

THE LIMITS OF INNOVATION IN MACHINE TRANSLATION

Margaret Masterman

Cambridge Language Research Unit

The three approaches to full machine translation which are now being implemented all have built into them their own limits to innovation. However, a far greater insight could be gained into where this limit really lies if a pattern recognition model could be constructed of how the human being actually translates. Then the question could arise as to how far the machine could simulate the model. The impressionistic suggestions for making such a model are given in an annex.

In the world outside, it is still the case that two extreme points of view exist with regard to machine translation. One viewpoint - the enthusiastic one - says that there would be no ultimate barrier to achieving machine translation at the very highest level - say, that of making a machine translate Shakespeare's sonnets - if existing artificial intelligence techniques were exploited to the full. The other extreme attitude - the iconoclastic one - says that, language being what it is, the very idea of translating one language into another by machine is derisory, since anyone who is sensitive to the nature of language can see a priori that high-level MT is impossible. Both these extreme attitudes are made even more extreme by the way they are put forward by the media; and by the fact that everyone claims the right to know all about the nature of language and to know nothing about the nature of machines.

Inside this conference, however, we have been dealing with the MT realities; and, by hearing detailed expositions and discussions of MT systems actually implemented, we have been enabled to enter a new, much more real, and quite different world. So this conference has successfully met a very real need.

There is immense sophistication of the position from three years ago, in that MT, in three variants, is actually being used far more. Nevertheless, the very fact that it is being used prompts the question: how far, by its nature, can MT go? Given the nature of machines and the nature of people, are there built-in limits to its improvability?

It seems to me that, among MT systems which have already been implemented, we can now discern three main approaches or trends. (I do not include among these, by the way, the interactive desktop translators' aids such as the Weidner, though I will have a word to say about these desktops later on.) And the overall point which I wish to make in this paper is that all of these three trends seem to me to have their own limits to further innovation built into them - and

that we have got to become clearer as to what these are. And the second point which I wish to make in this paper is that even if this is the case, there might be a new strategy which I will outline, which would help MT designers, over the longer term, to 'get out from under' these self-built-in limits to their own innovation, and which might also help MT to 'take a quantum jump' to 'make a new start'.

PART I: ARE WE RUNNING INTO A BLIND ALLEY?

I.1 The three already implemented approaches to machine translation

Now, as I have just said, we have been considering in some detail at this conference three approaches to MT which have actually been implemented. These are:

- i) Limited syntax, limited vocabulary MT, for a single field or a few related fields, such as we have heard is now being intensively developed at Xerox.
I will call this approach streamlined-input, single-field mechanical translation.
- ii) Secondly, there is free syntax, free vocabulary translation linguistically based, however, on a single pre-chosen text. The best possible example of this is the Aviation project of TAUM, at Montreal, and, interestingly, we have seen how this approach has developed to produce METEO.
I will call this approach linguistically based, single-text mechanical translation.
- iii) Thirdly, there is the far more open but also often far more defective approach of SYSTRAN, a system which can receive randomly chosen input but which, even when it does produce post-editable output, somehow still seems to be producing something less than translation. Following the European Commission, I will call this approach open-ended, random-input mechanical pre-translation.

Now I think it would be agreed that none of these approaches achieves, as yet, full translation. The live question is: how many of them are blind alleys?

I.2 The difference between having a built-in limit to innovation, and going down an MT blind alley

And here I must distinguish - and urge you who are here also to distinguish - the difference between an MT system, in general, having a built-in limit on its own innovation, and the more sophisticated phenomenon of going down an MT blind alley. A built-in limit to innovation will probably have to do with the size of the hardware; and it may or may not be remediable. For instance, there might now be on the market a handheld translators' aid with, say, 50 'frame' sentences stored in it, and, say, 25,000 words to fit into the frames; thus showing that this particular limit to innovation, namely, the limitation of having both far too few frame-sentences in the system and also far too few words, is being progressively overcome, as the system progressively profits by developments in the hardware.

Going down an MT blind alley, however, is something quite different,

and it is the central and important matter which we have to consider here. I define it as follows:

Definition of going down an MT blind alley

An approach to MT which probably already has had its own general but remediable limits to improvement built into it, runs into a blind alley when, as its efficiency builds up more and more, its capacity to develop and innovate becomes less and less.

I.3 Are the three already implemented approaches to MT running down blind alleys?

Now let us look in turn at the three actually implemented approaches to MT, the detail of which we have been hearing about at this conference.

i) Streamlined-input, single-field mechanical translation

This is the very interesting approach based on SYSTRAN, which has been implemented at Xerox. One of the many interesting features of this approach is that, just because the input is so severely streamlined, the system which handles it easily multilingualises. Thus the Xerox system is already processing four different language pairs; and it will surely go on to process many more. Another interesting feature claimed by Xerox system designers is that severely streamlining the Xerox MT input actually improves the English of the Xerox technical writers; and this is a claim which should be investigated further. Nevertheless, suppose it should be the case that the system designers at Xerox - and notably Mr Ruffino - wanted progressively to open up multinational customized English so as to make its input approximate more and more fully to complete English. If it turned out to be the case that the more they did this, the less the system could operate, then in so far as this actually happened, because the more efficiently this MT system program works, the less its designers can develop it or innovate.

ii) Linguistically based, single-text mechanical translation

Similarly, in so far as any system of the TAUM Q-system kind, which starts by producing perfect translation from a single text, develops, possibly even by multilingualising itself, in such a way that it becomes increasingly, and not decreasingly, tightly glued to its pre-chosen text - or, as in the case of METEO, its pre-chosen corpus of text - then, to the extent to which this happens, such an approach becomes, in my sense, a blind alley; in that the increasingly efficient extent to which it works, increasingly drives the linguistic model on which it is based to the point where - as M. Thouin himself has said - this linguistic model reaches its ceiling.

iii) Open-ended, random-input machine translation

The third approach - that of the more massive SYSTRANS - seen from this point of view, is clearly very different - because SYSTRAN is so very much more open-ended. Nevertheless, the masterly exposition which we have just heard from Dale Bostad of the current state of development of the USAF Russian-English SYSTRAN - claimed by him, I think, to be the most advanced MT

system in the world - shows the developers and maintenance-men of this system making a sustained and not always successful fight to prevent improvements to one aspect producing deterioration in another aspect of the system. That this is so can be seen by looking at the details of his paper; and, in so far as this SYSTRAN 'ageing' phenomenon has set in, even the Russian-English SYSTRAN is running into a blind alley.

However, and this is worth taking notice of, the EEC's SYSTRANS are being subjected to intensive linguistic development precisely in order to try and push back the limits to their improvability: so that Dale Bostad's claim for the pre-eminence of the USAF system is being challenged. For instance, in the EEC's English-French and French-English SYSTRANS detailed dictionary-entries inserted by Lawson, Pigott and Wheeler to make them better able to translate patent claims were so comprehensively designed that their insertion has improved the whole system's capacity to translate.

This is undoubtedly the kind of way to go forward: nevertheless, are the linguistic developers of the EEC SYSTRAN, or the EEC editors who are post-editing its output, prepared to say that this SYSTRAN also will never, like the others, run into a blind alley?

And finally, as a postscript, and to use Michael Hundt's own telling phrase used in discussion, in so far as sustained development-in-use of an interactive desktop translators' aid causes the human translator to have to do more and more, and the desktop machine less and less, so that the machine increasingly becomes 'a very expensive toy', in so far as this happens, the strategy for use of a desktop translators' aid has also run into its own form of blind alley.

PART II: IS THERE ANY WAY OUT FROM UNDER?

I perhaps should start by saying, with some firmness, that, even on the basis of the considerable amount which this conference has caused me to know about these differing approaches to MT, I do not wish to assert here categorically that any or all of them have actually, as a matter of fact, run into blind alleys.

What I do wish to assert, though, is that, in so far as the tendencies which I have characterised above really have set in, just in so far they have run into blind alleys.

And what I want to stress in this second part of my paper is my feeling that, even if this situation has set in, we need not for that reason just sit down under it; because I believe there is a way - possibly indeed several ways - in which we can break the straitjacket and get ourselves 'out from under'.

II.1 The three questions which I want to ask of the proponents of the three main MT trends

i) Streamlined-input, single-field mechanical translation

Since it would be admitted that, in this approach, the input is tailored to the point where the system can handle it, so that learning to write the streamlined texts which alone the machine

can handle becomes a special art, or at least a special skill, to which writers have to be trained: it seems to me that, if you take this approach, the question immediately arises: in the passage from the kind of unrestricted language which the machine cannot handle to the kind of limited-field, limited-syntax and limited-semantics language which the machine can handle, what is it about real language which gets lost?

ii) Linguistically based, single-text MT

Since it would be admitted that, in this Q-system approach to MT (which up to now has completely failed to generalise and tends to collapse into a 'tissue of particularities'), the distinction tends all the time to get blurred between a single dictionary-entry routine and a context-limited grammar-rule, the question, for those who adopted this computational-linguistic approach, immediately arises: what could be the basis for making a valid distinction between a context-limited grammatical rule and a dictionary-entry routine for handling a particular context in such a way that generality could be restored to the notion of a context?

iii) Open-ended, random-input mechanical pre-translation

In this approach, which up to now has been the main commercial one, it is not generality, but standard, which is lacking when the system fails, since the program incorporates within itself its own failsafe devices. But, just because the system will do something to handle any text, it can mistranslate, rightly translate by pure chance, or, unless backed by innumerable phrase-based dictionary-entries, produce only what looks like bald or uninspired target-language output ...

So the question arises: what is the 'translation' which this far more open, but also far more frequently defective approach overall and in general fails to attain?

Approaches i) and ii) do not need, over the short term, to ask this third question - since they both attain an intuitively acceptable standard of translation by severely limiting, though in different ways, the inputs which the programs are prepared to translate. However, since this third question is by far the most profound, it is the one with which to start: and note that it is the SYSTRAN approach, not the other two, which has provoked it. 'What is this "translation" which, overall and in general, this much more open but often defective approach fails to attain?'

II.2 Looking at this problem from the other end: can we make a model of how the human translator translates?

Note further that the question, as it arises out of this analysis, is one which requires a particular kind of answer. This fact can be brought out by rephrasing it as under: 'What conception of translation could serve as a guide to an Approach iii) type of open MT program to make it, when successful, translate more as a human being translates?'

The analogy is with chess. We thought, twenty years ago, that we knew all about how human beings could play chess - until we were

suddenly faced with the task of making the machine simulate the full chess-playing process. Then it suddenly became clear that, given the degree of explicitness which the simulation was going to require, we knew almost nothing about how human beings play chess. Similarly (with the honourable exceptions of Bachrach and Goetschalckx, who have made this point in repeated publications), we go on thinking we know all about how human beings translate - until, given the degree of explicitness which a simulation using MT Approach iii) is going to require, we suddenly come up against the fact that we know almost nothing about how human beings translate.

Moreover, again as in the chess case, a particular kind of explanation is required. Not a neuro-physiological explanation in terms of Arbib's 'top-down metaphorical brain'; not a self-conscious linguistic explanation in terms of what human beings say they do; but, as with analogical play in chess, as opposed to forcing play, an explanation in terms of the recognition of, and transformation of, pattern.

II.3 What follows is only the current state of the art of one such model

I have been working on such a model (see the annex to this paper), but this work is not sufficiently advanced to be worthy of publication; moreover, if it were, this conference would not be the right place to do more than mention that such a model may come into existence.

What can be said is that, if any such model of the way in which human translators translate could be constructed; and if the human translators on being presented with it came to feel 'Yes, this is indeed what we do when we translate, though until now we had not fully realised this', then the limits of innovation in MT could then be defined in a new and much more real way; namely, as limits to the extent to which the machine could simulate the model. Moreover, this kind of conception of the limits of innovation in MT would, at last, make MT approximate to other successful AI-based computer fields, and, by doing so, make MT itself, as a discipline, 'come of age'. Moreover again (always granted success with the human translators) such a model of translation might not only assist the further development of MT (Approach iii)). It might also unstick Approaches i) and ii) by enabling these to rebase themselves, and at a far more general level, on a conception of translation rather than on a particular corpus of text.

II.4 Two features of human translation brought out by the model - and which the machine may well never be able to simulate

Even from the preliminary glimpse of model-making activity shown in the annex, two features of translation-as-human-beings-do-it already come out; and it is interesting that, whereas one of these is already widely known, the other is not.

a) The extent to which the human translator reorganises the syntax

It will be seen, from the example given, that the model analyses text at three levels of depth. First, 'on the surface', there is the succession of breathgroups, to which syntactic and stress patterns can be attached. Secondly, 'lower down', there is a succession of

potentially reiterative semantic frames, which give an idea of what that breathgroup is doing in that sentence; and thirdly, 'lower still', there is a deep-semantic reiterative rhythm, which gives a simplified version of the 'theme-rheme' progression of the paragraph. Now, the first thing which has emerged from showing the model to a handful of translators (who, to an unexpected extent, received it favourably), is that it is to the two underlying levels, namely those of semantic pattern, as given in Levels 2 and 3 of the schema, that the human translators primarily react; not to the more detailed, superficial level of the syntactic and stress-patterning as given in Level 1. And this poses a problem for the future of MT, in so far as MT programmers set themselves to imitate the human translator's activity. For the human translator sticks very lightly to syntactic pattern; he or she does not hesitate to change the whole syntactic pattern of a sentence round, if this pattern is judged to obscure, rather than elicit, the basic underlying meaning of the paragraph. Human translators, of course, well know that they do this: they might well call this syntax-changing activity 'bringing out the underlying meaning of the text'. The only objection, from the systems analyst's point of view, to the use of this phrase is that there has not been, up to now, any underlying anything which the translators themselves could point to as being what it was they desired to bring out. It is this deficiency which the model exemplified in the annex endeavours to supply; though any other multi-level semantic model-maker is welcome to try and do better in this matter than I have. So I say here, yet once more to the translators: if you don't like my model, make a better model: but MAKE ONE, and when you have made it, send me a copy of it.

To return to my model; if all this is so - that is, if the human translators are primarily reaching down to Levels 2 and 3 of the analysis of the text which they translate, rather than confining their attention to Level 1 - then there are going to be early limits to innovation for any MT program which bases itself only upon Level 1. And, even if we can design a new kind of MT program which can form a 'picture' of Levels 2 and 3 of any input text (and to translate, say, Shakespeare's sonnets, there would need to be 'pictures' of N levels of the text), even then, its designer will be up against the problem of choosing between the very many ways in which the human translator might want to change round the syntax, 'in order to bring out the underlying meaning of the paragraph'.

In other words, as soon as we begin to make a deeper and more realistic model (but still a model) of how the human being really does translate, we come up against the fact that the machine may well not be able to simulate it.

b) The phenomenon of reinforcement of reiteration

And this brings me to the second feature of the model which may present a limit to MT innovation, and which, so far as I know, nobody has noticed or drawn attention to up to now. This second feature, which is not well known since it can only come to light from consideration of a reiterative model (which this is), is the human translator's tendency to bring out the underlying meaning of the text by using a mechanism of reinforcement of reiteration.

This mechanism consists of reorienting the syntactico-semantic 'frames' of Level 2, so as to bring out their relation to the underlying reiterative semantic 'theme', at Level 3. All I can do

here is to quote two or three very simple examples of this mechanism, which I have put in a separate section in the annex. But it is interesting that, as soon as one begins to notice the existence of this 'translational phenomenon', one becomes aware that, though one has never noticed it before, it is, in fact, going on everywhere; translators are doing it all the time. And of course, just because it is a device, a mechanism, there will be cases in which it can be incorporated into an MT program. With this proviso: the machine is not going easily to simulate a situation where the human translator produces a brilliant translation of some syntactic form of words of the source language - which is brilliant just because that form of words, in the whole long history of human translation has, equally evidently, never been given that particular translation before.

That already the attempt to make a new pattern-recognition model of the way in which human beings do, in fact, translate has highlighted two aspects in which the machine may well be unable to follow them, will be thought by many to be a reason for never making the model.

I disagree with this conclusion. If our knowledge of the nature of any human skill is ever to make progress by using the computer to simulate it, then the real nature of the skill in question has got to come out into the light.

It is because far too few mainstream academics have been willing to concede that the ability to translate well from one natural language to another is an exceedingly high-level creative human skill, that we have had so few attempts, up to now, really to analyse it.

- - - - -

ANNEX TO PAPER ON THE LIMITS OF INNOVATION IN MACHINE TRANSLATION

IMPRESSIONS OF PART OF A PATTERN-RECOGNITION

OF HUMAN TRANSLATION

Margaret Masterman

and

Bill Williams

Cambridge Language Research Unit

CONTENTS:

- 1 The impressionistic nature of this annex
- 2 The three levels of translational awareness
- 3 The model:
 - a) An example of what happens when the translator reorganises the syntax (from Table 1)
 - b) Examples of the reinforcement of reiteration in translation (from Tables 2 and 3)
 - c) An example of what happens to the prospects of human translation when there is no way at all of getting through to Level 3

Note on material

The authors wish to thank the Commission of the European Communities for permission to use for research purposes the English and French versions of the paragraph which has provided material for the three tables in this annex; and to thank David Shillan, sometime teacher of English as a foreign language, who made the first version of the structure of Table 1, and contributed many ideas towards Tables 2 and 3.

1) The impressionistic nature of this annex

This annex is not for systems analysts. It has been written for actual working translators. Its structure rests upon the layout of three tables; but, since the first and third tables only contain material from one paragraph, and the second table only material from one sentence from this paragraph, these structures cannot be evaluated on their own just by reading this annex. No claim whatever is advanced that any of the transformations exemplified in the tables is, in any straightforward way, mechanisable; many of the stages of a real MT program are left out, since human translators do not need to perform them; and, procedurally, the transformations exemplified in the tables are themselves impressionistic, since, in no table, is enough text processed to yield an algorithm.

Moreover, not only are these three tables impressionistic, but they are likely also to be inscrutable. For the technique which has inspired their layout does not come from the world of computing, but from that of telecommunications; a discipline which rejoices in multiplicity of levels and proliferation of means of access to these levels (because such a system is basically what a telephone system is), in order, when later processing the material, to open ways of evaluating all available speech-pattern options, and deciding between them.

What this annex does, therefore, is to present a first glimpse of a new way of looking at the phenomenon of translation; and it is presented here solely in the hope that some of the participants at this conference who actually do translations for their daily living will see fit to consider it. To them I ask: 'Can this model be considered to be to any extent a model of part of what you yourselves do when you translate - given that, for the purposes of making this model, translation is being considered, not in terms of the small-scale computing techniques of artificial intelligence, but of the large-scale pattern-recognition and pattern-transformation techniques which are used by telecommunication engineers when they handle spoken speech in order to disassemble, reassemble and distribute it?'

2) The three levels of translational awareness

After having been warned, the translator-reader is now invited to glance over the three tables.

At first sight, it may seem totally non-credible that these tables should depict or represent, in any way whatever, 'three levels of translational awareness'; or, alternatively, that they should represent three stages in the human operation of translation.

The tables, however, are not linguistically conventional tables.

Each of them represents both the stage before, and also the stage after, an operation of human comprehension has been achieved.

Thus, Table 1, starting from the two ends of this matter, simulates first (i.e. on the left) an English reader's instinctive tendency to read a text not in words, not in sentences, but in breathgroups. The right-hand column of the same table shows a French reader (i.e. of a text which translates the English) doing the same thing; and the arrow-symbolism of the table attempts (but fails) to connect the English breathgroup sequence to the French one. (For further explanation of this fact, see Section 3a of this annex.)

Table 2 simulates the next stage of the translator's penetration into meaning. Here the English breathgroup sequence of a sentence forms the centre column of the table, and a (traditional) semantico-syntactic description of the words of the breathgroups is on the left. The professional translator, unlike the machine, will not need this; because he will know, instantly and instinctively, that, for instance, an adverbial clause of time is an adverbial clause of time. However, the table assumes that, while the translator takes in all this (or alternatively, takes in some other, less traditional, grammatical 'picture' of the English), he is also 'echoing' across the breathgroup - being guided in the direction of the echo by that ineffable but real thing called 'the rhythm of the prose' - some subliminal reiterative signal of what the semantic 'tang' of the whole breathgroup is: and some signs that he really does do this are discussed in Section 3b.

However, even this is not all that he is doing: for, on this model, the translator is not like a photographer taking in a 'picture' of the text's meaning: he is much more like a musician absorbing a 'tune' made up of notes and of chords. For Table 3 shows, in the right-hand column, a sequence of signals which reiterate more slowly and more often and which may look merely like the slowly changing notes of a ground bass, but which also 'tell him what the whole paragraph is about'; and some evidence that he both gains, and profits by, this knowledge is advanced, though in a very preliminary way, in Sections 3b and 3c.

There is, almost certainly, at least one other 'deeper' level even than this last: namely, a more large-scale reiterative level which intuitively makes comparisons or contrasts between paragraphs. And the translational reiterative 'signals' displayed in the tables may well be much too few, much too crude, or even the wrong ones.

Nevertheless, the claim is here very seriously advanced - and, as I think, for the first time - that all this, and probably also much more, is what a human translator, largely without thinking about it does; and that all this, and probably also much more, is what the machine, if it is to simulate really creative translation, must learn to do; which prompts the thought that there may indeed be real limits to machine translation.

Machines of course can, and increasingly do, learn - that is, once their designers have realised what the nature of any learning-process must be.

Nevertheless, what a piece of work is man ...

- 3) The model
- a) An example of what happens when the translator reorganises the syntax
(from Table 1)

Of the three tables, Table 1 represents that overall transformation-process, from English to French, which starts by segmenting the English input text into stress-patterned breathgroups and then trying to translate it into French breathgroups. There is an increasing amount of evidence to suggest that (in simultaneous translation pre-eminently, but also for example in patent translation) the translator's primary unit is neither the word nor the sentence, nor the short phrase, but the whole breathgroup. So the structure of Table 1 presupposes a mechanism for simultaneously translating English to French, breathgroup by breathgroup.

However, the French text as shown does not fit this mechanism; for, far from being instantaneously emitted by a hard-pressed simultaneous translator panting in breathgroups, it was deliberately composed by a comparatively leisured EEC translator who strongly disapproved of the structure of the English syntax. (It is, in fact, the appalling contorted style of the English text (and, since this is an EEC text, only this fact) which prompts the judgment that the French was the original document.)

The table shows, therefore, how, when the translator reorganises the syntax, even quite a sophisticated MT attempt to simulate human translation breaks down. For (to take only one example of the breakdown) 'd'attribuer à Portugal' (French BG 1,3) does not correspond to the sequence of the two half breathgroups 'to allocate' (English BG 1,7) and 'in Portugal' (English BG 1,4). Moreover, 'une aide d'urgence de 100 000 UCE' is much more elegant than 'to allocate 100 000 EUA in emergency aid', but would not very easily be reached from it.

Now, of course, there are other ways to undertake the mechanical translation of this paragraph so as to produce from this input this output: the reader is invited to work one out for himself. But, given that the large-scale comparative analysis of English and French official Canadian texts, undertaken by CLRU in the 1960s, showed almost no sentence in which the translator did not basically reorganise the English syntax and/or split the original English sentence into two, three or even four, this predilection of human translators represents, in my view, a genuine limitation to the mechanisability of really high-level, deliberately composed, human translation.

The extreme Canadian example of such syntactic non-correspondence unearthed by CLRU was:

English - 'Serious consideration has been given by the government'
French - 'Les militaires ont décidé'

- b) Examples of the reinforcement of reiteration in translation
(from Tables 2 and 3)

Tables 2 and 3 open up an enormous subject: that of the correct reiterativeness of language.

It is not difficult to show that language is cardinally,

indefinitely, in surface and in depth, reiterative. (See indeed this last sentence as an example of this.) It would be agreed that it is always possible to rewrite and extend a surface text so as to bring out its latent capacity for reiteration - thus producing something between an incantation and a lilt. For example:

Text: The potters also made a contribution to the culture by their pottery.

Lilt: The potters,-
 yes,
 even the potters,
 the potters made a contribution,
 a contribution to the culture -
 And how did they make
 a contribution to the culture?
 They made a contribution
 to the culture
 By their pottery.
 (This pottery,
 etc.)

Now, normally, surface texts are not lilt: and normal computational linguistics does not allow for lilt. But my point here is - and the overall objective of constructing Tables 2 and 3 is to make this point explicit - as soon as you are endeavouring to recognise and simulate the pattern-recognition and pattern-transformations characteristic of the activity of translation - as opposed to those characteristic of unilingual analysis - the need both to draw on and to draw out language's latent capacity to reiterate at once becomes clear.

There are great difficulties in doing this, especially at Level 2. By comparison with those of Level 2, the simple, 'thudding' fourfold reiterations of Level 3 are much easier to find on the map. For instance, in Table 3, if the right-hand stressed words of each breathgroup are examined and counted, it will appear that 18 of the 22 of them fairly directly relate to the basic sequence of Level 3 reiterations of the paragraph's underlying semantic theme. (This paragraph turned out to be a happy example.) At Level 2, however, where the much smaller-scale reiterations of the auxiliary semantic elements have both to be drawn out by, and to be controlled by, the use of the deeper syntax, the extent to which this is done will vary with the model-maker.

Nevertheless, the extent to which this elicitation of reiterativeness is in fact done by translators - and however unlikely it may at first seem that it ever would be done - can be seen clearly by actually trying to use the tables; for the proper way to discover the potential and the errors in tables of this type is not indefinitely to comment on them or analyse them, but to use them.

I will start by drawing your attention to an extreme Level 3 example of what really happens in human translation; namely the fact that two occurrences of the same input phrase 'for disaster victims' (in 0,2 and in 2,3) have been given very considerably differing French translations. The first time the phrase occurs (in the title) it is translated as (see Table 1):

0,2 English - for disaster victims
 French - en faveur des personnes sinistrées

where the vague phrase 'personnes sinistrées' - 'people who have fallen on evil days' strongly suggests that such people (at Level 3) also need HELP; and, at Level 2, the idea of helping, by positively discriminating in such people's favour, is reinforced by the translation of 'for' not as 'pour' but as 'en faveur de'. By contrast, the English input phrase does not easily allow the covert reiterative idea of HELPing to break through to the surface: for a victim is a man who has been DISASTERed, and the DISASTER is what has made him into a victim: so the possibility of eliciting and reinforcing the idea of HELP is squeezed out by the predominance of the Level 3 reiteration of DISASTER. Contrast this context with that of 2,3:

2.3 English - for aid to disaster victims
 (2.4 " - in non-Community countries)
 French
 (des fonds
 destinés à des pays tiers)
victimes de catastrophes

Here the French reordering makes it vital for the translator to stress how deep was the CATASTROPHE, which made its victims, even in non-Community countries, qualify for HELP.

Now, the whole of the analysis given above is controversial: both because the French translation given here is not the only French translation, by any means, which could be made of this passage: and also because the Level 2 and Level 3 reiterations given here are not the only Level 2 and Level 3 reiterations which could be modelled. Nevertheless, in spite of all such doubts, and many others, I insist on two points. The first is that this translational device of reinforcing a Level 2 and Level 3 reiteration both can occur and does occur in creative human translation. And secondly, when it does occur, we immediately say 'Oh, what a brilliant translation', or 'Now, this is real translation', and that we particularly and instinctively admire the use of this device when this translation of, say, an English auxiliary phrase has never occurred in our experience of translation before. So it is up to us to complete, correct, improve and streamline our Level 2 and Level 3 tables: not to try and evade the fact that the reiterations are there.

As evidence for this, I will give now two Level 2 reiterations which are not given in any of the tables (in order to inspire the making of more and better tables). Another French translator, not the EEC one, tackling 1,2 produced the following:

1,2 English - caused by the storms and floods
 New French translation of 1,2
 - provoqués par les orages et les inondations
 1.3 - qui ont eu lieu au Portugal

Here the Level 3 theme of DISASTER reinforced by the Level 2 idea of natural happening has produced the further disaster-reinforcing idea of provocation by the violence of the floods. Likewise the translation (but not the original) reinforces the Level 2 reiterative idea that Portugal is a place.

A further example, from later in the same EEC passage but not given in the tables, is:

English - towards financing
 French - pour contribuer au financement
 ('to contribute to the contribution')

To conclude: there may be further need to stress the subliminal existence of the Level 2 and Level 3 phenomenon.

What there will be no need to stress is the fact that the difficulty of making a machine simulate it indicates the existence of a limit to innovation in machine translation.

c) An example of what happens to the prospects of human translation when there is no way at all of getting through to Level 3

This material, which is reproduced unaltered, comes from teaching material which was supplied for this paper by the kindness of the extra-mural department of Birmingham University:

'With hocked gems financing him, our hero (1) bravely defied all scornful laughter that tried to prevent his scheme. "Your eyes deceive", he had said. "An egg, not a table, correctly typifies this planet." Now three sturdy sisters (2) sought proof. Forging along, (3) sometimes through calm vastness, yet more often over turbulent peaks and valleys, days became weeks as many doubters spread fearful rumours about the (4) edge. At last, from nowhere welcome winged creatures appeared signifying momentous success (5).'

(See Dooling and Lachman, 1971, page 217.)

The absent Level 3 information can be found below.

- (1) - add ', Christopher Columbus'
- (2) - delete plural 's' of 'sisters' and add 'ships'
- (3) - add 'over the endlessly changing sea, '
- (4) - add 'nearness of the world's'
- (5) - add 'of the enterprise of finding land not behind but ahead.'

TABLE I
LEVEL 1 - Surface Syntactic Elements

ENGLISH		FRENCH SYNTHESIS		FRENCH	
SV	2				
BG 0,1	Emergency aid			Aide d'urgence	
WS	2				
SV	1				
BG 0,2	for disaster victims.			en faveur des personnes sinistrées.	
WS	2				
SV	1				
BG 1,1	On 11 April			La Commission a décidé	
WS	2				
SV	1				
BG 1,2	as a result of the serious damage			le 11 avril 1979	
WS	2				
SV	2				
BG 1,3	caused by the storms and floods			d'attribuer au Portugal	
WS	2				
SV	1				
BG 1,4	in Portugal			une aide d'urgence	
WS	2				
SV	1				
BG 1,5	in February and March,			de 100 000 UCE	
WS	2				
SV	0				
BG 1,6	the Commission decided			pour la fourniture de biens essentiels	
WS	2				

a b c d e f

TABLE I (cont'd.)

ENGLISH		FRENCH SYNTHESIS		FRENCH
SV	0	2		
RG 1.7	to allocate 100,000 ECU	2		à la suite des tempêtes et des inondations
WS	1	2		
SV	1	2		qui ont causé de graves dommages
RG 1.8	in emergency aid	2		
WS	1	2		
SV	1	2		en février et en mars.
RG 1.9	for essential supplies.	2		
WS	1	2		
SV	0	2		Cette aide,
RG 2.1	This aid,	2		
WS	1	2		
SV	2	0		qui provient des fonds
RG 2.2	granted from the appropriation	2		
WS	1	2		
SV	1	2		destinés à des pays tiers
RG 2.3	for aid to disaster victims	2		
WS	1	2		victimes de catastrophes,
SV	1	2		fait suite aux entretiens
RG 2.4	in non-Community countries,	2		
WS	1	2		
SV	2	2		que M. Natali a eus récemment
RG 2.5	follows Mr. Natali's recent talks	2		
WS	1	2		
SV	1	0		
RG 2.6	with the Portuguese authorities,	2		
WS	1	2		

TABLE I (cont'd.)

ENGLISH			FRENCH SYNTHESIS			FRENCH		
SV	1	2	0	0	2			
EG 2,7	who pointed to the extensive							
WS	1	2	3	4	5			avec les autorités portugaises
SV	2							
BG	infrastructure damage							
WS	6	7						
SV	2	0	0	2	4			
BG 2,8	caused by the floods.							
WS	1	2	3	4				qui ont indiqué
SV	1	2						
BG 2,9	This Commission aid operation							
WS	1	2	3	4				que ces inondations avaient endommagé
SV	1	0						
BG 2,10	will be implemented							
WS	1	2	3					un grand nombre
SV	0	2						
BG 2,11	by Secours Catholique Français.							
WS	1	2	3	4				d'infrastructures.
SV								
BG 2,12								
WS								l'aide de la Commission
SV								
BG 2,13								
WS								sera mise en oeuvre
SV								
BG 2,14								
WS								par le Secours catholique français.

TABLE I (cont'd.)LEGEND

SV - Stress Value
 BG - Breath Group
 WS - Word Sequence

----- - Direct translation of Breath Group.
 - - - - - Translation where Breath Group has to be split.
 -o-o-o-o- - Breakdown of the translation process and reassembling of information,
 as opposed to the grammar.

Note

When the Breath Group is stressed and segmented, a syntactic code is added at that time. This coding has been omitted from the table, because the human translator does not do this as a separate operation.

TABLE 2
LEVEL II - DEEP SYNTACTIC AND AUXILIARY SEMANTIC ELEMENTS

LEVEL II (A) SYNTACTIC SPECIFICATION OF LEVEL I OUTLINED BREATH GROUP	BREATH GROUP OVERALL SPECIFICATION	LEVEL II (B) SYNTACTIC/SEMANTIC REITERATIONS	
		HEAD STRESS	NUCLEAR STRESS
(i) Adjective of kind. (ii) Noun of appropriate action.	EMERGENCY AID (Phrase of subject of title). (i) Adjective of kind. Noun of appropriate action. (ii)	[HUMAN ACTION]	HUMAN ACTION
(i) Adjective of circumstance. (ii) Animate noun.	FOR DISASTER VICTIMS (Phrase of specification of subject of title) (i) Adjective of circumstance. (human being) noun. (ii)	[HUMAN BEING]	HUMAN BEING
i) Adverb of point of time. ii) Numeral (date). iii) Noun of period of time (month).	ON 11 APRIL (Adverbial phrase of time) i) Adverb of point in time. ii) Numeral date. iii) Noun of period of time.	[TIME]	TIME
i) Adverb of reason why. ii) Adjective of intensification. iii) Noun of change.	AS A RESULT OF THE SERIOUS DAMAGE (Conjunctive phrase of "reason why"). i) Compound adverb of reason why. ii) Pejorative adjective (of intensification). iii) Pejorative noun (of sudden change).	PEJORATIVE	PEJORATIVE

TABLE 2 (cont'd).

LEVEL II (A) SYNTACTIC SPECIFICATION OF LEVEL I CONTOURED BREATH GROUP	BREATHE GROUP OVERALL SPECIFICATION	LEVEL II (B) SYNTACTIC/SEMANTIC REITERATIONS	
		HEAD STRESS	NUCLEAR STRESS
1,3 i) Predictive adjectival indicator of cause. ii) Definite article. iii) Noun of natural happening. iv) Co-ordinate conjunction. v) Noun of natural happening.	CAUSED BY THE STORMS AND FLOODS (Dependent clause of cause). i) Predicate adjectival indicator of cause ii) Definite article. iii) Noun of natural happening. iv) Co-ordinate conjunction. v) Noun of natural happening.	NATURAL HAPPENING	NATURAL HAPPENING
1,4 i) Adverb of place. ii) Noun of region.	IN PORTUGAL (Specifying phrase of region where. i) Adverb of place. ii) Noun of region.	REGION	[REGION]
1,5 i) Adverb of time. ii) Noun of period. iii) Co-ordinating conjunction. iv) Noun of period.	IN FEBRUARY AND MARCH, (Specifying phrase of time during which). i) Adverb of time. ii) Noun of period of time (month). iii) Co-ordinating conjunction. iv) Noun of period of time (month).	PERIOD OF TIME	PERIOD OF TIME
1,6 i) Definite article. ii) Proper noun of corporate body. iii) Finite past verb of decision.	THE COMMISSION DECIDED (Subject/main verb). i) Definite article. ii) Proper noun of official corporate body. iii) Finite past verb of official decision.	OFFICIAL	[OFFICIAL]

TABLE 2 (cont'd.)

LEVEL II(A) SYNTACTIC SPECIFICATION OF LEVEL I CONTOURED BREATH GROUP	BREATHE GROUP OVERALL SPECIFICATION	LEVEL II(B) SYNTACTIC/SEMANTIC REITERATIONS	
		HEAD STRESS	NUCLEAR STRESS
1,7 i) Infinitive "to". ii) Infinitive verb of corporate donation. iii) Numeral of amount. iv) Proper name of international financial measure.	TO ALLOCATE 100 000 EUA (Complement; infinitive clause of corporate action). i) Infinitive "to". ii) Infinitive verb of corporate financial donation. iii) Numeral of amount. iv) Proper name of international financial measure.	[FINANCIAL]	→ FINANCIAL
1,8 i) Preposition of kind of action. ii) Adjective of kind of action. iii) Noun of approbative human action.	IN EMERGENCY AID (Prepositional phrase of kind of action). i) Preposition of kind of action. ii) Adjective of kind of action. iii) Noun of approbative human action.	KIND OF ACTION	→ [KIND OF] ACTION
1,9 i) Preposition of specification. ii) Adjective of intensification of need. iii) Concrete aggregate noun.	FOR ESSENTIAL SUPPLIES (Prepositional phrase of specification). i) Preposition of specification. ii) Adjective of intensification of need for life-support. iii) Noun of substance, concrete aggregate noun (of life-support).	[LIFE-SUPPORT]	→ LIFE-SUPPORT

TABLE 2 (cont'd.)

LEGEND

[] - Semantic echo

← - Direction of echo

↔ - Co-ordinate conjunction

NOTE 1(A) : Historical Note. To form the right hand column from the middle column, the descriptors which reiterated in the overall description were extracted - after which most of the second occurrences of the reiterating elements were removed from the overall description.

NOTE 1(B) : Logical Note. To form the left hand column from the middle column, all the non-reiterating elements are extracted - except where the Breath Group contains a co-ordinating conjunction, in which case the reiteration occurs in all three columns.

NOTE 2 : The absence of a "direction of echo" arrow indicates that the reiterative elements are of equal weight.

TABLE 3

LEVEL III - Influence of Basic Semantic Elements

	BREATH GROUP	BASIC SEMANTIC ELEMENT
0,1	<u>Emergency</u> +aid	<u>HELP</u>
0,2	for <u>disaster</u> +victims	<u>DISASTER</u>
1,1	On 11 April,	
1,2	as a result of the <u>serious</u> + <u>damage</u>	<u>DISASTER</u>
1,3	caused by the <u>storm</u> + <u>land</u> + <u>floods</u>	<u>DISASTER</u>
1,4	in <u>Portugal</u>	
1,5	in <u>February</u> and <u>March</u> ,	
1,6	the Commission <u>decided</u>	
1,7	to <u>allocate</u> 100 000 <u>EUA</u>	<u>HELP</u>
1,8	in <u>emergency</u> + <u>aid</u>	<u>HELP</u>
1,9	for essential <u>supplies</u> .	<u>HELP</u>
2,1	This <u>aid</u> ,	<u>HELP</u>
2,2	granted from the <u>appropriation</u>	<u>HELP</u>
2,3	for aid to <u>disaster</u> + <u>victims</u>	<u>DISASTER</u>
2,4	in non-Community <u>countries</u> (<u>communities</u>)	<u>DISASTER</u>
2,5	follows Mr. Natali's recent <u>talks</u>	<u>TALKS</u>
2,6	with the Portuguese <u>authorities</u>	<u>TALK(ER)S</u>
2,7	who <u>pointed</u> to the extensive <u>infrastructure</u> <u>damage</u>	<u>TALKS</u> <u>DISASTER</u>
2,8	* <u>caused</u> by the <u>floods</u> .	<u>DISASTER</u>
2,9	This Commission <u>aid</u> + <u>operator</u>	<u>HELP</u>

TABLE 3 (cont'd.)

	BREATH GROUP	BASIC SEMANTIC ELEMENT
2,10	will be <u>implemented</u>	<u>HELP</u>
2,11	by <u>Secours — Catholique — Francais.</u>	<u>HELP</u>

LEGEND

- HELP — Basic Semantic Element in a Breath Group.
- HELP — Basic Semantic Element referred, by the syntax-pattern, from one Breath Group to another.
- HELP — Basic Semantic Element "echoed" to a further adjoining Breath Group.

* 2,8 That this Breath Group, in this second context, is what people said, is shown by such facts as that, if the passage were to be translated into Latin, it would have to be in indirect speech.

Note 1: No indication is given by this table as to the nature of a pre-supposed process of semantic disambiguation.