Constraints on Sentence Complexity: A Model for Syntactic Processing

by

Elizabeth Ann Cowper

B.A., McGill University, 1972

A.M., Brown University, 1974

Thesis

submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in the Department of Linguistics at Brown University

June, 1976

This thesis by Elizabeth Cowper
is accepted in its present form by the Department of
Linguistics as satisfying the
thesis requirement for the degree of Doctor of Philosophy

Date: May 12, 1976 .... Henry Kucera ....

Recommended to the Graduate Council

Date: 12 May 1976 .... W. Nelson Francis ....

Department of Linguistics

Date: May 11, 1976 .... Susan Kuno ....

Harvard University

Approved by the Graduate Council

Date: ....

CURRICULUM VITAE

Personal: Born August 9, 1952; Deep River, Ontario, Canada.
Single

Education: PhD., Linguistics, Brown University, 1976
A.M., Linguistics, Brown University, 1974
B.A., French Language and Literature, McGill University, Montréal, Canada, 1972
Diplôme de Langue et Lettres Françaises, Faculté des Lettres, Aix-en-Provence, France, 1971

Teaching Experience: Teaching Assistant in Linguistics, Brown University, 1973-1976

Instructor of English Language, Brown University, 1974 &1975

Monitrice de Français, National Research Council of Canada, Ottawa, 1973

Awards: Brown University Fellowship, 1972-1973

ACKNOWLEDGEMENTS

There are many people to whom I am indebted for helpful discussion, advice and encouragement during the writing of this thesis. First among them is my advisor, mentor and friend, Henry Kučera. Throughout my entire graduate career, he has never lost faith in me, although I am sure that at times I gave him reason to. The many, many hours which he devoted to discussing my work with me, and the constant moral support and inspiration he provided made this thesis possible. Susumu Kuno and Nelson Francis both gave willingly of their time to read and comment on various drafts, and to listen to my often incoherent ideas and speculations Thanks to them, many of those speculations finally made sense.

I also owe a special thank you to Haj Ross, who gave encouragement, hours of discussion and hundreds of counterexamples and suggestions -- far beyond the call of duty -- and to David Perlmutter, who succeeded where many had failed, in teaching me how to "do syntax". Among my fellow students, Jacque Russom and John Goldsmith were especially helpful in discussing the work with me and reading parts of the draft.

To all my friends at Brown and in Cambridge, especially Barbara Moslin, Ellen Kaisse and Emmanuel Serusiaux, who provided moral support throughout the year and most importantly during the last few weeks, and Mary Louise Kean, in whose office most of the work was done, thank you for being there when I needed you.

Finally, I would like to thank Michael Szamosi, who was there on the evening in the spring of 1975 when it all began. Without his patience and attention as he helped me sort out what I was trying to say, and think, this thesis might never have

happened.

for my parents

TABLE OF CONTENTS

Chapter 1. INTRODUCTION

One of the basic assumptions underlying work in generative grammar is that sentences in natural language can be arbitrarily long, and of unlimited complexity. The argument supporting this assumption is analogous to the proof that there is an infinite number of integers, each one of which is finite. Just as any integer can be added to any other integer to produce a larger entity which is also an integer, any grammatical sentence can be conjoined to any other grammatical sentence to produce a longer sentence which is also grammatical. There is thus no such thing as the longest grammatical sentence in a natural language. Just as the number of integers is infinite, the number of grammatical sentences is also infinite.

In order for a finite grammar to enumerate an infinite set of sentences, it must incorporate the property of recursion; in other words, the output of some rule must at some point be able to serve as input to the same rule, thus creating a so-called loop in the derivation. The following grammar fragment, for example, is recursive.

(1) (a) S → NP VP

(b) VP → V NP

(c) NP → S

(d) NP → DET (ADJ) N

The output of rule (a) can serve as input to rule (c), whose output again serves as input to rule (a). A special case of a recursive grammar is a grammar containing a recursive rule, such as rule (e) or rule (f) below:

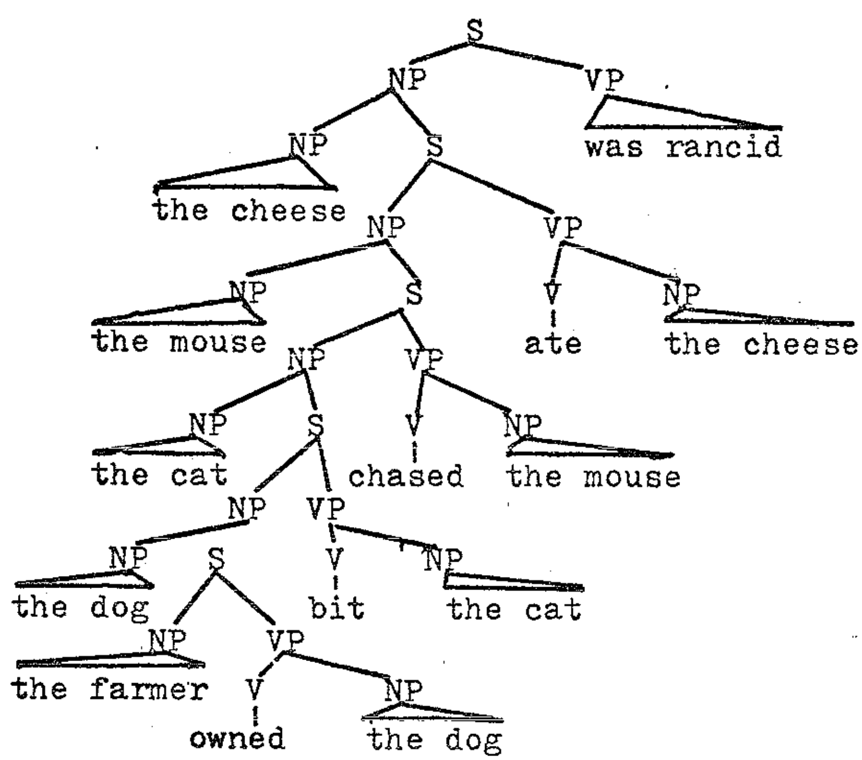
(e) NP → NP S

(f) S → S and S

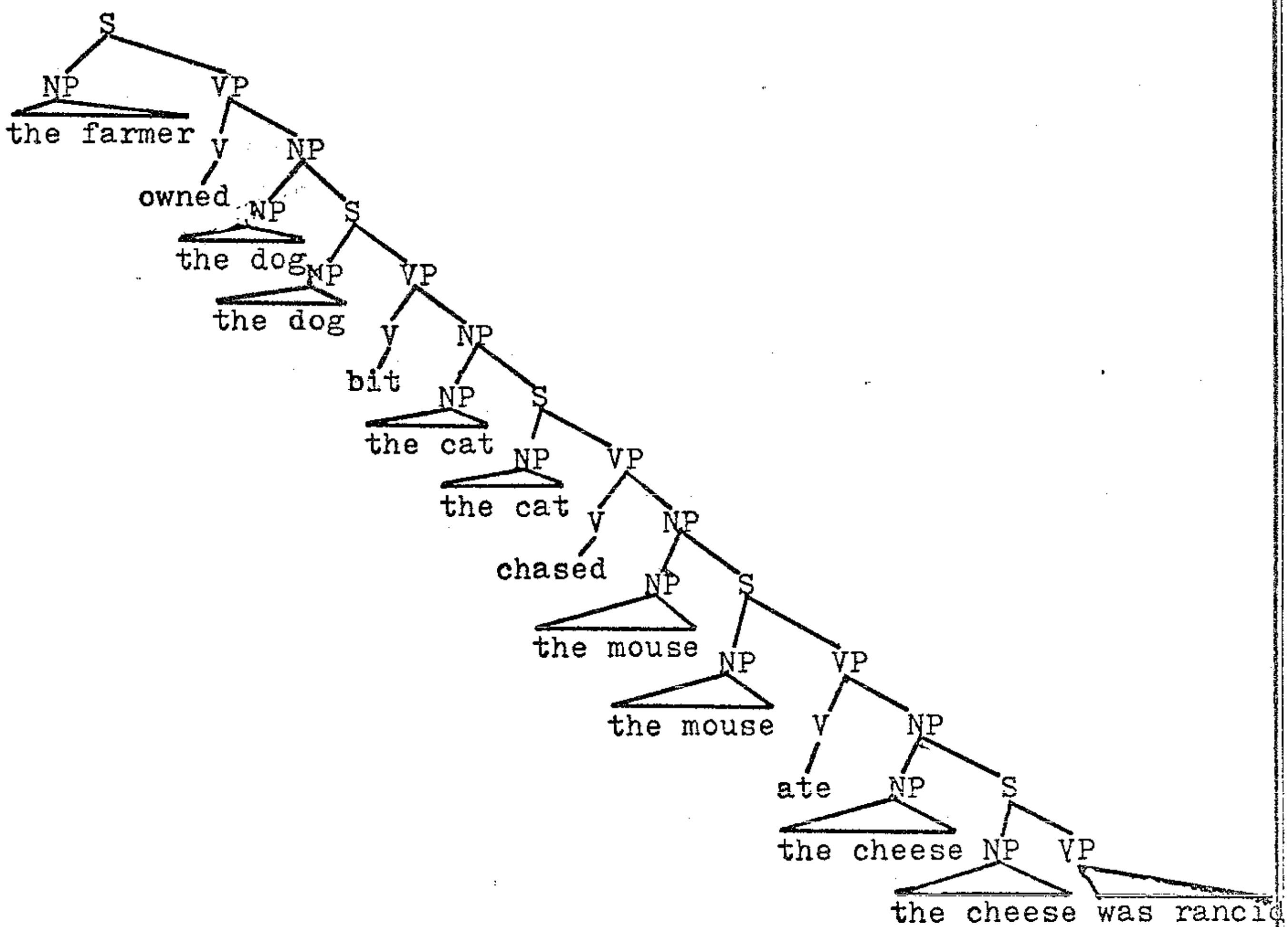
In this case, the symbol being rewritten, (<u>NP</u> in (e), <u>S</u> in (f)), shows up on the right-hand side of the same rule. Rules (e) and (f) can thus reapply immediately to their own output, without any other rules intervening.

Given that a grammar has the property of recursion, there is no limit to the number of times a derivation can loop. The following sentences, therefore, are both grammatical.

(2) The cheese that the mouse that the cat that the dog that the farmer owned bit chased ate was rancid.



(3) The farmer owned the dog that bit the cat that chased the mouse that ate the cheese that was rancid.



Although neither of these sentences is likely to be said by any native speaker of English, both are said to be within the speaker's linguistic competence; in other words, the speaker's internal grammar is capable of generating them. What prevents them from being actually produced has nothing to do with the language or the grammar of competence, but rather is due to limitations of the speaker, and to external factors. For example, no single human being could utter a sentence which would take three centuries to say. This inability, however, is due not to any characteristic of the language, but is a consequence of the limited life span of human beings. Other practical limitations on sentence length are imposed by the necessity for eating, sleeping, and in general doing other things besides speaking. Again, these limitations have nothing to do with language. There are many other factors such as interruptions, lapses of attention, slips of the tongue and false starts, all of which may cause a speaker's performance to deviate from the language of his competence.

Certainly, if one were attempting to give a coherent account of the grammar of a natural language, one would not want to be concerned with irregularities and idiosyncracies attributable to the above factors. Chomsky (1957) has restricted the domain of grammar by making the following idealizations:

The native speaker/hearer whose language we are concerned with lives in a completely homogeneous speech community, acquires language instantaneously, has no limitations on memory or attention, and is completely unaffected by external phenomena which might perturb his linguistic output.

As such, (2) and (3) above are both grammatical, even though real native speakers find (2) much less acceptable than (3).

The phenomena deemed irrelevant to the study of competence are basically of two types: those which are completely idiosyncratic and unpredictable, such as slips of the tongue, lapses of attention, and interruptions of various sorts, and those which are to some degree systematic and therefore predictable. The predictable phenomena can be further divided into two types: those which impose limits only on sentence length, and as such exclude sentences which take hours or days to say, and those which impose limits on sentence complexity. In this last group are the phenomena which are in all probability attributable to the limitations of short-term memory.

Chomsky (1965) has termed all of these phenomena performance factors, and claims that "investigation of performance will proceed only so far as understanding of underlying competence permits." (Chomsky, 1965, p. 10)

Implicit in the competence-performance distinction is the claim that it is possible to give a complete account of the grammar of a natural language without having to refer to any performance factors. In other words, it is not the case that performance factors will be needed in order to motivate any rules which occur in the grammar of competence. This is an empirical claim, and there is some evidence against it. Consider the following pairs of sentences:

(4a) Fred donated the money which his father's great-aunt Martha had left to him in her will to the United Way.

(4b) Fred donated to the United Way the money which his father's great-aunt Martha had left to him in her will.

(5a) That Mr. Jones always mows his lawn early Sunday morning for the sole purpose of annoying his neighbors is obvious.

(5b) It is obvious that Mr. Jones always mows his lawn early Sunday morning for the sole purpose of annoying his neighbors.

(6a) A truck carrying thirty-five tons of garbage to the dump outside of town hit my car.

(6b) My car was hit by a truck carrying thirty-five tons of garbage to the dump outside of town.

(7a) I found several messages and a huge pile of assignments to be corrected on my desk.

(7b) On my desk I found several messages and a huge pile of assignments to be corrected.

The (b) sentences above are derived from their (a) counterparts by the rules of Heavy NP Shift, Extraposition, Passive and Adverb Preposing. All of these rules are optional in these cases, and all of them have the effect of causing a so-called heavy constituent to occur at the end of the sentence. In all four cases, the (b) sentence is somewhat easier to understand than the untransformed (a) sentence. Since a grammar of competence is not concerned with how easy sentences are to understand, nothing can be said which would capture the common effect of these rules. In such a competence grammar, it would therefore be only a coincidence that there are at least four rules in the grammar which can cause a heavy constituent to occur at the end of the sentence.

Now consider (8a) and (8b) below, as compared to (4a) and (4b) above.

(8a) Fred donated his money to the charitable organization that his family had founded and supported for four generations.

(8b) Fred donated to the charitable organization that his family had founded and supported for four generations his money.

In both cases, the (b) sentences are derived from the (a) sentences by a rule moving an NP to the end of its clause. Let us call the rule deriving (8b) from (8a) Light NP Shift. If both (4b) and (8b) were derived by the same rule, it could be called simply NP Shift. In a grammar of competence, there are no grounds for predicting that a language will have a rule of Heavy NP shift, rather than a rule of Light NP Shift, or a rule of NP shift.[1] In terms of competence, therefore, it is a coincidence that many languages do, in fact, have Heavy NP Shift.

It should also be noted that in a traditional grammar, all of the rules mentioned above are optional; that is, grammatical sentences result whether or not the rules apply. It can be shown, however, that in most cases one of the two alternatives is decidedly more acceptable than the other. (4) - (7) above are such examples, as are the examples below.

(9a) That John will win the race which he decided to enter at the last minute is certain.

(9b) John is certain to win the race which he decided to enter at the last minute.

(10a) To convince John that this book should not be published until next year will be easy.

(10b) John will be easy to convince that this book should not be published until next year.

In practice, then, the application of optional rules often is not random, but rather can be predicted by the effect of the application of the rule on sentence acceptability.

This is not to say that the idealizations made by those working on competence grammars are invalid. One might conceive of a whole series of levels of abstraction, any one of which defines a possible domain of investigation. The most abstract level would correspond to the traditionally defined level of competence, or perhaps to something even more idealized. At the least abstract level, one would factor out none of the irregularities of human linguistic behavior. Moore (1967) has also discussed this question, and notes the following:

> It is not the case that a choice has to be made between levels of abstraction. It is sufficient that their existence and its relevance to a total account of language is noted. There is nothing reprehensible in

> investigating an unconstrained creative system such as the grammar of competence turns out in part to be. But given that a goal, admittedly a distant one, of linguistic enquiry is a theory that more closely simulates as much as is known or discovered about language behaviour, a theory is preferred that not only provides a systematic account of the rule-governed creativity of language, but one that also accounts for predictable bounds on certain directions that that creativity may take. (Moore, p.40)

The idealizations defining the level of abstraction of the present study are as follows: We are concerned with the linguistic behavior of a speaker/hearer who lives in a totally homogeneous speech community, who acquires language instantaneously, and who is unaffected by interruptions and distractions. He never tires, or needs to eat or sleep, etc., and as such is quite capable of uttering a sentence of any arbitrary length. Crucially, however, this speaker/hearer has the same short-term memory capacity as a normal human being, and therefore cannot utter sentences of unlimited complexity. For example, (2) above is unacceptable at this level, whereas sentences like (3) are acceptable.

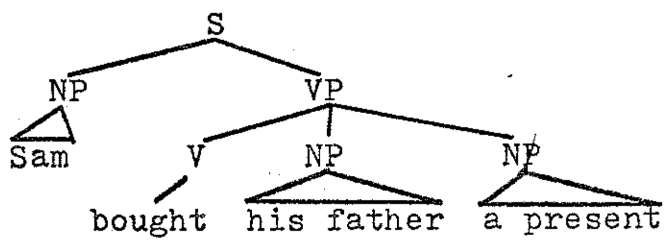
The aim of this thesis is to discover exactly what the limitations are on sentence complexity, and to construct a model to account for them. A priori, the only constraint on the model is that it must work in a way which reflects the temporal order of speech. We shall for convenience refer to this as working left-to-right. The reason for requiring that the model work left-to-right is that eventually, one would like to develop an actual account of sentence production, which would accurately reflect the way sentences are produced by human beings. No claim to psychological reality is made in this thesis, however; extensive experimental work would be required before any other such claims could be made. It should therefore be borne in mind that when statements are made about the speaker's capacity for processing sentences, what is meant is the capacity of the idealized speaker/hearer defined above.

The study is organized as follows. In chapter 2, an account is presented of three previous attempts to account for limitations on sentence complexity, which served as points of departure for the present work. A brief summary is then given of the major literature on the subject. In chapter 3, the first hypothesis is presented and shown to be inadequate. Chapter 4 develops the model in its entirety, using examples from English only. In chapter 5, another language type is introduced, and the model is shown to be adequate for these cases. In chapter 6, coordinate structures are accounted for. In chapter 7, several previously unsolved syntactic problems are shown to have simple solutions in terms of the model presented in chapter 4. Chapter 8 suggests some types of experiments which might be used to test the psychological reality of the model, and discusses the relationship between the model and accounts of competence.

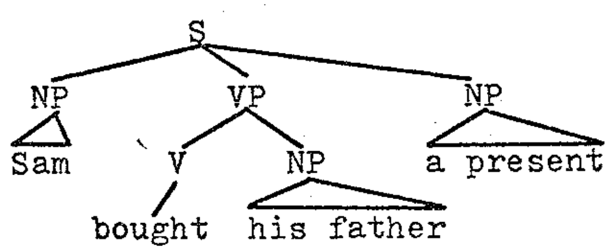
FOOTNOTES TO CHAPTER 1

1. Ross (1973) has proposed a rule of NP shift, which would derive (b) from (a) below:

(a)



(b)



This rule is needed to account for facts related to parentheticals. Since it does not change the linear order of constituents, it is unrelated to the rules under discussion here.

Chapter 2. DISCUSSION OF THE LITERATURE

2.0

Many scholars have worked on the problem of sentence acceptability, and on related problems, from many different perspectives. The extensive literature on machine translation and automatic parsing is relevant, since any automatic process which decodes sentences of natural language may provide some insights as to the problems faced by human listeners when they understand a sentence. There has also been a great deal of experimental work done which attempts to determine how syntactic processing works, and what syntactic factors affect the comprehensibility of sentences. In addition, several theoretical models have been proposed which take memory limitations into account.

This chapter is devoted to discussing in some detail the proposals of Moore (1967), Yngve (1960) and Kimball (1973), and a constraint proposed by Kučera (personal communication). Other literature on the subject of sentence acceptability is then discussed fairly briefly.

2.1

We have seen that one of the important characteristics of a traditional grammar of competence is that it allows for unlimited recursion. This allows sentences like (1) to be generated.

(1) #The prisoner that the lawyer that the judge that the President appointed reprimanded defended was convicted.

We have also claimed that an account of performance will have to provide some principled way of excluding such sentences.

Moore (1967) has proposed a model for doing this, which I will discuss insofar as it is relevant to this study. First of all, Moore claims that every simplex sentence has two major components: a topic (T) and a comment (C).[1] T is what the speaker intends the hearer to have in mind in order to understand the sentence; in other words, what the sentence is about. C is what the speaker says about T.

Topics in simplex sentences in English can be signalled in various ways. The following examples illustrate some of them.

(2) The library was destroyed by a terrible fire.

Here, the definite article in the subject NP shows that the library is known to the hearer, and the rest of the sentence tells the hearer something about the library.

(3) Fred? He's my brother.

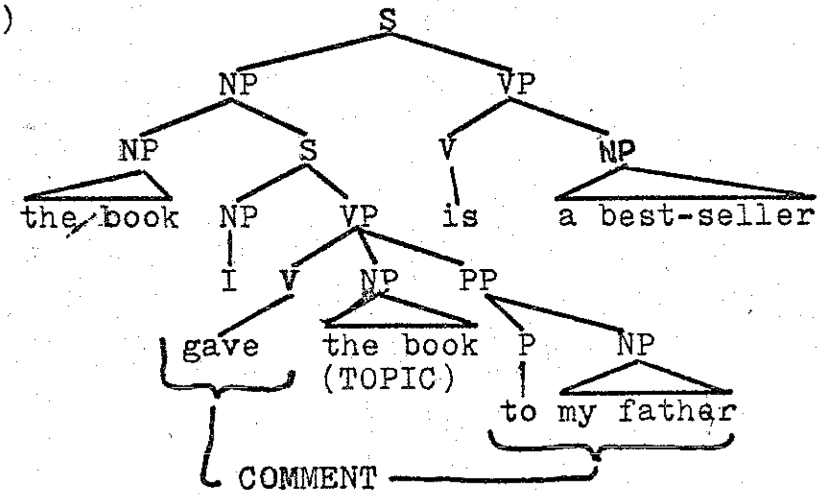
Clearly, this sentence is a reply to some sort of query about Fred.

In general, T seems to correspond roughly to the Praguian notion of <u>theme</u>, or <u>old information</u>. Moore refines the notion of topic, so that it is defined, not for each utterance, but rather for each clause. He deals exclusively with restrictive relative clause embeddings, and defines the topic of a relative clause as the NP which is co-referential with the head NP. In surface structure, then, the topic of a relative clause shows up as the relative pronoun. The comment is defined as everything in the clause except the topic. Clearly, C cannot be a deep structure node, since T is not necessarily either sentence-initial or sentence-final. C may therefore be discontinuous in underlying structure, as shown by the following example:

(4) The book which I gave to my father is a best-seller.

The underlying structure for (4) is represented in (5).

(5)



As we shall see, however, while the comment can be discontinuous in underlying structure, surface structure presents an entirely different picture. Consider the following sentences in terms of their topic-comment alignment:

(6) #The prisoner^{T1}/that^{T2}/the lawyer^{C2}/that^{T3}/the judge reprimanded^{C3}/defended^{C2}/was convicted.^{C1}/

(7) The judge^{T1}/that^{T2}/reprimanded the lawyer^{C2}/that^{T3}/ defended the prisoner^{C3}/was past retirement age.^{C1}/

(8) The prisoner^{T1}/that^{T2}/was defended by the lawyer^{C2}/that^{T3}/ the judge reprimanded^{C3}/was convicted.^{C1}/

(9) #The judge^{T1}/that^{T2}/the lawyer^{C2}/that^{T3}/defended the prisoner^{C3}/ was reprimanded by^{C2}/was past retirement age.^{C1}/

Notice that in the case of the unacceptable sentences (6) and (9), there is a disjunction in C2, whereas in the acceptable sentences (7) and (8), the T-C alignment is of the following form:

T1, T2, C2, T3, C3, ... Tn, Cn, C1

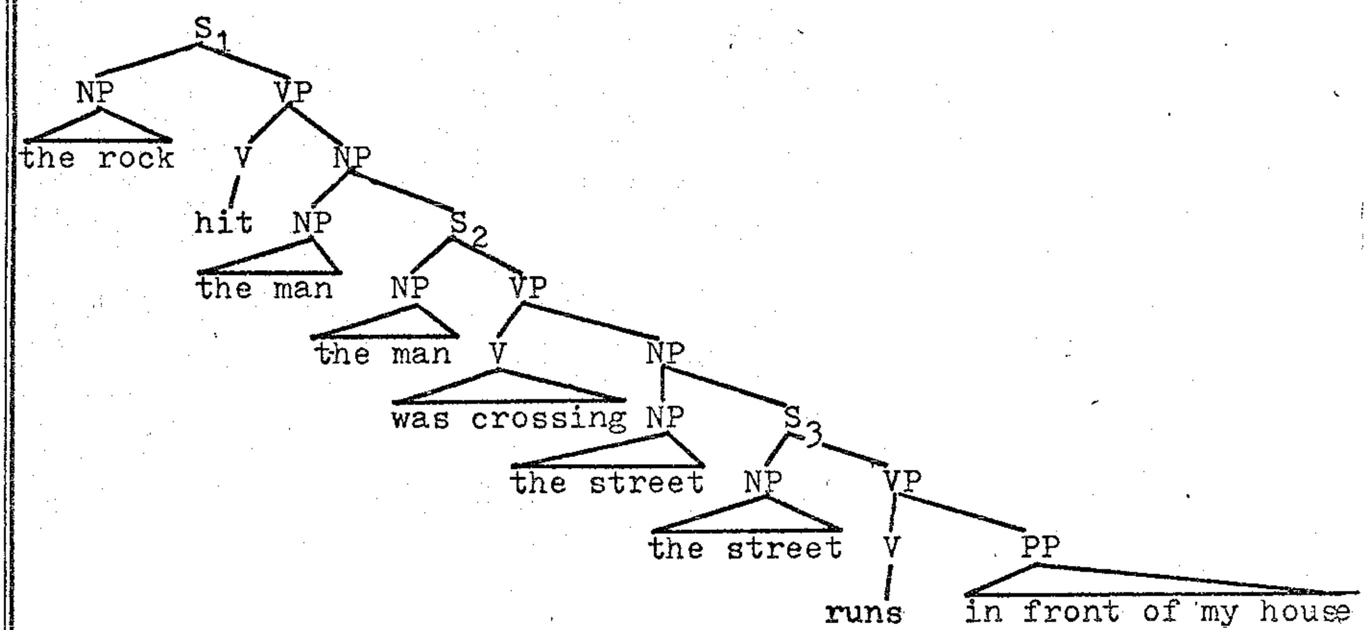
Moore claims that the unacceptability of sentences (6) and (9) is due precisely to the fact that the comments are disjoint on the surface. He gives many examples to support this claim, and I will not quote them here. (Moore, p.73ff) His next task is to provide some mechanism to prevent comments from showing up on the surface as disjoint. Since he wants to control the derivation to produce only acceptable sentences, rather than provide a filtering mechanism, or surface structure constraint, he must state his restrictions at the level of deep structure. As we have seen in (5), comment disjunction at the deep structure level does not necessarily result in surface structure comment disjunction. The statement, therefore, cannot take the form of a trivial condition on deep structures.

Moore's statement of the constraint is as follows:

> The configuration in S_2 of a comment functioning as a topic to S_3, followed by a copy topic, is an index that obligatorily requires the operation of the Passive transformation. (Moore, p. 73)

This statement is inadequate for two reasons, one of which Moore discusses. First, consider the following sentence, which is perfectly acceptable.

(10) The rock^{T1}/hit the man^{C1}/that^{T2}/was crossing the street^{C2}/ that^{T3}/runs in front of my house.^{C3}/



In this case, the comment in S_2 (street) is functioning as a topic to S_3, and it is followed by a copy topic. Clearly, nothing needs to be done to this sentence for it to be acceptable. In fact, if Passive operates in S_2, we obtain (11), which has comment disjunction in S_2, and is much worse than (10), if not completely unacceptable.

(11) #The rock^{T1}/hit the man^{C1}/that^{T2}/the street^{C2}/that^{T3}/runs in front of my house^{C3}/was being crossed by.^{C2}/

This problem can be solved if the original statement is modified as follows: (this formulation serves the same purpose as does Moore's (p. 75) but is simpler). The configuration in S_2 of a comment NP, which is also the subject of S_2, functioning as a topic to S_3, followed by a copy topic, is an index that obligatorily requires the operation of the Passive transformation.

The second problem with Moore's formulation, unfortunately, is not as easily solved. Notice that there are cases where the operation of Passive is required, but the structural description of Passive is not met:

(12) #The man^{T1}/that^{T2}/the girl^{C2}/that^{T3}/the dog was mad at^{C3}/ was fond of^{C2}/cried.^{C1}/[3]

There is a general convention that if the structural description for an obligatory rule is met, and that rule does not apply then the resulting sentence is ungrammatical.[4] (Lakoff, 1965 and Perlmutter, class discussion, 1975) This convention, however, will not help us here, since the structural description for Passive is not met at any point in the derivation. One might postulate an analogous principle, to the effect that if Passive is required by the T-C configuration, and Passive cannot apply, then the sentence will be unacceptable. This principle is less than satisfactory, since passivization is not the only means by which the correct T-C alignment can be achieved. In some cases, although not in the case of (12), Extraposition from NP can be used. Consider the following:

(13) #The book^{T1}/that^{T2}/the man^{C2}/whom^{T3}/the dog bit^{C3}/had^{C2}/was a best-seller.^{C1}/

(14) The book^{T1}/that^{T2}/the man had^{C2}/whom^{T3}/the dog bit^{C3}/was a best-seller.^{C1}/

Given the extreme simplicity of the major principle, -- comments must not be disjoint in the surface structure of a sentence -- and the complicated nature of the constraint Moore proposes, one is led to conclude that the problem might be more successfully handled if one were to look at it from another angle. Ideally, a constraint which exists solely to ensure the continuity of comments should be stated in terms of comment unity, rather than in terms of other facts which may turn out to be only accidentally related to comment disjunction.

Moore's model is also limited in that it deals only with relative clause embeddings. It is not clear from his work, for example, what the topic of the embedded sentence in (15) is:

(15) The proposal that the committee adjourn until the new year was accepted.

A model which accounted for the facts of restrictive relative clauses within a general theory of embeddings would clearly be more adequate.

2.2

Another way of stating Moore's principle has been proposed by Kučera (personal communication). This involves a global derivational constraint, as follows:

S_i -> (+Passive) if in P_n, C_i is disjoint, and $i > 1$.

where P_n is the surface structure obtained by applying the fewest transformations possible, and C_i is the comment in S_i, (following standard conventions for numbering S-nodes, with S_1 being the matrix.)

Kučera adds an additional statement to the constraint, as follows: If, in P_n, C_i and C_j are disjoint, and $j > i$, then a scale of acceptability obtains, as follows:

no passive < passive in S_j only < passive in S_i only < passive in S_i and S_j

where < means <u>less acceptable than</u>

This scale is illustrated by examples (16) through (19):

(16) ##The prisoner^{T1}/that^{T2}/the lawyer^{C2}/that^{T3}/the judge^{C3}/ that^{T4}/the President appointed^{C4}/reprimanded^{C3}/defended^{C2}/ likes peanut butter.^{C1}/

(17) #The prisoner^{T1}/that^{T2}/the lawyer^{C2}/that^{T3}/was reprimanded by the judge^{C3}/that^{T4}/the President appointed^{C4}/defended^{C2}/ likes peanut butter.^{C1}/

(18) ?The prisoner^{T1}/that^{T2}/was defended by the lawyer^{C2}/that^{T3}/ the judge^{C3}/that^{T4}/the President appointed^{C4}/reprimanded^{C3}/ likes peanut butter.^{C1}/

(19) The prisoner^{T1}/that^{T2}/was defended by the lawyer^{C2}/that^{T3}/ was reprimanded by the judge^{C3}/that^{T4}/the President

appointed[C4]/likes peanut butter.[C1]/

Kučera has further shown that his constraint, with a slight modification, can account for an analogous set of facts in the Slavic languages. Slavic languages have available the transformation of Word Order Permutation, which obviates the need to use Passive. Not surprisingly, non-reflexive passive sentences are relatively rare in the Slavic languages.[5] For Slavic, the constraint would be stated as follows:

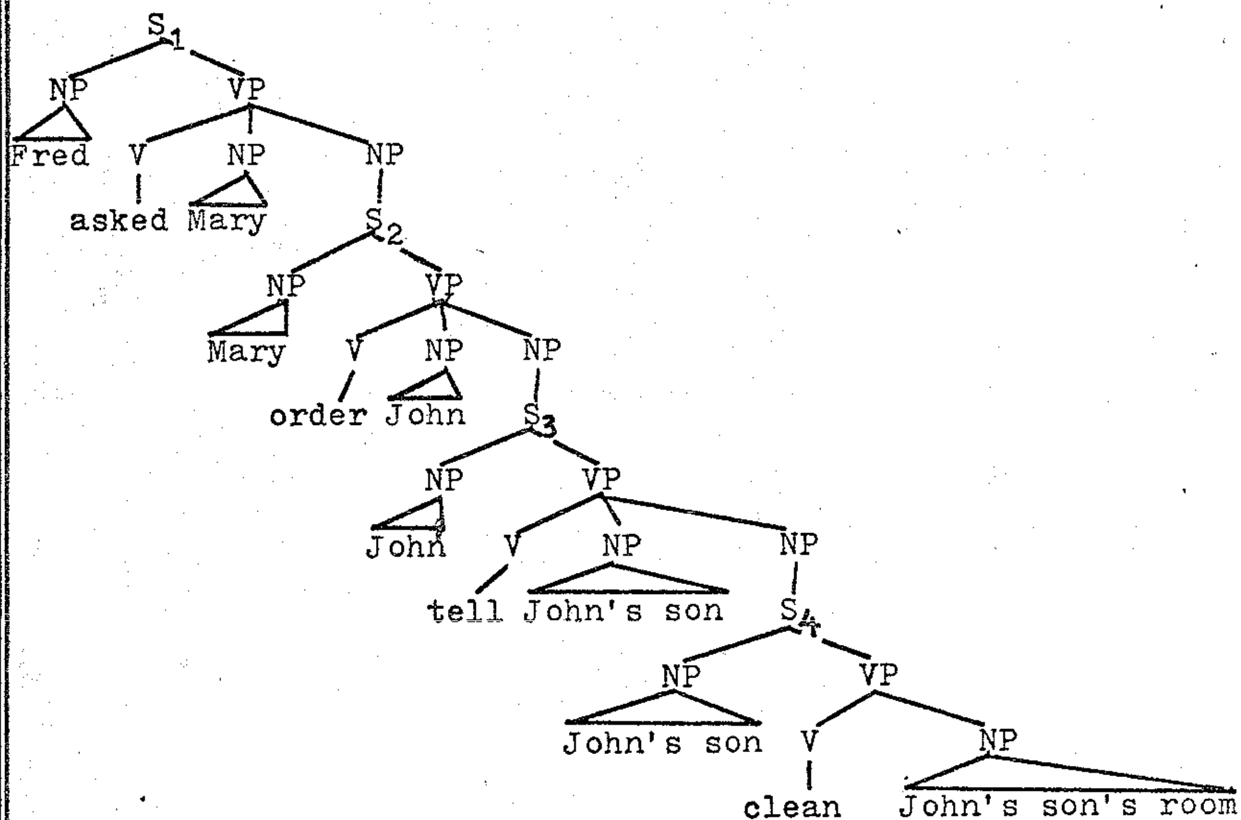
S_i --> (subject-final order) if, in P_n, C_i is disjoint.

The rest of the constraint would be exactly the same as the one stated for English.

This solution is more adequate than Moore's for the following reasons. First of all, it is stated in terms of comment disjunction, which is exactly what is at issue. Secondly, it has something to say about the varying degrees of acceptability of sentences (16) through (19), whereas Moore's model would predict that sentences (16) through (18) are completely unacceptable. On the other hand, one must ask the question of whether global derivational constraints can be incorporated into a model of performance which works in a left-to-right way. A model which generates surface structures directly, in a way which simulates the temporal order of speech, would not lend itself to a constraint which must be stated in terms of an entire syntactic derivation, complete with the cyclic application of rules.

Consider the following sentence:

(20) Fred asked Mary to order John to tell his son to clean his room.



Cyclic rules, working on this sentence, would apply first to S_4, which is in fact the rightmost sentence in the utterance. The rules would have to apply, then, in an essentially right-to-left order, the opposite of temporal order. A priori, this type of phenomenon seems to be something one would want to avoid in a left-to-right model of performance. I am of course taking no stand as to the merit of global derivational constraints in a grammar of competence. Nor am I sure at this point whether they will even be avoidable in a complete account of performance. So far they have not proved necessary, but it may later turn out that they are required.

2.3

So far, we have seen two proposed accounts for the unacceptability of sentences like (1). The third proposal which I shall discuss (Yngve, 1960), is somewhat more similar to the one which I will propose in Chapter 4, in that it was intended to generate surface structures directly, in a left-to-right manner. The original domain of Yngve's work was the field of machine translation, and his model, theoretically, could be applied to the automatic recognition and production of sentences. He proposes that sentences are generated from left to right, according to the expansion rules of a context-free phrase structure grammar. The leftmost node is always expanded first, and the nodes waiting to be expanded are stored in a "temporary memory" which is formally equivalent to a pushdown store. An example will more clearly illustrate the workings of this model.

Let us assume that the phrase structure grammar contains, among others, the following rules:

S --> NP VP

VP --> V NP

NP --> DET N

DET --> the[6]

N --> boy, man

V --> saw

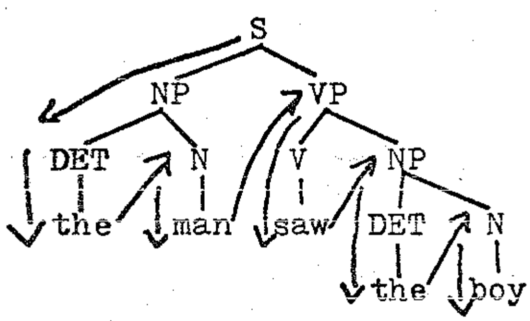
This model would generate the sentence <u>The man saw the boy</u> with an output like the following:

(21)

OUTPUT														REGISTER	STORE
														S	
													S	NP	VP
												S	NP	DET	N VP
											S	NP	DET	the	N VP
										S	NP	DET	the	N	VP
									S	NP	DET	the	N	man	VP
								S	NP	DET	the	N	man	VP	
							S	NP	DET	the	N	man	VP	V	NP
						S	NP	DET	the	N	man	VP	V	saw	NP
					S	NP	DET	the	N	man	VP	V	saw	NP	
				S	NP	DET	the	N	man	VP	V	saw	NP	DET	N
			S	NP	DET	the	N	man	VP	V	saw	NP	DET	the	N
		S	NP	DET	the	N	man	VP	V	saw	NP	DET	the	N	
	S	NP	DET	the	N	man	VP	V	saw	NP	DET	the	N	boy	
S	NP	DET	the	N	man	VP	V	saw	NP	DET	the	N	boy	--	--

The derivation thus proceeds through the tree in the following way:

(22)



Yngve then defines his "depth hypothesis" as follows: The depth (d) of a node _A_ is the number of items contained in the temporary memory at the point in the production when A is in the register. For example, d for _man_ in (21) is 1, since when _man_ is in the register, the store contains one item, _VP_. d for _DET_ in the subject NP, on the other hand, is 2, since when _DET_ is in the register, the store contains two items, _N_ and _VP_.

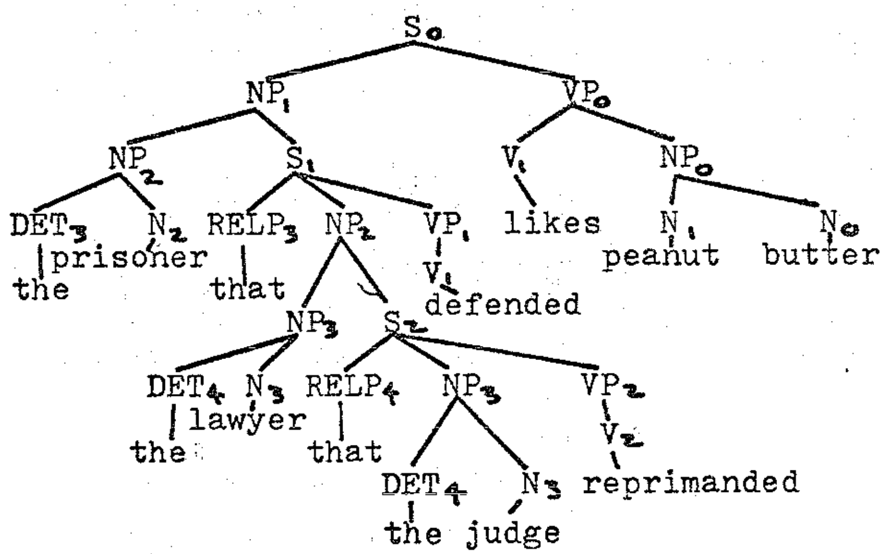
D, or d_{max}, is defined for a sentence as the maximum number of items contained in the temporary memory at any one point during the production of the sentence. In other words, D represents the storage capacity required for the production of the sentence. D for (21) is 2.

It can easily be demonstrated that in order to produce the set of well-formed algebraic expressions, a temporary memory of infinite capacity is required. Having done this, and given the obvious fact that human beings do not possess infinite memory capacity, Yngve asks whether his model, equipped with a temporary memory of some finite capacity, would be adequate to produce all the sentences of a natural language. He calls a grammar well-behaved if it generates a language all of whose sentences can be produced by a device with a temporary memory of finite capacity. If, in fact, the grammar of English is well-behaved, it is then interesting to ask what the capacity of the temporary memory must be. Given Miller's work on short-term memory (Miller, 1956) Yngve suggests that a capacity of 7 ± 2 items might be reasonable.

If Yngve's hypothesis is correct, then, the D-value will indicate the relative acceptability of a sentence, or in other words, how easy the sentence is to understand. Possibly, sentences with depth less than 7 ± 2 will be acceptable, while those with greater depth will be incomprehensible.

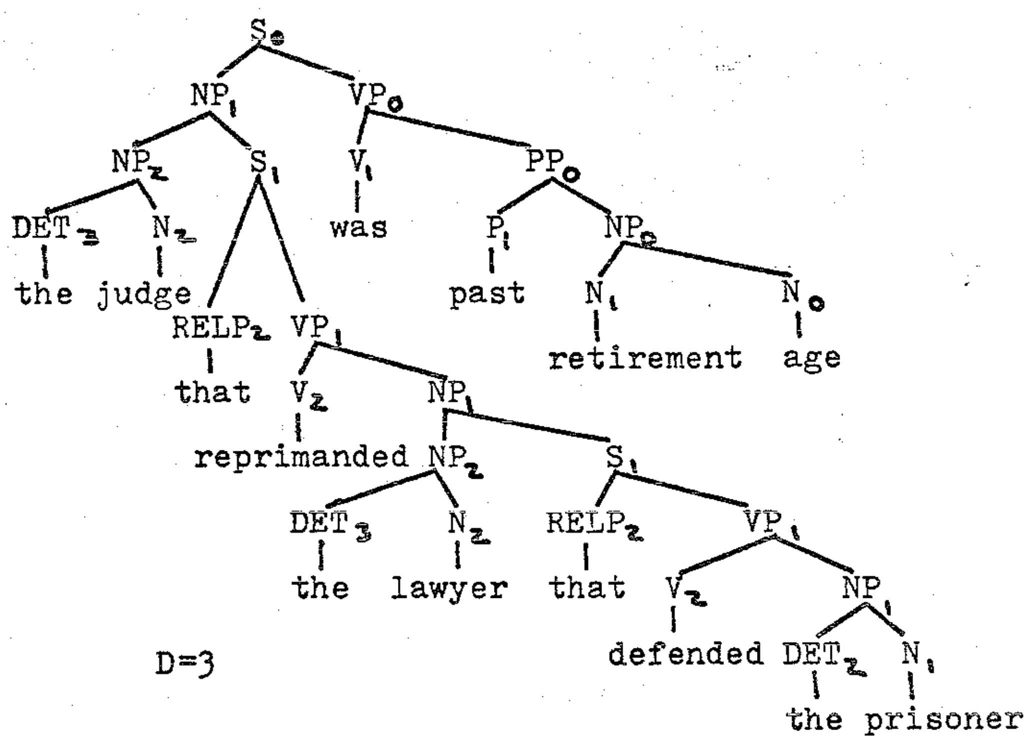
Let us test this hypothesis by examining some of Moore's examples, and some others, in terms of their depth. The subscript on each node indicates the depth (d) that that node would have if the sentence were produced by Yngve's model.

(23) #The prisoner that the lawyer that the judge reprimanded defended likes peanut butter.



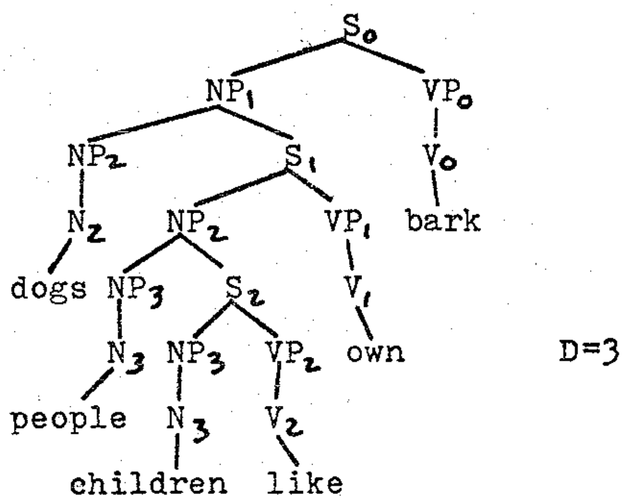
D= 4

(24) .The judge that reprimanded the lawyer that defended the prisoner was past retirement age.

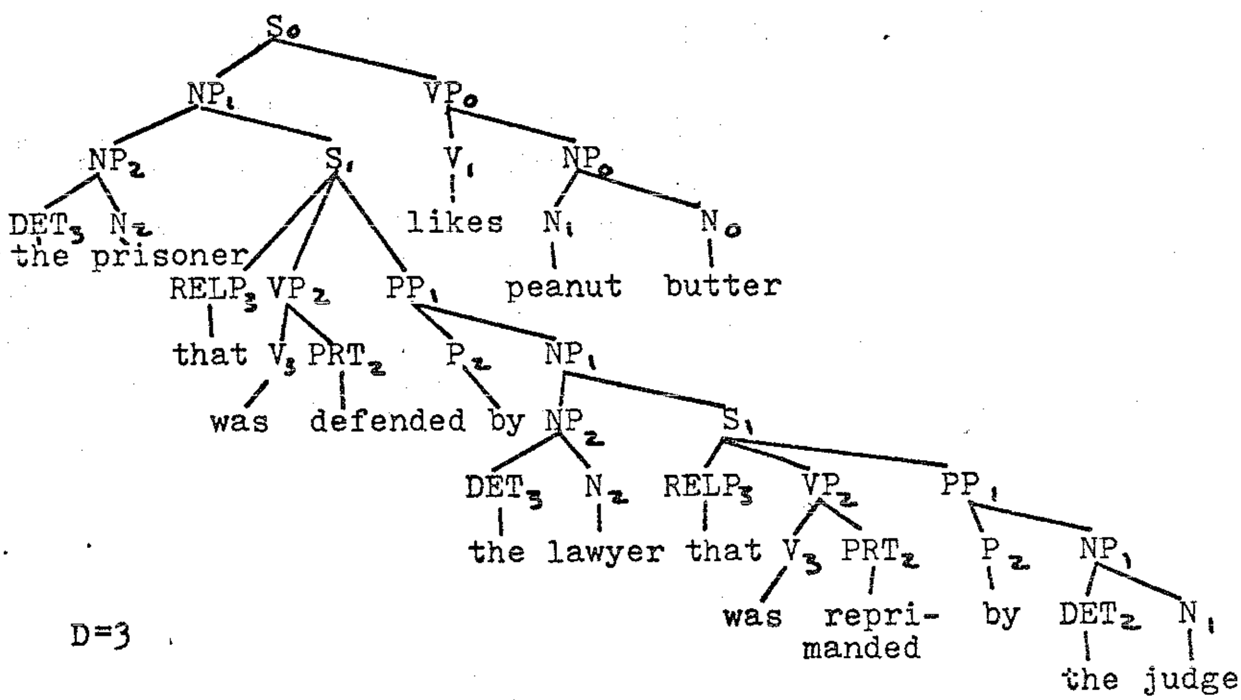


D=3

(25) #Dogs people children like own bark.



(26) The prisoner that was defended by the lawyer that was reprimanded by the judge likes peanut butter.



Notice that none of these sentences exceeds Yngve's tentative limit of 7 ± 2, and that (23) and (25) are both unacceptable. Clearly, then, 7 ± 2 is too high a limit. However, even if we decrease the depth limit to three, the predictions made by Yngve's model are still incorrect. The only difference between (23) and

(25) is the presence of determiners and relative pronouns in (23) and their absence in (25). The depth hypothesis predicts that (24), (25) and (26) should be of equal difficulty, and that all of them should be easier than (23). In fact, (25) is completely unacceptable, and perhaps even more difficult to understand than (23), whereas (24) and (26) are quite acceptable.

For a further test of the depth hypothesis, consider the following Japanese sentence, which is perfectly acceptable to native speakers.

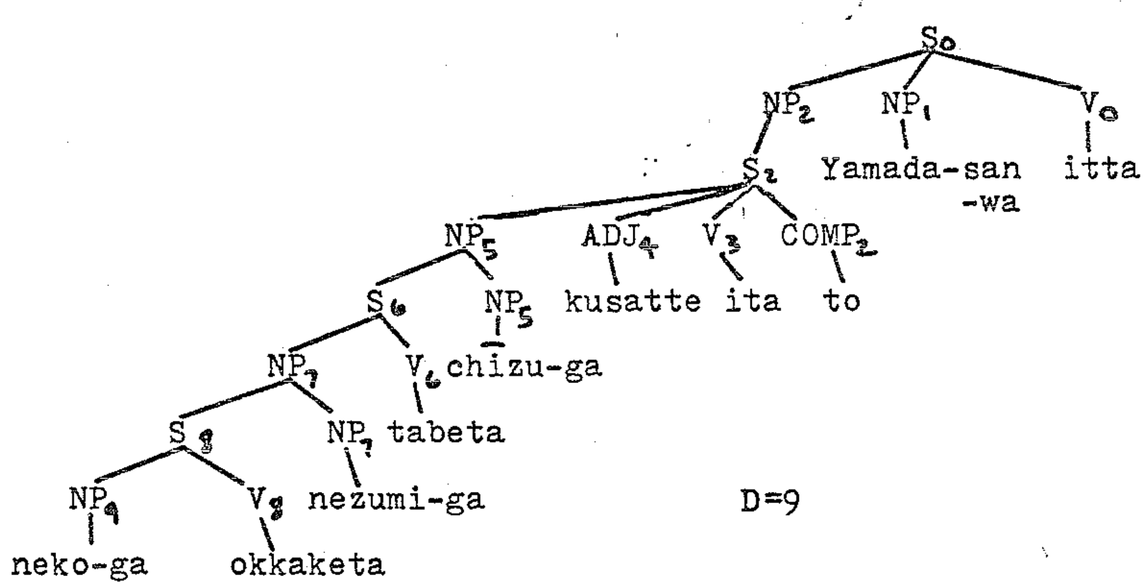
(27) [[[[Neko-ga okkaketa] nezumi-ga tabeta] chīzu-ga kusatte
cat chased mouse ate cheese rotten

ita to] Yamada-san-wa itta.][7]
was COMP Mr. Yamada said

(27a) Mr. Yamada said that the cheese that the mouse that the cat chased ate was rotten.

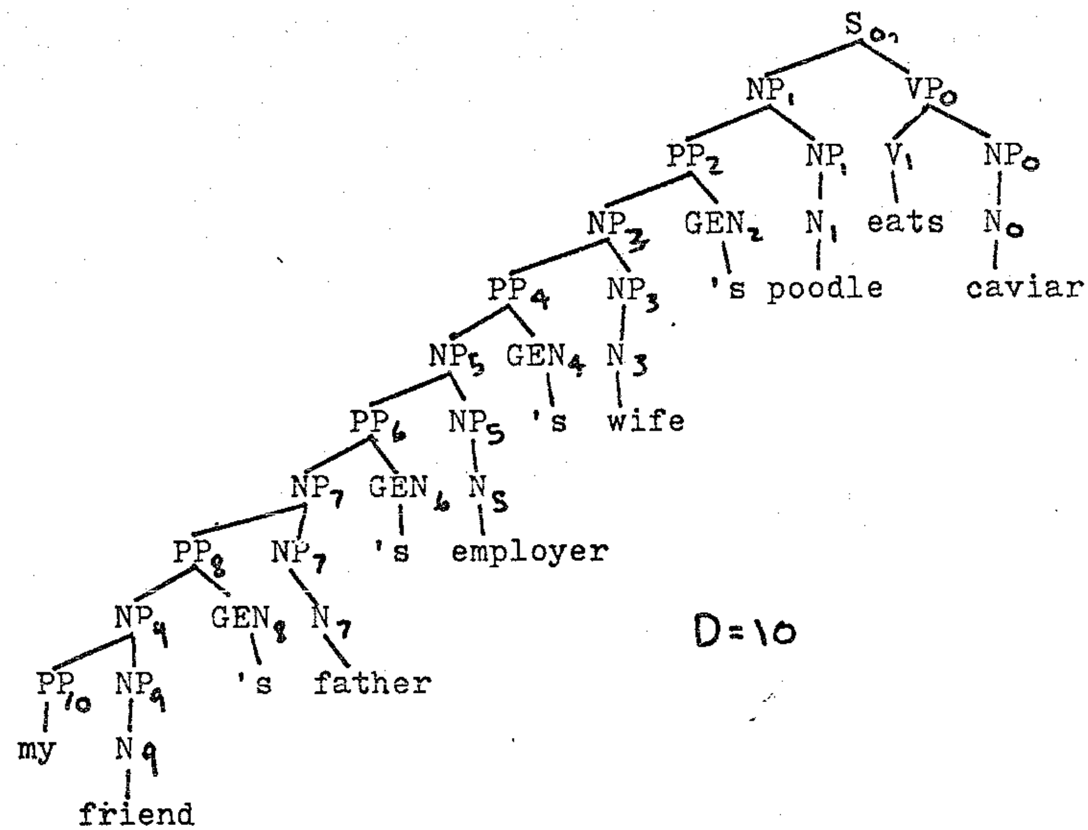
Unlike the English translation, which is center-embedded, the Japanese sentence has a left-branching structure, as represented in (28):

(28)

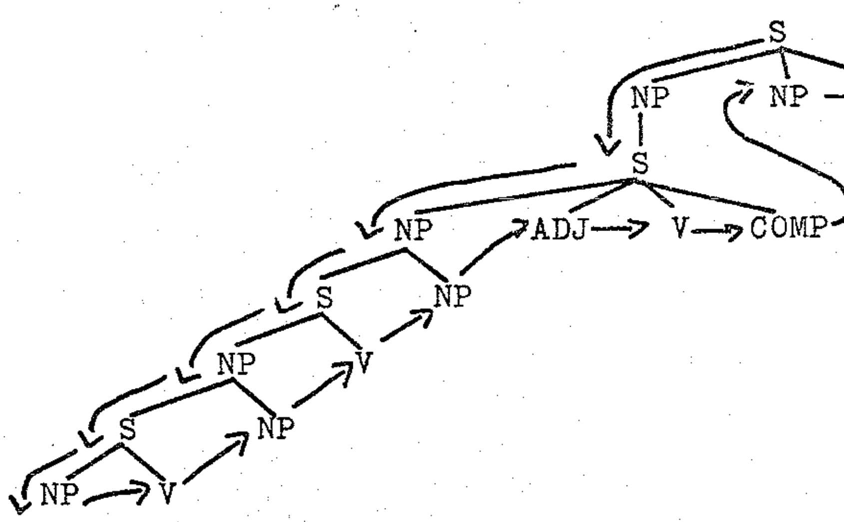


Given that the depth hypothesis is supposed to reflect a limitation on human memory, rather than a characteristic of a particular language, it must hold consistently for all languages. Clearly, any language which typically exhibits left-branching structures will provide counterexamples to the hypothesis.[8] Japanese and Turkish are two such languages, and English also has some left-branching structures, such as the one shown in (29):

(29) My friend's father's employer's wife's poodle eats caviar.



The reason that Yngve's hypothesis predicts such a high level of difficulty for left-branching structures is very simple: the point of departure for the calculation of depth is the matrix S-node, which, in a left-branching structure, is at the end, rather than at the beginning of the utterance. In Yngve's model, therefore, the production of (27) and (29) is not really left-to-right, but rather right-to-left-to-right, as shown below:



Any calculation of acceptability which uses the matrix S as its starting point will fail in the same way as does Yngve's. The model which will be developed in the next chapter, therefore, operates in a true left-to-right manner, starting with the first word spoken.

2.4

Kimball (1973) has proposed seven principles for parsing of natural language which are very similar to the initial hypothesis which I will propose in chapter 3. A point-by-point consideration of these principles will illustrate many of the problems involved in developing a left-to-right model of language.

<u>Principle one: Parsing in natural language proceeds according to a top-down algorithm</u>

Yngve (1960) also proposed a top-down algorithm, in this case for sentence production, and predictably, the problems with Yngve's model, discussed above, also present difficulties for Kimball. In a footnote, Kimball presents three ways of viewing the notion of Top-Down, or Over the Top Parsing:

> ...In fact, it may be suggested that the mechanism of parsing in fact utilised in natural language is this: Trees are not built down to single terminals but with regards to adjacent pairs of terminals (discriminant pairs). Given an initial member of a pair, a tree is built over-the-top down to the second member. This could be done in one of at least three ways: (1) The tree is built up only as far as the lowest common dominating node for the pair under consideration; (2) the tree is built up only as far as the lowest common dominating S node for the pair, and then down to the second member; or (3) the tree is built all the way up to the highest S node, and down to the second member. As I have given it in the paper, the parsing hypothesized for natural language corresponds to this third type of OTT parsing... (Kimball, p. 22)

This third way of parsing makes the assumption that the matrix S node is always determinable from the beginning of the sentence. This assumption seems to hold for English, since English sentences are in general right-branching or center-embedded. In the case of left-branching structures, however, the assumption is not valid, as shown by the following Japanese sentences:

(30) [$_S$ Neko-ga nezumi-o okkaketa]$_S$
cat mouse-acc chased

It was the cat that chased the mouse.

(31) [$_S$ [$_S$ Neko-ga okkaketa]$_S$ nezumi-ga chīzu-o tabeta]$_S$
cat chased mouse cheese-acc ate

The mouse that the cat chased ate the cheese.

(32) [$_S$ [$_S$ [$_S$ Neko-ga okkaketa]$_S$ nezumi-ga tabeta]$_S$ chīzu-wa kusatte ita.]$_S$
cat chased mouse ate cheese rotten was

The cheese that the mouse that the cat chased ate was rotten.

(33) [$_S$ [$_S$ [$_S$ [$_S$ Neko-ga okkaketa]$_S$ nezumi-ga tabeta]$_S$ chīzu-ga kusatte ita to]$_S$
cat chased mouse ate cheese rotten was COMP

Yamada-san-wa itta]$_S$
Mr. Yamada said

Mr. Yamada said that the cheese that the mouse that the cat chased ate was rotten.

(34) [$_S$ [$_S$ [$_S$ [$_S$ [$_S$ Neko-ga okkaketa]$_S$ nezumi-ga tabeta]$_S$ chīzu-ga kusatte ita to]$_S$
cat chased mouse ate cheese rotten was COMP

Yamada-san-ga itta to]$_S$ watasi-wa omoimasu]$_S$
Mr. Yamada said COMP I think

I think that Mr. Yamada said that the cheese that the mouse that the cat chased ate was rotten.

All of the above sentences are acceptable, and all begin in exactly the same way. From the listener's point of view, then, it is clearly impossible to know, from the first word, how many S nodes will intervene between the matrix S and the first word. Notice that only the third type of OTT parsing is inadequate in these cases; either the first or the second type can handle left-branching structures such as those in (31)-(34).

<u>Principle two: Terminal symbols optimally associate with the lowest non-terminal node.</u>

This principle is used to explain the preferred reading of (35) below:

(35) Joe figured that Sue wanted to take the cat out.

Here, <u>out</u> could be associated either with <u>take</u> or with <u>figured</u>, but listeners, upon hearing the sentence, naturally interpret <u>take the cat out</u> as a phrase. This principle seems to work for right-branching structures, but it is not clear whether it has anything to say about left-branching structures. Often, in the case of left-branching structures, a word will cause more nodes to be added at the <u>top</u> of the phrase-marker. The non-terminal with which the word will be associated often is not there until the uttering of the word makes it necessary for that non-terminal to be added. As such, the notion of lowest non-terminal node seems to have little meaning here.

<u>Principle three: The construction of a new node is signalled by the occurrence of a grammatical function word.</u>

Kimball claims that the absence of function words creates perceptual difficulties, so that (36) is more difficult than (37):

(36) The boy the girl the man saw kissed left.

(37) The boy who the girl who the man saw kissed left.

This claim may be true for some English relative clauses, but it is by no means universal. Japanese has no function word marking relative clauses, and the verb form in a Japanese relative clause is identical to the form used in matrix clauses. Also, in Japanese, complementizers are clause-final, so that the embedded sentence has already been uttered by the time the grammatical function word occurs. This is illustrated by (38).

(38) [$_S$Watasi-wa [$_S$Yamada-san-ga hon-o kaita to]$_S$ omoimasu.]$_S$
I Mr. Yamada book-acc. wrote COMP think

I think Mr. Yamada wrote the book.

In this case, the listener knows that <u>Yamada-san-ga</u> is in an embedded sentence because it has subject marking. Since the higher clause already has a subject, <u>watasi-wa</u>, <u>Yamada-san-ga</u> is assumed to be in another clause.[9]

Also, it is not clear that (39) is any more difficult to understand than (40).

(39) Joe said Peter thought Sue would leave by nine o'clock.

(40) Joe said that Peter thought that Sue would leave by nine o'clock.

<u>Principle four: Two sentences only can be processed at the same time</u>.

This principle is too strong, as shown by the following sentences, all of which are acceptable.

(41) [$_S$The proposal [$_S$that the man [$_S$who escaped]$_S$ should be shot]$_S$ was discussed.]$_S$

(42) [$_S$The book [$_S$that the man [$_S$who hired me]$_S$ wrote]$_S$ deals with politics.]$_S$

(43) [$_S$Watasi-wa [$_S$Yamada-san-wa [$_S$nezumi-ga chīzu-o tabeta to]$_S$
I Mr. Yamada mouse cheese ate COMP

itta to]$_S$ omoimasu.]$_S$
said COMP think

I think that Mr. Yamada said that the mouse ate the cheese.

Granted, most cases where more than two sentences are being processed at once are unacceptable, but one cannot ignore those cases where the sentence is acceptable. The model proposed in chapter 4 of this study will account for the above sentences.

<u>Principle five (Closure): A phrase is closed as soon as possible, i.e. unless the next node parsed is an immediate constituent of that phrase</u>.

This principle is roughly equivalent to the notion of <u>on the table</u> proposed in chapter 3.

<u>Principle six (fixed structure): When the last immediate constituent of a phrase has been formed and that phrase closed, it is costly in terms of perceptual complexity ever to have to go back to reorganize the constituents of that phrase</u>.

<u>Principle seven (Processing): When a phrase is closed, it is pushed down into a syntactic (possibly semantic) processing stage and cleared from short-term memory</u>.

All of these principles correspond to aspects of the hypothesis presented in chapter 3. That hypothesis was developed independently of Kimball's work. Both Kimball's principles and our hypothesis suffer from similar, rather severe, shortcomings, and the model presented in chapter 4 is intended to handle these problems.

2.5

There have been many other attempts to characterize the limitations imposed on sentence complexity by short-term memory. Kuno and Oettinger (1963) constructed an automatic model of natural language processing which utilized a series of push-down stores. Reich (1969) proposed a stratificationally-based model to account for limits on center-embedding. Unlike most treatments of this subject, Reich's article does not distinguish grammaticality from acceptability, and thus considers sentences (44) and (45) below to be both ungrammatical.

(44) *That Fred eats olives with peanut butter tends.

(45) #The house that the contractor that the architect paid built didn't satisfy the building code.

Langendoen (1976) shows that a limit on center-embedding allows a finite-state parser to be constructed for a language generated by a context-free phrase structure grammar. He also provides motivation for a reanalysis of surface structure such that in a left-branching or right-branching structure, a constituent is closed as soon as another constituent of the same type begins, even if the second is syntactically embedded in the first. This corresponds roughly to Kimball's principle of closure, and also to the concept of <u>on the table</u> which I will propose in the next chapter.

Thomas and Huff (1971) have done an experiment which shows convincingly that center-embedding is indeed what makes sentences hard to understand.

> A sentence will be difficult to the extent that the subject is compelled temporarily to ignore some parts of it (and perhaps to hold them in temporary storage) while dealing with other part. (p. 361)

They also sketch a model which is, in form, similar to the one which will be proposed in chapter 4. However, since their purpose was not to determine the capacity of hearers to process multiply embedded sentences, but rather to discover which factors increased processing difficulty, their model is extremely general.

Kuno (1972) proposes a model of sentence acceptability with a pushdown store of capacity 2, which accounts for the acceptability of right-branching and left-branching structures, and the unacceptability of center-embedded structures.

FOOTNOTES TO CHAPTER 2

1. The notation # is used to indicate that a sentence is grammatical, but unacceptable. * is used, as usual, to indicate ungrammaticality.

2. Hockett (1958, pp. 191-201) has also discussed the concepts of Topic and Comment.

3. This example was pointed out to me by J.R. Ross.

4. This type of convention is needed in a grammar with unordered rules. In a grammar with linearly ordered rules, when the turn of an obligatory rule comes up, if its structural description is met, then it must apply immediately. If the rules are unordered, then the following situation can arise: The structural descriptions for two rules, rule A and rule B, are met simultaneously. Rule A is obligatory, and rule B is optional. Rule B may or may not destroy the structural description for rule A. If rule A must apply immediately then rule B will never apply, thus preventing a well-formed string from being generated. If rule B applies first, then it may destroy the structural description for rule A, thus causing an ill-formed string to be generated.

5. Passives in the Slavic languages can also be formed with a reflexive verb, in the following way:

   This book reads itself by many people.

Since this reflexive construction can also be used for impersonal constructions and true reflexives, as well as passives, its frequency of occurrence is not a reliable indicator of the frequency of passivization.

6. I am begging the question of how lexical insertion happens, since it is not relevant to this study. It should be borne in

mind, however, that Yngve's work predates Chomsky (1965).

7. Judgements on this and all other Japanese examples were provided by Susumu Kuno.

8. Kuno (1974) points out that center-embedded structures, unlike left-branching or right-branching structures, cause perceptual difficulties, and notes that a simple pushdown store model such as Yngve's is inadequate to account for these facts.

9. In some cases, direct objects in Japanese can be marked with ga. This occurs only with a very restricted class of verbs, and, according to Kuno, it is not unreasonable to state that a sequence of two NP's marked with ga, or NP-wa followed by NP-ga causes the hearer to hypothesize that an embedded sentence has begun.

Chapter 3. AN INITIAL HYPOTHESIS

3.1

Retaining for the moment the terminology provided by Yngve, which we discussed in section 2.3, let us consider a variation of his model which would operate, not from the point of view of the speaker, but rather from that of the listener. Suppose that the temporary memory of the listener contains, not everything the speaker intends to say, but rather only the minimum number of constituents which must be uttered in order for the sentence to be complete and grammatical. For example, (1), which would have the production shown in (2) by Yngve's speaker-based model, would have, according to this new approach, the production shown in (3).

(1) Joe did his homework quickly and carelessly.

(2)

REGISTER	STORE
Joe	VP
did	NP ADVP
his	N ADVP
homework	ADVP
quickly	and ADV
and	ADV
carelessly	--

D = 2

(3)

REGISTER	STORE
Joe	VP
did	NP
his	N
homework	--
quickly	--
and	ADV
carelessly	--

D = 1

To test the predictions made by the listener-oriented model, let us reconsider examples (23)-(27) from chapter 2.

2:(23) #The prisoner that the lawyer that the judge reprimanded defended likes peanut butter.

REGISTER	STORE	
the	N VP	
prisoner	VP	
that	S VP	
the	N VP VP	
lawyer	VP VP	
that	S VP VP	
the	N VP VP VP	D=4
judge	VP VP VP	
reprimanded	VP VP	
defended	VP	
likes	NP	
peanut	N	
butter	--	

2:(24) The judge that reprimanded the lawyer that defended the prisoner was past retirement age.

REGISTER	STORE	
the	N VP	
judge	VP	
that	S VP	
reprimanded	NP VP	
the lawyer	VP	
that	S VP	D=2
defended	NP VP	
the	N VP	
prisoner	VP	
was	NP	
past	NP	
retirement	--	
age	--	

2:(25) #Dogs people children like own bark.

REGISTER	STORE	
dogs	VP	
people	VP VP	
children	VP VP VP	D=3
like	VP VP	
own	VP	
bark	--	

2:(26) The prisoner that was defended by the lawyer that was reprimanded by the judge likes peanut butter.

REGISTER	STORE
the	N VP
prisoner	VP
that	S VP
was	NP VP
defended	VP
by	NP VP
the	N VP
lawyer	VP
that	S VP
was	NP VP
reprimanded	VP
by	NP VP
the	N VP
judge	VP
likes	NP
peanut	N
butter	--

D=2

2:(27) [[[Neko-ga okkaketa] nezumi-ga tabeta] chīzu-ga kusatte ita to]
cat chased mouse ate cheese rotten was COMP

Yamada-san-wa itta.]
Mr. Yamada said

REGISTER	STORE
neko-ga	NP V
okkaketa	S
nezumi-ga	NP V
tabeta	S
chīzu-ga	NP V
kusatte	V
ita	--
to	S
Yamada-san-wa	V
itta	--

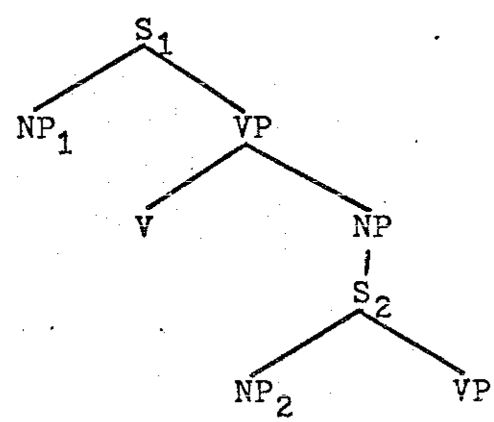
D=2

This new approach, which only stores the constituents which are required for a <u>minimally complete</u> utterance, allows us to make better predictions than those made by Yngve's, since according to the present metric, no acceptable sentence has a depth greater than 2. However, we still predict that (25) is more acceptable than (23), which is not the case. The problem is that the device is too sensitive, in that the internal structure of simple NP's is affecting the acceptability prediction. The model which I will present next is sensitive only to S nodes, and as such is not affected by what goes on inside each clause. It is called the Poker Principle.

THE POKER PRINCIPLE

Four basic definitions are necessary:

Strictly within: A node A is strictly within a sentence S_i if it is dominated, not necessarily immediately, by the node S_i and if no other S nodes intervene between S_i and A.



For example, NP_1 above is strictly within S_1, but NP_2 is not, due to the intervention of S_2. For the moment, we will say that, by definition, no S node can be strictly within any other S.

In play: A sentence S_i is in play if the word currently being uttered is strictly within S_i.

In hand: A sentence is in hand if words strictly within it have been uttered, and if there are words strictly within it which must still be uttered before the sentence can be complete and grammatical.

<u>On the table</u>: A sentence is on the table if all the words strictly within it have been uttered.

Example (4) below illustrates the definitions stated above.

(4) The man who arrived yesterday owns a Cadillac.

Word being uttered	In Play	In Hand	On the Table
the	S_1	S_1	
man	S_1	S_1	
who	S_2	S_2,S_1	
arrived	S_2	S_2,S_1	
yesterday	S_2	S_2,S_1	
owns	S_1	S_1	S_2
a	S_1	S_1	S_2
Cadillac	S_1	S_1	S_2
-------	--	--	S_1,S_2

By definition, if there are no sentences in hand, the sentence is considered to be finished.

<u>Hypothesis</u>: No more than two sentences can be in hand at once.

The following examples test the hypothesis.

(5) #The house that the man that my father hired bought burned down.

Word being uttered	In Play	In Hand	On the Table
the house	S_1	S_1	
that	S_2	S_2,S_1	
the man	S_2	S_2,S_1	
that	S_3	S_3,S_2,S_1	
my father hired	S_3	S_3,S_2,S_1	
bought	S_2	S_2,S_1	S_3
burned down	S_1	S_1	S_2,S_3
-------	--	--	S_1,S_2,S_3

(6) The man that bought the house that Fred used to live in works for my father.

Word being uttered	In Play	In Hand	On the Table
the man	S_1	S_1	
that	S_2	S_2, S_1	
bought the house	S_2	S_2, S_1	
that	S_3	S_3, S_1	S_2
Fred used to live in	S_3	S_3, S_1	S_2
works for my father	S_1	S_1	S_3, S_2
----------	--	--	S_1, S_3, S_2

(7) [[[[Neko-ga okkaketa] nezumi-ga tabeta] chīzu-ga kusatte ita to]
cat chased mouse ate cheese rotten was COMP
Yamada-san-wa itta.]
Mr. Yamada said

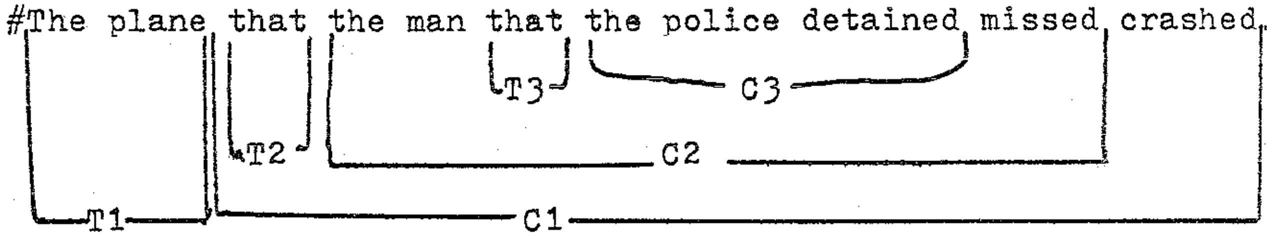
Word being uttered	In Play	In Hand	On the Table
neko-ga okkaketa	S_4	S_4	
nezumi-ga tabeta	S_3	S_3	S_4
chīzu-ga kusatte ita	S_2	S_2	S_3, S_4
to	S_2	$S_2(S_1)$	S_3, S_4
Yamada-san-wa itta	S_1	S_1	S_2, S_3, S_4
------------	--	--	S_1, S_2, S_3, S_4

Notice that if Yngve's model were altered so as to be sensitive only to S nodes, it would make the same predictions as the Poker Principle for examples (5) and (6) above. For (7) however, the altered version of Yngve's model would predict the following:

OUTPUT	REGISTER	STORE
	S_1	
S_1	S_2	S_1
S_1 S_2	S_3	S_2,S_1
S_1 S_2 S_3	S_4	S_3,S_2,S_1
S_1 S_2 S_3 S_4	neko-ga	"
S_1 S_2 S_3 S_4 neko-ga	okkaketa	"
S_1 S_2 S_3 S_4 neko-ga okkaketa	nezumi-ga	S_2 S_1
....S_3 S_4 neko-ga okkaketa nezumi-ga	tabeta	"
...neko-ga okkaketa nezumi-ga tabeta	chīzu-ga	S_1
..okkaketa nezumi-ga tabeta chīzu-ga	kusatte ita	"
...tabeta chīzu-ga kusatte ita	to	"
...tabeta chīzu-ga kusatte ita to	Yamada-san-wa	--
...kusatte ita to Yamada-san-wa	itta	--

So far, the Poker Principle seems to work, since none of the acceptable sentences has more than two S's in hand at once. Notice also that the concept of strictly within does not originate entirely with the Poker Principle. Recall Moore's claim that comment disjunction produces unacceptable sentences, and his definition of comment as everything in the sentence except the topic. If one were to take his definition of comment to pertain to complex sentences as well as to simplex ones, then there would be no comment disjunction in the following sentence.

(8) #The plane that the man that the police detained missed crashed



C1 would contain T2, C2, T3 and C3, and C2 would contain T3 and C3. Clearly, what Moore means by everything in the sentence except the topic is everything <u>strictly within</u> the sentence except the topic.

One further consequence of the Poker Principle is that it accounts for the strangeness of sentences like (9) and (10), as compared with (9a) and (10a).

(9) ?Tom said that Ed is a fool last night.

(9a) Tom said last night that Ed is a fool.

(10) ?Mary sang a song she had learned in Europe before the war to her children.

(10a) Mary sang to her children a song she had learned in Europe before the war.

The same is true in Czech, as shown by (11) and (11a).

(11) #Tom řekl, že Ed je hlupák, včera večer.
said that fool last night

(11a) Tom řekl včera večer že Ed je hlupák.

In a standard syntactic analysis, the (a) sentences would be derived by the optional rule of Heavy NP Shift, and no more could be said. According to the Poker Principle, however, (9) through (11) should be strange, for the following reason:
Recall that a sentence is in hand only as long as there are constituents strictly within it which still <u>must</u> be uttered before the sentence can be complete and grammatical. However, no clause is put on the table until one of the following two conditions is met:

1. Another clause begins.
2. The end of the sentence is reached.

In (13) below, for example, S_2 could be put on the table after learned, if the sentence ended, or another clause began, at that point. However, this does not happen, since neither of the two above conditions is satisfied. By the above criteria, then, since (12)-(14) are complete and grammatical, S_1 in (9)-(11) is no longer in hand once S_2 is in play.

(12) Tom said that Ed is a fool.

(13) Mary sang a song she had learned in Europe before the war.

(14) Tom řekl, že Ed je hlupák.

Since S_1 is no longer in hand when the final constituent shows up, it is no longer in an active state. It must be retrieved from some kind of less immediate memory storage in order for the last constituent to be added. This retrieval, we claim, adds more difficulty to the processing of the sentence.

3.2

There are many cases where the Poker Principle is too strong, or too simplistic. These will be mentioned here, and in the next chapter, a model will be proposed which, combined with the Poker Principle, handles them. First, consider the following sentences:

(15) #I met the man who the woman who lives next door loves.

(16) Joe said that the woman who lives next door speaks Indonesian.

According to the Poker Principle, these sentences should be equal in acceptability, but (15) is clearly less acceptable than (16). The problem is that the Poker Principle has no way of distinguishing between different types of embeddings. This same fault results in bad predictions in the case of (17), (18) and (19) below.

(17) The possibility that the man that I hired is a sexist frightens me.

(18) #The possibility that the idea that students are dangerous will be abandoned is faint.

(19) #The store that the man who I hired owns is going bankrupt.

It seems to be generally true that a sentence with two different types of embeddings is easier to understand than one with two instances of the same type of embedding. This is a well-known fact, which has been discussed by Chomsky (1965), Kuno (1974) and Yngve (1960).

Another problem is illustrated by (20) and (21) below:

(20) ?Mr. Smith, who my sister, who has a photographic memory, recognized, was wanted by the FBI.

(21) #The suspect who the witness who had a photographic memory recognized was wanted by the FBI.

The only difference between (20) and (21) is the fact that the relative clauses in (20) are non-restrictive, whereas in (21) they are restrictive. It is therefore clear that a model of sentence acceptability cannot be based only on general structural facts.

It is interesting to note that while speakers often claim that sentences like (20) are acceptable, they cannot correctly answer a question about the content of one of the embedded clauses. This is in marked contrast to sentences containing restrictive relative clauses, which are judged unacceptable if more than two sentences are in hand at once.

The above facts lead me to suggest the following as a plausible explanation for the difference between restrictive and non-restrictive relative clauses. Given that a non-restrictive relative clause does not serve to identify its head NP, but rather provides additional information about an entity already known to the hearer, one can understand the matrix sentence without understanding the non-restrictive relative clause embedded in it. I would like to suggest that in a sentence like (20), or to take a more extreme case, (22), if the hearer cannot successfully process all of the non-restrictive relative clauses, he will simply ignore some or all of them, and process the rest of the sentence as if the relative clause were not there.

(22) The question of whether Professor Jones, whom the chairman, who is a cousin of mine, recommended highly, should be given tenure, has been raised.

Restrictive relative clauses, on the other hand, serve to identify to the hearer the referent of the head NP. It is therefore impossible to understand a sentence containing a restrictive relative clause unless the relative clause can be processed.

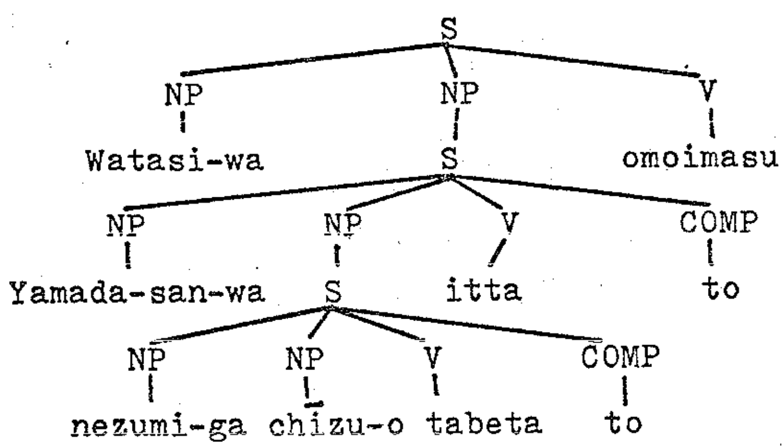
Another set of counterexamples to the Poker Principle has to do with embeddings of the type $[_{NP}\ S\ ]_{NP}$. I will call these sentential NP's, to distinguish them from complex NP's, which have the structure $[_{NP}\ NP\ S\ ]_{NP}$ or $[_{NP}\ S\ NP\ ]_{NP}$.

Sentential NP's do not generally center-embed in English, but they do in Japanese, as shown by (23) below:

(23) $[_S$ Watasi-wa $[_S$ Yamada-san-wa $[_S$ nezumi-ga chīzu-o tabeta to $]_S$
I Mr. Yamada mouse cheese ate COMP

itta to $]_S$ omoimasu. $]_S$
said COMP think

(I think Mr. Yamada said the mouse ate the cheese.)



According to Kuno, this sentence is marginally acceptable. The Poker Principle predicts that it should be no more acceptable than (24), which is completely incomprehensible.

(24) # $[_S$ Yamada-san-wa $[_S$ Taroo-ga $[_S$ sensei-ga sikatta $]_S$ kodomo-ni
Mr. Yamada Taroo teacher scolded child-to

yatta $]_S$ hon-o kaita. $]_S$
gave book-acc. wrote

Mr. Yamada wrote the book that Taroo gave to the child who the teacher scolded.

3.3

At this point we have proposed an initial hypothesis for a model of syntactic processing, and have become aware of many of the problems which present themselves. Before proceeding to the final hypothesis presented in this study, it would perhaps be good to define exactly what conditions a model of syntactic processing must satisfy in order to be adequate. We will also state the range of syntactic constructions that the model presented here will handle.

First, if two sentences differ in acceptability, and if the difference between them can be expressed in syntactic terms, then the model must correctly predict the acceptability distinction. If the two sentences are structurally identical; in other words, if the difference between them is semantic or pragmatic, then the model cannot be expected to account for any acceptability distinctions which may occur.

Secondly, the model must work left-to-right. It may have a limited amount of look-back; that is, it may be able to refer to elements which have already been processed. However, the constraints on the amount of look-back allowed must be statable in a principled way.

Thirdly, since the capacity of the model is supposed to reflect the limitations of human short-term memory, it must be universal. Any language for which consistently wrong predictions are made constitutes serious counter-evidence to the model.

The Poker Principle does not satisfy the above criteria, and is therefore inadequate. The model presented in the remainder of this study will satisfy the criteria of acceptability.

The syntactic constructions which will be used in developing the final hypothesis are the following:

1. Relative clauses

   The man who Fred hired arrived today.

2. NP-complements

   The probability that it will rain is very high.

3. Sentential complements

   Fred thinks that snow is green.

4. Wh-questions

   What did the committee recommend?

5. Topicalization

   Beans, I never eat.

6. Cleft

   It's the weather that bothers me.

7. Pseudo-cleft

   What Sue bought was a houseboat.

8. Non-subject raising (Tough-movement)

   John is easy for us to get along with.

9. Though-preposing

   Handsome though Fred is, Sue still doesn't like him.

In most cases, the example sentences will involve either several instances of one of the above rules, or a combination of several rules.

Chapter 4. THE PARALLEL PROCESSING MODEL

4.1

In section 3.2 it was shown that the Poker Principle, while it makes good predictions in some cases, is not adequate to account for all types of multiply-embedded sentences. The counter-examples were of the following form: Two sentences differed radically in acceptability, while the Poker Principle predicted that they should have equal status. The difference between the two sentences was describable in one of two ways: either in purely structural terms, or with reference to non-structural factors, such as topic configuration, or relative clause type.

Since the model presented so far is stated in purely structural terms, it would be unfair to expect it to account for distinctions which are not due to differences in structure. These will be discussed later, and for the moment we will be concerned only with the problems exemplified by the counter-examples stated below.

I Sentential versus Complex NP embeddings

(1) #[The man [that the lawyer [that the judge reprimanded] defended] won his appeal.]

(2) [[That the man [who said [that the book was finished]] didn't know his business] is clear to everyone.]

(1) clearly has three sentences in hand at once, and, as predicted by the Poker Principle, is unacceptable. The case of (2) is less clear. Since the first word in the sentence is a complementizer, the utterance cannot be complete and grammatical until the completion of the clause in which the first clause is embedded. Given that if there are no sentences in hand, an utterance is said to be

complete, we must therefore say that with the production of the first word, <u>that</u>, not only S_2 but also S_1 is in hand. In that case (2) also has three sentences in hand at once. According to the Poker Principle, then, (2) should be unacceptable, but it is clearly much more comprehensible than (1).

<u>II Location of Extraction Hole</u>

(3) #The book [that the man [who I hired ∅] wrote] deals with politics.

(4) The book [that the man [who ∅ hired me] wrote] deals with politics.

As far as the Poker Principle is concerned, these two sentences have identical production configurations, and as such should be of equal acceptability. However, since (4) is more acceptable than (3), we must conclude that the position of the extraction hole, something which the Poker Principle cannot even refer to, is relevant for the determination of acceptability.

<u>III NP-complements versus Restrictive Relative Clauses</u>

(5) The idea [that the man [who I hired] is incompetent] frightens me.

(6) #The woman [that the man [who I hired] married] screamed.

(7) ##The man [who the idea [that students are dangerous] frightens] lives next door.

(8) #The possibility [that the idea [that students are dangerous] will be abandoned] is faint.

The Poker Principle predicts that (5)-(8) should be equally unacceptable, since all of them have three sentences in hand. One cannot say simply that sentences containing two different types of embedding are easier to understand than those containing two occurrences of the same type of embedding. This statement would predict that (5) and (7) should both be easier than either (6) or (8), when in fact (7) is much less acceptable than (6) or (8).

The hypothesis which I will present next is intended to handle all of the so-called structural counter-examples to the Poker Principle. It will later be extended and refined to handle counter-examples of other types.

4.2

First of all, consider the status of (7)-(8) if the embedded sentence is simply deleted:

(7) The man who arrived yesterday brought a present.

(7a) The man brought a present.

(8) My sister says that blizzards are beautiful.

(8a)*My sister says.

Notice that (7a) is grammatical, while (8a) is not. The matrix sentence in (7) can therefore be said to be, in some sense, independent of the embedded sentence, whereas in (8), it is not.

Suppose that the production/recognition of a sentence proceeds in a linear fashion along some kind of track. An embedded sentence which is syntactically obligatory is processed in the same track as the sentence in which it is embedded, while non-obligatory embedded sentences must be processed on a separate track.

Let us further assume that one of the things a listener must do in order to understand a sentence is to hook up relative pronouns or questioned NP's with the predicates they are arguments of. As far as the present study is concerned, we will assume that if a left-shifted NP is replaced in the position it was moved from, then it can be processed with respect to its predicate. However, this replacement cannot take place entirely within the track on which the clause is being processed. If we allowed permutation of elements within a track, the power of the model would be greatly increased.

Relative pronouns and questioned words must therefore be held, apart from the clause they belong to, until their rightful positions are reached, at which point they are reincorporated into the sentence. We will call this separate storage place a hold cell, and we claim that the capacity of the hold cell is a single NP. It should be noted that we are not claiming that the held NP must be a relative pronoun or a question word; in fact it will later turn out that other types of NP's must also be held in this way.

The question then arises as to how many tracks and hold cells are available to the speaker for the processing of a sentence. Consider the following sentences in terms of how many tracks and hold cells they require.

(9) #The man that the judge that the President appointed sentenced won his appeal.

Track 1 The man won his appeal

Track 2 that the judge sentenced

Track 3 that the President appointed

HOLD1 that HOLD2 that

(10) That the man who said that the book was finished didn't know his business is clear to everyone.

TR1 That the man didn't know his business is clear to everyone

TR2 who said that the book was finished

(11) The book that the man who hired me wrote deals with politics.

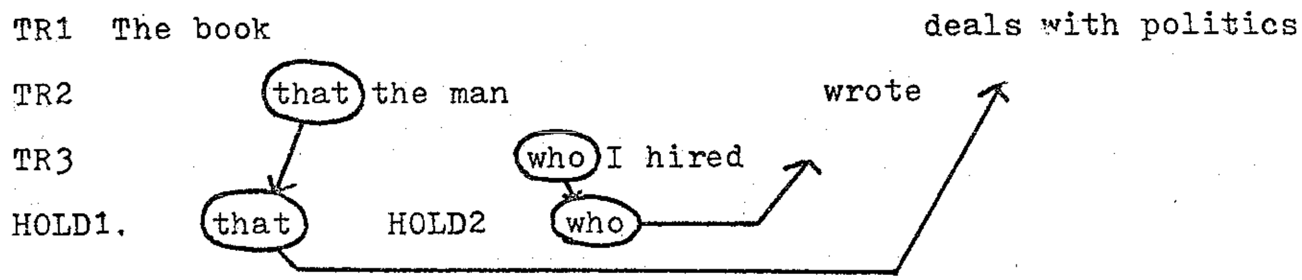
TR1 The book deals with politics

TR2 that the man wrote

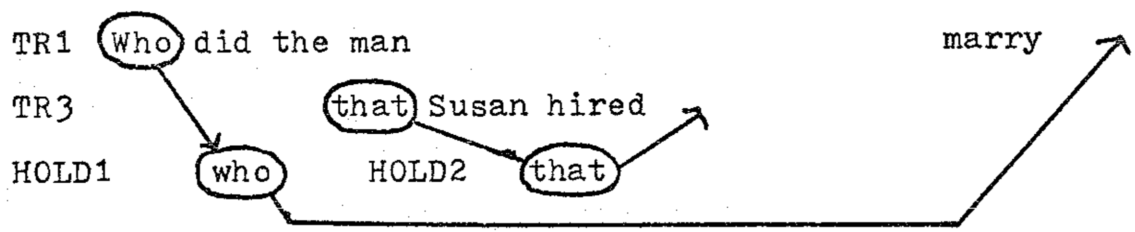
TR3 who hired me

HOLD1 that

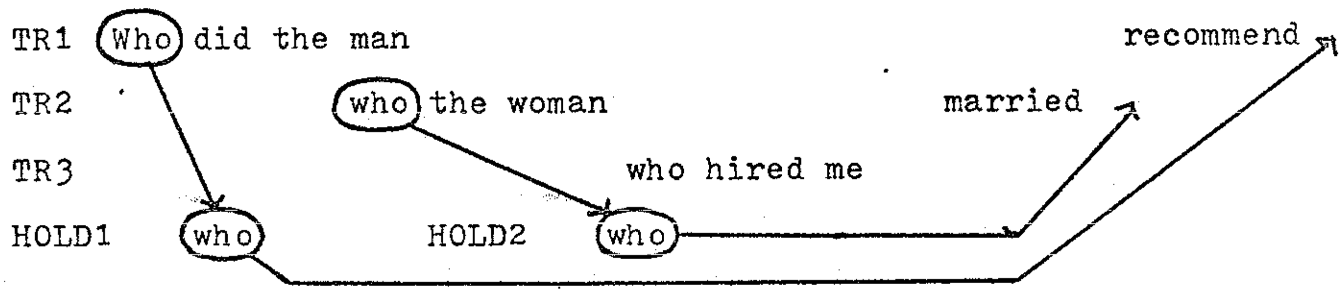
(12) #The book that the man who I hired wrote deals with politics.



(13) Who did the man that Susan hired marry?



(14) #Who did the man who the woman who hired me married recommend?



Notice that the unacceptable sentences above, (9), (12) and (14), all required three tracks and two hold cells. We therefore conclude that three tracks and two hold cells constitute too much processing space. However, there are some acceptable sentences, e.g. (11) which require three tracks, but only one hold cell, and others, e.g. (13), which require two hold cells, but only two tracks. Clearly, then, the number of tracks and the number of hold cells are not mutually independent.

Now, consider the following sentence:

(15) I know the man who married the woman who worked in the library that carries the journal that publishes the reports that Dr. Jones writes.

TR1 I know the man

TR2 who married the woman

TR3 who worked in the library

TR4 that carries the journal

TR5 that publishes the reports

TR6 that Dr. Jones writes

HOLD1 that

Given the number of tracks required for this sentence, and given the fact that (15) is much more acceptable than (12) or (14), it is clear that there is something wrong with the way the notion of track processing is formulated. Specifically, this approach fails to distinguish between simultaneous and sequential processing of clauses. In other words, it does not incorporate the notions of <u>in hand</u> and <u>on the table</u>, which were given in section 3.1. The following statement incorporates these notions, thus retaining the insights captured by the Poker Principle.

4.3

The capacity of a speaker/hearer to process utterances may be described in terms of the following model.[1]

THE PARALLEL PROCESSING MODEL

For sentence processing, there are three tracks and one hold cell. The first track (henceforth TR1) is global, in that it is opened at the beginning of each utterance, and not closed until the end of the utterance, even if the sentence it contains is on the table. The second track (TR2) is semi-local; it can be opened and closed several times during the course of one utterance. It is closed as soon as the sentence it contains is on the table, and is then free for the processing of another embedded sentence. TR2 may also be left open while another clause is processed on the third track (TR3). TR3 is strictly local. It can be opened and closed several times during the utterance, but unlike TR2, cannot be left open while another clause is processed elsewhere. In other words, sentences processed on TR3 must be either in play or on the table, while those on TR1 and TR2 may be in hand, in play or on the table. The hold cell (H) is used for storing a constituent which must be moved in order to be processed. H can contain no more than one constituent at a time. If H is full, and if another constituent needs to be stored, then it may be placed on TR2, if TR3 is empty. If a constituent is stored on TR3, then TR3 is unavailable for sentence processing until the stored constituent has been reintegrated into the clause from which it came. Note also that no constituent is stored in H or TR3 until it becomes clear that that constituent has to be moved. The following example illustrates how the model works.

(16) The candidate who the woman who hired me supports came to town.

Word being uttered	TR1	TR2	TR3	H
the	the			
candidate	candidate			
who		who		
the		the		who
woman		woman		
who			who	
hired			hired	
me			me	
supports		supports		
came	came	who		
to	to			
town	town			

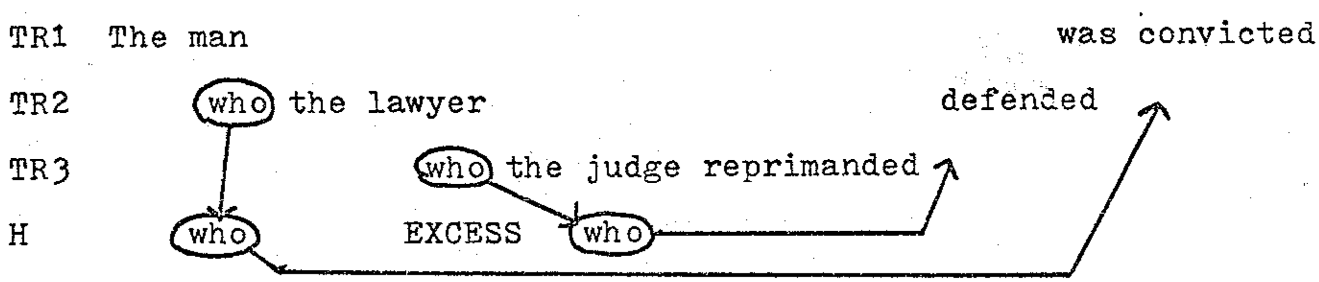
In future examples, a production like the one above will be represented as follows:

TR1	the candidate			came to town
TR2	who	the woman	supports	
TR3		who hired me		
H	who			

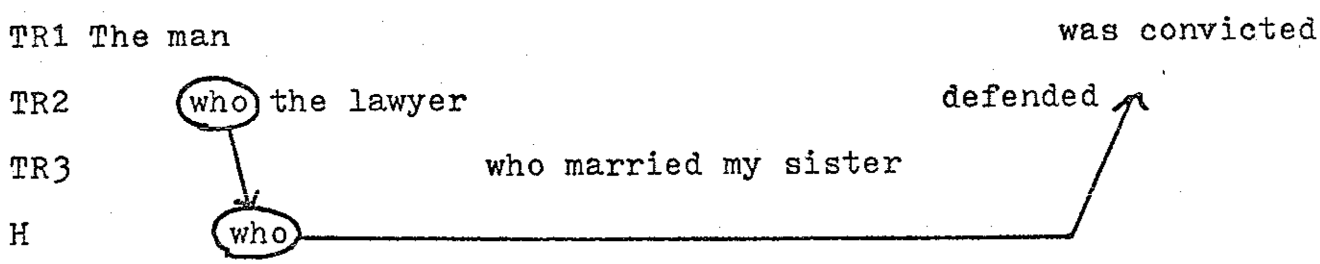
The following examples test the hypothesis stated above. They are grouped according to the type of embeddings they contain.

I Relative Clauses only

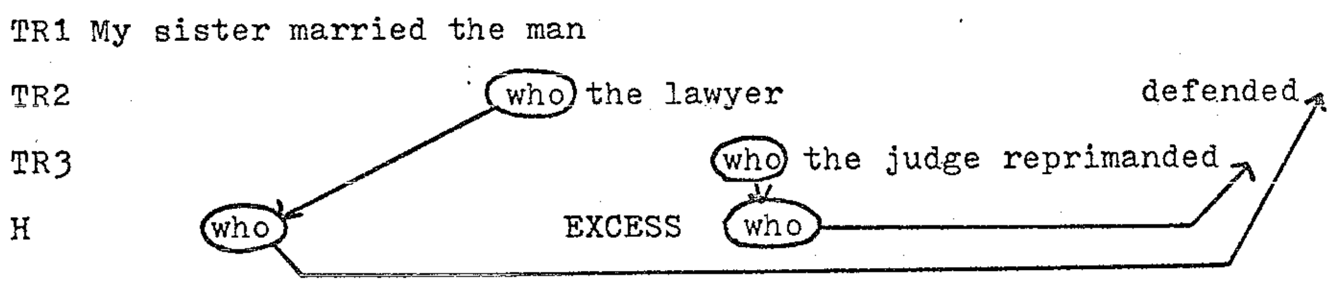
(17) #The man who the lawyer who the judge reprimanded defended was convicted.



(18) The man who the lawyer who married my sister defended was convicted.



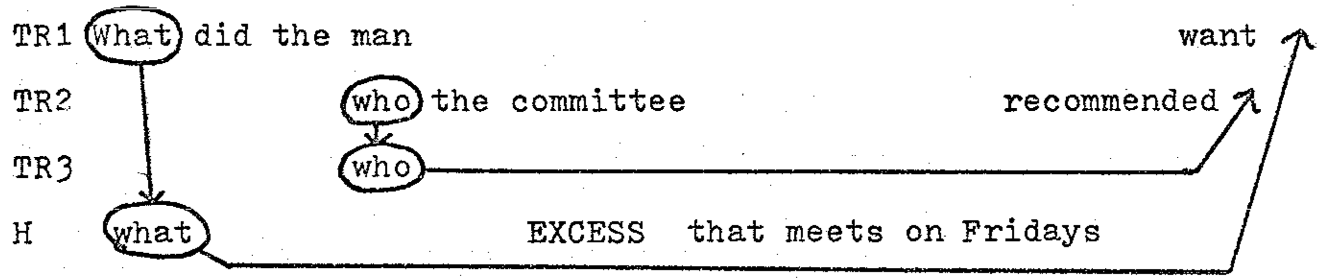
(19) #My sister married the man who the lawyer who the judge reprimanded defended.



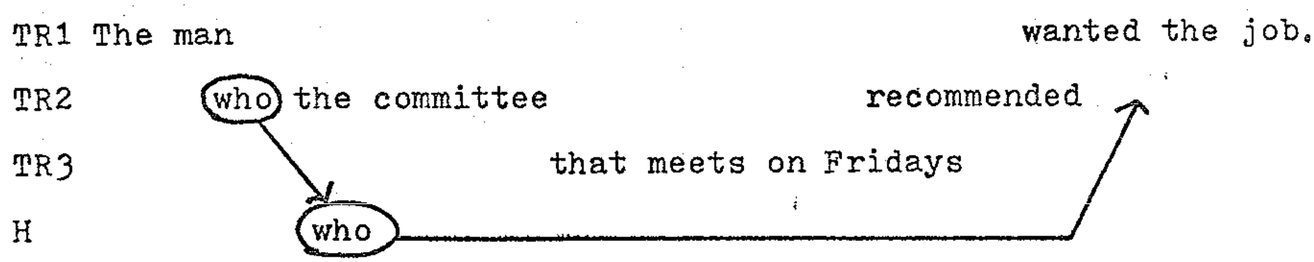
(17) shows that there is only one hold cell, and the unacceptability of (19) shows that even if the sentence occupying TR1 is on the table, that track is not available for storing embedded constituents.

II Relative clauses inside wh-questions

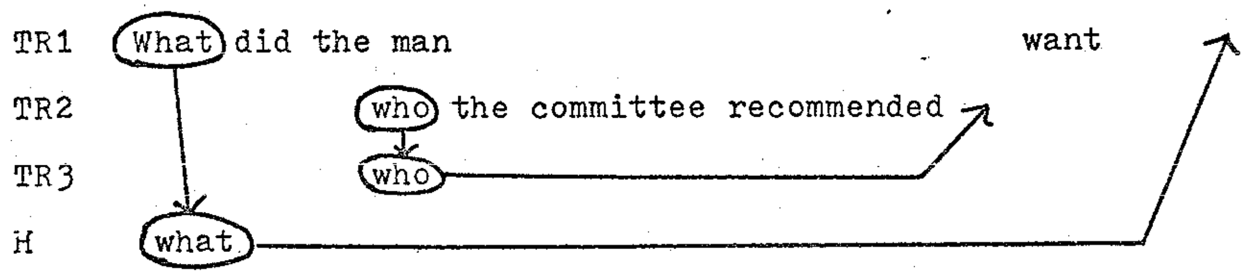
(20) #What did the man who the committee that meets on Fridays recommended want?



(21) The man who the committee that meets on Fridays recommended wanted the job.



(22) What did the man who the committee recommended want?

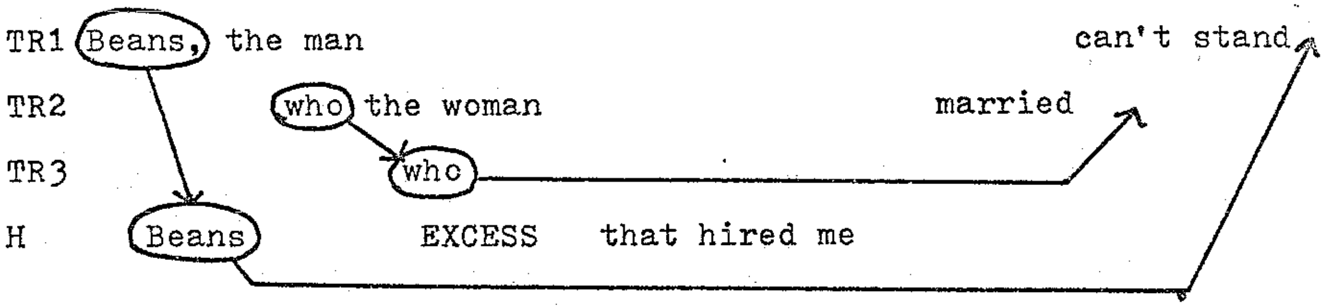


The difference in acceptability between (20) and (21) demonstrates that the question word does, in fact, occupy the hold cell, thus causing (20) to exceed the available processing space. The acceptability of (22) shows that a restrictive relative clause can occur between a preposed wh-word and its extraction hole.

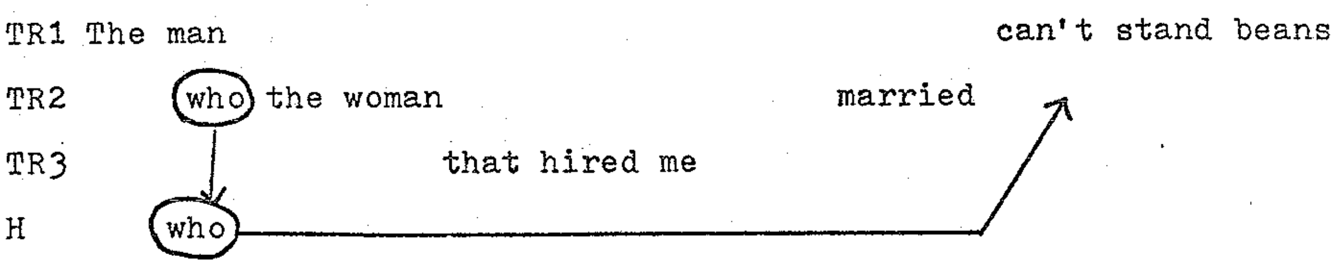
<u>III Topicalized sentences containing relative clauses</u>.

Some of the sentences below have been judged ungrammatical by native speakers; however, when pressed to judge the sentences according to the criterion of comprehensibility alone, they gave the judgements shown:

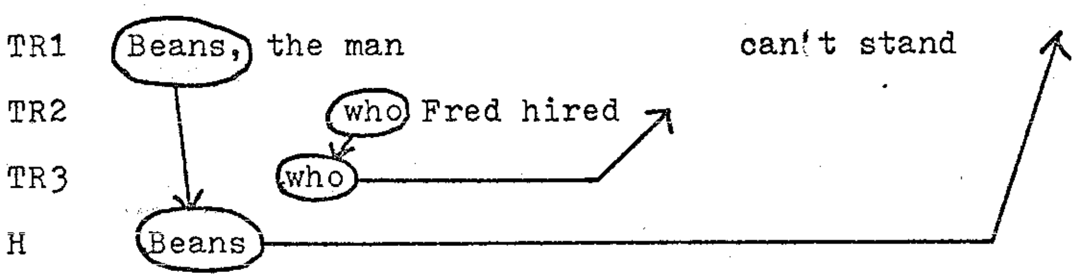
(23) #Beans, the man who the woman that hired me married can't stand.



(24) The man who the woman that hired me married can't stand beans.



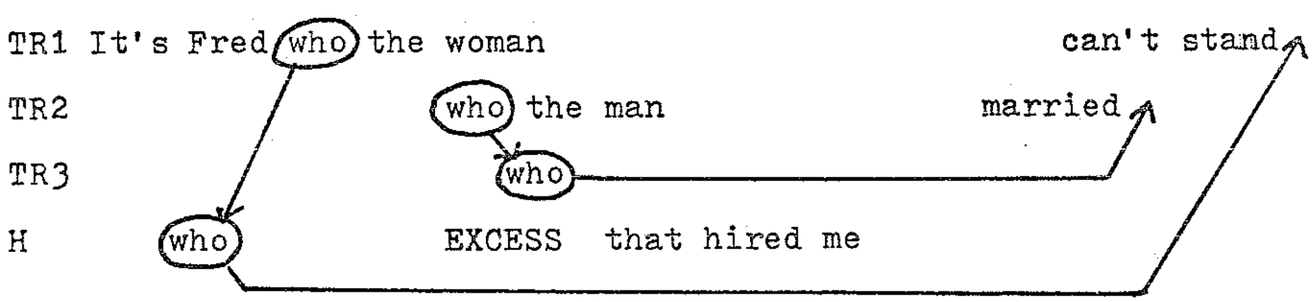
(25) Beans, the man who Fred hired can't stand.



The difference in acceptability between (23) and (24) demonstrates that the topicalized NP does occupy H, thus decreasing the amount of space available for processing the rest of the sentence.

IV Clefts and relative clauses

(26) #It's Fred who the woman who the man that hired me married can't stand.



(27) The woman who the man that hired me married can't stand Fred.

TR1 The woman can't stand Fred

TR2 who the man married

TR3 that hired me

H who

V Non-subject raising (Tough-movement) and Relative clauses

(28) #Fred is easy for the woman who the man who hired me married to please.

TR1 Fred is easy for the woman to please

TR2 who the man married

TR3 who

H Fred EXCESS who hired me

(29) It's easy for the woman who the man who hired me married to please Fred.

TR1 It's easy for the woman to please Fred

TR2 who the man married

TR3 who hired me

H who

VI Though-preposing and relative clauses

(30) #Handsome though the man who the woman who hired me married is, I still don't like him.

TR1 Handsome though the man is, I still don't like him

TR2 who the woman married

TR3 who

H handsome EXCESS who hired me

(31) Though the man who the woman who hired me married is handsome I still don't like him.

TR1 Though the man   is handsome, I still don't like him

TR2   who the woman   married

TR3   who hired me

H   who

The above sets of examples demonstrate that if a constituent occurs to the left of the clause to which it belongs, then that constituent occupies processing space (H, or TR3, as the case may be) until the extraction hole is reached, thus decreasing the amount of space available for the processing of other material.

4.4

The following examples, involving NP-complements, will require some discussion, and some refinement of the model. First, recall that syntactically obligatory embeddings are processed on the same track as the clause in which they are embedded, while non-obligatory embeddings are processed on a separate track. The question then arises as to whether the embedded sentence in (32) is syntactically obligatory.

(32) The idea that students are dangerous is ridiculous.

(32a)?The idea is ridiculous.

Unless there is some prior reference to an idea, in other words, unless <u>the idea</u> is discourse-anaphoric, (32a) is semantically incomplete. Nonetheless, it is syntactically well-formed in any context. In this way, then, we must say that the embedded sentence in (32) is not syntactically obligatory. We would therefore like to claim that NP-complement embeddings, like relative clauses, are processed on a separate track from the sentence in which they are embedded. Unfortunately, this simple statement leads us to make wrong predictions, as follows:

(33) The possibility that the man who I hired is incompetent worries me.

TR1	The possibility				worries me
TR2		that the man		is incompetent	
TR3			who I hired		
H		who			

(34) #The man who the possibility that students are dangerous frightens lives next door.

```
TR1  The man                                      lives next door
TR2       (who) the possibility           frightens   ↑
TR3                 that students are dangerous        |
H         (who) --------------------------------------+
```

As mentioned in section 3.2, it has been noted by various linguists that sentences containing two different types of embeddings are easier to understand than those containing two occurrences of the same type of embedding. However, the difference between (33) and (34) above, together with the fact that (34) is, if anything, less acceptable than (35), demonstrates that something else is involved.

(35) #The man who the woman that hired me married has a German Shepherd.

```
TR1  The man                                 has a German Shepherd
TR2      (who) the woman                married   ↑
TR3                that hired me                  |
H       (who) ------------------------------------+
```

The following group of sentences is intended to set the stage for the refinements which will be presented next. Each sentence involves an NP-complement, and one other type of construction which takes up processing space.

(36) #The man who the idea that students are dangerous frightens has an attack dog.

```
TR1  The man                                has an attack dog
TR2        who the idea               frightens
TR3                  that students are dangerous
H        who
```

(37) It's Fred who the idea that students are dangerous frightens.

```
TR1 It's Fred who the idea                                  frightens
TR2                  that students are dangerous
TR3
H          who
```

(38) What the claim that the earth is flat implies is that Columbus really fell off the edge.

```
TR1 What the claim                 implies is that Columbus really
                                                fell off the edge
TR2             that the earth is flat
TR3
H        what
```

(39) *Joe is hard to imagine the claim that the earth is flat convincing.[2]

```
TR1 Joe is hard to imagine the claim                  convincing
TR2                              that the earth is flat
TR3
H          Joe
```

(40) Crazy though the idea that students are dangerous may seem, a lot of people believe it.

TR1 Crazy though the idea may seem, a lot of people believe it

TR2 that students are dangerous

TR3

H Crazy

(41) What does the claim that English is an underlying tomato patch imply?

TR1 What does the claim imply

TR2 that English is an underlying tomato patch

TR3

H what

(42) Fred, the possibility that Sue won't show up worries.

TR1 Fred, the possibility worries

TR2 that Sue won't show up

TR3

H Fred

(43) The possibility that the man who I hired is incompetent worries me.

TR1 The possibility worries me.

TR2 that the man is incompetent

TR3 who I hired

H who

(44) The idea that it was a woman that they should hire never occurred to them.

TR1 The idea never occurred to them

TR2 that it was a woman (that) they should hire

TR3

H (that)

(45) The idea that what they should do was sell the car never occurred to them.

TR1 the idea never occurred to them

TR2 that (what) they should do was sell the car

TR3

H (what)

(46) The idea that Fred would be easy to talk to never occurred to me.

TR1 the idea never occurred to me

TR2 that (Fred) would be easy to talk to

TR3

H (Fred)

(47) The proposal that, crazy as Fred is, we should still hire him, will never be accepted.

TR1 The proposal will never be accepted.

TR2 that (crazy) as Fred is we should still hire him

TR3

H (crazy)

(48) The idea that Fred didn't know what he was in for appalls me.

TR1 the idea appalls me

TR2 that Fred didn't know what he was in for

TR3

H what

(49) #I work with the man who the idea that students are dangerous frightens.

TR1 I work with the man

TR2 who the idea frightens

TR3 that students are dangerous

H who

It would seem that whenever an NP-complement occurs such that the head NP is on TR2 and the complement S is on TR3, then the sentence is unacceptable, even if the amount of processing space available is not exceeded.[3] For the moment, then, let us simply say that NP complements must be processed on tracks 1 and 2. The implications of this statement will require a general discussion of the relative power of the different parts of the model, an elaboration of exactly how it works, and an evaluation of what type of machine it is.

4.5.0

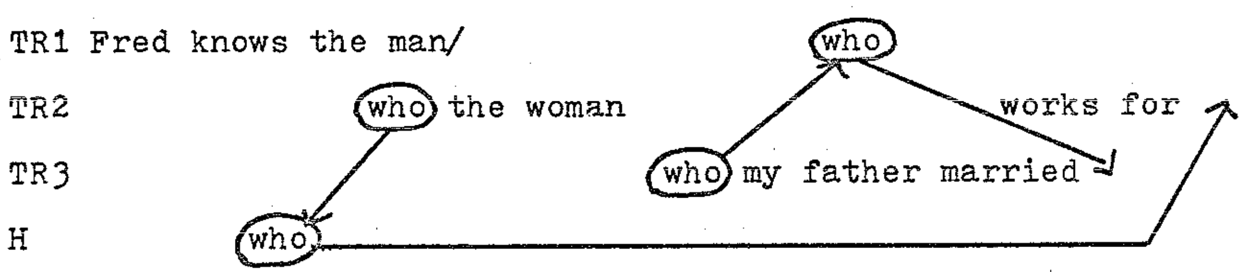
Up till now, I have been concerned with the general shape and size of the parallel processing machine. There are many questions about the detailed workings of the model which, in order to simplify matters, I have either ignored or assumed to have solved. This section is devoted to answering those questions.

So far, we have established that there are three processing tracks and one hold cell. We know that the tracks differ in power, and that TR2 and TR3 somehow become vacant when the sentences they contain go on the table. Given that, let us now turn to some of the detailed questions.

4.5.1

What does it mean for a clause to vacate a track? Recall that in section 4.3 we said that TR1 was global, and that even if the sentence it contained was on the table, it was not available for processing deeply embedded material. For example, the following production is ill-formed. (/ means that the material preceding it on that track is on the table.)

(50) #Fred knows the man who the woman who my father married works for.



I now claim that the apparent globality of TR1 is due to the fact that when a sentence has been completely processed, the whole sentence ends up on TR1. When a clause which has been processed on a lower track is on the table, it is then promoted to the track immediately above it, if that track is ready for it.

Now we must determine what the conditions are for the promotion of a clause to a higher track. For the purposes of this discussion, let us call the higher track TRa, and the clause it contains Sa. The next track down, TRb, contains the sentence Sb.

There are three possible conditions for the promotion of Sb from TRb to TRa:

1) Sa alone must be on the table.

2) Sb alone must be on the table.

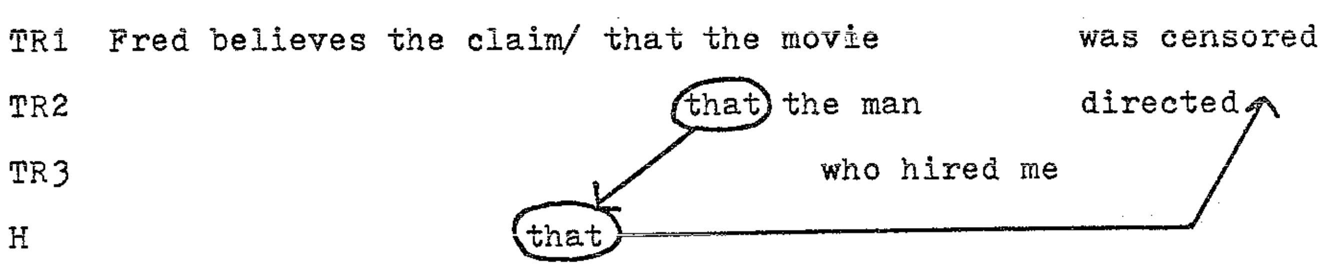
3) Both Sa and Sb must be on the table.

<u>Hypothesis A: Sa alone must be on the table</u>.

If Sb is promoted as soon as Sa is on the table, then for some center-embedded, and all right-branching structures, Sb will be promoted before it is processed. It will therefore be processed on the same track as Sa. For the moment, let us leave open the question of whether this situation could really be called one of promotion.

If Hypothesis A is correct, then sentences (51) and (52) below should both be acceptable.

(51) Fred believes the claim that the movie that the man who hired me directed was censored.



(52) #Fred wrote the book that the company that gave the man who hired me a job published.

```
TR1 Fred wrote the book/(that) the company                    published
TR2                                  that gave the man     a job
TR3                                           who hired me
H         (that)
```

Since (51) is acceptable to most speakers, it seems that if Sb is an NP-complement, then a sufficient condition for promotion is that Sa be on the table. The unacceptability of (52) shows that this is not a sufficient condition for the promotion of a relative clause.

<u>Hypothesis B: Sb alone must be on the table</u>.

According to this hypothesis, Sb is always processed on a different track from Sa, and is promoted as soon as it is on the table, whether or not Sa is on the table. Crucially, however, Sb is not promoted until it is on the table, even if Sa is on the table beforehand.

If this hypothesis is correct, (51) above should be unacceptable, since it would have the following production configuration:

```
TR1 Fred believed the claim
TR2                     that the movie                 was censored
TR3                               (that) the man      directed
H       (that)                  EXCESS  who hired me
```

Sb could not be promoted until after the production of <u>censored</u>, and therefore the lowest clause, <u>who hired me</u>, would exceed the available processing space. Since (51) is acceptable, it is clear that it is not a necessary condition for promotion that an

NP-complement sentence be on the table. (53) below tests whether it is a sufficient condition.

(53) The idea that the police should confiscate the hats that the students who live here were wearing is crazy.

TR1 The idea is crazy
TR2 that the police should confiscate the hats/(that) the students were wearing
TR3 who live here
H (that)

We know from (52) above that S_3 (that the students...were wearing) cannot be processed on the same track as S_2 (that the police... the hats). Since (53) is acceptable, we must therefore conclude that S_2 was promoted to TR1 as soon as it was on the table, thus leaving TR2 open for processing S_3.

We have now established that if Sb is an NP-complement, a sufficient condition for its promotion to TRa is that either Sb or Sa be on the table.

We will now test Hypothesis B for relative clauses.

(54) The man who confiscated the hats that the students who live here were wearing is crazy.

TR1 The man is crazy
TR2 who confiscated the hats/(that) the students were wearing
TR3 who live here
H (that)

For the same reasons as those stated concerning (53) above we must conclude that S_2(who confiscated the hats) was promoted as soon as it was on the table.

Now let us reconsider the question of whether the situation exemplified by (51) is in fact one of promotion. Recall that sentential NP's (object complements, for example) are claimed to be always processed on the same track as the sentence in which they are embedded.[4] If (51) is an example of promotion, then in order to be consistent, one would have to say that all cases of sentential NP embeddings are also examples of promotion. Further, it is not clear what the concept of promotion would then mean. In the case of sentences like (53) and (54), promotion means that a sentence is moved from one track to another, whereas with (51) no such movement takes place. It therefore seems reasonable to claim that there is no promotion involved with (51) at all.

The condition for promotion may now be stated simply, without reference to the type of embedding, as follows:

A sentence is promoted to the track immediately above it as soon as it is on the table.

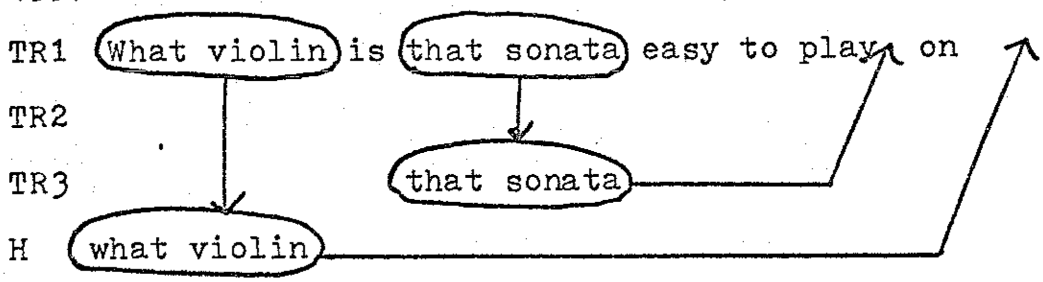
The following statement must also be made concerning the difference in behavior between relative clauses and NP-complements. A relative clause may never be processed on the same track as its head NP, whereas an NP-complement may be so processed, providing that the sentence containing the head NP is on the table.

4.5.2

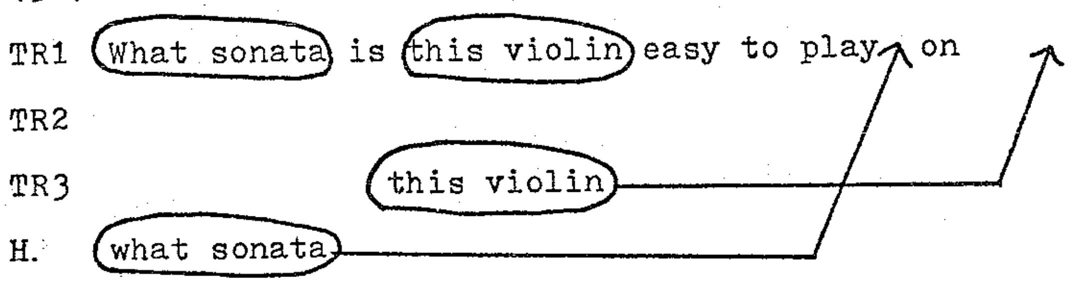
How much information is carried along with an NP when it is stored, either in H or in TR3?

If we assume that as little information as possible is carried along with an NP that needs to be moved, then in order to answer the above question, we must first determine what restrictions there are, if any, on the behavior of a stored NP. Consider the following sentences:

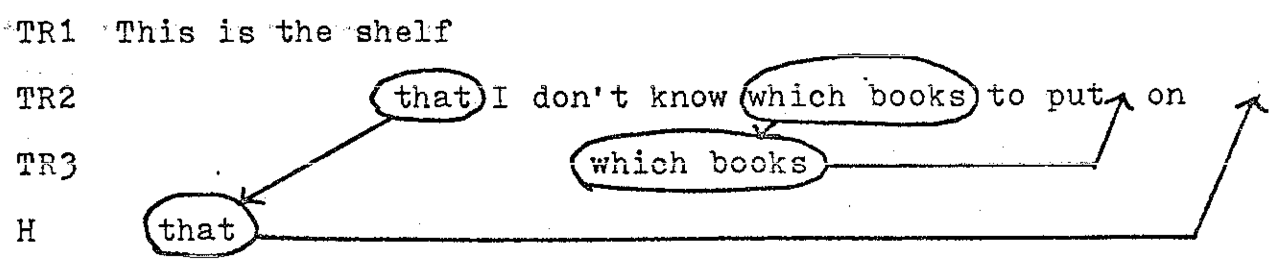
(55) What violin is that sonata easy to play on?



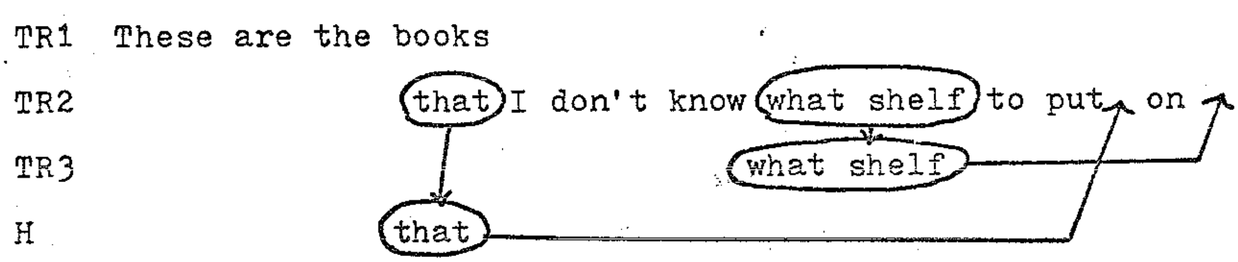
(56)*What sonata is this violin easy to play on?



(57) This is the shelf that I don't know which books to put on.



(58) *These are the books that I don't know what shelf to put on.



(59) It's Sue that Fred doesn't know what to say to.

TR1 It's Sue that Fred doesn't know what to say to

TR2

TR3 what

H that

(60) *It's Sue that Fred doesn't know what to talk to about.

TR1 It Sue that Fred doesn't know what to talk to about

TR2

TR3 what

H that

From the examples given above, it would seem that if two NP's are being held at the same time, then they must be reincorporated in the opposite order from that in which they were stored.[5] In other words, H and TR3, when used for holding NP's, combine to behave like a pushdown store of capacity 2.

Now, consider the following group of sentences.

(61) The man who Fred says Sue likes lives next door.

TR1 The man lives next door

TR2 who Fred says Sue likes

TR3

H who

(62) *The man who Fred knows the woman who likes lives next door.

TR1 The man lives next door

TR2 who Fred knows the woman

TR3 who likes

H who

(63) The man who the woman who hired me married eats horsemeat.

TR1 The man eats horsemeat
TR2 who the woman married
TR3 who hired me
H who

The ungrammaticality of sentences like (8) has been accounted for by Ross's Complex NP Constraint (Ross, 1967). In the context of this model, we can simply say that an NP which is being held in H or TR3 must return to the track from which it came. We can then claim that an NP which is held in H or TR3 is tagged for the track it came from and as soon as an extraction hole appears on that track, the NP is reincorporated.

The alert reader will immediately counter this claim by citing an example like (64) below.

(64) *The man who Fred doesn't believe the claim that Sue loves lives lives next door.

TR1 The man lives next door
TR2 who Fred doesn't believe the claim/that Sue loves
TR3
H who

If S_2 (<u>Fred doesn't believe the claim</u>) is on the table, then S_3 (<u>that Sue loves</u>)will be processed on the same track. The relative pronoun will thus be reincorporated on TR2, and the sentence should be grammatical.

To answer this we must further define the criteria for a sentence being on the table. In chapter 3, it was stated that a sentence is on the table as soon as all constituents strictly within it have been uttered. Given that if a sentence is on the table, it is no longer active, let us further require the following: A sentence on TRi is on the table if and only if all constituents strictly within it have been uttered, and if there are no NP's in H or TR3 which are tagged for TRi.

According to this criterion, then, in (64) above, S_2 is not on the table at the beginning of S_3, and therefore S_3 must be processed on TR3. The relative pronoun would thus have to be reincorporated on TR3, violating the condition stated above.

It would seem, therefore, that a stored NP must have with it some sort of tag which indicates the track it belongs to. Now, we must determine what other information is contained in the cell with the NP.

The reader will have noticed that when the NP to be stored was a pro-form of some sort, I simply placed the pro-form in the hold cell. No specific claim was meant by this. Clearly, the speaker must at some point make the connection between non-questioned wh-words and their antecedents. One might reasonably say that if the antecedent occurs before the wh-word, as in the case of English relative clauses, then the connection is made when the wh-word is uttered, and what is stored is not the wh-word, but a copy of the antecedent. If the antecedent does not occur before the NP is stored, then the wh-word is stored. The hold cell must therefore be able to hold information about reference or co-reference. If what is stored is a wh-word, then information as to whether or not it is a question word must also be stored.

FOOTNOTES TO CHAPTER 4

1. It is very important to keep in mind that no claim is made as to the psychological reality of any part of the model. This is a theoretical model, and no experiments have been conducted to test it. It may later turn out that the model does correctly represent the way speakers process sentences, but at present I merely claim that it makes predictions which are consistent with the acceptability judgements of native speakers.

2. The ungrammaticality of this sentence seems to have nothing to do with sentence complexity, since (a) is also ungrammatical, while (b) is grammatical:

(a) *Joe is hard to imagine that argument convincing.

(b) Joe is hard to imagine Sue liking.

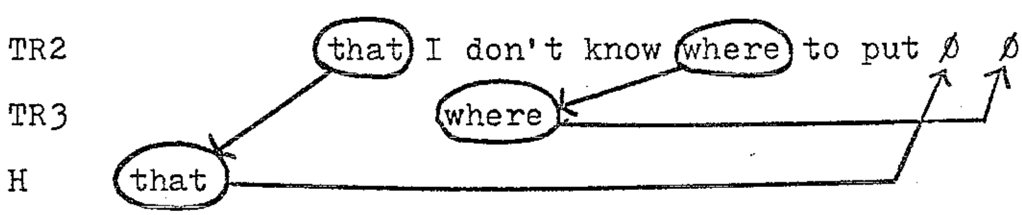
3. This fact was noticed by H. Kučera.

4. This is true for English, but may not be true for languages in which sentential NP's can be center-embedded. This question will be discussed in chapter 5.

5. There are some apparent counterexamples to this claim, having to do with adverbial wh-questions, as follows:

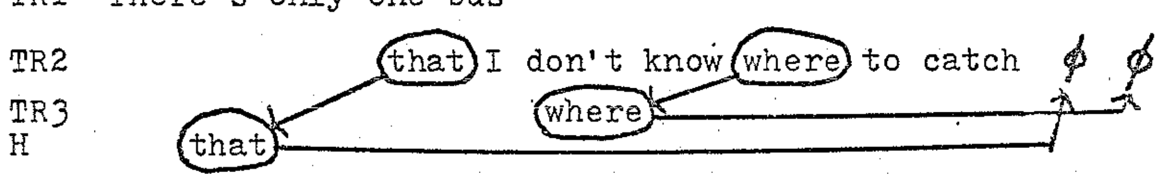
(a) These are the books that I don't know where to put.

TR1 These are the books



(b) There's only one bus that I don't know where to catch.

TR1 There's only one bus



Notice, however, that the following examples obey the constraint:

(c) *These are the books that I don't know what box to put in.

TR1 These are the books

TR2 that I don't know what box to put in

TR3 what box

H that

(d) *There's only one bus that I don't know what corner to wait for at.

TR1 There's only one bus

TR2 that I don't know what corner to wait for at

TR3 what corner

H that

It seems that if the two extraction holes are unambiguously marked, then the constraint applies. It also seems that adverbial wh-words are less closely attached to their original position than are wh-words which replace simple NP's.

Chapter 5. JAPANESE

5.0

Given that we originally rejected Yngve's depth hypothesis on the grounds that it failed to handle examples of left-branching structures, we must show that the parallel processing model can handle these cases.

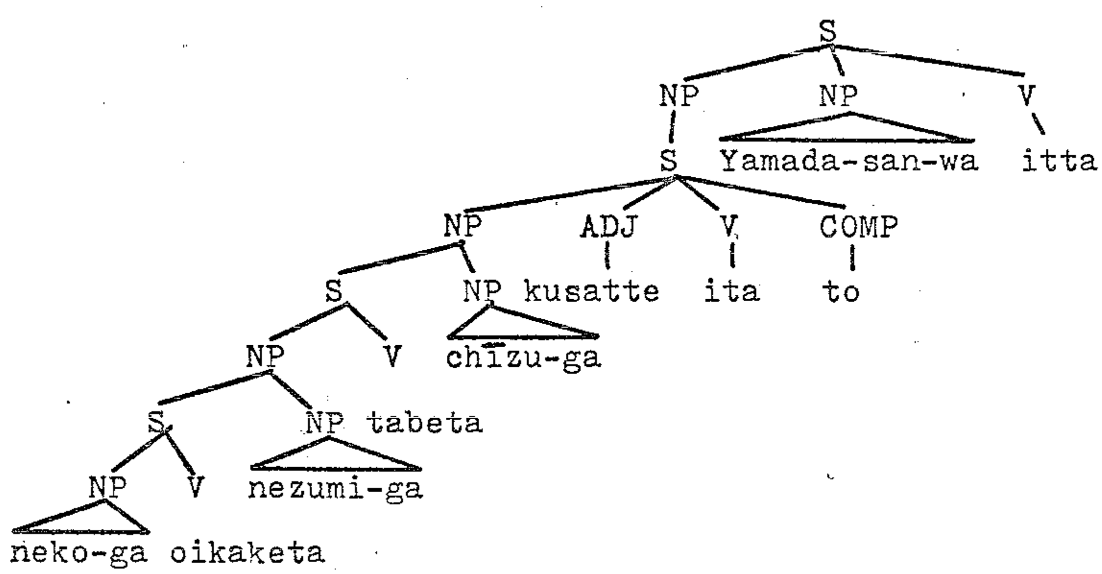
5.1 Left-branching structures

Consider the following examples from Japanese:

(1) Neko-ga oikaketa nezumi-ga tabeta chīzu-ga kusatte ita to
cat chased mouse ate cheese rotten was COMP

Yamada-san-wa itta
Mr. Yamada said

"Mr. Yamada said that the cheese that the mouse that the cat chased ate was rotten.



Recall that Yngve's model made the implicit assumption that the speaker knows, when he utters the first word, how many levels of embedding that word is from the matrix S node. The Poker Principle did not make this assumption, and made the right predictions for left-branching structures. As shown in section 2.4, there is good evidence that in a sentence like (1), it is not necessary for the speaker to know how far he is from the matrix S; (2)-(4) below are all possible sentences in Japanese.

(2) [Neko-ga nezumi-o oikaketa]
    cat mouse chased "The cat chased the mouse"

(3) [[Neko-ga oikaketa] nezumi-ga chīzu-o tabeta.]
    cat chased mouse cheese ate

"The mouse that the cat chased ate the cheese"

(4) [[[Neko-ga oikaketa] nezumi-ga tabeta] chīzu-ga kusatte ita]
    cat chased mouse ate cheese rotten was

"The cheese that the mouse that the cat chased ate was rotten."

This lack of certainty as to the level of embedding of a given word is in marked contrast to the situation in English. Unless the first word uttered in an English sentence is a complementizer, we know that the first word is strictly within the matrix. (For the moment, let us say that untensed clauses are not separate clauses, and that their constituents are strictly within the first tensed clause in which they are contained.) If the first word is a complementizer, we know that the first clause is a tensed clause immediately embedded in the matrix. Since the only kinds of embeddings which require a separate track are relative clauses and NP-complements, and since these types of embeddings follow their heads in English, there is no way an English sentence can begin with a clause which will have to be processed on TR2.

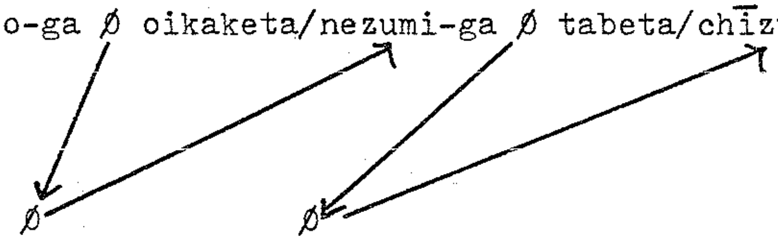
In contrast, it is conceivable that languages in which relative clauses and NP-complements precede their heads might begin sentences in this way. However, the question then arises as to whether pre-nominal relative clauses and NP-complements must be processed on a separate track from their heads. Recall that sentences are always promoted to the next higher track as soon as they are on the table. In languages with post-nominal complex NP embeddings, this means that the embedded sentence is promoted to the track which already contains the head NP. In the case of prenominal complex NP embeddings, however, the embedded sentence would be promoted <u>before</u> the head NP is uttered. If, in fact, the head NP is the first word in its own clause, then the higher track would be completely empty while the embedded sentence is processed on a lower track. In this case, it seems not unreasonable to hypothesize that the embedded clause is processed on the same track as the higher clause. A left branching structure would then have a production configuration like the one in (5) below:

(5) TR1 neko-ga ∅ oikaketa/nezumi-ga ∅ tabeta/chīzu-wa kusatte ita

TR2

TR3

H ∅ ∅



Having shown that the parallel processing model works for left-branching structures, we now turn to center-embedded structures in Japanese.

5.2 Center-embedded constructions

As in English, Japanese center-embedded structures cause processing difficulties. (6), for example, is completely unacceptable.

(6) #[John-ga [sinsi-ga [sensei-ga rakudaisaseta] seito-ni yatta]
John gentleman teacher flunked student-to gave
hon-o kaita]
book wrote

"John wrote the book that the gentleman gave to the student who the teacher flunked."

Clearly, then, a relative clause in Japanese which is center-embedded must require more processing space than one which occurs sentence-initially. Let us claim, therefore, that center-embedded Japanese relative clauses are processed on a separate track from the clause in which they are embedded. We must then ask how center-embedded sentential NP's and NP-complements are processed. Since relative clauses precede their heads, and complementizers occur clause-finally, there is no way to tell, at the beginning of an embedded clause, whether that clause is a relative clause, a sentential NP, or an NP-complement. This means that these three types of embeddings must be processed on the same track with respect to the clause they are embedded in. Since we have claimed that center-embedded relative clauses are processed on a separate track from the clause they are embedded in, we must claim that center-embedded sentential NP's and NP-complements are also processed on a separate track.

Before going any further, let us test the predictions made by the model as it now stands.

(6a) TR1 John-ga hon-o kaita

TR2 sinsi-ga seito-ni ∅ yatta

TR3 sensei-ga ∅ rakudaisaseta

H ∅ ∅

(7) Watasi-wa Yamada-san-wa nezumi-wa chīzu-o tabeta to itta to
I Mr. Yamada mouse cheese ate COMP said COMP

omoimasu
think

"I think Mr. Yamada said the mouse ate the cheese."

(7a) TR1 Watasi-wa omoimasu

TR2 Yamada-san-wa itta to

TR3 nezumi-wa chīzu-o tabeta to

H

(8) #Watasi-ga Yamada-san-ga nezumi-ga chīzu-o tabeta to itta to
I Yamada mouse cheese ate COMP said COMP

omoimasu

think

"It's I who think that it's Yamada who said that the mouse ate the cheese."

(8a) TR1 Watasi-ga omoimasu

TR2 Yamada-san-ga itta to

TR3 nezumi-ga chīzu-o tabeta to

H

So far, we can see some serious problems. The parallel processing model predicts that (6), (7) and (8) should be acceptable; however (6) and (8) are completely unacceptable. According to some speakers, by the time the listener hears the third NP-ga in (6) and (8), he is completely lost. Notice that the only difference between (7) and (8) is that in (7) the subjects of all three sentences have the topic-marker wa while in (8) they are marked with ga. This means that in the case of (7) all three NP's are themes; they have already occurred in previous discourse. It certainly seems reasonable that something which has already been established as a discourse topic should require less processing space than a previously unmentioned entity. The following constraint is an attempt to capture this, and at the same time correct the bad predictions made by the parallel processing model about (6) and (8).

The clause integrity constraint

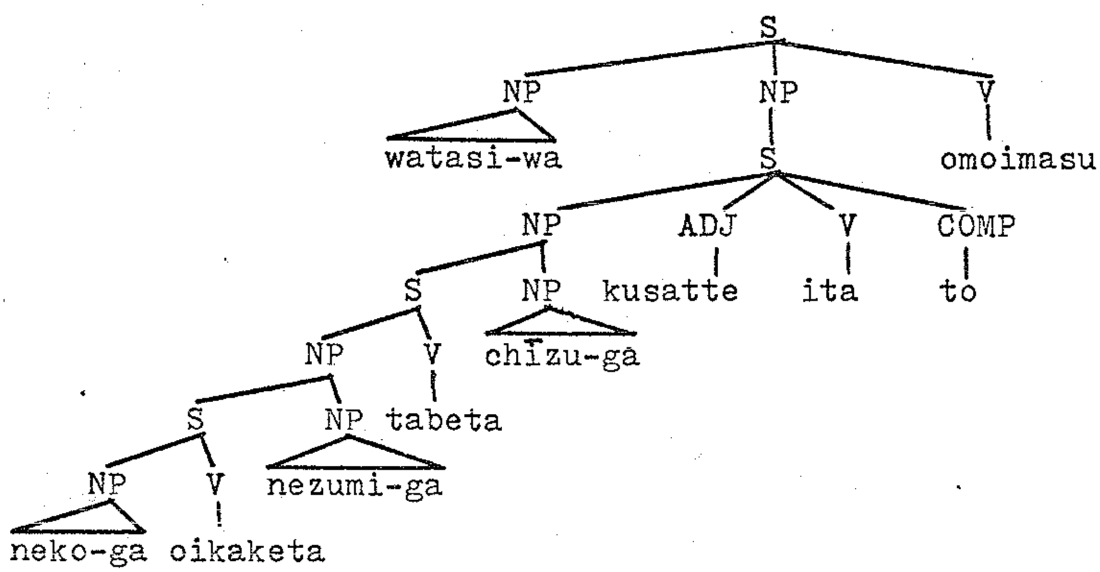
In languages with clause-final complementizers, any non-topical constituent which occurs to the left of its own clause, separated from that clause by one or more embedded sentences, must be held in H or TR3 until its own clause shows up.

The reason that this constraint applies only in languages with clause-final complementizers is the following: If the complementizer, or relative pronoun, occurs at the beginning of the clause, the embedded sentence is "hooked up" with the sentence in which it is embedded as soon as it begins to be uttered. The level of embedding of the embedded sentence, and in most cases the type of embedding as well, is thus determinable from the beginning. In languages like Japanese, however, it is entirely possible that any embedded clause is several levels of embedding away from the word that precedes it, as shown by (9) below.

(9) Watasi-wa neko-ga oikaketa nezumi-ga tabeta chīzu-ga kusatte
    I cat chased mouse ate cheese rotten

    ita to omoimasu
    was COMP think

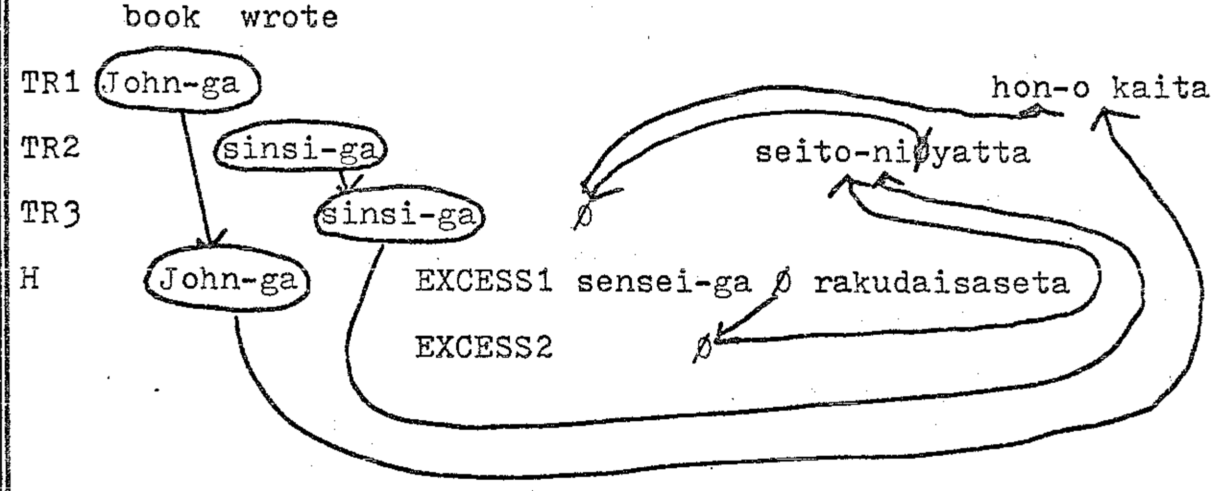
"I think the cheese that the mouse that the cat chased ate was rotten."



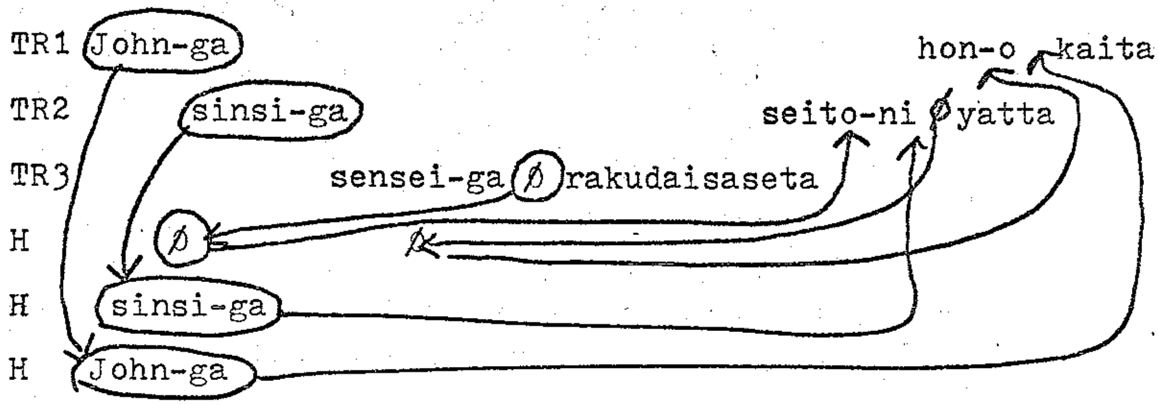
Let us now reconsider (6) as well as several more examples according to the clause integrity constraint.

(6) #John-ga sinsi-ga sensei-ga rakudaisaseta seito-ni yatta
    John gentleman teacher flunked student-to gave

    hon-o kaita
    book wrote



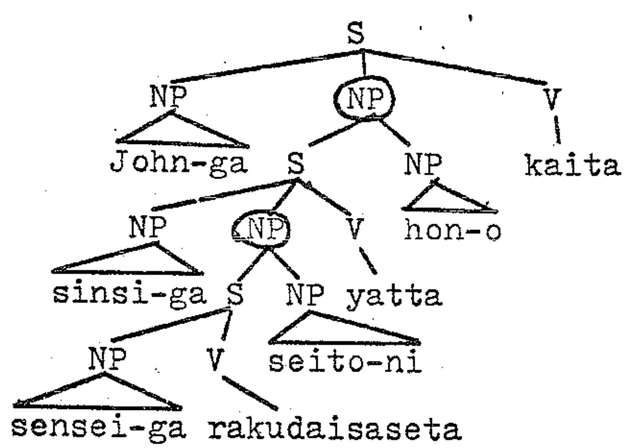
To clarify the above, consider how the production configuration would look if enough space were available.



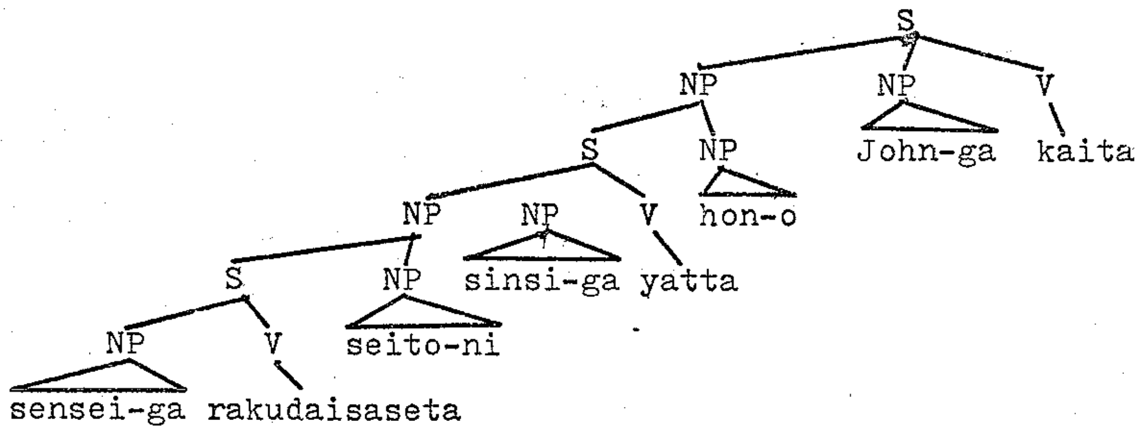
(10) Sensei-ga rakudaisaseta seito-ni sinsi-ga yatta hon-o
teacher flunked student-to gent. gave book

John-ga kaita
John wrote

This sentence is what was given by a native speaker of Japanese when asked to translate the English equivalent of (6) into Japanese. (7) is derivable from (6) by Heavy NP Shift, as shown below. In Japanese, Heavy NP Shift moves the NP to the beginning of its clause rather than to the end, as is the case in English.



The circled NP's are each moved to the beginning of their own clause, to produce the surface structure shown below.



Consequently, (7) would have the following production configuration

TR1 sensei-ga Ø rakudaisaseta/seito-ni sinsi-ga Ø yatta/hon-o J-ga kaita

TR2

TR3

H Ø Ø

In the case of topical NP's, however, the situation is very different, since these NP's do not need to be held until their clauses appear.

(11) Watasi-wa Yamada-san-wa sensei-wa tegami-o kaita to itta to
I Mr. Yamada teacher letter wrote COMP said C.
omoimasu
think

"I think Mr. Yamada said the teacher wrote the letter.

TR1 Watasi-wa omoimasu

TR2 Yamada-san-wa itta to

TR3 sensei-wa tegami-o kaita to

H

We have now shown that the parallel processing model is adequate to account for both left-branching and center-embedded sentences in Japanese. A further problem which presents itself is that of relative clauses with more than one head NP.

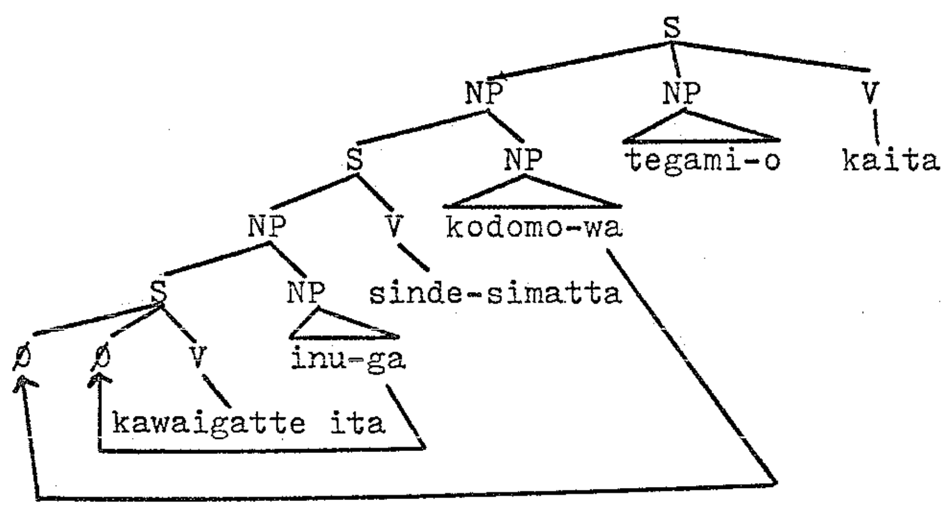
5.3 Multiple-headed relative clauses.

Sentences like (12) below, which are almost impossible to translate into English, are acceptable and grammatical in Japanese.

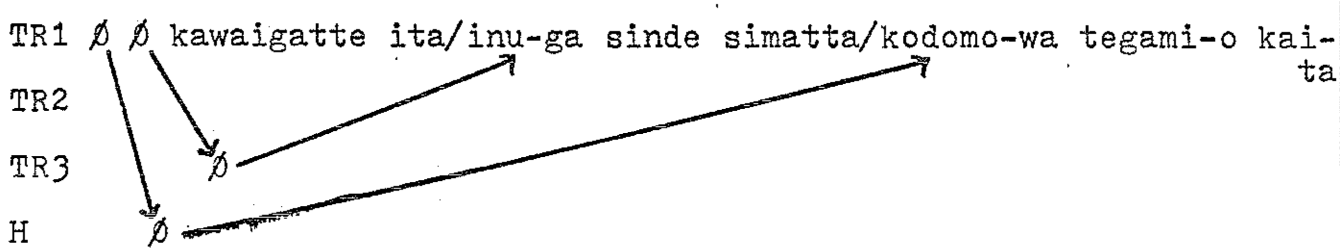
(12) [[[Kawaigatte ita] inu-ga sinde simatta] kodomo-wa tegami-o
was fond of   dog   died   child   letter

kaita]
wrote

"The child who the dog who (he) was fond of (it) died wrote the letter."



This sentence would have a production configuration as follows:

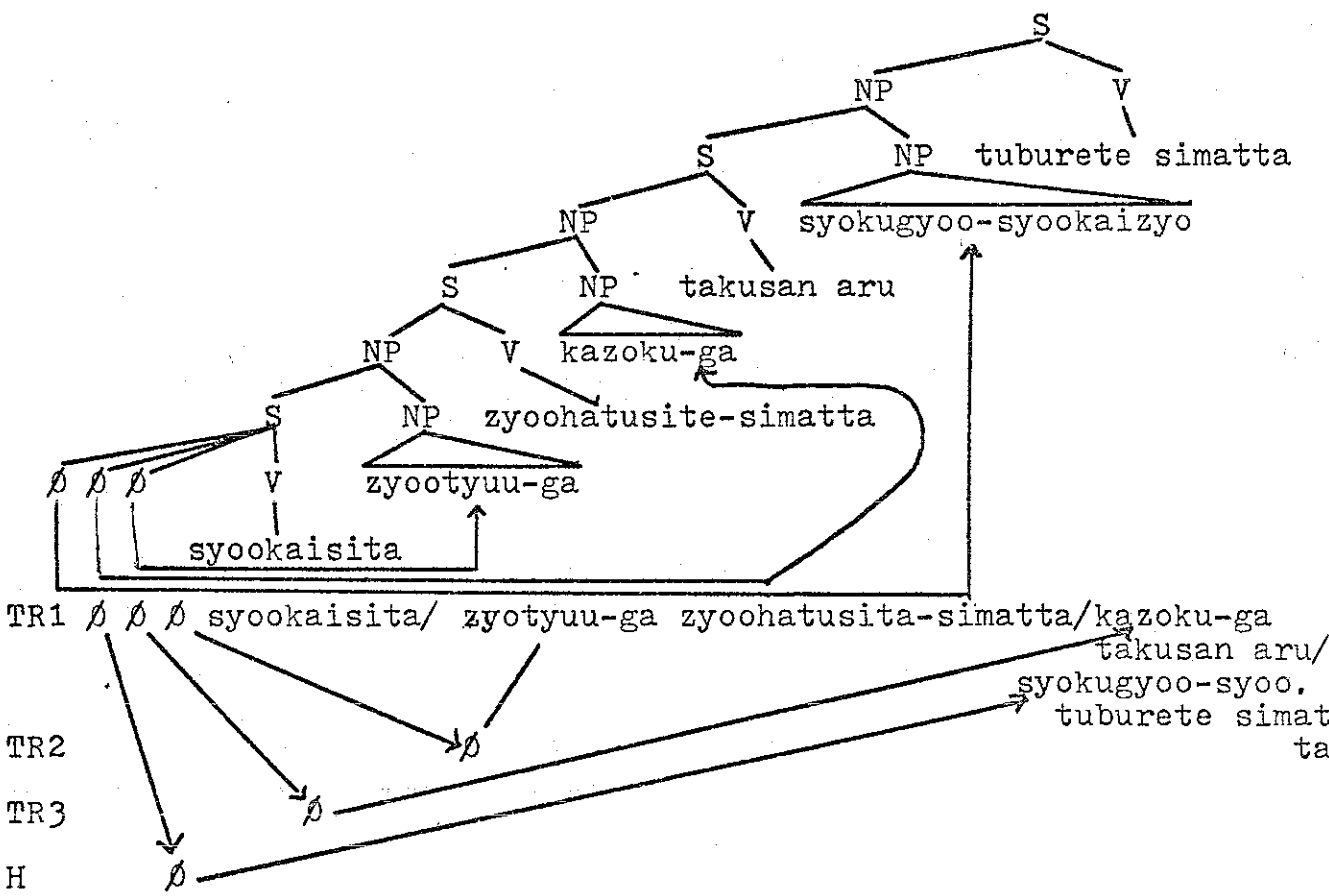


Relative clauses with three heads are also possible:

(13) Syookaisita zyotyuu-ga syoohatusite simatta kazoku-ga
introduced   maid   disappeared   family

takusan aru syokugyoo-syookaizyo tuburete simatta
are many   employment service   bankrupt went

"The employment service that the families that the maid that (they) introduced (her) to (them) disappeared are many went bankrupt."

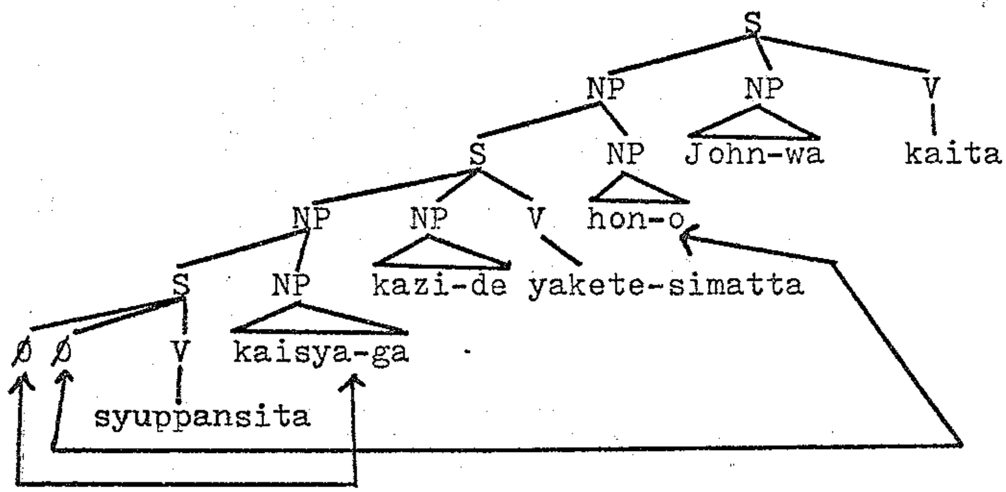


Since left-branching structures do not require more than one track for the processing of clauses, TR2 and TR3, as well as H, are available to store extraction holes until the NP occurs.

We would like to be able to say that this storage, like the storage of NP's in English, behaves like a pushdown store. However this is not the case, since the following sentences are grammatical

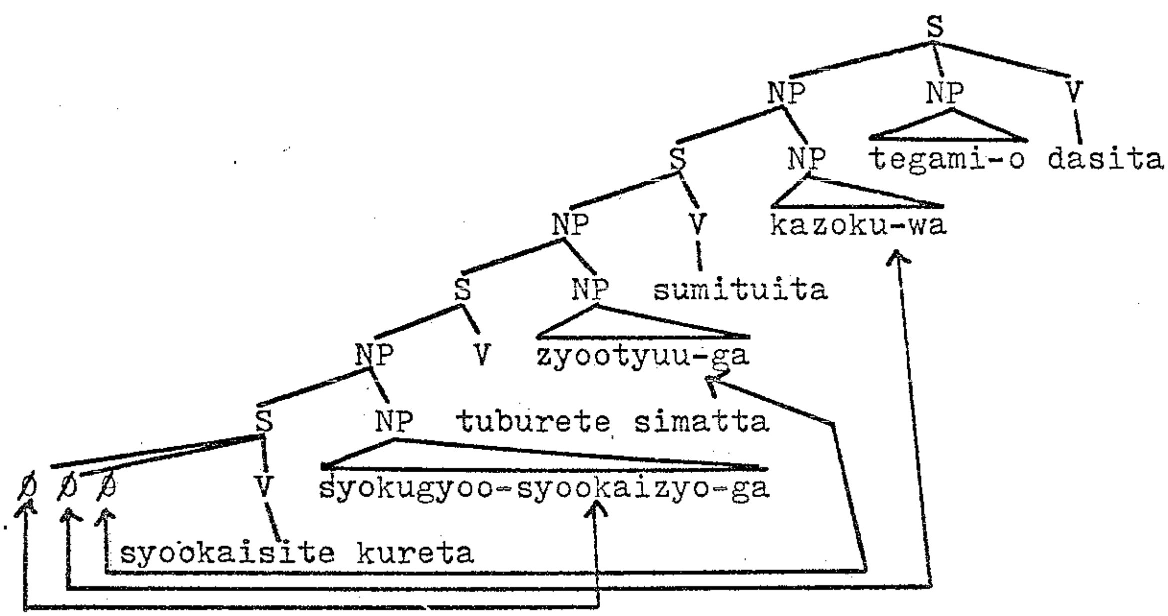
(14) Syuppansita kaisya-ga kazi-de yakete-simatta hon-o John-wa
published company fire-by was burned down book John
wrote
kaita

"John wrote the book that the company that published (it) was burned down by fire."



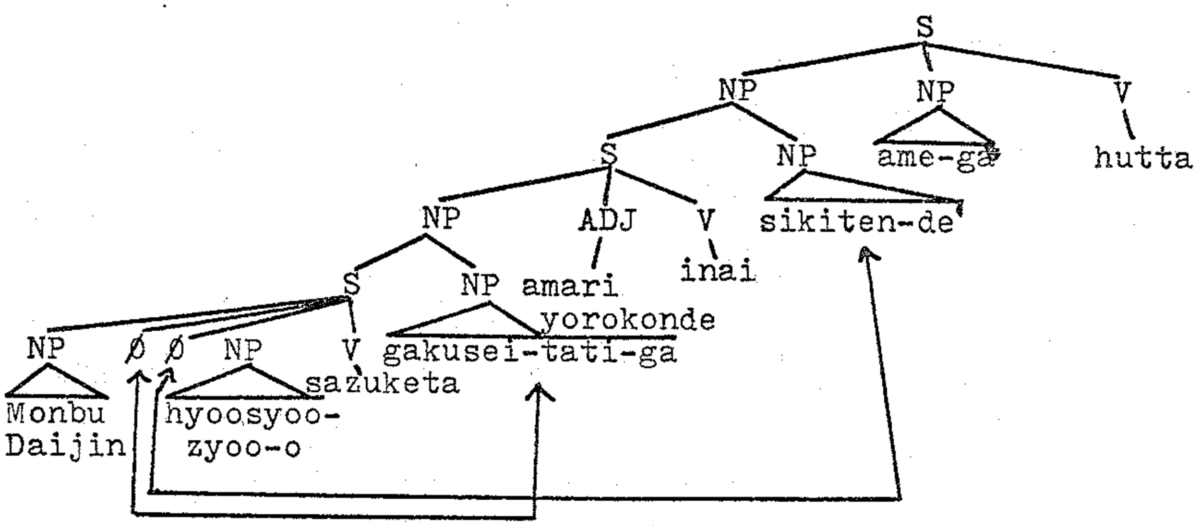
(15) Syookaisite kureta syokugyoo-syookaizyo-ga tuburete-simatta
introducing gave employment service bankrupt went

zyootyuu-ga sumituita kazoku-wa tegami-o dasita
maid settled in family letter sent



(16) Monbu Daijin-ga hyoosyoo-zyoo-o sazuketa gakusei-tati-ga
Education Minister citation awarded students

amari yorokonde inai sikiten - de ame-ga hutta
very pleased weren't ceremony at rain fell



Whitman (1976) has discussed this type of example in great detail, and has shown very convincingly that the factors affecting the grammaticality of these sentences are non-structural. Consider, for example, the following two sentences:

(17) *[[[Ø Ø katte ita] kodomo-ga sinda] inu-wa kanasi sooda]
keeping was child died dog sad looks

(18) [[[Ø Ø katte ita] kodomo-ga sinde simatta] inu-wa kanasi sooda]
keeping was child dying ended up dog sad looks

The only difference between (17) and (18) is that in (17) the verb in S_2 is <u>sinda</u>, "died", and in (18) it is <u>sinde simatta</u>, "ended up dying".

Also, since the order of NP's in a Japanese clause is relatively free, it is not entirely clear what the order of extraction holes has to be, or if there has to be a fixed order at all. Given that, the notion of pushdown store violation has very little meaning, since it depends crucially on the order in which the extraction holes were stored.

Since the acceptability of this type of sentence is affected by so many factors which are outside the scope of this study, I shall not make any attempt to account for it at present.

Ignoring the problems with pushdown store violations, however, the parallel processing model makes some good predictions for Japanese. Notice that all of the multiple-headed relative clauses cited above violate the complex NP constraint (Ross, 1967). The same-track constraint, stated in chapter 4, to the effect that a held NP must return to the track from which it came, was essentially equivalent to the complex NP constraint, as far as English was concerned. In Japanese, however, we can now see an empirical difference between the same-track constraint and the complex NP constraint. Even though the relative clause embeddings in (12)-(15 are complex NP's, they are not processed on a separate track from the sentence in which they are embedded. As such, the same-track constraint makes the right prediction, namely that NP's should be able to be extracted from these embedded sentences, or in this case, deleted under the control of an NP outside the complex NP. The same-track constraint is therefore more adequate than the complex NP constraint for these cases. More evidence bearing on this question will be raised in chapter 7.

Chapter 6. COORDINATE STRUCTURES

6.0

Until now, we have been concerned exclusively with embeddings which are subordinate to the clause they are embedded in. All constructions which could even loosely be called coordinate have been ignored. The reason for this is that coordinate structures seem to contribute much less to the difficulty of processing a sentence than do subordinate embedded sentences. This chapter is devoted to demonstrating how the parallel processing model handles coordinate structures.

Coordinate structures are of basically two types: those where the coordination is signalled before the first conjunct is uttered, and those where the coordination is not signalled until after the first conjunct is uttered. The first type includes sentences with both...and, either...or, the more...the more, not only...but also, and so on. The second type includes sentences with and, but, or, and in some cases because, after, before, etc.

I will claim that in the case of coordinate structures, a copy (M') is created of all or part of the processing machine (M).[1] Once both halves of the coordinate structure have been processed, the two are then stored in the original processing machine, and the copy is destroyed. Two main questions arise: first, when is the copy created, and second, how much of the machine is copied.

6.1 Initially-signalled coordinate structures

This type of coordination is the most straightforward. M' is created as soon as the first coordination signal is uttered. When the first half of the coordinate structure is completed and the second coordination signal is uttered, then M' is restricted so that the second half of the coordinate structure will be of the same syntactic nature as the first half.

Consider the following sentence.

(1) Sue said that both Fred and Peter left last night.

TR1 Sue said that both Fred and left last night
TR2
TR3
H TR1' Peter
TR2'
TR3'
H'

The restriction of M' in terms of the first conjunct explains why <u>Peter left last night</u> is not interpreted as a clause conjoined to <u>Fred</u>.

The following examples test whether, by allowing a copy of the processing machine, we have allowed too much space, that is, whether unacceptable sentences can be produced by the machine.

(2) Sue said that both the possibility that the man who hired her was incompetent, and the fact that the people who she worked with were obnoxious, had induced her to leave her job.

TR1 Sue said that both the possibility and had induced her to leave her job

TR2 that the man was incompetent

TR3 who hired her

H

TR1' the fact

TR2' that the people were obnoxious

TR3' who she worked with

H' who

It would seem that, given that (2) is acceptable, we have not created too large a processing machine.

(3) #Both the fact that it's raining and the idea that the professor who the dean who interviewed me recommended won't be there made me decide to stay home.

TR1 Both the fact and made me decide to stay home

TR2 that it's raining

TR3

H

TR1' the idea

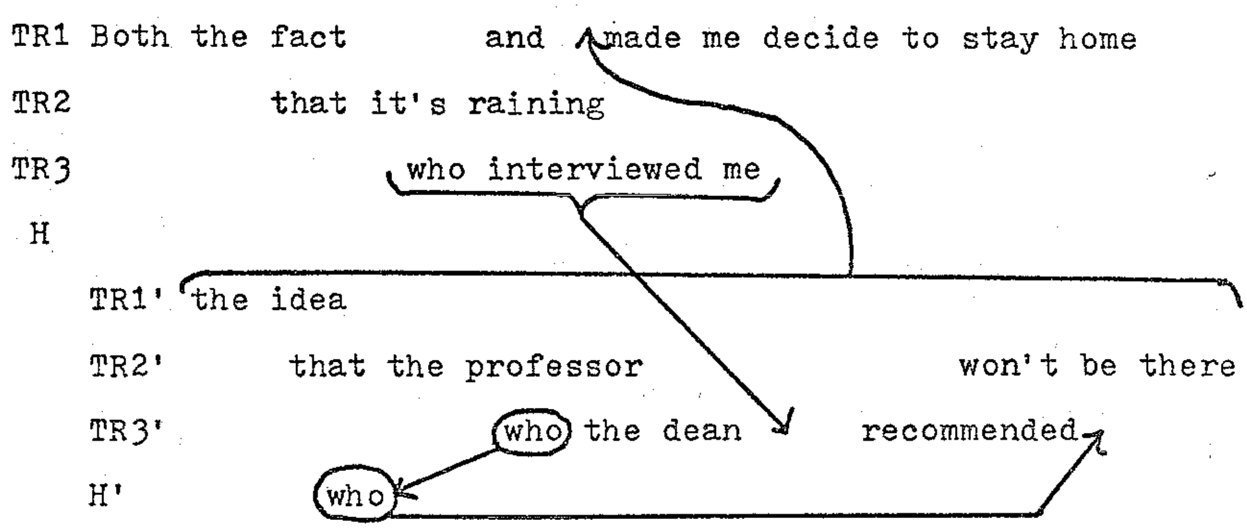
TR2' that the professor won't be there

TR3' who the dean recommended

H' who EXCESS who interviewed me

Notice that there is some unused processing space in M, which one might suspect would be available if the second conjunct proved to be too complex to be handled by M'. The unacceptability of (3) indicates that this space is not, in fact, available in this way. The following constraint prevents ill-formed productions of the type shown in (4).

(4)



The copy isolation constraint

If a copy (M') of the processing machine (M) is created to handle a coordinate structure, then control can pass to M' only at the beginning of the second conjunct, and back to M only after the second conjunct has been completely processed.

This constraint allows us to define exactly when M' is destroyed, as follows: When control passes from M' to M, the material processed in the copy is then stored in M, on the same track as the first conjunct. M' is thus destroyed as soon as control passes back to M.

Now, consider the following sentence.

(5) Both Fred and either Sue or Peter went to Chicago.

In this case, one of the conjuncts is itself a conjoined structure. If the embedded conjunction is to be handled in the same way as the main conjunction, then M' must be able to create a copy of itself (M'') as follows:

TR1 Both Fred and went to Chicago.

TR2

TR3

H

TR1' either Sue or

TR2'

TR3'

H'

TR1'' Peter

TR2''

TR3''

H''

The question then arises as to how many copies can be created.

(6) #Either Fred or both Sue and either Mike or Peter went to Chicago.

TR1 Either Fred or went to Chicago.

TR2

TR3

H TR1' both Sue and

TR2'

TR3'

H' TR1'' either Mike or

TR2''

TR3''

H'' TR1''' Peter

TR2'''

TR3'''

H'''

The unacceptability of (6) indicates that the number of separate machines which can exist at one time is three. Perhaps coincidentally, the maximum number of tracks in M is also three. This correspondence will be discussed later in this chapter.

6.2 Medially-signalled coordinate structures.

This type of sentence, exemplified by (7)-(9) below, is slightly more complicated than the initially-signalled type.

(7) Sue and Peter went to Chicago.

(8) Sue saw Mike and Fred at the movies.

(9) Sue saw Mike and Fred spoke to Mary.

Clearly, M' cannot be created until after the first conjunct has been uttered. As is demonstrated in (8) and (9), it is not always clear how much of the sentence is the constituent to be conjoined. The first conjunct is thus determined according to what the second conjunct turns out to be. It must therefore be the case that the end of the second conjunct is signalled in some way. Let us posit the following initial hypothesis:

The end of the second conjunct is signalled by a sequence which is ill-formed in that context.

This is the case in (8): Fred at the movies is not a possible constituent. However, the initial hypothesis can easily be shown to be too strong. Consider the following:

(10) Fred knows that his wife and their children will arrive at five o'clock.

In this case, the second conjunct ends after children, but their children will arrive at five o'clock is a possible sequence. It would seem, therefore, that while the second conjunct is being processed, some sort of matching is taking place between it and the material before the conjunction. I will not go into the details of how this matching could be done, since it is not essential to the thesis, and since the problem has been dealt with by others working on the automatic parsing of sentences. (Woods, Kaplan, among others).

The end of a conjunct can thus be signalled in one of two ways, first, by an impossible sequence, and second, by the failure to find a suitable match in the first conjunct.

(11) Fred and Sue and Peter and Michael and Joe went to the movies.

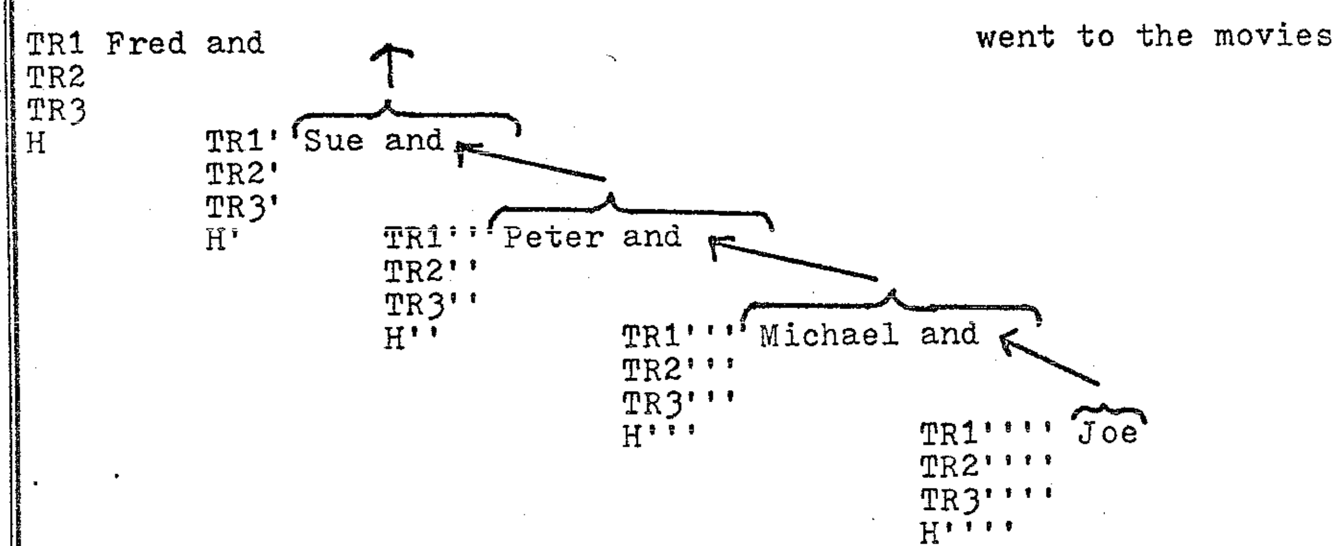
Theoretically, this conjoined series of NP's could be structured in many ways. A few of the possibilities are shown below:

(a) [Fred and [Sue and [Peter and [Michael and Joe]]]]

(b) [Fred and [Sue and Peter] and [Michael and Joe]]

(c) [[[[Fred and Sue] and Peter] and Michael] and Joe]

If (a) were the right structure, then the production configuration would be the following:



Given that this structure would require five copies of the processing machine, and we have said that only three can exist at one time, either the sentence is unacceptable, or we have posited the wrong structure. Since the sentence is quite acceptable, we must conclude that the conjunctions are not structured in this way. Consider the following alternative:

TR1 Fred and ^ and ^ and ^ and ^ went to the movies

TR2

TR3

H TR1' Sue / Peter / Michael / Joe /

TR2'

TR3'

H'

In this case, only two machines exist at any one time. M' is repeatedly created and destroyed, after each and NP sequence. Notice that M' is being used in a way which is analogous to the use of TR2 in (12) below.

(12) I know the man who owns the cow that kicked the dog that chased the cat that scratched my sister.

TR1 I know the man ^ ^ ^ ^

TR2 who owns the cow/that kicked the dog/that chased the cat/that scratched my sister

TR3

H

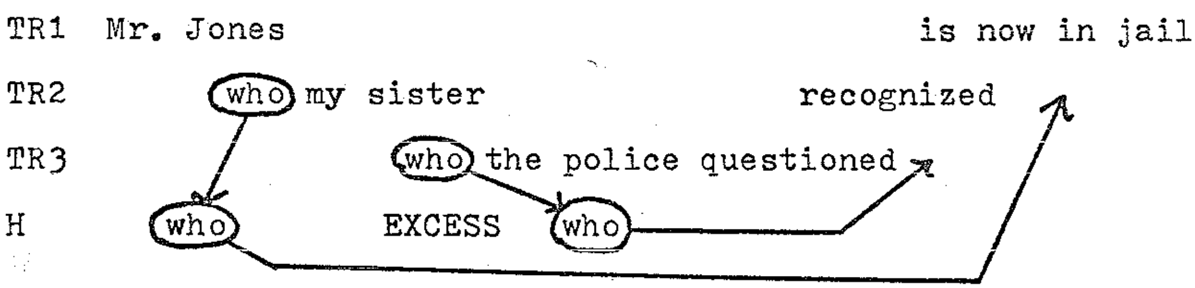
6.3 Non-restrictive relative clauses

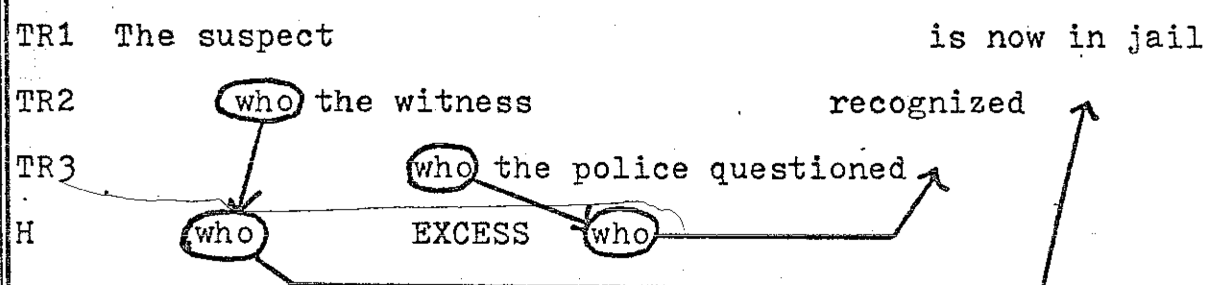
In this section it will be shown that non-restrictive relative clauses must be handled in approximately the same way as conjoined structures; that is, they are processed, not on a lower track of M, but rather in M'.[2] Consider the following sentences:

(13) Mr. Jones, who my sister, who the police questioned, recognized, is now in jail.

(14) #The suspect who the witness who the police questioned recognized is now in jail.

If non-restrictive relative clauses were processed in the same way as restrictive relative clauses, (13) and (14) would have identical production configurations, as shown below.



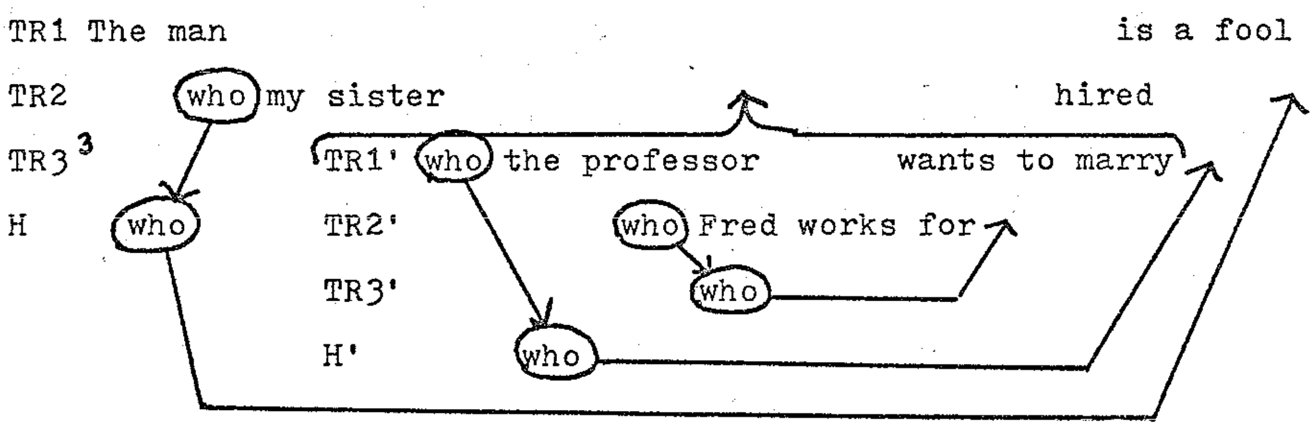


There would thus be no way of accounting for the difference in acceptability between (13) and (14).

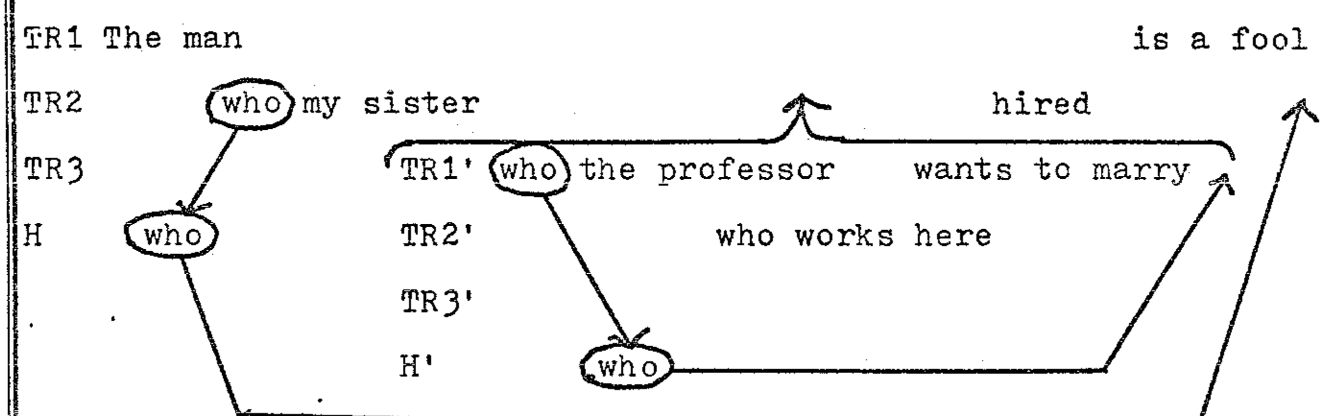
Given that a non-restrictive relative clause does not serve to identify the head NP, but rather provides additional information about an NP whose referent is presumed to be already known to the hearer, it does not seem unreasonable to suggest that this type of clause is processed apart from the sentence in which it is embedded.

The following examples test whether the processing of non-restrictive relative clauses in a separate machine makes the right predictions.

(15) #The man who my sister, who the professor who Fred works for wants to marry, hired, is a fool.



(16) #The man who my sister, who the professor who works here wants to marry, hired, is a fool.



Since (15) and (16) are both judged unacceptable by native speakers, and are predicted to be acceptable by the production configurations shown, it is clear that if an entire production machine is created as a copy, enough space is then available to process unacceptable sentences. Two possible alternatives come immediately to mind:

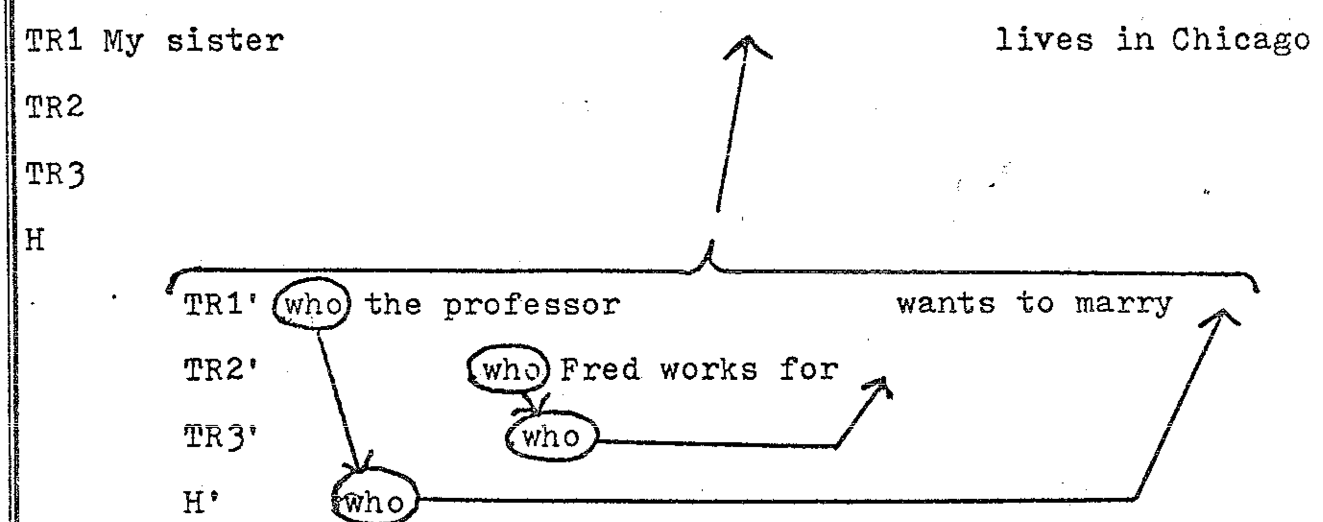
1. M' only includes the space which is subordinate or equal to

the track in M which the last word uttered is on.
For example, if the head NP of the non-restrictive relative clause was on TR2, then the copy would consist of TR2', TR3' and H'. If the head was on TR3, then the copy would consist only of TR3' and H'.

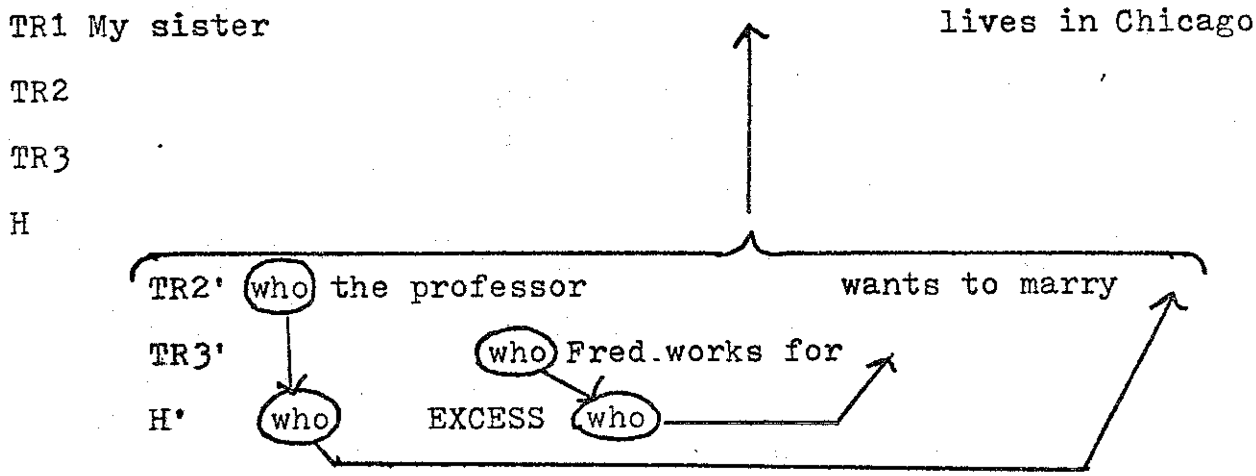
2. If the construction which causes the copy to be created is a two-track embedding, then the copy only contains those tracks which are subordinate to the track which the head NP is on. Otherwise, the copy contains those tracks subordinate or equal to the one which the material immediately preceding the copy creation is on. Consider now the following sentences with regard to the hypotheses stated above.

(17) My sister, who the professor who Fred works for wants to marry, lives in Chicago.

Hypothesis 1 predicts, correctly, that (17) is acceptable.



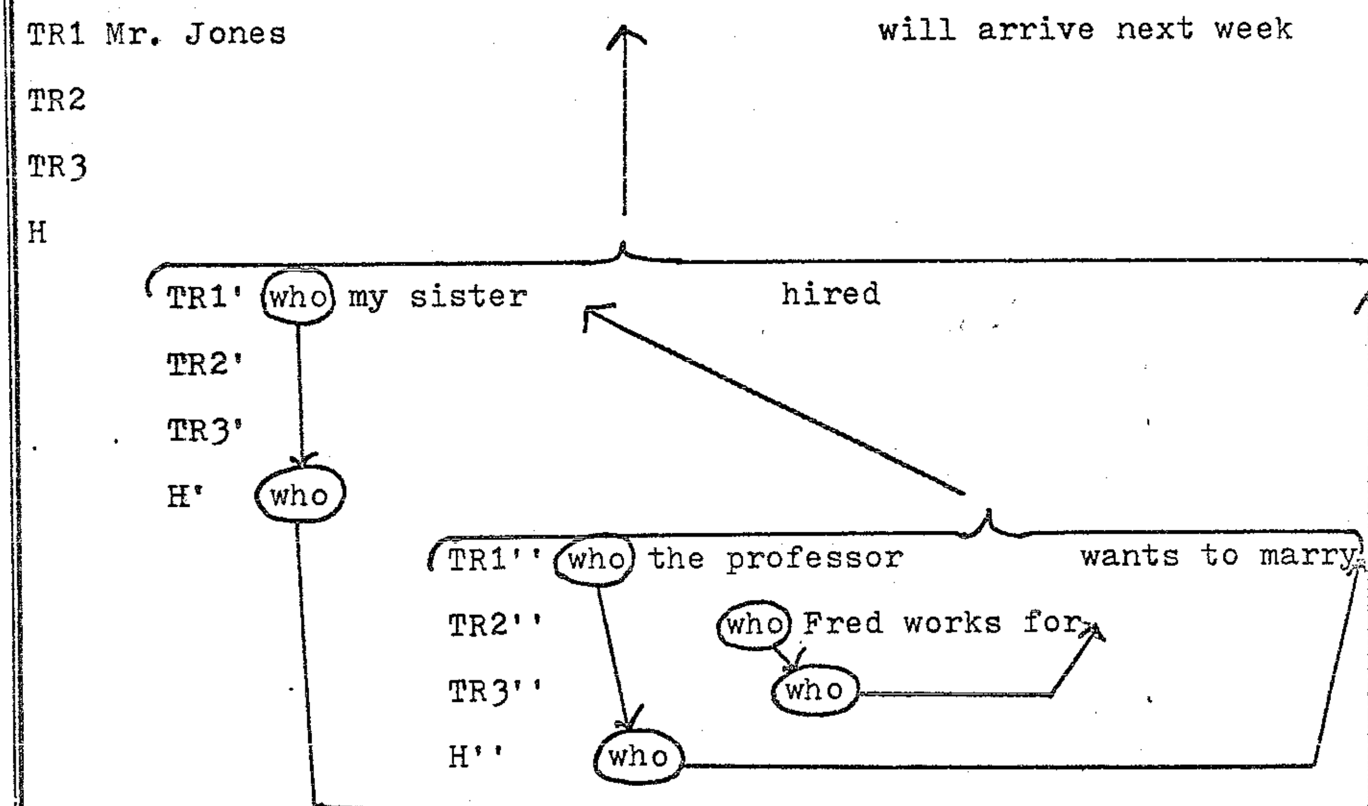
Hypothesis 2 predicts that it should be unacceptable.



So far, hypothesis 2 has been shown to be inadequate.

(18) #Mr. Jones, who my sister, who the professor who Fred works for wants to marry, hired, will arrive next week.

Hypothesis 1 predicts that this sentence should be acceptable, as follows:



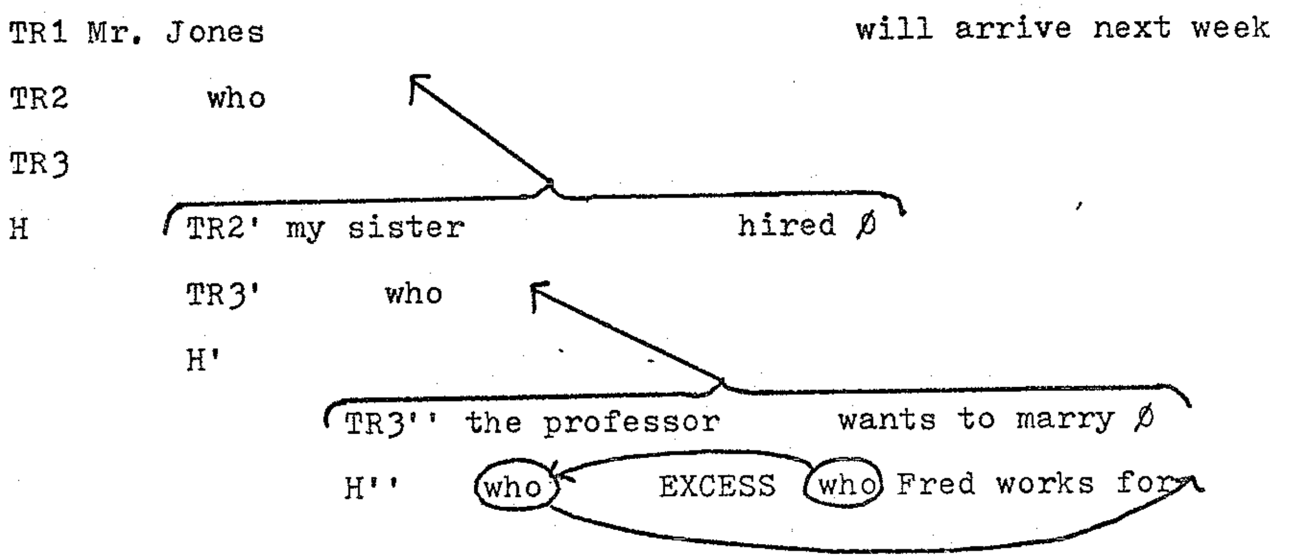
Both initial hypotheses have thus been shown to be inadequate; hypothesis 1 because it allows too much processing space, and hypothesis 2 because it allows too little. Notice that for each copy, hypothesis 1 allows just one more track than does hypothesis 2. Since the correct analysis must allow more space than hypothesis 2, and less than hypothesis 1, it must differ from either of them other than in the number of tracks allowed. Hypothesis 3, then, is a variation of hypothesis 2.

Hypothesis 3. The copy of the processing machine is not created until the relative pronoun beginning the non-restrictive relative clause is uttered. The copy consists of those tracks subordinate or equal to the track holding the relative pronoun. The relative pronoun is thus stored, not in the copy, but in the original. Since all relative clauses require a separate track, the relative pronoun is stored in the original, on the track below the head noun. Once the relative clause is returned to the original machine, it is stored on the same track as the relative pronoun, and then promoted with the relative pronoun to the track containing the head NP.

The fact that the relative pronoun is held in a different machine from the one in which the relative clause is processed means that the relative clause must be processed with a hole in it. Only after the processed relative clause is returned to the original machine is the relative pronoun, and therefore the head NP as well, hooked up with the relative clause.

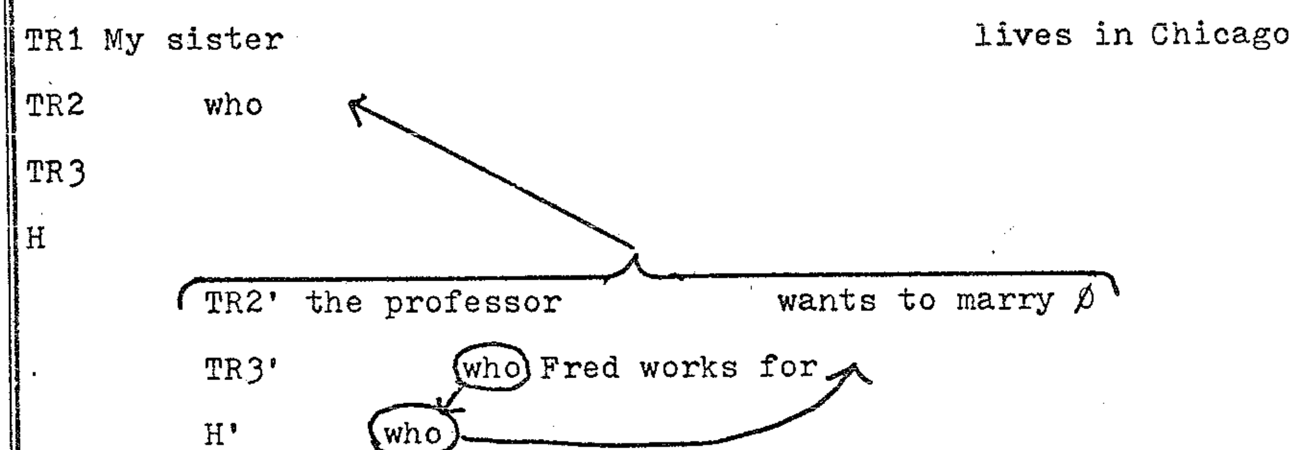
The following examples test the predictions made by Hypothesis 3.

(19) #Mr. Jones, who my sister, who the professor who Fred works for wants to marry, hired, will arrive next week.



Since the most deeply embedded clause, <u>who Fred works for</u>, exceeds the space available in M'', the sentence is correctly predicted to be unacceptable.

(20) My sister, who the professor who Fred works for wants to marry, lives in Chicago.



Since in the case of non-restrictive relative clauses, the relative pronoun does not occupy storage space in the machine in which the relative clause is processed, the location of the extraction hole in a non-restrictive relative clause should not affect acceptability. Recall that in the case of restrictive relative clauses, the location of the extraction hole did affect acceptability, as shown by the following two examples:

(21) #The man who the woman who Fred hired Ø married lives next door.

TR1 The man lives next door

TR2 who the woman married

TR3 who Fred hired

H who EXCESS who

(22) The man who the woman who hired me married lives next door.

TR1 The man lives next door

TR2 who the woman married

TR3 who hired me

H who

Equivalent examples in which the relative clauses are non-restrictive should not differ in acceptability, as follows:

(23) Mr. Jones, who my sister, who lives in Chicago, works with, is getting married.

TR1 Mr. Jones is getting married

TR2 who

TR3

H TR2' my sister works with Ø

TR3' who

H'

TR3'' Ø lives in Chicago

H''

(24) Mr. Jones, who my sister, who I saw yesterday, works for, lives in Chicago.

TR1 Mr. Jones lives in Chicago

TR2 who

TR3

H (TR2' my sister works for ∅)

TR3' who

H'

(TR3'' I saw ∅ yesterday.)

H''

Since (23) and (24) exhibit the same degree of acceptability, we conclude that hypothesis 3 is essentially correct.

6.5 Summary and Re-evaluation

In this chapter we have considerably developed and expanded the parallel processing model. Given these modifications, it seems profitable to reconsider the machine in its entirety.

The basic processing machine has three tracks and one hold cell. This basic machine may be copied, in whole or in part, to process what I have very loosely termed co-ordinate structures. There are several principles governing how the processing space is utilized, as follows:

1. Sentences invariably begin on TR1 of the original machine (M). Separate tracks are required only for post-nominal restrictive relative clauses, NP-complements, and center-embedded sentences in languages which have clause-final complementizers.
2. An NP which has been shifted to the left must be held, apart from its clause, until the extraction hole is reached.
3. (Pushdown store) If two NP's are being held at once, they must be restored to their extraction holes in the opposite order from that in which they were stored.
4. (Same track principle) An NP which is held must always return to the track from which it came.
5. (Promotion Principle) A sentence processed on a subordinate track (TR2 or TR3) is promoted to the track immediately above it as soon as it is complete (<u>on the table</u>).
6. (Sentence integrity constraint) In languages with clause-final complementizers, any non-topical constituent which occurs to the left of its own clause, separated from that clause by one or more embedded sentences, must be held until its own clause shows up.

7. (Copy Isolation Constraint) If a copy of the basic machine is created, control can pass to the copy from the original only at the beginning of the conjunct or relative clause, and can pass back to the original only after the constituent in the copy has been completely processed.

The question now arises as to how powerful the machine is. Clearly, if the manipulations required of the machine are excessively powerful, then the model loses much of its value.

Kuno (personal communication) has suggested that the capacity to create copies gives the machine excessive power, and that if it were possible to eliminate the copies entirely, the model would be more elegant. I agree entirely with his suggestion that a single processing machine would be preferable to the one which I have proposed. However, I have been unable to construct a single machine which makes the right predictions about the acceptability of sentences containing conjunctions and non-restrictive relative clauses.

One aspect of the model, as it is presently stated, gives the impression that the machine is extremely powerful. This is the notion that the machine is capable of creating copies of itself. Fortunately, this is not an essential part of the model. One could simply state that just as there are three tracks available, there are three machines available. We then have to explicitly state the constraints on how the three copies can be used. These are as follows:

1. If control is in TRi when it passes from M to M', it can only pass to TRi', and must return to TRi when the constituent processed

in M' is completely processed.

2. The same constraints apply in M' and M'' as in M, so that if control passes to TRi', any tracks higher than TRi' in M' are unavailable for processing deeply embedded material.

It would seem that the two statements of the model are functionally equivalent, although the first statement may be formally more powerful than the second.

Let us now examine the question of whether the principles stated above can be integrated into a more concise statement. First note the following similarities between subordinate tracks and subordinate machines.

1. Just as there are no more than three tracks, there are no more than three machines.

2. The promotion principle applies to subordinate machines, as well as to subordinate tracks. Material processed in a subordinate machine or track is promoted to the machine or track immediately above it as soon as it is processed.

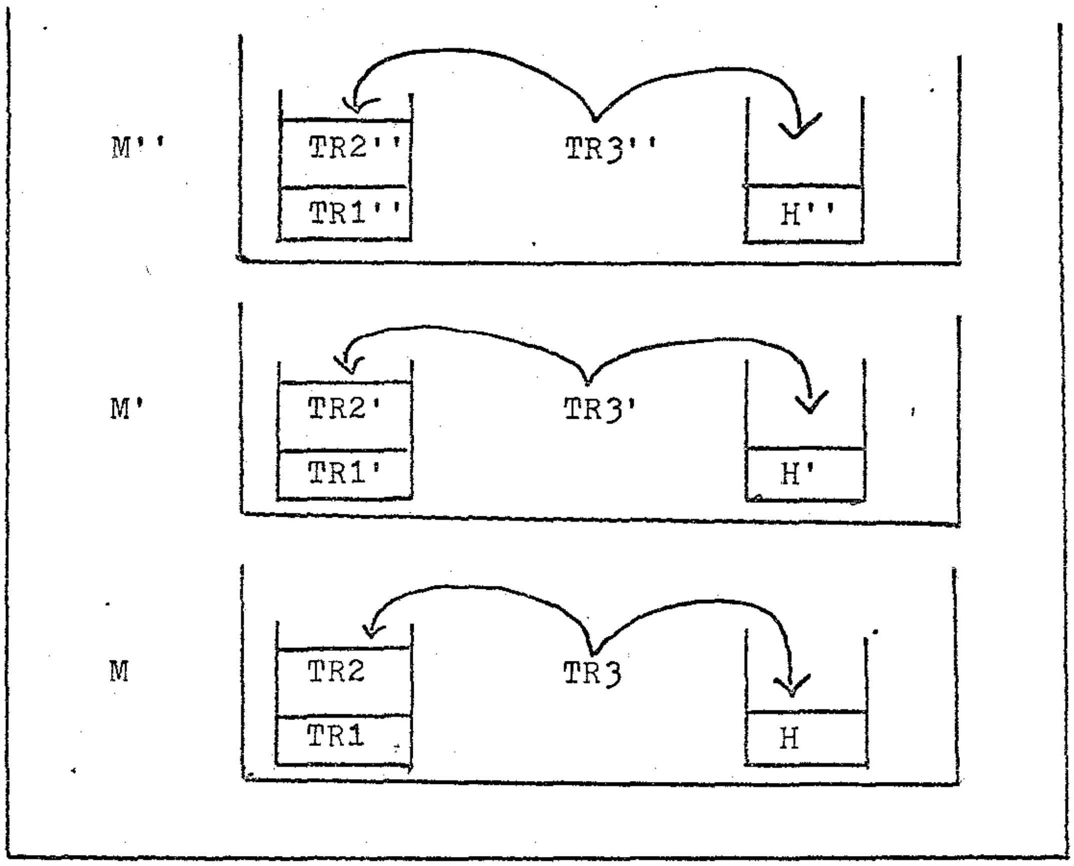
3. Just as subordinate tracks become available after the material on them has been promoted, subordinate machines also become available in this way.

4. The three tracks behave like a pushdown store, as follows: If there is material on TRi, then TRi-1 is inaccessible until the material on TRi has been promoted to TRi-1. The three machines behave in exactly the same way.

We can thus view the entire processing machine as a pushdown store of capacity 3. This pushdown store contains the machines, each of which is a pair of pushdown stores. The first member of the pair contains tracks, the second member hold cells. TR3 can be in either of these two pushdown stores, but cannot be in both of them at once. The question arises as to what each track is. Since I have not provided an actual parser which takes as input a sentence and provides as output a structure which could serve as input to a semantic processor, the answer to this question is not apparent. However, since no permutation of constituents can take place on a track without the use of a hold cell, I would like to claim that each track has at most the power of a pushdown store. It may even be the case that each track is a finite state device, but at present, I have no grounds for claiming that it is.

It is well known that for any pushdown device with finite capacity, there is a finite state device which is weakly equivalent. Given that all the pushdown stores in the parallel processing model are of finite capacity, we can conclude that, formally, the parallel processing model is no more powerful than a finite state device. However, the relevance of strictly formal notions of power to the evaluation of linguistic models has been questioned, and I do not wish to make any claims about the merit of the parallel processing model based solely on its formal power.

The following is a schematic representation of the parallel processing model.



The above schematic is somewhat misleading, in that it gives the impression that all of the space in the model can be filled at once. This is not the case, given the restrictions stated above on the ways in which control can pass from one machine to another, and from one track to another.

FOOTNOTES TO CHAPTER 6

1. The question of whether the creation of copies gives the machine too much power will be discussed later in this chapter.

2. This treatment of non-restrictive relative clauses implies no necessary agreement with the analysis of relative clauses proposed by Thompson (1971).

3. This notation does not mean that TR1' is contained in TR3. The two machines are collapsed here to save space.

Chapter 7. SYNTACTIC IMPLICATIONS

7.0

Having developed the parallel processing model primarily to account for the facts of sentence acceptability, it is interesting to examine what, if anything, the model can contribute to syntactic research; in other words, whether the model provides motivation for certain syntactic phenomena, or perhaps a simpler or more adequate statement of some syntactic rule, generalization or constraint. There are three specific areas where the parallel processing model seems to have something to contribute. This chapter is devoted to a discussion of these three areas in turn, first as they have traditionally been treated, and then in terms of the parallel processing model.

7.1 Non-restrictive relative clauses in Japanese

English non-restrictive relative clauses differ from restrictive relative clauses in a number of ways. First, they are invariably signalled by so-called comma intonation. Secondly, the relative pronoun, which can be deleted under certain circumstances in restrictive relative clauses, can never be deleted if the relative clause is non-restrictive. Third, non-restrictives, when center-embedded, are more acceptable than restrictives in the same situation. Fourth, it has been claimed that they are derived from an entirely different underlying structure than restrictive relative clauses.

Kuno (1973) has shown that none of the above distinctions obtain between restrictive and non-restrictive relative clauses in Japanese. In a standard syntactic analysis, one can only state that this is the case; transformational grammar provides no principled reason for it. A left-to-right model, on the other hand, allows a fairly straightforward explanation, as follows:

Since Japanese relative clauses are prenominal, the hearer has no way of knowing whether or not the head NP is already known to him until after the relative clause has been uttered. He therefore does not know whether the relative clause will contain information crucial to his identification of the referent of the head NP. Given this uncertainty, he must proceed under the assumption that the relative clause is essential; in other words, that it is restrictive.

Strictly in terms of the parallel processing model, we can also say the following:

In order for a machine copy to be used, there must be an overt trigger. In English non-restrictive relative clauses, the trigger is the comma intonation plus the relative pronoun. In Japanese, even if there were a relative pronoun, it would not occur at the beginning of the embedded clause. As such, there can be no overt trigger, and therefore no copy can be used. Non-restrictive relative clauses, must, therefore, be processed in the original machine, in the same way as restrictive relative clauses.

7.2 The Complex NP Constraint

This constraint is stated in Ross (1967, p. 70) as follows:

> No element contained in a sentence dominated by a noun phrase with a lexical head noun may be moved out of that noun phrase by a transformation.

In the parallel processing model, the constraint which serves the same purpose as the complex NP constraint is the same track constraint, which can be stated as follows:

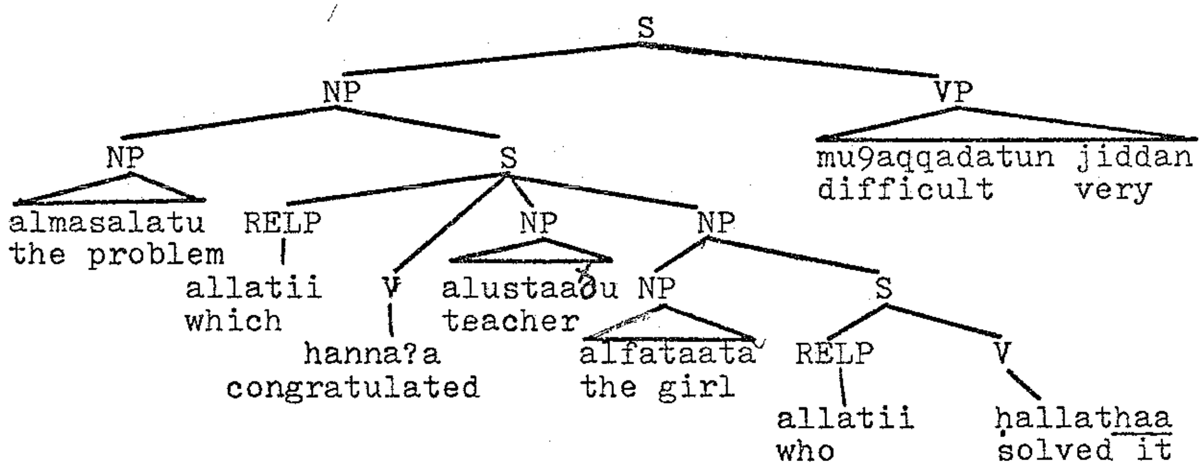
> A constituent which is removed from a track and stored in H or TR3 must be returned to the track from which it came.

The question arises as to whether these two constraints are empirically different, and if so, which one of the two is more adequate. Recall that in chapter 5, there were many examples of grammatical Japanese sentences which violated the complex NP constraint. The same-track constraint correctly predicted that these sentences should be grammatical.

Another class of counterexamples to the complex NP constraint was discussed by Awwad (1973). These involve wh-questions and relative clauses in Arabic.

(1) almasalatu allatii hanna?a alustaaðu alfataata allatii
the problem which congratulated the teacher the girl who
ḥallathaa mu9aqqadatun jiddan
solved it difficult very

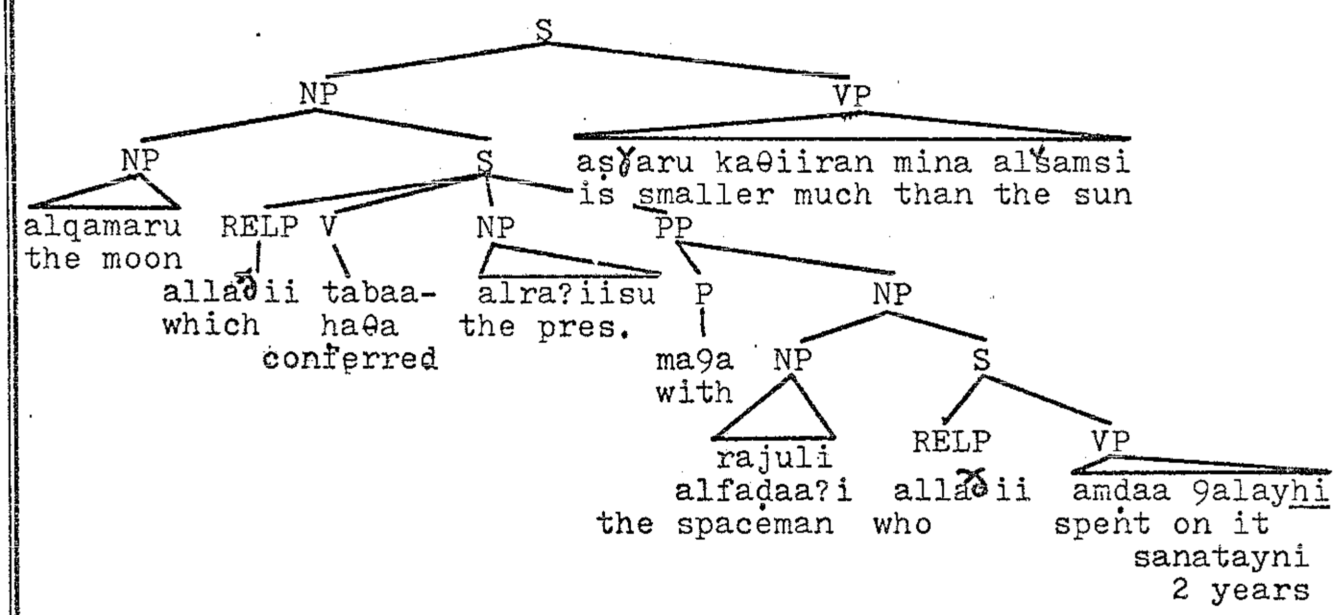
"The problem which the teacher congratulated the girl who solved (it) was very difficult.



(2) alqamaru allaðii tabaahaθa alra?iisu ma9a rajuli alfaḍaa?i
the moon which conferred the pres. with the spaceman

allaðii amḍaa 9alay<u>hi</u> sanatayni aṣɣaru kaθiiran mina alšamsi
who spent on it 2 years is smaller much than the sun

"The moon, which the president conferred with the spaceman who spent two years on (it), is much smaller than the sun."

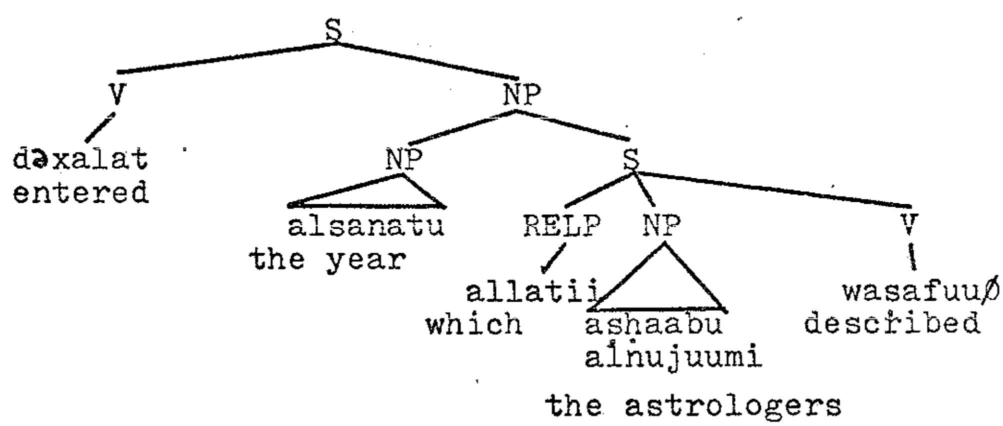


When an NP is relativized in Arabic, a so-called returning pronoun occurs; in other words, in addition to the relative pronoun, which occurs clause-initially, there is also a pronominal copy of the relativized NP inside the relative clause. These relative clauses, therefore, do not have extraction holes. In Modern Standard Arabic, although not in Palestinian or Lebanese colloquial Arabic, the returning pronoun may be deleted under certain conditions.[1] It is these conditions which are most interesting.

First, non-subject returning pronouns may only be deleted in those cases where the complex NP constraint is not violated. Consequently, (1) and (2) above would be ungrammatical if the returning pronouns haa and hi were deleted. The following examples, on the other hand, are grammatical without the returning pronoun, since there is no violation of the complex NP constraint.

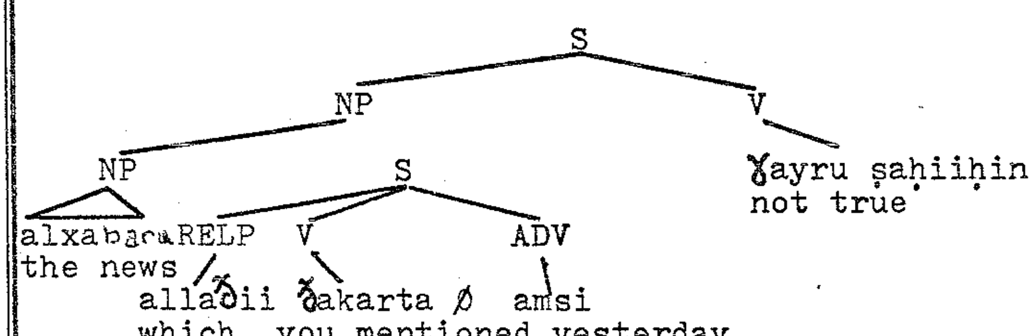
(3) dəxalat alsanatu allatii aṣḥaabu alnujuumi waṣafuuØ
entered the year which the owners of the described stars

"The year which the astrologers described started."



(4) alxabaru allaðii ðakarta Ø amsi ɣayru ṣaḥiiḥin
the news which you mentioned yesterday not true

"The news which you mentioned yesterday is not true."



Secondly, returning pronouns which are subjects can always be deleted. Awwad has the following to say about this:

> The deletion of the returning pronoun when it is in the nominative does not affect the grammaticality of sentences derived by relativizing an element of a relative clause. This is so because this pronoun is also realized as a person marker on the verb. (Awwad, 1973, p. 122)

Simply stated, then, the facts are as follows: Arabic relative clause formation is not subject to the complex NP constraint just in case there is some sort of pronominal copy of the relativized NP, either a person marker on the verb or a returning pronoun, left behind inside the relative clause.

Consider now how the parallel processing model would handle the situation. First, recall that a stored constituent is returned to its track only when its extraction hole is reached. Crucially, if there is a copy of the relativized NP inside the relative clause, then there is no extraction hole. The stored NP will thus not be returned to the relative clause while that clause is being processed.

To make this clearer, consider the production configuration for (1) above:

TR1 almasalatu mu9aqqadatun jiddan

TR2 allatii hanna?a alustaaðu alfataata

TR3 allatii hallathaa

H allatii

Recall the conditions for promotion of a sentence to a higher track:

1. The sentence must be on the table, i.e. it must be syntactically complete, and the first word of another clause (or the end of the sentence) must have been encountered.

2. There must be no stored NP's in H or TR3, tagged for the track which the sentence was processed on.

Given these two conditions, there is no reason why <u>allatii ḥalathaa</u> cannot be promoted to TR2 as soon as <u>mu9aqqadatun</u> is encountered. The contents of TR2 (<u>hanna?a alustaaðu alfataata allatii ḥallathaa</u>) cannot be promoted to TR1, however, until the relative pronoun in H, <u>allatii</u>, has been returned to its track. At this point, the returning pronoun, which is now on TR2, is matched up with the relative pronoun, thus removing the relative pronoun from H.

The returning pronoun, then, can be viewed as a means by which an embedded sentence can be made to satisfy the conditions for promotion. The restoring of the held NP to the clause in which it belongs is thus postponed until the clause is on the track which the held NP is tagged for.

Since Arabic has a productive rule of subject pronoun deletion, a sentence can be syntactically complete without a subject pronoun, and as such there is no need for a returning pronoun in subject position.

This analysis also predicts that if returning pronouns can occur in English, then they should allow speakers to violate the Complex NP Constraint. Many speakers do not allow returning pronouns at all, but for those speakers who do, the predictions made by the parallel processing model are correct, as follows:

(5) *Fred bought the book that I know the man who wrote.

TR1 Fred bought the book

TR2 that I know the man

TR3 who wrote

H hat

Since S_3 (who wrote) is not syntactically complete, it cannot be promoted to TR2. That cannot be retrieved from H, since it is tagged for TR2, and control is on TR3. The production is therefore blocked.

(6) Fred bought the book that I know the man who wrote it.

TR1 Fred bought the book

TR2 that I know the man

TR3 who wrote it

H that

In this case, S_3 (who wrote it) is syntactically complete, and as such can be promoted to TR2. The contents of TR2 (I know the man who wrote it) cannot be promoted until that is retrieved from H. That is matched with it, and then the whole clause is promoted to TR1, successfully completing the production.

Let us now examine some cases of wh-questions in Arabic, where a situation obtains which is similar to the case of relative clauses.

First of all, we need to give some facts about the rules forming wh-questions in Arabic. As Awwad states it, there are two alternatives: a chopping rule very much like the one forming wh-questions in English, and a copying rule involving a relativization process as well. These two possibilities are illustrated in examples (7) - (9) below. From a sentence like (7) we can question the direct object by the chopping rule as in (8), or by the copying rule, as in (9).

(7) šariba alqiṭṭu alḥaliiba
    drank the can the milk

    "The cat drank the milk."

(8) maaðaa šariba alqiṭṭu ?
    what drank the cat ?

    "What did the cat drink?"

(9) maa alšay?u [allaðii šaribahu alqiṭṭu]?
    what the thing which drank it the cat

    "What is the thing which the cat drank?"

Since the chopping rule does not allow for a returning pronoun, one might expect it to be subject to the complex NP constraint. This is, in fact, the case, as shown by the ungrammaticality of (10).

(10) *maaðaa hanna?a alustaaðu alfataata allatii ḥallat ∅
     what congratulated teacher girl who solved

     "What did the teacher congratulate the girl who solved?"

The copying rule, on the other hand, does have a returning pronoun and as such is not subject to the complex NP constraint.

(11) maa alšay?u [allaðii hanna?a alustaaðu alfataata [allatii
     what the thing which congrat. teacher girl who
     ḥallathu]]
     solved it

     "What is the thing which the teacher congratulated the girl who solved (it)?"

Just as the returning pronoun can sometimes be deleted in ordinary relative clauses, it can also be deleted in the relative clause which forms part of the copying type of wh-questions. This deletion is subject to the same conditions as in ordinary relative clauses; that is, it can only happen in cases where the complex NP constraint is not violated. The same track constraint correctly predicts these facts. The following are the production configurations for (8)-(11).

(8a) TR1 maaðaa šariba alqiṭṭu
TR2
TR3
H maaðaa

(9a) TR1 maa alšay?u
TR2 allaðii šaribahu alqiṭṭu
TR3
H allaðii

(10a) TR1 maaðaa hanna?a alustaaðu alfataata
TR2 allatii ḥallat Ø
TR3
H maaðaa

In this case, the production is blocked. S_2 (allatii ḥallat) cannot be promoted because it is syntactically incomplete. Maaðaa cannot be retrieved from H to complete it because it is tagged for TR1.

(11a) TR1 maa alšay?u
TR2 allaðii hanna?a alustaaðu alfataata
TR3 allatii ḥallathu
H allaðii

In this case, S_3 (allatii ḥallathu) is syntactically complete, and can be promoted to TR2. The contents of TR2 cannot be promoted to TR1 until allaðii is retrieved from H. Allaðii is matched up with the returning pronoun hu, allowing promotion to TR1 to take place. The production is therefore complete.

We have now shown that the same track constraint is empirically more adequate than the complex NP constraint, since it makes the right predictions for two classes of serious counter-examples to the complex NP constraint: Japanese multiple-headed relative clauses, and Arabic relative clauses and wh-questions.

7.3 Kitestring tangling

Consider the following sentences:

(12) It's John that I don't know what to say to.

(13) *It's John that I don't know what to talk to about.

Both of these sentences violate the double-hole constraint (Kuno & Robinson). The double hole constraint states that if an NP is extracted from a sentence, then that sentence becomes an island, and no further elements may be extracted by any later rule. Both (12)and (13) should therefore be ungrammatical. This problem has been discussed by Kuno & Robinson, who note that, in general, double hole violations where the second extraction does not extract an NP from between the hole left by the first NP and the place where that first NP ends up, are in general much more acceptable. This is the case with (12) and (13) above, as shown by the diagrams below:

(12a) It's John that I don't know what to say ∅ to ∅

(13a) *It's John that I don't know what to talk to ∅ about ∅

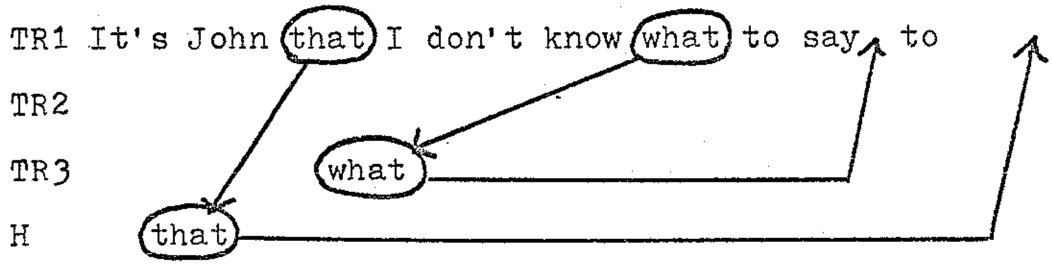
I shall refer to the situation in (13) as a case of kitestring tangling.

Various proposals have been advanced to handle this problem; however none of them can satisfactorily account for the grammaticality of (14) (again, only for those speakers who allow returning pronouns.)

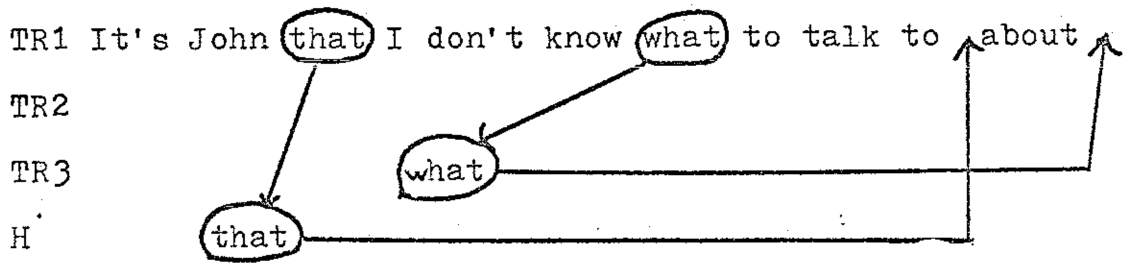
(14) I don't know which people Communism would be easy to talk to them about.

The parallel processing model provides a fairly simple account for the above facts. Recall that when two NP's are being held at the same time, they must be restored to their tracks in the opposite order from that in which they were stored. In other words, H and TR3 combine to behave like a pushdown store. The following are the production configurations for (12) and (13) above:

(15)

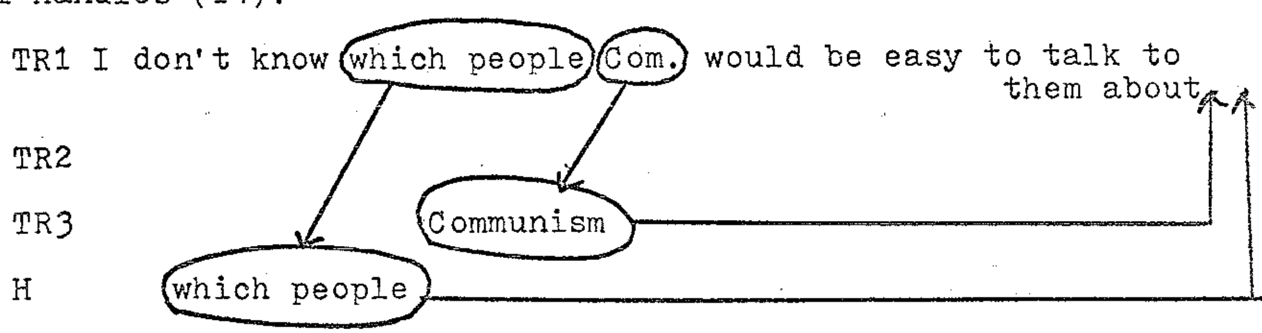


(16)



So far, the parallel processing model correctly predicts that (12) is grammatical and (13) is ungrammatical. Consider now how the model handles (14):

(17)



Since the extraction hole for <u>which people</u> is occupied by the returning pronoun <u>them</u>, there is no need to retrieve the held NP from H when <u>them</u> is reached. When the end of the sentence is reached, <u>Communism</u> has been retrieved from TR3. The sentence is not complete, however, until <u>which people</u> has been returned to its track. Now that <u>Communism</u> has been removed from TR3, <u>which people</u> is accessible. It is then matched with <u>them</u>, and the production is complete.

Consider now the contrast between (14) and (18) below:

(18) *I don't know which people Communism would be easy to talk to about it.

Clearly, it is not the case that if either one of the holes is occupied by a returning pronoun, the sentence will be grammatical. The parallel processing model again makes the correct predictions, as follows:

TR1 I don't know (which people) (Com.) would be easy to talk to $\emptyset$ about it

TR2

TR3 (Communism)

H (which people)

When the extraction hole after <u>to</u> is reached, the only NP accessible to fill it is <u>Communism</u>. The NP which belongs there, <u>which people</u>, cannot be retrieved as long as <u>Communism</u> is on TR3. The returning pronoun does no good, since it occurs after the production would already have been blocked.

The same situation also arises in French, as illustrated by the following examples:

(19) Voici les étagères sur lesquelles je ne sais pas quels livres
here the shelves on which I know not which books

je devrais mettre
I should put

"Here are the shelves that I don't know which books to put on."

TR1 Voici les étagères

TR2 (sur lesquelles) je ne sais pas (quels livres) je devrais mettre $\emptyset\emptyset$

TR3 (quels livres)

H (sur lesquelles)

(20) *Voici les livres que je ne sais pas sur quelles étagères je
here the books that I know not on which shelves I

devrais mettre.
should put

"Here are the books that I don't know what shelves to put on."

TR1 Voici les livres

TR2 que je ne sais pas sur quelles étagères je devrais mettre ∅ ∅

TR3 sur quelles étagères

H que

(21) Voici les livres que je ne sais pas sur quelles étagères je
here the books that I know not on which shelves I

devrais les mettre.
should them put

"Here are the books that I don't know what shelves to put them on."

TR1 Voici les livres

TR2 que je ne sais pas sur quelles étagères je devrais les mettre.

TR3 sur quelles étagères

H que

FOOTNOTES TO CHAPTER 7

1. Perlmutter (1972) has also discussed this question, and proposes that relativization be invariably a copying rule, with a subsequent rule of Shadow Pronoun Deletion, which is sensitive to the island constraints.

Chapter 8. CONCLUSION

In the preceding chapters, we have constructed a theoretical machine for the processing of sentences. Although no claim is made that the machine reflects the actual processing of sentences by human beings, we can claim that the machine allows us to predict correctly whether or not a sentence is acceptable.

We must now ask what the relationship is between a processing model of this type and the current theories of grammar. Clearly, the parallel processing model as it now stands is completely inadequate as a replacement for the syntactic component of a generative grammar. We have proposed no mechanism for enumerating the grammatical sentences of a language; we have provided a means for distinguishing acceptable sentences from unacceptable ones. Given that, we must therefore ask whether the parallel processing model is compatible with a generative grammar, and if so, how they would interact. The most obvious problem is the fact that while the parallel processing model works left-to-right, most transformational grammars operate cyclically, from the most deeply embedded clause to the matrix sentence. In most cases, unless the surface structure is entirely left-branching, these two orders of operation do not coincide. It is therefore the case that the parallel processing model could not be used as a framework for producing sentences according to the rules of a generative transformational grammar. This impossibility lends further support to the already widely accepted statement that a generative transformational grammar is not rule by rule, a real model of sentence production.

The parallel processing model could, on the other hand, serve as a filtering device, which would take as input the surface structures produced by a generative grammar, and filter out those which are unacceptable. However, this defeats the whole purpose of having a model which works left-to-right, since the ultimate goal of a left-to-right model is, in fact, to reflect the way speakers actually produce sentences. If the parallel processing model worked as a filtering device in this way, then the actual production of the sentence would not be left-to-right at all.

Clearly, then, in order to construct a complete model of sentence processing, there must be some kind of grammar which would work left-to-right. This grammar could not be a simple phrase-structure grammar, such as Yngve's, for the following reason. Just as cyclically applied rules do not work left-to-right for a right-branching structure, phrase structure rules do not work left-to-right for a left-branching structure. A left-to-right grammar must be able, not only to expand nodes, but also to build higher structure.

Several lines of inquiry are possible at this point. First, one might try to construct a left-to-right grammar which would be compatible with the processing model presented here. Secondly, one could explore further contributions that the model makes to syntactic theory. Third, one could try to test the psychological reality of parts of the model. Some aspects of the model which appear to be testable at this point are as follows:

1. The claim that once a sentence is on the table, it is no longer in an active state with respect to syntactic analysis.
2. The claim that no clause can be promoted if there remains in H or TR3 a constituent tagged for the track which that clause is on. This claim implies that the sentence should be rejected

at the point where promotion would have taken place. For example, sentences (1) and (2) below would be rejected at the points shown by /.

(1) The book that Fred read the story to me was/on the table.

(1a) The book that Fred read the story to me out of was on the table.

(2) It's Sue that Fred told me the dean wanted to see Mary. /

(2a) It's Sue that Fred told me the dean wanted to see Mary about.

(3) I don't know any people that Communism would be easy to talk to Fred about. /

(3a) I don't know any people that Communism would be easy to talk to them about.

Notice that in (3a) above, one could claim one of two things: either (i) that the returning pronoun <u>them</u> is matched with <u>people</u> as soon as <u>them</u> is uttered, or (ii) that the matching does not take place until the end of the sentence, when the presence of the stored NP tagged for TR1 forces an attempted match. If the first claim is correct, then (3) should be rejected when <u>Fred</u> is uttered, since the match would fail, but if the second claim is correct, then the sentence should not be rejected until the end.

3. The notion of varying memory load. The implicit claim is made that when all available processing space is occupied, the speaker's short-term memory is loaded to capacity, whereas if less space is used than is available, the speaker should have some memory "left over" for other tasks.

In sum, even though the problem of constructing a complete model of sentence processing has not been solved, we have discovered what the limitations are on sentence complexity. We have provided a means for stating these limitations in terms of a model which works left-to-right. This model has also allowed us to provide principled reasons for the existence of certain syntactic rules, to capture generalizations about rules previously thought to be unrelated, and to account for some previously unsolved syntactic phenomena. Since the model does work left-to-right, we have also provided a plausible framework for a psychological model of sentence processing.

REFERENCES

Awwad, M. 1973. Relativization and Related Matters in Classical, Modern Standard and Palestinian Colloquial Arabic. Brown University Dissertation.

Chomsky, N. 1957. Syntactic Structures. (Janua Linguarum, Series Minor, 4) The Hague: Mouton.

------. 1965. Aspects of the Theory of Syntax. Cambridge, Mass: MIT Press.

Hockett, C. F. 1958. A Course in Modern Linguistics. New York: Macmillan.

Kimball, J. 1973. Seven Principles of Surface Structure Parsing in Natural Language. Cognition 2. 15-47.

Kuno, S. 1972. Functional Sentence Perspective. Linguistic Inquiry 3. 269-320.

-----. 1972. Natural Explanations for Some Syntactic Universals. Mathematical Linguistics and Automatic Translation, Report NSF-28. The Aiken Computational Laboratory. Harvard University, Cambridge, Mass.

-----. 1973. Constraints on Internal Clauses and Sentential Subjects. Linguistic Inquiry 4. 363-386.

-----. 1973. The Structure of the Japanese Language. Cambridge, Mass.: MIT Press.

-----. 1974. The Position of Relative Clauses and Conjunctions. Linguistic Inquiry 5. 117-136.

-----, & A. Oettinger. 1963. Syntactic Structure and Ambiguity in English. AFIPS Proceedings of the 1963 Fall Joint Computer

Conference. Las Vegas, Nevada.

----, & J. Robinson. 1972. Multiple Wh Questions. Linguistic Inquiry 3. 463-487

Lakoff, G. 1965. On the Nature of Syntactic Irregularity. Mathematical Linguistics and Automatic Translation, Report NSF-16. The Aiken Computational Laboratory, Harvard University, Cambridge, Mass.

Langendoen, T. 1975. Finite-State Parsing of Phrase-Structure Languages and the Place of Readjustment Rules in the Grammar. Linguistic Inquiry 6. 533-554.

Miller, G. 1956. The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review 63. 81-97.

Moore, T. 1967. The topic-comment function: A performance constraint on a competence model. University of California at Los Angeles Dissertation.

Perlmutter, D. 1972. Evidence for Shadow Pronouns in French Relativization. The Chicago Which Hunt.

Reich, P. 1969. The finiteness of natural language. Language 45. 831-843.

Ross, J.R. 1967. Constraints on Variables in Syntax. MIT Dissertation.

-----. 1973. Slifting. in Gross, M., Halle, M., and Schützenberger, eds. The Formal Analysis of Natural Language. The Hague: Mouton.

Thomas, H., & Huff, F. 1971. Comprehension and recoding-time of transformed sentences. Language and Speech 14. 352-372.

Thompson, S. 1971. The Deep Structure of Relative Clauses. in Fillmore, C., and Langendoen, T., eds. Studies in Linguistic Semantics. New York: Holt, Rinehart and Winston, Inc.

Whitman, J. 1976. Rules with too much power: Constraints on Extraction Rules in Japanese. Harvard University Undergraduate thesis.

Yngve, V. 1960. A model and an hypothesis for Language Structure. Proceedings of the American Philosophical Society 104. 444-666