CHAPTER 4: Groups and projects in the United States 1950-66

4. 1: University of Washington (1949-1962)

This MT research team was one of the earliest and biggest of those set up during the 1950's. As we have seen (ch. 2.4), Erwin Reifler, who was Professor of Chinese at Washington, was one of the pioneers of MT.[1] Stimulated by Weaver's memorandum, he embarked on a number of studies principally concerned with German-English translation. He was supported between 1952 and 1953 by grants from the Rockefeller Foundation, enabling him to investigate general problems of MT (such as post-editing and dictionary construction) as well as more specific research on German compounds. The problem, as Reifler saw, was that it is simply impossible to provide dictionary entries for German nouns since so many compounds can be freely formed; automatic procedures are needed to divide compounds into those components which may appear in the dictionary, thus enabling plausible attempts at their translation. From his investigation Reifler (1955a) concluded that automatic procedures were feasible; the only problem was that some compounds could be split in two different ways (e.g. *Wachtraum* as *Wach/traum* (day dream) or *Wacht/raum* (guard room)); although most of these dissections would produce nonsense translations (e.g. *Literat/urkunde* (man of letters' document) instead of the correct *Literatur/kunde* (literary studies)) Reifler did not elaborate on how these components were to be translated.

Shortly afterwards Reifler began a collaborative project with members of the Department of Electrical Engineering at Washington (T.M.Stout, R.E.Wall, Professor W.R.Hill, and R.S.Wagner) on the development of a pilot MT model system for German-English translation. At the 1952 MIT conference, Reifler had agreed with Dostert of Georgetown University that what MT research needed next was to produce some prototype MT systems. The Washington prototype constructed during 1954 was extremely limited, just 60 German words with English equivalents, and none longer than 7 letters (*MT* 2(2) Nov 1955), but it provided the foundation for a MT project, later joined by Lew R. Micklesen (a Russian expert) and David L. Johnson (a computer expert). In this group Hill concentrated on the development of large-capacity, rapid-access storage while Micklesen investigated Russian compounds and problems of grammatical information.

In June 1956 the Washington MT group received a grant from the U.S.Air Force (Rome Air Development Center) to study the lexicographical, linguistic and engineering prerequisites for the automatic translation of scientific Russian into English, based on the photoscopic disc memory device being developed by Gilbert W. King at the International Telemeter Corporation in Los Angeles. It was charged specifically with the preparation of the Russian-English lexicon for the USAF system. At this time, according to Reifler (1958), the "U.S.Air Force (was) primarily not interested in machine translation, but in an efficient information retrieval system permitting quick access to the enormous amount of information stored in its files." In the initial project (1956-57) the team analyzed 111 Russian texts covering 40 scientific fields, building up a database of more than 14,000 Russian-English dictionary entries of technical and general-language vocabulary. An expanded project began in March 1957 to supplement this initial collection by adding all the inflected forms for this vocabulary (i.e. complete paradigms of nouns and verbs). This lexicographic phase concluded in June 1959 with a Russian-English MT lexicon of 170,563 entries stored on over half a million IBM punch cards. Further grants were received from the U.S.Air Force until March 1960 to enable the Washington group to investigate other aspects of Russian-English translation.

The emphasis of the Washington group's MT research was determined both by Reifler's conviction of the essential validity of the basic word-for-word approach (at least for intermediary

[1] See also L.W. Micklesen: 'Erwin Reifler and machine translation at the University of Washington', *Early years in machine translation: memoirs and biographies of pioneers*, ed. W.J.Hutchins (Amsterdam: John Benjamins, 2000), 21-38.

working systems) and by the physical capacities of the photoscopic disc. At a time when existing electronic computers had limited storage capacities, grossly inadequate for MT dictionaries and translation procedures, the impressive capacity of the photoscopic disc was very attractive. Many researchers had suggested that MT required purpose-built machines, and a translation machine including the photoscopic memory seemed a practical objective.

What had been developed at the International Telemeter Corporation under Gilbert King was a memory device consisting of a rotating glass disc, together with various mechanical, optical and electronic components. Information was stored photographically on the glass disc in binary coding in 600 concentric rows with 50,000 bits of information in each row, making possible the storage of large amounts of information in a limited space (30 million bits). Information was read by shining a light through the disc, rotating at 20 revolutions per second, and converting the resulting alternations of light and dark to electric signals by the use of a photocell; and these signals were then processed by computer (King et al. 1953, *MT* 3(2) Nov 1956, King 1959, Wall 1960)

In an investigation of the dictionary requirements for the Russian-English project, Wall (1960) estimated that the storage capacity needed was 130-180 million bits. The advanced model of the disc was intended to have a capacity of 100 million bits and a random access time of 1/20 seconds, so 2 photoscopic discs were considered easily sufficient for a large-scale translation system.

The distinctive features of MT research at Washington (Reifler 1961a) were its particular emphasis on lexicographic approaches rather than structural linguistic approaches, its emphasis on devising a system which could deal with the vocabulary of many fields of science rather than just one field or sub-field, and its emphasis on 'free-form' dictionary entries (i.e. all paradigmatic forms of words) rather separate entries of stems and endings. A major factor in adopting the lexicographic approach was that “since Dr. King’s automatic system did not yet include logical equipment for linguistic purposes, it was decided to attempt to solve as many bilingual linguistic problems as possible by purely lexicographical means.” However, it was discovered that many grammatical problems could also be solved by lexicography alone, without the necessity of syntactic analysis. Having been explicitly charged by their sponsors with the creation of a Russian-English lexicon it was, therefore, decided to aim for an “optimum of lexicography”, leaving unsolved problems to be dealt with later (Reifler 1960, 1961a)

The emphasis on a wide range of scientific vocabulary was also mainly determined by the large capacity of the photoscopic disc. It was believed that the earlier arguments for concentrating MT systems to scientific fields had lost their force: “If specialized glossaries, idio-glossaries, for one field or sub-field of science will still be used in MT in the future, it will not be because of any limitations in the storage capacity of the permanent memory device” (Reifler 1961a). However, Reifler claimed there were also good linguistic reasons: there were believed to be considerable overlaps in scientific vocabularies and it was held that the semantic problems of scientific texts were caused mainly by the general non-scientific vocabulary and by the specifically scientific vocabulary which was shared by a number of different fields. Therefore, the Washington group set about the compilation of a general scientific vocabulary and worked out procedures in order to pinpoint the intended meaning of those (homographic) terms occurring in more than one scientific field.

For this they took two approaches: classification and cover terms. Detailed statistical analyses of multiple meanings were conducted in order to establish a general classification of scientific vocabulary fields, with the intention of assigning these classes to items of Russian and English vocabulary. In the case of overlapping vocabularies, e.g. medicine and biology, and applied science and technology, a good deal of redundancy had to be tolerated (Micklesen 1960). In the other line which was pursued, the groups sought ‘cover forms’, single English equivalents which would give acceptable versions for Russian terms whatever the scientific field.

Although larger storage seemed also to make procedures for dividing entries into stems and endings less necessary (Reifler 1961b), this was not the principal argument used by Reifler in favour of ‘free-form’ entries. A stems-ending approach requires morphological analysis, often of some complexity and certainly of considerable programming size. Furthermore, the problem of recognizing homographs is increased because information carried by full forms has been lost (e.g. noun/verb homographs can be distinguished by their different inflections).

However, Reifler's goal of a MT system employing the “optimum of lexicography” went further. It meant not only an emphasis on a general-language vocabulary for all fields of science and the ‘free-form approach’, but also the treatment of many difficult phrases as if they were idioms. As Reifler (1961c) explained, any MT system has to cope with genuine ‘bilingual’ idioms, i.e. expressions which cannot be translated word-for-word (Reifler’s example was that while *The man is an ass* could be translated word for word into German *Der Mann ist ein Esel*, it cannot be so translated into Chinese). But in addition, MT systems can use the same technique for non-idiomatic cases. For example, while English *league* could be German *Liga, Bund* or *Bündnis* and *nations* could be *Nationen* or *Völker*, the ‘League of Nations’ should be treated as an idiom: *Völkerbund*. Likewise, compounds in which an element may have different translations according to its context (e.g. in *Denkarbeit*, *Denkart*, *Denkmünze*, etc. where ‘Denk-‘ can be ‘thinking’, ‘thought’, ‘commemorative’, etc.); these too can be treated as idiomatic terms.

The system developed under the USAF contract until 1960 was, therefore, an elaborated word-for-word system using the photoscopic store. The system had no problem with idioms, could cope with compounds and with some homographic problems by examination of immediate context. But it had no means of resolving syntactic ambiguities or of reordering the output to conform to English usage. An example translation from (Micklesen 1958) shows the kind of results achieved:

> Infection/corruption (by/with/as) nodular (by/with/as) bacteria comes/ advances/treads especially/peculiarly (it)(is)light/easy(ly) at/by/with/from (of) plants, (of) weakened/loosened (to/for)(by/with/as) nitrogen/nitrous (by/with/as) starvation, and/even/too (is)considerable/significant(ly) (is/are)more-difficult(ly) happens/comes-from at/by/with/from (of) plants, (is)energetic(ally) (of)growing on/in/at/to/for/by/with (of) rich (of)nitrogen/nitrous soils.

It was obvious to all that such ‘translations’ were, as members of the group admitted, “far from clear and far from being readily intelligible” (Wall 1960).

However, the project had always been seen as a two stage operation with the intention after the lexicographic work of developing ‘logical procedures’ to deal with divergences between source and target languages. This research began in June 1959 after the completion of the supplemented lexicon of 170,000 entries, and continued until the termination of the project in March 1960. Micklesen (1960) investigated a number of problems. One was the possibility of resolving verb homographs by a classification of object nouns (e.g. as animate, human, concrete, metals, plants, etc.) and a selection algorithm. For example, *dokhodyat*: (they)reach/ripen/are-done can only be 'ripen' if the following noun is classified as a 'fruit/vegetable'. Another problem was the practice of providing equivalents in the form of a ‘minimum correlation-specification’ (e.g. Russian *izmeneniya*: English (of)change(s)), with the result that the output included a clutter of superfluous grammatical information, as in the example above. It was the object of subsequent research to reduce this clutter substantially and, if possible, to reduce the number of English alternatives provided.

In the initial stages of the project ‘translations’ had been manually simulated; the availability of an IBM 650 enabled some testing of the algorithms. The output from the dictionary search (still manually simulated, the IBM 650’s capacity being only 2000 words) was processed in four stages (Micklesen & Wall 1961; Wall 1960). The first three resolved homographs by the identification of agreement relations (e.g. within noun phrases) and government relations (e.g. substantive and verb), and by the elimination of superfluous information (e.g. if a substantive had a preceding agreeing adjective the English preposition was eliminated: 'nervous (of)... impotence' became 'nervous... impotence'). The fourth stage inserted English equivalents for Russian case relations and translated Russian prepositions. An example of the 'improvement' given by Wall was:

about/against/with treatment (of)(to/for)(by/with/as)nerve/nervous (of)(to/for)impotence (by/with/as)novocain

which became:

about/against/with treatment of nerve/nervous impotence (by/with/as)novocain

As it transpired this work was the nearest the Washington team approached an actual MT system; the final stage, in which attention was to be paid to problems of producing acceptable word order in TL output, was not reached. There had been expressed some interest in transformational grammar for investigating problems of syntax but this could not be pursued. The Washington approach was throughout pragmatic: the aim was to produce something which however inadequate could be progressively refined. Although their own results were relatively meagre, they shared with others a general mood of optimism about MT: typical is Reifler's conviction that "it will not be very long before the remaining linguistic problems in machine translation will be solved for a number of important languages" (Reifler 1958).

Although most work was done on Russian-English translation, it is no surprise that Reifler did not neglect Chinese. A particular problem with Chinese is the identification of lexical units, since word boundaries are not indicated in written Chinese; for Reifler, then, the basic approach in Chinese MT had to be lexicographic. As in the USAF project, what was needed first was a study of the general-language vocabulary of Chinese scientific texts, the aim again being the establishment of a bilingual lexicon making possible idiomatic translations with least recourse to 'logical operations' for solving problems of ambiguity and differences in syntactic structures. However, because of the "much greater morphological and syntactic agreement between Chinese and English" than between Russian and English, Reifler expected word-for-word translations from Chinese to be more intelligible than those from Russian. On the other hand, the number of English alternatives for Chinese characters is on average much higher than in the case of Russian-English translation. As in the USAF project, the Chinese vocabulary was classified according to scientific fields, in order to assist homograph resolution and selection of English equivalents. Reifler (1967) gave an example of a simulated MT from Chinese, with simulated improvements of the purely lexicographic output.

By 1959 research on the development of the photoscopic store and with it, research on the USAF translation Russian-English system had been transferred to the IBM Research Center at Yorktown Heights, New York. Apart from individual important contributions by Reifler, Wall and Micklesen, the principal achievement of the Washington MT research group was its single-minded pursuit of the lexicographic approach and its demonstration of what could be achieved with bilingual (direct) word-for-word translation. Above all, from a historical perspective, it laid the foundations for the earliest operational MT system.

4. 2: IBM Research Center (1958-66)

By 1959 research on the development of the photoscopic store had been transferred to the IBM Corporation's T.J.Watson Research Center at Yorktown Heights, New York. As before in Los Angeles, the research was sponsored by the Rome Air Development Center at Griffiss Air Force Base, New York, and was directed towards the establishment of a Russian-English MT system for the U.S.Air Force. Considerable modifications of the equipment were introduced at IBM, particularly in respect of dictionary creation, Russian text input, and graphical output of English text.

Although called the USAF Automatic Translator Mark I (Shiner 1958), it could scarcely be called a MT system since it still comprised little more than the large bilingual dictionary created by the Washington MT group. It is evident, however, that the Washington entries were being amended at IBM. King (1959) remarked that "we are now incorporating useful lexicon entries by modifying lists which have only recently been made available to IBM by the Rome Air Development Center". Differences of approach were evident even before King moved to IBM in 1958.

King shared an empirical (engineering) approach to MT: "The program at IBM Research has been to examine the question... from an operational point of view" (King 1961).[2] He believed that a basically word-for-word approach could be made quite acceptable, if problems of multiple meaning are solved by contextual information and if choices of 'best' equivalents are based on probabilistic criteria (King 1956).

Improvements to the Mark I system were made by the use of local context in the form of 'stuffing' procedures (King 1961). For example, the word *coping* can be either a verbal form or a noun (part of a wall). In *I shall be coping with...* it is identified as a verb by the preceding *be*. In the IBM system information in the dictionary attached to *be* causes the modification of *coping* as *a-coping*, thus signalling treatment as a present participle. The same 'stuffing' procedure could be applied to semantic problems. The example given by King is: "*The curfew tolls the knell of parting day*, where a class of words of which *day* is a member stuffs 'de-' in front of *parting* to form *departing*, thus resolving the many interpretations of *parting*". Furthermore, stuffing could improve English output; a simple example is the formation of verb forms: *drag* plus '-ed' becomes *dragged* after a prefix 'g' is stuffed in front of the ending ('prefix' is King's term.)

A major departure from the Washington 'lexicographic' approach was the decision to adopt the 'stems and endings' approach for dictionary entries. The arguments were presented by Micklesen, a member of the Washington team who joined IBM in 1960. Despite the large-capacity storage, it was practical on economic grounds not to require dictionary compilers to enter all paradigmatic forms, both because many may never actually occur and because errors from human fallibility were increased. Micklesen (1962) argued, therefore, that Russian words should be entered as stems as far as possible, the only exceptions being made for homographic stems, e.g. *dn* (either *day* or *bottom*), where differences conveyed by their full forms would be lost.

From June 1959 the Mark I version was translating Russian newspapers (Macdonald 1960); according to King (1963) about 10,000 words were translated daily and "were found... to be quite useful by the Government". As an example of a Mark I translation there is a passage from a Russian newspaper article (*Izvestia*, 14 August 1960) about the IBM system itself:

Shakespeare Overspat/outdid...

> Begin one should from that that in United States appeared new translation immortal novel L.N. Tolstogo "War and World/peace". Truth, not all novel, but only several fragments out of it, even so few/little, that they occupy all one typewritten page. But nonetheless this achievement. Nevertheless culture not stands/costs on place...

Unlike some other MT researchers, King did not conceal some of the poor results of the IBM system; he never made high claims for his system and was always prepared to acknowledge its deficiencies. King (1961) admitted lack of syntactic parsing, and considerable problems of semantics. He thought, however, that improvements could be achieved by exploiting statistical procedures given the "50% redundancy in meaning content" in texts. He advocated also the production of 'pidgin' translations, output including constructed words of multiple-meaning which could be learnt as if a dialect of English by recipients of translations (i.e. a variant of the notion of MTese, ch 2.4.3 above); and related to the ideas of the Cambridge group on 'pidgin' translations

[2] For further information on King and the IBM system see J.Hutchins: 'Gilbert W. King and the IBM-USAF Translator', *Early years in machine translation: memoirs and biographies of pioneers*, ed. W.J.Hutchins (Amsterdam: John Benjamins, 2000), 171-176.

(5.2 below). Evidently, King was content to produce something however deficient; he thought there was no point in worrying about difficult problems, such as the translation of pronouns and the insertion of articles, if the text was 'readable' and conveyed the basic information.

Although there were some developments on the linguistic aspects of the system, the main improvements were technical, in particular a considerably improved photoscopic store (a large-capacity 10 in. glass disk, which revolving at 2400 rpm. was considerably faster than the Mark I version at 23 rpm.) In 1964 the IBM system was demonstrated at the New York World's Fair (Bowers & Fisk 1965). In the same year this 'Mark II' version was installed in the USAF's Foreign Technology Division at the Wright-Patterson Air Force Base, Dayton, Ohio, where it remained in daily operation until 1970 (when it was replaced by Systran, see ch. 12.1 below). The central feature was still the vast Russian-English dictionary on the photoscopic disk, updated periodically by new disks. The logical capabilities of the system remained, however, rudimentary (Kay, 1973; Bowers & Fisk 1965). Each Russian item (either stem or ending) in the dictionary was accompanied by its English equivalent and grammatical codes indicating the classes of stems and affixes that could occur before and after it. Translation was word by word, with occasional backtracking, e.g. in *s gruppami* the identification of the ending *ami* as 'instrumental' would determine the selection of *with* to translate *s*. But in general, the selection of English output for a given Russian item was determined solely by the choices made for the immediately preceding item. Prepositions, copulas and auxiliary verbs were inserted in well-defined circumstances only; but there was no insertion of English articles (Roberts & Zarechnak 1974).

In addition to lexical entries, processing instructions were also intermixed in the dictionary: 'control entries' relating to grammatical processes (forward and backward skips), and also instructions relating to loading and printing routines. There were some 25,000 such 'control entries' included in the dictionary. This contained 150,000 entries at the World's Fair demonstration, and 180,000 in the USAF version. A small store of 3500 high-frequency entries was located in core memory, the remainder was on the photoscopic disk. A third of the entries were phrases, and there was also an extensive system of microglossaries. An average translation speed of 20 words per second was claimed (Bowers & Fisk 1965; Roberts & Zarechnak 1974).

The examples of Russian-English translations at the World's Fair were reasonably impressive, e.g. the one cited by Bowers & Fisk (1965):

> All this page is machine translation of Russian text, which is printed on preceding page. This page of translation from Russian into English is not perfect due to unsolved problems of grammar. Before machine can translate from one language into another, linguists have to introduce in memory unit of machine large quantity of grammatical rules, which increase intelligibility of given translation. But because in languages exists significant variety and complexity, all grammatical rules of any language are not developed completely at present time for use by computers.

However, this level of quality was not normally achieved. The limitations of word by word translation are more evident in other examples of Mark II output, e.g. the one cited by ALPAC (1966):

> Biological experiments, conducted on different space aircraft/vehicles, astrophysical space research and flights of Soviet and American astronauts with/from sufficient convincingness showed that short-term orbital flights lower than radiation belts of earth in the absence of heightened solar activity in radiation ratio are safe. Obtained by astronauts of dose of radiation at the expense of primary cosmic radiation and radiation of external radiation belt are so small that cannot render harmful influence on organism of person.

The Russian-English translations produced by Mark II were often rather crude and sometimes far from satisfactory. An evaluation in 1965 by Pfafflin (1965) tested the 'comprehensibility' and 'clarity' of MT versions of Russian electrical engineering texts and their human versions. Although ratings for comprehension were only marginally less for the IBM versions, the clarity of meaning was significantly lower. The general conclusion of participants was that MT versions were adequate only as a guide to determine whether a human translation was necessary. A later evaluation by Orr & Small (1967) also tested the comprehensibility of IBM translations, and this time included post-edited versions as well as 'raw' output in comparisons with human versions. There was no doubt about the result: "With a clear and remarkable consistency from discipline to discipline and from subtest to subtest, the post-edited translation group scores were significantly lower statistically than the hand-translation group scores; and the machine-translation group scores were significantly lower than the post-edited translation group scores." Nevertheless, it was noted that "a great deal of information was obtainable through the machine translations."

As one of the first operational MT systems the IBM Russian-English system has a firm place in the history of MT. But its historical significance goes further, since it was partly on the basis of Mark II translations that the members of ALPAC came to their controversial conclusions (ch.8.9).

Research on MT at IBM was not restricted exclusively to Russian. Macdonald (1960) gave a brief example of French-English translation by Mark I (based on a dictionary of 23,000 French words):

> The algebraic logic which is the subject of this course/s is conceived here as the part the most elementary (of) the mathematical logic. Later we/us will specify what we/us hear/mean signify by the word "algebraic". But one needs indicate immediately in what consists the mathematical logic whose algebraic logic constitutes the first part.

At this stage, the French-English system could evidently not cope with certain noun-adjective inversions or distinguish subject and object forms of pronouns. Bowers & Fisk (1965) mention that "considerable progress" had been made (presumably with a Mark II photoscopic disk) but it is not known what operational capability was achieved. This research was possibly done by Ascher Opler at the Computer Usage Company Inc. in New York under contract from IBM (ch.4.13). Bowers & Fisk (1965) also mentioned work on German, but information on this research is lacking.

Evidently more substantial progress seems to have been achieved with Chinese. From 1960 the IBM Research Center experimented with a Chinese-English system, operating on the same principles as the Russian-English system and also using the photoscopic storage device (King & Chang 1963). Some of the work on Chinese was done at Yale University with IBM sponsorship by Fang Yu Wang (ch.4.13). For the input of Chinese characters a special machine, the Sinowriter, was developed jointly with the Mergenthaler Linotype Company. Operators were required to break up characters into combinations of typical strokes and identify 'families' of characters. As in the Russian-English system, searches in the dictionary were made for the longest matching sequence of lexical items (characters). The system evidently included a certain amount of syntactic analysis (particularly for the identification of discontinuous structures) and rearrangement of the word order for English output. An example translation was included by King & Chang (1963):

> MODERN GUIDED-MISSILE ALREADY POSSIBLE CARRY WITH WAR HEAD OF HYDROGEN BOMB AND ATOMIC BOMB. THEREFORE IT IS ONE KIND WEAPON WITH VERY BIG POWER OF DESTRUCTION.

It was accepted that, with the IBM approach, "real translation is impossible." But "the translation can nonetheless be good enough to convey as much information as the original..." (King

& Chang 1963). There is evidence that research on this system continued after King had moved in 1963 to the Itek Corporation (Chai 1967); however, it is not clear how much progress was made towards a viable operational system.

After 1964 research at IBM was led by Micklesen, Tarnawsky and Chang. Attention turned to general linguistic problems and to the possible application of transformational grammar in Russian-English translation (*CRDSD* 13 (1964)). MT research at the IBM Research Center ceased in 1966 (Roberts & Zarechnak 1974).

4.3: Georgetown University (1952-1963)

Research on MT began on Leon Dostert’s return from the MIT conference in June 1952. His consultations with linguists and engineers at Georgetown led to the setting up of the experimental system in collaboration with IBM, which was designed to demonstrate the technical feasibility of MT. As we have seen (ch.2.5 above), the demonstration of this small trial Russian-English system in January 1954 was one of the most significant events of early MT history, arousing a great deal of interest and helping to stimulate U.S. government funding in subsequent years.

Dostert did not believe in the need for Reifler’s pre-editor; his aim was to show that suitable codes assigned to dictionary entries could signal both grammatical and semantic contexts. The theoretical justification was the concept of translation as involving two basic types of decision processes: selection and manipulation (Dostert 1955). Selection deals primarily with lexical data, “choosing the item in the output language corresponding correctly to the item in the input language”. Manipulation deals mainly with syntactic structure, “the modification of the sequence of items in the input text to fit into the structural pattern of the output language”. All MT groups, of course, recognised the need for some rearrangement of TL texts, but Dostert was perhaps the first to give syntactic analysis an importance equal to that of lexical procedures in MT systems.

The participants in the Georgetown-IBM experiment were, on the IBM side, Cuthbert Hurd and Peter Sheridan, and on the Georgetown side, Leon Dostert and Paul Garvin. The computational aspects were the work of Sheridan, a programmer in the IBM Division of Applied Science, and the linguistic aspects were primarily the work of Garvin, a linguist who had attended the public opening meeting of the MIT conference (Bar-Hillel 1960). A detailed account of the programming of the IBM 701 was given by Sheridan (1955), and a full retrospective account of the linguistic basis of the experiment was later given by Garvin (1967a). Other accounts were given by Dostert (1955), and briefly by N. Macdonald (1954).[3]

Forty nine sentences in the field of chemistry were carefully selected for the experiment to illustrate a variety of different constructions and problems. From an analysis of the sentences it was concluded that translation of this small corpus could be effected with a dictionary of 250 Russian words and just six rules of grammar.

The words included in the dictionary were divided into two classes: those which could be split into stems and endings, and those which could not. Stems and endings were entered separately. Each Russian word was given one or two English equivalents, and a set of codes indicating the selection rules to be applied and defining the context required for choosing the correct English output.

The selection rules of the Georgetown-IBM system were restricted to the immediate environment of the SL text, no more than one word on either side of the word being currently translated. The six rules were numbered 0 to 6 as follows (Macdonald 1963; Dostert 1955; Pendergraft 1967; Garvin 1967a):

0. The order of the original text is to be followed.

[3] See also C.A. Montgomery: ‘Is FAHQ(M)T impossible? Memories of Paul Garvin and other MT colleagues’, *Early years in machine translation: memoirs and biographies of pioneers*, ed. W.J.Hutchins (Amsterdam: John Benjamins, 2000), 97-110.

1. There is to be a difference of order in the translation from the order in the original, and an inversion is necessary.
2. There is a problem of choice; the choice depends on an indication which follows the word under consideration.
3. There is a problem of choice; the choice depends on an indication which precedes the word under consideration.
4. A word appearing in the original text is to be dropped, and no equivalent will appear in the translation.
5. At a point where there is no equivalent word in the original text, a word is to be introduced into the translation.

The English sentences produced in the demonstration were impressively idiomatic (Macdonald 1954):

Starch is produced by mechanical methods from potatoes
Magnitude of angle is determined by the relation of length of arc to radius
Processing improves the quality of crude oil

However, this idiomaticity had been achieved by the incorporation of many *ad hoc* decisions, particularly on the insertion and omission of lexical items (rules 4 and 5) in order to deal with 'idiomatic' usages. The restriction of the rearrangement rules to information from the immediate context was accepted solely for the purposes of the demonstration. It was realised that they would be quite insufficient in a larger-scale system.

Nevertheless it was believed that the principal operations necessary for MT had been demonstrated. After the experiment, Dostert summarized what had been achieved: (i) authentic MT had been shown to be possible, (ii) results showed that neither pre-editing nor post-editing were necessary, (iii) the problem of MT was primarily that of linguistic analysis, (iv) systematic MT dictionaries must include semantic feature codes as well as grammatical codes, (v) there was a need for specialised dictionaries for dealing with problems of polysemy, (vi) the development of intermediate languages for multilingual systems seemed feasible.

Although the Georgetown-IBM experiment aroused considerable public interest it did not result immediately in any official support for further research at Georgetown. From 1954 until early 1956, work was continued by a small team under Dostert.[4] About this time, the Institute of Precision Mechanics and Computer Technology of the USSR Academy of Sciences announced that research on English-Russian translation had been started on the basis of the approach adopted in the Georgetown-IBM experiment (a demonstration of which had been seen by Panov, cf. 2.5 above). A substantial grant was awarded to Georgetown University in June 1956 by the National Science Foundation. In fact most of the grant had come from the Central Intelligence Agency, as did most of the later support for the Georgetown MT research team. In late 1956 the full-scale project started, with Dostert as director and more than twenty researchers, on work towards a Russian-English MT system in the field of organic chemistry (Macdonald 1963).

The project team was organized in two research groups: one for 'translation analysis', the other for 'linguistic analysis'. The translation analysis group was concerned with the transfer operations by which the English output would be produced, i.e. primarily with the preparation of translation equivalences as a basis for the Russian-English dictionary. As far as possible, stylistic variety was excluded; the aim was to find the best single translation which would be satisfactory in most instances. The linguistic analysis group (later called the 'experimental group') was concerned with operations for the recognition of Russian input. It began with a detailed analysis of the Georgetown-IBM experiment. As a result, it was decided: (a) that input was to be in the Cyrillic alphabet, (b) that dictionary entries were not to be split into base and ending forms, (c) that the

[4] See also M. Vasconcellos: 'The Georgetown project and Léon Dostert: recollections of a young assistant', *Early years in machine translation: memoirs and biographies of pioneers*, ed. W.J.Hutchins (Amsterdam: John Benjamins, 2000), 86-96.

coding method was to be expanded, (d) that the six selection rules were to be recast and subdivided, and (e) that a corpus of 80 sentences as a test for the system.

Initially problems were to be solved one at a time and solved completely, while other known problems were to be ignored until later; however, this piecemeal approach was soon found to be impractical as the linguistic problems were closely interrelated. An overall strategy was necessary. There were, however, considerable divergences of viewpoint within the linguistic analysis group and so, by January 1957, Dostert decided to set up four groups each to pursue their own ideas. The resulting methods were to be tested in 'open competition' on a prepared text taken from the Soviet *Journal of General Chemistry* and on a randomly selected text from the same source. The understanding was that the method "which responded best in a practical situation would be favoured over the others" (Macdonald 1963). A tentative date for the test was set for early 1958. Some idea of these early approaches to Russian syntax can be found in the papers of a conference at Georgetown in April 1957 (Dostert 1957), e.g. Garvin (1957) on coding of cases and agreement, Pyne (1957) on rudiments of structural relations, and Zarechnak (1957) on a typology of sentences.

The four methods to be tested were a 'code-matching' method proposed by Ariadne Lukjanow, a 'syntactic analysis' method by Paul Garvin, a 'sentence-by-sentence' method by A.F.R. Brown, and a method of 'general analysis' by Michael Zarechnak. The methods differed primarily in the treatment of syntactic relationships, but different practices arose also over whether lexical items were to be entered unsplit or as separate stems and endings; the general analysis and syntactic analysis groups used split entries, the code matching group used split ones.

In the code-matching method each dictionary entry was first assigned codes indicating grammatical functions and semantic classes. Analysis proceeded word by word through the text from left to right; the specific function applicable in the given text was selected by comparison of codes of contiguous words; mathematical processes were applied to strings of codes to select codes of the English output. The method was tested in August 1959. "The translation was excellent. No information was provided as to how it had been achieved. Other workers at Georgetown hazarded the guess that the procedures were almost entirely *ad hoc*... Miss Lukjanow was reticent in discussing her methods and did not produce a translation of either a random text or a prepared text of any greater length" (Macdonald 1963). The project reviewed the method and decided that, while some code-matching was necessary in MT, complete reliance on code-matching was unsophisticated and clumsy, and limitation to a single left-to-right pass was unnecessarily restrictive. Shortly afterwards, Lukjanow left Georgetown and joined the Corporation for Economic and Industrial Research (Arlington, Va.) where she continued her research for a while, reporting on its general principles at the Los Angeles conference in February 1960 (Lukjanow 1961).

The syntactic analysis method proposed by Garvin analysed sentences in terms of immediate constituents (Ch.3.4), concentrating at each level of analysis on the item conveying most grammatical information, called the 'fulcrum'. The method was not ready for testing by the date set, and in fact was not tested at all, since Garvin left Georgetown in March 1960 to go to Thompson Ramo Wooldridge (Los Angeles), where he developed further what later became known as the 'fulcrum technique' of syntactic analysis (see 4.6 below)

The 'general analysis' method identified relationships between elements at three levels: morphological analysis (word-formation, identification of word classes (case, number, person, etc.), identification of idioms), syntagmatic analysis (agreement of adjectives and nouns, government of nouns by verbs, modification of adjectives, verbs and adverbs), and syntactic analysis (identification of subjects and predicates, relationships of clauses). This method was the only one ready for any sort of test by early 1958, and then only on one prepared sentence, although admittedly of particular complexity. (The test was reported by Zarechnak (1959)). A second, more rigorous test, of the method was conducted in June 1959 on a prepared text of 100,000 words in the field of organic chemistry and on a random text of 1500 words. The results were examined by an independent chemist at Georgetown who concluded that "the texts conveyed the essential information although their style was clumsy and the reading of them was time-consuming". This was considered a success, and significantly, at this point the General Analysis Technique was renamed 'Georgetown Automatic Translation' (GAT); it was this method which was to be developed in later years.

There remained the 'sentence-by-sentence' method of A.F.R.Brown.[5] This was designed for French-English translation. In a lecture given to the Association for Computing Machinery in June 1957, Brown (1958) reported that by January of that year he had devised rules for dealing with 220 sentences in chemistry. He described his method thus: "I opened a recent French chemical journal at random, went to the beginning of the article, and set out to formulate verbal rules that would translate the first sentence. It had about forty words, and it took ten hours to work out the rules. Turning to the second sentence, I added new items to the dictionary, invented new rules, and modified existing rules until the system would handle both sentences. The third sentence was attacked in the same way, and so on up to 220." (There could be no better description of the 'pure' cyclic approach; cf. 4.4 and 8.2) Brown was confident that in this way "most of the major difficulties have been met and solved" for French, and that "further progress... should be very rapid." By June 1957 the program had been coded and tested on an ILLIAC computer. (However, dictionary lookup had not yet at this stage been mechanized.) In the programming for moving, substituting and rearranging elements much use was made of sub-routines which in Brown's view were "so general as to be almost independent of what languages are concerned", a feature which he emphasised in later developments.

Two years later, in June 1959, the system was ready for testing (at the same time as GAT). On a prepared French text of 200,000 words and a random text of 10,000 words the results were considered to be nearly as acceptable as those for GAT. Later the same month, at the Paris Unesco conference, Brown gave a demonstration of his French system; this was the first public demonstration of a MT system with an unprepared text. By this time, the method was developing definitely into a general programming system designed to provide facilities for various linguistic and MT operations under full control of the linguist, who was able to alter and expand data and rules whenever desirable. In recognition of this development, Brown's system was renamed the Simulated Linguistic Computer (SLC).

The computer implementation of the GAT method, the SERNA system, was largely the work of Peter Toma (1959), initially alone. Toma had joined the Georgetown project in June 1958 to work on dictionary searching and syntactic analysis in Zarechnak's group.[6] (Toma had worked previously at the California Institute of Technology and for the International Telemeter Corporation on the Mark I system under Gilbert King.). Toma and his colleagues obtained access to the Pentagon's IBM 705 computer during its 'servicing time', and between November 1958 and June 1959 worked continuously throughout every weekend (Toma 1984). According to Toma, the test of GAT in June 1959 was run on the Pentagon computer.

There is some controversy over the significance of Toma's contribution to the Georgetown system. Toma claims that SERNA, acronym of the Russian 'S Russkogo Na Angliskij' (*from Russian to English*), was entirely his own creation, but Zarechnak (1979: 31-32) contends that Toma's responsibility was limited to coordination of the programming efforts while Zarechnak had overall responsibility for the linguistic formulations. While this may be true, there is no denying that Toma's programming skills made possible the "first significant continuous outputs for Russian to English", as Dostert readily acknowledged (in the preface to Macdonald 1963).

---

[5] For Brown's account of his MT work see A.F.R.Brown: 'Machine translation: just a question of finding the right programming language?', *Early years in machine translation: memoirs and biographies of pioneers*, ed. W.J.Hutchins (Amsterdam: John Benjamins, 2000), 129-134.

[6] See also P.Toma: 'From SERNA to Systran', *Early years in machine translation: memoirs and biographies of pioneers*, ed. W.J.Hutchins (Amsterdam: John Benjamins, 2000), 135-145.

On 25th January 1960 a demonstration of GAT (SERNA) was staged at the Pentagon before representatives of government agencies, rerunning some of the earlier tests of the Russian-English translations of organic chemistry. Early in 1961 the programming system for GAT was converted for use on the IBM 709. The opportunity was taken to introduce certain improvements in the efficiency and accuracy of the operations. As a result, so many alterations of the SERNA programs were necessary that in effect there was a new system; it was now called the Direct Conversion programming system, and placed under the direction of John Moyne (1962).

Apart from Russian and French, research teams at Georgetown also examined other languages. Chinese was investigated by a team advised by John de Francis, producing in 1962 a Chinese-English MT dictionary using telegraphic code for Chinese characters, and starting work on a MT system for mathematics texts. There was some work on the comparative syntax of English and Turkish, and during 1961 some discussion about setting up a pilot project for English-Turkish translation (Macdonald 1963). Brown did a tentative study of Arabic-English MT on the SLC basis (Brown 1966). Much more substantial was the work of the Comparative Slavic Research Group set up in October 1961 under Milos Pacak. This group investigated Czech, Polish, Russian and Serbo-Croatian with the objective of establishing a common intermediary language, for use in MT systems for these languages into and from English.

By late 1961 the SLC French-English system had been adapted for Russian-English, and it could also now be run on the IBM 709. SLC was now no longer restricted to one specific language pair but it had become a generalized programming system (Brown 1966). As a MT system for French-English translation, the SLC method remained largely the special and sole concern of Dr. Brown (Zarechnak & Brown 1961); but as a programming system it was often used to support the GAT Russian-English system.[7] At the Teddington conference in September 1961, the demonstration of GAT was run on SLC only, since conversion of the SERNA programs to the IBM 709 was not yet complete. As a result of this demonstration, EURATOM (at Ispra, Italy) decided to install the Georgetown system using SLC programming, both for producing translations for their personnel and as a basis for further research (ch.11.1 below).

Another demonstration of GAT was conducted in October 1962 at the Oak Ridge National Laboratory, under the auspices of the U.S. Atomic Energy Commission. This time the texts were in the field of cybernetics, using both prepared and unprepared texts. An example translation is taken from Dostert (1963):

> By by one from the first practical applications of logical capabilities of machines was their utilization for the translation of texts from an one tongue on other. Linguistic differences represent the serious hindrance on a way for the development of cultural, social, political and scientific connections between nations. Automation of the process of a translation, the application of machines, with a help which possible to effect a translation without a knowledge of the corresponding foreign tongue, would be by an important step forward in the decision of this problem.

It was admitted that the system, developed primarily for the field of organic chemistry, had problems with the new vocabulary and style of cybernetics literature, but clearly there was confidence in the Georgetown team's ability to improve the programs and dictionaries, and the Oak Ridge authorities decided to install GAT for producing internal translations.

In the event, the GAT Russian-English systems were installed at Ispra in 1963 and at Oak Ridge in 1964 at or just after the termination of the Georgetown project in March 1963. So came to an end the largest MT project in the United States. Some MT research on a Russian-English system continued at Georgetown under the direction of R.Ross Macdonald after 1965 (Josselson 1971), but it was to be on a much smaller scale and without CIA sponsorship. The reasons for the unexpected

[7] For a description see M. Zarechnak: 'The early days of GAT-SLC', *Early years in machine translation: memoirs and biographies of pioneers*, ed. W.J.Hutchins (Amsterdam: John Benjamins, 2000), 111-128.

withdrawal of support in 1963 are unclear. Zarechnak (1979) believes the official version citing unsatisfactory quality was not completely honest, while Toma (1984) alludes to internal conflicts between linguists and programmers leading to wholesale resignations. Whatever the cause, there could be very little further experimental development of the Georgetown systems after their installation at Ispra and Oak Ridge. Indeed, they remained virtually unchanged until their replacements by Systran (ch.12.1) at Ispra in 1970 and at Oak Ridge in 1980.

However, there were also more intrinsic reasons for the end of the development of the linguistic aspects of the Georgetown systems, which are to be found in consequences of the general 'empirical' approach of Dostert. He outlined the basic working premisses of the project as being "text-focused in the sense that the lexical buildup and the structural inventory is essentially... text-derived", the analysis of actual data and the "progressive improvement of experimental runs by means of a feedback procedure". The developments of the GAT analysis programs were, therefore, cumulative: programs were tested on a particular corpus, amended and improved, tested on another corpus, amended again, and so forth. The result was a monolithic grammar of "monstrous size and complexity" with no clear separation of analysis and synthesis (Kay 1973). Despite its complexity, syntactic analysis was very rudimentary, devoted to nothing more than resolving ambiguities in the assignment of word-classes through examination of preceding and following sequences of grammatical categories.

Although the GAT system of analysis was described as consisting of three, later four, 'levels' (Zarechnak 1959; Zarechnak & Brown 1961; Zarechnak 1962), what was involved was essentially a number of 'passes' seeking to establish different types of interrelationship. After initial dictionary lookup (GAT used two dictionaries, one of unsplit common words such as prepositions, conjunctions, pronouns, but also some irregular nouns, and another, larger, dictionary of stems and endings) came 'morphemic analysis'. This established relationships between adjacent words (e.g. identification of word classes, cases, number, person, etc.), thereby resolving some homograph problems (Pacak 1961, 1966, 1967). The next stage was 'syntagmatic analysis', which identified three types of word-combination (agreement, government, and apposition). The third 'level' of analysis was 'syntactic analysis', which sought to establish the subject and predicate parts of the sentence. Finally, a fourth 'level' was added to refine the transfer of Russian cases in English by identifying certain phrase structure relationships. Though the methods made good use of such traditional grammatical notions as agreement, concord, subject-predicate, etc., there was yet no clear conception of grammatical rule or syntactic structure. The rules were consequently *ad hoc*, formulated to deal with particular structures as they were encountered in texts being examined. Above all, the rules were designed solely for one-directional Russian-English translation. Analysis was intended to provide the grammatical classifications which were to select from the possible English alternatives and to decide the eventual word order. "Such information about the structure of Russian and English as the program used was built into the very fabric of the program so that each attempt to modify or enhance the capabilities of the system was more difficult, and more treacherous than the last" (Kay 1973). Indeed, at Ispra it was decided to develop a new system rather than modify the existing one (ch.11.1) and at Oak Ridge the only changes made to GAT concerned those necessary for transfer from an IBM 7090 to an IBM 370; the linguistic aspects were left untouched (Zarechnak 1976).

Nevertheless, operators of the two systems have regularly reported the satisfaction of users with the output even when receiving unedited versions (Jordan et al. 1977; Kertesz 1974; Perschke 1968). In 1973 a questionnaire was sent to users of both the Ispra and the Oak Ridge installations, and used in conjunction with interviews and records of users' reactions to the unedited translations (Henisz-Dostert 1979). Over 90% of the 58 respondents rated the translations as 'good' or 'acceptable'; 82% found that information had been translated sufficiently comprehensively; 59% found the translations 'readable' (although they took 32% longer to read than human translations). Only 19% thought that there was a possibility of misinformativeness; while 80% did not believe they had experienced it. Respondents judged that 87% of the sentences had been correctly translated (even if understood only with difficulty) and that 76% of the technical words were correct or understandable. In sum, 96% of the users said they would recommend MT to colleagues. It is quite apparent from these perhaps surprisingly favourable results that many users would rather have low quality MT than no translation at all.

4. 4: RAND Corporation (1950-1960)

Research on MT at the RAND Corporation began early in 1950 with the very tentative experiments by Abraham Kaplan which Olaf Helmer reported on at the 1952 MIT conference (ch.2.4.3). As we have seen, Kaplan took up Weaver's suggestion that statistical studies of context might assist the resolution of multiple meaning. Despite its limitations and the tentativeness of the conclusions, this study encouraged the belief that problems of ambiguity in MT could be overcome and that statistical analyses were a valuable research method for MT.

However, a research team was not in fact set up at RAND until 1957, headed by David G. Hays[8] with Professor Kenneth Harper of UCLA as consultant. Harper's interest in MT had begun in 1953 with a paper on the morphology and syntax of Russian mathematics texts. This was also a statistical study, from which Harper concluded that Russian word order was roughly the same as in English, therefore word-for-word translation was feasible for this kind of text.

Also influential in the setting up of the project were the early papers by Oswald - also on the faculty of UCLA – firstly the paper with Fletcher (1951) on problems of syntactic analysis of German, which advocated the coding of grammatical functions for the identification of 'noun blocks' and 'verb blocks'; and secondly the paper presented at the 1952 MIT conference in which he put forward the concept of microglossaries for specific subject fields established from statistical analyses of the relevant literature.

These researchers shared a common conviction in the value of statistical analyses of grammatical data. Detailed empirical analysis was a dominant feature of research at RAND. From the beginning the team had a clearly formulated research methodology (Edmundson & Hays 1958; Harper & Hays 1960; Hays 1963): a large corpus of Russian physics texts was prepared, a text was punched up onto IBM cards, a glossary for every text word was prepared providing grammatical information and English equivalents, the computer (IBM 704) program matched text and glossary to produce a rough translation, the result was studied by post-editors who indicated errors of coding and translation, the post-edited text was analysed, the glossary and translation rules were refined in preparation for the next text, and so the process was to continue in cycles of translation and post-editing. Obviously the first rough translations were crude word-for-word versions and much post-editing was required; but it was hoped that the successive incorporation of refinements in the glossary and translation rules would result in a "gradual elimination of the post-editor". Until then, the post-editor was a linguistic analyst, supplying the correct English equivalent (preferring the first one offered if it was satisfactory, even if were not the best!), adding correct English inflections (e.g. plural forms), inserting pronouns, verb auxiliaries, and articles, and indicating grammatical relationships. (Edmundson et al. (1961) provided detailed instructions.)

Within a few years the RAND team had an impressively large corpus of texts, some 250,000 running words (Harper & Hays 1960; Harper 1961), compiled partly in cooperation with the University of Michigan project (4.5 below). It was a firm conviction of the RAND group that deficiencies in MT analysis programs were attributable to lack of data: "existing grammars of languages are incomplete, ill-adapted for purposes of machine translation, and prolix with regard to a specific area of discourse. In addition, we have felt that the determination of sentence structure would prove useful in dealing with both grammatical and semantic problems of machine

[8] For Hays' contribution to MT, to linguistic theory, and to computational linguistics see M.Kay: 'David G. Hays', *Early years in machine translation: memoirs and biographies of pioneers*, ed. W.J.Hutchins (Amsterdam: John Benjamins, 2000), 165-170.

translation" (Harper & Hays 1960). The team concentrated therefore on establishing word classes and sentence structures empirically, with no preconceptions about the 'correctness' of the results. Working on post-edited texts they employed two main methods: dependency analysis and distributional analysis.

Dependency analysis produces a tree-like representation of sentence structure, in which there is one word which is independent (usually the finite verb) and every other word depends on one and only one other word in the same sentence (ch.3.4 above.) In the RAND formulation, dependency is "partly syntactic, partly semantic. Syntactically, one occurrence depends on another if the inflection of the first depends on the nature of the second; semantically, one occurrence depends on another if the meaning of the first complements or modifies the meaning of the second" (Harper & Hays 1960). It is acknowledged that the concept is "closely akin to the immediate constituent analysis of Harris, and to the approach suggested by Oswald". When implemented in a computer program, the method searched for dependency relations in a series of passes, e.g. in the analysis of *I saw a red rose*, the dependency of *rose* on *saw* was established only after it had first recognized the dependency of *red* on rose and of *a* on *rose*.

Distributional analysis continued the study of Kaplan on contextual resolution of ambiguity. Harper (1956) held that even contiguous words can help in syntactic clarification: "there are no serious problems of syntax which cannot be resolved by reference to the grammatical features of pre- or post-words". They can help also in semantic clarification, e.g. the English equivalent for the Russian conjunction *i* (*and, but, also, even*) could be found, he claimed, in 90% of cases by reference to the grammatical classes of contiguous words; likewise for the preposition *po* by reference to preceding noun-classes. The objective of distributional analysis, therefore, was to subcategorize classes for greater specificity of syntactic codes. Analysis started from dependency pairs of very broad classes e.g. subject-verb, adjective-noun pairs; and examination of the data from post-editing lead to the identification of specific sub-classes. For example, subjects can be a) animate nouns or first-person pronouns, in the nominative case, or b) animate nouns, in the instrumental case; and verbs can be a) active verbs, or b) passive verbs or passive participles, or c) first-person verbs. The result would be a list of verbs which may have animate actor dependent nouns. Distributional analysis was employed by Hays (1960) for the identification of 'semi-idioms' or phraseological constructions such as *depend on, consist of, give up, play a role*; and as the basis for the identification of transformations: the recognition of a "pair of dependency types, linking different grammatical types, but equivalent in meaning".

For Harper distributional analysis was the essence of the empirical approach: "Distributional classes represent facts of language, but the optimum procedure for analysis of this very large number of facts must probably come from experience. We must simply begin..."; and it had to be accepted that membership of classes will change with new data. The task was seen as a long-term effort; by limiting their data to physics texts meant that the grammar produced by the RAND team might not be applicable to other subjects (i.e. Harper recognized the importance of sublanguage grammars, as well as subject-specific glossaries, ch.3.2).

Research at RAND was thus always basically devoted to fundamental linguistic data processing rather than operational MT systems; and its results, data and dictionaries were made freely available to other researchers in a spirit of scientific cooperation. By 1960 "it had become clear that the broader field of computational linguistics deserved general attention, and that machine translation should either be pursued as a developmental task, using existing techniques for a practical purpose, or set aside to await progress of a fundamental nature" (Hays 1967). Research continued on dependency grammar and on grammatical theory; a large text corpus of Russian was created, nearly four times the size of the previous one and with fuller syntactic information; and numerous improvements in computer handling of language data were developed. A later project at RAND relevant to MT was the development of the MIND system (ch.17.8)

4. 5: University of Michigan (1955-1962)

Research at the Willow Run Laboratories of the University of Michigan in Ann Arbor began in 1955 under the directorship of Andreas Koutsoudas and continued until 1962. (Bar-Hillel 1960, Josselson 1971). The Michigan group had close contacts with the RAND team, sharing a similar empirical approach to MT methodology and cooperating in the compilation of a large text corpus of Russian physics literature.

Like many other MT groups in the mid-1950's Michigan suffered from journalistic extravagances. In 1956 Koutsoudas felt obliged to respond to one such report in an Associated Press dispatch: "We wish to disclaim to our professional colleagues any responsibility for the publicity release on our mechanical translation project which was printed in today's newspapers. The implication that we are very far along toward practical translation was due to an overenthusiastic misinterpretation by a journalist" (*MT* 3(2) Nov 1956). In fact, the Michigan group concentrated primarily on statistical analyses of its Russian corpus and did not implement even an experimental MT program. Nevertheless the team was so sure that "within a generation machine translation will be a fait accompli, as will machine reading" that it could devote some of its time to the standardization of printing fonts for future OCR machines (Koutsoudas & Machol 1956).

As at RAND, the work of descriptive linguistics was treated with much scepticism. It was argued that not only were existing descriptions of individual languages incomplete "particularly where syntax is concerned" and deliberately avoided semantic considerations (for the sake of formal rigour), but also that the different aims of MT research required the development of rules or procedures to connect *pairs* of languages (Koutsoudas 1960). The aim of statistical analyses was to establish appropriate contextual clues for identifying word classes, discovering syntactic relationships resolving homonymy, and identifying idiomatic expressions. Its treatment of homonyms was typical of the empirical approach; it consisted in "translating small groups of words, listing in the dictionary multiple meanings under each word in the group, and finding algorithms which make it possible to choose the proper set of meanings for the group" (Koutsoudas & Machol 1956). The complexity of the approach, even for a small set of words, is clearly demonstrated in Koutsoudas & Korfhage 1956) where English equivalents are given for every context (i.e. every word with which it may occur in a phrase) even if the same translation has to be repeated many times and, in some cases, zero translations have to be noted, e.g. in a vocabulary subset of just 8 items *teori* is given 3 translations, *theory* twice and zero once.

A further example of the Michigan approach is to be found in Koutsoudas & Humecky (1957). Part of the results concerned an analysis of Russian forms ending in *-o*, *-e* or *-ee*. It was found that these could be adverbial modifiers if (i) preceded by period or comma and succeeded by a comma not followed by *sto* or *sem*, or (ii) followed by a non-infinitive verb form when not preceded by *kak*, or (iii) preceded by *esli* or auxiliary verbs and succeeded by an infinitive; and so on for another five examples of conditions. Attached to each statement was an indication of its frequency of occurrence in the corpus.

In the face of such masses of detail for even such small segments of Russian grammar it is not surprising that the Michigan group did not implement a MT system – getting no further than some tentative hand simulations – and so it turned increasingly to speculations on the design of a computer program which would learn to modify itself. Inspired by current research on information theory and on automata theory, the idea was that if a MT program were to translate all except one of a set of sentences correctly it should be possible to devise a self-modifying component which would automatically amend the MT program to translate all the sentences correctly, and without introducing excessive complexity (Koutsoudas 1961). Apparently, even an experimental prototype of such a learning program was not attempted, only the mathematical theory was elaborated, based on a categorial grammar approach (Smoke & Dubinsky 1961).

4. 6: Ramo-Wooldridge and Bunker-Ramo (1958-1967)

Research at the Ramo-Wooldridge Corporation (later the Bunker-Ramo Corporation and part of the Thompson-Ramo-Wooldridge group) in Canoga Park, California, started in 1955, receiving a grant for an initial study from the US Air Force in March 1957, and beginning full-scale research in March 1958 on a Russian-English system. Until 1960 the research group was directed by Don R.Swanson (Bar-Hillel 1960); in March 1960 it was joined by Paul Garvin from Georgetown University (Macdonald 1963), who as its director steered its activities towards development of his ‘fulcrum’ method of analysis. Before this, the Ramo-Wooldridge group had been most strongly influenced by the RAND approach (Professor Harper acted as a consultant in its early stages), sharing a strongly empirical attitude and a cyclic approach to the development of systems. In addition it shared the same interest initially in the corpus of Russian physics texts built up by RAND in cooperation with the Michigan group, although later it turned to experiments on translating newspaper articles from *Pravda*.

Like the RAND group, it saw its research work as “not simply the production of a machine translation program, but mainly to do linguistic research, primarily with emphasis on semantics, by use of computing machinery” (Ramo-Wooldridge 1958). It was, nevertheless much less committed than RAND to fundamental linguistic research per se and more determined to set up a working (albeit experimental) MT system. “It is not our belief that one attempts to solve all problems before going to the computer...we start by using the expert knowledge of both Russian and English that members of our research team have in order to gain initial understanding of the problem and its probable solution. We then mechanize the solution and test it out on our computer.” (Mersel 1961) The emphasis was on those problems which occurred frequently, not those of greatest intellectual difficulty: “the mere running of large amounts of text will by itself not solve any problem”, however, “the running of large amounts of text serves to test previous solutions, allows one to get a better feel for the most frequent problem areas... and gives sufficient examples of words in actual usage to allow an attack on their multiple meanings.” The research technique was “a cycle of observation, idea, mechanization, test, correction and new observation” using both humans and machines with the ultimate aim of mechanizing the whole cycle.

A brief outline of its basic system and an example translation from *Pravda* (7 September 1960) is to be found in a 1961 report (Ramo-Wooldridge 1961). With characteristic pragmatism, translations were produced by successive refinements of an initial rough word-by-word rendition: “With each step, our automatic process gains new insights into the meaning of the Russian text. Each new insight is used to improve the English translation. Thus, starting with the Russian text, we progress from a jumble of English words to a succession of coherent English sentences”. The article was first keypunched directly in Cyrillic script; then automatically transliterated; in the following dictionary lookup stage Russian words were replaced in sequence by English equivalents or by alternatives (if the Russian was a homograph), producing a crude word by word version. Any words not found by dictionary lookup were then tested against a list of endings. In the following stages, the English forms were inflected to conform with Russian endings (e.g. verbs put into past tense), and identifiable idioms were translated as units. The next stage was described as “the heart of our machine translation process” allowing “us to go beyond a crude word for word translation to capture the essential meaning of Russian sentences.” Syntactic routines detected ‘pivot’ words (such as finite verbs and heads of noun phrases) and formed ‘packages’ on the basis of the grammatical information provided by the dictionary. Then there followed stages in which some alternative translations of ‘multiple meaning’ words were eliminated, certain combinations were recognised as ‘idiomatic’, e.g. *depend* and *on*, articles were inserted and prepositions were added on the basis of information from Russian case markers. Lastly, some rearrangement of ‘packages’ was attempted to give closer convergence to English word order, e.g. the postposition of a modifying participial phrase (*These considered by us serious problems*), and the inversion of subject and object (*By Prof. Suvarov was considered the problem of radiation*). The result was the following ‘translation’ (extract):

> IN THIS MESSAGE, IN PARTICULAR IT MENTIONED, THAT BY FBI AND ORGANIZATIONS OF THE CIVIL POLICE PROCEEDED WITH SEARCHES OF TWO EMPLOYEES/SERVING OF THE TOP-SECRET NATIONAL AGENCY OF THE SECURITY/SAFETY, WHICH/WHO DISAPPEARED IN THE END OF JUNE. WAS REPORTED, THAT THESE EMPLOYEES/COLLABORATORS ARE BERNON MITCHELL AND WILLIAM MARTIN, WHICH/WHO LONG/LENGTHY TIME/PERIOD WORKED IN THE NATIONAL AGENCY OF THE SECURITY/SAFETY, THE USA. VANISHED EMPLOYEES/COLLABORATORS OF NSA ARE/ARE-FOUND NOW IN MOSCOW AND APPLIED WITH/FROM REQUEST TO APPEAR ON PRESS-CONFERENCE OF SOVIET AND FOREIGN CORRESPONDENTS. MITCHELL AND MARTIN SOLVED/DECIDED TO BREAK-WITH/TEAR-FROM THE USA AND REQUESTED POLITICAL ASYLUM FOR/AT/BY/FROM THE SOVIET GOVERNMENT IN/ON/WITH-RESPECT-TO POLITICAL MOTIVES.

The text still contained numerous alternative versions not yet resolved, as well as, of course, many stylistic infelicities. Nevertheless, it could claim to be no worse than many other MT efforts at the time, perhaps no less comprehensible to those who accepted its limitations, and the system appears to have been relatively fast (an average speed of 1/150 seconds per translated Russian word was claimed).

Reference to ‘packages’ in the above description indicates further points of similarity with the RAND approach, and, as will become apparent later, it also reveals the increasing influence of Garvin’s ideas on syntactic analysis.

The debt to Harper was explicitly acknowledged in a 1958 report (Ramo-Wooldridge 1958) when discussing the technique of organizing sentences into strings of ‘nominal word blocks’, i.e. a noun and its dependent modifiers forming a linear ‘package’. The influence of RAND is reflected further in such statements as: “Thus in essence we approach syntactic problems unhypnotized by their reputation for complexity and account for most of them with an elementary (but still largely effective) model. The price we pay for such irreverence is that we are left with a clutter of exceptions to our otherwise orderly set of rules.” And indeed, whereas some rules were quite simple, e.g. “insert “OF” before genitive noun coded C” (i.e. first word of nominal block), and “insert “TO” before dative noun coded C”, others had to be remarkably complex. The following as a rule for pluralization: “add (S) following any ‘potentially genitive’ feminine singular nouns carrying code O” (i.e. word not first in block) “provided the preceding word is neither a preposition nor a ‘potentially genitive’ feminine singular adjective”.

The complexity is attributable essentially to the linear conception of syntactic relations (ch.3.4), even though by this time the group was beginning to recognize non-linear dependencies. The problems of dealing with syntactic analysis without an adequate model are nicely illustrated in this quotation: “In pondering the prolixity of Russian inflectional forms, and at the same time the number of ambiguities involving the genitive case, it is at times tempting to wonder whether the entire Russian language might not be grossly simplifiable merely through borrowing, adopting, leasing, stealing, or otherwise appropriating the English preposition “OF”!” (Ramo-Wooldridge 1958). Such exasperation may well have been felt by others but at least this team was honest.

Further illustration of the approach at Ramo-Wooldridge is to be found in the treatment of ‘multiple meanings’ and dictionary lookup procedures. As far as homographs were concerned, the group concentrated for practical reasons on the most common ones in the corpus, principally prepositions, and for this study it built upon investigations by Harper, formulating such rules as: “if

DO follows within eight words of OT in any sentence, translate OT as FROM and the Russian DO as English TO", "if ZA is followed by a symbol, translate as PER, otherwise FOR+DURING", and more complex rules such as: "NET – if OF was inserted before a potential genitive at the beginning of a word block preceding or following NET, suppress the English equivalent for NET and substitute THERE IS NO for OF". In the routines for dictionary lookup no single method was adopted: some items were entered as units (stem plus ending), others separately as stems and endings; in order to find the latter a routine for splitting stems from endings (morphological analysis) was needed. The reason for having both types of entry was primarily historical – "not known how well or how easily the splitting procedure could be made to work", and expansion of the glossary entailed too much work. However, subsequent experience showed that there was a practical advantage in that words could be found either directly or by morphological analysis. (For similar reasons later MT projects have adopted the same procedure, e.g. TAUM, ch.13.1.)

Like many others, the Ramo-Wooldridge team was optimistic about the future of MT: "During the period of two years we will have seen a 32-fold increase in lookup rate with only a doubling of computer cost. The cost of dictionary lookup will no longer present an economic argument against the practicality of machine translation" (Ramo-Wooldridge 1958).

The influence of Garvin was already discernible by February 1960 (Mersel 1961); Garvin had been a consultant to the Ramo-Wooldridge project while still at Georgetown.[9] The syntactic routine was now described explicitly as a series of passes and as a series of searches for "a predicative or gerund to use as the pivot of the sentence" in which "great use is made of the government characteristics of the predicative." (The basic system was much the same but the theoretical foundation was more secure.)

With Garvin's assumption of leadership in March 1960, the group turned more specifically to the development of his 'fulcrum' approach to analysis. Its formulation received progressive refinements over the years (Garvin 1961, 1963, 1967b, 1968, 1972, 1980). Garvin characterized his approach as 'problem-solving' rather than 'theory-oriented', by which he meant the "solution of translation problems by computational means through utilization of linguistic and systems know-how, rather than first trying to develop a theoretical 'understanding' of the translation process and then applying this understanding through some form of computational implementation." (Garvin 1980). In this way he distanced himself both from the empirical 'brute-force' approach and from the theoretical 'perfectionist' approach (ch.8.2 below) A feature of what he called his 'engineering' solution was the inclusion of grammatical information within parsing programs, rather than the practice of other MT groups which separated linguistic data and processing algorithms (Garvin 1967; see also ch.9.9). It was his belief that this approach produced greater computational efficiency in syntactic analysis. The 'fulcrum' parser is essentially a linguistic pattern recognition algorithm which directs searches at different portions of sentences in order to identify grammatical patterns. The sequencing of searches is crucial; at each stage (or 'pass') the algorithm uses only as much grammatical information as appropriate to the particular pattern being sought. In other words, the complex system of rules making up the 'grammar' of a language is distributed over a complex algorithm.

One consequence was that the algorithms were language specific, i.e. in this case procedures were specifically designed for the analysis of Russian. Another feature of the practicality of the approach was that it was 'translation-oriented', i.e. the method sought to produce not all possible parsings of a sentence (as in 'theory-oriented' systems, such as at Harvard and MIT, ch.4.9 and 4.7) but only the most probable. As a consequence, probabilistic values were assigned to syntactic structures, e.g. a sequence which could in theory be either subject-predicate-object or

[9] For Garvin's resarch at Georgetown and at Ramo-Wooldrige see C.A. Montgomery: 'Is FAHQ(M)T impossible? Memories of Paul Garvin and other MT colleagues', *Early years in machine translation: memoirs and biographies of pioneers*, ed. W.J.Hutchins (Amsterdam: John Benjamins, 2000), 97-110.

object-predicate-subject was most likely to be the former and this interpretation was assumed to be correct.

The two basic principles of the 'fulcrum' approach are the concept of the fulcrum and the pass method. The fulcrum of a structure is the element carrying the maximum amount of information about the structure and its relationships; it is the 'pivot' or key element which governs the other dependent components of the structure, e.g. head nouns of noun phrases determine the forms of articles and adjectives (case, number, gender). Structures are built up in a series of passes: each pass is designed to identify a particular kind of fulcrum and the elements which depend on it. Initial passes identify nominal groups (noun phrases), prepositional phrases and predicates (verb groups); later passes establish dependency relations between groups. For example, in:

> These various compounds of copper have been treated in the technical literature on many occasions

the algorithm identifies the fulcrum *compound* and its modifiers *these various* and *of coppe*r, and then later recognises the whole structure as dependent on the verbal phrase *has been treated*. At each stage, attempts are made to resolve semantic ambiguities by the examination of elements related syntactically to the ambiguous element. Thus, the dependency of *of copper* confirms that *compound* has its chemical sense and not its 'association' sense.

As we have already indicated, the fulcrum model is essentially an advanced development of the dependency model which had already been adopted by the group; Garvin's contribution was to formulate it more precisely and to implement it more thoroughly. Its implementation in a Russian-English system at Bunker-Ramo Corporation (as Ramo-Wooldridge had now become) was developed with the sponsorship of the U.S.Air Force (Rome Air Development Center) and the National Science Foundation. The system served as the testbed for MT research and the foundation for an advanced version 'Fulcrum II' (also sponsored by the Rome Air Development Center).

The new version (Garvin 1967b) was to be characterised by more efficient algorithms, an updated sequence of passes and searches for fulcra, the inclusion of iterative search sequences, the capacity for producing English output text, and in particular the introduction of heuristic techniques. Heuristic methods were seen as natural extensions of the 'problem-solving' approach to language analysis (Garvin 1962, 1968). The heuristic component was intended to test whether all mandatory analytical processes had been correctly completed. Thus, if a verb in a predicate can only be transitive and the recognition program has not identified an appropriate object then a mistake of analysis has occurred. The origin of the mistake could have been an incorrect decision about case ambiguity (e.g. a noun form which could have been nominative or genitive); the heuristic component was to search back for this decision point, revise the analysis and complete the structure.

The method could also use the statistical information on the probabilities of particular structures given certain configurations of grammatical forms. In the first version of the fulcrum analyzer the most probable interpretation was adopted. In Fulcrum II the lower probabilities could be tested if further information on higher-level relationships indicated the necessity. Thus the initial identification of a Russian form as a genitive singular noun might be revised in a broader context as a nominative plural noun. As Garvin (1968) points out, the heuristic component of the revised fulcrum approach was deliberately modelled on contemporary AI research on problem-solving systems. In this respect, Garvin foreshadowed more recent applications of AI techniques in MT research.

Garvin was unable to implement the improved version of the fulcrum approach because the MT project at Bunker-Ramo "was terminated in 1967 for lack of funds". He remains convinced that his approach was basically correct, partly because he sees recent implementations of the fulcrum method in some AI research, but more specifically because "only a problem solving approach such as the one I have been advocating can be expected to yield worthwhile results in the long term." (Garvin 1980)

4.7: Massachusetts Institute of Technology (1953-1965)

The Massachusetts Institute of Technology made the first appointment of a research worker in the MT field. As we have seen (ch.2.4), this was the appointment in May 1951 of Yehoshua Bar-Hillel, a mathematician at the Hebrew University of Jerusalem. For two years Bar-Hillel investigated the possibilities of MT, instigated meetings, published reviews, and made some important theoretical contributions. However, he did not himself do any practical research. In July 1953 he returned to Israel. Victor H. Yngve took over and set up the research project in the Research Laboratory of Electronics at MIT. It was funded from 1954 until 1965 primarily by the National Science Foundation, although support was also given by the U.S.Army Signal Corps, the U.S.Air Force Office of Scientific Research and the U.S.Navy Office of Naval Research.

From the beginning, MT research at MIT had a strongly theoretical bias. No attempt was made to construct interim systems which though serving some practical needs were producing output of low quality. The goal of research at MIT was throughout "to work toward the achievement of language translation completely by machine and of a quality that rivals translations made by hand" (Yngve 1967). The researchers at MIT were "not looking for short-cut methods that might yield partially adequate translations at an early date" but for "definitive solutions that will constitute permanent advances in the field." (Yngve 1961). At MIT research was focussed on advances in linguistic theory, particularly the theory of transformational grammar, and on the development of programming tools for linguistic research. In both areas, researchers connected with the MIT project made important contributions of significance going well beyond the specific domain of MT research. (A complete bibliography of the MIT group is given in Yngve 1967.)[10]

The need for good programming tools became apparent very early at MIT, as it did elsewhere (cf. the development of SLC at Georgetown, ch.4.3). At this time there were no high-level languages such as Fortran and Algol, which in any case were designed primarily for mathematical work and were not suitable for non-numerical applications. Such were the complexities of early computers that linguists engaged in MT relied on expert programmers to implement their procedures in assembler code, they did not attempt to program themselves. As a result neither linguist nor programmer were fully effective. At MIT it was concluded that the solution was a programming system which allowed the linguist to write his procedures in a notation specially devised to fill his needs. The answer was COMIT, the first programming language devised for string-handling and pattern-matching. It was the product of collaborative work with the Computation Center at MIT. Research began in 1957 and a version was ready the same year, thus antedating by two years or more the first reasonably full implementation of the programming language LISP, the list-processing and symbol-manipulation language also devised for linguistic research and later adopted particularly in Artificial Intelligence (cf.ch.15 and 18.2 below).

Yngve claimed (1967) that COMIT was learnt very easily by linguists ("about six one-hour periods are sufficient to train a novice"), and enabled the team to formulate ideas clearly and concisely. ("The availability of the COMIT notation greatly increased the productivity of the group even before we could run a single program in COMIT.") Sammet (1969: 435) corroborates this opinion, comparing COMIT favourably to LISP: "One of the outstanding things about COMIT is the discrepancy between its apparent and surface difficult notation and the actual ease of writing and using the language. This contrasts sharply... to LISP, whose notation is inherently simpler than COMIT but which seems to be much harder to learn and use." Sammet (1969: 416-436) gives a good general description of COMIT, its development, technical characteristics and theoretical

[10] For a retrospective account see V.H. Yngve: 'Early research at M.I.T.: in search of adequate theory', *Early years in machine translation: memoirs and biographies of pioneers*, ed. W.J.Hutchins (Amsterdam: John Benjamins, 2000), 39-72.

foundations. Descriptions of COMIT with specific reference to its MT applications have been given by Yngve (1958, 1960a, 1967: 454-465).

Yngve began his linguistic MT research in the expectation, shared by many at the time, that statistical techniques could be applied to descriptions of syntax, and, more generally, as we have seen (ch.2.5), that good MT output could be achieved by progressive improvements refinements starting from the "surprisingly good" word-for-word translations. He believed also that the 'information theory' of Shannon (Shannon & Weaver 1949) could be applied to problems of MT (*MT* 1(1/2), 1954).

In 1955 four linguists joined the MIT Research Laboratory to work under Yngve (*MT* 2(1) July 1955). One of them was Noam Chomsky, who had studied under Zellig Harris at Pennsylvania and had been a member of the Society of Fellows at Harvard, and who was later to be appointed Professor of Linguistics at MIT. During the time he was connected with the MT project, Chomsky wrote a number of influential papers on mathematical linguistics and also his 'revolutionary' book on transformational generative grammar (Chomsky 1957). In subsequent years, the syntactic theory of Chomsky was to underpin the theoretical foundations of MT research at MIT.

Yngve was one of the first to recognise the crucial importance of syntactic analysis for the production of adequate MT (Yngve 1955, 1955a, 1957). Whereas other groups saw syntax as primarily an adjunct to lexicographic transfer, as a means of resolving SL ambiguities and rearranging TL output, Yngve placed syntactic transfer at the centre of his MT model. It was argued that syntactic analysis was basic not only to the resolution of word order problems but also to any procedures involving relationships of words in patterns; this would include idiomatic structures, semantic restrictions on cooccurrence (e.g. of nouns and verbs), pronouns and their antecedents, and even the recognition of the 'field of discourse'. Translation was seen as a two-step process: a decoding of the input text into a representation of the 'message' (a 'transition language') and its encoding into an output text in another language. The conception was inspired at least in part by 'information theory'. To go beyond word-for-word translation required a method of specifying the structures to be transferred. Initially Yngve saw this specification in terms of a Markov finite-state model (Yngve 1955), but when the more complete formulation of the 'transfer' approach appeared the model was clearly based on the syntactic theory of Chomsky (Yngve 1957). The 'recognition routine' applied rules of SL grammar to identify word classes and phrase structures of input sentences; the 'construction routine' produced the output text applying rules of TL grammar; between them was the 'structure transfer routine' which converted input structure into corresponding output structure. One point stressed by Yngve was the distinction between 'stored knowledge' (the grammar) and routines (of analysis and synthesis) which applied grammatical rules: 'abstract' knowledge about the structures of languages was to be separated from details of programming designed for a particular MT system (cf.ch.9.9 below).

The MIT method demanded detailed investigations of the syntax of source and target languages before routines of analysis, synthesis and structural transfer could be devised. Most research was devoted to German and English (the project was funded to develop a German-English MT system), but members of the MIT team also studied French, Arabic, Finnish and Chinese. Much of this work was purely theoretical and not immediately applicable to MT as such. A prime example is the substantial research of Edward Klima on negation in English (1964), which made an important contribution to the theory of transformational generative grammar. Other prominent linguists connected for a time with the MIT project included Robert B. Lees (English nominal compounds), James D. McCawley (Finnish syntax), and Gilbert H. Harman (mathematical linguistics). The work of others, though based on linguistic theoretical foundations, was intended to be more practical in so far as they were designing routines for MT systems; however, the routines were rarely fully tested in actual implementations.

The method of syntactic analysis proposed by Matthews & Rogovin (1958) was designed to identify the phrase structure of input sentences by discovering "those rules of grammar which must have been applied in order to produce that particular sentence". In practice, it was a bottom-up parser recognising immediate constituents and building up to a complete phrase structure representation (ch.3.4 above). A distinctive feature was the succession of alternating left-to-right and right-to-left passes through sentences, designed particularly to cope with problems of German syntax, e.g. the location of verbal prefixes and participles at the end of sentences (*Er kommt heute zurück; Er hat es nicht gesehen*) and of finite verbs at the end of subordinate clauses (*Er sagt daß er heute nicht zurückkommt*). After 'initialization' (recognition of word boundaries), dictionary lookup started at the end of the sentence: first the final word was sought in a small dictionary of separable prefixes and then the other words of the sentence were sought in the main dictionary in a right-to-left scan. Compounds were treated on the lines indicated by Reifler (1955, cf.4.1 above.) The next scan, left-to-right, located the finite verb and placed it at the end of the sentence (conjoined verbs and prefixes could now be sought in the dictionary). The following right-to-left scan identified dependent structures (subordinate and relative clauses, participial phrases), e.g. by punctuation and conjunctions, and placed them at the beginning of the sentence, marking their original position in the main clause (or, in the case of nested structures, in the dependent clause). The next scans identified the boundaries of noun phrases and prepositional phrases, and then the type of sentence (active or passive) and the subject noun phrase (on the basis of case endings). Finally, the structures of dependent clauses were analysed in the same way.

The program was written in COMIT for the MIT's IBM 704 computer. The authors did not claim that it could analyze all German sentences; but they thought the basic conception was sound, and improvements easily incorporated. Whether there was any more research on this method is not clear. In the event, by 1961 Matthews (1962) had formulated a more advanced parser based explicitly on the 'analysis by synthesis' approach, i.e. that phrase structures could be established by discovering the rules by which they had been generated (in the sense of Chomskyan generative rules). The basic idea was that all the possible phrase structures should be derived which terminate in the same number of symbols (word classes) as in the input string (the sentence to be analysed), and that then the derived phrase structures should be compared symbol-by-symbol with the input string. Those that matched were to be considered plausible analyses of the sentence. The method was admitted to be inefficient, but it was claimed that it could be improved by incorporating various short-cuts, e.g. by not deriving structures shorter than the input string, and by doing a 'preliminary analysis' of 'linguistically significant' features in order to define the class of derivations to be attempted. Nevertheless, the practical value of the approach remained uncertain.

The fullest exposition of the MIT approach was probably the detailed specification of an Arabic-English system by Satterthwait (1965). The complex problems of morphological analysis of Arabic were carefully examined. In syntactic analysis, Satterthwait returned to the 'immediate constituent' model, particular attention being paid to the identification of discontinuous structures (a major difficulty in Arabic). For the production of English sentences, Satterthwait adopted Yngve's mechanism for generating English sentences based on (i) information on the structural relations in the Arabic sentence, (ii) information from the bilingual dictionary about structural constraints on items of English vocabulary, and (iii) rules for the generation of English phrase structures.

Yngve's sentence-production routine for English operated on unordered lists of phrase structure rules, working from top to bottom, and motivated by specific requirements of vocabulary items (Yngve 1967). A by-product of this program was the formulation of what is perhaps Yngve's best-known contribution to linguistics, the 'depth hypothesis' (Yngve 1960). This is the discovery of the 'principle' that while there seems to be no limit to the number of embedded clauses in English if they are right-hand constituents (*This is the cat that killed the rat that ate the malt that...* etc.), there is a limit to the number of left-branching embeddings – a limitation determined by the restricted span of the human memory. The hypothesis explained various phenomena of English syntax: the difference in form and function of active and passive sentences, the function of the anticipatory *it*, the positions of interrogative pronouns, adverbs, direct and indirect objects, and so forth. It is an example of the reciprocal influence of MT research and linguistics that Yngve hoped for in the theoretical orientation of the MIT project.

Inevitably perhaps, the intensive exploration of the syntactic approach eventually discovered its limitations. By the mid-1960's it was clear to Yngve, as it was to most MT researchers (ch.8.8 below) that MT research had come up against the "semantic barrier" (Yngve 1964); syntactic analysis still left many ambiguities, which in Yngve's view could only be resolved if the machine can 'understand' what it is translating. As first steps in this direction, Yngve saw the work at MIT of Elinor K. Charney (1962) on the semantico-logical interpretation of sentence structures and of Jared L. Darlington on the translation of ordinary language into the terminology of symbolic logic. It is true that this work pointed ahead to the substantial research elsewhere in later years on the interrelations of logic and linguistic semantics, although as yet the findings have had little application in MT research.

There is no denying the importance of the contributions of the MIT project to the theoretical foundations of MT and of related spheres of linguistics and computer programming, but the prospect of any operational MT system receded more and more as the years passed. In 1964 research at MIT came to an end (along with other US projects) and in 1965 Yngve moved to Chicago. The NSF continued to fund his MT research for a while, but Yngve's interests turned increasingly and ever more strongly towards problems of theoretical linguistics. In the afterword to an article originally published in 1964, Yngve (1979) characterizes his current research as the development of "a linguistic theory... that focuses on people rather than on language", requiring "a complete restructuring of linguistics from the ground up". The goal of MT is not forgotten, however, since "the approach promises to provide a solid foundation for research in mechanical translation and computational linguistics." Evidently, Yngve has remained true to his guiding principle of MIT research, a conviction that MT must be built on theoretically sound linguistic foundations.

4. 8: National Bureau of Standards (1959-1963)

The Los Angeles NBS office had done some very tentative work in the early 1950's (ch.2.4.1), but there had been no follow up. A research project was set up at the Washington headquarters of the National Bureau of Standards in 1959. This was the project on Russian-English translation directed by Ida Rhodes (Rhodes & Alt 1962).

There had been interest in problems of syntactic analysis at NBS since 1957: Richard B. Thomas and P. I. Herzbrun had experimented on the SEAC computer with statistical analysis of syntactic patterns of English (Thomas 1957; NBS 1957). The initial impetus was an interest in the improved mechanized searching of patents. This statistical research led to the probabilistic aspects of the predictive syntactic analyzer, which is the principal achievement of the NBS research on MT.

These statistical studies were seen as preliminary investigations which had determined "the types of difficulties likely to be encountered" in a MT system. Ten categories of difficulties had been identified. The NBS team believed they could cope "with only the first five, which depend on syntactic analysis"; the others involved "semantic considerations" which were to be tackled later.

The basic conception of the NBS system of MT was the familiar one: look up SL words in dictionary, get SL grammatical information and TL equivalents, determine SL structural relationships, and produce TL output rearranged in light of this information. The originality and innovatory achievement of Ida Rhodes was the method of SL syntactic analysis. In essence the approach was very simple: taking each word in turn, on the basis of information about its potential grammatical categories (subject, verb, adjective, etc.) predictions are made about the categories of succeeding words and tests are made to see if earlier predictions can be fulfilled. Predictions could be grammatical (e.g. that the noun following an adjective should agree in number, gender, and case), or lexical (e.g. that particular verbs govern nouns in particular cases, genitive, dative, etc.), or physical (e.g. that capital letters indicate a name or abbreviation or new sentence). Some predictions had to be fulfilled immediately, others were optional. The aim was to establish a complete sentence analysis within just a single left to right pass. During analysis the system built up two 'pools': one of predictions, the 'foresight pool', and one of unexpected occurrences to be resolved later, the 'hindsight pool'. Every predictable occurrence was checked against the 'foresight pool' and any fulfilled predictions were erased. Any words occurring which had not been predicted (notably conjunctions, prepositions, adverbs, particles, punctuation marks) were put in a 'hindsight pool' for resolution by later items. At the end of the sentence no unfulfilled predictions should remain in the 'foresight pool' and no problems should remain in the 'hindsight pool'.

The system was first described in detail in a report widely distributed in 1959 (Rhodes 1959, later published as Rhodes 1961); later developments in the 'hindsight' technique were given by Rhodes & Alt (1961) and developments in the treatment of clauses and phrases were given by Alt & Rhodes (1961). From the first demonstration of 'predictive analysis' the NBS approach aroused considerable interest. Bar-Hillel (1960) was impressed by its practicality and efficiency, which he attributed to Rhodes' success in combining her long programming experience and her intuitions as a native Russian speaker; and he was surprised, in view of her evidently minimal knowledge of linguistics, by "how much of the practical aims of MT can be attained with so little use of structural linguistics". His enthusiastic endorsement of the NBS approach was shared by other MT researchers, in particular by the Harvard team which devoted considerable resources to the development of the 'predictive syntactic analyzer' after 1960 (ch.4.9).

By 1960, it was claimed (Rhodes 1961a) that the NBS system was capable of tackling the following problems: "(i) The stem of a source word is not listed in our glossary... (ii) The target sentence requires the insertion of key English words, which are not needed for grammatical completeness of the source sentences. For instance, the complete Russian sentence *on bednyi* (literally "He poor") must be translated as: "He (is) (a) poor (man)". (iii) The source sentence contains well-known idiomatic expressions. (iv) The occurrences of source sentence do not appear in the conventional order", e.g. problems of inversion, "(v) The source sentence contains more than one clause". The remaining unsolved problems were those of homonymy, polysemy, syntactic ambiguity in SL texts, misprints, incompleteness and grammatical errors in SL texts, and uncommon and local usages.

The NBS project was always severely practical and limited in its aims. Ida Rhodes (1966) was convinced that "a faithful translation cannot be achieved even when carried on by the most knowledgeable and competent human being." Computers could cope well with "a formal, systematic mental process", but "Language... is actually a notoriously lawless, arbitrary, capricious, wayward child of the human mind." Only parsing was mechanizable; the problems of semantics seemed insuperable: to achieve good quality translations the system would have to encompass the entire encyclopedic knowledge of the human translator. In her view, although MT research should strive for the utmost possible, the practical limitations should be accepted. Research at NBS continued until about 1963, but little advance was made on the basic design of the 'predictive syntactic analyzer' and the team did not develop a working MT system. Ida Rhodes (1967) saw "no royal road to MT. Only the rocky road of gruelling toil, unremitting attention to minutiae"; and at the end, only crude practical and essentially poor quality products.

4. 9: Harvard University (1954-1964)

Research on MT at Harvard University dates from 1950, when at the instigation of Professor Howard Aiken, one of the recipients of Weaver's memorandum, Anthony G. Oettinger began his doctoral dissertation on a computer-based dictionary. Soon after its presentation in 1954 (the first thesis on MT), and particularly from 1956, the important Harvard MT group began to be formed under Oettinger's leadership.[11] Funding by the National Science Foundation started in 1958 and the first of the massive Harvard research reports, all entitled *Mathematical linguistics and automatic translation*, appeared in January 1959. Research was to continue at Harvard until 1972, although after 1965, when Oettinger was no longer director, the emphasis moved to general computational linguistics and research specifically on MT virtually ceased.

Until 1960 research at Harvard concentrated on the development of a large-scale Russian-English dictionary and on its utilization for word-for-word translations into English. From 1959 the activities of the group turned primarily to problems of syntactic analysis. For this purpose, the Harvard team adopted the 'predictive syntactic analyzer' approach initiated at the National Bureau of Standards, and became the major developers of this method, later known as 'multiple-path analysis'. This change from lexical to syntactic interests marks a significant break in the activities of the Harvard group; the syntactic research will be treated after outlining the important Harvard contribution to dictionary construction.

The Harvard researchers were strong believers in empirical methods and in the value of using the computer as tool for the massive data handling required in MT work. The Harvard dictionary was seen to provide "(a) an immediately useful device for lightening the burden on professional translators, speeding up their work, and improving its accuracy and timeliness; (b) a system of automatic word-by-word translation system; (c) an experimental tool to facilitate the extensive basic research still necessary to develop methods for faithful smooth translation of technical Russian into English." (Oettinger et al. 1959). The first aim was scarcely achieved, but it was not surprising since, as later researchers realised, the requirements of an automatic dictionary for direct consultation by translators are quite distinct from those of a MT dictionary (ch.17.7). As for the second aim, word-for-word translations were indeed produced by the Harvard group, but they can hardly be considered particularly useful, since apart from keeping strictly to Russian word order, no attempts were made to select from the numerous alternative versions offered for many of the Russian words, and no effort was made to supply prepositions corresponding to Russian case endings or to insert articles. There remained the third purpose, an experimental tool, and there is no doubting the valuable contribution of the Harvard team in thoroughly exploring the problems of dictionary compilation.

Descriptions of the Harvard Automatic Dictionary are common (e.g. Oettinger et al. 1959), Giuliano & Oettinger 1960) but the most extended and detailed description is to be found in the nearly 400 pages of Oettinger's book *Automatic language translation* (Oettinger 1960). The first third of the book discussed the nature of computers and their programming (with examples mainly from UNIVAC I, the machine at Harvard) and the problems of applying computers to non-numerical tasks; the need for such a detailed introduction is indicative of the general ignorance of computers and their capabilities at this time among linguists and indeed the public at large. The next section was devoted to the 'problem of translation', first distinguishing it from transcription and transliteration, then discussing the possibility of translation in general and automatic translation in particular. Oettinger advocated a gradualist methodology, starting with good word-by-word approximations, adding some manipulation of word order as required, and dealing with idiomatic expressions. (By 1960, as he readily agreed, the Harvard system had gone no further than the first stage.) The next substantial sections investigated in detail the morphology of Russian (inflection, conjugation, and derivation) and the classification of Russian vocabulary, as a preliminary to discussion of the compilation of a dictionary in which stems and endings were entered separately, in order to reduce storage needs (on punched cards or magnetic tapes). To identify endings it was found necessary to devise a new system of classification of Russian inflections (8 classes of adjectives, 38 of nouns, and 46 of verbs). Compilation was partly manual and partly automatic. For its corpus of lexical items the Harvard group favoured selection from a general dictionary and from

[11] See also A.G. Oettinger: 'Machine translation at Harvard', *Early years in machine translation: memoirs and biographies of pioneers*, ed. W.J.Hutchins (Amsterdam: John Benjamins, 2000), 73-86.

specialized technical dictionaries of mathematics and electronics, followed by up-dating whenever new words were encountered in texts. By 1960 the Harvard dictionary file contained about 22,000 stem entries, representing more than 10,000 distinct Russian words or about 100,000 inflected forms (Giuliano & Oettinger 1960). The final third of Oettinger's book illustrated the operation of the dictionary, the lookup procedures and the production of word-by-word translations. Text words were first reduced to their stems, by the suffix-stripping program used in dictionary compilation. As in most systems of the time, lookup involved prior sorting of the words of a text alphabetically, temporary storage on magnetic tape and then item by item matching against the dictionary tape, and producing an augmented output tape from which the words could then be put back into text order. This final tape could then be printed out in full, for further analyses of dictionary entries, or it could form the basis of a word-for-word translation. The translation of a four page text, roughly 1000 words, required an hour-long computer run (Giuliano & Oettinger 1960); or "about 17 words per minute on a UNIVAC I, and about 25 words per minute on a UNIVAC II" according to Bar-Hillel (1960); but it was hoped that with more powerful machines, e.g. the IBM 709 or the UNIVAC 1105, and with the vast storage capacity of the photoscopic disk (see ch.4.2 above), the speed could be increased by 20 to 40 times (Bar-Hillel 1960). By 1961 about 40,000 running words had been translated (Giuliano 1961). These 'translations' included all the English variants for a given Russian form, without distinction, and so unravelling the sense took considerable effort. For example, one text begins:

> NEW/ MODERN/ NOVEL METHOD/ WAY MEASUREMENT/ METERING/ SOUNDING/ DIMENSION SPEED/ VELOCITY/ RATE/ RATIO LIGHT/ LUMINOSITY SHINE LUMINOUS PRESENTED/ INTRODUCED/ PRODUCED/ IMAGINED AKADEMIK-OM G.S.LANDSBERG

(where slashes indicate alternatives, and in which AKADEMIK was left untranslated). This was interpreted (by the post-editor) as:

> A NEW METHOD OF MEASUREMENT OF SPEED OF LIGHT PRESENTED BY ACADEMICIAN G.S.LANDSBERG.

Oettinger admitted the poor acceptability of these products: they were hardly any use to those who, while having a good knowledge of Russian, did not have the technical background; expert technical translators "seem hampered rather than assisted"; and only technical editors with either no or only rudimentary knowledge of Russian could find much use for them.

Some obvious developments were envisaged at this time in order to improve word-for-word translations (Giuliano & Oettinger 1960), such as providing English interpretations of Russian endings, and the "automatic inflection" of English noun and verb forms (instead of giving just the infinitive form of a verb, for example). More significant, however, were the plans to tackle syntactic problems. One suggestion was the implementation of a Trial Translator, which tested various transformation rules (the results to be evaluated by independent linguists and Slavists). Another was a proposed Formula Finder, which could deduce translation rules on the basis of annotations by analysts of texts and the final post-edited versions; the resulting formula could then be tested by the Trial Translator. The notion was developed in detail by Giuliano (1961): the idea was obviously attractive, but it both overestimated the algorithm-handling power of the computers of the time (and indeed, of many years to come), and grossly underestimated the human fallibilities of post-editors, linguists and analysts.

However, these plans were not pursued at Harvard. For a time they were continued by Giuliano who had moved in 1959 to the Arthur D.Little Corporation and initiated collaborative research on Russian-English MT with the Harvard group (*CRDSD* (Nov 1959)). As mentioned above, in 1960 the decision was made by the Harvard group to investigate the method of syntactic analysis developed at the National Bureau of Standards.

This was the 'predictive syntactic analyzer' (Rhodes 1959). It was believed that the approach "reveal(ed) that syntactic structures have an hitherto unsuspected degree of simplicity, regularity, and universality, and that, up to a certain point, they yield themselves to correspondingly simple and elegant, yet powerful, methods of analysis" (Oettinger & Sherry 1961). It is indeed true that the method was simple in its approach. Its advantages in the view of the Harvard researchers (Oettinger & Sherry 1961) were the use of a computationally simple push-down store, the fact that input was scanned once only and in one direction only, and that the amount of internal storage was relatively limited.

The basic premiss of predictive syntactic analysis was that on the basis of an identified grammatical category (article, adjective, noun, etc.) the following category or sequences of categories could be anticipated with an empirically determinable measure of probability. The system had the following characteristics (Oettinger & Sherry 1961, Sherry 1962): under the general control of a push-down store (i.e. last in first out) a sentence was parsed one word at a time left to right, the action taken for each word being determined by a set of predictions associated with the word class or classes to which the word had been assigned in dictionary lookup. At the beginning of the analysis certain sentence types were predicted in terms of word class sequences. Examination of each word was in two stages: first to test whether any of its word classes 'fulfilled' one of the predictions, starting from the most probable one, then either to modify existing predictions or to add further predictions. For example (Sherry 1962): "a noun assigned the preferred argument of *subject* would cause (1) the *subject* prediction to be wiped from the pool, (2) the *predicate head* prediction to be modified, so that only a predicate that agrees with the subject in person, number and gender can be accepted, and (3) two new predictions, a *compound subject* and a *noun complement*, to be entered at the top of the new pool. The *compound subject* is predicted because the noun was selected as the subject; the *noun complement*, a prediction of a genitive noun phrase, is predicted by every noun..."

Predictions were of different types: some predictions had to be fulfilled immediately (e.g. by the next word), others could be satisfied more than once (e.g. adjectives agreeing in number and case with a noun), and other predictions were obligatory, they had to be fulfilled if the analysis was to be complete otherwise the assigned structure would be ungrammatical. Any word class not selected by the first matching prediction in the 'pool' (i.e. the most probable one) was stored in a 'hindsight' store which could be consulted subsequently if the prediction selected was found to be incorrect. At the end of the sentence, analysis was complete if no obligatory predictions remained unfulfilled and if any component had not been included in the total parsing. The system included, therefore, a certain amount of backtracking, which was found to be necessary in particular for dealing with words which are tied to later occurrences, e.g. adverbs preceding verbs, and for identifying nested constructions, e.g. relative clauses.

However, certain problems emerged very quickly. Firstly there was the danger of exponential growth of predictions. Already in 1961 Sherry recognized that "on average, more than two new predictions are added for each analyzed word", and that therefore the system must impose limitations, it must immediately wipe those which it is known can no longer be fulfilled, and identify those which temporarily cannot be fulfilled.

It was also soon apparent that a single path system was not adequate: only the most probable path through the series of predictions was taken during parsing, but the enforced selection of one path meant that incorrect choices were made and the results were often unsatisfactory. Obviously, analysis had to take account of other paths and so the Harvard group developed the method known, appropriately, as multiple-path predictive syntactic analysis. Developed initially for English analysis (Kuno & Oettinger 1963a, Kuno & Oettinger 1963b, Kuno 1965, Kuno 1966), it was later adopted for Russian analysis on the IBM 7090 (Plath 1967). In the multiple-path technique all possible predictions were pursued. Whenever a word was a homograph, or had more than one grammatical category, or could initiate a number of constructions, each prediction was added to the pool. At the end of analysis, only those paths in which all predictions had been satisfied would be accepted as parsings of the sentence. To improve efficiency, so that analyses would not be pursued which could not possibly succeed, various methods of 'path elimination' were introduced. One approach was the ‘shaper test’: for example, a particular prediction might specify more words than were left in the sentence and so it could be rejected; equally if, for example, a prediction required a comma to be present in the remainder of the sentence then a quick test could establish whether there was one or not. Another test involved counting the number of nested structures predicted; a limit of seven was enforced on empirical grounds (following Yngve’s ‘depth hypothesis’ (cf.ch.4.7 above). More particularly, efforts were made to reduce the repetitive analysis of the same structure in the same way every time a new predictive path was attempted. For example, the phrase *the old man* would be reanalysed as a noun phrase every time a new attempt was made to establish the alternative parsings of the sentence *The police followed the old man in a car.* This feature was acknowledged by the Harvard group (Kuno 1965) as a serious disadvantage when compared with immediate constituent parsers (ch.3.4), but one which, it was hoped, could be reduced substantially by procedures to identify identical paths (Plath 1967). The eventual outcome of these developments was the plan for an Augmented Predictive Analyzer (Kuno 1966), equivalent in power to but less demanding in storage than other implementations of context-free grammars. Later these ideas were to be developed and improved by Woods in his Augmented Transition Network parser (Woods 1970; ch.9.13). These various improvements of the predictive analyzer undoubtedly increased its efficiency, and processing times for analysis decreased dramatically between the 1963 version (Kuno & Oettinger 1963b) and the 1965 version, which Harvard made available as a Russian analyzer to other researchers through the SHARE scheme (*TA Information* 1965). Nevertheless, one major unsatisfactory feature remained, the production of multiple parsings (Kuno & Oettinger 1963b, Kuno 1965). The designers hoped, of course, that multiple parsings would occur if and only if the sentence was genuinely ambiguous, but in practice, many more parsings were frequently produced than could possibly correspond to meaningful differences. For example, a table in Kuno (1965) reveals that as many as 136 analyses were generated for the sentence:

> Gravely concerned with spreading racial violence, President Kennedy used his press conference to issue counsel to both sides in the struggle.

Even the apparently innocuous sentence

> People who apply for marriage licenses wearing shorts or pedal pushers will be denied licenses.

produced 40 analyses (including presumably the humorous vision of ‘marriage licenses’ dressed in ‘shorts’). It is surely obvious that many of the parsings of these sentences could not possibly correspond to any acceptable interpretations. Furthermore, Kuno had to admit that there were other sentences in their corpus of 40 to 50 words in length, which “we could not even try to analyze ... for fear of excessive processing time”. Lastly, and most damaging of all, it was found not infrequently that the analyzer failed to produce the parsing corresponding to what would be accepted as the ‘normal’ interpretation of the sentence. The results for the Russian parser (Plath 1967) were somewhat better than these English examples: some 90 per cent of 73 sentences were provided with “semantically acceptable analyses”, i.e. 10 per cent either received no analyses at all or lacked the ‘correct’ analyses; the average number of analyses was 3.3 per sentence and a table suggests that the production of more than 20 analyses was uncommon even for quite long sentences. Many of the difficulties are, of course, common to all parsers lacking semantic controls, e.g. the problems of prepositional phrases (ch.3.6).

Just as before 1960 the Harvard project concentrated almost exclusively on problems of dictionary compilation, after 1960 most research was devoted to methods of analysis and very little to work on structural transfer and English synthesis. Plath (1967) admitted in discussing his own work on Russian-English MT that there was at Harvard “the absence of working systems for synthesis and structural transfer" so that "an accurate assessment of the potential adequacy of the current output as input to these terminal phases of the translation process is clearly impossible". From the evidence of his article, the versions produced offered little real advance on the earlier word-for-word translations, no attempt being made to select between alternative renderings either of lexical items or of syntactic constructions. The only genuine improvements being in the translation of idiomatic syntactic forms and in the indication of the word most likely to be 'correct'. Examples of the rules used for syntactic transfer were given at an early stage of Harvard activity in syntactic problems by Foust & Walker (1962). The syntactic transfer component was to produce an English sentence by inserting words, rearranging sequences or modifying elements. For example, a literal translation of a Russian 'if' construction as *if to neglect preliminary reduction* would be modified by inserting 'one' and 'the' and by changing the infinitive to a finite form: *if one neglects the preliminary reduction*. As an example of rearrangement, the output *in developed (by) us device* would become *the device developed by us*. The transfer rules were explicitly devised for Russian to English only, and founded on recognisably ad hoc sources: "syntactically analyzed texts, conventional grammar books... and the intuition of the authors". It is clear from Foust & Walker (1962) that this work was considered simply as an initial approximation, covering only the major and most frequent constructions, and it was hoped that additional rules for further texts could be added to the system without difficulty. Although Oettinger (1960:120) wrote of generalization to other languages and of developing a transfer notation of sufficient abstractness to constitute an 'intermediate language', it is clear from Plath (1967) that these developments could not be realised in practice within the Harvard approach.

Although most work at Harvard was concerned with Russian-English translation, a certain amount of interest was paid to other languages. In 1961, Kuno (1962) reported on ideas for Japanese-English MT. The main problems were seen as those of word identification and dictionary lookup, since Japanese script does not indicate word boundaries, and of syntactic analysis, since major functional indicators occur at the end of phrases and clauses. For the former, Kuno proposed 'longest-match' searches and a system of temporary stores; for the latter, he proposed a predictive analyzer operating from the end of the Japanese sentence, i.e. right to left, in order that the postpositional indicators (particles) could be used to predict preceding syntactic structures.

The predictive syntactic approach had been adopted initially with considerable enthusiasm for its practical computational virtues and not for any intrinsic theoretical properties. It was seen as complementary to Chomsky's theory of syntactic synthesis (ch.3.5), as an advance on Bar-Hillel's and Lambek's theories of categorial grammar (ch.3.4), in agreement with the observations of Yngve (ch.4.7), and, because it was not being limited to analysis of Russian, as applicable generally to all languages. Formally, the system was an implementation of a finite state grammar (ch.3.4). Later, as development progressed, the mathematical properties of predictive analyzers became apparent (Plath 1967 gives references), and the formal equivalence of such grammars to context-free grammars, specifically to (non-deterministic) pushdown store transducers, gave theoretical foundations to procedures developed empirically.

Interest in mathematical aspects of language processing increased and in the research pursued during the 1960's the Harvard group moved gradually away from any immediate intention to develop a working MT system. The unexpectedly slow development of a working system (as indicated above) certainly contributed to this change of emphasis. Typical are comments in which the MT goal is seen as only one possible application of the research, e.g. Kuno (1965) envisaged: "the designing of a system for information retrieval, question answering or mechanical translation developed around the analyzer with man-machine interaction incorporated at proper places". Already by 1965, work on Russian analysis had ceased 'temporarily' (*TA Information* 1965), and research concentrated increasingly on computational problems of English grammatical analysis, design of man-machine interactive systems and on theoretical questions of formal and mathematical linguistics. By 1966 Oettinger was no longer director of the project team; from 1964 he had been an influential member of the ALPAC committee whose recommendations in 1966 ended large scale funding of MT in the United States.

4. 10: University of California, Berkeley (1958-1964)

The research project in the Computer Center of the University of California, Berkeley, began officially in October 1958 with a grant from the National Science Foundation, primarily for the development of a Russian-English MT system (Lamb 1961). After a short period under Louis G.Henyey (*CRDSD* 3 (1958)), the director of the project was Sydney M. Lamb. The group adopted initially a fairly ‘empirical’ approach, but soon turned to a fundamentally ‘theoretical’ orientation (Bar-Hillel 1960). Elements of empiricism may possibly be seen in Lamb’s attitude to the question of accuracy. Because 100% correctness was not achievable and in order to avoid having to make guesses, translations would have to present alternative versions. In addition, accuracy demanded translations as close as possible to the original: “Departure from the wording of the input text should be allowed only to the extent necessary to insure readability and intelligibility. It is therefore unnecessary, and maybe even undesirable, that the English translation conform in all respects to the rules of English style” (Lamb 1961). In general, therefore, procedures should avoid syntactic analysis of features which have no bearing on the form of the output translation. Such considerations may have influenced the ‘word-orientation’ of much of the Berkeley research, but otherwise empiricism gave way completely to basic research.

From the beginning, Lamb stressed the need for full linguistic information before writing translation programs.[12] He ridiculed trial-and-error MT methods as “trying to cook tiger stew without having caught the tiger” (Lamb 1961, cf. ch.8.2 below). The Berkeley group concentrated much effort on the development of research tools and maximally efficient routines, e.g. a segmentation routine for splitting Russian words into component parts, a comprehensive coding system for Russian grammar, the development of a large Russian-English dictionary capable of covering a vocabulary of over 300,000 words, a lookup and segmentation routine working at a rate of 7,500 words a minute, a system of graphemic coding permitting direct input in any script (Cyrillic, Greek, Latin), and a system for obtaining relevant linguistic information about Russian scientific texts. An example of the latter was the analysis program for investigating the ‘tactic’ rules required on morphemic and lexemic strata (Lamb 1962), but whether it was implemented in practice is doubtful.

The dictionary system for an IBM 704 was described in detail by Lamb & Jacobsen (1961). Its major features were the incorporation of segmentation routines as part of dictionary lookup (Lamb rejected the Harvard approach of segmenting before lookup as it resulted in too many searches for impossible stems), and the coding of morphemes (stems, prefixes, suffixes, derivational affixes) in a tree structure. For computational efficiency the dictionary structure was stored in two parts: ‘letter tables’ and ‘truncate lists’. The first letter of a word referred directly to its entry in a ‘first-letter’ table, the contents of which were addresses in a ‘second-letter’ table for words beginning with that first letter; the second letter of the word determined which address of the table was to be consulted for access to the ‘third-letter’ table, and so forth. At any point, the search could be terminated if appropriate by the ‘truncate lists’ indicating segmentation cuts. As described, the lookup procedure was estimated to run at 8 milliseconds per word (tested on a 30,000 word text). By 1964, it was claimed that improvements had reduced times by a hundredfold (Lamb 1964)

The basic objective of Lamb’s approach was an “efficient MT system” which could be achieved, in his view, by applying the principles of good linguistic analysis, i.e. on “separating various things from one another on the basis of recurrent partial similarities”. The principle applied most obviously to segmentation of Russian noun and verb paradigms. But it applied also to the

[12] See also S.M.Lamb: ‘Translation and the structure of language’, *Early years in machine translation: memoirs and biographies of pioneers*, ed. W.J.Hutchins (Amsterdam: John Benjamins, 2000), 177-195.

design of MT systems, in the separation of algorithmic procedures (computer programs) from the linguistic data (cf. ch.9.9 below). For Lamb, it implied also separation of translation processes into stages in accordance with a stratificational view of linguistic systems.

Much of Lamb's work on the Berkeley MT project was devoted to the elaboration of his linguistic theory of 'stratificational grammar' (Lamb 1966, Lockwood 1972), which in certain respects may be considered a rival to Chomsky's transformational generative grammar. As a MT model the most detailed descriptions are to be found in papers by Lamb dating from 1962 (Lamb 1965) and 1964 (Lamb 1973).

Lamb's model of language posits a series of 'strata' (or levels) within which and between which linguistic units are related: phonemic (or graphemic), morphemic, lexemic and sememic. (The number and constituency of strata vary from version to version of the theory.) Elements of a stratum are combined according to the 'tactic' rules of the stratum, e.g. on the phonemic stratum of English, tactic rules permit the combination of /s/ and /k/, as in *skill* and *risk*, but not of /s/ and /d/, and on the morphemic stratum, tactic rules determine the applicability of derivational prefixes and suffixes. Tactic rules on the lexemic stratum are comparable to phrase structure rules and those on the sememic stratum are concerned with such semantic relationships as agent and action, modifier and modified. Realisational rules relate elements of different strata, e.g. that the morpheme /good/ is realised by the phoneme sequence /g u d/, and that the lexeme /understand/ is realised by the morphemes /under/ and /stand/. Realisations are not one-to-one: a lexeme such as /good/ is realised as /bet/ when followed by the comparative lexeme /er/, and /stand/ can be realised as a morpheme /stood/ if the lexeme /PAST/ is present. Above the lexemic stratum is the sememic, which deals with semantic components of lexemes (e.g. lexeme /ram/ as a realisation of the conjunction of sememes /sheep/ and /male/), with synonymy (e.g. lexemes /start/, /begin/, /commence/ as possible realisations of a single sememe /initiate/), with homonymy (e.g. lexeme /light/ as realisation of either /pale/ or /un heavy/) and with idioms (e.g. the lexeme combination /pass/ and /away/ as one realisation of /die/).

As a MT model, the stratificational approach envisages translation as a system of decoding and encoding, which proceeds from the graphemic stratum of the source language through its morphemic and lexemic strata to the sememic stratum. Here the links are made to the sememic stratum of the target language, and the process continues then from the sememic stratum through the lexemic and morphemic strata to the final graphemic output of the target language text. The basic linguistic units taken through the processes are the lexemes. Lamb characterised his own approach as word *by* word translation, examining each word within the broadest environment, not being limited by sentence boundaries or immediate contexts. The central objective was lexemic, or if possible sememic (semantic), transfer into target language texts, not sentence-for-sentence transfer. The emphasis of the Berkeley group's research was, therefore, on lexical and semantic aspects of translation and not on the development of syntactic analysis. (In crude terms, the approach was basically 'word-centred', but more sophisticated than 'lexicographic' approaches, ch.4.1-2 above)

Lamb acknowledged that there were certain practical problems in implementing the stratificational approach in MT systems. One involved the resolution of lexical ambiguities. When a particular morpheme can be the realisation of more than one lexeme, only consultation of the conditions attached to the realisation of each candidate lexeme can determine which is the correct analysis. In other words, the decoding of ambiguous morphemes (and lexemes) demands first the identification of possible realizates and then the testing of each by the tactics to determine which of them would encode in the form actually present. In effect it is a kind of 'analysis by synthesis' (cf.ch.4.7 above), except that it operates in a more restricted manner. Such a tactic decoding procedure was implemented in the Berkeley MT analysis program (Lamb 1964).

Apart from research on dictionary lookup, the contribution of the Berkeley group was, therefore, primarily of a theoretical nature. It was also mainly on Russian-English MT, but there was a cooperative project with the group in Mexico (ch.7.3) on a Russian-Spanish dictionary and, particularly towards the end of Lamb's leadership of the project (Lamb moved in 1964 to the Linguistic Automation Project at Yale University), there was a substantial amount of research on Chinese-English MT. A good deal of work was done on Chinese lexicography (Booth 1963, Lamb (1964); and research continued on Chinese grammar under Chang Yi Dougherty until the beginning of the POLA project (ch.11.2).

4. 11. University of Texas, Linguistics Research Center (1958-68)

Research on MT began at the University of Texas after a seminar in October 1958 conducted by Professor Winfred P. Lehmann and S.N. Werbow on problems of German and English syntax. Invited to the seminar was a group from Georgetown University which included Leon Dostert, A.F.R. Brown and Peter Toma (Werbow 1961; Toma 1984). After the seminar, informal studies on MT were continued by members of the Department of Germanic Languages. In May 1959 a contract was agreed with the U.S.Army (Signal Research and Development Laboratory) for research on a German-English MT system. Subsequently, in 1961, a grant was received from the National Science Foundation for the development of computer-based system for linguistic analysis. At this time the research group was established as the Linguistics Research Center (LRC) under the direction of Winfred Lehmann, with Werner Winter and Stanley Werbow heading research on linguistic analysis and Eugene Pendergraft the development of programming (Werbow 1961, Tosh 1965). In later years, most funding came from the Rome Air Development Center of the US Air Force.

From the beginning, the Texas group was concerned with more than research on German and English MT. As Werbow remarked (1961) interest in MT at Texas had grown from theoretical linguistic studies, "research in German syntax, historical work with Early New High German, with analysis of modern written German, and also with the study of contemporary spoken German" and "the syntax of English". It was natural, therefore, that research at LRC was closely involved in research on formal linguistic theories. Particular interest was paid to the developments of transformational grammar, both in theory and in MT applications (e.g. Yngve and others at MIT, ch.4.7 above), and to developments on the Discourse Analysis project at the University of Pennsylvania under Zellig S. Harris (ch.3.5). The LRC team took the long-term perspective on MT research; before attempting to develop translation programs they intended systematic detailed comparative studies of German and English syntax.

The achievements of the project up to 1968 are summarised by Lehmann (1965) and by Tosh (1968, 1969, 1970) and in greater detail by Tosh (1965)[13]. Translation was held to take place at three levels of increasing complexity: lexical, syntactic and semantic; to which corresponded three MT models: word-for-word translation, syntactic translation (incorporating lexical transfer and syntactic analysis and transfer), and the ultimate aim: transformational-semantic (incorporating semantic analysis and transfer). The LRC model was designed for 'semantic' translation. What had been achieved by 1965 was characterised as 'syntactic transfer', and was obviously modelled on Yngve's approach at MIT (ch.4.7). In conformity to the true 'transfer' model (ch.3.9), it had three clearly defined stages: SL analysis, transfer, and TL synthesis. The research strategy adopted was to establish phrase-structure analyses of a set of German sentences and separate independent phrase-structure analyses of their English equivalent sentences, and then to devise transfer rules to convert the structures into each other, both from German (SL) to English (TL) and from English (SL) to German (TL). Since the aim was a bidirectional MT system the syntactic transfer rules had to be reversible. For lexical transfer, it was proposed (Tosh 1965) that numerical codes would indicate equivalences of German and English lexical senses, e.g. 1009 for German *edel* and English

[13] See also W.P.Lehmann: 'My early years in machine translation', *Early years in machine translation: memoirs and biographies of pioneers*, ed. W.J.Hutchins (Amsterdam: John Benjamins, 2000), 147-164.

noble, 1008 for *gross* and *noble*, 1005 for *gross* and *large*, etc. The intention was to establish classes of bilingual sense equivalences, which would correspond in some respect to the classes of syntactic equivalences. There was some hope that these classes could be the basis for multilingual equivalences. Somewhat later, Tosh (1969) suggested the use of Roget's thesaurus to establish interlingual codes (presumably on the lines of the research at Cambridge, ch.5.2 below), but these ideas were not followed up.

As an example, Tosh (1965, 1968) describes in detail the rules of analysis, transfer and synthesis for the German sentence *Was Bewußsein ist, kann man nicht näher umschreiben* and its English equivalent *What consciousness is none cannot further circumscribe*. The analysis for the German sentence was (simplified):

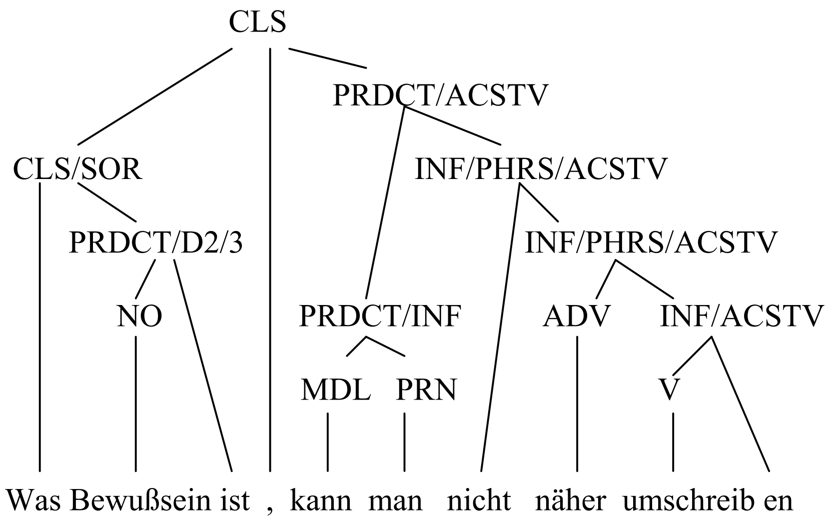

Fig.9: LRC analysis (German)

For the English sentence the analysis was (also simplified):

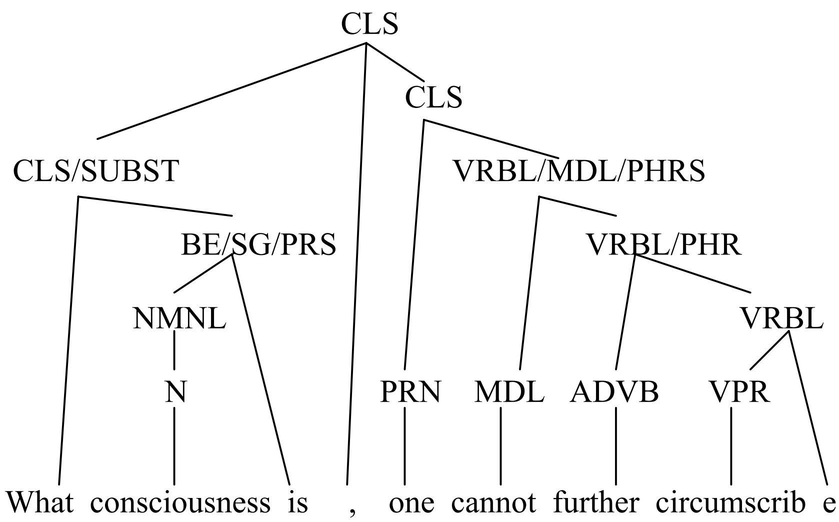

Fig.10: LRC analysis (English)

An example of a transfer rule was one converting the syntactic structure *kann (man) nicht (umschreiben*) to *(one) cannot (circumscribe*) and vice versa. It was formulated as the bidirectional rule:

kann+PRN+nicht+INF/PHRS/ACSTV ↔ PRN+cannot+VRBL/PHRS

where, of course, each non-terminal structure (capitals) is itself a phrase structure rule.

The first tentative testings took place in 1965. A small corpus of sentences had been selected, for which German-English lexical and syntactic transfer rules had been written. There were for German: 2,600 syntax rules, 41,800 rules for lexical items (i.e. dictionary items) and 6,200 rules for transfer; for English there were 800 syntax rules, 76,200 for lexical items and 3,000 for transfer. After an initial run which revealed numerous omissions from the dictionary and programming errors, the results were considered satisfactory (Lehmann 1965, Tosh 1968):

THE CONDITION OF THE READER'S BRAIN AT THIS MOMENT WHEN IT HAS DECIDED TO CONSIDER WITH THE AUTHOR SUCH A COMPLICATED SUBJECT IS THAT OF WAKEFUL ATTENTIVENESS.

The provisional nature of the results was clearly recognised (Tosh 1968). Only those rules necessary for the analysis and synthesis of the sentences of the corpus had been programmed; and, furthermore, only one possible syntactic output had been considered for each sentence (that which the human translator had provided). Despite the limitations of the linguistic data, there were "already sufficient implicit relationships to permit unplanned for... paraphrases." For the future, there was the problem of pronominal reference, and above all the development of a "more sophisticated model... in which there is a grammar of structural semantics."

These plans were not to be realised. After 1965 development of the German and English system continued at a somewhat lower level. On the other hand there were some small scale pilot projects to test the generality of the model on a number of other languages. Tosh (1969) mentions the compilation of lexicographic data for French, Spanish, Hebrew, Japanese and Chinese. There was also beginning some rather more extensive research on Russian. Around 1965, the Russian Master Dictionary, which had been developed through US Air Force support at the University of Washington (ch.4.1), had been incorporated into the LRC research program (Josselson 1971). By 1967 the LRC database contained some 140,000 dictionary items for Russian, on which the algorithms developed for the German and English system had also been tested, although perhaps only to a limited extent (Tosh 1970).

It was around this time, however, that LRC research plans were interrupted. MT research at Texas came to a temporary hiatus from about 1968 until 1970 when the German-English project was revived on a rather more ambitious level (ch.10.3)

4. 12: Wayne State University (1958-1972)

The Russian-English MT project at Wayne State University, Detroit, was set up in July 1958 under the direction of Harry H. Josselson of the Department of Slavic Languages and Arvid W. Jacobson of the Department of Mathematics. Its sponsor was the Office of Naval Research. The project took care not to duplicate the research of other MT groups, which was only wise in view of the number engaged on Russian-English translation at the time. It sought, therefore, both an individual methodology and areas of investigation not covered in depth by others (Josselson 1961).

It was decided to limit the subject field to mathematics, mainly in the area of differential equations, both to concentrate the vocabulary problems and to reduce structural diversity. Initially three Russian texts with extant English translations were selected in order to work with parallel texts in compiling dictionaries and programs. Eventually, the project's dictionary was based on fifteen mathematics articles and the Russian-English glossary of the American Mathematical Society (Wayne State Univ. 1968; Josselson 1972)

The emphasis was on "careful linguistic analysis of the prior to any effort to program", with the aim of "developing practical translation procedures yielding fluent and accurate text". As for overall methodology, the team adopted the 'cyclic' approach: prepare dictionary and programs, test on a text, update dictionary, revise program, run a new text, etc. (Josselson 1961)

For their specific area of MT study, the Wayne State team chose to concentrate on problems of ambiguity, both at the lexical and the morphological level, and on problems of syntactic structure analysis. The study of ambiguity reported by Janiotis & Josselson (1962) indicates a fairly standard approach: six types of ambiguity were identified, three could be resolved by structural analysis (homographs of different word classes, inflectional ambiguity, compound verb phrases), the others by contextual clues (idioms, homographs in same word class, polysemes); only with genuine polysemes could no obvious solution be seen.

More original was the Wayne State approach to grammatical analysis. The grammatical codings assigned to Russian lexical items was based not on traditional word classes (as in many MT systems) but on a reclassification in terms of syntactic functions and distribution. Nine classes were identified: nominal, predicative, modifier, infinitive, gerund, adverb-particle, preposition, conjunction, declined relative. The classes cut across traditional divisions; thus, 'modifiers' included adjectives, participials, numerals, demonstrative pronouns and possessive pronouns, while 'predicatives' included short-form adjectives, verbs and comparative adverbs (Josselson 1961)

Syntactic analysis was specifically "language oriented"; no general model was developed "although we used the fulcrum concept of Paul Garvin in the design of our routines". (Garvin had been consultant to the project while still at Georgetown and remained so after his move to Ramo-Wooldridge.) The first set of routines grouped immediate constituents into phrases consisting of a core word (Garvin's 'fulcrum') and its dependents. These 'blocking' routines formed noun phrases, prepositional phrases, verb phrases including adverbs, etc. The next stage, 'profiling', classified each sentence constituent according to potential functions, e.g. subject, predicate, adjunct. The third procedure, 'parsing', sought to identify the specific syntactic role of each 'block' in relation to the finite verb (the 'predicative') as fulcrum. The result, in many cases, was multiple analyses of sentences. In addition, after 14 years of development, Josselson (1972) had to admit grave weaknesses. "The principal weakness of the parsing routine is that it cannot parse sentences with more than one predicative" because routines did not "mark the boundaries of clauses and hence were unable to ascertain which predicative was in the main clause and which were in subordinate clauses, or whether the sentence contained more than one main clause." Whether such fundamental failings are attributable to the fulcrum approach as such must remain very much an open question, but they are unfortunately severe indictments of the limitations of the Wayne State methodology.

The project team saw the purpose of its research increasingly as contributing to Russian linguistics rather than to MT as such. A change in the project's title from 'machine translation' to 'computer-aided translation' was intended to reflect a shifted aim, from a "fully automatic high quality translation" system (cf. ch.8) to "an experimental system of computer oriented strategies for analyzing Russian sentences... We now envision a system of performing translation by means of man-machine symbiosis." (Josselson 1972). It was probably this expansion of horizons which ensured the continuation of the project until 1972, while other US projects of the same vintage did not survive the ALPAC report (ch.9.1).

4. 13: Other United States groups (1955-1965)

The early MT interest in California (ch.2.4) continued for some years with a number of generally short lived projects. Three Californian aviation and aerospace companies set up Russian-English projects. Norair, a division of Northrop Corporation, had a project under Ron Manly, which involved the "construction of an 'intelligence language'", i.e. an interlingua (*CRDSD* 7, Nov 1960). Autonetics, a division of North American Aviation, aimed for a system on their own RECOMP II computer which might produce "rough, understandable, but not publishable" translations. A "microglossary of approx. 100 scientific terms" was compiled; trials on 7 texts resulted in 69% of words correctly translated, 1.5% incorrect, 14% of the untranslated words made sense, but another 14% were "gibberish" (*CRDSD* 9, Nov 1961; 10, May 1962). At Palo Alto, the Lockheed Missiles

& Space Company, began a project in July 1960 led by B.D.Rudin on an 'empirical' research project towards a reciprocal Russian-English MT system (*CRDSD* 11, Nov 1962). Research continued until 1964 or later on categorial grammars, statistical analyses and structural linguistics (*CRDSD* 13, Nov 1964)

Victor Oswald, one of the MT pioneers (ch.2.4 and 2.5) continued his involvement in a project reported from the Planning Research Corporation in Los Angeles, on "the potential technical and linguistic system requirements for machine translation" including the "analysis of the probable requirements for accuracy and volume of both the commercial and governmental users of translation machines." This group included also Harold P. Edmundson and William Mitchell (*CRDSD* 5, Nov 1959)

The objective of the project set up in 1961 at the Summer Institute of Linguistics, Santa Ana, under Joseph E. Grimes was expressed as follows: "using SIL Concordance Program for the IBM 650 computer as a research tool, an exploration is being made of the possibilities of doing mechanical translation by comparing lexical strings in parallel texts, i.e. an original text and a translated version of it". The initial work was done on English and Comanche texts; later some cooperative studies were undertaken on Mexican languages with the group at the University of Mexico (ch.7.3) MT activities were suspended in 1962 (*CRDSD* 11, Nov 1962)

Peter Toma did MT research in Los Angeles both before and after going to the Georgetown project (ch.4.3).[14] In 1957 he did some research on Russian texts at the California Institute of Technology using a Datatron computer (*CRDSD* 1, Nov 1957; Bar-Hillel 1960); and he was also involved with the research at the International Telemeter Company for the USAF system (ch.4.2). After leaving the Georgetown project in 1962, he set up his own company Computer Concepts Inc. in Los Angeles (*CRDSD* 10, May 1962). The AUTOTRAN Russian-English system was claimed to be "fast efficient and accurate" A dictionary of 100,000 stem entries was compiled for the fields of atomic energy and medicine. The system, programmed for the IBM 7090, was said to be operational in July 1963 producing translations of "excellent quality" and a "very high degree of accuracy" (*CRDSD* 13, Nov 1964). Evidently the system was an early prototype of Systran (ch.12.1)

Another spin-off from Georgetown was the Russian-English system developed by Ariadne Lukjanow based on her code-matching technique (ch.4.3). A company named Machine Translation Inc. was set up in 1961 in Washington, D.C. to finance a "production system". It was intended also to market a Russian-German version. It was announced that the "complete production system is in the final stages of checkout and will be completed in the spring of 1962" (*CRDSD* 10, May 1962), but not long afterwards in November 1962 it was reported by the editor of *CRDSD* that "work on production system... has been discontinued" (*CRDSD* 11, Nov 1962).

In connection with its MT project, IBM sponsored MT research at a number of centres during the mid 1960's (ch.4.2). Fang Yu Wang began in 1960 at Yale University a "study of grammatical signals in Chinese-English machine translation" (*CRDSD* 11, Nov 1962). Asher Opler at The Computer Usage Company in New York was sponsored by IBM between 1962 and 1964 to develop a French-English system, based on the ideas of Sakai (ch.7.1 below), to be implemented on the IBM-USAF machine (*CRDSD* 11, Nov 1962; *CRDSD* 13, Nov 1964). And at Tufts University, Medford, Mass., Mason N. Crook began an evaluation study of the Russian and Chinese IBM systems (*CRDSD* 13, Nov 1964.)

The most important of the 'minor' Chinese-English MT projects was the one which began at Ohio State University in July 1961. This was a research project set up with a National Science Foundation grant under the leadership of William S-Y Wang (*CRDSD* 9, Nov 1961). The group was concerned principally with linguistic investigations on the syntax of Mandarin Chinese. Its theoretical framework was that of transformational grammar (*CRDSD* 10, May 1962). A focus of

---

[14] See also P.Toma: 'From SERNA to Systran', *Early years in machine translation: memoirs and biographies of pioneers*, ed. W.J.Hutchins (Amsterdam: John Benjamins, 2000), 135-145.

the research was a comparative study of the verb constructions of Chinese and English. From 1962 the research concentrated on Chinese phonology. One of the researchers of the team was Charles Fillmore, whose case grammar was to influence many MT projects (ch.9.16). In the mid 1960's Wang joined the Berkeley group and began research on the Chinese-English POLA project (ch.11.2)

In the early 1960's it was common for numerous projects to be described as being relevant to MT in some way. Many projects involving optical character readers and speech recognition were treated as MT projects, e.g. the OCR project at Baird Atomic (cf.ch.8.10) and the speech analysis projects at the Bell Telephone Company: "a limited vocabulary automatic word recognizer might be adapted to a 'word for word' translation between two languages" (*MT* 1(2), August 1954).

Rather more directly relevant, but nevertheless not now classifiable as MT projects, were a number of investigations of basic linguistic methods. The most well known at the time was the project under Zellig Harris at the University of Pennsylvania (ch.3.5), a group which had never had MT as such as its objective but whose method of 'string analysis' influenced a number of MT projects and remains an important technique to this day (cf. Sager 1981). Other projects during the period included one at the Georgia Institute of Technology, under B.J. Dasher, which was described as devoted to "linguistic problems of MT", specifically German and Spanish (*CRDSD* 10, May 1962, *CRDSD* 13, Nov 1964), and one at Indiana University under Fred W. Householder and John Lyons studying the "automation of general semantics" and "artificial regular languages... suitable for... MT..."(*CRDSD* 7, Nov 1960).