CHAPTER 2: The precursors and the pioneers

2. 1: Before the computer

The use of mechanical devices to overcome language barriers was suggested first in the seventeenth century. There were two stimulants: the demise of Latin as a universal language for scientific communication, and the supposed inadequacy of natural languages to express thought succinctly and unambiguously. The idea of universal languages arose from a desire both to improve international communication and to create a 'rational' or 'logical' means of scientific communication.

Suggestions for numerical codes to mediate among languages were common. Leibniz's proposals in the context of his monadic theory are perhaps the best known. Another proposal was made by Descartes in comments on the sixteenth proposition of his famous correspondent Anonymous. In a letter to Pierre Mersenne on 20 November 1629 (Mounin 1964:16) Descartes described a proposed universal language in the form of a cipher where the lexical equivalents of all known languages would be given the same code number. Descartes wrote: "Mettant en son dictionnaire un seul chiffre qui se rapporte à *aymer, amare, philein*, et tous les synonymes (of *aimer* in all languages) le livre qui sera écrit avec ces caractères (i.e. the code numbers) pourra être interprété par tous ceux qui auront ce dictionnaire". Actual examples of such mechanical dictionaries were published by Cave Beck in 1657, by Athanasius Kircher in 1663, and by Johann Joachim Becher in 1661. At the height of enthusiasm about machine translation in the early 1960's some writers saw these 17th proposals as genuine forerunners of machine translation. Becher's book, for example, was republished under the title *Zur mechanischen Sprachübersetzung: ein Programmierungversuch aus dem Jahre 1661* (Becher 1962), indicating the conviction of its editor that Becher's ideas foreshadowed certain principles of machine translation. Apart from an ingenious script, Becher's book is distinguished from others of this kind only by the size of the dictionary: 10,000 Latin words (stems and endings) were provided with codings. Like others, however, Becher failed to tackle the real difficulties of providing equivalent entries in other languages (Greek, Hebrew, German, French, Slav, and Arabic were proposed) and the necessary means to cope with syntactic differences.

The vast work by John Wilkins, *An Essay towards a Real Character and a Philosophical Language* (1668), was a more genuine attempt at a universal language in that it sought to provide a logical or rational basis for the establishment of inter-language equivalencies. Wilkins' aim was "a regular enumeration and description of all those things and notions, to which marks or names ought to be assigned according to their respective natures", i.e. a codification which embodied a universal classification of concepts and entities, a genuine interlingua.

All these writers recognised the problems of genuine differences between languages that could not be captured completely in dictionaries, however 'logically' constructed. Many of them like Kircher advised their fellows to write in a simple style and avoid rhetorical flourishes.

Suggestions for mechanical dictionaries on numerical bases continued to be made throughout the following centuries until the middle of the present century. Couturat and Leau in their Histoire *de la langue universelle* (1903) list numerous examples, including one by W.Rieger entitled *Zifferngrammatik, welche mit Hilfe der Wörterbücher ein mechanisches Uebersetzen aus einer Sprache in alle anderen ermöglicht* (Code-grammar, which with the help of dictionaries enables the mechanical translation from one language into all others); a title which links the present mechanical age to the 17th century.

As the reference to Couturat and Leau implies, all these apparent precursors of MT should be regarded more accurately as contributions to the ideal of a 'universal language' and to the development of international auxiliary languages (Large 1985), of which the best known is now Esperanto. Both concepts have in fact inspired many of those engaged in machine translation.

None of these proposals involved the construction of machines; all required the human translator to use the tools provided in a 'mechanical' fashion, i.e. for man to simulate a machine. It was not until the invention of mechanical calculators in the nineteenth and twentieth centuries (starting with the pioneer activities of Charles Babbage) that an automatic device could be envisaged which could perform some translating processes. In fact, the first explicit proposals for 'translating machines' did not appear until 1933, when two patents were issued independently in France and Russia. In both cases, the patents were for mechanical dictionaries.

A French engineer of Armenian extraction, Georges Artsrouni was issued a patent on 22nd July 1933 for a translation machine which he called a "Mechanical Brain" (Corbé 1960).[1] The invention consisted of a mechanical device worked by electric motor for recording and retrieving information on a broad band of paper which passed behind a keyboard. The storage device was capable of several thousand characters, and was envisaged by its inventor in use for railway timetables, bank accounts, commercial records of all sorts, and in particular as a mechanical dictionary. Each line of the broad tape would contain the entry word (SL word) and equivalents in several other languages (TL equivalents); corresponding to each entry were coded perforations on a second band, either paper or metal, which functioned as the selector mechanism. The required entry was punched at the keyboard, the perforations located and the full entry retrieved within, it was claimed, 10 to 15 seconds. A prototype machine was exhibited and demonstrated in 1937; the French railway administration and the post and telegraph services showed considerable interest, and only the start of the Second World War in 1940 prevented installation of Artsrouni's invention.

More important in retrospect was the patent issued in Moscow on 5 September 1933 to Petr Petrovich Smirnov-Troyanskii for the construction of a "machine for the selection and printing of words while translating from one language into another or into several others simultaneously." (quoted from Panov 1960). A brief account by Troyanskii himself written in February 1947 was published in 1959 (Troyanskii 1959). Troyanskii[2] envisaged three stages in the translation process; the machine was involved only in the second stage, performing as an automated dictionary. In the first stage a human editor knowing only the source language was to analyze the input text into a particular 'logical' form: all inflected words were to be replaced by their base forms (e.g. the nominative form of a noun, the infinitive form of a verb) and ascribed their syntactic functions in the sentence. For this process Troyanskii had devised his own 'logical analysis symbols'. In the second stage the machine was designed to transform sequences of base forms and 'logical symbols' of source texts into sequences of base forms and symbols of target languages. In the third stage an editor knowing only the target language was to convert this sequence into the normal forms of his own language. Troyanskii envisaged both bilingual translation and multilingual translation. Although the machine was assigned the task only of automating the dictionary, it is interesting to note that Troyanskii believed that "the process of logical analysis could itself be mechanized, by means of a machine specially constructed for the purpose" (quoted by Panov 1960a). It was this vision of the next steps beyond a simple mechanical dictionary that marks Troyanskii's proposal as a genuine precursor of machine translation.

In the 1933 patent, the technical implementation proposed was a purely mechanical device, a table over which passed a tape listing in vertical columns equivalent words from various languages (not dissimilar to Artsrouni's machine). But, by 1939 he had added an improved 'memory' device operating with photo-elements (Delavenay 1960; Mounin 1964), and by May 1941 it appears that an experimental machine was operational. Troyanskii in fact went further towards the electronic computer; in 1948 he had a project for an electro-mechanical machine

[1] For more on Artsrouni see my unpublished article 'Two precursors of machine translation: Artsrouni and Trojanskij'

[2] For a fuller account of Troyanskii and his proposal see: J. Hutchins and E. Lovtskii 'Petr Petrovich Troyanskii (1894-1950): a forgotten pioneer of machine translation', *Machine Translation* 15 (3), 187-221; and the article cited in footnote 1.

similar to the Harvard Mark I machine developed between 1938 and 1942, and which is regarded as a forerunner of the ENIAC computer (Augarten 1984).

Troyanskii was clearly ahead of his time; Soviet scientists and linguists failed to respond to his proposal when he sought their support in 1939, and later "the Institute of Automation and Telemechanics of the Academy of Sciences was equally unforthcoming in 1944" (Delavenay 1960) In retrospect, there seems to be no doubt that Troyanskii would have been the father of machine translation if the electronic digital calculator had been available and the necessary computer facilities had been ready (Mounin 1964; Panov 1960). History, however, has reserved for Troyanskii the fate of being an unrecognised precursor; his proposal was neglected in Russia and his ideas had no direct influence on later developments; it is only in hindsight that his vision has been recognised.

2.2: The first beginnings (1946-1949)[3]

The electronic digital computer was a creation of the Second World War: the ENIAC machine at the Moore School of Electrical Engineering in the University of Pennsylvania was built to calculate ballistic firing tables; the Colossus machine at Bletchley Park in England was built to decipher German military communications. Immediately after the war, projects to develop the new calculating machines were established at numerous centres in the United States and Great Britain (Augarten 1984). The first applications were naturally in the fields of mathematics and physics, but soon the enormous wider potential of the "electronic brain" were realized and nonnumeric applications began to be contemplated.

The first suggestion that electronic computers could be used to translate from one language into another seems to have been made during conversations in New York between Andrew D. Booth and Warren Weaver.

Warren Weaver was at this time vice president of the Rockefeller Foundation. During the war Weaver had served on a scientific mission to investigate Britain's weapons development, and at the Rockefeller Foundation he was closely involved in the sponsorship of computer research and development. Booth had become interested in automatic digital calculation while working at the British Rubber Producers' Association in Welwyn Garden City, and had started to build a machine for crystallographic calculations. In 1945 he was appointed a Nuffield Fellow in the Physics Department at Birkbeck College in the University of London under Professor J.D.Bernal, where he constructed a relay calculator during 1945 and 1946 and initiated plans for computational facilities in the University of London. As a consequence of this work and the efforts of Bernal he obtained funds to visit the United States in 1946 under the auspices of the Rockefeller Foundation. There he visited all the laboratories engaged in computer research and development, at Princeton, MIT, Harvard, and Pennsylvania (Booth 1980).

While he was in the United States he met Weaver for the first time on 20 June 1946 at the Rockefeller Foundation in New York (Weaver 1946). According to Booth (1985): "At that time neither they nor I had any interest in machine translation. The discussions that I had with Warren Weaver were entirely on the subject of coming over to look into the question of acquiring the techniques for building a machine for the University of London based on American experience."

At the end of his US visit in July Booth submitted a report on computer development with particular reference to x-ray crystallography, and he was offered a Rockefeller fellowship to enable him to work at an institution of his own choice in the United States the following year. Booth selected the von Neumann group at the Institute for Advanced Study, Princeton University,

[3] For a fuller account of the early history of MT see: J. Hutchins 'From first conception to first demonstration: the nascent years of machine translation, 1947-1954. A chronology', *Machine Translation* 12 (3), 195-252. A briefer version appered as: 'First steps in mechanical translation', *MT Summit VI: past, present, future, 29 October – 1 November, 1997, San Diego, California, USA: Proceedings*, ed. V.Teller and B.Sundheim (Washington, DC: AMTA, 1997), 14-23

recognised then and now as doing the most advanced research on computers at the time. On his return to the United Kingdom, Booth stopped work on the x-ray calculator and, together with Kathleen Britten, devoted himself to the problem of von Neumann type computers, and in particular to problems of large-scale storage (Booth 1980).

Booth met Weaver again at the Rockefeller Centre in New York on 6th March 1947. According to Booth (1985): "The discussion then was entirely on the question of the Rockefeller Foundation financing a computer for the University of London, and Weaver pointed out that there was very little hope that the Americans would fund a British computer to do number crunching, although they might be interested if we had any additional ideas for using the machine in a non-numerical context, and he suggested I thought about things of that type. I had already thought about non-numerical applications from conversations with A.M.Turing in the mid 1940's... One of these was in fact translation, although at that time I had thought only of using the machine as a dictionary." Weaver suggested treating translation as a cryptography problem.

Weaver had in fact already on 4th March 1947, just before this meeting with Booth, written to Norbert Wiener of the Massachusetts Institute of Technology, one of the pioneers in the mathematical theory of communication, about the possibility of MT. In his letter (Weaver 1947, quoted in Weaver 1949), after commenting on the problem of translation in the post-war world Weaver wrote:

> 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... 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."... As a linguist and expert on computers, do you think it is worth thinking about?

Wiener's response on the 30th April disappointed Weaver: "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." (Wiener 1947). Weaver wrote again on 9th May but failed to interest Wiener in the possibilities. He also tried to interest Ivor Richards, the literary critic and linguist, who collaborated with C.K.Ogden on the development of Basic English (Richards 1943, Ogden 1968), and who at this time was professor at Harvard University (Marshall 1947); but with equal lack of success it would appear.

At the Institute for Advanced Study, between March and September 1947, Booth worked with his assistant Kathleen Britten on many tentative ideas for the new London computer, including the recognition of sounds and of printed characters and also the details of a code, which would enable information from a dictionary stored in a computer memory to be retrieved from standard teletype input (Booth et al. 1958:1).

This work was reported in a memorandum sent to the Rockefeller Foundation early the next year, on 12th February 1948. In it, Booth mentioned as "a possible application of the electronic computer... that of translating from one language into another", adding that "We have considered this problem in some detail..." (Booth 1948). Booth admitted later (Booth et al.1958:2) that the program written in Princeton would probably have been of little practical use, nevertheless it was

obviously of sufficient encouragement for him to continue to develop his ideas on dictionary lookup procedures.

It is evident that the first serious discussions and investigations of the possibilities of machine translation took place during 1947, beginning with Weaver's letter to Wiener and his meeting with Booth in early March. However, at a later date in 1955 when writing the 'Historical introduction' to the MT collection he edited with Locke (Booth & Locke 1955), Booth recollected the first mention of MT as having occurred during his 1946 visit. This has been generally accepted as the 'birth' date of MT; however, in other later publications Booth gives the date 1947 (e.g. Booth 1956, 1958), and he has now confirmed the March 1947 meeting as the one when the MT discussion with Weaver occurred (Booth 1985). On the other hand, in a letter to Bernal on 20 May 1948, written in acknowledgement of Booth's February report, Weaver himself stated: "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." (Weaver 1948)

Alan Turing's influence in the development of computers and in their possibilities as 'intelligent machines' is well known. He may legitimately be regarded as the pioneer of what is now known as Artificial Intelligence. In an essay written during September 1947, he mentions a number of possible ways in which the new computers could demonstrate their 'intelligence': "(i) Various games, e.g. chess, noughts and crosses, bridge, poker; (ii) The learning of languages; (iii) Translation of languages; (iv) Cryptography; (v) Mathematics." (Turing 1947). Evidently, Weaver and Turing were thinking along similar lines independently; and probably, others too.

As there were no facilities available at Birkbeck College, Booth began construction of a small computer at the laboratories of the British Rubber Producers' Research Association in Welwyn Garden City near London. The machine was operational by 12th May 1948 and a demonstration was given on 25th May to Warren Weaver and Gerard Pomerat, also of the Rockefeller Foundation (Booth 1980, Weaver 1949). On this occasion Weaver met Richard H.Richens, with whom Booth had been collaborating in experiments on mechanical dictionaries.

Richens had first met Booth on the 11th November 1947. His interest in mechanical translation had arisen independently out of experiments with punched cards for storing information at the Commonwealth Bureau of Plant Breeding and Genetics, where he was Assistant Director. "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). Richens is to be credited with the first suggestion of the automatic grammatical analysis of word-endings (Delavenay 1960:28). He proposed the segmenting words into their stems (or base forms) and endings (e.g. inflections), both to reduce the size of dictionaries and to introduce grammatical information into a dictionary translation 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 provides 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 the approach was tested by hand and by using punched card machinery on a wide variety of languages; the texts were taken from abstracts in plant genetics. The result of this collaboration with Booth was a memorandum written during 1948.

From a later version of this memorandum, presented at the MIT conference in 1952 (see below), some idea of these early efforts in 'dictionary translation' can be gleaned (Richens & Booth 1955).

From the French text:

> 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 de ces substance*s.

(Where the stars (*) indicate automatic segmentations.)

The English translation:

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*.

(Where v indicates a French word not translated, *m* "multiple, plural or dual", *z* "unspecific", and slashes alternative translations.)

These tentative experiments by Booth and Richens were known to very few. Brief mention was made during the International Conference on Science Abstracting held in Paris, June 20-25, when R. B. Reid of Unesco told some delegates about what he heard of their work from Professor Bernal (Reid 1949). Shortly afterwards, a short account was given by Holmstrom in a report on scientific and technical dictionaries which was submitted to Unesco circulated in mimeograph form in July 1949 (Holmstrom 1951). However, it was undoubtedly the memorandum which Warren Weaver wrote on 15th July 1949 that had most widespread and profound influence (Weaver 1949). The memorandum was distributed to some 200 of Weaver's acquaintances who, it was thought, might have potential interest in the concept of machine translation (Weaver 1949). For probably all recipients this memorandum was literally the first suggestion they had seen that the new electronic computers could be used as translating machines. In effect this memorandum launched machine translation as a scientific enterprise in the United States and subsequently elsewhere. Its historic impact is unquestionable; and it was later published in the Locke & Booth collection (1955) of early contributions to MT.

In his memorandum Weaver dates the origin of his speculations about MT to his wartime experience with electronic computers and to stories of startling achievements in cryptanalysis using computers. Weaver remarks that "it was very natural ...to think, several years ago, of the possibility that such computers be used for translation." He then reproduces the correspondence with Wiener, and refers to Booth's mention of MT in his 1948 memorandum as being the "first instance known to W.W. subsequent to his own notion about it", and outlines the experiments by Richens.

Weaver knew of one other MT experiment at the time of his memorandum. Newspapers had recently reported the use of one of the California computers in a primitive kind of word-for-word translation. These reports had prompted a letter in the *Herald Tribune* of 13th June 1949 from a Max Zeldner demonstrating how ridiculous would be word-for-word translations of literary texts. Mounin (1964: 20) speculated that these reports referred probably to the very tentative experiments mentioned by Olaf Helmer of the RAND Corporation at the 1952 MIT conference (see below).

2.3: Weaver's memorandum (1949)[4]

Weaver's memorandum concentrated more on the general strategies and long-term objectives of MT than on the more technical problems Booth and Richens had been tackling. Because of its historic importance it is worth enumerating in some detail the issues and problems raised by Weaver. He raised four points: the problem of multiple meaning, the logical basis of language, the application of communication theory and cryptographic techniques, and the possibilities of language universals.

The problem of deciding which specific meaning an ambiguous word may have in a particular text was, he suggested, solvable in principle if a sufficient amount of the immediate context is taken into account. The practical question of how many contexts are necessary could be answered by a statistical study of different types of texts on a variety of subject matters. Weaver

[4] See also: J. Hutchins 'Warren Weaver memorandum: 50th anniversary of machine translation', *MT News International* 22 (July 1999), 5-6, 15; and 'Warren Weaver and the launching of MT: brief biographical note', *Early years in machine translation: memoirs and biographies of pioneers*, ed. W.J.Hutchins (Amsterdam: John Benjamins, 2000), 17-20.

explicitly rejected the idea of actually storing in dictionaries long sequences of words for this purpose, but did suggest that "some reasonable way could be found of using the micro context to settle the difficult cases of ambiguity."

He expressed optimism about finding logical aspects in languages. In connection with a remark by Hans Reichenbach that he "was amazed to discover that, for (apparently) widely varying languages, the basic logical structures have important common features", Weaver commented that Reichenbach's observation seems to be confirmed in the linguistics literature and he mentioned work by Erwin Reifler (below) on the comparative semantics of Chinese and English. Secondly, Weaver expressed the conviction that "insofar as written language is an expression of logical character" then the theorem of McCulloch and Pitts (1943) on the mathematical possibility of computing logical proofs appeared to offer a demonstration of the logical possibility of implementing the translation process on a computer, i.e. that the problem of translating with a computer is formally solvable.

On the applicability of communication theory Weaver pointed to the important work of Shannon (which he was instrumental in popularizing, e.g. Shannon & Weaver 1949) and its relation to cryptography. The success of computers in cryptanalysis had prompted Weaver's speculations about MT, as the letter to Wiener in 1947 illustrates. In the memorandum he recounted an anecdote about the deciphering of a short 100-word coded text in Turkish. The cryptographer did not know Turkish, nor did he know that the message was in Turkish. (Before distributing his memorandum, Weaver checked the veracity of the story with his informant, Professor W.Prager of Brown University: the decipherment had been done by R.E.Gilman of the Mathematics Department (Prager 1948).) For Weaver this achievement was sufficient demonstration of the power of cryptanalytic techniques independent of the languages involved. Allied to the probabilistic foundation of communication theory, Weaver believed that the translation problem could be largely solved by "statistical semantic studies".

For Weaver the "most promising approach of all" was the investigation of language invariants or universals. He linked this again with logical structures of language and with probabilistic uniformities. The analogy he suggested was "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... 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... Thus may it be true that the way to translate ... 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..."

2.4: From Weaver to the first MT conference (1950-1952)

Weaver's memorandum brought to the attention of a wide circle the possibilities of a new and exciting application of the computers whose potentialities were being discovered and proclaimed with enthusiasm and optimism at this time. But, it did more. It indicated potentially fruitful lines of research in statistical analyses of language, on the logical bases of language, and on semantic universals of language. In addition, it pointed to some actual, even if tentative, achievements in the work of Booth and Richens. It was, however, received with considerable scepticism by many linguists who rejected it for its naivety in linguistic matters and for its unfounded assumptions on the logicality of language, and they were naturally sceptical about the possibility of formalising language and translation processes.

The memorandum had also been noticed by the press. An account appeared in *Scientific American* in December 1949, which reported on the experiment by Booth and Richens. This in turn was picked up by the British newspaper the *News Chronicle* in the spring of 1950, and so appeared the first of what in coming years were to be frequent misunderstandings and exaggerations. Booth's

APEXC computer program was described as an “electronic translator”, at which an 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.”

2.4.1: First MT studies.

Weaver’s own favoured approach, the application of cryptanalytic techniques, was immediately recognised as mistaken (see Mounin (1964: 31-39) for a detailed discussion). Confusion between the activities of deciphering and translation arise whenever the same person does both. Obviously, no translating is involved when an English message is deciphered into English by an English-speaking recipient. Likewise, the decipherment of the highly complex Enigma code used by Germany in the Second World War, with its immensely complex sequences of manipulations and transpositions, was not translation; it was only after the German texts had been deciphered that they were translated. The Colossus computers at Bletchley Park were applied to cracking the cipher, not to translating the German text into English. In practice, the cryptanalyst generally knows what the language is of the texts to be deciphered and often what their content is likely to be and the circumstances in which the message was transmitted. All this helps him to guess which letters and words are likely to be most frequent in the text. In the case cited by Weaver, the decipherment was based on “the frequencies of the letters, pairs of letters, etc. in English” (Prager 1948); fortunately they were much the same in Turkish and the original could be interpreted.

Though the cryptanalytic approach was mistaken, there were sufficient stimulating ideas in Weaver’s paper to launch MT as a serious line of research in the United States. During the next two years, individuals and groups began MT studies at a number of locations, the Massachusetts Institute of Technology (MIT), the University of Washington in Seattle, the University of California at Los Angeles (UCLA), the National Bureau of Standards (NBS) also in Los Angeles and the RAND Corporation nearby at Santa Monica.

On 10th January 1950, Erwin Reifler circulated privately the first of a series of studies on MT (Reifler 1950). Reifler was a Sinologist of German origin, head of the Department of Far Eastern and Slavic Languages and Literature at the University of Washington in Seattle. Recognising the problem of multiple meanings as an obstacle to word-for-word translation of the kind attempted by Booth and Richens, Reifler introduced the concepts of ‘pre-editor’ and ‘post-editor’. The human ‘pre-editor’ would prepare the text for input to the computer and the ‘post-editor’ would resolve residual problems and tidy up the style of the translation. One suggestion was that the pre-editor should indicate the grammatical category of each word in the source language (SL) text by adding symbols or diacritic marks, e.g. to distinguish between the noun *cónvict* and the verb *convíct*. The post-editor’s role was to select the correct translation from the possibilities found by the computer dictionary and to rearrange the word order to suit the target language. As we shall see, the concepts of pre-editor and post-editor recur in one form or another throughout the development of MT research.

Following Weaver’s suggestion for statistical studies of microcontext for resolving problems of multiple meaning, Abraham Kaplan at the RAND Corporation investigated polysemy in mathematics texts. (The study was completed on 30th November 1950.) A group of test subjects were presented with a set of words, each with a number of possible meanings, and asked to select the most applicable sense. Kaplan limited the test to nouns, verbs and adjectives on the assumption 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 preceding and following words (up to two before and after), and finally the whole sentence. It was 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 (Kaplan 1955). Despite its limitations and deficiencies (Kaplan excluded homographs: words of different grammatical categories having the

same form) and the tentativeness of the conclusions, this study encouraged hopes that problems of ambiguity were tractable and that statistical analyses could contribute useful linguistic data for MT.

In the latter half of 1950, a survey was conducted by W.F.Loomis on behalf of Weaver to find out all those who were interested in MT and what research was underway. The survey revealed a surprising amount of activity already (Loomis 1951): apart from Booth, Richens and Reifler, two groups had been set up in California. One was at the RAND Corporation in Santa Monica under J.D.Williams, and Kaplan's paper was to be the first of a series of MT studies. The other had been formed by Harry D.Huskey of the National Bureau of Standards in Los Angeles, with the intention of using the SWAC (Standards Western Automatic Computer) for MT research. The group included Victor A. Oswald of the German Department at UCLA and William E.Bull of the UCLA Spanish Department, and was soon joined by Kenneth E.Harper of the UCLA Slavic Languages Department. In support of its work, the group received some funds from the Rockefeller Foundation in July 1951.

From this group appeared during 1951 the first published journal article devoted to MT (Oswald and Fletcher 1951). Up to this time papers on MT had been mimeographed memoranda and reports of limited circulation. (It is true to say that to the present day information about MT research is largely contained in report literature of limited distribution; journal and book publication has on the whole been considered secondary.) The article by Victor Oswald and Stuart L. Fletcher, of UCLA and NBS respectively, was also the first attempt to investigate syntactic aspects of MT. Arguing that word-for-word translation of a language like German would produce obviously unsatisfactory results, Oswald and Fletcher proposed a detailed grammatical coding of German sentences indicating syntactic functions of nouns and verb forms in clauses and enabling the identification of 'noun blocks' and 'verb blocks'. On the basis of the codes, certain sequences were identifiable as candidates for rearrangement when the output was to be in English. The procedures were framed as 'instructions' for the SWAC computer at NBS, but were not in fact implemented in any way. In an appendix, the authors illustrated how a mathematics text by Cantor would be coded, reordered and then translated word by word into English, e.g. the original:

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

reordered:

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

English:

> 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.

Oswald and Fletcher concluded that syntax "does not constitute, as had been thought by some, a barrier to mechanical translations"; they stressed the problems of solving the "lexicographic difficulties" of MT.

2.4.2: Bar-Hillel's survey (1951)

At the Massachusetts Institute of Technology a meeting on MT was held soon after the appearance of Weaver's memorandum. According to Locke and Yngve (1958: 511): "In January 1950 Dr. Weaver met at M.I.T. 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". In the next year, 1951, Yehoshua Bar-Hillel became the first person to be appointed full-time specifically for research on MT; the appointment in the Research

Laboratory of Electronics was made possible by a grant form the National Science Foundation (quite possibly with the influence of Weaver who was a director of NSF at the time.) Bar-Hillel's task was to study the possibilities of MT and to plan future research at MIT.

At the end of 1951 he produced a survey of the current position (Bar-Hillel 1951). The paper[5] raised many of the issues which dominated discussion of MT in the following years: the feasibility of fully automatic MT, the role of post-editing, the objectives of syntactic analysis, the role of statistical information, the possibility of universal grammar, the logical foundations of language, and restricted vocabularies.

Bar-Hillel argued that since problems of semantic ambiguities could not be resolved at present, "high-accuracy, fully-automatic MT is not achievable in the foreseeable future". This fact, however, should not discourage MT research: a less ambitious target is feasible, "a mixed MT... 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." Bar-Hillel saw no alternative to the post-editor whose task would be the "elimination of semantical ambiguities, in addition, of course, to stylistic smoothing."

As for the machine processes, Bar-Hillel saw these as comprising the morphological analysis of each word into its stem and grammatical category, the "mechanical identification of small syntactical units" and the "transformation of the given sentence into another that is logically equivalent to it." The second stage required what Bar-Hillel called an "operational syntax", an explicit programmable method for syntactic analysis, the rudiments of which Bar-Hillel recognised in Oswald and Fletcher's paper.

While recognising some value in statistical analyses, he was sceptical of proposals to include only high frequency vocabulary in dictionaries since the residue are likely to be those "highly loaded with information" in a given text. In any case, he felt that the current limitations of computer storage would be solved by hardware developments.

At the end of the paper he considered the possibilities of constructing a "universal, or at least general grammar, perhaps even ... a whole artificial exchange-language" on the basis of work by logicians such as Ajdukiewicz, Carnap, Reichenbach and himself, and by the linguist Zellig S.Harris (cf.ch.3.4-5 below), or, less ambitiously, "transfer-grammars... in which the grammar of one language is stated in categories appropriate to some other language". Finally, he proposed the application of MT to situations where restricted vocabularies are used, e.g. by pilots and meteorologists, and where "the theoretical difficulties of such a type of MT are clearly less formidable". To these he added regularized languages such as Basic English and auxiliary languages such as Esperanto and Interlingua.

As we shall see, not only had Bar-Hillel raised many of the major MT issues recurring in the following years and to some extent to the present day, but he also stated views which he was to repeat at greater length and with great impact in the early 1960's (ch.8.3)

2.4.3: First MT conference (1952)[6]

By 1952 interest in MT had grown enough for the Rockefeller Foundation to sponsor the first conference on MT. It was held at the Massachusetts Institute of Technology from 17th to 20th June 1952, and was organised by Bar-Hillel. Eighteen individuals interested in MT attended. As might be expected, MIT was well represented, by seven participants: Yehoshua Bar-Hillel; Jay W.Forrester, head of the Computer Laboratory; William N.Locke, Dept. of Modern Languages; James W.Perry, Center of International Studies; Vernon Tate, Director of Libraries; Jerome

---

[5] See also: J. Hutchins 'Bar-Hillel's survey, 1951', *Language Today* 8 (May 1998), 22-23.

[6] For more on the conference see: J. Hutchins 'Looking back to 1952: the first MT conference', *TMI-97: proceedings of the 7th Interantional Conference on Theoretical and Methodological Issues in Machine Translation, July 23-25, 1997, St.John's College, Santa Fe, New Mexico, USA* (Las Cruces: Computing Research Laboratory, New Mexico State University), 19-30.

B.Wiesner, director of the Research Laboratory of Electronics; and Dudley A.Buck, Electrical Engineering Dept. The Los Angeles area was represented by four participants: Victor Oswald and William E.Bull from UCLA; Olaf Helmer from RAND; and Harry D.Huskey from NBS. Erwin Reifler came from the University of Washington, as did Stuart C.Dodd, the director of the Washington Public Opinion Laboratory. The remainder were: Andrew D.Booth from Birkbeck College (the sole British participant), Leon Dostert from the Institute of Languages and Linguistics of Georgetown University, Duncan Harkin from the Department of Defense, Victor H.Yngve from the University of Chicago, and A.Craig Reynolds from the Endicott Laboratories of IBM. The backgrounds of participants reflect already some of the principal sources of interest in MT: electronic engineering and computing, linguistics, librarianship and information science, military and governmental bodies. Although the proceedings were not published at the time, accounts of the conference appeared subsequently by Reynolds (1954) and Reifler (1954), both compiled immediately after the conference; and, somewhat later, a number of the papers were printed in the collection by Locke and Booth (1955).

The conference opened with a public session. Bar-Hillel enunciated the need and possibilities for MT, particularly to cover the immense and growing volume of scientific research and popular periodical literature of a country. Leon Dostert spoke on his experience in setting up the simultaneous translation systems at the Nuremberg trials, at the United Nations and other international conferences. He was sceptical of the contribution of MT except for the processing of material currently not touched in specialized fields. Olaf Helmer mentioned the tentative experiments at RAND, and Perry spoke of the relationship of MT to automatic indexing and retrieval systems.

The presentations on the following days included Reifler and Bar-Hillel on pre-editing and post-editing respectively, and Booth on his work with Richens on mechanizing a dictionary. For Reifler (1954) the low intelligibility of the output from the latter supported his argument for pre- and post-editing. Oswald presented his proposals for treating German syntax, and then with Bull suggested that problems of ambiguity could be overcome by constructing micro-glossaries for particular subject fields; the glossaries could be established from statistical analyses of the relevant literature, and an investigation of the vocabulary of brain surgery had already been undertaken (later published in Oswald and Lawson, 1953). Bull warned that no scientific vocabulary constituted a closed domain and the lexical ambiguities in general vocabulary would still remain. Dodd outlined his ideas of regularizing English syntax and morphology for use in MT, e.g. “I will send he to she”, “he have”, “she did be loved” (Dodd 1955); and Reifler suggested encouraging writers to write with MT in mind (i.e. to write in MTese, as later researchers put it). Dostert put forward the advantages of a “pivot language” in the context of multilingual MT. Reifler (1954) thought this would most likely be English and he also thought that MT output should be regularized on the lines suggested by Dodd; studies of language universals could contribute to this regularization in the MT context.

According to Reynolds (1954), Bar-Hillel's operational syntax 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.” There was naturally some discussion on the problems of writing computer programs for MT, on computer costing and on the use of punched card machinery. The conference concluded with statements from participants about what research they hoped to do. The most challenging proposal came from Dostert, who had been converted from his original scepticism in the course of the conference. He 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 1954).

2. 5: From the MIT conference to the Georgetown-IBM demonstration (1952-1954)

The conference was an undoubted success, the participants were enthusiastic about the prospects and the general public had been made aware of the possibilities. Although no formal conclusions were drawn, it seems there was general agreement on what the next stages of MT research should be (Booth & Locke 1955: 6-7): "word frequency and word translation studies" on micro-glossaries and investigation of suitable storage, input and output methods for "an operating automatic dictionary at the end of approximately 12 months"; immediate studies on "operational analysis of syntax"; later longer-term work would come on multilingual MT, universal grammar and interlinguas. In retrospect, the expectations were far too optimistic, but they provided a stimulating framework for further MT research.

In the following year, articles presenting MT to the general public began to appear, e.g. Booth (1953) and Bar-Hillel (1953) in Computers *and Automation*, and MT appeared in a textbook for the first time, in a chapter on 'Some applications of computing machines' in the book by Booth and his wife, the former Kathleen Britten (Booth & Booth 1953).

During 1953 Bar-Hillel left MIT to return to the Hebrew University of Jerusalem; he was replaced by Victor Yngve, who had participated at the 1952 conference. At this time, Yngve shared the view 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" (Yngve 1955). A partial translation of a German mathematics text was simulated in which function words and endings were left untranslated:

> 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...

Although readers with some familiarity with German could work out the general drift it was clear to Yngve that syntactic analysis was essential for better MT output; and this was to be the central emphasis of the MIT group which he formed in the coming years (ch.4.7 below)

One of the most significant outcomes of the 1952 conference was the establishment of a MT research team at Georgetown University by Leon Dostert to start work towards the pilot experiment he had advocated to demonstrate the practical feasibility of machine translation. Dostert was fully aware of the considerable linguistic problems of MT but had concluded 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" and that the test should serve as the basis for a series of progressively larger and more complex tests. In brief, Dostert advocated what was later to be generally known as the 'empirical' approach to MT research. Collaboration was agreed with IBM, under the aegis of the company's programme of endowed research in computation; Paul Garvin was appointed at Georgetown to work on the linguistic procedures for the trial translation of Russian texts; and Peter Sheridan of the IBM Scientific Computing Service was given the task of implementing the procedures on the IBM 701 machine.

The program was ready by the end of 1953 and on the 7th January 1954 a public demonstration of the Georgetown-I.B.M. experiment took place at IBM's Technical Computing Bureau in New York.[7] This small-scale experiment in Russian-English translation was one of the

[7] For more details of the demonstration see: J. Hutchins 'The Georgetown-IBM demonstration, 7th January 1954', *MT News International* 8, 15-18.

most important events in the early history of MT. It was the first real demonstration of MT on a computer; earlier experiments had either been hand simulations or had used punched card machinery. No pre-editing of text was required and the output appeared to be fully intelligible. Furthermore, it was the first implementation which went beyond word-for-word translation. Nevertheless, its limitations were clearly recognised: a vocabulary of just 250 Russian words, only six rules of grammar and a carefully selected sample of Russian sentences (ch. 4.3 below). For the general public and the scientific community at large, however, the experiment showed that MT was a feasible objective, and it undoubtedly helped to stimulate the funding of MT research by U.S. governmental agencies in the following decade.

The demonstration received wide publicity; Sheridan and Dostert lectured widely on the system and the potential future of MT. Further demonstrations were given during the year, for example on the 14th September at a New York meeting of the American Chemical Society (*MT* 1(3) Dec 1954). Another demonstration was given to D.Y. Panov from the USSR Academy of Sciences, who was visiting computer installations in the United States. On his return, research on MT began in the Soviet Union.

2. 6: From the Georgetown-IBM experiment to the first international conference (1954-1956)

During 1954 two further MT groups were founded, an informal group at Cambridge in Great Britain by Margaret Masterman, and a research team at Harvard University by Anthony Oettinger who also in the same year Oettinger presented the first doctoral dissertation on MT (ch.4.9 below). In March appeared the first issue of the journal *Mechanical Translation* issued from M.I.T. under the editorship of Victor Yngve and supported by a grant from the Rockefeller Foundation. This journal was to carry in subsequent years many of the most important articles on MT.

During 1955 the first news of Russian activity became known. As a result of his visit to the demonstration at IBM, Professor Panov began MT experiments in January 1955 on the BESM computer at the Institute of Precision Mechanics and Computer Technology in Moscow. Within the next two years other MT groups were formed in the USSR at the Steklov Mathematical Institute, at the Institute of Linguistics and at the Laboratory of Electrical Modelling, all in Moscow, and at the University of Leningrad (ch.6 below).

In August 1955 the Cambridge Language Research Group held a meeting at King's College, Cambridge. Participants included R.H.Richens, the mathematicians A.F.Parker-Rhodes and E.W.Bastin and the linguists J.R.Firth, M.A.K.Halliday and R.A.Crossland. Already in the abstracts of the proceedings (CLRU 1956) there is evidence of the distinctive emphasis of CLRU on interlinguas, pidgin languages, logical and semantic foundations, and lattice theory (ch.5.2 below). At about this time also, Silvio Ceccato started a MT project in Milan, adopting a distinctive 'philosophical' approach to MT grammar and semantics (ch.5.3 below)

Publications on MT were now beginning to appear in larger numbers, shortly voluminous. A major publication of 1955 was the collection of articles edited by Locke and Booth (1955), the first book to be published devoted entirely to MT research. In it appear many of the most significant papers of the period up to 1955, including Weaver's memorandum, the report by Richens and Booth on their early work, and contributions to the 1952 conference. It includes also a valuable 'Historical introduction' and early papers by Oettinger (1955) on automatic dictionaries (predating his dissertation), by Yngve (1955) on syntactic analysis, by Booth (1955) on available computer hardware, and by Bar-Hillel (1955) on the problems of idiomatic usage.

It was clear that MT was growing fast. In October 1956 the first international conference was organised by MIT attended by 30 workers from the United Sates, Great Britain and Canada, and papers were received from the Soviet Union. The presentations by MT groups of their current research activity revealed the general pattern of approaches which was to dominate the next five

years or so. On the one hand there was the emphasis on dictionary, lexicographic and semantic problems, e.g. at Washington University, at Los Angeles (UCLA and RAND), at Michigan, at Harvard and at Cambridge; and on the other hand, there was the greater emphasis on syntactic problems at Georgetown and MIT. Another major division was between the 'empirical' approach, typified by RAND and Georgetown, and the 'theoretical' approach, typified by MIT and Cambridge. A third division was between those groups aiming to produce operational systems in the near future, however limited (e.g. Birkbeck, Washington and to some extent Georgetown), and those with the longer term objective of producing a good quality product (e.g. Harvard and MIT).

From 1956 research on MT was pursued with great vigour all over the world by numerous groups. In the United States MT research received increasingly large grants from governmental, military and private sources. By 1964, however, the promise of operational MT systems still seemed distant and the sponsors set up a committee, which recommended in 1966 that funding of MT should be reduced. It brought to an end a decade of intensive MT research activity. The work of the various groups during these years will be described in chapters 4 to 7, after a preliminary chapter outlining some of the basic MT methods and strategies.