Preprint No. 18 Classification: TG 3.1

An Application of an Extended Generative Semantic Model of Language to Man-machine Interaction

- Robert I. Binnick -

INTERNATIONAL CONFERENCE

ON

COMPUTATIONAL LINGUISTICS

COLING

1969

KVAL
Forskningsgruppen
för
KVANTITATIV
LINGVISTIK

RESEARCH GROUP FOR QUANTITATIVE LINGUISTICS

Address: Fack Stockholm 40, SWEDEN

International Conference on
COMPUTATIONAL LINGUISTICS
1-4 September 1969
Sånga-Säby, Sweden

Preprint No. 18
Classification: TG 3.1

An Application of an Extended Generative Semantic Model of Language to Man-machine Interaction

- Robert I. Binnick -

Contents:

Author's address:

Department of Linguistics
University of Chicago
Chicago, Ill., 606 37
U.S.A.

FEASIBILITY

This paper discusses the feasibility of applying a model of language use based on a modification and extension (to be discussed below) of the generative semantic (transformational) theory of language competence recently developed by Paul Postal, George Lakoff, John Robert Ross, Eames D. McCawley, and others, to problems of computational linguistics.

The theory of generative semantics, to be discussed in section II, is an outgrowth of, and reaction to, Chomsky's 1965 theory of transformational linguistics. It is a radical theory which deals with a very great range of problems with very abstract methods. Those working in this paradigm hold that there is a linguistic level reflecting conceptual or semantic structure which is directly convertible into surface syntax by a single set of garden-variety transformations, with no significant intermediary level, that is, no "deep structure". Those of us working in generative semantics believe that methods substantially those long familiar in linguistics can achieve very absract, very general results which treat semantics in a more serious and enlightening way than ever before. I do not, I think, support this very strong claim very well in section II, but I provide summaries of several studies and a lengthy bibliogrpahy of works which when consulted will hopefully give some feeling for what is being attempted, I think not without results.

But generative semantics is a model, or rather, a theory, of competence, like most serious theories of language now held to by American linguists. Even if, as might be claimed, our semantic structures are to be merely variants of the structures long familiar from formal logic, so that if our assumptions are correct, we will ultimately be able to directly transform surface structures into underlying semantic structures, the majority of actual sentences, as well as all hypersentential structures, the treatment of which has been swept under the rug of "performance", will remain unhandleable.

Accordingly, I propose initially certain extensions and modifications of the theory to make it in some sense a model of performance. But if we are to apply it to the computer, a major component must still be added. The impetus to this application is the possibility of creating an understanding machine, described in section IV below. Since the actual human interpretation of language depends on past knowledge (consider which of these sentences is good and why:

As for Albuquerque, the Eiffel Tower is pretty.
As for Paris, the Eiffel Tower is pretty.

And these:

Shirley is a blonde and Susan is Nordic-looking too.
Shirley is a linguist and Susan is Nordic-lloking too.)

the old split between semantics, syntax, and pragmatics must be revised, and our model closely linked with a memory and possibly a logic component as well.

Obviously this defines a very difficult task, but insofar as such goals as MT, artificial intelligence, and machine reading of handwritten material or writing of spoken material involve comprehension on the part of the machine, of which there seems to be no doubt, these important goals will continue to elude us until such time as we can devise such an understanding machine as I have described below.

I believe that generative semantics lays the foundation for studies relevant to such a development, and it is in this context that my proposals are made.

* * *

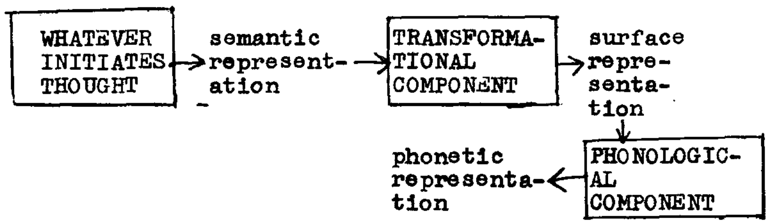
In section II I will discuss generative semantics.
In section III I will discuss the body of my proposals here.
In section IV I will discuss what should be required of a generalized "understanding" machine.

Part II. The theory of Generative Semantics.

The theory of generative semantics is an outgrowth and reaction to the theory of transformational grammar as represented in Chomsky's 1965 book, Aspects of the Theory of Syntax (MIT Press). To a very large extent, this theory has been the development of a small group of former students of Chomsky's or their close colleagues. John (Haj) Ross has said that the theory is really just an attempt to explicate Paul Postal's work of five years ago to date. If Postal was the founder of this school, if you can call it that, its main workers have been Haj Ross and George Lakoff, who between 1965 and 1968 swept aside most of transformational linguistics as it then was. But perhaps best known of the group is James McCawley, who graduated from MIT in 1965 with a Ph.D. based on work in phonology, not syntax or semantics. He promptly amazed Lakoff and Ross by some very substantive work in the latter areas as well as phonology. As a student of McCawley's I will be emphasizing his contributions here, and those of my colleagues at Chicago, Jerry L. Morgan and Georgia M. Green, but it should be kept in mind that people like Ross, Lakoff, Postal, Arnold Zwicky, David Perlmutter, Emmon Bach, Robin Lakoff, and several others, have made the current theory possible, and that many others, such as Robert Wall, Lauri Kartunnen, Ronald Langacker, and others, have contributed as well. It should also be kept in mind that the Case Grammar of Fillmore and the work done by Gruber, while differing from generative semantics, have contributed a great deal to it.

The basic theory of generative semantics is built upon an attempt to relate the underlying semantic structure of language to the surface, phonetic manifestation of that underlying structure. That is, a phonetic reality is recognized, and a semantic reality is recognized. But unlike other versions of transformational grammar, this theory assigns no special status to syntax; syntax is subsumed in the semantics. McCawley has jokingly referred to his theory as being one of either "semantax" or "synantics". The name generative semantics is not a particularly good one, since it implies that the goal of the theory is, as with the work of Chomsky, to "separate the grammatical sequences" of a language "from the ungrammatical sequences." (Chomsky, Syntactic Structures, 13.) In other words, to generate all and only grammatical sentences of a language. This is not at all the goal of generative semantics. Rather, what we want to do is in some rigorous way specify the correlations of underlying semantic entities and surface phonetic entities: to specify for any underlying semantic structure what its possible phonetic realizations in some language are, and for some phonetic structure what underlying semantic structures it can represent. Naturally, some descriptive ability is predicated as well, that is, we want to be able to define ambiguity in some algorithmic fashion, we want to be able to define levels or classes of ill-correlation between structures on different levels, etc. Chomsky would say that a sentence like "Golf plays John" is eminently deserving of a star; we would say (1) if it's supposed to mean 'John plays golf', it doesn't succeed in conveying the message; (2) if it's supposed to mean 'John loves Marsha', then it's <u>really</u> bad; and (3) if Golf is a man's name and John the name of a game or role, it's a good sentence --- indeed, one can very well imagine arcane circumstances under which one might utter that sentence with the intent of saying that the game plays John, that the tail wags the dog, as it were. Suppose, for example, that John's wife were tired of him spending all his free time playing golf and she grumbled to a neighbor about it, and the neighbor rather unfeelingly replied, "Oh well, John plays golf." I can very well imagine John's wife complaining bitterly,"Oh no, golf plays John." In any case, it is for hus unimaginative approach to language that Chomsky has been jokingly called a "bourgeois formalist". Even when we use stars, we try to keep in mind that just about any valid phonological string of a language conveys one or more meanings in some context, and that it is artificial to take a string out of context and declare it good or bad. So "generative semantics" is a bad name.

The following diagram of the components of the theory is based on McCawley's paper in the proceedings of the 4th Regional Meeting of the Chicago Linguistic Society (1968). A theory very similar is discussed in Ronald Langacker's book <u>Language and its Structure</u> (Harbrace, 1968), pp. 114-34.



The above diagram comes from a report prepared by myself, Jerry Morgan, and Georgia Green, called the Camelot Report, which attempted to describe the current state of transformational research in the Summer of 1968, particularly in reference to the LSA Summer Linguistic Institute at the University of Illinois, where Haj Ross, George Lakoff, and Jim McCawley had lectured to large groups on a huge number of very "hairy" (i.e., difficult and tickleishly novel) topics.

In that report (which was prepared for Victor Yngve), we raised several questions concerning the above representation. We asked:

1. What will an adequate semantic representation have to include? What form will it have?

2. What can a transformation do? What does one look like?

3. At what stage and in what manner are semantic representations converted into words of real languages?

These were by no means all of the questions asked. Needless to say, the answering of these questions has hardly begun and will undoubtedly guarantee linguists a few good centuries of work at least. It is only in the last decade that syntax has been the subject of serious work, and we are still only discovering how ignorant we are. Semantics is even newer, less than a decade old. If anyone doubts that this is true, consider a) what the above 3 questions would have meant to a linguist in (say) 1955, and b) why he would have been wrong in his (lack of) comprehension of them. One of the great contributions of Postal and Ross has been their constant critical look at transformational grammar. One of the things they saw was that our transformations were (and are) extremely powerful devices, with practically no constraints placed on their formulation.

What I will do here is summarize some of the attempts at partial answers to the three above questions. In this way I can delimit and explicate generative semantics best.

I will start by abstracting parts of two papers by McCawley that deal with the nature of semantic representation. In a paper in the Japanese journal Kotoba no Uchu (World of Language) in 1967, McCawley argued that semantic representation would be similar to syntactic representation as familiar from Aspects-type grammar, but that it would also be quite similar to symbolic logic as familiar from the tons of work that have followed Principia and such studies. That semantic representation should resemble syntactic representation makes sense if only because we are arguing for a single set of rules that transforms (i.e., relates) the underlying structure into (to) the surface structures. There will be more about that later.

McCawley argues as follows: the following devices have all had a role in symbolic logic:

1. propositional connectives: 'and', 'or', 'not'.

2. constants denoting individuals.

3. predicates, denoting properties and relationships.

4. set symbols and the quantifiers 'all' and 'there exists'.

5. descriptions of sets and individuals.

The following devices play a role in natural languages:

1. all lgs. have words for 'and', 'or', and 'not'. (he notes however that these words in natural lgs. may connect more than sentences)

2. "indices" denoting individuals; John loves John might be represented as x_1 loves x_2, but John loves

himself is x_1 loves x_1.

3. predicates are expressed in natural lgs. (by verbs, adjectives, nouns, etc.)
4. "Words such as all and at least one are two members of a rather large class of expressions which are used to indicate not only the existence of an individual or a set but the absolute or relative number of members in that set."
5. sets and individuals can be expressed as descriptions using modified noun phrases.

McCawley then gives further reasons for supposing symbolic logic representation to be proper for semantic representation. (See the bibliography to this section where this and other papers that can be consulted for these arguments in detail are listed.)

In a paper prepared for the symposium on "Cognitive Studies and Artificial Intelligence Research" held by the Wenner-Gren Foundation at the University of Chicago in March of this year, McCawley discussed semantic representation at length. Some of what he had to say there should be noted. He claimed, "semantic representation must indicate the immediate constituent structure of the elements involved in it (i.e. examples showing that different meanings can consist of the same semantic elements combined in different ways [are easy to come by])" (p.1) He gave the example of

John doesn't beat his wife because he loves her.

If the negation applies to John beats his wife, the sentence means 'the reason that John doesn't beat his wife is that he loves her', whereas if it applies to the John beats his wife because he loves her., the mg. is 'the reason that John beats his wife is not that he loves her.' Notice that here a surface form represents at least two different underlying structures which nonetheless contain precisely the same semantic elements-- grouped differently, however.

Another point made is that "semantic representations must include ..., some indication of presupposed coreference." (p.2) That is, the following sentence in neutral (i.e. null) context is ambiguous three-ways:

John told Harry that his wife was pretty.

Whose wife? John's? Harry's? or a third's? It could be any. However, if we know who his refers to, there is no such ambiguity. This may seem trivial, but it is a point often ignored.

McCawley then gives an argument for referential indices being different from expressions used to describe. The sentence

Max denied that he kissed the girl he kissed.

is not contradictory if "the girl he kissed" is the speaker's description.

Another notion is that of presupposed set membership. In

Max is more intelligent than most Americans.

said with primary stress on most, the sentence is good if and only if Max is presupposed to be American, that is, the sentence implies Max is American. With primary stress on Americans, however, Max is presupposed not to be American. Presupposition is in general a very hairy topic which was recently the subject of an entire conference (at the Ohio State University). We know very little about the nuamces of implication and are only beginning even to identify the problems. But if a machine is ever to read Catcher in the Rye catching all the nuances of the italicized words, we had better find out how stress is used to alter the presuppositional set of a sentence. I need not be so unsubtle as to suggest the extreme value of such researches to psychology. Perhaps they already know about all this, for all I know. In any case I cannot restrain myself from including McCawley's beautiful example

CIA Agents are more stupid than most Americans.

He had primary stress on the most, but I prefer to think of it as going on the Americans.

I would like to interject at this point a minor apology. I have been rather fan-clubish here and have waved my hand a lot. Frankly I see no value in rehearsing here all the arguments available elsewhere. But I would like the reader to bear in mind my skimpy resume in no way reflects the quality of the original. Let me also note, lest I seem unduly credulous towards the thoughts of Chairman Quang [mild-mannered linguist J. D. McCawley is in reality Q. P. Dong, Chairman of Unamerican Studies at an unknown university], that most of us working within the paradigm of generative semantics would be the first to admit that our theories haven't a prayer of being right, that is, they hardly approach even a partially realistic and naturalistic theory of language. If we like it better than other paradigms it is because we believe that no other current theory is any better and that this one at least has a good chance of self-improvement. (End of apologia.)

If semantic representation looks much like logical representation, it also differs from it. In the Kotoba no Uchu paper McCawley noted the following differences:

1. "It is necessary to admit predicates which assert properties not only of individuals but also of sets and propositions."

2. "In mathematics one enumerates certain objects which [one] will talk about, defines other objects in terms of these objects, and confines [oneself] to a discussion of objects which [one] has either postulated or defined.... However, one does not begin a conversation by giving a list of postulates and definitions. ....people often talk about things which either do not exist or which they have identified incorrectly. Indices exist in the minds of the speaker rather than in the real world; they are conceptual entities which the individual speaker creates in interpreting his experience."

In the Wenner-Gren symposium, McCawley had more to say about the difference between logic and language.

1. Immediate constituent structure (trees) rather than parentheses are basic. First, "semantic representations are to form the input to a system of transformations that relate meaning to superficial form; to the

extent that these transformations have been formulated and justified, they appear to be stateable only in terms of constituent structure and constituent type, rather than in terms of configurations of parentheses and terminal symbols." Secondly, "it may be necessary to operate in terms of semantic representations in which symbols have no left-to-right ordering...."

2. There will have to be more 'logical operators', such as most, almost all, and many.

3. "And and ... or ... cannot be regarded as just binary operators but must be allowed to take an arbitrary number of operands."

4. The quantifiers must be restricted rather than unrestricted as in most logical systems. Some quantifiers imply existence: All dogs like to bite postmen. involves the presupposition that dogs exist, whereas the unrestricted quantifiers logicians use have no such presupposition.

5. "Adequate semantic representation of sentences involving 'shifters' (Jakobson, 1957) such as I, you, here, now, ..., gestures and deictic words like this and that, and tenses, will have to include reference to the speech act. The most promising approach to this aspect of semantic representation ... is Ross's (1969) elaboration of Austin's (1962) notion of 'performative verb'." (See now too Searle's book, Speech Acts, CUP, 1969 ---RIB)

6. "The range of indices will have to be enormous. In particular, it will have to include not only indices that purport to refer to physical objects, but also indices corresponding to mythical or literary objects, so that one can represent the meaning of sentences such as

   The Trobriand Islanders believe in Santa Claus, but they call him Ubu Ubu."

7. McC. rejects "the traditional distinction between 'predicate' and 'logical operator' and treat[s] such 'logical operators' as quantifiers, conjunctions, and negation as predicates...."

To clarify the relationship of semantic to syntactic representations let me quote here from McCawley's Kotoba no Uchu paper:

> Since the rules for combining items into larger units in symbolic logic formulas must be stated in terms of categories such as 'preposition', 'predicate', and 'index', these categories can be regarded as labels on the nodes of these trees. And since ... these categories all appear to correspond to syntactic categories, the same symbols (S, V, NP, etc.) may be used as node labels in semantic representations as are used in syntactic representations. Accordingly, semantic representations appear to be extremely close in formal nature to syntactic representations, so close in fact that it becomes possible to catalogue the conceivable formal differences and determine whether those differences are real or apparent.

Among such differences he lists:

1. "The items in a syntactic representation must be assigned a linear order, whereas it is not obvious that linear ordering of items in a semantic representation makes any sense."

2. "Syntactic representations involve lexical items from the language as their terminal nodes, whereas the terminal nodes in a semantic representation are semantic units rather than lexical units."

3. "There are many syntactic categories which appear to play no role in semantic representation, for ex., verb-phrase, preposition, and prepositional phrase." (At the 5th Regional Meeting of the CLS, April of this year, A. L. Becker of the University of Michigan presented a paper in which he argued prepositions are underlying predicates; prepositional phrases are accordingly verb-phrases.)

McCawley concluded nonetheless that these differences do not provide an argument that semantic representations are different in formal nature from syntactic representations. Again, I will omit his reasons for that conclusion.

I might summarize all this by saying:

1. Semantic representation is a modification of the representations long familiar from formal logic.

2. Such representations do not radically differ from the surface syntactic representations of <u>Aspects</u>-type grammar.

Let me close by posing more problems. McCawley asks the following questions at the end of his Kotoba no Uchu paper. While they do not specifically relate to semantic structure, I include them to give some idea of what we believe to be the sort of questions that a serious theory of language should provide justifiable answers for:

1. How do the mgs. of words change as a language evolves?

2. How does a child learn mgs. in learning to speak his native language?

3. What mechanisms are involved in phenomena such as metaphor....? (Dorothy Lambert has written a Ph.D. thesis at Michigan on the subject of metaphor within the paradigm of Case Grammar. This 500 page dissertation is probably one of the best studies of the subject to date from a linguistic point of view.)--RIB)

4. To what extent are the units of semantic representationq universal?

5. To what extent does the lexicon of a language have a structure?

6. Can all languages express the same ideas?

7. To what extent doe's one's language affect his thinking?

8. To what extent is one's ability to learn lexical items conditioned by his knowledge of the world?

I will now turn to the second question raised above on p. II-3. This question has as yet received little study. It is a very difficult topic, but a very important one. I will confine myself here to a few brief comments and a few references.

One of the important studies underway now is about syntactic variables. This was the subject of Ross' 1967 dissertation. Variables such as X and Y are familiar from transformational grammars, but no one had attempted before to specify in general what the notion of syntactic variable entailed. While Ross' study was important, and he came up with several important constraints on the form of transformations, much work remains. Lakoff and Postal are also working on related questions. Let me list here some of the constraints Ross gave in his thesis:

1) The complex NP constraint.
   No element contained in a sentence dominated by a noun phrase withxa lexical head noun may be moved out of that noun phrase by a transformation. (p, 127)

2) The cross-over condition.
   No NP mentioned in the structural index of a transformation may be reordered by that rule in such a way as to cross over a coreferential NP. (p. 132)

3) The coordinate structure constraint.
   In a coordinate structure, no conjunct may be moved, nor may any element contained in a conjunct be moved out of that conjunct. (p. 161)

4) The pied piping convention.
   Any transformation which is stated in such a way as to effect the reordering of some specified node NP, where this node is preceded and followed by variables in the structural index of the rule, may apply to this NP or to any non-coordinate NP which dominates it, as long as there are no occurences of any coordinate node, nor of the node S, on the branch connecting the higher node and the specified node.

(That is,:

> ...any NP above some specified one may be reordered, instead of the specified one, but there are environments where the lower NP may not be moved, and only some higher one can, consonant with the conditions imposed /In the convention./) (p.206)

5) The sentential subject constraint.
No element dominated by an S may be moved out of that S if that node S is dominated by an NP which itself is immediately dominated by S. (p. 243)

6) The frozen structure constraint.
If a clause has been extraposed from a noun phrase whose head noun is lexical, this noun phrase may not be moved, nor may any element of the clause be moved out of that clause. (p. 295)

7) Definition of identity.
Constituents are identical if they have the same constituent structure and are identical morpheme-for-morpheme, or if they differ only as to pronouns, where the pronouns in each of the identical constituents are commanded by antecedents in the non-identical portions of the phrase-marker. (p. 348)

A very important constraint occurs on p. 480 of the thesis, but I omit it here because it contains many terms I would not care to define here. I reccomend Ross' dissertation for anyone with doubts about any deep principles of language organization emerging from our studies in transformational grammar. He will be cured.

Recently George Lakoff has studied the notion of "derivational constraint". This study is quite recent and still very very hairy, but hints in his 1969 CLS paper, and comments by Postal on it suggest that rule odering is merely a special case or manifestation of a deeper principle of grammar organization. The next revolution effected by generative semantics may well be to drop rule ordering from our canons.

For various reasons (partly that it interests me more) I will have much more to say here about lexical insertion than I will about constraints on transformations, although undoubtedly the latter is ultimately of much greater importance.

Until 1965 or so, it was assumed that the terminal symbols of a P-marker are lexical items; the lexicon merely assigns properties to these items. Gruber in his 1965 dissertation argued that certain transformations had to occur before lexical items entered trees: that is, that there were pre-lexical transformations.

Before Gruber, the system of semantics was one in which T-rules generated from deep structures surface structures and P-rules generated semantic representations for those deep structures. This was the theory of intepretive semantics (as in Katz and Postal, for ex.) Gruber proposed a derivational semantics. Gruber intended to "show various consistently recurrent semantic relationships among parts of the sentence and among different sentences, which can best be explained by the existence of some underlying pattern of which the syntactic structure is a particular manifestation." (p.1) He concluded that "a level at which semantic interpretation will be relevant will ... be deeper than the level of 'deep structure' in syntax." (p.2) Later Lakoff showed evidence that in fact the level of semantic interpretation was that of deep structure, but argued that (as Gruber said) "syntax and semantics will have the same representation at the prelexical level"(p. 3): a single set of rules would transform semantic structures containing no lexical items into surface syntactic representations containing them.

The study of lexical insertion, the process by which the underlying semantic elements are grouped into units replaceable by surface lexical items has led to a large literature containing a great many questions, and some positive answers. An important paper was McCawley's 1968 paper, "Lexical insertion in a transformational grammar without deep structure."

There he started by assuming various points concluded in other papers of his. He very clearly presents some of the tenets of generative semantics, so with some repetition from above I quote these points here:

1. Syntactic and semantic representations are of the same formal nature....

2. There is a single system of rules ... which relates semantic representation to surface structure through intermediate stages.

3. In the earlier stages of the conversion from semantic

> representation to surface structure, terminal nodes may have for labels 'referential indices' such as were introduced in Chomsky 1965.... In semantic representation, only indices and 'predicates' are terminal node labels....

McCawley then defined 'dictionary entry' as a transformation which replaced part of a tree by a surface lexical item. He expressed doubt these rules could be ordered internally or external, since it would hardly be possible, for example, that some question would arise as to the relative ordering of the transformation introducing the word horse and that extraposing NP's in two dialects, that is, the ordering could not possibly matter.

He then raised several possibilities as to the relative ordering of the lexical rules vis-a-vis other rules. Are the lexical rules last, first, or where? McCawley argued for the lexical rules applying just before the post-cyclic rules, and adduced evidence for several rules, predicate-raising, equi-NP deletion, etc., being pre-lexical.

In his 1968 LSA paper, Jerry L. Morgan of the University of Chicago added to this. He pointed out "the rather strong assumption that lexical items only 'replace' constituents." (p.3) He wrote, "the process of syntactic derivation begins with semantic representation in terms of trees containing very highly abstract semantic terms, operating upon this by means of rules permuting, deleting, and collapsing parts of the representation, finally deriving a structure whose constituents are replaced by lexical items." (p. 3) He then stated a very strong claim of the theory:

> Given the set of universal pre-lexical rules, the set of universal semantic primitives, and the set of universal constraints on the operation of rules, such as those described by Ross 1967, these define the universal set of possible lexical items in their semantic aspect; that is, they rule out as impossible an infinite classof a priori possible "meanings" a lexical item could have. (p.4)

A second very strong claim of the theory is:

> Insofar as the selection from, and details of implementation of, the universal set of rules is language-specific, the idiosyncracies of a given language in this respect will also be reflected by systematic gaps in the lexicon. The same is true for the set of semantic primitives and the set of constraints on rules. .... (p.4-5)

Morgan came up with some restrictions on lexical items:

1) "lexical items only can replace a constituent which is not labelled S." (p.6)

2) "verbs cannot incorporate referential indices."(p.6)

One further point to be made is that lexical items can only replace well-formed subtrees.

My own work has been concerned with specifying classes of possible lexical items and accounting for the syntactic properties of verbs in terms of their semantics, thereby attempting to capture the intuition long familiar from traditional grammar that certain semantic classes of verbs, such as "verbs of giving and taking" or "verbs of motion" also form syntactic classes and hence their syntactic properties can be regarded as derived from their semantics.

Georgia Green of the University of Chicago has presented a paper (1969) which is also interesting in terms of lexical insertion. She tends to regard lexical insertion as fairly divorced from morphology, and views lexical insertion as the replacement of an entire sub-tree by a surface lexical item which may contain more than one morpheme as classically defined. This position is somewhat different from my own, as I regard lexical insertion as primarily involving the replacement of items on a 1-1 basis. However, this is an empirical question and only future research will decide which of us is more nearly correct.

So far I have discussed lexical insertion in terms of sweeping, general principles of the organization of the grammar. In order to more clearly specify what lexical insertion is all about, I ought to present some of the kinds of problems which have generated all of this interest in the subject.

At the Texas Conference on Universals in 1967, the proceedings of which were published in 1968 as Universals in Linguistic Theory, McCawley raised the Universals in Linguistic Theory, McCawley raised the question of dictionary organization anew. He opted for a "Weinreichian" lexicon in which lexical items were combinations of semantic, syntactic, and phonological information. McCawley supported this with this evidence: the reason John is sadder than that book. is bad is that the two sads in the underlying structure of the sentence are different lexical items. They therefore cannot participate in comparison:

> *John is as sad as that book he read yesterday.
> *He exploits his employees more than the oppurtunity to please.
> *Is Brazil as independent as the continuum hypothesis? (exx. of Chomsky's.)

McCawley called for a theory of "implicational relations", since in cases such as the ambiguity of warm the ambiguity is not a property of the item itself but of a class of items, and therefore such an ambiguity must be specified in terms of general principles. McCawley was not clear about the nature of these implicational relations, so that the nature of the relationship of the various sads was more or less left open. I have discussed the notion of systematic ambiguity, where the ambiguities of an entire class of verbs is specified in terms of the derivational process underlying them all, not just in terms of a descriptive statement. Thus we are seeking to explain lexical gaps in terms of statements such as

> "The reason some language L lacks a verb V glossing the verb W in the language M is that M, but not L, has the transformation T."

Anyone familiar with the lexicons of French, English, and German, for example, knows that there are certain kinds of verb which are not typical of one or another of these languages which nonetheless readily occur in the others. Such verbs are derived by processes occuring in one but not another language, and our task is to discover and describe such processes. Thus we may ultimately be able to tell how the class of French verbs, say, differs from the class of all possible verbs.

I have attempted in these few pages to present a digest of some works in the paradigm of generative semantics. I have not really attempted to provide even an elementary guide to the methods of generative semantics or to its conclusions, its findings, but I hope I have explicated somewhat its goals and given some insight into the direction in which it is moving. Some very strong claims are forthcoming on the nature of grammars and languages and hence of language itself. A tremendous amount of work needs to be done, but one can see clearly that one possible end point of this work will be a very comprehensive, very strong theory of language competence that has a great deal to say about human beings.

One perhaps minor point, though, looms up large here: generative semantics relates semantic structures to surface sentences by a single set of rules. There are several versions of transformational grammar that do this, but generative semantics is perhaps the most-developed of these. But as the saying goes, what goes up must come down: we may paraphrase this as: what can be generated, can be analyzed. The theory permits, ideally, an algorithmic translation of a surface string into one or more underlying semantic structures. For computational linguistics, that may be its most appealing feature.

BIBLIOGRAPHY

A short, select bibliography of recent works in and on generative semantics.

Austin, J. L.
1962. How to do things with words. ed. J.O. Urmson. London: OUP. (1965 paper.)

Bach, Emmon.
1967. "Have and be in English syntax." Lg. 43.462-85.
1967. "Problominalization I-II." Mimeo.
1968. "Nouns and noun phrases." in Bach & Harms.
1969. "Anti-pronominalization." Mimeo.
Forthcoming. "Binding."

Bach, Emmon, and Peters, Stanley.
1968. "Pseudo-cleft sentences." Mimeo.

Bach, Emmon, and Harms, Robert, edd.
1968. Universals in Linguistic Theory. NYC: Holt, Rinehart, Winston.

Becker, A. L.
1969. "Prepositions as Predicates." in Binnick et al.

Benwick, Launcelot de, the Green Knight, and Morgan le Faye.
(=R. Binnick, G. M. Green, J. L. Morgan.)
1968. Camelot 1968. Internal memo., MT Group, UC, mimeo.

Bierwisch, Manfred, and Heidolph (edd)
to appear. Recent Advances in Linguistics. Mouton.

Binnick, Robert I.
1967. "Semantic and syntactic classes of verbs." Mimeo.
1967. "The lexicon in a derivational semantic theory of transformational linguistics." In Chicago Journal of Linguistics 1/1, available from University Microfilms, Ann Arbor, Michigan, as serial S-372.
1968. "On the nature of the 'lexical item'", in Darden et al.
1968. "On transformationally derived verbs in a grammar of English", Ditto, read at LSA.
1968. "The characterization of abstract lexical entities", Ditto, read at ACL.
1968. "Transitive verbs and lexical insertion", dittoed, read at Kansas and CLS.
1969. "Predicative structure." Unpublished Ph.D. diss.

Binnick, R., G. Green, J. Morgan, & A. Davison, edd.
1969. Papers from the 5th Regional Meeting, Chicago Linguistic Society. Chicago: Department of Linguistics, University of Chicago. (Advt.: dirt cheap at $5.)

Camelot. see Benwick.

Darden, B. J., C. J. Bailey, A. Davison, edd.
1968. Papers from the 4th Regional Meeting, Chicago Linguistic Society. Chicago: Department of Linguistics, University of Chicago. (Advt.: still dirt cheap at $3.)

DeRijk, Rudolph.
1968. "A note on prelexical predicate raising." Dittoed.

Donnellen, Keith.
1966. "Reference and definite descriptions." Philosophical Review 75. 281-304.

Green, Georgia M.
1968. "On <u>too</u> and <u>either</u>, and not just <u>too</u> and <u>either</u>, either." in Darden et al.
1969. "Some theoretical implications of the lexical expression of emphatic conjunction." Unpublished M.A. thesis, dittoed.
1969. "On the notion 'related lexical entry.'" in Binnick et al.
Forthcoming. Review of R. Lakoff, Abstract Syntax and Latin Complementation. To appear in <u>Lg</u>.

Gruber, Jeffrey S.
1965. Studies in lexical relations. Unpublished diss.
1967. "Look and see", in <u>Lg</u>. 43.937-47.
1967. "The functions of the lexicon in formal descriptive grammars." Systems Development Corporation document TM-3770/000/00.

Jakobovits and Steinberg, edd.
to appear. Semantics: an interdisciplinary reader....

Jakobson, Roman
1957. "Shifters, verbal categories and the Russian verb." Slavic Dept., Harvard Univ.

Kartunnen, Lauri.
1968. "Co-reference and discourse." Read at LSA.
1968. "What do referential indices refer to?", RAND Corporation Report.
1969. "Migs and pilots." Mimeo.
1969. "Pronouns and variables." In Binnick et al.

Katz, J. and Postal, Paul.
1964. An integrated theory of linguistic descriptions. MIT Press.

Kiparsky, Paul.
1968. "Linguistic Universals and Linguistic Change." in Bach and Harms.

Kiparsky, P. & C.
1968. "Fact." To appear in Bierwisch and Heidolph.

Lakoff, George.
1965. "On the nature of syntactic irregularity." Indiana Univ. diss. =NSF-16 report, ed. A. Oettinger. Available from Clearing House for Federal Scientific and Technical Information, Springfield, Va., as document PB 169 252 ($3). See also Oettinger below.(Alias NSF-17).
1966. "Stative adjectives and verbs in English." "A note on negation." Both in NSF-17.
1967. "Pronominalization, negation, and the analysis of adverbs." ms.
1966. "Deep and surface grammar." ms. Also several of these papers currently available from Linguistics Club, Linguistics Department, Indiana University.
1968. "Counterparts." Read at LSA. Mimeo.
nd. "Pronouns and reference". ms.
1969. "Some semantic considerations in syntax." In Binnick and al.
add:
1968. "Repartee". To appear in Foundations of Language.
1968. "Instrumental adverbs and the Concept of Deep structure." in Foundations of Language.
1969. "Presuppositions and relative grammaticality." Read at LSA in 1968.
to appear. "On Generative Semantics" in Jakobovits and Steinberg.

Lakoff, George, and Peters, Stanley.
1966. "Phrasal conjunction and symmetric predicates", in NSF-17.

Lakoff, George, and Ross, John R.
1966. "A criterion for verb phrase constituency." In NSF-17.

Lakoff, Robin T.
1968. Abstract syntax and Latin complementation. MIT Press.
1968. "Some reasons why there can't be a some-any rule." Read at LSA, mimeo.
1969. "Syntactic arguments for not-transportation." In Binnick et al.
to appear. Review of Grammaire Générale et Raisonnée, 1660. To appear in Lg.

Langacker, Ronald.
1968. Language and its Structure. Harbrace.
to appear. "On pronominalization and the chain of command", mimeo 1966, to appear in an anthology by Reibel and Schane to appear.

McCawley, James D.
nd. "The annotated respective." Mimeo.
1968. "On the role of semantics in a grammar." In Bach & Harms.
1966. Review of Cooper's Set Theory and Syntactic Description. in Foundations of Language.
1968. "Concerning the base component of a transformational grammar." in Foundations of Lg.
1969. "A note on multiple negations." mimeo.
1969. "English as a VSO Lg." mimeo, read at LSA 1968.
1969. "Tense and time reference in English." Read at Ohio State Semantics funfest; to appear in Working Papers in Linguistics 4 of OSU Ling. dept.; mimeo.
add:
1967. "Meaning and the description of lg." Kotoba no Uchu, Tokyo, nos. 9,10,11.
1969. "Semantic representation." Read at Symposium on Cognition etc.
1968. "On lexical insertion in a transformational grammar without deep structure." In Darden et al.

Morgan, J. L.
1968. "Some strange aspects of "it"." In Darden et al.
1968. "On the notion 'possible lexical item'". Read at LSA. Ditto.
1969. "Irving." Dittoed.

1968. "Three notes on Irving and other matters." Ditto.
1969. "On arguing about semantics." Ditto, read at SECOL.
1969. "On the semantic representation of lexical items." In Binnick et al.

Oettinger, A. ed.
1966. NSF-17. Available from Clearinghouse as document PB 173 630 ($3).

Perlmutter, David.
1968. The two verbs 'begin'." to appear in an anthology by Jacobs and Rosenbaum to appear.
1968. "Deep and Surface Structure Constraints in Syntax." MIT diss.

Postal, Paul.
1966. "On so-called 'pronouns' in English." in Georgetown University Monography Series on Languages... 19, 177-206.
1968. "Cross-over constraints." ms.
to appear-a. "On coreferential complement subject deletion" in Jakobovits and Steinberg.
to appear-b. "On the derivation of surface nouns." in Linguistic Inquiry /a new journal./

1969. "On 'remind'", read at Ohio State.
1969. "Anaphoric Islands." in Binnick et al.

Quang Phuc Dong.
1968. "English sentences without overt grammatical subject." Mimeo.
1968. "A note on conjoined NP's", mimeo.

Ross, John R.
1966. "A proposed rule of tree pruning."
"Relativization in Extraposed Clauses."
Both in NSF-17.
1967. "Auxiliaries as main verbs." Dittoed.
1966. "On the cyclic nature of English pronominalization." in Jakobson Festschrift, Mouton.
1966. "Adjectives as NP's". Dittoed.
1967. "Gapping and the order of constituents." PEGS.
1968. "On declarative sentences." Dittoed.
1967. "Constraints on variables in syntax." Diss.
1969. several forthcoming papers.

Vendler, Zeno.
1967. Linguistics in Philosophy. Cornell Univ. press.

Wall, Robert.
1967. "Selectional restrictions on subjects and objects of transitive verbs." Mimeo.

Zwicky, Arnold.
1968. "Naturalness arguments in Syntax." In Darden et al.

EXTENSION AND MODIFICATION OF THE THEORY

Paul Postal, in a 1964 paper, "Underlying and superficial linguistic structure", seemed to rule out any principled approach to the study of performance. But it seems clear to me that performance has merely been a catch-all term used by linguists with a lot of nasty facts on their hands they had no way of handling. In section II I mentioned the treatment of semi-grammatical sentences as they used to be called. Now I think we should be able to treat so-called sentence fragments as being part of language proper. I see no reason, once we get over our hang-ups with sharp categorization of grammaticality and judgments of grammaticality in null context, why we cannot have a principled treatment of sentence fragments.

Another area usually relegated to Never-never land is that of the structure of discourse. Obviously the sentence pairs

Harry is a fool. He voted for Richard Nixon.
He voted for Richard Nixon. Harry is a vote.

are not equivalent. Imagine if we take every other sentence on a page, say, the beginning of Matthew 2. The result is hardly a well-formed discourse.

> Now the birth of Jesus came about in this way. But her husband, Joseph, was an upright man and did not wish to disgrace her, and he decided to break off the engagement privately. "Joseph, descendent of David, do not fear to take Mary, your wife, to your home, for it is through the influence of the holy Spirit that she is to become a mother." All this happened in fulfillment of what the Lord said through the prophet.... But he did not live with her as a husband until she had had a son, and he named the child Jesus. "Where is the newly born king of the Jews?"

To now, it has generally been held that the structure of discourse is linear, that is, sentences are strung together one after the other and well-formedness is based on how well these sentences string. But the context is vital to the form of a sentence. Similarly, whether two clauses are united or put into separate sentences depends on context: by context we cannot mean merely the two sentences on either side of the sentence in question, nor can we mean the n sentences to either side. This is quite as mad as the folly of the early 50's that syntax was a matter of which words had what probability of occuring n words to either side of a given word. What we need is a grammar, a generative grammar, a transformational grammar, of discourse, based on the same methods that have been developed in syntax over the last decade. This work was pioneered by George Lakoff's 1964 study of Russian folk-tales, in which he revised Propp's phrase structure grammar of the "morphology" of Russian folk-tales. I subsequently re-modified Lakoff's work and programmed it in COMIT for a 7090-7094 machine to generate plot outlines of Russian folktales. The results were partly abominable and partly amusing, but the point is that while hardly any discourse is as steretyped as Russian folktales or US patents, that certain structures nonetheless occur which are larger than the sentence. The notions of subordination and coordination of sentences and even whole discourses are quite valid and quite amenable to investigation.

A third class of problems concern logic. The implications of a sentence may be quite as important as the statements made by it. We linguists are only beginning to investigate presupposition, implication, insinuation, assertion, etc., but philosophers have been aware of these problems for a long time and a large literature exists. We want a machine to get as much information out of a sentence as a human would.

A fourth class of problems concern memory. Any program must involve knowledge. Humans do not use language in vacuo. Suppose I know that Sherlock Holmes is a tall, thin man. Suppose further that a fat, short man comes up to me and tells me he is Sherlock Holmes. If my memory and logic components are going full blast I immediately suggest to the gentleman that a)he is either lying, or b)could use a good psychiatrist, or c)he has a bad sense of humor. We would not like the computer to read a sarcastic sentence, such as "Surely they have a right to do unto others what they would not want others to do unto them" and file it away neatly. We need to give the computer a certain amount of linguistic sophistocation as far as irony, insinuation, and such go. This might seem overly optimistic, since most human beings lack this ability, but let me suggest that the goal of computational linguistics is to understand human capabilities, not reproduce them, something which can be done far cheaper by producing new human beings thru natural means than producing software in our labs. The only thing keeping us from programming computers to, for example, have a sense of human, is our peculiar delusion that we can't do it.

So these are the problems that have not been the subject of serious research. Note that I do not mean by this that no one has ever looked at them and found anything out. Even Newton was Platonian enough to realize that nothing new is ever discovered under the sun. But no linguist operating in terms of a formalized or quasi-formalized system has studied these problems very much. This is not to say that certain conclusions about the future construction of a theory of language use cannot be drawn from our present ignorance. The rest of this section will be devoted to how we with our Neanderthalic knowledge of language can outline a decent formal theory of 'la parole', something that we would want to do, I think, even had the computer never been invented. (End of sermon.)

One question which arises here is what the nature of underlying semantic structures is. Do people think in trees? McCawley in his article on the base rejected the notion of derivation. Instead he instituted a system of "node-admissibility conditions". These are actually conditions on the well-formedness of trees. Any object meeting these requirements is a w ll-formed tree, otherwise it is not (although I have yet to settle in my own mind whether an ill-formed tree is still a tree, just as I have been confused about whether an ill-formed sentence of English is still a sentence of English at all.) Each NAC has the form

$$< A;\ B\ C >$$

which is read, "a node A is admissible if it immediately and exclusively dominates a node labelled B and a node labelled C." NAC's generate trees directly, as opposed to rewriting rules which, in Chomsky's system, first go through a derivation, from which trees are then constructed. But the important point here is "Grammars are written by fools like me, but only God can make a tree": meaning that linguists need not concern themselves with the origin of trees to discover their properties.

Of course, if we are to be manipulating semantic structures, we are going to have to be concerned with where trees come from. A more basic question is whether the kinds of trees generative semantics claims to be semantic are reasonable semantic structures, that is, whether the investigator in artificial intelligence, for example, could live with them. I think there is a very good chance that this is the case. The basic elements of these trees are as follows. We have referential indices referencing individuals. I think that in any system we will need a device such as this. Both these indices and larger entities called senetences or S's can be dominated by the category N. I think again that any system will heed to consider sentences recursive in this way. Then we will need predicates of arbitrary "weight", tho' in natural language the number of N's associated with any predicate V will undoubtedly be rather small. One possible counter to this is obviated if we assure that we have ways of referring to sets. Then we can define S as a V and associated N's. This is not really a bad scheme.

Where it does fall down is in its failure to reflect hyper-propositional relations. The conceptual universe of a person is not a bunch of unrelated trees or sentences (propositions). We will want ways to connect the Napoleon of "Napoleon ate cheese" with that of "Napoleon hated Elba". Thus the conceptual universe is a network, with a far more complex structure than our underlying semantic trees. We therefore need some set of rules for isolating part of this network to serve as the underlying tree for some surface sentence or set of sentences, since it may turn out from our study of the structure of discourse that the unit of generation is larger than the sentence.

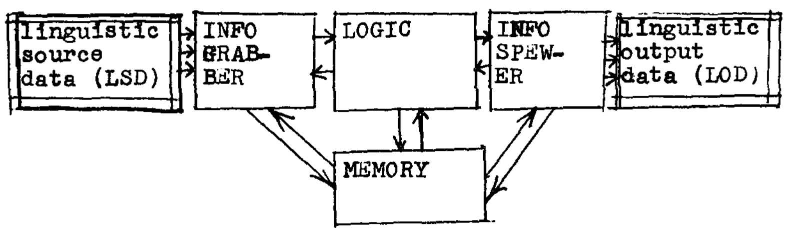
More will be said on these matters in section IV.

PART IV. The understanding machine.

One basic goal of research into computational linguistics might be to investigate how information is extracted from linguistic source data. (ultimately this ties into such questions as that of automated abstracting.) That component of our projected understanding machine which will model the information abstracting process let us dub the "info grabber".

The info grabber of course is not isolated. It will have to be connected with a logic component and a memory with which it will interact.

Nor is this the whole picture. As shown below one needs also a way of encoding the semantic output of the logic component for later output as linguistic data. Therefore the whole system will look like:



Notice that I have dubbed that component which synthesizes the LOD the "infox spewer".

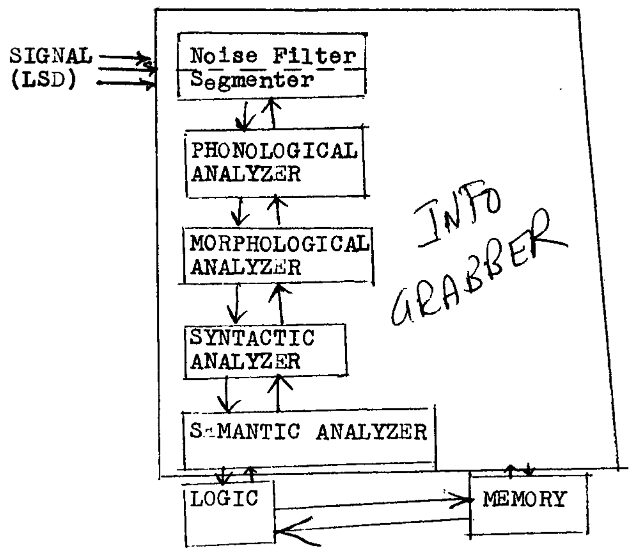
We can regard the above as a reasonable model not only of an understanding machine, but of the speaker. The above model would certainly be of use in the study of the use o f a natural language as a computer input-output language both for programming and for other applications, such as interaction between student and teaching machine in an educational program. I have made some study of such a system, which I called EASIOL (English as an Input-output Language), taking into account the results of the two studies I know of which approximated what I was after, namely Daniel Bobrow's STUDENT program, reported on in "A Question Answering System for High School Algebra Word Problems" Proc. FJCC 25, 1964, and in Scientific American September 1966, pp. 252-260, which was Bobrow's research for his doctorate. Bobrow modified LISP in the direction of COMIT, calling the hybrid METEOR. The system he evolved has a fair amount of flexibility and generality, and can deal with many kinds of problems expressed in stylized language. I might criticize Bobrow for his naivete over natural language, but since I am even more naive about information processing I will not do so.

A second system which I have heard later evolved into a more general system, is the BASEBALL program reported on by Green, Wolf, Carol Chomsky, and Laughery in the Feigenbaum-Feldman volume, Computers and Thought. This system bases itself on a rather stylized type of data structure. I have not followed the progress of either of these projects, but both betray inherent faults that made it unlikely that either could form the basis for a more general system operating on actual discourse. Nonetheless, these systems are very convincing for those who think that language is the sacrosanct birthright of human beings and that computers will never be able to handle such tasks as writing abstracts of articles.

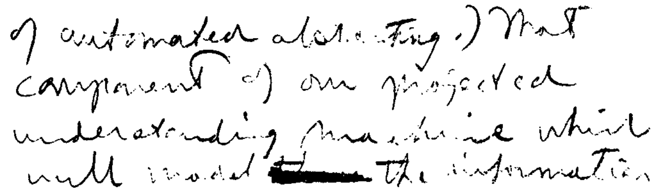
The above model is also a reasonable model of human speakers (if we forget that people differ from machines in essential ways -- vive la différence!)

The first part of the "info grab" is the read-in. Hopefully this will someday be done by the machine itself, via optical reader or speech analyzer. I think that research on readers and analyzers has in general been unhappy because of a failure to realize how complex recognition by humans is. Recognition is not simply an optical or auditory problem. All levels of language must interact in the process: it is well-known that real speech is more easily handled than approximants to speech, this can only be due to the recognition process being cyclic and operating simultaneously on all levels. The simplest recognition routines would involve something like:

IV-3



Indeed, we have to connect up the logic and the memory to this system. Below is a real sample of my handwriting when writing rapidly.



No recognition routine, not even my own human one can at all times decipher this garbage. Redundancy is pretty near nil and such words as "of", "as", "a", and "or" tend to be homologous. What a human reader can't do, we can hardly expect a machine to do. But humans can guess from context what a word must be, and then see if the squiggle on the page is close enough. This involves both syntactic and semantic recognition, and if we ever want machine reading of handwriting, we must give the machine this capability. But suppose the reader still can't handle the writing? I suppose then we want it to get the logic component to intiate a question such as," What is that?"

That is, we want the computer to be able to go thru the whole set of levels. This will necessitate a much more complicated program than those around today, incorporating a greater amount of linguistic expertise, but undoubtedly it is necessary.

Let us assume that the info grabber has grabbed the info, it will have (1) to store this information in the memory, and (2) let the logic component examine the information. Suppose I know that Richard Daley is the mayor of Chicago, and I read in a Chicago newspaper that the Mayor of Chicago is the greatest man in the world. The LSD must somehow be so stored that I can retrieve from my memory the fact that Richard Daley is thought to be the greatest man in the world by that newspaper.

This raises the question of how to convert underlying semantic trees into subnets of the semantic network of which memory probably consists. Many of the features incorporated into Sidney Lamb's conceptual networks will, I think, be incorporatable into the model. In particular, all occurences of a particular entity (concept) will have to be linked in some way or identified.

In a sense info grabbing starts by analyzing the LSD into semantic structures, and ends by synthesizing these structures and those already in memory into a new memory network.

One point that should be made clear is that all information will have to be represented on the same level. That is, both the program and the data will reside in the same memory net, as in a computer. Reading an algorithm in a book, the machine will store this in its memory just as it stores part of its own program, and it will be able to either quote the algorithm later as linguistic material as part of info retrieval, or use that algorithm as part of its own logical operations. There is some question tho as to whether this quite ideal machine could actually function in this way. But human beings are like this in some ways, and it is part of their language capability that they should or could.

The process of info spewing is a reverse of the info grab. The logic component will initiate the spew, using part of the memory net and selecting one or more underlying trees to spew out. It will then go through the derivational process and ultimately generate an actual string of sentences. Perhaps feedback will enter here, so that the machine can utilize part of its own spewings as immediate LSD, although it is hard to see why the machine would need to do so, altho humans are constantly correcting themselves mid-sentence.

An obvious question is what the role of generative semantics in all this. I think the experience of CL has been in general that ad hoc programs don't work. We need a basic linguistic theory. I think generative semantics is the best bet. But as I noted, it is a theory of cometence. We will need to modify it. I think we need to

1) admit rules of non-recoverable deletion,

2) admit rules for hypersentential constructs,

and

3) build strong interactions with logic and memory components.

In particular, the relationship of underlying semantic structures to conceptual networks will have to be investigated in depth.

If the hypotheses of the GS linguists are correct, then we have a simple but powerful basis for programs directly transforming language source materials into semantic information usuable by programs. For example, if the semantic structures turn out to be universal, they can serve as a pivot or intermediary for the currently out of fashion goal of MT.