

A computer model for the perception of syntactic structure

J. P. THORNE

English Language Research Unit, Edinburgh University

In his paper 'The problem of serial order in behavior' Karl Lashley (1951, p. 113) points out that 'language presents in a most striking form the integrative functions that are characteristic of the cerebral cortex' adding '...the problems raised by the organization of language seem to me to be characteristic of almost all other cerebral activity'. Some idea of the complexity of the integrative processes involved in speech can be gained from the fact that the adult speaker's ability to produce syllables at an average speed of 210 to 220 a minute (or roughly 14 phonemes per second) means individual muscular events occurring throughout the speech apparatus at a rate of several hundred every second; in the case of some phonemes the total time required to activate the muscles involved in their production being as much as twice as long as the duration of the sound itself. Not very much is known at present about what this involves on the neuronal level, where the rate at which individual events occur must be greater by a large factor, but it is a point of considerable interest that there is at least some evidence to suggest that in some instances the order of neuronal events might be different from that of the muscular events with which they are correlated.*

The point Lashley is making in his paper is that any form of behaviour revealing this degree of complexity in its organization cannot be analysed as an associative chain of reflexes. But, as he points out, in the case of speech the evidence against the associative chain hypothesis is particularly compelling. This arises from considerations of two kinds. The first is the fact that the character of certain sounds is determined not only by the sounds that precede them but also by those that follow them. The second is the fact that the character of certain sounds is determined not only by the sounds in their immediate environment but also by the position they occupy with respect to the syntactic structure of the utterance. To take just one example, the speech of Standard English speakers contains at least twelve varieties (allophones) of the phoneme *t*. But whenever this is the first sound in a word and is immediately followed by a vowel they will always use the aspirated allophone never any of the others. This is clear evidence that in producing utterances speakers follow out principles of organization relating to syntactic structure. To produce a plausible model for speech we have to postulate not only principles of organization more complex than the Markov processes of associative chain theories but hierarchies of organization, elements on one level corresponding to what Lashley calls 'generalized schemata of action' and Miller, Galanter & Pribram (1960) call 'plans' which are carried out on the level below. Evidence in favour of such a model can be obtained from a study of speech disorders, ranging from the transpositions occurring in the speech of a tired or nervous

* For an excellent summary of what is presently known about the physiological and neurophysiological basis of language see Lenneberg (1967, Chapter 3).

speaker to remarks of aphasics indicating that although for the most part they can only produce strings of unintelligible sounds they still 'know what they want to say'. All these disorders can be viewed as involving in some degree a breakdown in integrative functions, an inability to carry out successfully plans for utterances.

Plans for utterances (particularly plans at the syntactic level) have been studied for over 3000 years. Essentially this is what grammatical studies are studies of. Grammatical rules represent specifications that plans for utterances must meet if they are to be well-formed. During this time the main achievement of grammarians has been the development of a terminology for describing these structures. This includes terms like *word*, *sentence*, *noun*, *subject* and *object*. It is important to bear in mind (particularly as it is precisely this point that most grammar books obscure) that the facts about the structure of a language which are conveyed through the use of this terminology must (if they are facts at all) be known to all speakers of that language. For example, the statement that in English noun phrases the definite article precedes the noun states something that must be known to anyone who speaks English, even though, of course, they may have never heard of the terms 'definite article' and 'noun' (or even the term 'word' for that matter). A grammar of a language is a statement of what the native speaker knows about its structure.

Despite the antiquity of grammatical studies it was not until recently that an attempt was made to set out the formal conditions that a grammar must satisfy if it is to count as a reasonable description of the native speaker's knowledge of the structure of his language. This was first done by Noam Chomsky in 1957 in his book *Syntactic structures*. Chomsky shows that a grammar must take the form of a set of effectively computable rules, which must be finite (because it is a model of knowledge 'internalized' by a human being) but which is capable of generating an infinite number of sentences together with their analyses (because the number of well-formed sentences in a natural language is infinite). Chomsky was also the first to set out clearly the empirical conditions that a grammar of a natural language must satisfy. The most important of these is the requirement that the grammar should be capable of generating all and only the well-formed ('grammatical') sentences of the language and that the analyses it assigns to them should accord with the native speaker's intuitions about the structure of the language. They also include the requirement that in the case of an ambiguous sentence, such as *I hate boring students*, the grammar should reflect the ambiguity by assigning two analyses to the sentence—that is, it should contain two sets of rules for generating the sentence.

Not all the information about the syntactic structure of an utterance that is available to the native speaker is directly relatable to what he actually hears. Consider the three sentences (1) *The boy hit the girl*, (2) *The girl was hit by the boy*, (3) *It was the girl that the boy hit*. In sentence 1 *the boy* comes in front of *hit* and *the girl* comes after. In sentence 2 *the girl* comes in front of *hit* and *the boy* comes after. In sentence 3 both noun phrases come in front of *hit*. Yet in all three sentences *the boy* is the subject and *the girl* is the object. Indeed it is clear that

notwithstanding the obvious differences between these three sentences they all have essentially the same syntactic structure (a fact intimately connected with their all being paraphrases of each other). Chomsky has shown that in order to account for all the syntactic information about a sentence available to the native speaker a complete analysis must consist of two components, a surface structure analysis and a deep structure analysis, the former, which is more directly relatable to what is actually heard when the sentence is spoken, often having a quite different organization from that of the latter. The essential similarity in the structure of the three sentences cited above is explained by their all having the same deep structure. This is roughly equivalent to the surface structure of sentence 1. But in sentences 2 and 3 the order of the elements in the surface structure is quite different from their order in the deep structure. In addition they contain elements, like the *it* at the beginning of sentence 3, which have no correlation with elements in the deep structure. But it can also happen that elements can form part of the deep structure which have no correlation with elements in the surface structure. Consider the sentence *The girl I liked left*. Any English speaker hearing this sentence possesses the information contained in the statement that *the girl I liked* is the subject of the sentence and that *left* is the predicate, and that *the* is a definite article, *the girl* a noun phrase, etc. This is all information about the surface structure of the sentence. But he also knows that in this sentence *the girl* is not only the subject of *left* but is also the object of *liked*, even though, of course, the surface structure is *The girl I liked left* and not *The girl I liked the girl left*. Notice that we derive this information from the sentence without making any reference to the context (as here where it is given merely as an example). It is clear therefore that we derive this information not from the context, which is sometimes suggested) but from our perception of the structure of the sentence. In fact it forms a part of the deep structure which is not realized in the surface form of the sentence. A grammar that is intended to be a complete statement of the native speaker's knowledge of the structure of his language must be able to assign deep and surface structure analyses to all well-formed sentences in it. The last few years have seen considerable advances in the construction of grammars for natural languages which satisfy this condition. It must, however, be emphasized that there are many structural characteristics of language which are still not properly understood.

The statement that not all the information about the syntactic structure of utterances available to the hearer is relatable to what he hears must not be taken as referring only to deep structure information. Even surface structure information is not directly relatable to what we actually hear. There is, for instance, nothing about the pronunciation of the sentence *The girl I liked left* that indicates to us that *liked* is the same part of speech as *left* but a different part of speech from *girl*, or that indicates that *the girl* is a phrase but *girl I* is not. There is not even necessarily anything to indicate that *girl* is one word and *girl I* two. The point is so obvious that it is easy to miss but it provides the basis for any reasonable theory of the perception of syntactic structure. For it follows from this that the perception of syntactic structure (and it is indisputable that this is an essential part of the process of understanding an utterance) cannot be merely a matter of what we

hear but must also be a matter of what we know about the structure of our language.

The question therefore arises as to how this knowledge is brought to bear. This paper describes one approach to this problem. It concerns the construction of a computer program designed to simulate the perception of syntactic structure. The potential interest of any such program written at the present time is severely limited by the fact that it has to be designed to work with orthographic rather than spoken input. Even so it seems possible that something can be learned about the perception of syntactic structure from a study of the capacities that have to be built into a program that will assign deep and surface structure analyses to arbitrary sentences in a natural language presented to it in orthographic form.

Before discussing what these capacities are let us first briefly consider the reasons for discarding one obvious way to construct such a program, which is to take a generative grammar as a flow-diagram for a program which, each time a sentence is presented to it, will generate sentences until it produces a match; a print-out of the rules which have produced the matching sentence providing an analysis. A simple calculation shows that this approach is quite impracticable. It has in fact been calculated that even allowing only 1 s for the generation of each sentence it could still take up to 10^{42} s (longer than the time that has elapsed since the creation of the Earth) to produce the right match for an English sentence of 20 words.* More important, it is obvious that such a program would contribute nothing to answering the question posed above. It fails to reflect an essential characteristic of way in which human beings perceive syntactic structure. Just as with producing utterances so with understanding utterances what is demanded is not just a knowledge of the grammar of the language but also a capacity for forming plans.

If one is listening to someone talking who is interrupted one is usually left in no doubt as to whether at the point at which he broke off he had finished a sentence or not. Moreover, one can usually make a guess about how the sentence would have been completed. Certainly anyone can play the game of making lists of possible ways of completing an unfinished sentence or of choosing from a list those endings that are possible and those which are not. In making these decisions an indispensable criterion is whether or not a particular ending is syntactically possible, that is, whether or not adding it would produce a grammatical sentence. Given the first part of a sentence, even just the first word, on the basis of its syntactic structure one can state the structure of elements that could legitimately succeed it—because one knows the grammar of the language. It seems likely that it is in this ability to predict the structure of elements to come on the basis of the structure of those already heard that the answer to the question ‘how do we bring our knowledge of the grammar to bear in recognizing the structure of utterances?’ is to be sought. The process of perceiving syntactic structure can be viewed as a

* See Mathews (1962). Mathews suggests several strategies which would improve the performance of an analysis-by-synthesis device, such as getting the program to count the number of words in the sentence under analysis before starting to generate sentences in order to prevent the generation of sentences that are longer than it. None of these improvements makes analysis-by-synthesis seem any more plausible as a model for human behaviour.

process whereby each element of the sentence as it is read in is used to test certain predictions, those which prove correct forming the basis for further predictions about the next element. Essentially what this kind of procedure involves is planning. As Miller, Galanter & Pribram (1960, p. 38) put it 'Planning can be thought of as constructing a list of tests to perform'.

Assuming that this account (in as far as it goes) bears some resemblance to the processes of syntactic perception a program intended to simulate this activity must be able to use the grammar (which, of course, forms an essential component of such a program) to make predictions about the structure of sentences it is given to analyse, to construct at any point in the sentence it is analysing a plan of its possible continuations.* Or rather (and the point is an important one) sets of alternative plans. It is very likely that the first word in a sentence will turn out to satisfy the requirements for being the initial element not just in one but in several sentence structures. In which case after the first word has been processed not one but several predictions concerning the structure of the next word are possible. The next word will almost certainly fail to satisfy some of these predictions, but it is equally likely in its turn to satisfy more than one and, therefore, be the origin of more than one prediction about the immediate continuation of the sentence structure. And so on up to the end of the sentence. In the case of a sentence that is syntactically unambiguous every sequence of predictions, except one, will fail somewhere before the end of the sentence is reached. In an ambiguous sentence there will be two or more sequences extending to the end of the sentence.

In view of the efficiency with which we perceive the syntactic structure of utterances it is obvious that if the procedure just outlined is to be seriously considered as a model for this activity then we must be prepared to assume that it involves the development of several different sets of predictions simultaneously. The kind of process we have to postulate is not the formulation of one plan at a time each involving a separate pass through the sentence (a process which cannot be halted even when one has proved correct because until all have been tested there is no guarantee that the sentence is not ambiguous), but the progressive evolution in the course of one scan of the sentence of one or more plans through the simultaneous development and testing of many possible plans. By the same token the program must realize a strategy for developing simultaneously at any point in a sentence all the analyses possible up to that point. Of course, the whole purpose of this line of research lies in devising such a strategy (or strategies). Describing perception of syntactic structure in terms of a task may provide some insight into the process, but it really adds very little to our understanding unless we can also give some indication as to what accomplishing such a task involves. There is no way of judging from this kind of description how complicated such a task is or how complex a mechanism would have to be to be able to carry it out.

A program for analysing English sentences which operates in the way specified has been constructed. Sentences are read in a word at a time and the program

* The predictive approach is used in several automatic analysers not intended to simulate the perception of syntactic structure, notably the Harvard Analyser (Sherry 1961). It was originally proposed by Ida Rhodes (1959).

constructs progressively a labelled tree-like structure each branch corresponding to a successful prediction. When none of the predictions for the development of a particular branch are satisfied no new branches are added to it. When a word satisfies identical predictions from more than one branch the branches are conflated. If the sentence is unambiguous there will be only one continuous path through the completed structure representing an analysis of the sentence. If it is ambiguous, there will be two or more. The program has been described in detail elsewhere (Bratley, Dewar & Thorne 1967, 1968. Examples of output from this program are given in Appendix I). But one feature of it is particularly interesting. It is basically very simple. This is reflected in the fact that it consists of about a thousand lines of a high level programming language; that is roughly ten thousand instructions. It would, of course, be absurd to claim that the brain actually employs the same strategy as the program but the fact that it is possible to devise a relatively simple strategy for accomplishing this task gives some grounds for thinking that this might be the task that is accomplished in the perception of syntactic structures.

Two other features of the program are worth commenting upon. The first is the feature which most obviously distinguishes it from other automatic syntactic analysers. This is its ability to analyse sentences without having to have incorporated into it a dictionary giving part-of-speech information about every word it encounters. Clearly a program which is intended to simulate the way humans perceive syntactic structure must possess this ability in view of the fact that we are able to recognize the syntactic properties of words we have never heard before. It is impossible to construct a program which does not derive some of its information by a process of dictionary look-up but this program works with a dictionary limited to less than 2000 entries, the most important of these being the grammatical formatives, that is articles, prepositions, auxiliary verbs, etc., all of which are word classes with a finite membership, together with grammatical affixes, like *-ing* and *-ed*. On the basis of the information about which elements in the sentence under analysis belong to these classes and information about sentence structure derived from the grammar the program is able to work out what part-of-speech all the other words in the sentence are. Take, for example, the sentence *The boy smokes*. The only information the program derives by dictionary look-up is that *the* is a definite article and that *-s* is a possible verb inflexion. But from the grammar it derives the information that English sentences can consist of a noun phrase followed by a verb phrase and that noun phrases can consist of a definite article followed by a noun and that a verb phrase can consist of a single intransitive verb. On the basis of this information it is able to work out that in this sentence *boy* is in fact a noun and *smoke* an intransitive verb, all the other analyses it tries (taking *boy* as an adjective and *smokes* as a noun, for example) having to be discarded on the grounds that they are not analyses of well-formed English sentences. The interesting point is that the operation of this program which does not derive information about each word by dictionary look-up is actually simpler than one that does. This in turn suggests that human perception of syntactic structure might also be a predominantly 'top-to-bottom process'; that is, that we recognize

the structure of individual elements in relation to the overall structure rather than deriving the overall structure from that of the individual elements in it.

The other interesting feature of the program is that it undertakes surface and deep structure analyses simultaneously. That is to say that at the same time as identifying surface structure elements it indicates which of them occupy a different position in the surface structure from that which they occupy in the deep structure and in addition identifies deep structure elements that have been deleted, specifying their position with respect to the other deep structure elements. The requirements that the program should make only one pass through the sentence and that it should work with a limited dictionary are really *a priori* conditions for a program which is intended to simulate the perception of syntactic structure. Unless it meets these conditions hardly any claims can be made for it as a model of human behaviour. But the reasons for requiring that it should undertake deep and surface structure analysis simultaneously are not obvious. The alternative hypothesis, that is, that perception of syntactic structure involves two processes, the recognition of surface structure and the derivation of deep structure information from it, appears quite plausible. Two stage analysers comprising a surface structure analyser the output of which is input to a deep structure analyser have been constructed (see Zwicky *et al.* 1965; Thorne *et al.* 1966). They run into the difficulty, however, that in the case of some kinds of sentences (notably those involving deletion of deep structure elements; sentences with embedded relative clauses, or with conjunctions like *and* and *but*, for example) the surface structure analyser is likely to produce a very large number of wrong analyses, any *ad hoc* device for reducing their number having the unfortunate consequence of usually also removing the correct analysis. As a result the function of the second-stage analyser has to be extended to cover not only the production of a deep structure analysis corresponding to the correct surface structure analysis but the discarding of incorrect surface structure analyses. The production at one stage of a large number of analyses which have to be discarded at another is obviously inefficient. And from this point of view a program which can produce both surface and deep structure analyses simultaneously is at a considerable advantage. It also suggests that in the human perception of syntactic structure both kinds of information are obtained together.

REFERENCES (Thorne)

- Bratley, P., Dewar, H. & Thorne, J. P. 1967 Recognition of syntactic structure by computer, *Nature, Lond.* **216**, 969-973.
- Chomsky, N. 1957 *Syntactic structures*, The Hague: Mouton.
- Lashley, K. 1951 The problem of serial order in behavior. In *Cerebral mechanisms in behavior* (L. A. Jeffress ed.), pp. 112-136. New York: John Wiley and Sons.
- Lenneberg, E. 1967 *Biological foundations of language*. New York: John Wiley and Sons.
- Mathews, G. 1962 Analysis by synthesis of sentences in natural languages. In *Proceedings of International Congress on Machine Translation and Applied Language Analysis*. London: H.M.S.O.
- Miller, G., Galanter, E. & Pribram, K. 1960 *Plans and the structure of behavior*. New York: Holt, Rinehart and Winston.
- Rhodes, I. 1959 A new approach to the mechanical translation of Russian. *National Bureau of Standards Report No. 6295*.

Sherry, M. 1961 Comprehensive report on predictive syntactic analysis. In *Mathematical linguistics and automatic translation*, Report No. N57-7, Harvard Computation Laboratory.

Thorne, D., Dewar, H., Whitfield, H. & Bratley, P. 1966 A model for the perception of syntactic structure. In *Colloque International Sur L'Informatique*, Toulouse.

Thorne, J., Bratley, P. & Dewar, H. 1968 The syntactic analysis of English by machine. In *Machine intelligence 3* (D. Michie ed.). Edinburgh University Press.

Zwicky, A., Friedman, J., Hall, B. & Walker, D. 1965 The MITRE syntactic analysis procedure for transformational grammars *AFIPS 27 Fall J.A.A.* Washington D.C. Spartan Books.

APPENDIX I. EXAMPLES OF OUTPUT

The computing time for each sentence and the total number of nodes in the analysis structure are recorded on the line following the sentence. The timing figure (in seconds to three decimal places) includes the time spent in consulting the CCD; it does not include the time spent in outputting the analyses. The arrangement of the printout is hierarchical with the elements entering into the analysis of each category being printed out under the category name.

The abbreviations used are listed below:

SRMs (syntactic relation markers)		category names	
[SE	sentence]	STAT	statement
TE	terminator	QUES	question
SU	subject	IMP	imperative
AV	active verb	INDS	indirect statement
OB	object	INDQ	indirect question
MO	modifier	INFC	infinitive clause
AU	auxiliary	NOMC	nominal clause
DE	determiner	PARC	participial clause
HE	head (of noun phrase)	SUBC	subordinate clause
AT	attribute	GER	gerund
IN	indirect object	REL	relative
LI	link (preposition or conjunction)	PREC	prepositional clause
PO	prepositional object	CNP	complex noun phrase
PA	particle		
CO	complement		
OO	INVERTED ELEMENT		

* Marks the position from which an INVERTED ELEMENT has been displaced. (The inverted element in question always being the nearest to the left, taking the most deeply nested asterisked position first.) In sentences with no inverted elements it marks the position of a deleted PROFORM (= Somebody).

- 1 Whitehead and Russell devised this method
(time taken: 0.960 s, nodes used: 15)
- 1 SE: STAT
2 (SU: Whitehead and SU: Russell) AV: devised OB: CNP
3 DE: this HE: method
TE: 1 2 3
- 2 why have they avoided the issue?
(time taken: 1.026 s, nodes used: 23)
- 1 SE: QUES
2 OO: why OO: have SU: they AU: * AV: avoided OB: CNP
3 DE: the HE: issue
TE: ? 1 2 3
MO: *
- 3 they have avoided it because they hate confrontations
(time taken: 1.192 s, nodes used: 22)
- 1 SE: STAT
2 SU: they AU: have AV: avoided OB: it MO: SUBC
3 LI: because SU: they AV: hate OB: CNP
TE: 1 2 3 4
HE: confrontations
- 4 observing butterflies and moths amuses them
(time taken: 1.754 s, nodes used: 61)
- 1 SE: STAT
2 SU: GER
3 SU: * AV: observing OB: CNP
4 (HE: butterflies and HE: moths)
TE: 1 2 3 4
AV: amuses OB: them
- 5 Christopher denies that he likes conferences
(time taken: 1.102 s, nodes used: 16)
- 1 SE: STAT
2 SU: Christopher AV: denies OB: INDS
3 LI: that SU: he AV: likes OB: CNP
4 HE: conferences
TE: 1 2 3 4
- 6 when are they taking their holidays?
(time taken: 1.229 s, nodes used: 31)

1 SE: QUES
2 OO: when OO: are SU: they AU:* AV: taking OB: CNP
3 DE: their HE: holidays

TE: ? 1
2
3

7 offer the examiners a fat fee
(time taken: 1.53 s, nodes used: 31)

1 SE: IMP
2 SU: * AV: offer IN: CNP OB: CNP
DE: the examiners DE: a AT: fat HE: fee

TE: 1
2
3

8 can we have a break when we have finished this experiment?
(time taken: 1.445 s, nodes used: 37)

1 SE: QUES
2 OO: can SU: we AU:* AV: have OB: CNP MO: SUBC
3 DE: a HE: break LI: when SU: we AU: have AV: finished OB: CNP
4 DE: this HE: experiment

TE: ? 1
2
3
4

9 stop immediately
(time taken: 0.532 s, nodes used: 17)

1 SE: IMP
2 SU: * AV: stop MO: immediately

TE: 1
2

10 must visiting professors be a bore?
(time taken: 1.275 s, nodes used: 35)

1 SE: QUES
2 OO: must SU: CNP AU:* AV: be CO: CNP
3 AT: visiting HE: professors DE: a HE: bore

TE: ? 1
2
3

1 SE: QUES
2 OO: must SU: GER AU:* AV: be CO: CNP
3 SU: * AV: visiting OB: CNP DE: a HE: bore
4 HE: professors

TE: ? 1
2
3
4