FIRST STEPS IN MECHANICAL TRANSLATION

John Hutchins
(University of East Anglia, Norwich, UK)

Abstract

Although the first ideas for mechanical translation were made in the seventeenth century, it was not until this century that means became available for realisation with the appearance of the electronic computer in the mid 1940s. Fifty years ago, in March 1947 Warren Weaver wrote to Norbert Wiener and met Andrew Booth, mentioning to both the use of computers for translation. The possibilities were investigated during the next seven years, until in January 1954 the first prototype program was demonstrated. This article is a chronicle of these early years of mechanised translation.

1. The beginnings to 1947

The idea of mechanising translation can be traced back to the seventeenth century, almost always in association with the idea of inventing a universal language. In the previous century Francis Bacon had stressed the inadequacies of natural languages for the description of scientific discoveries. By the early seventeenth century the need was more widely felt, and the rapid growth of international trade throughout the world and the decline of knowledge of Latin made the dream of a universal language even more desirable. As a result many proposals for 'universal characters' were put forward; the best known being that of John Wilkins in 1668 for the newly founded Royal Society of London. These schemes for universal characters or 'common writing' were invariably furnished with dictionaries from a vernacular language or from Latin, which were intended to be applied by users in a 'mechanical' fashion. It was in this spirit that Descartes made a comment in a letter to Marin Mersenne on 20th November 1629. It may be freely translated:

> Thus the sole utility I can see arising from this invention is for writing: if a large dictionary is printed in all the languages in which one wants to be understood, where common characters are put for each primitive word corresponding to the sense and not to the syllables, such as the same character for *aymer*, *amare* and *filein*; [so that] those having the dictionary and knowing the grammar would be able, through seeking all these characters one after another, to understand in their language what is being written.

Descartes' comment can be read as a prefiguration of how an interlingual 'mechanical' dictionary might work. In this sense, the glossaries attached to universal languages might be seen as forerunners of 'mechanical dictionaries'. This was indeed been the case with a scheme published in 1661 by a German chemist Johann Joachim Becher. At the height of enthusiasm for MT in the 1960s, his work was re-published as an form of pre-computer program for machine translation (Becher 1962).

However, as the century progressed, the search for an ideal language of communication -- a goal pursued by many famous philosophers, notably by Leibniz -- was recognised to be unrealistic. The search was virtually abandoned (apart from a brief revival in the decade after the French Revolution) until the last years of the nineteenth century. It is then that there first appeared various schemes for international languages, the best known being of course Esperanto.

However, it was not until the twentieth century that the means became available to develop practical devices for mechanising translation. In 1933 two inventions were patented independently. One was in France by Georges Artsrouni, the other in Russia by Petr Troyanskii. In both cases, the patents were essentially electrical devices for mechanical dictionaries (Hutchins 1993). Artsrouni's invention was exhibited in 1937 at the World Exhibition in Paris, and attracted much attention. Apparently under development since 1929 Artsrouni's invention was intended for a wide variety of tasks - the production of timetables and telephone books, for accounting, and for deciphering and encrypting messages, as well as for translating. In essence, the device comprised four components: a 'memory' of dictionary entries on a paper tape, a keyboard for entering words, a search mechanism, and a means of displaying results in a series of windows on the keyboard.

While Artsrouni's invention was no more than a mechanical dictionary, Troyanskii went further. His devise consisted of a mechanically operated table containing a multilingual 'glossary field' by which words of the source language were presented through an aperture and equivalent words of a target language were printed out. The words appeared not in full forms but as stems, and Troyanskii devised a set of 'logical symbols' common to all languages for expressing grammatical relationships. These symbolic characters were clearly based on Esperanto. For example, the suffix -aj indicated a plural noun, -n was used to express a direct object case, -ir a verb form, etc. Troyanskii envisaged translation as a three-stage process. A human 'editor' would prepare the source text by reducing words to stems and by identifying the relevant 'logical' relationships. The prepared text would then be converted wholly mechanically by his device into an equivalent sequence of words in a target language, while retaining the logical symbols. A human 'editor' would then convert the string of stems and symbols into the target language.

In the years following his 1933 patent, Troyanskii worked on various technical improvements and at some point suggested that the processes of analysis and generation might also be mechanised. Troyanskii clearly underestimated the amount of work involved in analysing texts into stems and logical symbols, or in generating texts from them. He was aware that homonyms and lexical differences between languages had to be dealt with, but he seems to have been quite unaware of the problems of syntactic restructuring. Nevertheless, there is no doubt that had he lived to see the first computers, Troyanskii would have been one of the pioneers of machine translation.

2. March 1947

These precursors were, however, quite unknown in the mid 1940s, when the first electronic computers were being developed and when the first suggestions were made to apply their formidable powers to the task of translating languages. (A fuller account of the MT pioneer years is given in Hutchins, 1997), where details of all the sources cited here will also be found.)

We can therefore date the true beginning of machine translation (i.e. the application of computers to translation) as the 4th of March 1947. On this day, Warren Weaver, director of the Natural Sciences Division of the Rockefeller Foundation, wrote to the cyberneticist Norbert Wiener:

> Recognizing fully, even though necessarily vaguely, the semantic difficulties because of multiple meanings, etc., I have wondered if it were unthinkable to design a computer which would translate. Even if it would translate only scientific material

> (where the semantic difficulties are very notably less), and even if it did produce an inelegant (but intelligible) result, it would seem to me worth while.

Weaver, who had heard about the success of computers in cryptography, wondered

> if the problem of translation could conceivably be treated as a problem in cryptography. When I look at an article in Russian, I say 'This is really written in English, but it has been coded in some strange symbols. I will now proceed to decode.'

Wiener's reply was disappointing, arguing that the problems of semantics were too great for successful mechanisation.

However, just two days after writing to Wiener, Weaver met Andrew Booth, a British crystallographer, who was in the United States investigating the newly invented computers. He had come to Weaver to discuss possible American funding for a British computer at the University of London. Weaver did not think there would be any funds for numerical applications but he thought that there might be an interest in supporting new non-numerical applications of computers, and he suggested translation.

The idea of MT had come to Weaver some time earlier, quite possibly in 1945, and he may well have spoken to people about it. One of these was probably J.Desmond Bernal, who was head of Booth's department at Birkbeck College and was a regular visitor to the US and the Rockefeller Foundation.

At a later date, Booth stated that the discussion about MT took place in June 1946 (Booth and Locke 1955) - and this statement has been taken as accurate by many later writers. It is certainly true that he met Weaver on June 1946 and that there was discussion of the funding of a British computer. However, there is no documentary evidence that Weaver raised the idea of MT on this occasion. Indeed, Booth himself stated in other papers of the time that the meeting had been in March 1947.

3. Booth and Richens

Weaver's discussion with Booth bore fruit. On his return to England, Booth did not forget about MT. In his report to the Rockefeller Foundation in February 1948 he included MT in a list of possible applications for the proposed US-funded computer at the University of London:

> ...We have considered this problem in some detail and it transpires that a machine of the type envisaged could perform this function without any modification in its design.

He was probably referring to the programming of a mechanical dictionary, which he had begun with Kathleen Britten (his future wife) during 1947 while at the Institute for Advanced Study in Princeton. However, Booth had subsequently met Richard Richens.

Richens was Assistant Director of the Commonwealth Bureau of Plant Breeding and Genetics in Cambridge. His interest in mechanical translation had arisen out of experiments with punched cards for storing information:

> The idea of using punched cards for automatic translation arose as a spin-off, fuelled by my realisation as editor of an abstract journal (*Plant Breeding Abstracts*) that linguists conversant with the grammar of a foreign language and ignorant of the subject matter provided much worse translations than scientists conversant with the subject matter but hazy about the grammar

He proposed the segmenting of words into their base forms (stems) and endings (e.g. inflections), both to reduce the size of dictionaries and to use grammatical annotations to augment a strictly word-by-word dictionary 'translation'.

The validity of Richens' method was tested by hand and by using punched card machinery on a variety of languages. It was not apparently implemented as a computer program, primarily because the authors did not have access to a computer of sufficient size and reliability. The experiment was written up in a memorandum during 1948, although not published until much later, in a revised version (Richens and Booth 1955). From this account, the crudity of the output is obvious:

> (French) Il n'est pas étonn*ant de constat*er que les hormone*s de croissance ag*issent sur certain*es espèce*s, alors qu'elles sont in*opér*antes sur d'autre*s, si l'on song*e à la grand*e spécificité des ces substance*s.
>
> (English) *v* not is not/step astonish *v* of establish *v* that/which? *v* hormone *m* of growth act *m* on certain *m* species *m*, then that/which? *v* not operate *m* on of other *m* if *v* one dream/consider *z* to *v* great *v* specificity of those substance *m*.

(The asterisks in the French indicate automatic segmentations. in the English 'translation' *v* indicates a French word not translated, *m* "multiple, plural or dual", *z* "unspecific", and slashes alternative translations.)

Although, only the very first stage of a true translation system, Richens is rightly credited as being the first person to propose a method of automatically identifying word-endings, albeit not using a computer but a punched card machine.

4. Developments in Los Angeles

By this time there were signs of the start of research in the United States. In May 1948 the National Bureau of Standards (NBS) decided to build computers in Washington and in Los Angeles at its Institute of Numerical Analysis. Harry D.Huskey, who had worked for a year in England at the National Physical Laboratory (NPL), led the team building the computer which was eventually called the Standards Western Automatic Computer (SWAC). Either independently or perhaps from conversations with Turing at NPL, Huskey proposed to use SWAC for translation. In May 1949, the machine was demonstrated and a report appeared in the New York Times, the first newspaper article on MT:

> A new type of "electric brain" calculating machine capable not only of performing complex mathematical problems but even of translating foreign languages, is under construction here... While the exact scope of the machine will have in the translating field has not been decided, the scientists working on it say it would be quite possible to make it encompass the 60,000 words of the Webster Collegiate Dictionary with equivalents for each word in as many as three foreign languages.

After describing how the computers of the time operated, the report continued:

> When a foreign word for translation is fed into the machine, in the form of an electro-mathematical symbol or a tape or card, the machine will run through its "memory" and if it finds that symbol as record, will automatically emit a predetermined equivalent - the English word.

The reporter recognised the limitations of "a crude word-for-word translation" but saw the potential value as "scientists' translations of foreign technical papers". An editorial the next day (one of the few ever devoted to MT) expressed doubts about the accuracy:

> How is the machine to decide if the French word "pont" is to be translated as "bridge" or "deck" or to know that "operation" in German means a surgical operation? All the machine can do is to simplify the task of looking up words in a dictionary and setting down their English equivalents on a tape, so that the translator still has to frame the proper sentences and give the words their contextual meaning.

Quite wisely, it commented that:

> We are still far from the machine into which we will pour cards and pull out great poetry or great novels. In fact we shall never reach that stage.

Nevertheless, the editor shared the common assumption that the awesome mathematical power of computers made them capable of anything.

5. Weaver's memorandum, July 1949

In July 1949, two years after writing to Wiener and meeting Booth, Weaver wrote the memorandum (Weaver 1949) which was to launch MT as a serious subject of research in the United States, and subsequently throughout the world. His aim was to summarise the work which had already been done and to outline the directions he thought research could take.

He begins by describing the origins of his own interest in the topic. First he recounts a story about decoding Turkish, which had succeeded without the decipherer having any knowledge of the original language. The success of cryptography lay, he believed, on "frequencies of letters, letter combinations, intervals between letters and letter combinations, letter patterns, etc., *which are to some significant degree independent of the language used*"; and he supposed, therefore, that:

> there are certain invariant properties which are, again not precisely but to some statistically useful degree, common to all languages.

In support he cites the findings of logicians such as Hans Reichenbach and linguists such as Erwin Reifler; the latter having noted that "the Chinese words for 'to shoot' and 'to dismiss' show a remarkable phonological and graphic agreement" -- a range of meaning found also in English 'fire', which Weaver thought was more than a coincidence.

After reproducing his letter to Norbert Wiener and Wiener's disappointing reply, Weaver reports what had been done so far. He had visited England and heard about Richens' work with Booth; and he outlines Richens' ideas on suffix-cutting. He mentions also the newspaper reports of activity in California, although strangely without saying where it was undertaken or who was directing the work.

There had been no actual translation by computer -- the work of Booth and Richens was simulated on punched cards only, and from California it was merely a proposal to apply a computer still under construction. Weaver's main concern was to indicate how MT might go beyond the limitations of word-for-word translation, and he proposed, as he puts it, "at the risk of being foolishly naïve... four types of attack".

The first approach was to tackle multiple meanings by examination of immediate context. How much context would be required could vary from subject to subject, and from one word to another, and he thought an investigation would be useful. His second proposal assumed that there are (at least in non-literary works) logical elements in language. He believed that MT could be based on the theorem proved in 1943 by McCulloch and Pitts on logical deduction of proofs

by automata, which suggested that "insofar as written language is an expression of logical character" the problem of translation by computer is formally solvable.

It was linked to a third proposal involving the recently developed 'information theory' of Claude Shannon, which was concerned with the basic statistical properties of communication. Although enthusiastic for this approach, Weaver did stress the fundamentally probabilistic nature of language, and the inevitable imperfection of any translation:

> "Perfect" translation is almost surely unattainable. Processes, which at stated confidence levels will produce a translation which contains only X per cent "error," are almost certainly attainable.

But in his fourth proposal, he was more utopian. The "most promising approach of all" is to investigate the fundamental structures of languages:

> Think, by analogy, of individuals living in a series of tall closed towers, all erected over a common foundation. When they try to communicate with one another, they shout back and forth, each from his own closed tower. It is difficult to make the sound penetrate even the nearest towers, and communication proceeds very poorly indeed. But, when an individual goes down his tower, he finds himself in a great open basement, common to all the towers. Here he establishes easy and useful communication with the persons who have also descended from their towers.
>
> Thus it may be true that the way to translate from Chinese to Arabic, or from Russian to Portuguese, is not to attempt the direct route, shouting from tower to tower. Perhaps the way is to descend, from each language, down to the common base of human communication -- the real but as yet undiscovered universal language -- and then re-emerge by whatever particular route is convenient.

He readily admits that this approach involves a "tremendous amount of work in the logical structures of languages before one would be ready for any mechanization", although he believes that Ogden and Richard's work on Basic English represented some steps towards it.

6. Contemporary proposals

While we can credit Weaver with the most explicit proposals about MT during the early years of the computer, it is quite clear that the idea of mechanised translation was an idea that had come to others at this time.

There are a number of instances where proposals for translation by computer were put forward independently. In 1946, Duncan Harkin had been inspired by studies of word counts to think of a system based on the most frequent vocabulary, and in collaboration with an electronics engineer, planned the outline of a machine for mechanical translation. In September 1947 Alan Turing wrote a report for the National Physical Laboratory about his plans for constructing a computer, in which he mentioned MT as a possible demonstration of the 'intelligence' of computers. In August 1949, the *New York Times* reported from Salerno that an Italian named Federico Pucci, had invented a machine to translate, saying that it would be exhibited at a Paris Fair; but no more was to be heard of it. In February 1950 Charles C.Holt of Chicago revealed in a letter to the Rockefeller Foundation that he had thought of using cathode ray tubes as a storage device for a mechanical dictionary -- in some respects, it was similar in concept to the mechanism put forward by Troyanskii in the 1930s.

There may well have been others. The new electronic computers had caught the imagination of many people. Reports on the 'electronic brains' -- the term regularly used by

journalists -- appeared almost daily in national newspapers throughout the world. Using them for translation occurred to a number of people.

7. Reactions to the memorandum

Some recipients of Weaver's memorandum were inspired to take up the challenges of MT immediately. First on the scene was the sinologist mentioned in the memorandum itself, Erwin Reifler. Already by 10th January 1950, Reifler produced a 55 page study, in which he put forward the first formulated conceptions of pre- and post-editing. Assuming that 'mechanical' processes could only be word for word one-to-one substitutions, there were obvious inadequacies. Reifler suggested that in order to remove ambiguities from the source text a human "pre-editor" could add extra symbols for grammatical and logical explicitness. Theoretically a pre-editor need not know the output language. The task of the post-editor was to render the machine output into a reasonably literate form -- and, theoretically too, the post-editor need not know the source language. Reifler argued that without knowledge of the source language a post-editor would have great difficulty faced with lists of alternatives for each word; the ambiguities of the original could be more easily resolved by the pre-editor who would know the linguistic and cultural context. Reifler envisaged that the pre-editor would have access to a monolingual dictionary presenting all the alternative interpretations for those words with more than one possible translation in the target language.

Weaver's memorandum had stimulated interest also at the Rand Corporation in Santa Monica. Following up Weaver's suggested statistical approach to resolving problems of multiple meaning, Abraham Kaplan investigated the micro contexts of polysemy in mathematical texts. His study was completed on 30 November 1950, concluding that the "most practical context is ... one word on each side, increased to two if one of the context words is a particle", i.e. an article, preposition or conjunction. Despite its limitations and the tentativeness of its conclusions, this study was to encourage hopes that problems of ambiguity could be resolved, and that statistical analyses could contribute useful data for MT systems.

Also in California, 1950 saw the beginning of research at UCLA, initiated by Huskey. In the Department of Spanish, William E.Bull began his statistical studies of vocabulary, and in the Department of German Victor A.Oswald, Jr. researched syntactic problems in collaboration with Stuart L.Fletcher of the NBS. This was the first research in MT devoted to syntactic questions, and it resulted in the first article on MT to be published in an academic journal (Oswald and Fletcher 1951). The authors proposed the coding of German sentences to identify 'noun blocks' and 'verb blocks' and to determine which blocks were candidates for rearrangement after translation into English. For example, using a passage from a mathematics text by Cantor:

> Bevor wir diese Definition im Einzelnen zergliedern, wollen wir einige Beispiele von Mengen betrachten, die uns anschauliches Material zum Verständnis der Definition liefern sollen.

This sentence was to be rearranged, on the basis of identified 'blocks', as:

> Bevor wir zergliedern diese Definition im Einzelnen, wir wollen betrachten einige Beispiele von Mengen, die sollen liefern uns anschauliches Material zum Verständnis der Definition.

In this sequence, the German could be given a potential word-for-word translation into English as:

> Before we analyze this definition in detail we want-to regard some examples of sets, which shall furnish us perceptible material for-the understanding of-the definition.

Their conclusion was that syntax "does not constitute, as had been thought by some, a barrier to mechanical translations." it was a rather optimistic conclusion, but, we should remember, it was not based on any test by a computer; although the procedures had been formulated as 'instructions' for the SWAC, they were not implemented.

8. Bar-Hillel's survey 1951

In May 1951, Yehoshua Bar-Hillel was appointed as a research assistant in the Research Laboratory for Electronics at the Massachusetts Institute of Technology (MIT). His task was to study the possibilities of MT and to plan future research, and in October 1951 he visited the few places in the US where MT research was going on. After the tour it was decided to hold a conference at MIT in June 1952, and in preparation for it Bar-Hillel wrote a state-of-the-art report which represents a milestone in the development of MT.

His survey (Bar-Hillel 1951) begins by summarizing the presumed benefits of MT to satisfy translation demands, particularly in science, finance and diplomacy; and to provide means for scanning at high speed ("though perhaps low-accuracy") the huge printed output in newspapers, journals, leaflets of actual or potential enemies. As a side-product MT could also throw light on aspects of linguistics and communication.

Although he could envisage no hope of achieving "high-accuracy, fully automatic MT... in the foreseeable future", he stressed that this did not mean computers could not be applied to translation: "with a lowering of the target, there appear less ambitious aims the achievement of which is still theoretically and practically viable." Therefore, Bar-Hillel reviewed various options for "mixed MT, i.e. a translation process in which a human brain intervenes."

A post-editor was considered indispensable to eliminate "semantical ambiguities", but he stressed the importance of not printing out all possible translations of every word in the sentence order of the original. What was clearly needed was for the grammatical ambiguities to be resolved automatically, and he made suggestions for an "operational syntax". He did not agree with Oswald and Fletcher that syntactic analysis must await the availability of complete statistical data for morphological analysis in order to overcome the limitations of computer storage. He was already profoundly sceptical of statistical approaches -- a view which he was to repeat in subsequent years. Statistics might well identify the most frequent words and endings and enable 90 per cent of an average text to be translated, but the result would be unsatisfactory since the post-editor would be faced with translating the words which are "the least predictable and highly loaded with information".

He then briefly considered the role of a pre-editor as envisaged by Erwin Reifler. The problem with Reifler's proposals was the construction of a monolingual dictionary which the pre-editor could use to select appropriate translation equivalents in an unknown target language. Bar-Hillel agreed that it might be possible where only two languages are concerned, but not for "*general* MT, where translation from any language into any other is considered."

Whereas, in Bar-Hillel's view, "*specific* MT" could be pursued on a simple trial-and-error basis, "general MT will require establishment of a *universal*, or at least *general grammar*, perhaps even the construction of a whole artificial exchange-language." He was aware of failed

attempts in previous centuries to construct universal languages, but believed that "mathematical logic, and modern structural linguistics" may prove a better foundation.

In the final pages of his survey, he mentions the possibility of MT of texts with "a restricted vocabulary or a restricted number of sentence-patterns." Here he had in mind Basic English, artificial languages such as Esperanto, and the codes used by pilots and meteorologists. He thought it might be possible in such cases to translate whole units or sentences, and "the theoretical difficulties... are clearly less formidable". But he went further: there was also the "possibility of restricting, by voluntary convention, the richness of expression", i.e. by establishing what are now known as 'controlled languages'.

Bar-Hillel's paper was written before any MT research on even a reasonably modest scale had begun, in particular before any computer had been programmed to do even word-for-word translation, and it is therefore remarkable that he was able to identify so clearly the main problem areas and to outline many of the basic strategies which continue to be valid to the present day.

9. The first MT conference, June 1952

Bar-Hillel organised the first conference on machine translation at the Massachusetts Institute of Technology from the 17th to the 20th June 1952. Reports of this milestone event are to be found in papers by Reynolds (1954) and Reifler (1954).

At its opening public session Bar-Hillel outlined the history of MT, from Weaver's first ideas to the establishment of the research groups represented at the conference. After pointing out the obvious potential of MT to cope with the immense and growing volume of materials to be translated, Bar-Hillel went on to stress the complexities of mechanising translation processes, concluding that completely automatic and autonomous mechanical translation with unique correlates to the original text is, in general, practically excluded, even with respect to scientific texts... This being so, machine translation means no more than *mechanical aids to translation*. Only some kind of *brain-machine partnership* is envisaged.

As in his 1951 review, Bar-Hillel mentioned the partnerships of post-editors and pre-editors. He conceded that economically MT was not yet viable:

> Even if it should turn out that none of the possible machine-brain partnerships would be more effective than a human translator, in the sense that they will be neither quicker nor cheaper nor more exact than the human translator, under conditions existing today or in the near future, I would strongly advocate a continuation of this research. Electronic machines will doubtless become cheaper, human brains probably more expensive. A partnership that could not stand free competition today may well outbid its human competitors in some not too remote future.

All participants agreed on the need for post-editing. At the conference, Bar-Hillel repeated what he had said on the topic in his survey. Then in another paper he argued that post-editors should be given every possible version. This paper, which was concerned with idioms, suggested that 'idiomatic' phrases should either be included as units in the lexicon or that post-editors should be alerted to the fact that certain output forms might be replaced by phrases. His example was *Es gibt einen Unterschied*, where the correct translation should be *There is a difference*. Inclusion of *es gibt* as a phrase in the dictionary with the automatic translation *there is/are* would prevent a literal translation when required, e.g. *es* might refer to *Mädchen*, in which case *she gives...* would

be correct. He argued therefore that post-editors should be presented with all the alternatives. In fact, Bar-Hillel was of the opinion that if MT is limited to Western languages closely related to English and to scientific publications, and if the dictionary is limited to 'non-overlapping synonyms' then multiple meanings can all be dealt with satisfactorily by a post-editor.

Many agreed with Reifler that pre-editing of some kind was also essential to minimise ambiguities and syntactic complexities of source texts. Alternatively, it was thought that authors could be trained to write with MT in mind, e.g. by using a 'controlled language'. At the conference, Stuart Dodd of the Washington Public Opinion Laboratory presented his ideas for the "standardization of English syntax as a means of simplifying the use of English either as a source language or as a target language". This included the regularisation of verb forms (*She did be loved* instead of *She was loved*) and of pronouns (using only nominative forms: *I will send he to she*) and, of course, the use of words in one meaning only.

Reifler elaborated on the pre-editing codes. These were to be "a graphic supplementation of the conventional form of the foreign message which raises its graphic-semantic explicitness to the level necessary for a mechanical translation." The coding would take into account not just the multiple meanings in the source language but also lexical and semantic differences in the target language. He thought the codes could be inserted by writers themselves using special monolingual dictionaries with symbols for distinguishing homographs. But he had also a more radical proposal: a new orthography for all languages which would distinguish grammatical categories: "all nouns would have... a capital first letter..., all principal verbs with a capital second letter and all attributive adjectives with a capital third letter..."; so that for example, the German *er hegt die fromme Hoffnung* would be written "er hEgt die frOmme Hoffnung." The idea was that this would ease the specification of the context in which Kaplan's method of disambiguation could operate. Furthermore, in conjunction with a regularisation of the target 'model' language, it would "either restrict post-editorial interpretation to a minimum, or it may even make it completely superfluous."

For translation into many target languages Reifler agreed with Bar-Hillel that this raised the question of a 'universal grammar'. He believed that comparative-historical linguistics could help to identify real universals, but there may also be 'pseudo-universals'. He thought that words of source languages could be assigned grammatical categories arbitrarily aligned to categories of target languages. For example, the Mandarin Chinese version of English *he walks quickly* is *t'a1 tsou^{3}-ti k'uai^{4}*, where *k'uai^{4}* corresponds to something like "to be quickness" or "to be quick" and *tsou^{3}-ti* corresponds to "walk's" or "of walk". But in other contexts *tsou^{3}-ti* can be translated as "walking"; if there were an arbitrary equation of *-ti* and English *-ing*, the sentence could be rendered as "he walk-ing quick" -- a kind of pidgin English.

Although Weaver had suggested investigations of universal languages for MT, the topic was not discussed in the conference. For translation into many languages, where now we would expect proposals for some kind of interlingua, there was instead a suggestion from Leon Dostert of Georgetown University that:

> general MT (mechanical translation from one into *many* languages)... should be so developed that one translates first from the input language into one 'pivot' language (which in our case will, most likely, be English) and from that pivot language into any one of the output languages desired.

The conference was attracted by the ideas of Victor Oswald for domain-specific dictionaries ('microglossaries' he called them) to minimise the problems of "multiple meanings". The 'sublanguage' vocabulary (as it would now be called) was to be identified by statistical analysis of corpora on the basis of the observation that "the data of all frequency counts fall into the same pattern, which means that a frequency count of any micro-segment of any language -- say the nouns in German contexts pertaining to brain surgery -- should give a parabolic curve where high-frequency elements ought to dispose of eighty-percent of all running nouns." Not only was this found to be true, and that familiarity with 80% of the technical words for any article was alone enough to make sense of the article, but Oswald found a similar frequency distribution for the non-technical words.

William E.Bull was sceptical about the value of frequency analyses for constructing micro-glossaries:

> There exists no scientific method of establishing a limited vocabulary which will translate any predictable percentage of the content (not the volume) of heterogeneous material... A micro-vocabulary appears feasible only if one is dealing with a micro-subject, a field in which the number of objective entities and the number of possible actions are extremely limited. The number of such fields is, probably, insignificant...

Indeed, Bull went on to pinpoint a basic problem for all MT systems to the present day:

> The limitations of machine translation which we must face are, vocabularywise, the inadequacy of a closed and rigid system operating as the medium of translation with an ever-expanding, open continuum.

Syntax was something quite new for many participants. Victor Oswald spoke about his analysis of German syntax: his 'blocks' approach was in essence Oswald a form of constituency analysis, which was becoming familiar at the time. More of a revelation was Bar-Hillel's 'operational syntax'. In place of Oswald's method of restructuring and conversion, he put forward his categorial grammar based on the work of the Polish logician Ajdukiewicz.

At this conference, there was virtually no practical experience of programming to be reported, or indeed of using computers. Nevertheless, their was enthusiastic talk of future research plans. Leon Dostert suggested "the early creation of a pilot machine or of pilot machines proving to the world not only the possibility but also the practicality of MT". In the realisation of this pilot in January 1954, Dostert made the transition from speculation about 'mechanical translation' to the reality of 'machine translation'.

10. Further progress in 1952 and 1953

Before this historic demonstration, however, Bar-Hillel, Reifler and others published further articles giving MT a wider audience. Reifler produced further studies on pre-editing and post-editing, on the treatment of German compound nouns and methods for dealing with ambiguity of homographs within sentences. Also in September 1952 James Perry at MIT reported a simulation of word-for-word translation of a Russian text into English (Perry 1955). The Russian words were written on separate slips of paper, drawn at random, checked in a dictionary and English translations found, and then restored to their original order. One simulation produced:

> On/Onto/At Fig.12 traced/mapped-out/drawn parabola according-to/along/in-accord-with which move thrown/deserted with/from velocity 10m/sec. under/below angle to/toward vertical line into/in/at 15^{o}, 30^{o}, 45^{o}, 60^{o}.

It was claimed that "the rough translations exhibited a high degree of intelligibility." To prove it, two assistants with training in physics and chemistry wrote interpretations, e.g. for the above:

> On Fig.12 a parabola is drawn according to which a body moves, thrown with the velocity of 10m/sec and making angles of 15^{o}, 30^{o}, 45^{o}, 60^{o} with the vertical line.

Recognising possible improvements if Russian grammar were taken into account, Perry did a follow-up simulation of a dictionary look-up procedure, which searched for the longest stem match and then identified the ending and its grammatical function.

There continued to be reports of the SWAC computer in Los Angeles, which implied that MT was close to realisation. For example, there was another report in the *New York Times* on 5th October 1952, which was as optimistic as in 1949; and in December 1952 there was an article in the scientific journal *Industrial and Engineering Chemistry* stating that "the next scientific bottleneck to open up before the mental capacity of electronic equipment may well be the problem of language barriers in international communication." Then it reported Perry's experiments, quoting a simulated translation which the reporter thought was "not unmeaningful to an individual having some knowledge of the science." A major impediment was the painfully slow input of the documents before translation, and a photographic scanner was suggested. In all, the report was a reasonably balanced assessment of the future potential of MT. However, the reporter could not refrain from an amusing coda, alluding to the current Cold War witchhunting mentality in parts of the United States:

> One wonders, though, of the outcome, should one of our more ardent and less responsible hunters of subversives get wind of these machines that may blithely swallow the words of an alien and repugnant tongue, and then spew forth what is possibly a strange and, therefore, suspect gibberish. Could any culpable and innocuous computor withstand the taunts and accusations leveled, perhaps, in its direction without manifesting an incoherent hum, suffering a twitch, or developing at least a mild paranoia? Our machines may now also need a built-in psychoanalytical circuit and couch for purposes of self-confession when mental doubts becloud their electronic tubes and snarl their recording devices.

He was not to know that some 20 years later, Colby was to write a computer program PARRY, which simulated the behaviour of a paranoid patient (Colby 1975)

During 1953, Bar-Hillel published three articles on various linguistic aspects of MT, and two articles appeared from the UCLA researchers in the same issue of *Modern Language Forum*. One was by Victor Oswald and Richard Lawson on microglossaries, the other was by Kenneth Harper, on the importance of morphological analysis when translating Russian. In the same year also, Anthony Oettinger from Harvard University produced a progress report on his doctorate study for the automation of a Russian-English dictionary; his thesis was to be approved in 1954 (the first in the field of MT).

Bar-Hillel ended his two-year research post at MIT in July 1953 and returned to Jerusalem. He was replaced by Victor Yngve, who proceeded to set up the MT research project in the Research Laboratory of Electronics. The first progress report appeared in October. Yngve (1955) described a simulated word-for-word translation from German into English, in which articles and other function words were left untranslated. He reproduced an extract:

> Die CONVINCINGe CRITIQUE des CLASSICALen IDEA-OF-PROBABILITY IS eine der REMARKABLEen WORKS des AUTHORs. Er HAS BOTHen LAWe der GREATen NUMBERen ein DOUBLEes TO SHOWen: (1) wie sie IN seinem SYSTEM TO INTERPRETen ARE, (2) THAT sie THROUGH THISe INTERPRETATION NOT den CHARACTER von NOT-TRIVIALen DEMONSTRABLE PRPOSITIONen LOSEen. CORRESPONDS der EMPLOYEDen TROUBLE? I AM NOT SAFE, THAT es dem AUTHOR SUCCEEDED IS, den FIRSTen POINT so IN CLEARNESS TO SETen, THAT ALSO der UNEDUCATED READER WITH dem DESIRABLEen DEGREE-OF-EXACTNESS INFORMS wird.

The intention was to test what could be achieved from such crude beginnings. Yngve was surprised to find how much was comprehensible, and concluded that since "word-for-word translations are surprisingly good, it seems reasonable to accept a word-for-word translation as a first approximation and then see what can be done to improve it." The most obvious need was some syntactic analysis, and this aspect of MT was to be the principal focus of research at MIT in subsequent years.

11. The Georgetown-IBM demonstration, January 1954

The most significant outcome of the MIT conference in June 1952 was undoubtedly the decision by Leon Dostert to start work on a computer program to demonstrate the feasibility of MT. At Georgetown University he appointed Paul Garvin, a linguist with knowledge of Russian, and established links with the Cuthbert Hurd and Peter Sheridan of the IBM Corporation. A small-scale system for translating some Russian sentences into English was developed and on 7th January 1954 a demonstration took place at the New York headquarters of IBM.

The next day, the *New York Times* carried a front-page report of what it believed to be "the cumulation of centuries of search by scholars for 'a mechanical translator'." Although stressing the limited nature of the demonstration, the reporter retailed Dostert's optimism that

> Such a device should be ready within three to five years, when the Georgetown scholars believe they can complete the "literary" end of the system.

It was thought that other languages would be easier than Russian, and the dictionary data (i.e. the punched cards) for German, French and "other Slavic, Germanic and Romance languages can be set up at will."

Reports appeared in many other newspapers. Nearly all stressed the speed of the operation and the fact that the punched card operator did not know a word of Russian. Spectators were also impressed by the wide range of subject matter translated, and that the output was so fluent.

The limitations of the experiment (a 250-word vocabulary and six grammar rules) were freely admitted by the Georgetown researchers, but the general public had been impressed: MT was now seen as a feasible objective, US government agencies were encouraged to support research on a large scale for the next decade, and MT groups were established in other countries, notably in the USSR. On the other hand, unrealistic expectations had been raised which did not materialise for many years to come.

The IBM-Georgetown demonstration was the first actual implementation of a translation system performing on the still primitive electronic computer. All previous work on MT had been theoretical in the sense that none of the proposals had actually been implemented as computer

programs. The SWAC demonstrations of 1949 and 1952 had not produced translations; Huskey and his colleagues had merely asserted that MT was possible and shown the machine on which they believed it could be done. All other previous research in the 1950s had been either 'thought-experiments' or simulations on punched cards or paper slips. Before the demonstration in January 1954, the mechanisation of translation had been largely speculative. From now on, real research began.

12. References

Bar-Hillel, Y. (1951): The present state of research on mechanical translation. *American Documentation* 2 (4), 229-237.

Becher, J.J. (1962): *Zur mechanischen Sprachübersetzung, ein Programmierungsversuch aus dem Jahre 1661*. J.J.Becher: Allgemeine Verschlüsselung der Sprachen (Character, pro Notitia Linguarum Universali). Stuttgart: Kohlhammer, 1962.

Booth, A.D. and Locke, W.N. (1955): Historical introduction. In: Locke and Booth (1955), 1-14.

Colby, K. (1975): *Artificial paranoia*. New York: Pergamon Press.

Hutchins, W.J. (1993): The first MT patents. *MT News International* 5, 14-16.

Hutchins, W.J. (1994): The Georgetown-IBM demonstration, 7th January 1954. *MT News International* 8, 15-18.

Hutchins,W.J. (1997): From first conception to first demonstration: the nascent years of machine translation, 1947-1954. A chronology. *Machine Translation* (to appear)

Locke, W.N. and Booth, A.D. (1955), eds.: *Machine translation of languages: fourteen essays*. Cambridge, Mass.: M.I.T.Press.

Oswald, V.A. and Fletcher, S.L. (1951): Proposals for the mechanical resolution of German syntax patterns. *Modern Language Forum* 36 (3/4), 1-24.

Perry, J.W. (1955): Translation of Russian technical literature by machine: notes on preliminary experiments. *Mechanical Translation* 2 (1), 15-24.

Reifler, E. (1954): The first conference on mechanical translation. *Mechanical Translation* 1 (2), 23-32.

Reynolds, A.C. (1954): The conference on mechanical translation held at M.I.T., June 17-20, 1952. *Mechanical Translation* 1 (3), 47-55.

Richens, R.H. and Booth, A.D. (1955): Some methods of mechanized translation. In: Locke and Booth (1955), 24-46.

Weaver, W. (1949): Translation. Repr.in: Locke and Booth (1955), 15-23.

Yngve, V.H. (1955): Syntax and the problem of multiple meaning. In: Locke and Booth (1955), 208-266.