of weakness that bear closer attention. It is in the nature of heuristics to fail once in a while, and it is no criticism of Brown and Hanlon's suggested heuristics to observe that they are obviously meant

64

only to permit rough analyses.

Thus, while the thoroughgoing uncertainties of dealing with the linguistic productions of children urge the development and formalization of as much method as possible, on the other hand what we could list at present would be only a small set of rather thin heuristics,^ and it follows that pending availability of deeper theoretical understanding and, equally, completion of the suggested Methods text, the uncertainties discussed above will continue in full force.

2.9. In conclusion, we have in this Chapter on the stronger inclusion hypotheses established that none of these hypotheses has either formal or natural-language warrant, and we have observed [2.6.] that, in addition, a child's sentence can be illformed with respect to the English LG at any level but still be, at the same level, wellformed with respect to the

<sup>64</sup> Their own description of their set of heuristics shows, I think, the proper mixture of confidence and caution: "We propose to treat the child's production of a construction, in quantity and over a good part of its proper range, together with evidence that he understands the construction when others use it, as signs [italics theirs] that he has the grammatical knowledge represented formally by the derivation of that construction in the adult grammar [the adult MG - wcw]."

<sup>65</sup> It must be noted that the heuristics in question are 'thin' in a respect other than that discussed above - their not being foolproof - for their already-imperfect reliability declines steadily in proportion as contact with a sentence's surface is less informative about underlying structures. Thus compositional command over a sentence and its related sentences, together with details observable on the surface, suffice to reassure the analyst that the sentence is stringwise and surface-structurally wellformed (or illformed); but 'understanding' is impossible to gauge with much accuracy when one cannot, as one can with adults, elicit paraphrase-sets. Behavior is but an uncertain indication of understanding, as may be seen in the fact that two utterly-dissimilar sentences can effect the same action: e.g., "Please pass the ketchup" and "If you don't pass the ketchup I'll poison your coffee". (For comments on this problem in an experimental situation where it is much reduced, see Shipley, Smith, and Gleitman, 1967.) So a child's Deep Structures are more inaccessible than his surface-structures; and his transformations (which as we have seen can be LG-illformed independently of both deep and surface structures) are less accessible still. With all of this I am sure Brown and Hanlon would agree, and my aim in mentioning these points has, again, only been to elucidate further the uncertainty they voice.

[Postscript: since completing this paper I have learned of the availability of what seems to be the 'Methods' text called for above: Slobin, D.I. (Ed.) A Field Manual for Cross-Cultural Study of the Acquisition of Communicative Competence, obtainable from the ASUC Store, University of California, Berkeley.]

child's own MG. The dissociations between E and C, \*E and \*C, we exhibited graphically in Figure 5. We have made special note [2.7.] of the distinctive ways in which children's performances can disguise the true nature of their MG's generative capacity, to such an extent that, predictively, it is as great a problem to determine whether or not a given production betokens MG-generation of the string in question, as it is, given that the sentence is at least stringwise wellformed with respect to the child's MG, to determine what structure the MG assigns that sentence at the levels beneath the surface. This discussion was summed up in the 'decision diagram' of Figure 6. Having seen the logical consequences of the fall of the inclusion hypotheses together with the nature of human linguistic performance (that of children in particular), we passed [2.8.] to the practical consequences: namely, the need for more and better discovery heuristics to aid in analyzing these refractory data.

This brings us to the end of our two tasks of elucidating the unspoken assumptions which, so I have argued, underlie contemporary developmental psycholinguistics to a greater or lesser degree, and of setting forth the most essential details of what follows when those assumptions, revealed as groundless, are withdrawn. It remains to say that our treatment of these problems has not been exhaustive, and certainly there are other problems in this area that we have not even touched on;<sup>66</sup> here, as in linguistics as a whole, no study much reduces the amount of work yet to be done.

### 3: General Conclusion

In the two preceding chapters we have treated as separate problems the relationship between the adult MG and the LG of English [1.] and the



<sup>66</sup> As examples of questions that must eventually be taken up but which we have not touched on here we might cite these three: (1) To what extent can a child's speech (MG) be influenced by an idiosyncratic family or parental dialect, if there is one, if such a dialect reduces the generality of the rules of English (by introducing special exceptions) and so impedes generalization? (2) To what extent can childish language-play like that reported by Weir (1962) become conscious with the child, producing, if not poetry, at least conscious linguistic play like that reported by Stene (1934) for adults? (I have never been able to find out whether the happy coinage "porculant"--said of the author by Language Acquisition Device #5 at 5,2--was the product of accident or design.) And (3) to what extent is it accurate to think of a child's linguistic development as passing from one MG to another (larger and better) MG, and thence to another, rather than as passing from one MG to a quasi-grammatical (hence unstable?) state, thence to a new MG?

relationship between both LG and adult MG on the one hand and the child's MG on the other [2.]. Now, in conclusion, we pause very briefly to knit these two strands together. We see that we have been studying but two aspects of a single problem: the nature (broadly considered) of the human linguistic capacity--both MG and archival linguistic faculty--as this relates to the LG of English. In the first chapter we treated mainly of adult competence and performance; in the second, mainly those of the child: but of course the MG that the child forms is the MG that he will have as an adult, and the two competences and performances are concomitantly close. If performative criteria of economy play a role in shaping the adult APG, as we have hypothesized, then they play that role by shaping the way in which the child's developing APG takes form. Indeed, there is no other way in which such criteria could have any influence.

This said, it seems proper to examine one sort of doubt concerning the most basic characteristic assumed of the adult MG: its status as a grammar. We know that the English LG must include some marginal peculiarities (as see n. 53, above); but some MG's clearly betray peculiarities that are more central. This fact alone is not astonishing; there is no reason to assume that the gradual process of improving one's grammar, observable in the child, terminates in a grammar amenable to no further improvement. (No one has ever claimed that it does: in fact, I know of no one who has addressed himself to this question, with the partial exception of Halle, as see just below). But, while suboptimality does not run counter to what we 'know' about acquisition, they might seem to run counter to what we 'know' about grammars as such, depending on how extensively we expect grammars to obey competence criteria of economy. Let us, then, inspect a particular MG suboptimality and try to estimate its adverse effect on the MG's status as a grammar.

In fact, examples of general adult suboptimality are exceedingly easy to find, and we have one ready to hand. In (1.2.2.1.), above, we considered the case of the Benefactive Objects of English, after Lakoff and Ross (1966); we observed that while in the dialect they analyzed the Benefactive Object must not be within the Verb Phrase, in my own MG the Benefactive Object, just as clearly, is within the Verb Phrase. The criteria for establishing the structural locus of the Benefactive Object were two: the Object is in the Verb Phrase if it can become the superficial Subject after passivization and cannot be referred to in a "do so" construction; it is outside the Verb Phrase if it cannot become the Passive's Subject and can be referred to in a "do so" construction. As I made clear in the initial presentation, the facts in the matter seem quite unequivocal: in my own speech I can freely form Benefactive-Object passives with any number of different verbs ("buy", "obtain", "procure", "steal", "find", and many others); and on the other hand the "do so" construction with the Benefactive Object is to me utterly ungrammatical. (To me it has roughly the befuddling quality of Ross's well-known "Relativization in extraposed clauses [is] a problem which evidence is presented that help is needed to solve" [1966, p. 1].)<sup>67</sup> Presumably the speakers on whom Lakoff and Ross based their contrary analysis were just as fixed in their MG. In short, given the facts, two rather different MG's must be postulated. Of course both can be considered to be, for its language, of optimum simplicity or economy.

But there are speakers who readily accept "John was bought a new Cord by his doting father" yet who cannot accept such a Passive with any

<sup>67</sup> Naturally this is not the whole story. While it seems that any transitive V with Benefactive Object can, in my dialect, spawn a Benefactive-Object Passive, nonetheless I would certainly avoid such a generation for any V that could take both a "to" Indirect Object and a "for" Benefactive Object, since in the case of the resultant ambiguity my interpretation would so favor IO as to eclipse BO, unless something else in the sentence were disambiguating. Thus, "John peddled an old Chevy for me last week" and "John peddled an old Chevy to me last week" can both, according to my MG, be Passivized to "I was peddled an old Chevy by John last week", but my interpretation of this Passive would heavily favor an IO reading (the second given). It is, I suppose, possible that I would disambiguate in favor of BO rather than of IO if the remainder of the sentence leaned that way, as in e.g. "I was written a fine letter by my amanuensis last week". As is plain, the matter is far from exhausted.

other verb: not with "obtain", not with "procure", not with "steal" or "find". Such speakers, it is my experience, do find acceptable the to-me forbidden "John's father bought a new Cord for John in 1938 and will do so for Gertrude next Saturday". That is, they can passivize on the Benefactive Object with the commonest verb "buy" (perhaps with some others that I have failed to uncover), but this ability has not affected their inability to form other Passives on that sort of Object or their ability to "do so" with such Objects. Then, clearly, their MG's are inconsistent: they have a special rule for the (otherwise-undistinguished) verb "buy". This being so, it seems entirely appropriate to say that, by any likely criteria of economy, these MG's will be suboptimal, and this in a rather non-marginal case: formation of the Passive.

How does this happen? Quite speculatively, we might conjecture that while there is indeed a general tendency, as one acquires a language, to optimize the MG one is building,<sup>68</sup> on the other hand the acquisition process does not go on indefinitely, and it need not stop at just that point when every aspect of the grammar is maximally economical. We have no particular reason to suppose that every rule, however late, is absorbed all at once for all possible cases, and so it is quite conceivable that some late rules are left, when the grammar has stabilized, in an arrested state of development. Such, if accession of the Benefactive-Object rules is late, might be the case with the "inconsistent" MG's just mentioned: once started with "buy" they should have moved all the way to my dialect or else dropped "buy"-Passives for Benefactive Objects and receded to the dialect analyzed by Lakoff and Ross. But, it may be hypothesized,

It might seem that we have maintained a systematic ambiguity in these pages by speaking of the process of optimizing one's MG both with regard to coming up with the optimally-economical account of the language (English) one is acquiring, and with regard to optimizing one's MG beyond that point, instituting changes in English to conform with the more highly-valued MG thus formulated. But the ambiguity is only apparent, in my view: only at the very earliest stages (in fact, when there are no recursive rules in the MG) can it be said that an MG is being formulated so as to generate with maximum simplicity a language actually observed: the 'ambiguity' disappears when we realize that the child's MG, at a relatively early stage, generates a set of sentences that ranges far beyond anything the child has actually heard. Thus it seems quite reasonable to say that the child does indeed attempt to optimize his MG and have that MG generate what he 'knows' to be in English. When these two goals conflict, as when English is sub-optimal, then one of two things happens: (1) the fact that English is suboptimal is manifest, and the over-optimized grammar is adulterated to fit the facts; or (2) the fact that English is suboptimal is not manifest, and the over-optimized grammar is allowed to stand, with—if all speakers over-optimize--a consequent increase in the economy of the LG itself. As an example of (1) we may take the fact that children invariably over-generalize (over-optimize) their rule for forming the past and past-participles of English 'strong' verbs (e.g. "break/breaked/breaked" instead of "break/broke/broken") and then, at least in the usual environment, adulterate their verb-formation rules to allow for the English irregularities (Ervin, 1964, pp. 177-179; McNeill, 1966, pp. 70-72). As an example of (2) we may suppose that a failure of English (i.e., the speech of familiar speakers) to optimize might tend to go unnoticed of its effect were either very rare ("do so" with Benefactive Objects) or else purely negative (non-occurrence of Passives on Benefactive Objects with verbs other than "buy"); thus, for instance, an adult speaking the "inconsistent" dialect noted earlier might still have a child who, purely through contact with the parent's speech, could" through optimization derive the more powerful MG.

That children are strongly impelled to achieve a maximally economical MG has been stated forthrightly in various places, most succinctly perhaps in Halle (1964, p. 344, after Chomsky): "...language acquisition by a child may best be pictured as a process of constructing the simplest (optimal) grammar capable of generating the set of utterances [sentences], of which the utterances heard by the child are a representative sample". (The rest of Halle's discussion on this and the following page is well worth re-reading.) This comment does not directly suggest that languages change in the direction of permitting simpler grammars, but the inference is a natural one and it fits neatly with our general apprehension of many changes: with, for example, our apprehension that the uniquely-inflected verb "dive" ("dive/dove/dived" in the United States) may well change, receding to "dive/dived/dived" or else advancing to "dive/dove/diven".

It is perhaps worth remarking in this connection that, conceivably, some well-known diachronic phenomena may illustrate a language's changing to optimize performative economy rather than, as in these examples, to optimize competence economy. Epenthesis and hypercharacterization (Malkiel, 1957-1958) seem to be steps toward optimizing production and reception, respectively, rather than steps toward optimizing competence economy of rule-statement; and it might be that some of these changes are better explained under the APG hypothesis than under the conventional hypothesis supporting the CG. But further speculation along these lines would be otiose at this time.

For a comment on the inconsistency of children's MG's, and on the fact that this inconsistency gradually diminishes with increasing age, see Turner and Rommetveit, 1967a, pp.656f.

they did neither, and so were left with an inconsistent and suboptimal grammar.

Whether or not the Benefactive Object rules are indeed late--and what "late" means<sup>69</sup>--must remain subjects for further inquiry. At least the fact of suboptimality of many adult MG's seems irrefragable; and the explanation proposed above is consistent with that fact and is not inconsistent with what little has been hypothesized about the acquisition of language.

But we must be perfectly clear about what this finding means. It does not mean that what speakers have in the way of a linguistic faculty is a disorderly congeries of rules; what it means, from all we know, is that speakers do have MG's that are basically "grammars" in the full sense in which the English LG is a "grammar"--see again all of the discussion of this point in [1.], above--but that these MG's can be, at various places, suboptimal by either CG or APG criteria of simplicity. But NB that these restricted effects are conspicuous precisely because they are exceptions to the overwhelmingly general case: the coherence and (given its scope) the simplicity of the grammar.

Thus even the view that insists on allowing for every plausible way in which the MG can depart from the LG--first by being an APG, then by being a suboptimal APG--ends by postulating a grammar much like the familiar one (but one optimized at least in part in accordance with criteria of performative simplicity). In the end, though now from a new angle, we return to an over-all view more like the orthodox one than might at first have seemed likely. Whether or not the new angle is a more correct one, of course, is a judgment that we cannot make with any finality until we can avail ourselves of the results of further experiments--including the several outlined in these pages--and of further insights into the domain of grammar.



<sup>69</sup> Rather little effort has been directed at determining the age at which people stop making significant additions to their MG's. (Of course non-significant additions, like new lexical items or idioms, may continue to be acquired up to senility, at whatever age that commences.) That is, when is what we may call one's "linguistic majority" reached? The standard answer appears to be "12" (as see Lenneberg, 1967, pp. 164-181), but clearly, for some speakers, the figure is as late as 14. Or this at least is a reasonable guess based on the fact that at 12 and 13 these speakers appear to lack little of the full English apparatus and to be acquiring at a slow rate. (But they are still acquiring.)

One might suppose that these last two (or more) years of acquisition are optional, and that they are at least partly responsible for the greater grammatical richness observable in some people's capacity.

References

1. Brown, R., and Bellugi, U. Three processes in the Child's Acquisition of Syntax. Harvard Educational Review, 1964, 34, 133-151.
  2. Brown, R., and Fraser, C. The Acquisition of Syntax. In Bellugi, U., and Brown, R., (Eds.) The Acquisition of Language. (Monographs of the Society for Research in Child Development, 92). Lafayette, Ind.: Purdue University, 1964.
  3. Chomsky, N. Transformational Analysis. Unpublished Doctoral Dissertation, University of Pennsylvania, Philadelphia, 1955.
  4. Chomsky, N. Syntactic Structures. (Janua Linguarum 4). The Hague: Mouton and Co., 1957.
  5. Chomsky, N. Explanatory Models in Linguistics. In Nagel, E., Suppes, P., and Tarski, A. (Eds.), Logic, Methodology, and Philosophy of Science. Stanford: Stanford University, 1962.
  6. Chomsky, N. Formal Properties of Grammars. In Luce, R. D., Bush, R. R., and Galanter, E., (Eds.) Handbook of Mathematical Psychology, Vol. II. New York: John Wiley and Sons, 1963.
  7. Chomsky, N. Current Issues in Linguistic Theory (Janua Linguarum 38). The Hague: Mouton and Co., 1964.
  8. Chomsky, N. Aspects of the Theory of Syntax. Cambridge: M.I.T., 1965.
  9. Chomsky, N., Cartesian Linguistics. New York: Harper and Row, 1966.
  10. Chomsky, N. The Formal Nature of Language. In Lenneberg, E. H., Biological Foundations of Language. New York: John Wiley and Sons, 1967.
  11. Chomsky, N. The General Properties of Language. In Darley, F. L. (Ed.) Brain Mechanisms Underlying Speech and Language. (Proceedings of the Princeton Conference on Brain Mechanisms Underlying Speech and Language, 1965.) New York: Grune and Stratton, 1967, a.
  12. Chomsky, N., and Miller, G. A. Introduction to the Formal Analysis of Natural Languages. In Luce, R. D., Bush, R. R., and Galanter, E., (Eds.) Handbook of Mathematical Psychology, Vol. II. New York: John Wiley and Sons, 1963.
  13. Clark, H. H., and Clark, E. V. Semantic Distinctions and Memory for Complex Sentences. Quarterly Journal of Experimental Psychology, 1968, 20, 129-138.
  14. Clark, H. H., and Stafford, R. A. Memory for Semantic Features in the Verb. Journal of Experimental Psychology (in press).
  15. Ervin, S. M. Imitation and Structural Change in Children's Language. In Lenneberg, E. H. (Ed.) New Directions in the Study of Language. Cambridge: M.I.T., 1964.
-

16. Fillmore, C. J. The Case for Case. In Bach, E., and Harms, R. T. (Eds.) Universals in Linguistic Theory. New York: Holt, Rinehart and Winston, 1968.
17. Firbas, J. Thoughts on the Communicative Function of the Verb in English, German, and Czech. Brno Studies in English, 1959, 1, 39-68.
18. Firbas, J. From Comparative Word-Order Studies. Brno Studies in English, 1964, 4, 11-126.
19. Fodor, J. A., and Garrett, M. Some Reflections on Competence and Performance. In Lyons, J., and Wales, R. J. (Eds.) Psycholinguistics Papers. Edinburgh: University of Edinburgh, 1966.
20. Fodor, J. A., and Garrett, M. Some Syntactic Determinants of Sentential Complexity. Perception and Psychophysics, 1967, 2, 289-296.
21. Foss, D. J., Bever, T. G., and Silver, M. The Comprehension and Verification of Ambiguous Sentences. Perception and Psychophysics, 1968, 4, 304-306.
22. Fraser, C., Bellugi, U., and Brown, R. Control of Grammar in Imitation, Comprehension, and Production. Journal of Verbal Learning and Verbal Behavior, 1963, 2, 121-135.
23. Gough, P.B. The Verification of Sentences: The Effects of Delay of Evidence and Sentence Length. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 492-496.
24. Grice, H. P. Utterer's Meaning, Sentence-Meaning, and Word-Meaning. Foundations of Language, 1968, 4, 225-242.
25. Grunig, B. Les théories transformationnelles. La Linguistique, 1965, #2, 1-24; 1966, #1, 31-101.
26. Halle, M. Phonology in Generative Grammar. Word, 1962, 18, 54-72; reprinted in Fodor, J. A., and Katz, J. J. (Eds.) The Structure of Language: Readings in the Philosophy of Language. Englewood Cliffs, New Jersey: Prentice-Hall, 1964.
27. Halliday, M.A.K. Notes on Transitivity and Theme in English, Part 2. Journal of Linguistics, 1967, 3, 199-244.
28. Harris, Z. S. Methods in Structural Linguistics. Chicago: University of Chicago, 1951.
29. Harris, Z. Distributional Structure. Word, 1954, 10, 146-162.
30. Harris, Z. S. Co-occurrence and Transformation in Linguistic Structure. Language, 1957, 33, 283-340.
31. Harris, Z. S. Transformational Theory. Language, 1965, 41, 363-401.
32. Harris, Z. S. Mathematical Structures of Language. (Interscience Tracts in Pure and Applied Mathematics, 21). New York: John Wiley and Sons, 1968.

33. Harwood, F. W. Quantitative Study of the Speech of Australian Children. Language and Speech, 1959, 2, 236-271.
  34. Hiž, H. Congrammaticality, Batteries of Transformations, and Grammatical Categories. In Jakobson, R. (Ed.) Proceedings of Symposia in Applied Mathematics, XII. Providence: American Mathematical Society, 1961.
  35. Hiž, H. The Role of Paraphrase in Grammar. In C.I.J.M. Stuart (Ed.) Monograph Series on Languages and Linguistics 17. Washington: Georgetown University, 1964.
  36. Katz, J. J., Mentalism in Linguistics. Language, 1964, 40, 124-137.
  37. Katz, J. J., and Postal, P. M. An Integrated Theory of Linguistic Descriptions. (Research Monograph 26). Cambridge: M.I.T., 1964.
  38. Kotarbinska, J. On Ostensive Definitions. Philosophy of Science, 1960, 27, 1-22.
  39. Lakoff, G. On the Nature of Syntactic Irregularity (Doctoral Dissertation, Indiana University). Mathematical Linguistics and Automatic Translation Report NSF-16. Cambridge: The Computation Laboratory of Harvard University, 1965.
  40. Lakoff, G. Instrumental Adverbs and the Concept of Deep Structure. Foundations of Language, 1968, 4, 4-29.
  41. Lakoff, G., and Ross, J. R. A Criterion for Verb Phrase Constituency. In Kuno, S. (Ed.), Mathematical Linguistics and Automatic Translation Report NSF-17. Cambridge: The Computation Laboratory of Harvard University, 1966.
  42. Lenneberg, E. H. Cognition in Ethnolinguistics. Language, 1953, 29, 463-471.
  43. Lenneberg, E. H. Biological Foundations of Language. New York: John Wiley and Sons, 1967.
  44. MacKay, D. G. To End Ambiguous Sentences. Perception and Psychophysics, 1966, 1, 426-436.
  45. Malkiel, Y. Diachronic Hypercharacterization in Romance. Archivum Linguisticum, 1957, 9, 79-113; 1958, 10, 1-30.
  46. McCawley, J. D. Concerning the Base Component of a Transformational Grammar. Foundations of Language, 1968, 4, 243-269.
  47. McNeill, D. Developmental Psycholinguistics. In Smith, F., and Miller, G. A. (Eds.) The Genesis of Language. Cambridge: M.I.T., 1966.
  48. Mehler, J. Some Effects of Grammatical Transformations on the Recall of English Sentences. Journal of Verbal Learning and Verbal Behavior, 1963, 2, 346-351.
-

49. Miller, G. A. Some Psychological Studies of Grammar. American Psychologist, 1962, 17, 748-762.
50. Miller, G. A., and Chomsky, N. Finitary Models of Language Users. In Luce, R. D., Bush, R. R., and Galanter, E., (Eds.) Handbook of Mathematical Psychology, Vol. II. New York: John Wiley and Sons, 1963.
51. Miller, G. A., and McKean, K. A Chronometric Study of Some Relations Between Sentences. Quarterly Journal of Experimental Psychology, 1964, 16, 297-308.
52. Postal, P. M. On So-Called 'Pronouns' in English. In Dinneen, F. P. (Ed.), Monograph Series on Languages and Linguistics 19. Washington: Georgetown University, 1966.
53. Postal, P. M. A Note on 'Understood Transitively'. International Journal of American Linguistics, 1966, a, 32, 90-93.
54. Ross, J. R. Relativization in Extraposed Clauses. In Kuno, S. (Ed.) Mathematical Linguistics and Automatic Translation Report NSF-17. Cambridge: The Computation Laboratory of Harvard University, 1966.
55. Ross, J. R. Constraints on Variables in Syntax, Doctoral Dissertation (dittoed), Massachusetts Institute of Technology, Cambridge, 1967.
56. Savin, H., and Perchonock, E. Grammatical Structure and the Immediate Recall of English Sentences. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 348-353.
57. Shipley, E.F., Smith, C.S., and Gleitman, L.R. A Study in the Acquisition of Language: Free Responses to Commands. (Technical Report VIII). Philadelphia: Eastern Pennsylvania Psychiatric Institute, n.d. [1967].
58. Slobin, D.I. Grammatical Transformations in Childhood and Adulthood. Unpublished Doctoral Dissertation, Harvard University, Cambridge: 1963.
59. Slobin, D.I. Grammatical Transformations and Sentence Comprehension in Childhood and Adulthood. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 219-227.
60. Slobin, D.I. Recall of Full and Truncated Passive Sentences in Connected Discourse. Journal of Verbal Learning and Verbal Behavior, 1968, 7, 876-881.
61. Stene, A. The Animate Gender in Modern Colloquial English. Norsk Tidsskrift for Sprogvidenskap, 1934, 7, 350-355.
62. Svartvik, J. On Voice in the English Verb. The Hague: Mouton and Co., 1966.
63. Turner, E.A., and Rommetveit, R. Experimental Manipulation of the Production of Active and Passive Voice in Children. Language and Speech, 1967, 10, 169-180.

64. Turner, E.A., and Rommetveit, R. The Acquisition of Sentence Voice and Reversibility. Child Development, 1967, a. 3J3, 649-660.
65. Turner, E.A., and Rommetveit, R. Focus of Attention in Recall of Active and Passive Sentences. Journal of Verbal Learning and Verbal Behavior, 1968, 2, 543 - 548.
66. Valdman, A. Review of Guiraud, P., Le français populaire. Language, 1968, 123-127.
67. Watt, W.C. English Locative Sentences. Paper given at the Forty-Second Annual Meeting, Linguistic Society of America, 1967. Abstract in Handbook of the Meeting, p.62.
68. Watt, W.C. English Reduplication. Journal of English Linguistics, 1968, 2, 96-129.
69. Watt, W.C. Linguistic Recoverability. Forthcoming.
70. Weir, R.H. Language in the Crib. The Hague: Mouton and Co., 1962.