

APPENDIX II

SOME LINGUISTIC OBSTACLES TO MACHINE TRANSLATION*)

Y. BAR-HILLEL

Hebrew University, Jerusalem, Israel

For certain pairs of languages it has experimentally been shown that word-for-word machine translation leads to an output which can often be transformed by an expert post-editor into a quite satisfactory translation of the source text. However, if one is interested in reducing the burden of the post-editor, or if one has to do with pairs of languages for which word-for-word translation is not by itself a satisfactory basis for post-editing, it is natural to think of mechanizing the determination of the syntactic structure of the source sentences. It is a priori clear, and has again been experimentally verified, that knowledge of the syntactic structure of the sentences to be translated does considerably simplify the task of the post-editor. It is obvious, for instance, that this knowledge tends to reduce, and in the limit to eliminate, those syntactical ambiguities which are created by the word-for-word translation and which are non-existent for the human translator who treats the sentences as wholes. The task of the post-editor would then consist solely in eliminating the semantical ambiguities and in polishing up the style of the machine output. Whether these steps, too, can be taken over by machines of today or of the foreseeable future is still controversial; I myself believe that I have strong reasons for regarding it as hopeless, in general, but this is not the point I would like to discuss here.

A few years ago, I proposed what I called a quasi-arithmetical notation for syntactic description¹⁾ whose employment should allow, after some refinements, for a mechanical determination of the constituent structure of any given sentence. At that time, I actually demonstrated the effectiveness of the method for relatively simple sentences only but cherished the hope that it might also work for more complex sentences, perhaps for all kinds of sentences. I am now quite convinced that this hope will not come true. As a consequence, the road to machine translation can be shown to contain more obstacles than was realized a few years ago. I think that this should be of sufficient interest to warrant some more detailed exhibition, especially since this insight is due to an important new, not to say revolutionary, view of the structure of language, recently outlined by the American linguist and logician Noam Chomsky²⁾, and should, in its turn and in due time, be turned into a new method of machine translation, which would be more complex than the known ones but also more effective.

Since I cannot assume acquaintance with the paper in which I introduced the quasi-arithmetical syntactical notation mentioned above, let me present it here again very briefly, with some slight modifications³⁾; for a full presentation, the paper should be consulted.

*) A revised version of a talk given before the Second International Congress of Cybernetics, Namur, September 1958. It is to be printed in the Proceedings of this congress.

The basic assumption is that all words of a given language belong to one or more of the members of an infinite hierarchy of syntactic categories. Among these categories two are regarded as fundamental, viz. the categories of nominals and of sentences, denoted by n and s, respectively; the remainder are operator categories whose members, the operators, are considered as forming out of their arguments, always occurring to their immediate left or immediate right, more complex expressions. To illustrate: In the English sentence

John slept,
John is a nominal⁴⁾ and slept an intransitive verbal, i.e. an operator which out of a nominal to its left forms a sentence. We shall therefore denote the category of this operator by

$\underline{n/s}$
(read: n sub s). In the sentence
Little John slept,

John and slept would belong to the mentioned categories, whereas little would be adjectival, i.e. an operator that out of a nominal to its right forms again a nominal, hence be assigned to category

$\underline{n/n}$
(read: n super n). In

Little John slept soundly,
soundly would be an (intransitive verbal) adverbial, i.e. an operator that out of a left operator that out of a left nominal forms a sentence forms an operator that out of a left nominal forms a sentence, hence be assigned to category

$\underline{(n/s)\ (n/s)}$,
or rather, to use a self-explanatory additional notational convention, to
 $\underline{n/s\ \backslash\ n/s}$.

Most English words, perhaps all, would belong, of course, to more than one syntactical category. Soundly, for instance, would belong also to $\underline{n/s}/\underline{n/s}$, to $\underline{(n/s)/n}/\underline{(n/s)/n}$ (think of Belgium soundly defeated the Netherlands), etc.

Assuming, then, that a category "dictionary" listing for each English word all its categories stands at our disposal, the task of finding out whether a given word sequence is a sentence or, more generally, a well-formed (or connex) expression and, if so, what its constituent structure is, could now be solved according to the following utterly mechanical procedure: we would write under each word of the given word sequence the symbols for all the categories to which it belongs and then start "cancelling" in all possible ways, according to either of the two following rules:

$$\alpha^{\wedge}\alpha\backslash\beta\text{--}\rightarrow\beta \text{ and } \alpha/\beta^{\wedge}\beta\text{--}\rightarrow\alpha. \quad 5)$$

A series of such symbol sequences where each sequence results from its predecessor by one application of a cancellation rule is called a derivation. The last line of a derivation is its exponent. When the exponent consists of a single, simple or complex, symbol, the word sequence with this exponent, and with the constituent structure given by the derivation, is well-formed; if the exponent is, more specifically, s, the sequence is a sentence.

To illustrate, let us start with the last analyzed expression:

Little John slept soundly.

Let us assume (contrary to fact) that consultation of the category dictionary would have resulted in the following category symbol sequence:

(1) n/n n n/s $n/s\ \backslash\ n/s$.

It is easy to see that there are exactly three different ways of performing the first cancellation, starting off three different derivations, viz.:

(2) n n/s $n/s\ \backslash\ n/s$,

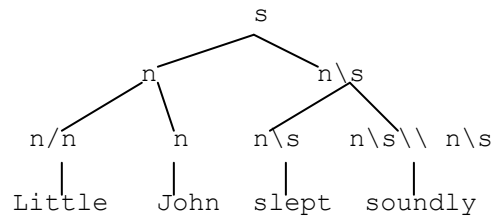
(2') n/n s n\s\\n\s,

(2'') n/n n n\s.

(2') leads into a blind alley. The other two lines, (2) and (2''), allow each for two continuations, of which one again leads into a blind alley, whereas the other allows for just one more derivation, with both exponents being s. Let me write down one of these derivations:

(1) n/n n n\s n\s\\n\s
 (2) n n\s n\s\\n\s
 (3) n n\s
 (4) s

The other derivation differs from the one just presented only in that the two cancellation steps in (2) and (3) occur in the opposite order. These two derivations are therefore equivalent, in an important sense; in fact, they correspond both to the same tree expansion:



Our second and final example will be:

Paul thought that John slept soundly.

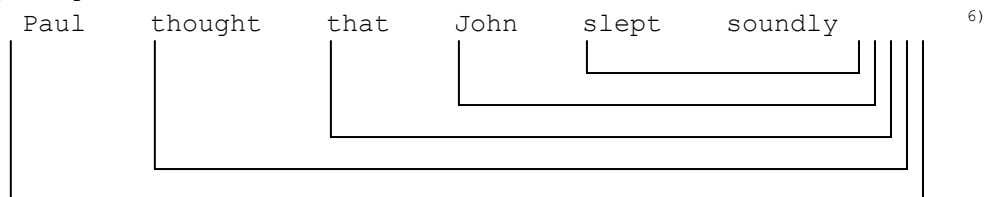
(I hope that the somewhat shaky English of this example will be forgiven; it simplifies making the point without falsifying it.) Copying only the first entry under each word in our fictitious category dictionary, we arrive at Paul thought that John slept soundly

n (n\s)/n n/s n n\s n\s\\n\s.

There are two non-equivalent derivations with a single exponent. I shall again write down only one of these derivations:

(11) n (n\s)/n n/s n n\s n\s\\n\s
 (12) n (n\s)/n n/s n n\s
 (13) n (n\s)/n n/s s
 (14) n (n\s)/n n
 (15) n n\s
 (16) s

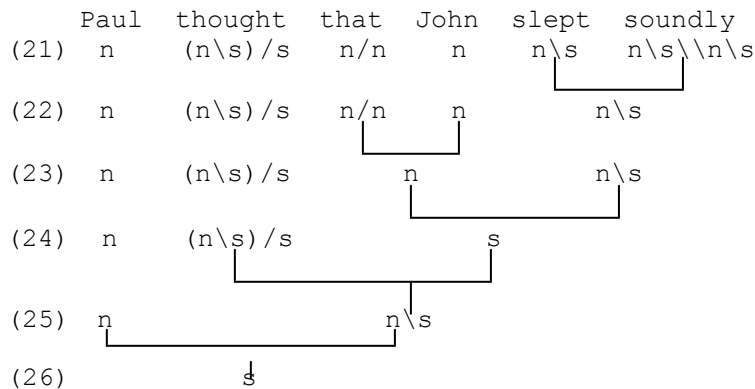
The constituent structure corresponding to this derivation can be pictured in the following parsing diagram:



As said before, the situation actually is more complicated. An adequate category dictionary would contain in general more than one entry per word. That e.g., is often a nominal, n, and even more often an adjectival, n/n, soundly could as well be an n\s//n\s or an ((n\s)/n)/((n\s)/n) (as mentioned above) and thought, finally belongs also to categories n, n\s, (n\s)/s (Paul thought John was asleep) and, qua participle, to still others. It can nevertheless readily be seen that our method is capable, at least in certain cases, to determine by purely mechanical operations the specific category to which a given word belongs in its given linguistic context. In our example, e.g., listing all the mentioned categories in column form yields the following scheme:

Paul	thought	that	John	slept	soundly
n	(n\s)/n	n/s	n	n\s	n\s\\n\s
	n	n			n\s//n\s
	n\s	n/n			((n\s)/n)/((n\s)/n)
	(n\s)/s				

It would be a tedious but wholly routine exercise to determine that out of the very *many* derivations corresponding to this word sequence – notice that there are 36 initial lines alone! – there exist only three essentially different ones with a single exponent, namely, in addition to the two above-mentioned derivations just



I still remember my surprise a few years ago when I discovered that this constituent structure is doubtless grammatical, however wildly implausible the conditions under which it would be uttered.

So far, so good, then. But, unfortunately, the actual situation is still much more complicated. It will be necessary to distinguish various kinds of nominals, for instance, singular and plural, animate and inanimate. Some additional notational means will have to be found from which it will follow that John slept, The boy slept, Boys slept, The boys slept are well-formed but that Boy slept, The John slept are not, that The little boy slept is connex but not Little the boy slept. These, and thousands other additional refinements,⁷⁾ can probably still be introduced without blowing the whole method up. But there are many features which make it highly doubtful whether English grammar – or that of any other natural language, for that matter – can at all be forced into the straitjacket of the immediate-constituent model and remain workable and revealing. Since the arguments against such a possibility have already been presented elsewhere with great force,⁸⁾ I shall not repeat them here in all their generality but restrict myself to the point of view of machine translation.

It takes but little to realize that the four categories mentioned above for thought are far from being exhaustive. In addition to its being a participle, which has already been mentioned, there are such phrases as thought processes, thought thirsty (not common but definitely grammatical), thought provoking, etc. In order to take care of the first two contexts, e.g., we would have to assign thought also to the categories n/n and n/n//n/n. ("In these contexts, thought occurs in the function of an adjective or an adverb, respectively" would have been one traditional way of putting the issue.) The third context would have raised the notoriously difficult problem of the status of the participle present, in addition. The task of preparing a category list that would work for all these and innumerable many other contexts is certainly much harder than the first successful analyses caused us to believe. Would not the required list become so long that the mechanical determination of the constituent structure of say, a 30-word sentence with three or four categories per word, on the average, might well require trillions of machine operations, hence be totally impractical for machines of today as well as of tomorrow?

And what with a sentence such as Playing cards is fun? On first sight, it seems that one has to arrive at the category n for the phrase playing cards. However, it is intuitively clear that this should not be derived from cards being an n and playing being an n/n (and not only intuitively so: notice that the next word is is and not are; playing cards is in our context a singular nominal). There are, of course, many other ways of enforcing an assignment of n to playing cards, but none of these, to my knowledge, is such that it would not introduce unwarranted and counter-intuitive syntactical resolutions of other sentences. "Hocus-pocus" linguistics – as certain linguistic methods were called whose only purpose was to save certain phenomena, without regard to any intuitive (or psychological) realities – would in our case definitely refute itself by saving also phenomena that are non-existing.

And what about a sentence like He gave it up? What category would up have to be assigned to in order that this sequence should turn out to be connex? We all feel that gave and up somehow belong together and that this is so without regard to the length of the expression that separates them. This, however, is definitely beyond the reach of the immediate constituent model in which the immediate constituents of a connex expression are always contiguous or, to put it in a different terminology, where modified expression and modifying expression have to stand one directly after, or before, the other.⁹⁾

If now the immediate constituent model is not good enough to serve as a general model for the whole grammar of a given language, the method of mechanical structure determination outlined above can no longer be assumed to be of general validity, either. As a matter of fact, I had noticed already six years ago that complex sentences could not be analyzed well by this method as it stood then but I had rather hoped that this was due only to lack of refinement. I have now come to realize that its failure in the more complex cases has a much deeper cause: the linguistic model on which this method was based is just not good enough.

Since the thinking of the linguists working on machine translation was mostly governed by the immediate constituent model, unless they were working with a still more primitive model, a communication-theoretical finite-state Markov process model (or, of course, working without any model), it should not be really surprising that so little progress was made during the last years in the mechanization of the syntactic analysis of languages. I, for one, am satisfied with

this explanation of the present stagnation in this respect.

Having identified the nature of this obstacle to machine translation, we must, of course, ask ourselves what consequences are to be drawn from this identification for future work on MT. The answer is rather simple as such, though its exact implications are far from being clear. A better model for the working of grammar, i.e. for the synthesis of well-formed expressions, especially sentences, out of the linguistic elements – which, for MT purposes, are the letters and other elemental graphic signs such as numerals, punctuation marks, etc. – has first to be set up and then turned around to allow for the mechanical analysis of the resulting large units. Chomsky and Harris¹⁰⁾ have shown us outlines of a third, more powerful model for linguistic synthesis, the so-called transformational model. It does not discard the immediate-constituent model but rather supplements it. The former model remains intact for a certain kind of simple sentences, the so-called kernel sentences (or rather for their underlying terminal strings)¹¹⁾ – and our method of mechanical structure determination remains therefore valid for these sentences –, but has to be supplemented by additional procedures, the so-called transformations, in order to account for the synthesis of all sentences.

The answer to the question, "What is the constituent structure of the sentence, He gave it up?", is now: this sentence has no proper constituent structure; it is the result of a certain transformation on the terminal string, He gave up it, which has indeed a rather simple and perspicuous constituent structure. The answer to the question, "What is the subject of the sentence, Playing cards is fun?", is now – whatever grammarians had to say on this topic until now (and what they had to say was highly unsatisfactory and often contradictory) – that this sentence, not being a terminal string, has no proper subject but is rather the result of certain transformations on certain terminal strings.¹²⁾ (The actual situation is too complicated to be treated in the space at my disposal.)

Each sentence, according to our last model, is then the result of a series of one or more transformations performed one after the other on one or more terminal strings – unless, of course, it is a terminal string itself. A complete analysis, mechanical or otherwise, of a given sentence has to tell us what its basic terminal strings are, together with their constituent structure, and what transformations, and in what order, were performed upon them. Assuming that a complete transformational grammar, for some given language, has been prepared, the preparation of a corresponding analytical (or operational) grammar is a formidable, though perhaps not necessarily an impossible task. So far, of course, no transformational grammar exists for any language, to any serious degree of completeness.

The recognition that immediate constituent grammars have to be supplemented by transformational grammars makes the task of mechanizing translation look much harder, but the resulting picture is not at all uniformly black. On the contrary, there are reasons to suppose that the additional insight we get on the basis of this model will not only be of decisive importance for theoretical linguistics but may well turn out to facilitate the mechanization of translation from new angles.

First: you remember that one of our previously analyzed sentences was
Paul thought that John slept soundly
and the troubles we foresaw in its mechanical analysis. It is obvious, however,

that in a transformational grammar this sentence will not be a terminal string but rather (1) either the result of a certain kind of "fusing" transformation on the sequence of the two terminal strings

Paul thought this: John slept soundly
or (2) the result of two transformations, the first being the same "fusing" transformation performed, however, on somewhat different terminal strings

Paul thought this: That John slept soundly,
the result of which would be

Paul thought that that John slept soundly,
the second transformation being a certain kind of "elliptic" transformation causing, in our case, the omission of the first that.

No longer, then, will $n \setminus s / s$ be regarded as one of the categories of thought, nor n/n and $n/n//n/n$, as thought processes and thought thirsty will now be treated as resulting from processes of thought and thirsty for thought by certain transformations.

The first gain consists, then, in that the number of categories per word will almost always be less, sometimes much less, than under the former method. For some words this number will now be zero, indicating that no sentence containing such words is a terminal string. To give an example: sleeping will not be assigned to any category, any sentence containing this word being considered as the result of a transformation. (Interesting, however, will be assigned to the category n/n .)¹⁴⁾ That there might be words which do not belong to any syntactic category will strike many linguists as rather queer, but I am convinced that on second sight they will realize the enormous advantages of such an attitude; innumerable pseudo-problems have in the past been created by the search for the syntactic category (the traditional term is, of course, "part of speech") of certain words or phrases which – under the new model – just don't belong to any category. This is – if I may be allowed one generalization – just one more instance of the very common class of situations where the attempt of applying a model which is very useful within certain limits leads, when pushed beyond these limits, to pseudo-problems and their pseudo-solutions.

The second gain is somewhat more speculative: it seems likely, but has so far not been seriously tested, that languages will be much more similar with regard to their terminal string structure than with regard to the structure of the totality of their sentences. Word-for-word translation of terminal strings, with some occasional permuting, seems to yield satisfactory results for many pairs of languages, including those for which this kind of translation does not work at all with regard to more complex sentences.

The most remarkable gain, however, would be achieved when it turned out that between the sets of transformation of two languages there existed a close semantic relationship. Should it happen that for certain two languages, L_1 and L_2 , there exist two transformations, say t_1 and t_2 , such that for any semantically equivalent terminal strings of these languages, k_1 and k_2 , $t_1(k_1)$ is semantically equivalent to $t_2(k_2)$, this would allow for a relatively simple mechanization of the translation, provided, of course, that the syntactic analysis of L_1 has been mechanized, whereas a word-for-word translation of $t_1(k_1)$ into L_2 might be highly unsatisfactory.

Of course, there is but little hope that the sets of transformations of two languages which do not stand in any close genetical relationship will do us the favor of exhibiting isomorphism or near-isomorphism with regard to seman-

tic equivalence. So far, there exists to my knowledge no general theory of machine translation which would ensure that, if only the precepts of this theory are followed, the target language counterpart (or counterparts) of any sentence of a given source language will be no more and no less syntactically ambiguous than the original sentence itself. Current statements to the contrary seem to me palpably false, and any hope for an imminent establishment of such a theory – unsubstantiated. Great progress has been made in this respect with regard to certain ordered pairs of languages, such as French-English, German-English, Russian-English, English-Russian, German-Russian and French-Russian, partly prior to the appearance of the transformational model and without any conscious use of its methods, and more progress may be expected in the future through a conscious use of these methods. As one necessary condition for further success I regard the recognition on behalf of the workers on MT that the model with which they were working, consciously or unconsciously, during the first decade of their endeavors was too crude and has to be replaced by a much more complex but also much better fitting model of linguistic structure.

NOTES

1) "A quasi-arithmetical notation for syntactic description", Language 29:47-58 (1953).

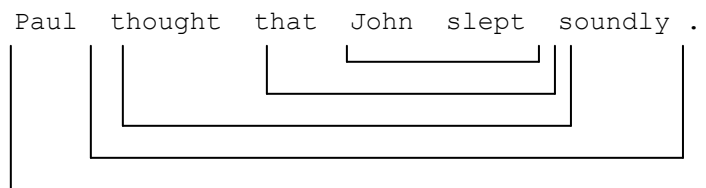
2) Noam Chomsky, Syntactic Structures, 's-Gravenhage, 1957.

3) These modifications refer both to terminology and to notation. The latter are influenced by J. Lambek, "The mathematics of sentence structure", American Mathematical Monthly 65:154 (1958).

4) Nominals, verbals, adjectivals, etc., in my usage, are syntactical categories. They should not be confused with nouns, verbs, adjectives, etc., which are morphological (paradigmatic) categories in my usage. The connection between these two classifications, as the choice of terms is intended to indicate, is that nouns usually, though by no means always, belong to the syntactical category of nominals, etc., and that most expressions belonging to the syntactical category of nominals of course only if they are single words, are nouns.

5) The reading of these rules should be self-explanatory. The first, for instance, reads: Replace the sequence of two category symbols, the first of which is any category symbol whatsoever and the second of which consists of the first symbol followed by a left diagonal stroke followed by any category symbol whatsoever, by this last category symbol.

6) The other single exponent derivation yields a constituent structure whose diagram is



If this structure is regarded as unacceptable, the notation will have to be considerably refined in order to exclude this derivation.

With regard to the problems arising in connection with the fact that the notation $(n \setminus s) / n$ creates an arbitrary-looking referential reading of what should "naturally" have been written $n \setminus s / n$, see op.cit. in note 1, p. 55 and op.cit. in note 3. Both treatments do not yet cover all aspects of the problem.

7) Such as the one discussed in the preceding note.

8) In op.cit. in note 2, as well as in other recent publications by the same author.

9) Discontinuous constituents were occasionally discussed in theoretical linguistics, but not before Chomsky was it realized what a difference this makes as against continuous and contiguous constituents.

10) Zellig S. Harris, "Co-occurrence and transformation in linguistic structure", Language 33:283-340 (1957).

11) Cf. op.cit. in note 2, p.45.

12) With appropriate safeguards, but only with such safeguards, one might also answer the first question by saying that the sentence, He gave it up, has He and

gave it up as its immediate constituents, and that its second component has the discontinuous expression gave...up and it as its immediate constituents. Likewise, the answer to the second question could also be formulated by saying that Playing cards is its quasi-subject, but this requires, of course, a prior definition of 'quasi-subject'.

13) This is only a first approximation. Actually, a satisfactory description will have to be much more complex.

14) Why? Hint: we have very interesting but not very sleeping.