18

SUGGESTIONS FOR MECHANICAL TRANSLATION*

SILVIO CECCATO and ENRICO MARETTI

To philosophize is hard, but not to philosophize is even harder.

E.Rogge.

THE MAIN OBSTACLE

IN OUR opinion, linguists and engineers who have so far turned their attention to mechanical translation have been delayed by the lack of a workable description of the operations which men perform on translating. That description still seems elusive. We believe the philosophical tradition is against it; and lay people as well as scientists and linguists conform to that tradition whenever interpreting and describing human activities relevant to thought, language and knowledge.

For this reason we believe that the research going on in the Italian Operational School will be of some use in this connexion. This research is carried on mainly with the view to constructing an apparatus which performs some of our mental operations and gives them verbal expression. This programme requires that the philosophical tradition be reconsidered and that we adopt some new views and results.

TRANSLATING

Translating consists in replacing one language with another while maintaining the same nominata. Illustrating the scheme:

Source-language ≠ Target-language

Nominata   =   Nominata

In passing from the source-language to the things which are named, and from these to the target-language, the translator follows the correlations which in every language are established between the 'sound-scratches' which are used as names and the things of any type which are used as nominata.

In doing so, nevertheless, the translator meets with some difficulties. This will happen if we do not have an ideal situation in which we may translate name for name, and nominatum for nominatum, or in which the builder of the translating machine may reduce his construction to a set of connexions directly substituting sound-scratch for sound-scratch.

First difficulty

This difficulty arises from a possible lack or nominata in the target-language in comparison with the nominata in the source-language. It may become the most serious difficulty confronting translation.

The translator must define the thing which is named in the source-language in terms which are known and named by the community of the target-language.

Example: Suppose 'hat' has to be translated for a community which never put on and saw any headgear. The translator will begin, for instance, analysing the hat into a certain shape, as a roundish one, of a certain material, considering its place on the head and its functions in protecting from sun and rain, and so on. If that community knows and names all the correlations and the correlata occurring in the definition it can be translated intelligibly, leading to the construction of something like the nominatum in the source-language.

This provides little difficulty, because it can be tackled by bilingual dictionaries, and so becomes one of the many connexions directly substituting one sound-scratch for another.

The compilers of dictionaries are not always able to supply the needed definitions. In order to work as a criterion for construction, a definition must neither be too wide nor too marrow, nor be formulated in negative, meta-phorical, contradictory, or synonymous terms. For instance, 'that which has no-parts' will never lead to constructing a 'point'. Our actual dictionaries as well as our actual philosophic and scientific treatises show, instead, many definitions of that type. We shall see that this is one consequence of the philosophical tradition.

Second difficulty

This difficulty arises when two or more things which, in the target-language, have each a name of its own, have only one name for all of them in the source-language. Before giving rise to a difficulty in translating, this multiplicity of meanings produces the well known difficulty of interpreting a text.

Among the nominata we must single out that which is intended by the author. We act by taking into account the situation in which the name occurs, either the linguistic, or a wider situation. For instance, one meaning in place of another may lead to violating some grammatical rule, or to making a statement false, metaphorical, contradictory, or simply uncommon. Thus that meaning is rejected; likewise if it leads out of the subject-matter. The field of inquiry enlarges when we consider, for instance, the habit of thinking and speaking of the author and the place and time in which the name is said. All the interpreter's and translator's culture may be called into play.

At all events, we begin by representing all the nominata which the name suggests, and we keep them alive until we find a reason for eliminating this or that nominatum, or for selecting a certain nominatum. Only when one remains can the trial be settled.

However, at present we do not possess any machines for executing such considerations. Hence students of mechanical translation follow another

way; so far as possible they try to have these considerations expressed by arrangement and occurrence of the words of a context[2].

Example: The English 'mole' is, in Italian, 'neo', an excrescence of the skin; 'talpa', a small burrowing mammal; 'molo', a stone pier. The occurrence in the context of words such as 'ship' or 'sailors' determines 'molo'.

Third difficulty

This difficulty arises because languages show two main discrepancies from one another. The first discrepancy is a matter of separation and number. In one language the way in which we separate (isolate) the sound-scratches from one another, and the way we separate the names from each other and the nominata from each other, may not coincide; and so also the number of the single sound-scratches, on the one side, and the number of the single names and of the single nominata, on the other side, may not coincide: also, since this discrepancy may not be the same in the various languages, a one-to-one correspondence between the words of the languages may be lacking.

(a) What in one language is named by means of only one sound-scratch, is named in another language by means of two or more sound-scratches; and *vice versa.* (a1) These sound-scratches are separated by their respective intervals alone. (a2) These sound-scratches are separated by some other sound-scratches.

(b) What in one language is named by means of the whole of a sound-scratch is named in another language by means of a part of a sound-scratch, and *vice versa.* Thus we come to find an Italian 'per', an English 'in order to', a German 'um . . . zu' which name the same nominatum; or an English 'in order to sleep' and a Latin 'cub-itum'.

(c) What in one language is. named by means of the shape of the sound-scratches in another language is named by means of a relation between two or more sound-scratches. Languages make this relation recognisable in two ways. (1) By putting a sound-scratch before or after a sound-scratch which is individuated owing to its shape—example: the Latin 'canem' is in English '... dog', where the '...' takes the place of a verb. (2) By adding a differentiating process to the correlated sound-scratches, for instance, by raising or lowering the voice.

The second discrepancy is a matter of order of occurrence. In one language the order of occurrence of the words may show a certain independence from the order of presentation of the content: this is shown by the fact that in two languages the same content can be represented by words which do not follow the same order of occurrence. For instance, the Latin 'Canem (1) feles (2) timet (3)', is in English 'The cat (2) fears (3) the dog (1)'.

All this prevents us from translating successfully sound-scratch for sound-scratch; and therefore we cannot translate by merely applying a bilingual dictionary or by merely using a set of directly substituting connexions. The translator ought to take a unit of translation longer than a single sound-scratch; but even so, the number of sound-scratches in input and output in the dictionary and of substituting connexions in the machine increases beyond measure.

When the interest in mechanical translation arose, students soon realized

that the correspondence between the single words in the various languages had something queer about it. They were faced by a somewhat odd situation; men know how to overcome this difficulty because they know grammar, but what is grammar ?

We shall examine now why philosophical tradition hinders us in giving a useful answer to this question.

THE PHILOSOPHICAL TRADITION

The core of philosophical tradition appeared in explicit form with the Greek speculators Alkmeon, Empedocles, Leucippus and Democritus. It is represented by a confusion between the state of being a name and the material used as a name, which are not distinguished in the word. Thus, it became impossible to understand how it is that the activity of correlating makes anything be a name or a nominatum. On the contrary, the correla-tion between the names and the nominata was explained by introducing a third thing between the two, such as the effluvia, eidola or particles. This thing passes from the nominata to the speakers and conveys to them the presence and the characteristics of the nominata:

$$\textit{Speakers} \quad |... \leftarrow ...| \quad \textit{Nominata}$$

This mixture of activity and passivity, of entities ready-made and made to measure, took the name of 'knowledge'; the speakers became the knowing subjects and the nominata the things given and to be known. It would require hundreds of pages to collect the consequences, and the consequences of the consequences, of this cognitive situation. The whole of philosophy and that part of science which could not become a technique feed upon them. Here we pause only to look at one direct consequence.

In order to be the origin of the thing sent, every nominatum has to be located in space and time. Besides, every nominatum has to be isolated and differentiated before it can send something, as a message, informing us of its *characteristics.* Thus, every nominatum became a hybrid of a physical and a psychical thing: this is appreciated if we consider that a thing is physical when it is located in space, owing to its distinction from something else in another place, and that a thing is psychical when it is located in time owing to a distinction from something else in another moment. The two types of things can be distinguished only on condition that an irreducibly meta-phorical 'core' is assigned to the psychical things.

At the same time the activities were excluded from the nominata: this is easily realized if we consider that an activity cannot be located in space and in time except as a succession, and that the differences constitutive of an activity are internal to the activity itself. An activity contains both the terms of a difference, and this is not something external, as in the case of physical or psychical phenomena, when the terms of the difference are shared between two things, each one containing one term of it.

THE ITALIAN OPERATIONAL SCHOOL

the i.o.s. became aware of the cognitive situation after almost ten years of trials and errors (from 1938 to 1948). We were helped most by investigating

certain nominata which we attempted to individuate and to analyse, for instance, 'part' and 'whole'. We observed that anything whatever can be made to name 'part' or 'whole', exactly, without changing any of its possible physical and psychical characteristics. It suffices to correlate, in a certain way, one thing with something else. Take, for instance, a 'cup'. Without changing any of the characteristics through the use of which we name something 'cup', the cup is made a 'part' when it is put in relation with a coffee-service, and a 'whole' when it is put in relation with a handle *etc.* Thus, we could draw the conclusion that: (a) not all the words are names for physical or psychical things; (b) some words are reserved to name activities; (c) those words sometimes take the form of verbs.

Our conviction was strengthened when we succeeded in individuating and analysing as activities such nominata as 'cause', 'effect', 'number', 'surface', 'point', 'line', 'law', 'phenomenon', 'and' 'or', 'not', 'time', 'space', and so on. The individuation, and their analysis as activities, of 'name', or 'symbol' and of 'nominatum', or 'symbolisatum' allowed us to master the cognitive situation.

[For the first steps of the I.O.S. away from the philosophical tradition, see reference 3.]

BEARING UPON THE DICTIONARY

Attempts to individuate and to define all the nominata as physical or psychical things lead to consequences which might be foreseen.

(a) If the activities are named by a verb, something physical or psychical is added to them as their unfailing subject. The infinitive mood is considered 'unreal', a 'grammatical fiction'. Anyway, they can be positively defined.

(b) If the activities are named by words which are not verbs, there are several cases:

(bl) If the word names both an activity and something which is correlated by that activity, the philosopher, instead of isolating the activity and defining it according to its characteristics, tries to find these characteristics in some possible physical or psychical correlated thing. Then, the individuation is distorted and the definition seems to be pertinent but always too narrow.

(b2) If the word names an activity which is performed on constructing the physical things (as the activity named 'space' is) or the psychical ones (as the activity named 'time' is), the philosopher starts from the physical and psychical things and tries to find the characteristics of the activity (i) by denying something physical or psychical or (ii) by opposing to one another something physical or psychical. In the former case the definition turns out in negative terms, in the latter one in contradictory terms.

(b3) If the word names an activity which is constitutive of language itself (the activities named 'thing' and 'to be', in one of the meanings of those words, for instance) the philosopher tries to find the characteristics of the activity by depriving the physical and psychical thing of all its characteristics. Then, insofar as he does not look for anything else, he gets what he wants by naming, or speaking ('thing' and 'to be' in the case of 'thing' and 'to be'). The definition turns out to be in synonymous terms.

(b4) If the word names, together, a physical or psychical thing and some inexecutable activity (*e.g.* 'sense', when that word names not what is required

175

by anatomy and physiology, but when it is shifted from its original nominatum and used by philosophers, psychologists, or physicists, as indicating a cognitive act), the philosopher tries to individuate and to define the nominatum by introducing some other physical or psychical thing which he loads with the same inexecutable activity. The definition is in irreducibly metaphorical terms.

(b5) If the word names a correlating activity alone *(e.g.* the activities named 'and', 'or', ',' , 'but', ':' ,) the philosopher, not being able to introduce anything physical or psychical, loses the nominatum. He remains, then, with something which he continues to consider linguistically, but which is contradictorily deprived of the nominatum. He talks about 'signs', a name which means both sound-scratches and symbols, and so covers the equivocal situation. The definitions he proposes are in terms which appeal to other 'signs' which are equally deprived of the nominatum e.g. 'syntactical definitions', or 'axiomatic definitions'. Therefore, we must learn from somebody else how to employ these words.

Here is, then, the reason why in actual dictionaries and in philosophic and scientific treatises many definitions are not usable as criteria for recognition and construction.

This understanding of the cognitive situation removes any limitation to the admission and individuation of nominata. We have now what can be called and recognized as activities. Thus, they can be defined in telling and positive terms.

This is the contribution which the I.O.S. offers towards solving the first difficulty of translating.

BEARING UPON GRAMMAR

Had languages reserved one isolated sound-scratch for every correlation they establish between the single names and the single nominata, then to know the established correlations would suffice in order to master a language, just as the use of the bilingual dictionary would suffice in order to translate. No grammar would be required. But we saw that languages may present two discrepancies in this connexion. The reasons for these discrepancies are easy to find.

Among our nominata. some occur with a frequency which is enormously greater than that of others. During a whole day we may have little occasion to name, for instance, a cat. a table, a pillar, a walk, or a study. During that day we certainly have often to name a singular or plural in connexion with the cat, table and pillar, as well as to name a present, past or future in connexion with walking and studying: and so in connexion with a lot of things which are often interchanged.

Thus, our languages have found a way for saving somewhat on the material used in naming, thus being quicker in following, expressing and indicating what is to be named. This way consists in depriving a nominatum, which occurs very often, of an isolated sound-scratch as its name. This happens when it figures, at the same time as another nominatum, giving rise to a sort of polyphony. In a rough musical writing *(Figure 1)* two ways are followed. (a) The frequent nominatum receives as its name a very short sound-scratch that is blended with the sound-scratch representing the other

176

nominatum. For instance, 'cat', 'singular' is simply 'cat'; 'cat', 'plural' is 'cats' (too expeditious a way, indeed, because English, French, Italian, German, for instance, are now in difficulty when naming 'cat' without involving its singularity or plurality: 'cat-ness', perhaps).

(b) The frequent nominatum loses completely the sound-scratch which is used as its name, and is named by means of a relation put between two other sound-scratches which represent two other nominata. For instance, 'man' and 'subject' is 'man . . .', where ' . . . ' takes the place of a verb.

One single sound-scratch may become the bearer also of several names. The sound-scratches bearing three, four, or five names are very common in our languages. [The Chinese language is perhaps the least inclined to follow this method.]
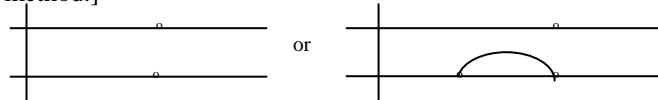

Figure 1.

The second discrepancy arises because, among the frequent nominata, some are activities used in order to correlate, and because these activities are named in four ways at least which differ from the one-to-one correspondence. This variety is allowed by the characteristics of the correlative situation. In every correlation three items occur in a fixed order; and they must attain this fixed order whatever be the order in which they enter into the correlation. The two moments of the correlating activity occur each one together with one of the two things which are correlated. In our rough musical writing *(Figure 2)* we can be concerned with two chairs, for instance, as carpenters or as observers, and we may have them in every respect independent of each other. They are in a priori order of occurrence when we have them *e.g.* one on-the-left-of the other, or one taller-than, or lower-than, the other. If then one, two or all of the items or the correlation as a whole are in turn correlated, the number of the items which occur in a definite order increases.
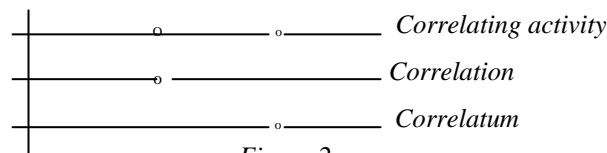

Figure 2.

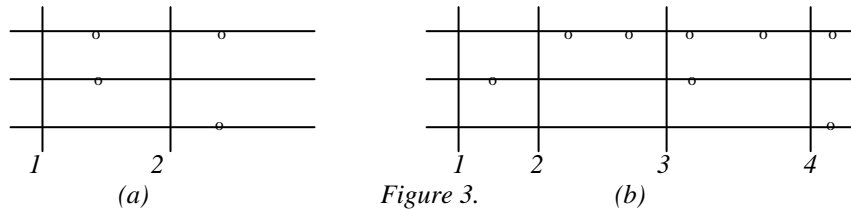We are often thinking in structures of 30, 40, 50 items.

Our language makes use of that fixed order of occurrence, breaking the one-to-one correspondence, either for economy or in order to maintain as far as possible the temporal correspondence between the occurrence of the nominata and the occurrence of the names. They face here a polyphonic situation; and language is only monodic.

The one-to-one correspondence between the sound-scratches, on the one hand, and the names and the nominate, on the other (which is the case when the correlating activity is named for instance 'and', 'or', 'of'), makes the twofold construction of the correlation in *Figure 3a* be the fourfold one of *3b.*

Another way of naming a correlation consists of giving two names to the correlating activity, by means of two isolated sound-scratches, putting one

177

before the first thing correlated and the other before the second thing correlated (which is the case, for instance, with 'if... then . . .', 'either . .. or...'). Shortening and blending the sound-scratches allows us to name the correlating activity by one or two short sound-scratches. Thus we can blend, respectively, the one with one of the sound-scratches which names the correlated things ('fear-ful child', 'child-ish fear'), and the other with one of those sound-scratches ('he (nominative case) sing-s') which names the correlating activity. The shortest way of naming allows us to make use of a particular relation between the sound-scratches which name one cor-related thing, in place of any sound-scratch, used to name the correlating

Figure 3.

(a)   (b)

activity ('were I you', subjunctive mood before the subject). These five ways of naming the correlations are very often mixed together.

Given such a variety of naming, only a fixed order in the occurrence of the nominata, which is independent of the linguistic form of any language, can assure the sameness of the content. That variety can only change the steps which the speaker or writer, and the reader or hearer, take on passing from the structures to their names and from the names to their structures. The reader or hearer operates with a continual revision and a continual anticipation. He revises in order to take up again what occurred out of correlation and now has to be correlated; he anticipates if the correlating activity is already named, but one or both of the things to be correlated are still lacking. They are such frequent and quick operations that they may pass unnoticed, but they are easily brought to light. 'I divide the money between . . .', and you expect certain correlata; 'I see a nice . . .', and you expect the correlated substantive which the adjective announces since it is already the other correlatum and bears the entire correlating activity. The speaker or writer operates also in at least two ways when he names cor-relations. The structure of the correlation must first be performed as an activity and borne in mind with its double articulation before the particular things to be correlated and named can be added to it. Then, the speaker or writer takes up what he has already thought and proceeds again with addi-tions. The correlating activity is consequently performed twice. If, then, a structure is enlarged by performing a new correlating activity on one of the previously correlated things, the back-and-forth movements and the repe-titions increase. The correlation which is coupled with a first instant of another correlation keeps this latter alive and in suspense until the former is terminated; the correlation which is coupled with a second instant of another correlation keeps the former alive in its determinate state until the latter is terminated.

All this is in action when a sentence is being uttered. The full stop between two sentences is the name for a particular correlating activity the instants of which are an ending and a beginning; so that this name empha-sizes a split in the correlation.

178

Insofar as the speaker or the writer starts naming the single items of a sequence in a fixed order owing to a structure which he has already decided upon, he can run over this sequence and select the item with which he prefers to begin. He can only profit by the freedom which language allows him, within the limits of possible equivocation. Thus he can originate, for instance, a gradation of importance among the items. The reader or the hearer will follow him by putting the items back in their fixed order. Only his movements back and forth and his repetitions increase or reduce.

Grammar is the ensemble of the passages which do not relate directly one single nominatum and one single sound-scratch; and the dictionary is the ensemble of the passages which do relate directly. At present this distinction, however much it is felt, cannot appear sharply.

We think that all this could have been made clear from the beginning if students had been able to confront the occurrence of the words with the occurrence of the single nominata; but to make this possible, all the nominata must be present. Philosophical tradition, however, by eliminating some nominata and by distorting the individuation of some others, had opened too large gaps in their ranks.

To fill those gaps is the contribution which the I.O.S. offers to the solution of the third difficulty of translating.

MAN OR MACHINE TRANSLATION

'Translator' implies both reader or hearer and speaker or writer; and to these is added the translator's knowledge of two languages, namely, of two dictionaries and of two grammars; the first in order to pass from the words of the source-language to their nominata, and the second in order to pass from these nominata to the words of the target-language.

Anyone who learns two languages from infancy, of course, does not want any dictionary or grammar: both languages are learnt in the presence of the nominata. Students of translation have to consider also how one language is learned by starting from another language; and then the philosophical tradition hinders them from realizing that such learning can and must happen in a similar way, namely by providing the same nominata, though this time through one language, and not directly with the nominatum. If many words have no nominata ....

Thus they had recourse to a stratagem. They worked on the words themselves, and devised a classification of them, dividing the words into classes: the familiar grammatical categories, conjugations, with moods, tenses *etc* declensions, with cases *etc.* These classes had to function as a union between the two languages, taking the place of the nominata they have in common. The words of a certain class in the target-language correspond to the words of a certain class in the source language.

Now, note that the teaching of language succeeds by making the classes known through giving many examples (at least two) of the words belonging to them: and the stratagem works.

It works because anyone who knows a language passes automatically from the names to the nominata; and then the classification together with the examples forces him to separate the names and the nominata which are

179

blended into one word and to bear them in mind while the words of the new language are exhibited. Every classification, if it is plurally exemplified, acts in the same manner, since it makes us hang on to one thing which persists unchanged, while other, and different, things are interchanged.

This is the first reason why the actual grammars did not suffice to give the builders of translating machines any workable suggestion: their machines do not speak any language; exemplification has no effect on them!

Let us assume that the builder of translating machines succeeds in separating the single nominata which are blended into a word, and in committing each one of them to a substituting connexion. Still the machine cannot fulfill its task without the unavoidable revision and anticipation which we saw is compelled by the polyphony of the correlations into the monody of language.

[In the full version of this paper is included an Appendix, giving illustrations of the 'defining of nominata' and 'recovering them as activities', together with demonstrations of the authors' system of notation—ED.]

REFERENCES

[1] CECCATO, S. and MARETTI, E. 'Construction of a Translating Machine', *Methodos,* 8 (1956)

— 'La grammatica insegnata alle macchine'. *Civiltà delle Macchine* 4 (1956) 1-2

[2] REIFLER, E. 'The Mechanical Determination of Meaning', in *Machine Translating of Languages,* ed. N. Locke and D. Booth, pp. 136-164. New York; Wiley and Sons, and Massachusetts Institute of Technology: London; Chapman and Hall, 1955

[3] CECCATO, S. *Language and the Table of Ceccatieff*, Paris; Hermann, 1951

— 'L'Ecole Opérationelle et la Rupture de la Tradition Cognitive', *Bull. Soc. Franc. Philos.,* Mar. 1952-May 1953

Articles published by representatives of the School from 1948 in *Methodos* quarterly review