MACHINE TRANSLATION OF LANGUAGES

by

Dr. A.D. Booth
Birkbeck College,
University of London

It is now nearly 10 years since the discussions between myself and Warren Weaver of the Rockefeller Foundation, which led to the suggestion that mechanical translation might be a possibility. During this period much work has been done to reduce the naive ideas of our original discussions to a practical system whereby language can be translated, and it is appropriate to look back now and to see what has been achieved and to venture a prediction as to what is likely to be achieved in the next few years.

In the first place, it is necessary to disentangle the subject from the mass of speculation with which it has been clothed since professional linguists became interested in it. In its original form machine translation was merely to be another application of a high-speed computing machine. The principles upon which it was based were well understood and the philosophical problems in so far as they existed at all, were not discussed. The original proposition was simply that a computing machine, because it contains a large unit known as the store or memory, would be able to hold in that store a dictionary and since access to the store of the computer is an extremely rapid operation, it might be that by using a computing machine the time wasted in looking up words when translating from a foreign language could be saved. This is as far as the original proposals went.

There are no difficulties in principle in carrying out this scheme, and the reason that no practical trial was made was simply that in 1947 no computing machine was inadequate store or in fact no computing machine at all existed. By the time computing machines became more readily available, and certainly by the time that their stores became adequate, mechanical translation became so invested with complexity, that the original scheme was not deemed worthy of a trial and some of the more exciting of the modern developments were unfortunately still too difficult to achieve.

It has been said that the early proposals merely consisted in looking up words in a dictionary, yet it is evident, on examining this proposition, that certain difficulties arise. Chief among these is the fact that when a human being examines a dictionary to find a foreign language word, he is using not only the dictionary itself, but also certain knowledge of the structure of the foreign language which he himself possesses. Without this knowledge it would be almost impossible to translate a significant proportion of the words from a foreign language text using a dictionary alone, and the fact that it has been plain that translation is possible by means of a dictionary, is simply due to the fact that many of the foreign languages are similar in structure to the languages with which most translators are familiar. This situation is brought clearly to my mind by an example which occurred very early in my own school days when I attempted to look up the German word "grösse" in a German dictionary. Two difficulties arose - firstly, that the dictionary was not written in the archaic script in which the text I was using appeared, and secondly, that the letter "ss" could not be found in it. In any case, this particular word would be given, but in the simpler form "gross".

The minimum requirements of the computing machine when applied as a mechanical dictionary are that it should be at least as good as a real dictionary and preferably that it should be as good as a real dictionary taken in conjunction with the user, who is not a professional linguist, but at least has a smattering of the language. The practical realisation of a plan of this sort is not difficult, and in the years between 1947 and 1950, I, together with one or two collaborators, evolved a system whereby a relatively small dictionary used in a high speed computer, together with certain rules of procedure would enable a significant number of words in the foreign language text to be looked up.

The developments which made this possible were briefly as follows. First, the realisation that dictionary requirements could be greatly reduced if the dictionary itself were constructed on a new plan. This was simply that instead of one of the complete parts of the words, which are normally given in a foreign language dictionary, there should instead be given merely the stem or possibly stems of the foreign language word. The definition of stem in this context is simply that it is the longest segment of a word which is common to different forms of that word to which different endings can be added, in order to produce real words in the foreign language, and to forestall questions I may perhaps mention that in the case of irregularly manipulated words, for example, in French "avoir" and "etre", not one stem is stored, but a whole number of stems, and possibly complete forms of words in difficult cases. At this early stage, it was considered that the provision of stems would be sufficient. Endings could be dealt with either by looking up in a special ending dictionary and adding a few grammatical notes or by ignoring them altogether. It will be seen that this division of words into stems and endings greatly reduces the size of the dictionary required to recognise a significant fraction of the words in a foreign language text.

Since the store of all existing high speed computers is limited to a few thousand words each having about 10 letters of alphabetic type, it is not possible to store anything like a complete dictionary of the foreign language concerned. To overcome this difficulty, it was proposed originally to store only those words which were relevant to the context of the passage being translated, and it was assumed that in most cases the person requiring the translation would be sufficiently familiar with the subject matter to know which types of words were going to be encountered in a given context. Sets of words peculiar to given subjects, such as Plant Genetics, Nuclear Physics, X-Ray Crystallography, and so on, were known as microglossaries. Apart from these words, which were specific to the subject in hand, it was proposed to store also a number of words of general utility. These are usually called by linguists "cement words". Statistical investigations seem to indicate that about a thousand words of general utility and about a thousand words specific to the subject under translation, enable about 90% of the words in most scientific texts to be translated, and it was, and is, in fact, now maintained that this proportion is sufficient to justify the operation of translating by a special machine or even on a computing machine. It may perhaps be argued here that the "cement words" are unnecessary. Many scientists read foreign language texts by the simple process of looking up or recognising the nouns and special verbs and also the symbols and diagrams which occur in the passage being translated. Translation made by the stem ending procedure together with microglossaries would of course fall short of the sort of thing which the skilled human translator could produce.

These ideas received a small practical trial on a punched card installation. The results were not regarded as sufficiently important to justify extensive publication, but the experiments were reported to a Conference on

Mechanical Translation which was held at the M.I.T. in 1953. During the period under consideration, work had been carried out on the other side of the Atlantic, noticeably by Byhillo, Reifler, Oswald and Fletcher and by Perry. This work parallelled that which we were doing in England, but to a large extent bypassed the simple procedures which we advocated and proceeded at once to far more sophisticated situations, which it was hoped might lead to the translation of language in a more literary fashion.

Among the ideas which were produced during this period, it is perhaps worth mentioning those of Reifler on pre- and post-editing. The first stumbling block encountered when any translation of the simplest form is envisaged lies in the difficulties of ambiguity which are inherent in the original text. It may perhaps be argued that a good author would not commit ambiguities. This however, is not true. It is quite easy for an author to write passages which are to such an extent peculiar to the language in which he is writing that it perhaps is not immediately recognised that they are ambiguous. Examples in English are the expressions "she cannot bear children", in which of course "bear" is ambiguous, and "these men are revolting", in which "revolting" has a quasi-humorous commutation. These sentences as they stand have two completely distinct meanings and it is quite impossible to decide which of them is intended without reference to a more extensive context. At the same time, the sentences concerned are neither badly written, nor are they particularly unfortunate examples of their kind.

Reifler's pre-editor was to be a person versed only in the foreign language from which it was required to effect translation. Its function would be to remove all the known ambiguities which the language itself contained. Since he would not have to be an expert in any language but his own, it should not be difficult to obtain a sufficient number of such people to do the work concerned, and it was considered that editors might perform the function of pre-editing before papers appeared in their journals. The post editor was to be a native of the language into which translation was to be effected. He would not necessarily know the foreign language although, of course, it would be an advantage that he should, and his function would simply be to render the output of the translating machine into acceptable prose in the target language. Both Reifler and his colleagues now appear to have abandoned the ideas or pre- and post-editing. In my original proposals the post editor was to be the scientist for whom the translation was to be made, and it was assumed that if he had sufficient knowledge of the subject of translation he could resolve the ambiguities in the machine output, either on his own, or possibly with certain auxiliary reference to the machine.

Up to the 1953 Conference when linguists had deigned to be interested in machine translation at all, they tended to be extremely sceptical of its possibility. It is still true that many workers in the field appear from their publications and utterances to be more interested in demonstrating the difficulties of the subject than putting the matter to a practical test, and this attitude perhaps marks the differences between the American and the British schools. In England it was proposed to conduct limited tests and see what can be done and thereby to obtain encouragement for further research. In America on the other hand it appears to be thought not worth while to conduct any tests at all unless they can produce a completely literate form of output. I find myself completely in disagreement with the American point of view. It seems to me, that unless the experimental method, as it is known and understood in the exact sciences, can be applied to this new science of mechanical translation, then it is unlikely that translation by machine can become really practicable. It is easy to object that it is known a priori that the

output will be imperfect, and therefore that nothing at all should be done. This is likely eventually to bring the whole subject into disrepute.

After the 1953 Conference in America, one practical experiment on translation from Russian was conducted. It made use of a limited vocabulary of 250 words and had the feature that certain grammatical rules were imposed upon the output of the machine so that the resulting prose was in acceptable form. Apart from this experiment, little practical work has been done in the United States, although of course, the extremely valuable survey of Oswald and his collaborators on German syntax should be mentioned.

In England a steady progress of the simple experiments was greatly assisted by a grant by the Nuffield Foundation in 1955 which enabled the Birkbeck College group to take on the services of a full-time professional linguist, then at Brandwood in my laboratory. When we had the linguist, of course, we had to choose the language to start operations on, and I in fact chose the French language for two reasons:- Firstly that it is structually similar to English, and secondly, the purely utilitarian reason that it is a language in which most scientists, and in particular myself, are reasonably well versed. This latter condition is important since at the moment there appears to exist in England no linguist sufficiently familiar with the operations of high-speed computing machines to enable him to prepare his material for machine use. If the intermediary of a mathematician or physicist has to be used, then it is, I think, rather important that both the linguist and the machine programmer should understand one another, otherwise it is unlikely that an efficient programme will be produced.

Just to bring this historical survey up to date, I might mention the experiment which was announced at the Institution of Electrical Engineers Conference in London in 1956. This was a Russian experiment conducted by Mukin who claimed to have translated on the Moscow machine "Besm" from English into Russian. And finally, I hear just recently, that the Americans at Rome base are constructing a machine specially for linguistic research sufficient to give an enormous impetus to the work which is going on there in the more abstract fields.

Well, now I shall come to a more technical discussion of the way in which we effect translation on the machines. Calculating machines are designed essentially to deal with numbers. In many cases these numbers are represented in the binary scale, and this of course, is no disadvantage when translation is concerned, but all the form of input of a message for translation is via a teletype tape and it is well known that teletype apparatus automatically converts alphabetic symbols into a five-hole binary code. This five-hole code is suitable for input directly to an automatic computer, and once the foreign language word has been transcribed into teletype code by means of the teletype keyboard, one can regard the aggregate of coded impressions which go to make up a word as a binary number. The machine can then operate on this number in the normal way.

The basis of any translation is a dictionary and for this purpose it is necessary to restrict the dictionary to a size which can be contained within the storage organ of the machine. With the machines which are at present available, and in the Birkbeck machine as well, this restriction implies that the size of the store dictionary can only at the best be a couple of thousand words. This is because for one thing the computer word is very much shorter than the word of real language. It may be argued that a dictionary of such a size will be quite inadequate for general translation even from a simple language, and this is of course true. At the start of the considerations of machine translation, which I mentioned before, its limitation was realised

and we have the statistics for translating with limited vocabularies. In the earliest experiments the only words which could be recognised were those which were exactly matched by a word contained in the dictionary, and it was therefore quite impossible to translate anything other than the microscopial proportion of the words in the text of an inflected language, and that is of course, most languages other than English. The method of overcoming this, however, was not difficult to find. The ideas of stems and endings must be familiar to anyone who studied Latin, French or any other European language. From the machine point of view a new sort of dictionary is constructed. This consists of stems of words, which can be defined as the longest segment, as I said before, common to all of the parts, for example, with the Latin word "amo" the stem is "am", which can take the endings "o", "as", "at", "bo", "bis", "bit", and so on.

Associated with the stem dictionary is a completely separate dictionary to contain the endings. There are three main methods of dictionary search which are available using a computing machine. The first completely impracticable, the second practicable but undesirable, and the third, the one we use. In the first method the translation of a word is stored in that storage location whose position number is the code number of the foreign language word. Now you can see that even if we restrict words to having a maximum of ten letters this would imply that we should require 26^{10}, or about 1.4×10^{14} storage locations to contain our dictionary, and this of course is quite ludicrous. The second searching method which is in fact one which was used for some time and which is still used by some groups, works as follows:-

The dictionary is stored in the store with the words in ascending order numerical magnitude. The search procedure involves starting at the last or numerically greatest word in the dictionary and subtracting this and these words in turn from the foreign language word. The result of the subtraction will be negative until the stem translation is reached, and after this point it will become positive. It is clear that if we choose a word at random, on the average it will be necessary to examine about half of the dictionary words before the translation is found. This is highly undesirable because, of course, even with fast modern machines a dictionary of 1,000 words would probably require of the order of 10 seconds for locating any given word, and with a really large dictionary very much longer times would be needed. A human operator could in fact do better using an ordinary dictionary.

This highly undesirable situation is improved by what I call the bracketing technique. In this, the foreign language word is first subtracted from a word which is roughly halfway between the start and finish of the dictionary. If the result of the subtraction is positive, then it is clear that the word to be translated lies in the first half of the dictionary. If, on the other hand, the result is negative, then the word is known to lie in the second half. This process is repeated using this type of word which is either one quarter or three quarters of the way through the dictionary. We found that the foreign language lies by this procedure in some particular quarter of the dictionary. Repetitions of the process will limit the possible position of the translation to one eighth, one sixteenth and one thirty-second, and so on, of the possible dictionary words. It is quite easy to show that if the number of words in the dictionary is D, then in fact you require on the average about $\log_2 D$ look-up operations. To put this into numerical perspective, I can mention that

if you take a dictionary of the size of the Concise Oxford Dictionary, which has about 50,000 words, the partitioning method will require about fourteen or fifteen looking-up operations.

Were I here in person I usually conduct a small experiment with the dictionary at this point to show the validity of the procedure which I mentioned. And if we had a dictionary of a million words we should still only need just over twenty look-up operations at any time. Now it is of course possible to complicate this dictionary procedure by partition ing the dictionary into sections containing the most frequent words, and that containing the less frequent. It turns out, however, that the programming required to operate a dictionary of this sort makes the time saving which it could produce unimportant, and therefore we do not in fact do it. Having looked up the word in the stem dictionary, various courses of action are possible. I mentioned that in the early experiments when the stem was discovered, the meaning was immediately typed out at the output. Thus for example, if we took the Latin word "amat", we type out of the output the word "love". The next problem is solved by comparing the residue which is left after subtraction of the stem with the contents of a completely separate dictionary. This dictionary gives not translations of the endings, but some grammatical notes, for example in the case of "amat" it gives a note to the effect that the word "love" is in fact third person singular, present tense, active mood. Translations which result from this sort of process are difficult to read and are in many ways unsatisfactory. The essential point, however, of the stem ending procedure is that if "N" words are contained in the dictionary and each of these words can take "M" stem endings, then the number of possible combinations is "N" x "M", whereas in fact the storage required is simply "N" + "M", and this gives a considerable saving.

As the subject developed these early translations were considered less and less satisfactory, and recently Brandwood, myself and John Cleve, all at Birkbeck College, have given attention to rendering text into respectable prose. The means by which this is achieved is really quite simple. Again the stem ending procedure is invoked, but instead of the machine translating in our case, say, "amat is love", this time outputs "lov" which you will see is the English stem. The machine then, by an examination of the Latin ending, finds not a set of grammatical notes, but an appropriate ending which is to be put on to the English stem. I may mention here of course that endings are not only terminal letters, but may in fact include some prefixes. Thus in the case of "amat" the ending would be possibly "he", and the single letter "s", so that the whole thing would be "he loves". The point which arises here is in fact that the translation might be "he loves" or "she loves" or "it loves", so that in fact the machine does not in practice output "lov" at the early stage, but assembles the whole phrase in its own temporary store and then outputs the expression at a suitable time.

This idea of hoarding up the actual typing of the translation until a later stage is of course important when the more complicated problem of rearranging word order in accord with a set of grammatical rules is considered, and the simple example which I just quoted is sufficient to show that the machine is quite capable of doing this. But now one or two other points.

First I might mention the question of the resolution of ambiguity. Let us take the French word "noyeau". This may mean nut, kernel, centre or nucleus in English. Which of these meanings is the correct one would de-

pend on the context. Thus in Nuclear Physics we might encounter the word "nucleus" in the experimental sense, or the word "kernel" if we are dealing with the mathematical theory of the nucleus in considering an integral equation. On the other hand in Plant Genetics the correct translation might be "kernel", or of course "nut", and yet again in Sociology we might have "centre" or "nucleus". The idea of a microglossary which is specific to a given subject, I have already mentioned, and this removes most of the ambiguity. Those which still exist can be dealt with in two ways, either by outputing alternative meanings which we can then sort out as human beings, or alternatively, the machine can keep an account of certain non-ambiguous words which can give it a clue to the way in which the particular word must be translated in the passage under consideration. In this way the machine can select between output meanings and produce a unique one, and although there is no technical difficulty in carrying out this operation, it does of course require considerable storage facilities on the part of the machine.

Now just before I conclude this brief statement about the sort of things which a machine can do with language, let me just mention the question of idioms. An idiom is usually based on some key word and in the machine to deal with idioms, we arrange that any keyword which can form the basis of an idiom is marked by a number which characterises that particular idiom. The example I like to take is the old French chestnut "boite de nuit", which I am sure all of you know means a night club. What in fact happens here is that when the machine sees the word "boite", it finds associated with it, firstly the meaning of "box", secondly an indication that the word may form the basis of an idiom, let us suppose idiom number 27. The machine next looks up the succeeding word, having stored the meaning "box", and finds that this is "de" which has its usual translation meanings in the dictionary but also an indication that it may form a part of idiom 27. The machine then proceeds, having stored the meaning of "de", in this case "of", to the further word which is "nuit". This word contains two meanings, first of all "night" and secondly "night club". It also contains an indication that it is the end word of idiom number 27. Thus, having recognised the three words "boite de nuit", all of category 27, the machine ignores the translations of the first two and outputs the meaning "night club" from the third. If of course we had had a third word as, let us say, "savon", "savon" is not the concluding word of an idiom as far as I know and the machine would then output directly the whole expression "box of soap".

This idea of having various forms of code number associated with dictionary entities is vitally important when we come to consider a language like German. German has very little relation between the word order in the German language and the word order in English. This means that the machine has to make in fact a grammatical analysis of the German sentence, and to do this it makes use of the various code numbers indicating nouns, adjectives verbs and so on from the German text, after processing through the dictionary. When the code number word has been assembled for a complete German sentence, and we process German in sentences, this is compared with a dictionary of such code numbers which tells the machine the way in which the word order must be re-arranged, in order to produce a sensible output in the English language.

At this point, I should appreciate it if someone would either show on an epidiascope or hand round the example of French prose passage which has been translated using these sorts of ruleson the Birkbeck College machine. Those

of you who are pure mathematicians will see that the passage concerned comes from the opening of the Third volume of Goursat Cours d'Analyse Mathematique I think most of you will admit that it would be difficult to tell that the translation had been made by a machine and not by perhaps a terribly talented human being.

"L etude des fonctions definies par une equation differentielle dans tout leur domaine d existence, est un probleme dont la solution complete dans le cas general, depasse actuellement la puissance de l analyse. On a cependent obtenu des resultats du plus haut interet en se limitant a l etude des integrales infiniment voisines d une integrale connue. C est ainsi que, dans ses memorables travaux sur le probleme des trois corps, H. Poincare a pu demontrer l existence d une infinite de solutions periodiques et de solutions asymptotiques a une solution periodique."

"The study of functions defined by a differential equation in all their sphere of existence, is a problem of which the complete solution, in the general case, surpasses at present the power of analysis. One has nevertheless obtained results of highest interest while limiting oneself to the study of infinitely adjacent integrals of a known integral. It is just so that, in his memorable works on the problem of three bodies, H. Poincare has been able to demonstrate the existence of an infinity of periodic solutions and of asymptotic solutions to a periodic solution."

Now I want to continue and in a sense conclude this talk by making one or two remarks about ancillary equipment which is going to be necessary if we have a special purpose translating machine built in the future, and the first problem of course is that of input-output. Here of course it is necessary to be able to insert data into the machine and it has already been mentioned that we do this by means of a teletype apparatus. This cannot be supposed for a single moment to be an efficient means of input because it implies that any document which is presented to the machine must be re-typed before the machine can deal with it. Such a re-typing is to say the least a bit costly, and may, in the case of translations from languages which do not employ Roman script, be extremely difficult, and this leads naturally to the idea that an ideal translating machine would have as part of its equipment some means of reading directly from an original document.

There are three major methods which I can mention for a character recognition. The first makes use of a set of masks upon which are cut representations of the characters to be recognised. By means of an optical system a bright image of the character is projected on to the unknown character and only when the image and the character correspond completely is there a minimum of reflected illumination, and this can be used to indicate to the machine the nature of the character on the original document. The second class of character recognising device is based upon a logical analysis of character shape and this is normally conducted by means of the scanning pattern whose intersections with the character, lead to a set of discriminations which enable the character to be recognised. A large experiment based on these principles has been carried out by the I.B.M. Corporation in the United States, and it was reported at the I.E.E. Conference last year. The third method derives I suppose formally from the second, and it is the one which we have been working on at Birkbeck. In this the unknown character

is scanned by means of the projected image of a flying spot projection tube. Times of intersection of each line of the pattern are obtained and these are fed to the computer itself. Previous to this operation of course, the computer makes a trial run over a sample of the type found of the passage to be recognised. In this way a set of test numbers are obtained and these are stored in a form of dictionary. When the unknown character appears, its number is calculated or its set of numbers are calculated, these are compared with the dictionary entries and we would be able to show that with quite a wide degree of tolerance one gets a minimum only when exact recognition occurs. From the dictionary of course, we can now immediately output the coded form of the character for machine use.

I perhaps might mention the sorts of speed which are available. The character disc recognition, that is the obscuring of the actual character by a bright image, operates in some of the trials at about 50 characters per second, the other two methods do not appear to have any very valid upper limit and the speeds which have been spoken of are of the order of 1,000 characters per second. The real importance of character recognition as an input lies of course in its application to rather odd languages, like Russian, Chinese, Japanese, Arabic and Greek and so on, because without some form of recognition device it is very unlikely that Western computing machine people and translating machine people are ever likely to be able to process languages of this sort. Typists are extremely slow when operating in type founts other than those to which they are accustomed.

The further aspect of input is that of speech. A certain amount of work has been carried out on this branch of the art, chiefly pioneered by the Bell Laboratories in the United States. Speech recognition systems at present available do not appear to be very reliable. One varient which we have tried at Birkbeck makes use of a similar sampling technique to that mentioned in connection with the recognition of printed characters. Effectively spoken sound is applied to a microphone, amplified and turned into a voltage waveform. This voltage waveform is then passed through three filter channels of which the first effects no change; the second is a low-pass and the third is a high-pass filter. The outputs of these channels are now standardised as far as amplitude is concerned and the number of axis crossings in certain sampling intervals of time are measured by means of standard binary counters. The numbers thus obtained are fed to the computer and recognition is achieved again by comparing these numbers with those obtained by a previous trial run with the subject whose voice is to be recognised. The results of this work suggest that if care is taken to standardise the duration of each word, and if the computer works from a record made from the person who is going to use the thing, it is possible to recognise the digits 0-9 with something like one error in a thousand. The question of recognising actual words is far more difficult although some progress has been made along these lines, noticeably at University College, London by Fry and Denish, but it cannot be said at the present time that there is much hope of applying this sort of technique to machine translation.

And last but not least, it is necessary for a translation machine of course to produce an output. The outputs at present available are either on punched teletype tape or by direct imprint on a teletypewriter or on the form of electrical typing apparatus. The speeds obtainable are of the order of 7 letters per second for direct typing or 25 letters per second on punched tape. There is no reason, however, why these speeds cannot be greatly improved by the use of one of the parallel printing devices which are now obtainable. It is not beyond the realms of possibility of course

that at some future time machine output may be produced in the direct form of spoken words and in this way the dream of science fiction authors of the machine talking to its operator may sometime be realised.

Well now something about the economics of machine translation. At the present time translation is limited by the fact that the machines have an inadequate store, and this means either that a very restricted microglossaries have to be used, or alternatively that the translation cannot be made without some intermediate form of input, such as the punched paper or a magnetic tape. Both of the latter forms of input for dictionaries are very slow. With this restriction however, the speed which can be obtained from present-day machines in translating French is of the order of 1,000 to 3,000 words per hour, depending entirely on the complexity of the grammatical analysis which is required. The implications of this as far as any practicable use of translating machines is concerned is interesting. The cost of time on a computing machine is, let us say, of the order of £30 per hour. A human translator can translate a thousand words of text into extremely presentable English in this period and his remuneration would not be likely to exceed 10/- or £1, so that it is clear that at the present time the machine is uneconomical by a factor of at least thirty. This does not mean, however, that the work on translation by computing machines is worthless. It is by no means the case. Many machines are coming into operation at the present time and programmes once prepared for a given machine are readily transformable into a code suitable for any of the other machines. Machines frequently have small amounts of time available and this, because it is not normally processible by the Accounts Department, could be used for translating papers for the scientific staff of the institution housing the machine.

A further great advantage occurs when we have to translate exotic languages such as Russian, Chinese and so on, because it is rather unlikely that a given laboratory would have available personnel who are familiar with these languages or any extensive number of them.

Finally, a peep into the future, let us just see what sort of machine we could build supposing that we start now and project the completion of a translating machine in, let us say, two or three years' time. The first thing is storage. It seems to me at the moment very likely that magnetic drum storage or possibly one of the photograph disc stores being developed in America would be the ideal medium for storage for machine translation. The capacity should be between 20,000 and 60,000 words of real language. The word of "real language" in this context I would define as being an aggregate of anything between 50 and 100 letters. There does not seem to me any impossibility of constructing a store of this size at the present time and I think that the 60,000 words store should not exceed in cost, let us say, £3,000 or £4,000. I have mentioned already the fact that the computer word length would have to be considerably greater than that which obtains in most machines at the moment. If one allows not only for the foreign language word, but also for various alternatives in the target language, for grammatical notes and for idiom base numbers, then it is quite likely that a computer word might have to have as many as 250 binary digits. As far as arithmetic operations are concerned, the machine really needs only subtract, it is probably an advantage that it should be able to add, but it is certainly quite unnecessary for the machine to have facilities for multiplication or any of the more exotic arithmetical operations. I do perhaps make a plea for one or two special operations, for example, a conditional

transfer on groups of digits within a word. This would be extremely useful in the recognition programme. Further, a special instruction which would enable a word to be hunted out in a stem dictionary in a single order; and finally, a rather exotic instruction which would enable the position of a word in a sequence of words to be ascertained, the words to be interposed or inter-polated into that sequence and all of the following words which were already present there to be displaced up one stage at a time. This would be extremely valuable in re-arrangement programmes. The input to the machine which we are considering will probably be in the form of a device for the recognition of printed characters. Solartron in England and International Business Machines in America have already gone a great way to producing such a device on a commercial scale, although I regret to say that the sort of cost involved appears to be about £20,000. The output of the machine could either be on to magnetic tape, which would be fast, or alternatively, directly on to a parallel printer which might be capable of, let us say, a thousand lines of type per minute.

Finally, what about the speed? Well, using the existing techniques, and envisaging the machine which as I say we should be able to finish in two years to be quite simple to construct a device with all of the characteristics except that of character recognition, and possibly the spoken word input. It seems to me that the speed which I have mentioned, 1,000 to 3,000 words per hour could be increased by at least a factor of ten, and in this case our cost would be reduced to anything between 6/- and £3 per thousand words, all of which is eminently reasonable.

Well I hope that in this short talk I have been able to show you some of the possibilities of machine translation, and some of the techniques which we have developed. This is a rapidly growing subject, more and more workers are coming into the field, and I anticipate that in the next five years we shall see revolutions in technique.