[Corrected version (2005) of paper in: *Machine Translation*, vol.12 no.3, 1997, p.195-252]

From first conception to first demonstration: the nascent years of machine translation, 1947-1954 A chronology

John Hutchins
(http://ourworld.compuserve.com/homepages/WJHutchins)

1. The beginnings in March 1947

On 4th March 1947, Warren Weaver, director of the Natural Sciences Division of the Rockefeller Foundation, wrote to the cyberneticist Norbert Wiener:

> One thing I wanted to ask you about is this. A most serious problem, for UNESCO and for the constructive and peaceful future of the planet, is the problem of translation, as it unavoidably affects the communication between peoples. Huxley has recently told me that they are appalled by the magnitude and the importance of the translation job.
>
> 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.
>
> Also knowing nothing official about, but having guessed and inferred considerable about, powerful new mechanized methods in cryptography - methods which I believe succeed even when one does not know what language has been coded - one naturally wonders 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."
>
> Have you ever thought about this? As a linguist and expert on computers, do you think it is worth thinking about?

Weaver had been involved in ballistics work during the War, and it was through this activity that he had met Wiener and had learnt of his work on cybernetics[1] and on the simulation and computerisation of human behaviour. Weaver had thus good reason to believe that Wiener would be interested in the problem. Weaver's wartime activity had also brought him in contact with those engaged in the application of computers to the decipherment of enemy messages.

Just two days after writing to Wiener (6th March 1947), Weaver met Andrew Booth, a British crystallographer from Birkbeck College (University of London), who was in the United States investigating the newly invented computers at the Moore School of Engineering, Pennsylvania, at Princeton University, and at other locations. Booth had seen Weaver briefly the previous year (20th June 1946) before visiting various computer installation. On this occasion he came to discuss with Weaver the possibilities of 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 did believe that the Americans might be interested in new non-numerical uses for computers, and he suggested translation as an idea for Booth to think about.

These two events in March 1947 mark the beginning of interest in the application of computers to the translation of natural languages – or machine translation (MT) as it was to be called. Computers were still at an early stage of development. The first electronic computers had been constructed during the war (the ENIAC at the Moore School, the Colossus at Bletchley Park, UK, and by Zuse in Germany); the first 'true' stored-program computer (the EDVAC) had been conceived only two years before, in 1945; and the earliest large computers were still under construction at this time, many not operating until 1949 (Augarten 1984).

2. Some earlier ideas

The idea of MT had come to Weaver some time earlier. In his autobiography (Weaver 1970: 106) he states: "Early in 1947, having pondered the matter for nearly two years, I started to formulate some ideas about using computers to translate..." It is quite possible that he spoke about it to others. In his introduction to the MT conference at MIT in June 1952, Bar-Hillel states that the application of computers to non-numerical tasks "challenged the thinkers' imagination as early as 1945... In that year, Dr. Warren Weaver... started thinking and talking about the possibility of using computer-like machinery for carrying out part or all of the operations involved in translating from one language to another" (Bar-Hillel 1952a).

There is, however, no evidence that Weaver had mentioned MT to anyone before writing to Wiener and meeting Booth in March 1947. At a later date, Booth stated that his discussion with Weaver about MT took place in June 1946 (Booth and Locke 1955). It is certainly true that he met Weaver on 20th June 1946 and that he referred to the funding of a British computer. There is, however, no documentary evidence that Weaver raised the idea of MT at this meeting[2]. Indeed, Booth's report to the Rockefeller Foundation in July 1946 detailing at length possible applications of computers makes no mention of translation (Booth 1946). Nevertheless, Booth's report of this meeting in his chronology (Booth and Locke 1955) has been taken as historically accurate by subsequent historians (Delavenay 1960, Mounin 1964, Locke 1975, Zarechnak 1979, Bruderer 1982). It should be noted that Booth had already given the 1947 date in an earlier paper (Booth 1953), which he was to repeat explicitly later (e.g. Booth 1956, 1959).[3]

On the other hand, in May 1948 Weaver wrote to J.Desmond Bernal (Booth's senior at Birkbeck College) that "I think I may have been the first to suggest this possible use. At any rate I did discuss it two or three years ago, and talked with both you and Booth about it, I believe." This does not necessarily imply that he did in fact mention MT to Bernal or Booth in 1946 (or even, less plausibly, in 1945), only that he believed he mentioned it to someone or other at that time. This person might possibly have been Desmond Bernal. It is known that Bernal had received grants from the Rockefeller Foundation since 1941, that it was through Bernal's contact with Weaver that Booth first went to the United States and that Bernal did meet Weaver in April 1946.[4]

Booth also claimed later (Booth 1985) that already before 1947 he had had some ideas about MT from conversations with Alan Turing. This is plausible, since Booth would have met Turing at computer meetings in London, and Turing mentioned MT as a possible demonstration of the 'intelligence' of computers in a report for the National Physical Laboratory written in September 1947 – although not submitted until July 1948 (Turing 1969). However, Turing may not have first thought of MT until sometime during 1947, since an earlier lecture given by him on 10

February 1947 does not mention MT as an application of ‘machine intelligence’ (Turing 1986). In this case, it may also be that Turing did not think of MT independently, but he had got the idea directly from Booth sometime between Booth’s meeting with Weaver in March and the writing of his NPL report in September. Whatever the truth, Booth's recollection of having spoken to Turing about MT before 1947 must be treated with caution.

In fact, the idea of mechanising translation can be traced back much earlier, to the seventeenth century (Salmon 1972), even though it was not until the twentieth that the means became available to put it into practice. In 1924, for example, a model for a ‘typewriter-translator’ was demonstrated in Estonia (Mel’chuk and Ravich 1967: 26)[5]; and in 1933, two patents appeared independently, one in France from Georges Artsrouni, the other in Russia from Petr Troyanskii. In both cases, the principal focus was on mechanising multilingual dictionaries using electrical devices (Hutchins 1993).[6] However, Troyanskii continued to develop his ideas until the end of the Second World War and in about 1947 he had gone much further in also proposing the automation of the whole translation process and not just the dictionary. However, these precursors were not known to the MT pioneers of the 1940s and early 1950s, and indeed they remained unknown until the late 1950s (Corbé 1960; Panov et al. 1956; Bel’skaya et al. 1959).

3. Reactions to Weaver's letter

Norbert Wiener's reply on 30 April 1947 to his letter was disappointing for Weaver:

> ....as to the problem of mechanical translation, I frankly am afraid the boundaries of words in different languages are too vague and the emotional and international connotations are too extensive to make any quasi mechanical translation scheme very hopeful.

However, Wiener continued:

> I will admit that basic English seems to indicate that we can go further than we have generally done in the mechanization of speech, but you must remember that in certain respects basic English is the reverse of mechanical and throws upon such words as “get”, a burden which is much greater than most words carry in conventional English.

Wiener’s reference to Basic English reflects the widespread interest of the time in this simplified version of English as a means of international communication, proposed by Charles Kay Ogden (1930). The attraction to Weaver, Wiener and many contemporary scientists was the notion that such a small vocabulary of just 850 words could suffice for communication at quite complex levels.

Weaver wrote again to Wiener on 9th May. While expressing his disappointment, Weaver was more hopeful about Basic English:

> The difficulty you mention concerning Basic seems to me to have a rather easy answer. It is, of course, true that Basic puts multiple use on an action verb such as “get.” But even so, the two-word combinations such as “get up”, “get over”, “get back”, etc., are, in Basic, not really very numerous. Suppose we take a vocabulary of 2,000 words, and admit for good measure all the two-word combinations as if they were single words. The vocabulary is still only four million: and that is not so formidable a number to a modern computer, is it?

Prompted perhaps by this reference to Basic English, one of Weaver’s staff John Marshall wrote on 3 September 1947 on Weaver's behalf to Ivor A. Richards, one of its promoters

(Richards (1943), a British literary critic who had been professor at Harvard University since 1939:

> When you are next in New York Warren Weaver, our Director of the Natural Sciences, would very much like to have a chance to talk with you about the possibility of utilizing the new super calculating machines in translation. I have talked with Warren a number of times about this and it seems to us that the feasibility of any such use involves primarily what is known as "briefing" the machine. According to Warren there can hardly be any doubt about its capacity for this kind of thing.

Apparently, the response from Richards (which has not been preserved) was as unenthusiastic as Wiener's.

However, the idea was discussed by Weaver's colleagues at the Rockefeller Foundation. One, commenting on Marshall's letter to Richards, asked in an internal memorandum (17 September 1947):

> Does "briefing" mean pre-setting the machine *in toto*? The variables that need formulating under even a limited vocabulary would surely call for more labor than that required of the same mind for direct translation. The real fallacy, I should say, is in believing any mechanical weapon is as efficient as the mind that made it in meeting variations. Translation is so largely a question of associated meaning according to the time of writing a particular word that at the very best this device might do something of rudimentary kind in simple exposition.

Such doubts would have been natural – indeed, they are still prevalent in translators' circles in the 1990s – and Weaver was feeling already the need to elaborate. In a hand-written note attached to the memorandum he wrote: "I am going to write a memo which will answer this and, I hope, other questions." In fact, Weaver's memorandum was not to appear until July 1949.

Before doing so, however, he checked his facts on cryptography. On 11 March 1948 he wrote to Professor W. Prager at Brown University. Weaver recollected that

> ... one of your colleagues (I think a mathematician) asked you if you would be interested to cook up some coded message on which he could try a cryptographic scheme, or process, or device: that you wrote out a message in Turkish and then coded it by some rather complicated code into a column of five digit numbers: and that he brought this back to you within a short time (actually, it seems to me that you said the next day) with the remark that they had apparently not had much success: but that what he handed you was, when properly broken up into words, and when corrected mildly, your original message in Turkish.

Weaver added that he was not interested in the deciphering problem itself but because of "the possibility that the written languages have certain invariant properties which are independent of the language in question", and because, if true, "it has... an interesting bearing on the problem of translation..."

Prager confirmed the story (15 March 1948):

> The colleague in question was Professor R.E.Gilman of our Mathematics Department... it was based on the frequencies of the letters, pairs of letters, etc. in English... before coding, I somewhat simplified the spelling of the original Turkish message by replacing ç, ğ, î, ö, ş, ü by c, g, i, o, s, u, respectively, though these letters are pronounced differently. Gilman's result would have been read correctly and without difficulty by anyone familiar with the language.

4. Progress by Booth and Richens in 1948

On his return to England, Booth did not forget about MT. In a report to the Rockefeller Foundation in February 1948, after listing a number of the possible applications for the proposed US-funded computer at the University of London, he wrote:

> A concluding example, of possible application of the electronic computer, is that of translating from one language into another. 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. (Booth 1948)

It was in response to this comment that Weaver wrote the letter to Bernal in May 1948 cited above (section 2).

From later accounts it is clear that Booth was referring primarily to the work which he and Kathleen Britten (his future wife) had done while at the Institute for Advanced Study in Princeton during 1947 (Booth and Locke 1955). There they had looked at the problems of programming a mechanical dictionary for translation, but without actual realisation on a computer. It was this work which he had continued on returning to England.

By this time Booth had, in addition, met Richard H.Richens – on 11 November 1947 through Douglas R.Hartree (Richens 1984), who had known Booth since 1945. Hartree was one of the principal figures in early computer developments; in June 1945 he had paid an official visit on behalf of the British government to the United States to report on developments; in October 1946 he moved to the Cavendish Laboratory at Cambridge University, where he supported Maurice Wilkes in work on the influential EDSAC computer. 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 (Richens 1951):

> 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 (Richens 1984)

Shortly after receiving Booth's report, Weaver travelled to England, meeting both Bernal and Booth, primarily to discuss the funding of a computer at Birkbeck College. On 23 May 1948, Weaver went with his Rockefeller colleague Gerard Pomerat to the Laboratory of the British Rubber Producers Research Association in Welwyn Garden City near London, to see the computer installation set up by Booth but now run by Kathleen Britten. Weaver asked Booth about his MT experiments, and Booth told him about his collaboration with Richens. With a tone of some disappointment, Weaver summarised the British MT effort as follows:

> So far, it turns out, they have not considered at all the problem of multiple meaning (!), and have been concerned only with the mechanics of looking up words in a dictionary. First, you sense the first letter of a word, and then have the machine see whether or not the memory contains precisely the word in question. If so, the machine simply produces the translation (which is the rub; of course "the" translation doesn't exist) of this word. If this exact word is not contained in the memory, then the machine discards the last letter of the word, and tries over. If this fails, it discards another letter, and tries again. After it has found the largest initial combination of letters which <u>is</u> in the dictionary, it "looks up" the whole discarded portion in a special "grammatical annex" of the dictionary. Thus confronted by "running," it might find "run" and then find out what the ending (n)ing does to "run."

> This may be useful, but it is certainly only a very small part – almost a trivial part – of the real translating problem.

The method had been developed by Richens using the punched cards familiar from his work in indexing (Richens 1951). 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 introduce grammatical information into a dictionary-based system. For example, in the case of the Latin verb *amat* a search is made for the longest matching stem, i.e. 'am-', and for the ending '-at'. The stem indicates the English translation *love* and the ending gives the grammatical information '3rd person singular'. In this way, grammatical annotations augment a strict 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 does not appear to have been implemented as a computer program, primarily because the authors did not have access to a computer of sufficient size and reliability.[7] The experiment was written up in a memorandum during 1948, although not published until much later in a revised version (Richens and Booth 1952). From this account, the crudity of the output is obvious.

> (German) Wenn in ein*em gröss*eren Gebiet zwei Form*en neben*einander leb*en, ohne sich zu vermisch*en, so gehör*en sie verschieden*en Form*en*kreis*en an.
>
> (English) if in a/one *d* large (more) area two form *m* beside one another live *z* without self to/too mix *z*, so belong/hear *p z z* different *m* form *m* circle *m* at.

The asterisks in the German indicate automatic segmentations. In the English 'translation' *d* indicates a German word in the dative case, *p* a past verb form, *m* "multiple, plural or dual", *z* "unspecific", and slashes alternative translations.

Many other example translations are to be seen in the memorandum – from Albanian, Danish, Dutch, Finnish, French[8], Hungarian, Indonesian, Italian, Latin, Latvian, Norwegian, Polish, Portuguese, Romanian, Spanish, Swedish, Turkish, Arabic, and Japanese – demonstrating a commendable enthusiasm and confidence. Although, as Weaver remarked, it was only the very first stage of a true translation system, the step had been made, and Richens is rightly credited as being the first to propose a method to automatically identify word-endings (Delavenay 1960:28).

From 21 June to 2 July 1948 the Royal Society Scientific Information Conference took place in London, in which Bernal took a prominent part. Included in the proceedings (Royal Society 1948) is a note on "The mechanized distribution of information". This report includes what is probably the first mention of MT in print. It reads in full:

> The basis of translation being largely mechanical, in that it consists of the transference of ideas from their expression in one set of symbols to another set, the construction of devices to accomplish this mechanically would not be impossible. Practical research has, in fact, already been started.

The research referred to was almost certainly that of Booth and Richens, since both attended working parties of the conference (Booth the Working Group on Translation, and Richens the Working Group on Mechanical Indexing).

5. Developments in Los Angeles

By this time, however, some research had begun in the United States. This was the work based on the SWAC computer in Los Angeles under the direction of Harry D.Huskey.

Huskey had been involved in both the ENIAC and the EDVAC projects for developing computers in the United States, and in January 1947, on the recommendation of Douglas R.Hartree, he went for a year to the National Physical Laboratory (NPL) at Teddington near London where he joined the section which was developing the ACE computer under Alan Turing (Huskey 1984, Wilkinson 1980). Before his arrival, it had been decided to contract out the building of the computer; Huskey persuaded the NPL to change this policy and led the team in building the 'pilot ACE' based on a simplification of Turing's designs – Turing himself had virtually left the NPL by this time (Wilkinson 1980). However, the project was not completed by the end of Huskey's sabbatical year and he returned to the United States in disillusion.

In January 1948 Huskey returned from England and took up an appointment at the National Bureau of Standards (NBS) in the National Applied Mathematics Laboratories. Shortly afterwards, in May 1948, the NBS decided to build its own computer in Washington, and then, in October it decided to build another at the Institute for Numerical Analysis. This was a newly founded branch of the NBS located on the campus of the University of California in Los Angeles. Huskey transferred to Los Angeles and began work in January 1949 on the computer, which was eventually called the Standards Western Automatic Computer (SWAC).

It was perhaps from conversations with Hartree or Turing about ACE and computing in general during his stay at NPL that Huskey thought of using SWAC to translate. However, Hartree was also acting director of the NBS Institute in Los Angeles in early 1949; and, as we have seen, Hartree knew both Booth and Richens as well as Turing. On the other hand, he may have had the idea quite independently.

Although the SWAC computer was not fully finished until July 1950 (Huskey 1980), it must have been sufficiently advanced in May 1949 for a public demonstration. On 31 May 1949 the *New York Times* carried a report bearing what must surely have been at the time a most startling headline "'Electric Brain' able to translate foreign languages is built" – it was the very first public announcement of the possibility of machine translation. The report by Gladwin Hill from Los Angeles began:

> 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 at the United States Bureau of Standards Laboratory at the University of California's Institute of Numerical Analysis. While the exact scope 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, and after describing various types then existing including the SWAC type, the report continues:

> When a foreign word for translation is fed into the machine, in the form of an electro-mathematical symbol on 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, however:

> This admittedly will amount to a crude word-for-word translation, lacking syntax, but will nevertheless be extremely valuable, the designers say, for such purposes as

> scientists' translations of foreign technical papers in which vocabulary is far more of a problem than syntax. An obvious corollary is its possible use in cryptography.

The comment on cryptography is indicative of the current interest in the achievements of computers. It is unlikely that the reporter knew of Weaver's opinion of the analogy between translation and cryptography, although Huskey may well have had similar ideas.

A much shorter report on SWAC appeared the same day in the *New York Herald Tribune*, entitled "'Brain' developed to do translations". This contained the following passages:

> Development of an "electric brain" which will automatically translate foreign languages was disclosed today by Dr.Harry Huskey... He said the electronic computer primarily was a tool to aid in advanced mathematical research, but an auxiliary problem in the work was language translation and the machine would be able to do that function....
>
> Dr. Huskey... said the "brain" would give "word for word substitution" and those making use of it would then have to interpret it.

On the following day, an editorial in the *New York Times* expressed greater scepticism and doubts – it was, however, one of the few occasions when MT has received such prominence in a newspaper:

> We have our misgivings about the accuracy of every translation. 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 mistake of assuming that the awesome mathematical power of computers made them capable of anything:

> ... differential equations can be handled by machinery. So can thousands of biological facts – facts that it would take years, even a lifetime, to compile and interpret. Translating 5,000 technical words from French into English – what's that compared with the making of discoveries that will tell us what happens in the nervous system when we are worried or angry or when a normal cell in the body goes wild and grows into a tumor?

6. Weaver's memorandum in July 1949

In July 1949, two years after writing to Wiener and meeting Booth, Weaver wrote the memorandum entitled simply "Translation" which was to launch MT as a serious subject of research in the United States, and subsequently throughout the world (Weaver 1949). The memorandum, dated 15 July 1949 and written from "Carlsbad, New Mexico", was sent, according to Weaver himself (Weaver 1970: 106), to "twenty or thirty persons - students of linguistics, logicians, and mathematicians." However, the editors of the collection in which it was reprinted (Locke and Booth 1955) gave the more inflated figure of 200 which is the one frequently quoted. It is, of course, quite possible that this was the size of the eventual readership in the next few months.

Weaver's 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 the story about decoding Turkish, which he had checked with

Professor Prager, adding the biographical detail that Prager "had spent some time at the University of Istanbul and had learnt Turkish there." He stresses that:

> The most important point... is that the decoding was done by someone who did not know Turkish, and did not know that the message was in Turkish.

and he then asserts that the success of cryptography depends 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 this observation leads him to suppose that

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

The reason for this, Weaver attributes to the fact that "languages were invented and developed by *men*; and all men... have essentially the same equipment to bring to bear on this problem." To buttress his argument, he then turns to the findings of the logicians – referring particularly to Hans Reichenbach (who had also spent some time in Turkey) – who had observed the same "basic logical structures" in widely diverging languages[9]. Lastly, he notes that in the journal *Science* had appeared a summary of a paper by Erwin Reifler (1948) which included the comment that "the Chinese words for 'to shoot' and 'to dismiss' show a remarkable phonological and graphic agreement". It leads Weaver to comment:

> This all seems very strange until one thinks of the two meanings of "to fire" in English. Is this only happenstance? How widespread are such correlations?

At this point in the memorandum, Weaver recalls his exposure to computers in the War and states that "it was very natural... to think, several years ago, of the possibility that such computers be used for translation." He then reproduces in full his correspondence with Norbert Wiener – the letter of 4th March 1947, Wiener's reply of 30th April, and his response of 9th May.

Although disappointed by Wiener, Weaver is encouraged by evidence of interest in MT elsewhere. He quotes the extract from Booth's memorandum of February 1948, recounts his visit to England and reproduces his description of the work of Richens (Section 4 above); and adds that he has "no more recent news of the affair." His second example of activity is the newspaper reports from California. Strangely he does not mention where it was undertaken or who was directing the work[10]. He points out, quite rightly, the limitations of the proposed method in much the same terms as he criticised Booth and Richens' work:

> The published reports do not indicate much more than a word-into-word sort of translation, and there has been no indication... of the proposed manner of handling the problems of multiple meaning, context, word order, etc.

With some approval he summarises the comments by Max Zeldner in a letter to the *New York Herald Tribune* dated 13 June 1949, prompted by the reports of the SWAC demonstration:[11]

> the most you could expect of a machine translation of the fifty-five Hebrew words which form the 23rd Psalm would start out "Lord my shepherd no I will lack", and would close "But good and kindness he will chase me all days of my life; and I shall rest in the house of Lord to length days."

Weaver is quite ready to admit the uselessness of the method for literary translation, but he does stress that even word-for-word translation could be valuable for technical material – primarily because "the problem of multiple meaning is enormously simpler. In mathematics... one can very nearly say that each word, within the general context of a mathematical article, has one and only one meaning."

Weaver could obviously not point to any actual translation by computer – the work of Booth and Richens was simulated on punched cards and not yet implemented in a computer program, and the work of Huskey was merely a proposed application of a computer still under construction. Weaver's main concern in this memorandum is, therefore, to indicate how MT might go beyond the mechanical dictionary methods and beyond the severe limitations of word-for-word substitution. He proposes therefore, as he puts it, "at the risk of being foolishly naïve... four types of attack".

His first proposal is that the problem of multiple meanings might be tackled by examination of immediate context:

> If one examines the words in a book, one at a time as through an opaque mask with a hole in it one word wide, then it is obviously impossible to determine, one at a time, the meaning of words. "Fast" may mean "rapid"; or it may mean "motionless"; and there is no way of telling which.
>
> But, if one lengthens the slit in the opaque mask, until one can see not only the central word in question but also say N words on either side, then, if N is large enough one can unambiguously decide the meaning...

The question is how much context is required. For Weaver it relates to the "statistical semantic character of language", to which he has already referred in the memorandum. The size of the context may vary from subject to subject, and may vary from one word to another. However, on the assumption that "relatively few nouns, verbs and adjectives" are actually ambiguous. Weaver is led to "the concept of a translation process in which, in determining meaning for a word, account is taken of the immediate (2N word) context".

Weaver's second proposal starts from the assumption that there are (at least in non-literary works) logical elements in language. He draws attention to the theorem proved in 1943 by McCulloch and Pitts (1943), which "states that a robot (or a computer) constructed with regenerative loops of a certain formal character is capable of deducing any legitimate conclusion from a finite set of premises." The mathematical possibility of computing logical proofs – related to Turing's 'thought-experiment' (Turing 1936) – suggests to Weaver that "insofar as written language is an expression of logical character" the problem of translation by computer is formally solvable.

The idea was no doubt linked in Weaver's mind with his third proposal which involves the recently developed 'information theory' of Claude Shannon. Weaver himself was a mathematician and at the time of writing this memorandum, he was in fact collaborating with Shannon on a seminal book-length treatment of the subject (Shannon and Weaver 1949). Shannon's theory is concerned with the basic statistical properties of communication, and embraces "the whole field of cryptography"[12]. Weaver concedes that the validity of this approach was difficult to assess (admitting that probably only Shannon himself could do so); but nevertheless, Weaver was obviously very attracted[13] - and he repeated, changing the language, the observation made in his letter to Wiener:

> it is very tempting to say that a book written in Chinese is simply a book written in English which was coded into the "Chinese code." If we have useful methods for solving almost any cryptographic problem, may it not be that with proper interpretation we already have useful methods for translation?

Weaver makes at this point an observation which was to be overlooked by many of those who were inspired to take up the challenge of MT research in the following years. He stresses the fundamentally probabilistic nature of language in use, 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.

In other words, he emphasises that even imperfect translation can be useful – another observation which was too often forgotten in subsequent years.

However, in his fourth and final proposal, he becomes more utopian. The assumed universal applicability of the cryptographic-statistical approach to translation brings him back to the belief there are linguistic universals common to all human languages. The "most promising approach of all" is to investigate the fundamental structures of languages. He provides a metaphor:

> 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 repeats his belief that Ogden and Richard's work on Basic English represents some steps towards it – despite the grave, and fully justified, reservations of Weaver and his own colleagues in the passages cited above.

When writing his memorandum, Weaver was not only, for obvious reasons, unaware of the previous research by Troyanskii (section 2 above); he seems also to have been unaware of the long history of unsuccessful attempts to devise universal languages, including those based on logical (and philosophical) foundations. Instead he holds confidently to a belief in the reality of a 'universal language' and that it can be discovered.

In this memorandum Weaver presented in outline some of the basic features and issues of MT which remain valid to the present day. He stressed statistical methods, the problem of multiple meaning and ambiguity, the need to look at context, the desirability of concentrating on the language of specific subjects, the two basic approaches (direct translation, and indirect translation via an intermediate form), and the search for language features common to more than one pair of languages. In one aspect only was he on the wrong track, and this was his advocacy of cryptographic methods.[14] However, it was not long before the mistakenness of this proposal was recognised (cf. Delavenay 1960: 8; Mounin 1964: 31-39). It came from a confusion between the activities of deciphering and translation which arise whenever the same person does both. Decipherment is monolingual and involves frequency analysis of the language and text to be deciphered; in the case cited by Weaver, it was based on "the frequencies of the letters, pairs of letters, etc. in English"; fortunately, in this instance, they were much the same in Turkish and the original could be interpreted.

7. Other independent 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 occurred to others at this time. The possibly independent suggestions of Turing and Huskey have been mentioned already; and it is obvious that Richens must have thought of MT independently before meeting Booth. Two other examples are known.

The first was revealed some years after Weaver's 1949 memorandum and long after MT research had begun in earnest. In 1957 Duncan Harkin wrote:

> In 1946, the present author happened to come upon a copy of the Bell Laboratories word-count on spoken English, which included graphical presentation of the statistics on both written and spoken English. It showed that 500 words account for 77% of all the words we use in writing and 93% of the words we use in speaking. This suggested the feasibility of mechanical translation on an economic basis – that it was not necessary to have hundreds of thousands of words available in storage. The following year, in collaboration with an electronics engineer, the essential outline of a machine for mechanical translation was designed. This was the first recognition of actual feasibility on an economic basis and the first actual design proposed for mechanical translation. The author later learned of the ideas of Dr. Warren Weaver and of his suggestion of the desirability of automatic translation, however without any enabling proposals. (Harkin 1957)

Although no further details are provided it is evident that the proposal would have done no more than give a word for word translation of a simple text, with the most significant words not being translated at all since the set of 500 would be function words and common vocabulary. The proposal was less ambitious than Troyanskii's (Hutchins 1993), and whether Harkin intended a computer implementation is uncertain.

The second consists of an intriguing announcement from Italy just three months after the newspaper report of SWAC. On 26 August 1949, the *New York Times* reported (page 9) from Salerno:

> Federico Pucci announced today that he had invented a machine that could translate copy from any language into any other language. He said that the machine was electrically operated, but refused to disclose details. He said that he would enter it in the Paris International Fair of Inventions next month.

It is uncertain whether Pucci had any knowledge of Huskey's proposals, and it seems most unlikely he knew about Weaver's memorandum or the British experiments. In any event, there is no trace of any demonstration at the Paris fair; and nothing more is known about Pucci.

It is not known how many others at this time had similar ideas about translating machines. 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. Translation was then, and often still is, regarded by those unfamiliar with its difficulties as essentially a question of finding equivalent words in another language. To use a computer in such a task seemed trivial.

8. The Unesco conference

In fact, Booth and Richens' ideas were already becoming more widely known before Weaver's memorandum. On 30 December 1949, J.B.Reid of Unesco wrote to Weaver from Paris acknowledging its receipt – a copy had been sent to Pierre Auger, an eminent French professor of physics, who was at the time head of the Department of Science at Unesco. Reid comments:

> I wish very much that I had had your memorandum on 21 June this year. In the International Conference on Science Abstracting, we arranged one committee on the effects on abstracting of new technical devices for handling information... When the discussion lagged, I introduced the mechanical translation idea, reproducing what I had learned of Booth's work from Bernal. Most people thought that I was joking and the interpreters seemed to be supremely unconcerned at this potential threat to their jobs. Not having your knowledge of the subject, I was unable to press the idea any further and allowed the matter to pass off as a comic interlude.

Reid himself did not think MT was feasible, except in the very long term, and he was sceptical about universal languages – even the wider adoption of Esperanto to aid international communication seemed too remote. However, Unesco was interested in the mechanisation of terminology and multilingual dictionaries and had asked J.E. Holmström to write a report. It was circulated in mimeograph form dated 20 July 1949, although not published until 1951. The author speculated about the possibility of MT, basing his remarks on the proposals by Richens (Holmström 1951: 21-22):

> Mr.R.H.Richens... has tentatively thrown out the suggestion (but has not published anything on the subject) that Dr.Booth's machine might be developed to serve not merely as a dictionary of isolated terms but as a means of doing translations automatically.

Holmström was highly sceptical:

> ...whereas arithmetic is utterly logical translation is not so... translation is an art; something which at every step involves personal choice between uncodifiable alternatives; not merely direct substitutions of equated sets of symbols but choices of values dependent for their soundness on the whole antecedent education and personality of the translator.

It was an opinion to be echoed repeatedly by translators over the next decades. However ingenious the analysis of grammar and syntax, Holmström was convinced that:

> the resulting literary style would be atrocious and fuller of "howlers" and false values than the worst that any human translator produces.

He conceded that Richens was not making extravagant claims, merely proposing that computers "might be made to turn out a rough draft which a competent editor versed in the subject matter, though unacquainted with the *foreign language*, could then pull into shape." But he countered with the comment that "the machine would need to be given an enormous memory not only for variants of terms but for rules of accidence and syntax." Invited to respond, Richens claimed that:

> Syntax is not a very serious matter, and although a skeleton knowledge of syntax may be necessary for a few languages, in using such a mechanical translator, this will not make the translation of the machine any less intelligible.

This view also was to be echoed in the following few years by the early MT researchers who over-estimated the intelligibility of word-for-word renditions and who grossly under-estimated the problems of syntactic analysis.

9. Responses to Weaver's memorandum

Booth acknowledged receipt of Weaver's memorandum on 14 September 1949. He was intrigued by his description of a powerful cryptanalytic method:

> I hope that it is not of such a nature as to be withheld from the general scientific public for very long.

and then added:

> We have not forgotten the problem of translation in my group but rather have had to hold it in abeyance due to the pressure of so many other and more urgent activities.

Booth was primarily involved in designing and constructing a computer for Birkbeck College (Booth 1980). His resources were small and the problems of acquiring the necessary materials considerable. Not surprisingly, he could not do much on MT for several years.

Many of the initial responses to Weaver's memorandum were as negative as Wiener's had been to his letter in March 1947. But others were enthusiastic. For example, Vannevar Bush – developer of the mechanical and electronic analogue calculator, and the *eminence grise* of US science during and after the War – responded in October 1949 (quoted in Booth and Locke 1955):

> Far from the multiple meaning of words being a barrier to your project, it seems to me it is an opportunity... I do not think it would be difficult at all to make the translating machine exercise as good a judgment in picking the right word as is exercised by many human translators, particularly when they get to translating in a subject which they do not understand. Another thing is word order... it would not be beyond the realm of reasonable possibility to force it [the machine] to follow certain rules of word order before it emitted the result.

Within the Rockefeller Foundation, Weaver received support from his assistant director in the Natural Sciences division, William F.Loomis, who wrote on 3 November 1949. He began by noting the problems of translating between widely diverging languages and cultures: the feasibility of translation by "word substitution" entailed that "the two languages must have similar structures" so that translation from English into Chinese is likely to be "of a different order of difficulty" than translation into French. But there were more serious semantic differences:

> the civilizations parent to the two languages must have some close resemblances if the referent for the translated word is to have meaning in the language it is translated into. Thus the word "democracy" in America can have no extensional meaning in China where the illiteracy rate, etc., etc., etc., is so different from the American.

Thus, Loomis argued that the languages would have to have similar categorisations of vocabulary ("ladders of abstraction" he called them), such as "species - genus - class - family - kingdom, etc." Translatability seemed to be related to uniqueness of reference; consequently "translation in the sciences with definite referents and operationally defined concepts is amenable to translation by electronic computers far beyond works of art, such as novels, poems, etc." He noted, however, that "analogies, parables, etc., survive translation with surprising effectiveness", and in scientific documents the presence of graphs and diagrams is of great assistance in conveying the intended sense. He ended on a positive note:

> Since translation can be accomplished by humans even of relatively poor education but who are familiar with both languages it is indeed fascinating to picture a machine that can reproduce the relatively low order of mental processes that occur in, say, a student translating his daily passage of Caesar. WFL can see no theoretical impossibility in making electronic computers that could translate material, at least of the more factual variety, at least as well as some humans do...

Finally he suggested that there may be some statistical work of relevance, e.g. "the frequency distributions of multiple meanings as recorded in, say, two dictionaries in two languages".

10. Wider circulation

Weaver's memorandum was soon to receive a wider readership through a brief item in the December 1949 issue of *Scientific American* (vol.181 no.6, p.30). Under the heading "Translating Machines?" the anonymous article begins:

> If machines can be built to count, calculate, play chess, even "think", why not a machine to translate one language into another? Scientists have been pondering this possibility.

There follows a summary, evidently taken from the memorandum itself (although not acknowledged), of the decoding of the Turkish text and Weaver's suggested use of deciphering techniques, and of the work by Booth and Richens on a mechanical dictionary. However, the reporter had evidently contacted Richens as well, since he continues:

> While translating machines might not be able to do justice to the linguistic niceties of literary or diplomatic English, Richens is convinced they could translate scientific articles: "The resultant translation would be highly artificial and would be what I would call standardized pidgin English. This, though no doubt highly repugnant to those whose main interest ... is esthetic, will constitute no obstacle to those whose chief purpose in using such a machine is to find out what the original is about."[15]

The article came to the notice of a journalist in London, who interviewed Booth. On the 20th December appeared the first report in a British newspaper to mention the possibility of using computers to translate. The headline in the *News Chronicle* (20 December 1949, p.3) was: "Mathematician hasn't enough time to add up, so – His machine knows all the answers, and it knows languages, too." The reporter describes Booth's plans for "a 'thinking' machine" and the financial backing he had received from the Rockefeller Foundation, and the "strictly utilitarian purpose... to provide a machine to help in research into the structure of proteins." Then it describes plans for translation:

> ... In its final form the electronic translator will consist of an ordinary typewriter keyboard connected by cable to a cabinet about the size of a large refrigerator. A cable will connect the "brain" to a teleprinter.
>
> So the operator could select which of a dozen or more languages he desired to translate. As fast as he could type the words, say, in French, the equivalent in Hungarian or Russian would issue on the tape.

After describing the electronics involved[16], the reporter ends by quoting Booth as saying: "There is nothing comparable to or better than my machine yet completed in the United States." Evidently, Booth intended only dictionary translation (as in his and Richens' proposals) and it was clear that a system still lay in the future, but many readers would have received the impression that automatic translation was close to realisation.

The article in *Scientific American* prompted also Charles C.Holt to write from Chicago on 9th February 1950 to the Rockefeller Foundation:

> In the last six months I have become very interested in the feasibility from a technological point of view of designing an electronic machine which would be capable of making word for word translations from one language to another. I have arrived at the conclusion that such a machine could be built at reasonable cost by the use of techniques which have been thoroughly tested in radar, television and microfilm applications. Enormous energy and resources are being devoted to the development of electronic calculating machines, but evidently very little is being done to develop translating machines.

Holt was under the impression from the *Scientific American* article that Weaver was already working on MT and believed the Foundation might be interested in his ideas. Holt elaborated:

> A television camera tube is the first component of my proposed translator. Each time a word is to be translated this tube can read the whole dictionary of words spelled out in the code of black and white spots. Now how can a whole dictionary be riffled before the television eye in one second? This can be accomplished by recording the coded words on a strip of microfilm and mounting it on the circumference of a wheel which spins in front the television camera tube. There would be two columns of coded words on this microfilm strip, the original words being translated and the corresponding translations in the two columns respectively. The wheel turns once a second and the television tube looks at the lines of microfilm spots which represent the words being translated. Each word is read as it goes by.
>
> The material to be translated is typed, only stenographic skill being required, on a machine which punches holes in a paper tape according to the code. This paper tape is now passed before a second television camera tube which reads the rows of spots (holes) of the word being translated. The two television camera tubes are synchronized so that two words of black and white spots are read simultaneously. Every time a new dictionary word is read the word being translated is reread. Electronic circuits compare the two words and when the two are identical indicating that the word to be translated has been located in the dictionary a signal is produced. This signal causes the dictionary television camera tube to shift to the next column of coded spots which spell out the translated version of the word. The electrical signal produced by reading this row of spots is transmitted to an electric typewriter which decodes and types out the translated word.

In some respects, Holt's proposed system is a version of the mechanism put forward by Troyanskii from 1933 to the early 1940s (Hutchins 1993). Both proposed machines which would have a dictionary rotating in front of a input tape of the words of a text. Holt had the advantage of more advanced electronics, but unlike Troyanskii he did not suggest any way of dealing with morphology and syntax. In any case, Holt did not get as far as constructing his device.

In his reply (28 February 1950), after pointing out that he was not doing MT research, Weaver wrote that:

> there is no real difficulty in connection with the actual instrumentation of the translating computer, provided one knew just what it was that he would like this computer to do. Your comments concerning television tubes are interesting, but I think that there is a considerable gap between what a television tube does at present, and the use which you imply...

He remarked that attempts to use cathode ray tubes for computer storage, originally developed F.C.Williams at the Manchester University in 1948 (Augarten 1984: 149) and investigated by a number of US groups subsequently, had not proved to be as satisfactory as anticipated. Even so, Weaver was right to comment that the problems were not technical but "arise in connection with the more linguistic aspects of the problem... multiple meaning, sentence structure, etc."

11. Research begins

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 of the University of Washington, Seattle. As Reifler (1962) put it, "Dr.Weaver sent me a copy of his memorandum because he referred in it to a paper which I read in April 1947, before the American Philosophical Society in Philadelphia on "The Chinese language in the light of comparative semantics," in which I gave numerous examples demonstrating agreements between unrelated languages which are not due to borrowing or cross-fertilization." An abstract of the paper was published in *Science* (Reifler 1948).

Already by 10th January 1950, Reifler produced a 55 page study (Reifler 1950), in which he put forward the first fully formulated conceptions of pre- and post-editing. He conceived the 'mechanical' process purely in terms of word for word one-to-one substitutions. There were obvious inadequacies. He suggested that in order to remove ambiguities from the source text a human "pre-editor" could add extra symbols for grammatical and logical explicitness, i.e. to distinguish both grammatical ambiguity (one word form with different potential syntactic categories) and homographic and polysemic ambiguity (one word form with different senses in the source language.) Theoretically a pre-editor need not know the output language. The task of the post-editor was to resolve residual ambiguities and to render the machine output into a reasonably literate form – and, theoretically too, the post-editor need not know the source language. (The need was perceived earlier by Richens, as evident from the quotations above, but Reifler was the first to coin the term "post-editing.") Reifler conceded that without knowledge of the source language a post-editor would have great difficulty faced with lists of alternatives for each word; ambiguities in the original would 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.

Early the next year interest was kindled at the Massachusetts Institute of Technology by a visit from Warren Weaver (Locke and Yngve 1958: 511):

> In January 1950 Dr Weaver met at MIT with a dozen men from nearly as many different fields, including the heads of our Research Laboratory of Electronics, of our Digital Computer Laboratory, and of the Department of Modern Languages... The conclusion was cautious: the possibility of translation by machine was worth examining.

However, nothing practical seems to have been done at MIT until the appointment the following year of Yehoshua Bar-Hillel.

At the neighbouring Harvard University, Anthony Oettinger began work on his doctoral thesis at the suggestion of Howard Aiken, director of the Computation Laboratory, who had undoubtedly been a recipient of Weaver's memorandum. Oettinger's thesis was an investigation of the requirements for an automatic Russian-English dictionary. As part of his study, he visited the University of London during the following year, where he probably met Booth.

The memorandum had stimulated interest also at the Rand Corporation in Santa Monica, where preliminary studies by Abraham Kaplan followed up Weaver's suggested statistical approach to resolving problems of multiple meaning. In his report completed on 30 November 1950, Kaplan described an investigation of the micro contexts of polysemy in mathematical texts (Kaplan 1950). A group of test subjects (translators with mathematical training) were presented with a set of words, each with a number of possible meanings, and asked to select the most applicable sense. The test was limited to nouns, verbs, and adjectives on the assumption (echoing Weaver's thoughts in his memorandum) that "these are the major carriers of the content of any discourse, and probably more markedly exhibit ambiguities." Each word was presented first in isolation, then together with the preceding and following words (up to two before and after), and finally the whole sentence. Kaplan found 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 – Kaplan excluded homographs (words of different grammatical categories having the same form) – 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.

During 1950, activities in Los Angeles expanded with the initiation, under the direction of Huskey, of preliminary studies by linguists at the University of California at Los Angeles (UCLA). In the Department of Spanish, William E.Bull began some "vocabulary studies", and in the Department of German Victor A.Oswald, Jr. researched syntactic problems in collaboration with Stuart L.Fletcher of the NBS. This was in fact the first research in MT devoted to syntactic questions in MT – and it resulted in the first academic journal article on MT (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 to be candidates for rearrangement before word-for-word translation into English. An illustration of the method was provided on 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.

which was rearranged, on the basis of the 'blocks' identified, 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.

This would have permitted 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.

Although the procedures were formulated as 'instructions' for the SWAC, they were not implemented in any way. Nevertheless, the authors concluded that syntax "does not constitute, as had been thought by some, a barrier to mechanical translations." By contrast, they stressed the greater problems of solving the "lexicographic difficulties".

Reports of demonstrations of the SWAC continued to mention MT as a potential application. On 18th August 1950 there appeared another article in the *New York Times*. Again there was scepticism:

> To what extent the machine can be spared for such work is undecided. The translations, it is admitted, would be very rough and not much better than a boy could do in a language examination with the aid of a "pony".

But, like the British newspaper report earlier in the year, it served to keep the public aware of MT as a possible application of the new computers.

The *News Chronicle* report did not come to the notice of Warren Weaver until September - he was under the impression that it had been "printed some time last spring". In a letter to Booth (11th September 1950), he wondered whether there was more to it than journalistic exaggeration. After reproducing the passage quoted above, he commented:

> I have had too much experience with that sort of thing to take very seriously a newspaper account of this particular character, but the reporter does seem to emphasize a good deal the possible use of your computer for translation.... Is this all a matter of the reporter's imagination and enthusiasm, or have you made some definite progress in the use of your machine for translating purposes?

Booth's reply on the 21st September confirmed, as Weaver suspected, that the reporter had exaggerated[17]. While he could not claim any further progress in his MT activities, he was confident that "there is no technical difficulty whatever in making the dictionary translation...

[although] this, of course, is a very different thing from making an accurate grammatical translation."

12. Loomis' survey of activity

Weaver's colleague at the Rockefeller Foundation, William F.Loomis had maintained an interest in what Weaver was proposing since commenting on the memorandum the previous year (3 November 1949, section 9 above.) In October 1950, Loomis sent Weaver a note commenting on the correspondence resulting from the memorandum:

> To my mind the job is really not as difficult as some of your correspondents have made it out to be, as it essentially consists in mechanizing technical dictionaries such as "German-English Dictionary for Chemists." Assigning some sort of serial number system to each word in the German language would then be expected to produce its corresponding counterpart in English, typed out automatically at the end in German word order. In other words, working with the presently available dictionaries it is possible to mechanically envision the inelegant translation of technical data.

After then conceding that 'multiple meanings' were a problem ("the shortest and most common words are found to carry an inordinate load of different usages"), although presuming that statistics could help, he then remarked:

> The real difficulty, as I see it, is mechanical in that according to J.Forrester of... M.I.T., present machines lack the capacity for internal memory storage by a factor of approximately 100... It would seem that in 1950 the bottleneck in this field lies in the present electronic weaknesses of such machines to store information which is not arranged in the order in which it is to be used....

Such a view was to be shared by many in the coming years, particularly those coming from engineering and mathematical backgrounds – and initially, these people predominated because they alone were able to program the earliest computers.

On 8th November 1950, Loomis wrote to all those with whom there had been contact over the past two years. The object was "to find out what persons are actively interested in this problem, and what actual projects, if any, are being carried out." Secondly:

> If the amount and level of interest is at all encouraging, then we would wish to consider a further question – namely, in what way can such efforts be stimulated or helped? Should a conference, for example, be arranged between certain isolated groups that at present may be working quite independently on different aspects of the problem?

On 18th January 1951 Loomis was able to report back to Weaver[18]. Three locations of MT activity were identified: two in Los Angeles at NBS/UCLA and at Rand, and one in Britain.

From the Rand Corporation Williams reported about the encouraging work of Abraham Kaplan; and (according to Shannon) "M.Flood of the Rand Project [had outlined] some ideas in mechanical translation." From NBS Huskey reported on the theoretical work by Bull and Oswald – the latter had sent Loomis a draft of his paper – and said he was interested in "pilot tests" on the SWAC even though it was "not designed for non-numerical work"; J.D. Williams commented that "Huskey... is looking for jobs for his SWAC."

From Britain, there were replies from Booth and Richens; and Calvin Mooers (Zator Company) confirmed that Richens had "actually operated a tabulating machine to do translation by printing the multiple equivalents." This had been reported to him by Robert Fairthorne (Royal Aircraft Establishment, Farnborough, UK). Mooers' letter is in fact the first concrete confirmation that translation had been achieved mechanically, albeit not on a computer.

No other locations were identified as being currently active. However, but it was noted that there was some considerable interest in the Research Laboratory of Electronics at MIT. (J.B.Wiesner's reply had mentioned Jay Forrester, head of the Computing Laboratory, Dr.Straus and Professor Fano in Wiesner's laboratory, and a graduate student Anatole Holt "interested in the logic of language") – this was probably the result of Weaver's visit to MIT the previous January (section 11 above.) MT had also been talked about at the Bell Laboratories, and, in Britain, there was said to be "long-term" interest expressed by D.M.MacKay (King's College), D.B.Fry (University College), Dennis Gabor (Imperial College) and Dr.G.Timms (not apparently connected with any university).

But the name most often occurring was that of Alan Turing, who was mentioned by J.B.Wiesner (MIT Research Laboratory of Electronics), Claude Shannon (Bell Telephone Laboratories), and Donald MacKay (King's College London). All knew that he had been thinking about MT and suspected he might be doing research. As we know, he was not; but it is further circumstantial confirmation that Turing was recognised as one who had earlier influenced speculation in this field.

Loomis ended his report to Weaver with recommendations of financial support to the two groups at the National Bureau of Standards and the Rand Corporation, and he proposed that a conference could be held in London, with a second meeting two years later in Los Angeles. (In a later memorandum of 19th March 1951, Loomis suggested Cambridge (England) instead of London.)

A one-year grant of $5,000 to Huskey at the NBS was approved on 12th July 1951 "to test the idea of automatic mechanical translation... but also for the supporting studies" of Oswald and Bull. This is the first grant specifically for MT research. (Any previous grants to Booth would have been for computer equipment.)

13. Reifler's research

Two names not mentioned in Loomis' report, nor it seems by any of his respondents, were Anthony Oettinger at Harvard University and Erwin Reifler at the University of Washington. The omission of Oettinger is understandable since he was still at an early stage of his doctorate and had not yet written any paper on MT.

The omission of Reifler's name is, however, much more surprising as Reifler had been probably the first to respond to Weaver's memorandum by starting actual MT research. His first MT study had appeared in early 1950, and his second came in April 1951, comprising a discussion of Weaver's memorandum and a brief summary of his first study.

In late 1951, Reifler approached the Rockefeller Foundation for financial support in his investigations. In a note to Weaver (4 October 1951), Loomis expressed his doubts. After reading Reifler's papers and meeting him in Seattle, he considered him "a crank". Loomis cautioned against "endless expenditure of funds for very questionable 'semantic studies'". He preferred to support just one researcher and to "back his judgment about when to do what"; his nomination was Huskey at UCLA. (As we have seen, Huskey had already received $5,000 from the Foundation.) If Huskey's work was not fruitful then the Foundation could turn to Booth or "someone at M.I.T." Loomis had acquired a poor opinion of theorists, and "[d]espite my feelings about Korzybski, somehow I find that I instinctively trust the electronics boys more than I do the semanticists".[19]

Nevertheless, Reifler did receive a grant from the Rockefeller Foundation in February 1952. The grant was for the preparation of "meticulous translations into idiomatic Mandarin, classical Chinese, Japanese, German, French, Russian, Hebrew, and so on, of a lengthy article in model English", i.e. the regularised simplified English proposed by Stuart Dodd (also at the University of Washington), at this date still unpublished (see Section 11 below.) Reifler's stated aim was to devise codes "to insure the conveyance of exact semantic meaning" in dictionaries for automatic translation.

14. Bar-Hillel appointed at MIT

In May 1951, Yehoshua Bar-Hillel took up his appointment at the Massachusetts Institute of Technology (MIT) in the Research Laboratory for Electronics. He had come to the United States on a fellowship from the Hebrew University of Jerusalem in the autumn of 1950, first to the philosophy department of the University of Chicago (Bar-Hillel 1964b), where in conversations with Rudolf Carnap he learnt of information theory, i.e. the statistical theory of communication as developed by Shannon and Weaver (1949). In March 1951 he left Chicago for Cambridge, Mass., in order to meet Willard V.Quine and Norbert Wiener – the latter's book on *Cybernetics* (Wiener 1948) had been a great influence on Bar-Hillel (1970b). Through Walter Pitts (a former student of Carnap's and, as we have seen, an influence on Weaver through his and McCulloch's automata theory) he had been introduced to Jerome B.Wiesner, an associate director of MIT's Research Laboratory for Electronics (RLE), who offered him the position "just when my fellowship money was touching bottom" (Bar-Hillel 1964b). Bar-Hillel was to spend two years and three months at RLE, which he described as "doubtless the most stimulating and creative in [his] life so far [1964]". The appointment was made with the assistance of a grant from the National Science Foundation, quite possibly with the influence of Weaver who was a director of the Foundation at this time.

Bar-Hillel's task was to study the possibilities of MT and to plan future research at MIT. The attraction of MT for Bar-Hillel was as "a testing ground for the validity of the attempts at formalizing natural languages" and "its possible effect on international communication". Less than three months after arriving he was delivering lectures on MT at an MIT Summer Course on Communication Theory, which included "a not too successful repetition" of Kaplan's experiment on reducing semantic ambiguity through context, some general ideas on formal linguistics and a review of the then meagre literature of MT (Bar-Hillel 1964b). It was at this time as well, in the autumn of 1951, that he first met Noam Chomsky, who was to remain a life-long friend and whose linguistic formalism was to have considerable influence on him. An acquaintanceship with Zellig Harris was also renewed – Bar-Hillel had met him in 1947 during a visit of Palestine by Harris – he was regularly informed about Harris' work on 'discourse analysis' which was to lead later to the concept of grammatical transformations (Harris 1954).

In October 1951 Bar-Hillel visited the few places in the US where MT research was going on. In Los Angeles he met Victor Oswald and William Bull at UCLA, and Abraham Kaplan and Olaf Helmer at RAND; in Seattle, he met Erwin Reifler; and he lectured on MT at the University of California (Berkeley), and at the University of Michigan (Ann Arbor). 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.

Bar-Hillel's activity in 1951, undoubtedly greater and more influential than those of any other person at the time, are the justification for the often repeated statement that he was "the first full-time paid research worker in the field" (Booth and Locke 1955: 5). Bar-Hillel himself, however, always stressed that he "never wrote a program for MT, never collaborated with a group that designed mechanical translators, and never induced a student to write a thesis on MT." (Bar-Hillel 1964b)

Bar-Hillel begins with a summary of the presumed future benefits of MT in satisfying demands for translations, particularly in science, finance and diplomacy; in providing means for scanning at high speed ("though perhaps low-accuracy") the huge printed output in newspapers, journals, and leaflets of actual or potential enemies; and in throwing light on aspects of linguistic communication.

He discusses first the possibility of "fully automatic MT" with no human intervention. He concluded that at the present time it would be achievable "only at the price of inaccuracy". The primary problem was that he could see no method available to eliminate "semantical ambiguities." The only approach he could envisage would require "a knowledge of the relative frequencies of all word digrams... trigrams, etc., or a knowledge of equivalent conditional frequencies of the foreign language... or else it would presuppose a "learning" organ..." It is undoubtedly indicative of the then current state of linguistics (in the United States at least, and probably in general) that in 1951 Bar-Hillel could not conceive any non-statistical approach to ambiguity resolution.

Although there was no hope of achieving "high-accuracy, fully automatic MT... in the foreseeable future" he stresses that this did not mean computers could not be applied to translation, since "with a lowering of the target, there appear less ambitious aims the achievement of which is still theoretically and practically viable." Bar-Hillel thus reviews the various options for "mixed MT, i.e. a translation process in which a human brain intervenes... either at the beginning of the translation process or the end, perhaps at both, but preferably not somewhere in the midst of it."

It was his view that a post-editor was indispensable to eliminate ambiguities, and he considers in detail how MT in this mode might work. His assumed starting point was word-for-word translation, or direct substitution of source words by target words by means of an automated dictionary. It was obviously impracticable for a post-editor to be faced with all the 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 illustrated how much easier the post-editor's task would then be.

The question was, however, how the automatic resolution of ambiguities was to be achieved. First, he reviews Kaplan's experiments in contextual disambiguation. He interprets the results not as supporting a statistical approach but that

> the ambiguity is not so much reduced by cumulative effect as by the occurrence of some one particular word which has a great chance of appearing in the given context.

For Bar-Hillel this greatly simplified the task of post-editing, and

> he should be able to produce out of the raw output of this hypothetical machine a readable translation in a fraction of the time it would take a bilingual expert to produce a translation with the conventional procedure.

Bar-Hillel then considers grammatical analysis, both "in order to eliminate grammatical ambiguities and to re-arranging the words of the FL sentence in accordance with the standard order of the TL". He identifies three stages (the italics are Bar-Hillel's):

> 1. *Mechanical analysis* of each *word* in the FL into the *stem (lexical unit)* and *morphological category...*

2. *Mechanical identification* of small syntactical units within the given sentence on the basis of the morphological categories...
3. *Transformation* of the given sentence into another that is logically equivalent to it, and *rearrangement* of the parts of the transformed sentence in accordance with some standard order of the TL.

His comments are revealing in hindsight. The challenge for the first stage was seen to be the practical difficulties of constructing large enough dictionaries and the resolution of problems of morphology which had not so far been treated in depth. The second stage required the elaboration of what Bar-Hillel referred to as an "operational syntax". Despite the mass of information about languages, there was nothing suitable for analysis by computers. An important step was the research of Oswald and Fletcher, but they had concentrated primarily on the resolution of homonyms and also, in Bar-Hillel's view, they "apparently were not sufficiently aware of the necessity of combining the different routines... into *one* sequential system", i.e. as a computer program. He did not agree with their view that syntactic analysis must await the availability of complete statistical data for morphological analysis (stage 1 in Bar-Hillel's terms) in order to overcome the limitations of computer storage. He was 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". If the machine is to do syntactic analysis the problem of storage had to be solved. However, Bar-Hillel was confident that it would be solved, and he mentioned "gravitational non-scanning devices, ... magnetic, electronic, or photo-electric devices." As for his third stage (transformation into the target language), he does not discuss it at all – either because he saw it as a trivial task once the problems of morphological and syntactic analysis had been overcome, or because he could not envisage at this date how it might operate.

He then briefly considered the role of a pre-editor as envisaged by Erwin Reifler (section 12 above). The problem with the proposals was the construction of the dictionary to be used by the pre-editor. It had to provide information, in the editor's own language, which would allow the selection of translation equivalents into an unknown target language. Bar-Hillel agreed that it might be possible where only two languages are concerned ("specific MT"), but not for "general MT, where translation from any language into any other is considered."

Bar-Hillel thought that "specific MT" could be pursued on a simple trial-and-error basis. But (in implicit agreement with Reifler), he held that "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 not encouraged by the failed attempts in previous centuries to construct universal languages, based on "metaphysical preconceptions, Aristotelian logic" and ignorance of non-European languages; but he believed that "empirical open-mindedness, mathematical logic, and modern structural linguistics" may prove a better foundation. He cited the work of Kazimierz Ajdukiewicz (1935) and Zellig Harris (1951) for "the beginnings of a universal system of syntactical categories", and the contributions of Carnap (1934), Reichenbach (1947) on the logical foundations of syntax. He thought also that "Stuart C.Dodd's system... highly interesting." This was the proposal, seen by Bar-Hillel in a 1949 mimeographed form (no doubt from Reifler), to use a restricted and regularised 'Model English' in MT systems. Dodd was to present his ideas at the first MT conference held at MIT in June 1952.

Less ambitious than universal grammar was the investigation of "*transfer-grammars*, i.e. systems in which the grammar of one language is stated in categories appropriate to some other language." For this also he mentioned the work of Zellig Harris, who had not at this time published his proposals – they were not to appear until three years later (Harris 1954). As conceived by Harris and Bar-Hillel transfer grammars were rather different from what is now understood by the term. They did not comprise sets of rules for converting structures of one language into those of another, but methods of analysing sentences of source languages into syntactic categories and representations congruent with those of the target language.

Given the problems of ambiguity and syntactic complexity, Bar-Hillel discussed 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. It might be possible in such cases to translate whole units or sentences, and "the theoretical difficulties... are clearly less formidable". But it was not only restricted languages which could aid MT, 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', combining terminological control and simplification of syntactic constructions in texts to be translated. With such voluntary restriction, the translation of technical abstracts could be facilitated into "some exchange language, natural or artificial".

In conclusion, Bar-Hillel looked briefly at the question of hardware, and in particular whether general-purpose computers or special translation machines would be more satisfactory in the long term. Since arithmetic operations were not required and economic reasons determined that MT would have to be done on a large scale, his conclusion was that special-purpose computers should be developed. Nevertheless, even though "not ideally suited for MT", preliminary experimentation should be performed on computers in their present form.

Bar-Hillel's survey, written in late 1951 (Bar-Hillel 1951), was available for the conference at MIT in June 1952, although it did not appear in print until 1953 (Bar-Hillel 1964c: 153) – confusingly, the issue of the journal in which it was published was dated 1951. It was written before any MT research on even a reasonably modest scale had begun. What is remarkable, therefore, is the perspicuity with which he identified the main problem areas (although not all, of course) and outlined many basic strategies which continue to be valid to the present day: use of post-editing, investigation of interlinguas and transfer procedures, research on morphological and syntactic analysis, and use of restricted (or controlled) languages.

15. The first MT conference, June 1952

In April 1952, the Rockefeller Foundation approved a grant to Bar-Hillel for the continuation of his appointment at MIT until June 1953, and for the organisation of the first conference on machine translation. It took place at the Massachusetts Institute of Technology from the 17th to the 20th June 1952. The original intention was to publish the proceedings verbatim but, as Bar-Hillel puts it, "after we got the mimeographed version of the stenographic record, we decided against it" – no reasons are elaborated! – "instead, in 1955, the first collection of essays on MT was published, which included most of the formal contributions". This is the collection edited by Locke and Booth (1955)[20]. However, although we do not have the original proceedings, we do have contemporary accounts of the conference by Erwin Reifler (1952c) and by A.Craig Reynolds (1954).

The participants comprised many of those mentioned already in this chronology, plus some new names from institutions known to have an interest. The MIT was represented by Jerome B.Wiesner, Yehoshua Bar-Hillel, Jay W.Forrester (Digital Computing Laboratory), James W.Perry (Center of International Studies) and William N.Locke (Department of Modern Languages), Vernon Tate (Director of Libraries), and Dudley A.Buck (a research assistant in the Electrical Engineering Department). From the Los Angeles area, the Rand Corporation was represented by Olaf Helmer, the Institute for Numerical Analysis (NBS) by Harry Huskey, and the UCLA by William E.Bull and Victor A.Oswald. From the University of Washington came Erwin Reifler and Stuart Dodd, and from England came Andrew Booth as the sole non-American participant. In addition there was Duncan Harkin (Department of Defense, Washington), who, as we have seen (section 7 above), that he had thought of MT in 1946 or 1947; A.Craig Reynolds from the IBM Endicott Laboratories); Leon Dostert from Georgetown University, invited as an expert on simultaneous translation; and Victor H.Yngve of the University of Chicago, a cosmic ray physicist.

Some MT pioneers were absent, most notably of course Warren Weaver - who, having set the MT research in motion, did not contribute anything more – with one exception: his introduction to the 1955 collection of Locke and Booth (Weaver 1955). Other absentees were Anthony Oettinger from Harvard University, Richard Richens from Cambridge, Abraham Kaplan from Rand and Stuart Fletcher from NBS. Both Kaplan and Fletcher had been short-term consultants and were no longer active in the field.[21]

The conference opened with a public session on the evening of 17th June under Wiesner's chairmanship. Bar-Hillel (1952a) 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.

The public session continued with Leon Dostert (Georgetown University), who spoke of his experience in setting up simultaneous translation services at the Nuremburg trials and at the United Nations. He was sceptical about MT, believing that there was no shortage of human translators, and that MT could only contribute in those highly specialised areas where material was not being translated at all. Then Olaf Helmer described the need to resolve the problems of idiomatic, contextual and syntactic ambiguities and reported on the tentative MT experiments which had been undertaken already at the Rand Corporation.[22] Andrew Booth discussed popular misconceptions of MT ("How intelligent can a machine translator be?"), stressing that no ‘intelligence’ is required of computers for MT purposes. James Perry (MIT) spoke about the relation between MT and automatic indexing and retrieval systems, the development of punched cards at IBM for dealing with vast amounts of information, and expressed a belief that MT was not only feasible but closer to realisation than the audience might think. The public session ended with a lively discussion reported by Reynolds (1954) in these terms:

> There was general agreement on the part of both the panel and the audience that mechanical translation was feasible. It was interesting to note that the computer engineers present presented all of the difficulties standing in the way of producing a mechanical translator from the engineering standpoint; the linguist, from his standpoint; and the psychologists and philosophers from the standpoint of their respective disciplines. Each agreed, however, that, if the other two groups did their work, we could in the near future produce adequate and intelligible machine programmed translations.

This optimism was characteristic of the whole conference despite the abundant evidence of the linguistic and computational difficulties ahead. The first session provided an opportunity for Bar-Hillel and Reifler to state their differing views of pre- and post-editing. For Bar-Hillel a post-editor would be essential to choose among alternative target language variants, and he repeated his earlier observations (Bar-Hillel 1951). “For a particular sentence, say of 10 words length, this can easily result in possible combinations of words in the target language extending to several thousands of more or less meaningful combinations” (as cited in Reynolds' report of his paper.) Bar-Hillel was concerned at the implications for storage capacity.

He illustrated his ideas on post-editing in a paper on idioms (Bar-Hillel 1952c). “One of the standard objections... often raised against the possibility of MT was its alleged inability to cope with idioms.” He saw two solutions: either to include phrases in the lexicon or to inform the post-editor 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 translation *there is* would prevent the literal translation when required, e.g. *it gives* or *she gives (*if *es* refers to *ein Mädchen)*. Bar-Hillel thought also that the phrase dictionary would have to be uneconomically large. In any case, he preferred the post-editor solution, where “the raw translation of *es gibt* would still be *it (he, she) gives*, but the English reader or, preferably, the English post-editor, would be instructed to replace, or at least to consider a possible placement of, *it (he, she) gives* by *there is (are)*.”[23] 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.

A number of participants agreed with Reifler that it would be essential to minimise the ambiguities and complexities of syntax by preparing source texts for MT, either by pre-editing, or by training authors (or “more particularly their secretaries”, as Reynolds commented) to write with MT in mind. Reifler supported therefore the proposals by Stuart Dodd for a ‘controlled language’ to be used in MT, the “standardization of English syntax as a means of simplifying the use of English either as a source language or as a target language.” Dodd (1952) proposed ten “ideal rules”: the model language should “include the most usage, i.e. ... maximize the number of persons who can use it”; it should include “the most used words” and the “most useful words”; every word should have “a fixed position (i.e. one sequence in a sentence)”, “just one meaning”, “just one grammatical form (i.e. uninflected particles)”, “just one pronunciation”, and “just one spelling (preferably phonetic)”; and ever letter should have “just one shape” and “just one sound”. As a result verb forms would be regularising (*She did be loved* instead of *She was* *loved*), there would be unique pronoun forms (nominatives: *I will send he to she*), regularised word orders in clauses (subject, verb, direct object, indirect object), adjectives and adverbs would come always before whatever they qualify, etc., and, of course, the semantic rule ("one symbol, one meaning") would be adhered to, requiring sometimes the addition of adjectives to distinguish meanings of homonyms (*tank* as a military vehicle would always have to be given as *army tank*.)

Reifler (1952a) elaborated on the pre-editing proposals which he had first proposed in 1950 (see section 11 above). It is clear that Reifler shared the view common at the time that mechanisation of translation meant only the process of one-to-one substitution of mainly dictionary forms (either full or stem entries). By contrast Bar-Hillel and Oswald already envisaged the mechanical processing of structures, the analysis of input sentences and the generation of idiomatic output.

For Reifler, then, pre-editing was essential. Its purpose was "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 have to take into account not just the multiple meanings in the source language but also the lexical and semantic differences in the target language – if semantic interpretation were omitted then a post-editor would be essential (as Bar-Hillel envisaged). But pre-editing could be reduced if "people desirous of a MT" wrote in a regularised language such as Dodd's model English, or if they inserted the codes themselves using special monolingual dictionaries with symbols for distinguishing homographs. Reifler went further to suggest – in anticipation of later procedures of interactive analysis, e.g. in the MIND system (Kay 1973) and later systems – that code insertion itself could be mechanised:

> When the pre-editor dials the conventional graphic form of the foreign message into the translation mechanism, it would first pass through the mechanical dictionary. Whenever in terms of the target language no multiple meanings are involved, the dictionary mechanism would not intervene and the dialled material would move on to the next stage in the translation process. Otherwise a device would call the attention of the pre-editor to the fact that multiple meanings are involved and the dictionary entry concerned would appear on a screen. The pre-editor would then select the meaning required by the context and dial the distinctive graphic symbol representative of this meaning and supplied by the dictionary. (Reifler 1952a)

However, Reifler thought that none of these solutions was practical: "The burden on the supply side is too great, the extent of human intervention too large." Instead he made the proposal (eccentric then as now) of 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. Reifler believed that, in conjunction with a regularisation of the target 'model' language (on the lines proposed by Dodd), it would "either restrict post-editorial interpretation to a minimum, or it may even make it completely superfluous."

For translation into many target languages ('general MT'), however, "one and the same preparation of the code text" would be necessary and sufficient for translation. Reifler (1952b) 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': "by arbitrarily attributing grammatical meanings to linguistic forms which they, in fact, do not have, namely by changing the structure of a language, we may, for instance, within the limitations of intelligibility, so modify the grammar of a language as to bring it more in line with the grammar of other languages" – an idea akin to Harris' ideas on transfer grammar (Harris 1954). For example, the Mandarin Chinese version of English *he walks quickly* is *t'a*^{1} *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"; so a literal translation might be "he is quickness of walk". But in other contexts *tsou*^{3}*-ti* can be freely translated as "walking"; therefore, Reifler argued, we could make an arbitrary equation of *-ti* and English *-ing*:

> We may therefore render the Mandarin sentence by "he walk-ing quick". This is bad English, but perfectly intelligible and, because it permits a word-to-word translation, has the great advantage of simplifying the mechanical correlation problem.

The proposal was put forward as a further elaboration of Dodd's model English. Reifler conceived the outputs of MT systems as texts in "adjusted model target languages" which readers would learn to interpret. As we have seen the idea of 'pidgin' translation (which is effectively Reifler's proposal) had been made earlier by Richens in 1949 (section 10 above).

Evidently, universals were not seen to imply intermediate languages (interlinguas), nor vice versa. Leon Dostert proposed 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" (quoted by Reifler 1952c). After the conference, Reifler (1952c) remarked that he foresaw this as an idea "which will certainly become an important feature of future MT."[24]

The conference was attracted by the ideas of Victor Oswald and William Bull for domain-specific dictionaries ("microglossaries" they called them) established on the basis of statistical frequency analyses of vocabulary. As a reminder of the state of technology at the time, Reynolds comment is illuminating: "A discussion of the means required to further extend the investigations showed clearly that the analysis could be facilitated by the use of punched cards." It seems the possibility of using computers for statistical analysis of language did not occur to participants.

Victor Oswald (1952b) proposed microglossaries as one means of overcoming the need for pre-editors, post-editors and a battery of experts in all the subjects of translated material: "an alternative arrangement is possible: to replace the battery of specialists by a series of permanent micro-glossaries, each of which would provide no more than two-to-one, and a preponderance of one-to-one, TL equivalents." 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:

> In other words, brain surgeons writing on brain surgery are not only compelled to choose their technical nouns from a limited vocabulary, but their patterns of communication are so limited by practice and convention that even the range of non-technical nouns is predictable.

Although encouraged, Oswald was cautious. "Frankly, I do not know today just how significant our findings are for MT. They indubitably indicate that micro-glossaries can be constructed... Their ultimate efficiency remains untested, however, and it is possible that it might be prohibitively expensive to produce them."

While agreeing that microglossaries could reduce problems of ambiguity, William E.Bull (1952) was more sceptical about the value of frequency analyses for constructing microglossaries:

> 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 the basic problem for all MT systems to the present day:

> The limitations of machine translation which we must face are, vocabulary-wise, the inadequacy of a closed and rigid system operating as the medium of translation with an ever-expanding, open continuum. (Bull 1952)

While semantics (or rather, 'multiple meanings') had been a concern from the beginning, syntax was something quite new for many participants. Victor Oswald (1952a) spoke about his analysis of German syntax, from which he had concluded that word-by-word translation was "literally impossible." Following his already published article (Oswald and Fletcher 1951), he proposed that structural 'blocks' could be isolated and treated as units in the rearrangement required for translation into English – in essence a constituency analysis. But for the participants even more of a revelation was Bar-Hillel's presentation of 'operational syntax' (Bar-Hillel 1952b) – an idea briefly introduced earlier in his survey paper (Bar-Hillel 1951):

> [it was] a completely new concept to the linguists of the conference who had intuitively felt that such a structure did exist but without the tools of symbolic logic had been unable to isolate the essential features that lead to the exceedingly simple arithmetic operations. The engineers immediately recognized the extreme advantages and the simplicity of the computing loops necessary to give the machine the ability to recognize word block functions and programmed reorganization of sentence structure (Reynolds 1954)

In constructing an operational syntax Bar-Hillel believed that, for German, a combination of the methods of C.V. Pollard (*A key to rapid translation of German*, University of Texas, 1947) and of Oswald and Fletcher (1951) would provide a good foundation. He envisaged a semi-automatic analysis of a text corpus for identifying grammatical categories and phrase structures. In place of Oswald's and Fletcher's rules of analysis and conversion, Bar-Hillel introduced what was later to be called categorial grammar, based on the work of the Polish logician Ajdukiewicz (1935). The concept was illustrated by an analysis of *John thought that Paul lied*. The word category for *John* was to be a simple noun *n*; *thought* could be either a noun *n* or an intransitive verb with noun to the left *s/(n)*, or a transitive verb with a noun to the left and either a noun or a phrase to the right: *s/(n)[n]* or *s/(n)[s]*; *that* could be either a noun *n* or an adjective *n/[n]* or a conjunction *n/[s]*; *Paul* was a noun *n*; and *lied* was an intransitive verb *s/(n)*. Bar-Hillel then demonstrated that simple 'cancellation' rules – the index-sequence β α/(β) is replaced by α; α/[β] β is replaced by α; and β α/(β)[ω] ω is replaced by α – could reduce the possible resolutions to just two: one with the interpretation of *that* as an adjective (contrasting with, e.g., *...this Paul...*) and one as a conjunction. The latter would be the 'normal' interpretation which would be selected by the post-editor. Bar-Hillel claimed that the same approach could deal with the nominal and verbal 'blocks' in the analyses of Oswald and Fletcher. However, he admitted that the “word-category-list” would have “for English... some million and a half entries [and] The preparation of such a list is certainly not a simple task, since all possible occurrences of these words in all kinds of syntactic construction have to be envisaged.”

Despite these reservations, the apparent simplicity of the approach appealed to the computer engineers present at the conference. The linguists, however, were quick to point out the immensity of the task of covering the whole syntax of a language; although some thought 75% or 90% coverage would suffice. In his paper, Bar-Hillel believed that “a linguist with a staff of a few assistants and clerks should be able to provide [an operational system] for any language that has already been more or less exhaustively described – like English, German, or Russian – within a year or two.” As for the extra time required for multiple analyses, Reifler wondered whether they could be done simultaneously; and Bar-Hillel agreed that it would be possible in principle, although it would mean “two or more machines working in synchronization” (an interesting anticipation of parallel processing.)

There was, in fact, almost no practical experience of programming to be reported, or indeed of using computers (a highlight of the conference was a visit to the large Whirlwind computer under construction at MIT.) Huskey outlined the memory and programming requirements for future machines, based on his experience with the SWAC computer[25]; Jay Forrester (director of the MIT computer laboratory) presented figures on the costs of magnetic drums, magnetic tapes, and electrostatic storage devices; and, in an extended account of his and Richard Richens’ experiments in England (Richens and Booth 1952), Andrew Booth compared times and costs of using punched cards and the potentially lower costs and higher speeds of using digital computers.[26]

The conference included some wishful crystal-gazing: William Locke looked forward to the development of voice input and output devices (in a paper published later, Locke 1955). And the final paper no doubt supported the aspirations of some participants that research undertaken for MT could have wider applications. In a talk on “Teaching foreign languages” William Bull showed that concepts from ‘operational syntax’ could aid language learners. One example (quoted by Reifler, 1952c) was the distinction in Spanish between *estar* and *haber*, both possible translations of English *to be*: “... even the educated native does not know what determines his choice... The problem, however, can be solved both for the machine and the student by isolating the fact that “the” in English takes *estar* and “a” takes *haber*.”

At the end of the conference on June 20th, participants were invited to outline their future research plans. The computer engineers were optimistic and were keen to seek financial support for building MT systems; and Duncan Harkin of the Department of Defense promised that backing would be forthcoming for “a concrete proposal for such a translation and subsequent demonstration” (reported by Reynolds 1954). Leon Dostert, who had now been converted from his initial scepticism, 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” (Reifler 1952c). Various language pairs were proposed: most participants believed that quick results could be obtained with translation between German and English, although Reifler (1952c) believed that “both in morphology and syntax Chinese and English happen to have more in common than German (or any other European language) and English.” However, it was conceded that “certain government agencies may be readier to supply the funds necessary for further research and improvements if the first pilot machine is designed for mechanical translation from Russian into English.”

Some years later, Bar-Hillel (1964b) regarded this as "one of the most exciting conferences I had ever participated in, and I shall never forget the pervasive feeling of euphoria which I felt and thought that everybody else shared." The important contribution of Bar-Hillel himself was well expressed by Reifler (1952c):

> There can be no doubt that much of the success of the Conference was due to Dr.Bar-Hillel's efforts, and it is, I believe, no overstatement to say that MT, if and when it materializes, will be very much indebted to him.

It is a sentiment which nobody would dispute to the present day.

16. Research continues: 1952

Within the next eighteen months, MT activity in the United States witnessed a marked acceleration. In July 1952, Huskey's group received a further grant from the Rockefeller Foundation. This time it was to support the work of Kenneth E.Harper (Department of Slavic Languages, UCLA) on investigations of Russian syntax, along much the same lines as the research of Oswald and Bull on German and Spanish respectively. Harper's report was to appear the next year.

On the 4th September Bar-Hillel was invited to speak at the 7th International Congress of Linguists in London, his participation being supported by a grant from the Rockefeller Foundation (Bar-Hillel 1964b). The presentation would appear to have been made at the last minute – itself an indication of the growing public interest in MT – since Bar-Hillel's paper is not recorded in the published proceedings (Norman and Ganz 1956), even as an abstract. It may be assumed, however, that it would have been based on his 1951 survey and the outcomes of the June conference, and was probably on the lines of an article published the next year (Bar-Hillel 1953b).

Reifler continued to produce further MT studies during 1952. His first in July was a report of the MIT conference (Reifler 1952c) – from which quotations have been made above. The second in August was a further discussion of pre-editing, overturning his opinion at the MIT conference. He argued now that a pre-editor was not essential for translating German into English (Reifler 1952d), summarising arguments to be elaborated in the next papers. In one of these, his third study completed in September (Reifler 1952e), he discussed the treatment of German noun compounds and how their constituents might be identified in an automatic dictionary. Reifler had taken up the challenge of Oswald, who had stated at the conference: "We know of no mechanical process by which this could be accomplished..." (Oswald 1952b). Reifler proposed "a very simple mechanical solution" comprising a 'capital memory' – a separate store or file – of German nouns and compounds whose meaning cannot be inferred from their constituents. Other compounds would be dissected by the recognition of linkage patterns which Reifler had identified. A fourth study appeared in October 1952 on the use of grammatical 'pinpointers' and 'operational form classes' to aid the disambiguation of individual words (Reifler 1952f). For example, a genitive noun would be 'pinpointed' by the nominative noun it modifies (the 'pinpointer') – in effect, Reifler was proposing a kind of dependency grammar – and, in defining such dependencies ('pinpointing') Reifler found the need to establish 'form classes' distinct from traditional grammatical categories.

In September 1952 appeared the first report of an MT experiment at MIT. James Perry (1952a) did a manual simulation of word-for-word translation of a Russian text into English. 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 establish this point, two of the writer's assistants who had had training in physics and chemistry were requested to edit the rough translation produced by simulated machine operations so as to indicate how they would interpret its meaning." For the above sentence, the following was produced:

> 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 that improvements would be possible if Russian grammar were taken into account, Perry produced a follow-up report in November (Perry 1952b) putting forward a method of dictionary look-up which searched for the longest stem match and then identified the ending and its grammatical function. As before, this was a simulation experiment.

These reports by Reifler and Perry were, it must be remembered, the first investigations in some detail for the MT problems they were addressing. Today they may appear simplistic and obvious, perhaps even naive and trivial; but in this pioneer period of MT they represented substantial contributions. What should be stressed once more is that researchers had no access to computers, or even in some cases to punched card equipment; they had no idea whether their methods could be realised in practice; they tried to be as 'mechanistic' as possible, but their work could be no more than theoretical.

On 5th October 1952 there was a further newspaper report on the SWAC computer. As in 1949, the *New York Times* reported the imminent appearance of a translating machine ("Bilingual Machine Electronic Device Will Translate Foreign-Language Data"):

> Dr.Harry Huskey, Victor Oswald and William Bull of the University of California at Los Angeles have come forth with a machine that will translate...
>
> Don't expect a literary production from the translating machine, however. The California professors have in mind no more than a rapid exchange of information between scientists of different nationalities. Facts, not style, are the main consideration.

The reporter described the prospects as follows:

> In any cybernetic machine the problem must be stated on perforated tape by some competent person... In the case of the translating machine the problem can be stated only if there is a sound linguistic theory. Dr.Oswald says the linguistic theory is ready, and Dr.Huskey adds, that the Institute of Numerical Analysis... can set up any electronic apparatus to meet linguistic requirements.
>
> An actual demonstration can be made when special scientific glossaries have been prepared and the problems of making them mechanically accessible has been solved.

As usual, it seemed to be just a matter of time – and not long either! The researchers appeared to be as optimistic about the imminence of MT as the reporters were. Two months later, MT was the subject of an item in a widely read scientific journal. Under the title "Lingua ex machina" the December 1952 issue of *Industrial and Engineering Chemistry* (vol.44 no.12, p.11A and 13A) reported 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." It went on to summarise the paper-slip simulation of Russian translation by James Perry at MIT (Perry 1952a), citing some of the crude results:

> Saccharification cellulose begin use (verb) in technology. For that waste product wood working plant heat (verb) under pressure with 0.1% solution sulfuric acid; obtained such means syrup process (verb) in wine alcohol.

which the reporter thought was "not unmeaningful to an individual having some knowledge of the science." There was some natural caution:

> From the above cited example it is apparent that translations in elegant style can scarcely be expected. Equipment of great complexity and cost would be needed to cope with all the difficulties of case, gender, and the like. Literary versions, however, are not needed. The standards for a satisfactory translation depend on the use to which it is put... A rough approximation of the original will suffice... for a person conversant with a particular subject to select from a mass of potentially available references those of actual value...

A major impediment was the painfully slow input of the documents for translation, and the report stressed the need for a scanner consisting of "a photographic or television device able to distinguish each printed letter by form and supply the information to the machine's magnetic memory." The reporter stressed that MT itself did not yet exist: "no computor [sic] has been specially designed for the purpose, although the Bureau of Standards' studies [by Oswald and Fletcher] have been carried out along lines applicable to its SWAC instrument at Los Angeles"; and he was aware also that Richens and Booth had "applied translation by semantic units to such strange languages as Indonesian, Japanese and Arabic."

The reporter assessed future possibilities in these words:

> In its simplest form such a device would serve as a super mechanical dictionary in aiding a translator in the more routine phases of his practice. More highly developed, it would act as a means by which foreign texts could be rapidly digested and roughly evaluated. Another possibility: a tool for use in coding and indexing the content of foreign publications, in a form immediately suitable for high speed electronic searching.

In all, this was a realistic and balanced assessment of the future potential of MT, stressing the limitations but also pointing to useful applications, in a manner not always followed by reports in subsequent years. However, the reporter (signing himself D.M.K.) could not refrain from an amusing coda, alluding to the current Cold War witch-hunting mentality in parts of the United States. It is worthy to be quoted in full:

> 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.

Little was he to suspect that some 20 years later, Colby was to write a computer program PARRY, which simulated the behaviour of a paranoid patient. (Colby 1975)

For some time, the SWAC research appeared to be the only implementation of a MT system[27], although in fact from all the reports it is clear that the MT application was only in its planning stages. Nevertheless, it was widely reported as fact. For example, Albert Ducrocq, in his book *L'ère des robots* (Ducrocq 1953: 227-258), describes *le langage binaire* of computers and its universal applicability as the future language of communication. He considered that one such application was MT, particularly for scientific and technical texts where he argued the complexities of faithful communication are less severe (and even, Ducrocq believed, for economic and political documents). As an illustration he was able to cite the SWAC and UCLA experiments.[28] This publication was quite probably the first in any language other than English to mention MT – apart, of course, from the patents of Artsrouni and Troyanskii, which were still unknown at this time. It was certainly the first in French known to Mounin (1964: 22), and there were definitely no Russian publications on MT as yet.[29]

In February 1953 the Rockefeller Foundation showed its confidence in Huskey's group by approving a further extension of its grant to December 1954; it noted, however, that progress had been delayed "because of a scarcity of mathematicians" – presumably to program the SWAC for linguistic work – but it believed that the theoretical investigations of Oswald and Bull were showing "considerable promise" and deserved continued backing.[30]

17. Progress in 1953

During 1953, the volume of articles about MT showed a marked increase. Yehoshua Bar-Hillel published three articles during the year. In *Language* appeared the elaboration of his proposed categorial syntax (Bar-Hillel 1953a), which had been presented in briefer form at the 1952 MIT conference. In the *Philosophy of Science* he wrote an article on "Some linguistic problems connected with machine translation" (Bar-Hillel 1953b). This article (possibly the one given at the Congress of Linguists in London the previous year) was divided into four parts. The first outlined again his ideas on an 'operational syntax', with the comments that "operational syntax for any or all languages is... a task which should prove highly rewarding even for the most theoretically minded linguist." The second treated the philosophical question of the 'intertranslatability of natural languages.' The third section discussed the problem of 'idioms' in a revised version of his 1952 conference paper. The final part dealt with the question of 'universal syntactic categories', not whether in theory there are any, but how useful it might be in practice to apply them. The third article, in *Computers and Automation* (Bar-Hillel 1953c), was a summary of his 1951 overview with an outline of the contributions at the 1952 conference and summaries of publications since then.[31]

Andrew Booth also contributed to the latter journal in this year. His article (Booth 1953) concentrated primarily on the problems of storing and searching dictionaries, suggesting the division of words into stems and endings, and briefly sketching his and Richens' approach. In this year, Booth was also the first to bring MT to the general attention of computer scientists in a chapter ("Some applications of computing machines") of a book written with his wife, the former Kathleen Britten (Booth & Booth 1953).

By the end of the year, two articles also appeared from the UCLA researchers in the same issue of *Modern Language Forum*. Victor Oswald published a paper on microglossaries for MT in collaboration with Richard H.Lawson (Oswald and Lawson 1953). This was an elaboration of the report given by Oswald at the MIT conference on the study of the German vocabulary in the field of brain surgery, containing the optimistic statement that "the major linguistic problems of pragmatic mechanical translation can now be regarded as solved." The second paper was by Kenneth Harper, who wrote on the translation of Russian (Harper 1953), stressing the importance of dealing with morphology in such a highly inflected language. Syntax was considered less important because "the basic sentence structure of scientific Russian is quite similar to that of English... The necessity for rearranging the Russian sentence into a rigidly fixed English pattern seems doubtful... [it] will be intelligible... when the word order is left undisturbed.” By contrast, it was “imperative” to solve the problem of Russian morphology, and Harper undertook detailed basic examinations of noun and verb paradigms for coding in an MT program. Finally, he illustrated an “intelligible” (simulated) word-for-word translation:

> Essential (by) characteristic of segment is its length. Recall briefly, what/that understand under this term. Take any arbitrary, but definite segment in quality of unity of length and this (by) segment measure given segment. (by) Length of given segment is called number, indicating, how many time segment, taken for unity, is contained in given segment. It can, of course, to happen, what/that segment-unity not is included whole number of time, in given. Then length given (of) segment (it) will be not whole number, but fractional or irrational; last event, proper speaking, and is most common...

Unpublished reports continued to appear as well. A new entrant was Luitgard Wundheiler of the Illinois Institute of Technology, who advocated an ‘invariant’ syntax for languages, based on the work of logicians such as Ajdukiewicz, which would enable ‘normalised’ language to be translated with the aid of a dictionary alone (Wundheiler 1953). Another name relatively unfamiliar before this date was Anthony Oettinger from Harvard University. In February 1953 he produced a ‘progress report’ (Oettinger 1953) on his doctorate study for the automation of a Russian-English dictionary; his thesis (Oettinger 1954) was to be approved in April 1954 (the first in the field of MT), after which his work received wider circulation (e.g. Oettinger 1955).

In March 1953 Reifler applied for financial support from the Rockefeller Foundation “to begin the instrumental development of a mechanical translator.” However, the Foundation declined – perhaps Loomis’ earlier opinion of Reifler (section 14 above) had some influence – although it did agree to help him conclude the study based on translating an article in Model English into a number of languages, for which he had received a grant in February 1952.[32]

Bar-Hillel ended his two-year research post at MIT in July and returned to Jerusalem. He was replaced by Victor Yngve (one of the participants in the MIT conference), who proceeded to set up the MT research project in the Research Laboratory of Electronics funded by the National Science Foundation. The first quarterly progress report appeared in October 1953 – republished two years later (Yngve 1955). In it, Yngve described a simulated word-for-word translation from German into English, in which articles and other function words were left untranslated. The intention was to test what could be achieved from such crude beginnings. 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 PROPOSITIONen 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.

He reported that “Those who knew no German at all were able to grasp only the subject matter from the translated stems. They were generally unable to get much idea of just what was being said about the subject matter. On the other hand, people who knew a little German grammar, after they had recovered from their mirth, demonstrated that they were able to understand quite well and fairly rapidly what was being said.”

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 (Hutchins 1986: 87-92).

18. The first demonstration, January 1954

One of the most significant outcomes of the MIT conference in June 1952 was the decision by Leon Dostert to start work on a program to demonstrate the feasibility of MT. He had left the conference convinced that "rather than attempt to resolve theoretically a rather vast segment of the problem, it would be more fruitful to make an actual experiment, limited in scope but significant in terms of broader implications". He identified "the primary problem [as] one of linguistic analysis, leading to the formulation in mechanical terms of the bilingual transfer operations, lexical or syntactic." The aim was a system requiring no pre-editing of the input, and producing "clear, complete statements in intelligible language at the output", although "certain stylistic revisions may...be required..., just as when the translation is done by human beings." (Dostert 1955)

On his return to Georgetown University, Dostert appointed a linguist with knowledge of Russian, and established links with Cuthbert Hurd of the IBM Corporation. The two principal researchers were Paul L.Garvin at Georgetown, who did the linguistic work, and Peter Sheridan at IBM, who did the programming. Together they developed a small-scale system for translating some Russian sentences into English. On 7th January 1954 the demonstration took place at the New York headquarters of IBM. It was to strike the imagination of the invited journalists and have a profound impact, possibly the most influential publicity that MT has ever received, and was to lead ultimately in subsequent years to the substantial funding of MT research in the United States.[33]

The next day (8th January 1954), the *New York Times* carried a front-page report. The impact of this account was undoubtedly considerable.[34] The MT demonstration was reported as perhaps "the cumulation of centuries of search by scholars for 'a mechanical translator'." It stressed that it had been accomplished on "a standard commercial model of the largest International Business Machines "stock" computer. This device, called the IBM Type 701 Electronic Data Processing Machine, was put on the market last April." The report included a detailed account of how the sentences for translation were punched onto IBM 80-column cards, stressing that the operator did not know Russian and that the results were printed out "almost simultaneously":

> Several short messages, within the 250-word range of the device, were tried. Included were brief statements in Russian about politics, law, mathematics, chemistry, metallurgy, communications and military affairs. The sentences were turned into good English without human intervention.

The reporter (Robert K.Plumb) rightly stressed the limited nature of the demonstration (the restriction to a 250-word vocabulary and six grammar rules), but he attempted to describe for the general reader how syntactic changes and lexical transfer worked:

> In translating, for instance, a word "A" which precedes a word "B" in Russian, may be reversed in some cases in English. Each of the 250 words is coded for this inversion. Sometimes words must be inserted in the English text, sometimes they must be omitted, following code instructions.
>
> When there are several possible English meanings for a Russian word, the instructions tell the machine to pick out the meaning that best fits the context.

Having stressed that they had used a general-purpose computer (the IBM 701) and not some special machine for translation only, Hurd was reported as saying that "the corporation would now design a machine particularly fit for translating rather than for general computing utility." The prediction made (as were many to be during the 1950s and 1960s when forecasting the future of MT) was 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." In conclusion, optimistic predictions were made that other languages would be easier than Russian, and the dictionary data (i.e. punched cards) for German, French and "other Slavic, Germanic and Romance languages can be set up at will."

Other reports appeared in US newspapers. In the *New York Herald Tribune* (8th January) the front-page report by Earl Ubell ("It's all done by machine") marvelled at the achievement[35]:

> A huge electronic "brain" with a 250-word vocabulary translated mouth-filling Russian sentences yesterday into simple English in less than ten seconds.
>
> As lights flashed and motors whirred inside the "brain" the instrument's automatic type-writer swiftly translated statements on politics, law, science and military affairs. Once the Russian words were fed to the machine no human mind intervened.

The report recorded predictions of machines "freely translating all languages" within a few years. However, it too stressed the vocabulary limitations and also that "before the machine can work on the Russian sentences they have to be encoded on punched cards... This slows up translation." However, Dostert predicted "automatic text-reading machines", also within three to five years. As in the *New York Times*, the reporter made an effort to convey how the translation was achieved by the "collaboration among engineers, mathematical logicians and linguists." He described how the information on punched cards was stored as "electrical impulses on a magnetic drum" and how the syntax rules were derived. A further prediction from Dostert was that "100 rules would be needed to govern 20,000 words for free translation."

The report in the *Los Angeles Times* on the same day (8th January) was much less informative than those in the New York newspapers. The reporter noted the researchers' stress on the limitations of a 250 vocabulary, but still expressed the expectation that "with improvements to come, the device could run the gamut of all language dictionaries." The report in the *Christian Science Monitor*, which appeared a few days later (11th January 1954), was probably a faithful reflection of the exaggerated popular expectations of computers at the time.[36] Like the New York reporters, much was made of the speed and intelligence of the computer:

> The girl who operated 701 did not understand a word of Soviet speech and yet more than 60 Soviet sentences were given to the "brain" which translated smoothly at the rate of about 2½ lines a second...
>
> The "brain" didn't even strain its superlative versatility and flicked out its interpretation with a nonchalant attitude of assumed intellectual achievement.

Although stressing that these were early days – the demonstration was only "a Kitty Hawk of electronic translation" – the report continued with optimistic statements from Dostert:

> "Those in charge of this experiment... now consider it to be definitely established that meaning conversion through electronic language translation is feasible." Although he emphasised it is not yet possible "to insert a Russian book at one end and come out with an English book at the other", the professor forecast that "five, perhaps three, years hence, interlingual meaning conversion by electronic process in important functional areas of several languages may well be an accomplished fact."

More realistically (as it has subsequently emerged) the prediction was made (as reported also in the *New York Herald Tribune*) “that the day will arrive when a simpler and cheaper machine will be available for less than the present $500,000 supercalculator...”

Two days later (13th January 1954), the *Christian Science Monitor* carried an editorial on the demonstration – one of the very few editorial statements about MT ever – which puts it into context. While the two reports of the demonstration emphasised the imminence of translation machines, the editorial gave a more sober (and, in hindsight, more realistic) assessment:

> Such an accomplishment, of course, is far from encompassing the several hundred thousand words which constitute a language. And with all the preparations for coping with syntax, one wonders if the results will not sometimes suggest the stiffness of the starch mentioned in one of the sentences as being produced by mechanical methods. Nevertheless, anything which gives promise of melting some of the difficulty which writers and speakers of different languages encounter in understanding each other – particularly as between English and Russian today – is certainly welcome.

A concern with costs was the main theme of the report in the popular magazine *Newsweek* (18th January 1954, p.83):

> Obviously it was a tour de force for a $500,000 machine to spend its energies translating selected sentences. And it would be equally wasteful to assign any of the other 100-or-so big general-purpose computers to simple linguistic tasks. But computer experts at the National Bureau of Standards, the Massachusetts Institute of Technology, and several other centers, as well as IBM, have started to design machines specifically equipped for translating at superhuman rates of speed.

There was concern also about the slow speed, attributed to the time required for reading through punched cards and so Dostert's prediction of “automatic text-reading machines” in the near future, was encouraging. But, this report was the only one for the general public which mentioned other MT research groups – even if mistakenly it believed that the future lay with special-purpose machines.

The February 1954 issue of *Computers and Automation* expressed perhaps the greatest enthusiasm for future possibilities within a more sober assessment of what had actually been achieved. Neil Macdonald reported:

> Many exciting possible developments are indicated by the success of the trial... Linguists will be able to study a language in the way that a physicist studies material in physics, with very few human prejudices and preconceptions... The technical literature of Germany, Russia, France, and the English-speaking countries will be made available to scientists of other countries as it emerges from the presses... But of course, it must be emphasized that a vast amount of work is still needed, to render mechanically translatable more languages and wider areas of a language. For 250 words and 6 syntactical structures are simply a “Kitty Hawk” flight. (Macdonald 1954)

Some of the technical problems were well covered in Macdonald’s report, and also in a review of MT activity by Booth in the July issue of *Discovery* (Booth 1954), which includes photographs taken at the demonstration. However, for much more detail of the technical side we must go to the fascinating account given by Peter Sheridan (1955). As the first substantial attempt at non-numerical programming, every aspect of the process had involved entering quite unknown territory. Decisions had to be made on how alphabetic characters were to be coded, how the Russian letters were to be transliterated, how the Russian vocabulary was to be stored on the magnetic drum, how the ‘syntactic’ codes were to operate and how they were to be stored, how much information was to go on each punched card, etc. Detailed flow charts were drawn up for what today would be simple and straightforward operations, such as the identification of words and their matching against dictionary entries.

Details of the linguistic aspects of the experiment were given in a contemporary account by Dostert (1955) and at a later date by Garvin (1967) and Zarechnak (1979), and have been summarised in Hutchins (1994, 2004b). The limitations and artificiality of the experiment were freely admitted by the Georgetown researchers, but the general public had been impressed: MT was now seen as a feasible objective, and 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, the first ones at the Steklov Mathematical Institute and at the Institute of Precision Mechanics and Computer Technology (Panov et al. 1956).

The true significance of the demonstration was that it was an actual implementation of MT. All previous work on MT had been theoretical in the sense that none of the proposals had in fact been implemented as computer programs. Even the SWAC demonstrations of 1949 and 1952 had not shown the computer actually producing translations; Huskey and his colleagues had merely asserted that MT was possible and shown the machine on which they believed it could be done. Considering the state of the art of electronic computation at the time (as seen in any contemporary publications, e.g. in journals such as *Computers and Automation*), it is remarkable that anything resembling automatic translation could be achieved at all. Whatever its limitations, the demonstration in January 1954 marked the definitive beginning of MT as a reality and as a research field worthy of support. Just seven years after Weaver had first proposed the use of the newly invented computers to translate, there was now some concrete evidence that MT was indeed possible.

19. References

Ajdukiewicz, Kazimierz: 1935, Die syntaktische Konnexität, *Studia Philosophica* (Lvov) 1, 1-27. Transl. as 'Syntactic connexion' in: McCall, Storrs (ed.) *Polish logic 1920-1939,* Oxford, Clarendon Press, 1967, 207-231.

Augarten, Stan: 1984 *Bit by bit: an illustrated history of computers*, London, George Allen & Unwin.

Bar-Hillel, Yehoshua: 1951, The present state of research on mechanical translation, A*merican Documentation* 2 (4), 229-237. Repr. as Bar-Hillel (1964c).

Bar-Hillel, Yehoshua: 1952a, Mechanical translation: needs and possibilities. [Cambridge, Mass.: Research Laboratory of Electronics.] 6 pp.

Bar-Hillel, Yehoshua: 1952b, Operational grammar. [Cambridge, Mass.: Research Laboratory of Electronics.] 9 pp., with 2 pp. of discussion.

Bar-Hillel, Yehoshua: 1952c, The treatment of 'idioms' by a translating machine. [Cambridge, Mass.: Research Laboratory of Electronics.] 5 pp., with 3 pp. of discussion.

Bar-Hillel, Yehoshua: 1953a, A quasi-arithmetical notation for syntactic description, *Language* 29, 47-58. Repr. in: Bar- Hillel (1964a), 61-74

Bar-Hillel, Yehoshua: 1953b, Some linguistic problems connected with machine translation, *Philosophy of Science* 20, 217-225. Sections 1, 3, and 4 repr. in: Bar-Hillel (1970a), 308-316; section 2 repr. in: Bar-Hillel (1964a), 56-58

Bar-Hillel, Yehoshua: 1953c, Machine translation, *Computers and Automation* 2 (5), 1-6

Bar-Hillel, Yehoshua: 1954, Can translation be mechanized? *American Scientist* 42, 248-260.

Bar-Hillel, Yehoshua: 1955, Idioms. In: Locke and Booth (1955),183-193

Bar-Hillel, Yehoshua: 1964a, *Language and information*, Reading, Mass., Addison-Wesley.

Bar-Hillel, Yehoshua: 1964b, Introduction. To: Bar-Hillel (1964a), 1-16.

Bar-Hillel, Yehoshua: 1964c, The state of machine translation in 1951. In: Bar-Hillel (1964a), 153-165.

Bar-Hillel, Yehoshua: 1970a, *Aspects of language*: essays and lectures on philosophy of language, linguistic philosophy and methodology of linguistics. Amsterdam: North-Holland.

Bar-Hillel, Yehoshua: 1970b, Cybernetics and linguistics. In: Bar-Hillel (1970a), 289-301.

Bel'skaya, I.K., Korolev, L.N. and Panov, D.Y., eds., 1959, *Perevodnaya mashina P.P.Troyanskogo: sbornik materialov... predlozhennoi P.P.Troyanskim v 1933 g.* Moskva: Izd.Akad.Nauk SSSR.

Booth, Andrew D.: 1946, Report on visit to American laboratories. 3rd July 1946 (Rockefeller Foundation Archives)

Booth, Andrew D.: 1948, Proposals for the establishment of a centre for the construction and use of an electronic computer, February 12, 1948. (Rockefeller Foundation Archives)

Booth, Andrew D.: 1953, Mechanical translation, *Computers and Automation* 2(4), 6-8.

Booth, Andrew D. and Booth, Kathleen H.V.: 1953, *Automatic digital calculators*, London, Butterworth.

Booth, Andrew D.: 1954, Calculating machines and mechanical translation, *Discovery* 15 (7), 280-285.

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

Booth, Andrew D.: 1955, Storage devices. In: Locke and Booth (1955), 119-123.

Booth, Andrew D.: 1956, The nature of a translating machine, *Engineering* 182, 302-304.

Booth, Andrew D.: 1959, Mechanical translation and its implications in data processing, *Automatic Data Processing* 1, 18-23

Booth, Andrew D., ed.: 1967, *Machine translation*, Amsterdam, North-Holland.

Booth, Andrew D.: 1980, Computers in the University of London, 1945-1962. In: Metropolis (1980), 551-561

Booth, Andrew D.: 1985, Letter to W.J.Hutchins, 16 May 1985.

Booth, Andrew D.: 1997, An autobiographical sketch, *IEEE Annals of the History of Computing* 19(4), 57-63.

Bruderer, Herbert E.: 1982, Einleitung. In: Bruderer, H.E., ed. *Automatische Sprachübersetzung* (Darmstadt: Wissenschaftliche Buchgesellschaft), 1-6.

Bull, William E.: 1952, Problems of vocabulary frequency and distribution. [UCLA]. 8pp.

Carnap, Rudolf: 1934, *Die logische Syntax der Sprache,* Wien, Springer. [Translated as: *The logical syntax of language*, London, Kegan Paul, 1937.]

Colby, Kenneth: 1975, *Artificial paranoia*, New York, Pergamon Press.

Corbé, Michel: 1960, La machine à traduire française aura bientôt trente ans, *Automatisme* 5(3), 87-91.

Delavenay, Emile: 1960, *An introduction to machine translation*, London, Thames and Hudson. [Translation of *La machine à traduire*, Paris, P.U.F., 1959.]

Dodd, Stuart C.: 1952, Model English for mechanical translation: an example of a national language regularized for electronic translators. [Seattle: Washington Public Opinion Laboratory.] 7pp. Repr. as 'Model English' in: Locke and Booth (1955), 165-173.

Dostert, Léon: 1955, The Georgetown-I.B.M. experiment. In: Locke and Booth (1955), 124-135.

Ducrocq, Albert: 1953, *L'ère des robots,* Paris, Juillard.

Fairthorne, Robert A.: 1952, Automata and information, *Journal of Documentation* 8 (3), 164-172

Garvin, Paul: 1967, The Georgetown-IBM experiment of 1954: an evaluation in retrospect. In: *Papers in linguistics in honor of Léon Dostert* (The Hague: Mouton, 1967), 46-56; repr. in his *On machine translation* (The Hague: Mouton, 1972), 51-64

Harkin, Duncan: 1957, The history of word counts, *Babel* 3(3), 113-124.

Harper, Kenneth E.: 1953, The mechanical translation of Russian: preliminary report, *Modern Language Forum* 38 (3-4), 12-29. Rev. repr.in: Locke and Booth (1955), 66-85

Harris, Zellig S.: 1951, *Methods in structural linguistics*, Chicago, University of Chicago Press.

Harris, Zellig S.: 1954, Transfer grammar, *International Journal of American Linguistics* 20 (4), 259-270.

Holmström, J. E.: 1951, *Report on interlingual scientific and technical dictionaries*, Paris, Unesco. (Unesco Publication 884)

Huskey, Harry D.: 1980, The SWAC: the National Bureau of Standards Western Automatic Computer. In: Metropolis (1980), 419-431

Huskey, Harry D.: 1984, From ACE to the G-15, *Annals of the History of Computing* 6 (4), 350-371.

Hutchins, W.John: 1986, *Machine translation: past, present, future,* Chichester, Ellis Horwood, 1986.

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

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

Hutchins, W.John and Lovtskii, Evgenii: 2000, Petr Petrovich Troyanskii (1894-1950): a forgotten pioneer of mechanical translation, *Machine Translation* 15 (3), 187-221.

Hutchins, W.John: 2004a, Two precursors of machine translation: Artsrouni and Trojanskij, *International Journal of Translation* 16 (1), 11-31

Hutchins, W.John: 2004b, The Georgetown-IBM experiment demonstrated in January 1954, *Machine translation: from real users to research: 6th conference of the Association for Machine Translation in the Americas, AMTA 2004, Washington, DC, USA, September 28 – October 2, 2004.* Proceedings [ed.] Robert E.frederking [and] Kathryn B.Taylor. (Heidelberg: Springer, 2004), 102-114.

Kaplan, Abraham: 1950, An experimental study of ambiguity and context (Report P-187), Santa Monica, The RAND Corporation, 1950. 18pp. Repr. in: *Mechanical Translation* 2 (2), 1955, 39-46.

Kay, Martin: 1973, The MIND system. In: Rustin, R.ed. *Natural language processing* (New York: Algorithmics Press), 155-188

Korzybski, Alfred: 1958, *Science and sanity,* 4th ed., Lakeville, KT., International Non-Aristotelian Library.

Lavington, S.H.: 1980, *Early British computers*. Manchester: University Press.

Locke, William N.: 1955, Speech input. In: Locke and Booth (1955), 104-118.

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

Locke, William N. and Yngve, Victor H.: 1958, Research in translation by machine at M.I.T. In: *Proceedings of the Eighth International Congress of Linguists* (Oslo: Univ.Press), 510-514.

Locke, William N.: 1975, Machine translation. In: Kent, A. et al., eds. *Encyclopedia of library and information science* (New York: Dekker), vol.16, 414-444

McCulloch, Warren S. and Pitts, Walter: 1943, A logical calculus of the ideas immanent in nervous activity, *Bulletin of Mathematical Biophysics* 5, 115-133.

Macdonald, Neil: 1954, Language translation by machine - a report of the first successful trial, *Computers and Automation* 3 (2), 6-10.

Masterman, Margaret and Kay, Martin: 1960, Mechanical pidgin translation (ML 133) Cambridge Language Research Unit. Rev.repr. by Margaret Masterman in: Booth (1967), 195-227.

Mel'chuk, Igor A. and Ravich, R.D.: 1967, *Avtomaticheskii perevod 1949-1963: kritiko-bibliograficheskii spravochnik*, Moskva, Vsesoyuznii Institut Nauchnoi i Tekhnicheskoi Informatsii, Institut Yazykoznaniya.

Metropolis, N. et al., eds.: 1980, *A history of computing in the twentieth century: a collection of essays*, New York, Academic Press, 1980

Mounin, Georges: 1964, *La machine à traduire: histoire des problèmes linguistiques,* The Hague, Mouton, 1964.

Norman, Frank, and Ganz, Peter F., eds.: 1956, *Proceedings of the Seventh International Congress of Linguists*, London, 1-6 September 1952. London: C.I.P.L.

Oettinger, Anthony G.: 1953, A study for the design of an automatic dictionary. Cambridge, Mass.: Harvard University Computation Laboratory. 31pp. (Progress report 26.)

Oettinger, Anthony G.: 1954, *A study for the design of an automatic dictionary*. Doctoral thesis, Harvard University, April 1954.

Oettinger, Anthony G.: 1955, The design of an automatic Russian-English technical dictionary. In: Locke and Booth (1955), 47-65.

Ogden, Charles Kay: 1930, *Basic English: a general introduction with rules and grammar*, London, Kegan Paul, Trench, Trubner.

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

Oswald, Victor A.: 1952a, Word-by-word translation. [UCLA]. 7pp.

Oswald, Victor A.: 1952b, Microsemantics. [UCLA] 10pp.

Oswald, Victor A. and Lawson, Richard H.: 1953, An idioglossary for mechanical translation, *Modern Language Forum* 38 (3/4), 1-11.

Panov, D.Y., Lapounov, A.A. and Mukhin, I.S.: 1956, *Avtomatizatsiya perevoda s odnogo yazyka na drugoi*. Moskva: Izd. Akademii Nauk SSSR. (Sessiya po Nauchnym Problemam Avtomatizatsii Proizvodstva: Plenarnoe zasedanie.)

Perry, James W.: 1952a, Machine translation of Russian technical literature: notes on preliminary experiments. [Cambridge, Mass.: M.I.T.] 16pp. Repr. as 'Translation of Russian technical literature by machine: notes on preliminary experiments' in: *Mechanical Translation* 2 (1), 1955, 15-24.

Perry, James W.: 1952b, Machine translation of Russian technical literature: notes on exploitation of the Russian grammar. [Cambridge, Mass.: M.I.T.] 11pp.

Reichenbach, Hans: 1947, *Elements of symbolic logic*, New York, Macmillan.

Reifler, Erwin: 1948, The Chinese language in the light of comparative semantics (a new approach in historical linguistics), *Science* 107 (4 June 1948), 586
Reifler, Erwin: 1950, MT: studies in mechanical translation, no.1. [Seattle: University of Washington.] 51pp.
Reifler, Erwin: 1952a, MT with a pre-editor and writing for MT (Studies in Mechanical Translation no.3). [Seattle: University of Washington.] 16pp.
Reifler, Erwin: 1952b, General MT and universal grammar (Studies in Mechanical Translation no.4). [Seattle: University of Washington.] 6pp.
Reifler, Erwin: 1952c, Report on the first conference on mechanical translation at the M.I.T., Cambridge, Mass., June 17-20, 1952 (Studies in Mechanical Translation no.5). [Seattle: University of Washington.] 22pp. Repr. with minor revisions as 'The first conference on mechanical translation' in *Mechanical Translation* 1(2), 1954, 23-32.
Reifler, Erwin: 1952d, Report on research results for the summer quarter 1952. (Studies in Mechanical Translation no.6). [Seattle: University of Washington.] 7pp.
Reifler, Erwin: 1952e, The mechanical determination of the constituents of German substantive composita. (Studies in Mechanical Translation no.7). [Seattle: University of Washington.] 27pp. Repr. in: *Mechanical Translation* 2(1), 1955, 3-14.
Reifler, Erwin: 1952f, The MT form-class filtering system. (Studies in Mechanical Translation no.8). [Seattle: University of Washington.] 20pp. [This item no longer traceable.]
Reifler, Erwin: 1955, The mechanical determination of meaning. In: Locke and Booth 1955), 136-164.
Reifler, Erwin: 1962, Machine translation, *Bulletin of the Medical Library Association* 50 (3), 473-480
Reifler, Erwin: 1967, Chinese-English machine translation: its lexicographic and linguistic problems. In: Booth (1967), 317-428.
Reynolds, A.Craig: 1954, The conference on mechanical translation held at M.I.T., June 17-20, 1952, *Mechanical Translation* 1(3), 47-55.
Richards, Ivor A.: 1943, *Basic English and its uses*, London, Kegan Paul, Trench, Trubner.
Richens, Richard H.: 1951, An abstracting and information service for plant breeding and genetics. In: Casey, R.S. and Perry, J.W. eds.: *Punched cards: their applications to science and industry*, New York, Reinhold, 191-204.
Richens, Richard H. and Booth, Andrew D.: 1952, Some methods of mechanized translation. [Cambridge.] 31pp. Repr.in: Locke and Booth (1955), 24-46.
Richens, Richard H.: 1984, Letter to W.J.Hutchins, 13 May 1984.
Royal Society: 1948, *The Royal Society Scientific Information Conference, 21 June-2 July 1948*: Report and papers submitted. London, Royal Society.
Salmon, Vivian: 1972, *The works of Francis Lodwick: a study of his writings in the intellectual context of the seventeenth century*, London, Longman.
Shannon, Claude and Weaver, Warren: 1949, *The mathematical theory of communication*, Urbana, University of Illinois Press.
Sheridan, Peter: 1955, Research in language translation on the IBM type 701, *IBM Technical Newsletter* 9, 5-24.
Turing, Alan M.: 1936, On computable numbers, with an application to the Entscheidungsproblem, *Proceedings of the London Mathematical Society*, Ser.2, 42: 230-265.
Turing, Alan M.: 1969, Intelligent machinery. In: Meltzer, B. and Michie, D. eds. *Machine intelligence 5*, Edinburgh, University Press, 3-23.
Turing, Alan M.: 1986, Lecture to the London Mathematical Society on 20 February 1947. In: Carpenter, B.E. and Doran, R.W. eds.: *A.M.Turing's ACE report of 1946 and other papers* (Charles Babbage Institute Reprint Series for the History of Computing 10), Cambridge, Mass., MIT Press, 106-124.
Weaver, Warren: 1949, Translation. Repr. in: Locke and Booth (1955), 15-23.
Weaver, Warren: 1955, Foreword: the new tower. In: Locke and Booth (1955), v-vii.
Weaver, Warren: 1970, *Scene of change: a lifetime in American science,* New York, Scribner.
Wiener, Norbert: 1948, *Cybernetics, or control and communication in the animal and the machine*, 2nd edition, Cambridge, Mass., M.I.T.Press, 1961.
Wilkinson, J.H.: 1980, Turing's work at the National Physical Laboratory and the construction of Pilot ACE, DEUCE, and ACE. In: Metropolis (1980), 101-114.
Wundheiler, Luitgard: 1953, Invariant syntax as a prerequisite of all translation. [Illinois Institute of Technology.] 13pp. Expanded as Luitgard and Alex Wundheiler: ‘Some logical concepts for syntax.’ In: Locke and Booth (1955), 194-207.

Yngve, Victor H.: 1955, Syntax and the problem of multiple meaning. In: Locke and Booth (1955), 208-226.
Zarechnak, Michael: 1979, The history of machine translation. In: B. Henisz-Dostert, R.Ross Macdonald and M.Zarechnak *Machine translation,* The Hague, Mouton, 1979, 20-28.

Notes:

[1] Weaver's contacts with Wiener date back to wartime research in this field. In his book on cybernetics, Wiener wrote: "A considerable period before this [1944], the war research group conducted by Dr.Warren Weaver had published a document, first secret and later restricted, covering the work of Mr.Bigelow and myself on predictors and wave filters." (Wiener 1948: 15) It was this research which was to lead to mathematical prediction theory and the statistical approach to communication engineering (Shannon and Weaver 1949). After the war in 1946, Wiener obtained a substantial grant from the Rockefeller Foundation (through the office of Warren Weaver) to pursue research on the physiological and mathematical foundations of cybernetics (Wiener 1948: 21).

[2] The sources for this information come from documents held in the archives of the Rockefeller Foundation. This is also the source for the letters to and from Weaver and the memoranda between Weaver and his colleagues which are quoted in subsequent sections. The author is particularly grateful to the Rockefeller Archive Center (Pocantico Hills, 15 Dayton Avenue, North Tarrytown, New York 10591-1598) for permission to quote extensively from these papers.

[3] More recently, Booth (1997: 61) has stated that the meeting with Weaver at which MT had been raised had taken place on 11 August 1947.

[4] On 27 May 1946 Weaver wrote to Bernal:

> I am happy to inform you that an allocation of $400 has been provided toward the living and/or travel expenses of Dr.A.D.Booth in the United States. It is our understanding that he plans to attend the annual meeting of the American Society for X-ray and Electron Diffraction at Lake George, New York.

As we have seen, Booth took the opportunity also to visit centres where electronic computers were being built. Other information about J.D.Bernal is from the Bernal archives at Cambridge University Library.

[5] The entry no. 69 in the Mel'chuk and Ravich bibliography reads:

> *Perevodnaya mashina v Talline v 1924g.* - V kn.: Soobshcheniya po mashinnomu perevodu. Sb.1. Tallin, 1962, 95-96.
>
> Privoditcya soobshchenie iz estonskoi gazety "Vaba maa" ot 24 fevralya 1924g. ob izobretenii A.Vakherom pishushchei mashiny-perevodchika, model' kotoroi on prodemonstriroval.

[6] For a full account of Troyanskii's proposals see Hutchins and Lovtskii (2000), and for a description of the work of both Artsrouni and Troyanskii see Hutchins (2004a)

[7] Booth's computer ARC at the British Rubber Producers Research Association was not fully operational until late 1949, when it was transferred to Birkbeck College (Booth 1980, Lavington 1980: 62-67).

[8] An example from French is the following:

> (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 indicate automatic segmentations, *v* indicates a French word not translated, *m* "multiple, plural or dual", *z* "unspecific", and slashes alternative translations.

[9] Weaver was probably referring to Hans Reichenbach's 1947 book *Elements of symbolic logic* (Reichenbach 1947), where in the seventh chapter ('Analysis of conversational language') he attempted to construct a method of analysing natural language based on logical classifications of term, function and argument instead of traditional grammatical categories (noun, verb, adjective). His approach has remained influential, particularly in the analysis of temporal expressions.

[10] This omission in probably the most widely read of all the earliest papers in MT must surely explain why the role of Huskey in the early period of MT has been almost totally unrecognised. Huskey's work on SWAC (and his later links with UCLA researchers, Oswald, Bull, Harper, etc.) was also omitted from the historical introduction by Booth and Locke (1955), widely regarded as a reliable source for the period, and from the early accounts by Delavaney (1960) and Mounin (1964).

[11] The letter appeared in the *New York Herald Tribune* on 26 June 1949. The full text is as follows:

> Recently there appeared several news items regarding a new invention – a machine that translates from one language into another. One need not be a language specialist to realize the futility of translating intelligently from one language to another by machine. Time and again language experts have declared that the art of translation is one of the most difficult to master, or, as a great Hebrew poet once remarked: "It is like kissing your sweetheart through a veil." And if that is true of common, every-day words, how much more so is it of idioms and of great literary masterpieces. Let us take to the machine, for example, the fifty-five Hebrew words that make up the famous 23rd Psalm. Let us also assume that the intricate and efficient mechanism could translate correctly all the Hebrew verb forms, and every noun with its pronominal suffixes, the possessive and construct state; and that it could, moreover, give the correct meaning of the jussive, cohortative and consecutive. We then press the proper buttons on the translating machine and out comes –
>
> "Lord my shepherd no I will lack. In dwellings of grass he will cause me to lie down; on waters of resting places he will guide me. My life he will turn back; he will lead me in circles of justness so that his name. Also because I will go in valley great darkness no I will fear bad because Thou with me your tribe and your support they will console me. Thou shalt arrange before me a table opposite my enemies; Thou madest fat with oil my head my goblet saturation. But good and kindness he will chase me all days of my life; and I shall rest in the house of Lord to length days."
>
> Surely, only those in a "snake pit" shall follow such machine-made translations. But as for as we shall always read and enjoy the beautiful man-made translation beginning with "The Lord is my shepherd, I shall not want." Max Zeldner, New York, June 13, 1949.

[12] Shannon was the author of one of the most influential papers on the topic: 'Communication theory of secrecy systems', *Bell System Technical Journal* 28 (4), October 1949, 656-715. The content of the article had previously appeared in a classified report "A mathematical theory of cryptography" dated 1 September 1945. It seems very likely that Weaver knew of Shannon's work on cryptography since 1945, if not earlier.

[13] At the same time as he was writing the memorandum, Weaver published an article in the *Scientific American* on "The mathematics of communication" (vol.181 (1), 11-15), which introduced Shannon's theory to a wider audience, and which contains the passage:

> ... In a similar way, the theory contributes to the problem of translation from one language into another, although the complete story here clearly requires consideration of meaning, as well as of information.

[14] It is interesting to note that just as Weaver's idea of seeking a universal language to realise MT harks back to the seventeenth century, so too are there precedents in that century for his cryptographic proposal:

> ... if you once understand the Rules for *Decyphering*, in *one* Language, you may really, and without any Reservation, in a few hours, understand as much of *any* Language, as is needful to reduce it out of *Cypher* (John Falconer, Cryptomenysis patefacta, 1685; quoted by Salmon 1972)

[15] The idea of producing pidgin translations as an intermediary stage towards more polished translations was taken up later by the Cambridge Language Research Unit, the MT group to which Richens later belonged. The most extensive treatment of the proposal was made by Margaret Masterman and Martin Kay (1960).

[16] The report includes details of Booth's experience with computers:

> Mathematician as he is, Dr.A.D.Booth does not want to spend the best years of his life working out sums. He thinks, instead, that he should invent a machine to do the job.
>
> So he sets about making a "thinking" machine in a disused brewery near his parents' home at Fenny Compton, Warwickshire. And he uses some war surplus material from the scrap heap. And it takes two years to make the model.
>
> By this time the machine begins to add up to something. But the cost adds up, too. Dr.Booth has spent £3,000 on material – so for the £10,000 perfection of the mechanical brain Dr.Booth is receiving some financial backing from the Rockefeller Foundation.
>
> The robot brain grows. It will work out in a mere three weeks what it would take the scientists at London University 100 years to calculate with accuracy. And – almost as a sideline – it is a world dictionary and linguist, as well.

After the passage on how translation might work, it goes on:

> This machine which Dr.Booth is perfecting has a "memory" – a nickel-plated brass cylinder on which are magnetised many millions of combinations of electronic impulses.
>
> When corresponding series of impulses representing words in one language are transmitted to this memory, the machine reacts by delivering the equivalent in the other language.

[17] From his letter, it is clear that Booth had not seen the *Scientific American* article, since he wrote:

> The occasion for the News Chronicle article was a report from what the reporter led me to understand was an American article by yourself, which I have been unable to trace, and dealing with translation by mechanical means.

[18] Bar-Hillel (1951, 1964c: 153) reported that Weaver wrote a memorandum dated 6 March 1951, which apparently summarised the range of MT activity revealed by Loomis' survey. However, this memorandum has not been traced in the archives of the Rockefeller Foundation, or elsewhere.

[19] Loomis' scepticism about semantics stemmed probably from the contemporary popular interest in 'general semantics', the movement initiated by Alfred Korzybski in the 1930s (Korzybski 1958), which sought to explain social, political and international misunderstandings as the results of errors in the use of language and in the use of supposedly mistaken Aristotelian logic. The movement was an embarrassment to academic linguistics for its vacuous claims, ill-informed philosophy and ignorance of linguistics. However, Reifler's semantics had nothing to do with this movement, being based on traditional historical-comparative linguistics.

[20] In fact, only two of the papers in the collection were unchanged reprints of conference contributions. One was the paper presented by Booth on his and Richens' research (Richens and Booth 1952). The other was Dodd's paper on "Model English" (Dodd 1952). Other papers in the collection were reworkings: the article by Locke on "Speech input" (Locke 1955) was an expansion of his conference paper on "MT of printed and spoken material"; and Reifler's contribution "The mechanical determination of meaning" (Reifler 1955) was an elaboration of some of the ideas presented at the conference and in later papers (Reifler 1952d, 1952e). Bar-Hillel's paper on "Idioms" (Bar-Hillel 1955) was completely new.

Photocopies of some of the presentations, those by Bar-Hillel (1952a, 1952b, 1952c), Oswald (1952b), and Reifler (1952a, 1952b, 1952f), have been preserved on microfilm (roll 799 "Papers on mechanical translation") at the MIT Libraries Document Services, Room 14-0551, 77 Massachusetts Avenue, Cambridge, Mass. 02139-4307, USA.

According to Reynolds (1954), the presentations by the two computer engineers, Huskey on "Basic machine operations in mechanical translation" and Forrester on "Problems of storage and cost", were oral talks and no written papers had been prepared.

[21] After his 1950 paper, Kaplan did no more work in MT. He and was professor in the UCLA Department of Philosophy from 1946 to 1963, when he moved to Michigan. He had acted as a consultant for the Rand Corporation since 1947; and during 1951-52 was also a research associate on the Rockefeller project Language and Symbolism.

[22] Mounin (1964: 20) mistakenly thought that this research had been the occasion for Zeldner's letter in June 1949:

> Qu'il ait existé, dès cette date, d'autres essais sporadiques indépendants, cela résulte de la mention que fait Weaver lui-même d'une lettre de Max Zeldner... concernant un essai de traduction par une machine (il s'agit peut-être des essais mentionés par Olaf Helmer à la conférence du M.I.T. de juin 1952, essais faits pour la Rand Corporation de Santa Monica, Californie?)

In fact, as we have seen, Max Zeldner's letter was prompted by a report of the SWAC computer.

[23] Bar-Hillel's paper on idioms in the Locke and Booth collection (Bar-Hillel 1955) made much the same argument in favour of the post-editor solution. His example this time was translation into German of the (more readily accepted) idiom *red herring*. As in 1952, he was concerned that the literal translation should be available (*ein roter Hering*) when required (e.g. "when used in describing a painting by Marc Chagall"). He argued that the 'literal' translation should always be produced, with the translation into the German idiom (*Finte*) as an option for the post-editor.

[24] In his report of the conference, Reifler (1952c) advocated a pivot language which would be 'regularized' as this would greatly simplify translation from the pivot into another regularized language. As noted earlier, Reifler had always favoured MT output in the form of an "adjusted model target language".

[25] It may be noted that Reynolds (1954) believed that the SWAC computer had actually been used in word-for-word translations. Reporting Oswald's research, he wrote:

> The micro-glossary was employed in programming translations on the SWAC in cooperation with Dr.Harry D.Huskey... The translations so obtained conveyed the meaning of the original article with correlations of meaning better than 90 per cent, on the assumption that the problems of syntax and contextual modification had previously been solved.

It is, however, clear from Oswald (1952b) that the calculations were purely manual and assumed all problems of syntax and compound segmentation were resolvable. Reynolds had been confused, as many were to be in future years, between researchers' hopes and expectations and the realities of actual computational implementations.

[26] They were later published by Booth in 1955, when presumably they were still of relevance (Booth 1955).

[27] An article by Fairthorne (1952) on the mechanisation of information retrieval could point only to the NBS experiments to illustrate MT.

[28] Ducrocq wrote:

> Déjà, plusieurs cerveaux électroniques (notamment la SWAC) ont été convertis en machine à traduire. Si l'on sait ce qu'il faut penser de la traduction artificielle alors que l'homme, faisant appel au plus riche bagage d'images-souvenirs, éprouve la plus grande peine à exprimer des idées fidèles lorsqu'il change de langue, le problème prend en fait un aspect très différent lorsque les textes traduits concernent non plus des descriptions poétiques ou des analyses psychologiques, mais bien des éléments logiques, dûment classifiés et pour lesquels on peut obtenir immédiatement l'équivalent exact d'une langue à l'autre. C'est particulièrement le cas des textes techniques ou scientifiques, comme ce devrait être aussi bien le cas d'informations économiques, législatives ou politiques. Or plutôt que de concevoir des machines assurent toutes les traductions possibles selon les diverses langues prises deux à deux, il apparaît que le binaire pourrait représenter entre toutes ces langues l'agent naturel de conversion.

[29] Consultation of the bibliography produced by Russian researchers covering 1949 to 1963 (Mel'chuk and Ravich 1967) confirms that there were no publications on MT in the Soviet Union until 1955. Delavenay (1960: 30) supports: "The year 1955 also brought the first news of Russian activity in the field." This was in the form of two accounts of the Georgetown-IBM experiment (section 18) – see extended version of Hutchins (2004b)

[30] It may be noted, however, that after the approval of this grant extension no more was to be heard of the MT research team under Huskey. In 1953 Oswald and Bull returned to their full-time duties at UCLA, and Lawson obtained a position at Washington State University; none wrote again on MT. In 1954 Huskey himself left the Institute for Numerical Analysis to be a professor at the University of California in Berkeley. The SWAC computer was not used for MT research, and indeed Huskey never mentioned MT in accounts of his activities (Huskey 1980, 1984) . The only one to continue in the field was Kenneth Harper, who was later to be a leading member of the Rand Corporation MT group at Santa Monica (Hutchins 1986: 78-81).

[31] From internal evidence it would seem that in this year he wrote a fourth more substantial article surveying the achievements and prospects for MT, and summarising in some detail arguments which he had put forward in his 1951 survey and his 1952 conference papers. This article appeared in *American Scientist* in April 1954 (Bar-Hillel 1954). It must have been completed before the demonstration at IBM in January 1954 (section 19) since he mentions the demonstration only briefly in a final footnote.

[32] Although the results of some of this research were reported at the MIT conference (Reifler 1952b), there is no documentary evidence that he produced a full report, cf. his own bibliography (Reifler 1967).

[33] For a full account of the demonstration, how the system worked, and its impact at the time see Hutchins (2004b), and the expanded version on my website.

[34] The full report is reproduced in Hutchins (1994).

[35] The full report was as follows:

It's all done by machine Words go in in Russian, English sentence comes out

by Earl Ubell

A huge electronic "brain" with a 250-word vocabulary translated mouth-filling Russian sentences yesterday into simple English in less than ten seconds.

As lights flashed and motors whirred inside the "brain" the instrument's automatic type-writer swiftly translated statements on politics, law, science and military affairs. once the Russian words were fed to the machine no human mind intervened.

In demonstrating this feat for the first time scientists of the International Business Machine Corp. and Georgetown University, which collaborated in the project, said they hoped that within a few years such machines would be freely translating all languages.

At the demonstration in I.B.M. offices at 590 Madison Ave., Peter Sheridan, an I.B.M. mathematician, fed into the machine, filling a room as big as a tennis court, a series of cards carrying the Russian sentence:

"Myezhdunarodnoye ponyimanyiye yavlyayetsya vazhnim faktorom v ryesnyenyiyi polyityichyeskyix voprosov."

The machine blinked its lights, was quiet for a moment as if thinking and within nine seconds the automatic typewriter clacked and out came:

"International understanding constitutes an important factor in decision of political questions."

Even though the machine can translate such complex sentences, it is limited by its vocabulary, and by its "knowledge" of grammar, according to Dr.Leon Dostert, Georgetown language scholar, who originated the project. Another drawback to the machine's operation is that before the machine can work on the Russian sentences they have to be encoded on punched cards similar to those on which government checks are printed and which the Internal Revenue Bureau uses for taxpayer's names and addresses. This slows up translation.

Dr.Dostert said that it will not be too long – possibly three to five years – when automatic text-reading machines will feed in Russian sentences automatically into the machines without punched-card intervention.

Then, Dr.Dostert said, complete libraries of Western technical works could be made available to non-industrial backward nations to give them the "Know-how" of western technology. "At present," he added, "we are at the 'Kitty Hawk' state."

Dr.Dostert also foresaw the day when simpler and cheaper machines than the $500,000 I.B.M. super-calculator – the 701 – could be used. He said the 701 is "overdesigned" for the language translation problem and has many functions not necessary in this project but which were built in to solve problems in astronomy and physics.

How does the machine work?

I.B.M. engineers gave assurances that there was no pint-sized bilingual Russian inside the instrument pulling the right levers.

The success was the result of a collaboration among engineers, mathematical logicians and linguists. It was Dr.Paul Garvin, Indiana University graduate, who with Dr.Dostert, analyzed the Russian and English languages. And it was Mr.Sheridan, a City College graduate, who worked six months setting up the machine.

Formulated Rules

What Dr.Garvin did was to formulate rules that govern the translation of particular Russian words into English when those words appear in a particular context. Mr.Sheridan set up the machine to receive the rules in electrical form.

For example:

The word root "ugl" in Russian means either "angle" or "coal" depending upon its suffix. This root is stored in the form of electrical impulses on a magnetic drum together with its English meanings and the Garvin rules of syntax and context which determine its meaning.

The code is so set up so that when the machine gets electrical impulses via the punched cards that read "ugla" it translates it as "angle", when "uglya" the translation is "coal". Electrical code impulses activate the typewriter keys.

So far Dr.Garvin has formulated some six rules to govern the 250 words and their various English meanings. Dr.Dostert estimates that 100 rules would be needed to govern 20,000 words for free translation.

Eventually, the machine will be able to translate from Russian:

"She taxied her plane on the apron and then went home to do housework."

In such a sentence with double-meaning words, the machine will be able to tell what meaning of apron and taxi would be needed in that particular context.

To do all this the machine performs about 60,000 operations a sentence. During the demonstration yesterday it had two "nervous breakdowns". Random errors crept in which automatically shut the 701 down. "She" didn't cry.

[36] The full report is reproduced in Hutchins (1994).