W. John Hutchins

Machine Translation over Fifty Years

Abstract

The history of machine translation is described from its beginnings in the 1940s to the present day. In the earliest years, efforts were concentrated either on developing immediately useful systems, however crude in their translation quality, or on fundamental research for high quality translation systems. After the ALPAC report in 1966, which virtually ended MT research in the US for more than a decade, research focussed on the development of systems requiring human assistance for producing translations of technical documentation, on translation tools for direct use by translators themselves, and, in recent years, on systems for translating email, Web pages and other Internet documentation, where poor quality is acceptable in the interest of rapid results.

Keywords: **machine translation, translation tools, history, research, commercial systems**

1. Introduction

This paper traces the history of efforts to develop computer programs (software) for the translation of natural languages, commonly and traditionally called ‘machine translation’ (MT), or, in non-English-speaking countries, ‘automatic translation’ (*traduction automatique*, *avtomaticheskij perevod*). Translators (particularly translators of literature) have generally regarded the task as misguided if not impossible. From the beginning, however, MT researchers have concentrated almost exclusively on the translation of scientific and technical documents, where the difficulties of cultural differences and variable contexts are less acute, but where (in particular) the demand for translation has almost always exceeded the capacity of the translation profession. What is often needed is not a ‘perfectly’ accurate rendition but something that can be produced quickly (sometimes immediately) conveying the essence of the original text, however grammatically imperfect, lexically awkward and stylistically crude. For other purposes, e.g. the production of multilingual documentation for large companies,

the outputs of MT systems serve as time-saving and cost-saving draft translations which are then edited for publication.

This brief history can of course mention only the most significant research systems and projects and only the most important operating and commercial systems (and none in any detail). It concentrates on the research and development of MT systems; it does not attempt to discuss all the various theoretical discussions and disputes among researchers; and, in particular, it does not cover – except in passing – the potential relevance of MT research to, and its impact on, fields such as: computer programming, theoretical and formal linguistics, natural language processing in general, computational semantics, artificial intelligence, automatic indexing and summarization, message understanding, information extraction, controlled languages, speech recognition and synthesis, translation studies, cross-language information retrieval, etc. Neither does it cover the worldwide impact of MT systems and computer-based translation tools on: the translation profession, multi-national corporations, the software localization industry, Internet service providers, etc. For such information and much more about the systems mentioned and the issues raised – and for substantial bibliographies of the voluminous MT literature – readers are referred to the selection of histories, surveys and collections listed at the end of this paper.

2. Precursors and pioneers, 1933-1956

Although we may trace the origins of machine translation (MT) back to seventeenth century ideas of universal and philosophical languages, and of ‘mechanical’ dictionaries, it was not until the twentieth century that the first practical suggestions could be made – in 1933 with two patents issued in France and Russia to Georges Artsrouni and Petr Trojanskij respectively. Artsrouni’s patent was for a general-purpose machine which could also function as a mechanical multilingual dictionary. Trojanskij’s patent, also basically for a mechanical dictionary, went further with proposals for coding and interpreting grammatical functions using ‘universal’ (Esperanto-based) symbols (Hutchins and Lovtsky, forthcoming).

Neither of these precursors was known to Andrew Booth and Warren Weaver when they met in 1946 and 1947 and put forward the first tentative ideas for using the newly invented computers for translating natural languages. In 1948 Booth worked with Richard H. Richens on morphological analysis for a mechanical dictionary. By this time, the idea of mechanical translation (as it was known almost invariably in the period up to the early 1960s) had occurred independently to a number of people, and in July 1949 Warren Weaver put forward specific proposals for tackling the obvious problems of ambiguity (or 'multiple meanings'), based on his knowledge of cryptography, statistics, information theory, logic and language universals. This memorandum was the stimulus for MT research in the United States (Hutchins 1997).

Then, in May 1951 Yehoshua Bar-Hillel was appointed to do research at the Massachusetts Institute of Technology (MIT). After visiting all those interested in the subject he wrote a state-of-the-art report, in which he outlined some of the basic approaches to MT questions; and in June 1952, he convened the first MT conference (at MIT), which was attended by nearly everyone already active in the field. It was already clear that full automation of good quality translation was a virtual impossibility, and that human intervention either before or after computer processes (known from the beginning as pre- and post-editing respectively) would be essential; some hoped this would be only an interim measure, but most expected that it would always be needed. At the conference, various ideas were put forward for pre-editing and post-editing, for micro-glossaries as means of reducing ambiguity problems (selecting appropriate target lexical items), and for some kind of syntactic structure analysis. Various suggestions for future activity were proposed; in particular, Leon Dostert from Georgetown University, who had come as a sceptic, argued that what was required was a demonstration of the feasibility of MT in order to attract research funding.

Accordingly, he collaborated with IBM on a project which resulted in the first public demonstration of a MT system on 7th January 1954. It was the joint effort of Peter Sheridan of IBM and Paul Garvin at Georgetown. A carefully selected sample of 49 Russian sentences was translated into English, using a very restricted vocabulary of 250 words and just 6 grammar rules. Although it had little scientific

value, it was sufficiently impressive to stimulate the large-scale funding of MT research in the USA and to inspire the initiation of MT projects elsewhere in the world, notably in the USSR.

In the same year, the first journal was founded by William Locke and by Victor Yngve, who had succeeded Bar-Hillel at MIT in 1953 – the journal "Mechanical Translation" was to carry some of the most significant papers until its eventual demise in 1970 – and also in this year, the first doctoral thesis in MT, Anthony G. Oettinger's study for a Russian mechanical dictionary. The years 1954 and 1955 saw the foundation of the Cambridge group under Margaret Masterman, the Milan group under Silvio Ceccato, the first Russian groups at the Institute of Precise Mechanics and Computer Technology, the Institute of Applied Mathematics, Leningrad University, etc. and the start of various Chinese and Japanese projects. And in 1955 the first MT book appeared, a collection edited by Locke and Booth, including Weaver's 1949 memorandum, Booth and Richens' experiments, some papers given at the 1952 conference, and other contributions from Bar-Hillel, Dostert, Oettinger, Reifler, and Yngve.

3. The decade of high expectation and disillusion, 1956-1966

When MT research began, there was little help to be had from current linguistics. As a consequence, in the 1950s and 1960s, the research methods tended to polarize between, on the one hand, empirical trial-and-error approaches, which often adopted statistical methods for the 'discovery' of grammatical and lexical regularities which could be applied computationally, and, on the other hand, theoretical approaches involving fundamental linguistic research. The contrastive methods were usually described at the time as 'brute-force' and 'perfectionist' respectively; the aim of the former being 'useful' if crude quality systems operating in the near future, and that of the latter being eventually systems producing output requiring little or no editing.

In most cases, the 'empiricists' adopted the 'direct translation' approach, i.e. the development of programming rules to translate specifically from one source language (SL) into a particular target language (TL) with the minimal amount of analysis and syntactic reorganisation. Many researchers sought to simplify

bilingual dictionaries by providing single equivalents for SL words that would 'cover' most senses, i.e. would not require analysis of contexts, and that would maintain the word order of the SL original as much as possible.

Some groups taking the more long-term view investigated the possibilities of interlingual representations (codes or symbols independent of both SL and TL), where translation would then be in two stages, from SL to interlingua and from interlingua to TL. Most of them undertook basic linguistic research, but some believed that only fundamental research on human thought process (what would later be called 'artificial intelligence' or 'cognitive science') would 'solve' the problems of automatic translation. Less ambitious that the interlingua approach were the beginnings at this time of what later became known as the 'transfer approach', where conversion was through a transfer stage from abstract (i.e. disambiguated) representations of SL texts to equivalent TL representations; in this case, translation comprised three stages: analysis, transfer, and generation (or synthesis).

Any evaluation of the period must remember that computer facilities were frequently inadequate; much effort was devoted to improving basic hardware (paper tapes, magnetic media, access speeds, etc.) and to devising programming tools suitable for language processing – in particular, COMIT developed at MIT by Yngve's team. Some groups were inevitably forced to concentrate on theoretical issues, particularly in Europe and the Soviet Union. For political and military reasons, most US research was for Russian-English translation, and most Soviet research on English-Russian systems, but the multilingual policy of the Soviet Union inspired research there on a much wider range of languages than elsewhere.

The research under Erwin Reifler at the University of Washington (Seattle) epitomized the dictionary-based 'direct' approach; it involved the construction of large bilingual dictionaries where lexicographic information was used not only for selecting lexical equivalents but also for solving grammatical problems without the use of syntactic analysis. Entries gave English translations with rules for local reordering of output. The huge lexicon made extensive use of English 'cover terms' for Russian polysemes, the inclusion of phrases and clauses and the

classification of vocabulary into sublanguages. After initial work on German and English, the group was engaged on the foundations of a Russian-English system for the 'photoscopic store', a large memory device. From 1958 practical development was directed by Gilbert King at the IBM Corporation (Yorktown Heights, New York), and a system installed for the US Air Force produced (crude but apparently usable) translations for many years, until replaced in 1970 by Systran.

Many researchers at the time distrusted linguistic theory – the formal linguistics of Zellig Harris and Noam Chomsky had scarcely begun – and preferred to develop methods based on the analysis of language corpora. For example, researchers at the RAND Corporation undertook statistical analyses of a large corpus of Russian physics texts, to extract bilingual glossaries and grammatical information. On this basis, a computer program was written for a rough translation; the result was studied by post-editors; the glossaries and the rules were revised; the corpus was translated again; and so it continued in cycles of translation and post-editing. The main method of analysis was initially statistical distribution, although it was at RAND that David Hays later developed the first parser based on dependency grammar.

The research under Leon Dostert at Georgetown University had a more eclectic approach, undertaking empirical analyses of texts only when traditional grammatical information was inadequate. Initially there were several groups at Georgetown, for many years the largest in the USA. One group was led by Paul Garvin, who later left to found his own group at the Bunker-Ramo Corporation, developing his 'fulcrum' method, essentially a dependency parser. Antony Brown was a one-man 'group' experimenting with a pure example of the cyclical method on a French-English system. The method eventually adopted was developed by a group under Michael Zarechnak. This, the Georgetown Automatic Translation (GAT) system, had three levels of analysis: morphological (including identification of idioms), syntagmatic (agreement of nouns and adjectives, government of verbs, modification of adjectives, etc.), and syntactic (subjects and predicates, clause relationships, etc.) GAT was initially implemented on the SERNA system, largely the work of Peter Toma, and then with the programming

method developed by Brown. In this form it was successfully installed by Euratom in Ispra (Italy) in 1963 and by the US Atomic Energy Commission in 1964.

Anthony Oettinger at Harvard University believed in a gradualist approach. From 1954 to 1960 his group concentrated on the compilation of a massive Russian-English dictionary, to serve as an aid for translators (a forerunner of the now common computer-based dictionary aids), to produce crude word-for-word translations for scientists familiar with the subject, and as the basis for more advanced experimental work. From 1959 research focused on the 'predictive syntactic analyzer' – originally developed at the National Bureau of Standards under Ida Rhodes – a system for the identification of permissible sequences of grammatical categories (nouns, verbs, adjectives, etc.) and the probabilistic prediction of following categories. However, the results were often unsatisfactory, caused primarily by the enforced selection at every stage of the 'most probable' prediction. (Nevertheless, an improved version, the Multiple-path Predictive Analyzer, led later to William Woods' familiar Augmented Transition Network parser.)

The research at MIT was directed by Victor Yngve from 1953 until its end in 1965. Here syntax was placed at the centre: a SL grammar analyzed input sentences as phrase structure representations, a 'structure transfer routine' converted them into equivalent TL phrase structures, and the TL grammar rules produced output text. But in the end, a 'semantic barrier' to adequate output was reached. It may be noted that despite Chomsky's association with the group for a short time, transformational grammar had little influence – indeed there is virtually no evidence of Chomskyan approaches in any MT research at this time.

The Linguistic Research Center (LRC) at the University of Texas, founded by Winfried Lehmann in 1958, concentrated also on basic syntactic research. Efforts were made to devise reversible grammars to achieve bi-directional translation within an essentially 'syntactic transfer' approach, laying foundations for the later development of the METAL system.

At the University of California, Berkeley, the project under the direction of Sydney Lamb stressed the importance of developing maximally efficient

dictionary routines and a linguistic theory appropriate for MT. This was Lamb's stratificational grammar, with networks, nodes and relations paralleling the architecture of computers. Translation was seen as a series of decoding and encoding processes, via a series of strata (graphemic, morphemic, lexemic, sememic).

There were no American groups taking the interlingua approach; they tended to adopt less speculative approaches despite Weaver's earlier advocacy. Interlinguas were the focus of projects elsewhere. At the Cambridge Language Research Unit, Margaret Masterman and her colleagues adopted two basic lines of research: the development of a prototype interlingua producing crude 'pidgin' (essentially word-for-word) translations, and the development of tools for improving and refining MT output, primarily by means of the rich semantic networks of a thesaurus (utilizing mathematical lattice theory as a basis.) At Milan, Silvio Ceccato concentrated on the development of an interlingua based on 'cognitive' processes, specifically on the conceptual analysis of words (species, genus, activity type, physical properties, etc.) and their possible correlations with other words in texts.

In the Soviet Union research was as vigorous as in the United States and showed a similar mix of empirical and basic theoretical approaches. At the Institute of Precision Mechanics research under D.Y. Panov on English-Russian translation was on lines similar to that at Georgetown, but with less practical success. More basic research was undertaken at the Steklov Mathematical Institute by Ljapunov, Kulagina and Mel'chuk (of the Institute of Linguistics) – the latter working on an interlingua approach that led to his 'meaning-text' model. This combined a stratificational dependency approach (six strata: phonetic, phonemic, morphemic, surface syntactic, deep syntactic, semantic) with a strong emphasis on lexicographic aspects of an interlingua. Fifty universal 'lexical functions' were identified at the deep syntactic stratum covering paradigmatic relations (e.g. synonyms, antonyms, verbs and their corresponding agentive nouns, etc.) and a great variety of syntagmatic relations (e.g. inceptive verbs associated with given nouns, *conference: open, war: break out*; idiomatic causatives, *compile: dictionary, lay: foundations*, etc.)

Interlingua investigations were consonant with the multilingual needs of the Soviet Union and undertaken at a number of other centres (Archaimbault and Léon 1997). One of them was at Leningrad State University, where a team under Nikolaj Andreev conceived an interlingua not as an abstract intermediary representation but as an artificial language complete in itself with its own morphology and syntax.

By the mid-1960s MT research groups had been established in many countries throughout the world, including most European countries (Hungary, Czechoslovakia, Bulgaria, Belgium, Germany, France, etc.), China, Mexico, and Japan. Many of these were short-lived; an exception was the project which begun in 1960 at Grenoble University.

Throughout this period, research on MT became an ‘umbrella’ for much contemporary work in structural and formal linguistics (particularly in the Soviet Union), semiotics, logical semantics, mathematical linguistics, quantitative linguistics, and nearly all of what would now be called computational linguistics and language engineering (terms already in use since early 1960s). Initially, there were also close ties with cybernetics and information theory, at least in theory (Léon 1997). In general, throughout the early period, work on MT (both theoretical and practical) was seen to be of wide relevance in many fields concerned with the application of computers to ‘intellectual’ tasks; this was true in particular for the research on ‘interlingual’ aspects of MT, regarded as significant for the development of ‘information languages’ to be used in document retrieval systems.

4. The ALPAC report and its consequences

In the 1950s optimism was high; developments in computing and in formal linguistics, particularly in the area of syntax, seemed to promise great improvements in quality. There were many predictions of imminent breakthroughs and of fully automatic systems operating within a few years. However, disillusion grew as the complexity of the linguistic problems became more and more apparent, and many agreed that research had reached an apparently insuperable ‘semantic barrier’. In an influential survey, Bar-Hillel (1960) criticized the

prevailing assumption that the goal of MT research should be the creation of fully automatic high quality translation (FAHQT) systems producing results indistinguishable from those of human translators. He argued that it was not merely unrealistic, given the current state of linguistic knowledge and computer systems, but impossible in principle. He demonstrated his argument with the word *pen.* It can have at least two meanings (a container for animals or children, and a writing implement). In the sentence *The box was in the pen* we know that only the first meaning is plausible; the second meaning is excluded by our knowledge of the normal sizes of (writing) pens and boxes. Bar-Hillel contended that no computer program could conceivably deal with such 'real world' knowledge without recourse to a vast encyclopedic store. His argument carried much weight at the time, although some later developments in artificial intelligence and (within the MT community) the research on knowledge-based systems (section 7 below) have demonstrated that his pessimism was not completely justified.

In 1964 the government sponsors of MT in the United States asked the National Science Foundation to set up the Automatic Language Processing Advisory Committee (ALPAC) to examine the prospects. In its famous 1966 report it concluded that MT was slower, less accurate and twice as expensive as human translation and that "there is no immediate or predictable prospect of useful machine translation" (ALPAC 1966). It saw no need for further investment in MT research; instead it recommended the development of machine aids for translators, such as automatic dictionaries, and the continued support of basic research in computational linguistics. Although widely condemned at the time as biased and short-sighted, its influence was profound, bringing a virtual end to MT research in the USA for over a decade. In addition, it put an end to the widespread perception of MT as a leading area of research in the investigation of computers and natural language; from this time on, for example, researchers in information retrieval dropped all interest in MT and indeed in computational linguistics as a whole.

5. The quiet decade, 1967-1976.

In the United States the main activity had concentrated on English translations of Russian scientific and technical materials. In Canada and Europe the needs were quite different. The Canadian government's bicultural policy created a demand for

English-French (and to a less extent French-English) translation beyond the capacity of the market. The problems of translation were equally acute in Europe, with growing demands for translations of scientific, technical, administrative and legal documentation from and into all the European Community languages. While in the United States MT was not revived for many years, in Canada and Europe (and later in Japan, and elsewhere) its need did not cease to be recognised, and development continued.

At Montreal, research began in 1970 on a syntactic transfer system for English-French translation. The TAUM project (Traduction Automatique de l'Université de Montréal) had two major achievements: firstly, the Q-system formalism for manipulating linguistic strings and trees (later developed as the Prolog programming language), and secondly, the Météo system for translating weather forecasts. Designed specifically for the restricted vocabulary and limited syntax of meteorological reports, Météo has been successfully operating since 1976 (since 1984 in a new version). An attempt to repeat this success with another sublanguage, that of aviation manuals, failed to overcome the problems of complex noun compounds and phrases, and TAUM ended in 1981.

The principal innovative experiments of the decade focused on essentially interlingua approaches. Between 1960 and 1971 the group established by Bernard Vauquois at Grenoble University developed a system for translating Russian mathematics and physics texts into French. Its ‘pivot language’ (influenced to some extent by the research of Kulagina and Mel’chuk) represented only the logical properties of syntactic relationships; it was not a pure interlingua as it did not provide interlingual representations for lexical items – these were translated by a bilingual transfer mechanism. Analysis and generation involved three levels: phrase-structure (context-free) representation, a dependency structure, and a ‘pivot language’ representation in terms of predicates and arguments. A similar model was adopted at the University of Texas during the 1970s in its METAL system for German and English: sentences were analyzed into ‘normal forms’, i.e. semantic propositional dependency structures with no interlingual lexical elements. While in the Soviet Union, Mel’chuk continued his research on a ‘meaning-text’ model for application in MT (see above).

However, by the mid-1970s, the future of the interlingua approach seemed to be in doubt. The main problems identifed were attributed by the Grenoble and Texas groups to the rigidity of the levels of analysis (failure at any stage meant failure to produce any output), the inefficiency of parsers (too many partial analyses which had to be 'filtered' out), and in particular loss of information about surface forms of the SL input which might have been used to guide the selection of TL forms and the construction of acceptable TL sentence structures. As a consequence, it seemed to many at the time that the less ambitious transfer approach offered better prospects.

6. Operational and commercial systems, 1976-1989

In the decade after ALPAC, more systems were coming into operational use and attracting public attention. The Georgetown systems had been operating since the mid-1960s. As well as Météo, two other sublanguage systems had appeared: in 1970 the Institut Textile de France introduced TITUS, a multilingual system for translating abstracts written in a controlled language, and in 1972 came CULT (Chinese University of Hong Kong) specifically designed for translating mathematics texts from Chinese into English.

More significant, however, were the first Systran installations. Developed by Peter Toma, its oldest version is the Russian-English system at the USAF Foreign Technology Division (Dayton, Ohio) installed in 1970. The Commission of the European Communities purchased an English-French version in 1976 and followed it by systems for translation of most other languages of the European Communities (now European Union). Over the years, the original ('direct translation') design has been greatly modified, with increased modularity and greater compatibility of the analysis and synthesis components of different versions, permitting cost reductions when developing new language pairs. Elsewhere, Systran has been installed at numerous intergovernmental institutions, e.g. NATO and the International Atomic Energy Authority, and at many major companies, e.g. General Motors of Canada, Dornier, and Aérospatiale. The application at the Xerox Corporation is particularly noteworthy: post-editing has been virtually eliminated by controlling the vocabulary and structures of technical

manuals for translation from English into French, German, Italian, Spanish, Portuguese, and Scandinavian languages.

From the early 1980s until recently, the main rival of Systran was the system from the Logos Corporation, developed initially by Bernard E.Scott as an English-Vietnamese system for translating aircraft manuals during the 1970s. Experience gained in this project was applied to the development of a German-English system which appeared on the market in 1982; during the 1980s other language pairs were developed.

At the end of the 1980s appeared the commercial METAL German-English system, which had originated from the research at the University of Texas University. After its interlingua experiments in the mid 1970s this group adopted an essentially transfer approach, with research funded since 1978 by the Siemens company in Munich (Germany). Other language pairs were later marketed for Dutch, French and Spanish as well as English and German.

Systems such as Systran, Logos and METAL were in principle designed for general application, although in practice their dictionaries have been adapted for particular subject domains. Special-purpose systems, designed for one particular environment, were also developed during the 1970s and 1980s. The Pan American Health Organization in Washington built two mainframe systems, one for Spanish into English (SPANAM) and the other for English into Spanish (ENGSPAN), both essentially by just two researchers, Muriel Vasconcellos and Marjorie León. Large tailor-made systems have been the speciality of the Smart Corporation (New York) since the early 1980s. Customers have included Citicorp, Ford, and largest of all, the Canadian Department of Employment and Immigration. The principal feature of the Smart systems is (as at Xerox) strict control of input (English) vocabulary and syntax so that minimal revision of output is needed.

During the 1980s, the greatest commercial activity was in Japan, where most of the computer companies (Fujitsu, Hitachi, NEC, Sharp, Toshiba) developed software for computer-aided translation, mainly for the Japanese-English and English-Japanese markets, although they did not ignore the needs for translation to and from Korean, Chinese and other languages. Many of these systems were low-

level direct or transfer systems with analysis limited to morphological and syntactic information and with little or no attempt to resolve lexical ambiguities. Often restricted to specific subject fields (computer science and information technology were popular choices), they relied on substantial human assistance at both the preparatory (pre-editing) and the revision (post-editing) stages.

Some of the Japanese systems were designed for microcomputers. But they were not the first on the market. The earliest were the American Weidner and ALPS systems in 1981 and 1983 respectively. The ALPS system offered three levels of assistance: multilingual word-processing, automatic dictionary and terminology consultation, and interactive translation. In the latter case, translators could work with MT-produced rough drafts. The Weidner systems offered packages for a large number of language pairs, with its Japanese-English systems being particularly popular. In the late 1980s Weidner was acquired by Bravice but shortly afterwards the company was wound up. By this time, however, other systems for personal computers had come onto the market (PC-Translator from Linguistic Products, GTS from Globalink and the Language Assistant series from MicroTac).

7. Research from 1976 to 1989

After the disappointment of its interlingua system, the Grenoble group (GETA, Groupe d'Etudes pour la Traduction Automatique) began development of its influential Ariane system. Regarded as the paradigm of the "second generation" linguistics-based transfer systems, Ariane influenced projects throughout the world in the 1980s. Of particular note were its flexibility and modularity, its algorithms for manipulating tree representations, and its conception of static and dynamic grammars. Different levels and types of representation (dependency, phrase structure, logical) could be incorporated on single labelled tree structures and thus provide considerable flexibility in multilevel transfer representations. Disappointingly, Ariane did not become an operational system (despite involvement in a French national MT project), and active research on the system ceased in the late 1980s.

Similar in conception to the GETA-Ariane design was the Mu system developed at the University of Kyoto under Makoto Nagao. Prominent features of Mu were the use of case grammar analysis and dependency tree representations, and the development of a programming environment for grammar writing (GRADE). The Kyoto research has had great influence on many Japanese MT research projects and on many of the Japanese commercial systems. Since 1986, the research prototype has been converted to an operational system for use by the Japanese Information Center for Science and Technology for the translation of abstracts.

Experimental research at Saarbrücken (Germany) began in 1967, developing from the mid 1970s a multilingual transfer system SUSY (Saarbrücker Übersetzungssystem), displaying a heterogeneity of techniques: phrase structure rules, transformational rules, case grammar and valency frames, dependency grammar, the use of statistical data, etc. Its main focus was the in-depth treatment of inflected languages such as Russian and German, but other languages were also investigated, including English and French. The group also developed a generator (SEMSYN) to convert output from the Fujitsu ATLAS system in order to translate titles of Japanese scientific articles into German.

One of the best known projects of the 1980s was the Eurotra project of the European Communities. Its aim was the construction of an advanced multilingual transfer system for translation among all the Community languages. Like GETA-Ariane and SUSY the design combined lexical, logico-syntactic and semantic information in multilevel interfaces at a high degree of abstractness. No direct use of extra-linguistic knowledge bases or of inference mechanisms was made, and no facilities for human assistance or intervention during translation processes were to be incorporated. However, by the end of the 1980s no operational system was in prospect and the project ended, having however achieved its secondary aim of stimulating cross-national research in computational linguistics.

During the latter half of the 1980s there was a general revival of interest in interlingua systems, motivated in part by contemporary research in artificial intelligence and cognitive linguistics. The DLT (Distributed Language Translation) system at the BSO software company in Utrecht (The Netherlands), under the direction of Toon Witkam, was intended as a multilingual interactive

system operating over computer networks, where each terminal was to be a translating machine from and into one language only. Texts were to be transmitted between terminals in an intermediary language, a modified form of Esperanto. Analysis was restricted primarily to morphological and syntactic features (formalised in a dependency grammar). There was no semantic processing; disambiguation took place in the central interlingua component. The project made a significant effort in the construction of large lexical databases, and in its final years proposed the building of a Bilingual Knowledge Bank from a corpus of (human) translated texts – in this respect anticipating later example-based systems.

A second interlingua project in the Netherlands, innovative in another respect, was the Rosetta project at Philips (Eindhoven) directed by Jan Landsbergen. The aim was to explore the use of Montague grammar in interlingual representations – semantic representations were derived from the syntactic structure of expressions, following the principle of compositionality; for each syntactic derivation tree there was to be a corresponding semantic derivation tree, and these semantic derivation trees were the interlingual representations. A second important feature was the exploration of the reversibility of grammars, i.e. the compilation of grammatical rules and transformations that would work in one direction for syntactic and semantic analysis of a language and in the other direction for the generation (production) of correct sentences in that language. Reversibility became a feature of many subsequent MT projects.

MT research in Japan, initially greatly influenced by the Mu project at Kyoto University, showed a wide variety of approaches. While transfer systems predominated there were also a number of interlingua systems (e.g. the PIVOT system from NEC) and knowledge-based experiments (e.g. the LUTE project at NTT). Japan also launched its own multilingual multinational project in the mid 1980s, with participants from China, Indonesia, Malaysia and Thailand and the involvement of major Japanese research institutes.

During the 1980s many research projects were established outside North America, Western Europe, and Japan – in Korea (sometimes in collaborative projects with Japanese and American groups), in Taiwan (e.g. the ArchTran system), in mainland China at a number of institutions, and in Southeast Asia, particularly in

Malaysia. And there was also an increase in activity in the Soviet Union. From 1976 most research was concentrated at the All-Union Centre for Translation in Moscow. Systems for English-Russian (AMPAR) and German-Russian translation (NERPA) were developed based on the direct approach, but there was also work under the direction of Yurij Apres'jan based on Mel'chuk's 'meaning-text' model – Mel'chuk himself had been obliged to leave the Soviet Union in 1977. This led to the advanced transfer systems FRAP (for French-Russian), and ETAP (for English-Russian). Apart from this group, however, most activity in the Soviet Union focused on the production of relatively low-level operational systems, often involving the use of statistical analyses – where the influence of the 'Speech Statistics' group under Raimund Piotrowski at Leningrad State University was significant.

During the 1980s, many researchers believed that the most likely means for improving MT quality would come from natural language processing research within the context of artificial intelligence (AI). Investigations of AI methods in MT began in the mid-1970s with Yorick Wilks' work on 'preference semantics' and 'semantic templates' (i.e. means for identifying the most common or most favoured collocations of entities, concepts, activities in particular structural relationships, such as subject-verb, verb-direct object, etc.) Further inspiration came from the research of Roger Schank at Yale University, and particularly from the development of expert systems and knowledge-based approaches to text 'understanding'.

A number of projects applied knowledge-based approaches – some in Japan (e.g. the ETL research for the Japanese multilingual project), others in Europe (e.g. at Saarbrücken and Stuttgart), and many in North America. The most important group was at Carnegie-Mellon University in Pittsburgh under Jaime Carbonell and Sergei Nirenburg, which experimented with a number of knowledge-based MT systems. Basic system components were a small concept lexicon for the domain, analysis and generation lexicons for the two languages, a syntactic parser with semantic constraints, a semantic mapper (for semantic interpretation), an interactive 'augmentor', a semantic generator producing TL syntactic structures with lexical selection, and a syntactic generator for producing target sentences.

The concept lexicon and the semantic information in the analysis and generation lexicons (i.e. defining semantic constraints) were language-independent but specific to the domain. The core of the system was the interlingual representation of texts, in the form of networks of propositions, derived from the processes of semantic analysis and of interactive disambiguation performed by the 'augmentor' with reference to the domain knowledge of the 'concept lexicon'. By the end of the 1980s, the Carnegie-Mellon team had fully elaborated its KANT prototype system and was ready to develop an operational knowledge-based system.

8. Translation tools and the translator's workstation

During the 1980s, translators were becoming familiar with the benefits of computers for their work – word processing, creation of individual glossaries, facilities for on-line access and transmission of documents. They were not, however, satisfied with the quality of output of MT systems as such. It was clear already that translators wanted to have computer aids where they are in control of processes, and not to be 'slaves' of automatic systems. Many tools were developed, notably for concordancing, dictionary creation, terminology management, and document transmission. In the early 1990s, however, came the most significant development, the marketing of integrated tools in the 'translator's workstation' (or 'workbench').

The original ideas for integrating various computer-based facilities for translators go back to the early 1980s (Hutchins 1998). Translation workstations combine multilingual word processing, OCR facilities, terminology management software, facilities for concordancing, and in particular 'translation memories'. The latter facility enables translators to store original texts and their translated versions side by side, i.e. so that corresponding sentences of the source and target are aligned. The translator can thus search for phrases or even full sentences in one language in the translation memory and have displayed corresponding phrases in the other language, either exact matches or approximations. In addition, translation workstations often provide full MT programs (for translating segments, paragraphs or whole texts), to be used or adapted by translators as appropriate.

There main vendors of workstations are Trados (probably the most successful), STAR AG in Germany (Transit), IBM (the TranslationManager), LANT in Belgium (the Eurolang Optimizer), and Atril (Déjà Vu). In addition, the Translation Service of the European Commission has developed its own powerful workstation facility (EURAMIS), combining access to its massive dictionary resources and to its own Systran systems. The translation workstation has revolutionised the use of computers by translators; they have now a tool where they are in full control, using any (or none) of the facilities as they choose.

9. MT research since 1989

The dominant framework of MT research until the end of the 1980s was based on essentially linguistic rules of various kinds: rules for syntactic analysis, lexical rules, rules for lexical transfer, rules for syntactic generation, rules for morphology, etc. The rule-based approach was most obvious in the dominant transfer systems (Ariane, Metal, SUSY, Mu and Eurotra), but it was also at the basis of the various interlingua systems, both those which were essentially linguistics-oriented (DLT and Rosetta), and those which were knowledge-based (KANT).

Since 1989, however, the dominance of the rule-based approach has been broken by the emergence of new methods and strategies which are now loosely called ‘corpus-based’ methods (for more details see Hutchins 1994, 1999). The most dramatic development has been the revival of the statistics-based approaches. With the success of stochastic techniques in speech recognition, a group at IBM (Yorktown Heights, New York) began to look again at their application to MT. The distinctive feature of their Candide system was that statistical methods were used as virtually the sole means of analysis and generation; no linguistic rules were applied. The IBM research was based on the corpus of French and English texts contained in the reports of Canadian parliamentary debates (the Canadian Hansard). The essence of the method was first to align phrases, word groups and individual words of the parallel texts, and then to calculate the probabilities that any one word in a sentence of one language corresponds to a word or words in the translated sentence with which it is aligned in the other language. What surprised most researchers was that the results were so acceptable: almost half the phrases

translated either matched exactly the translations in the corpus, or expressed the same sense in slightly different words, or offered other equally legitimate translations.

The second major 'corpus-based' approach – benefiting likewise from improved rapid access to large databanks of text corpora – was what is known as the 'example-based' (or 'memory-based') approach. Although first proposed in 1981 by Makoto Nagao, it was only towards the end of the 1980s that experiments began, initially in some Japanese groups and during the DLT project (section 7 above). The underlying hypothesis is that translation often involves the finding or recalling of analogous examples, i.e. how a particular expression or some similar phrase has been translated before. The example-based approach is founded on processes of extracting and selecting equivalent phrases or word groups from a databank of parallel bilingual texts, which have been aligned either by statistical methods (similar perhaps to those used by the IBM group) or by more traditional rule-based methods. For calculating matches, some groups use semantic methods, e.g. a semantic network or a hierarchy (thesaurus) of domain terms, other groups use statistical information about lexical frequencies in the target language. The main advantage of the approach is that since the texts have been extracted from databanks of actual translations produced by professional translators there is an assurance that the results will be accurate and idiomatic.

Although the main innovation since 1990 has been the growth of corpus-based approaches, rule-based research continued in both transfer and interlingua systems. A number of researchers involved in Eurotra worked within the theoretical approach developed, e.g. the CAT2 system at Saarbrücken; and one of the fruits of Eurotra research was the PaTrans transfer-based system developed in Denmark for Danish/English translation of patents.

Another example of the linguistics-based transfer approach was the LMT project, which had begun under Michael McCord in the mid-1980s, at a number of IBM research centres in Germany, Spain, Israel and the USA. Implemented in Prolog, LMT ('Logic-programming Machine Translation') has the traditional four stages: lexical analysis; syntactic analysis of source texts, producing representations of both surface and deep (logical) relations; transfer, involving both isomorphic

structural transfer and restructuring transformations; and morphological generation of target texts.

The interlingua approach continued at Carnegie Mellon University (CMU). In 1992, it began a collaboration with the Caterpillar company on a large-scale knowledge-based and controlled-language system CATALYST for multilingual translation of technical manuals. Towards the end of the decade, the knowledge-based approach at CMU was combined with developments in statistical analysis of text corpora for the rapid prototyping and implementation of special-purpose systems (DIPLOMAT), e.g. for translation of Serbo-Croatian in military operations.

In the mid 1990s other interlingua-based systems were started, e.g., the ULTRA system at the New Mexico State University developed by Sergei Nirenburg, the UNITRAN system based on the linguistic theory of Principles and Parameters, and the Pangloss project, a collaborative project involving the universities of Southern California, New Mexico State and Carnegie Mellon.

Pangloss itself was one of three MT projects supported by ARPA (Advanced Research Projects Agency), the others being the IBM statistics-based project mentioned above, and a system developed by Dragon Systems. The restitution of US government support for MT research signalled the end of the damaging impact of the ALPAC report (section 5 above). The three systems were subjected to extensive comparative evaluations for ARPA in the mid 1990s.

Since the mid 1980s there has been a trend towards the adoption of ‘unification’ and ‘constraint-based’ formalisms (e.g. Lexical-Functional Grammar, Head-Driven Phrase Structure Grammar, Categorial Grammar, etc.) In "second generation" rule-based systems (such as Eurotra and Ariane) there had been series of complex multi-level representations and large sets of rules for the transformation, mapping and testing of labelled tree representations. Many rules applied only in very specific circumstances and to specific representations, i.e. grammars and mapping rules defined the ‘constraints’ determining transfer from one level to another and hence from SL text to TL text. The introduction of the unification and constraint-based approaches has led to the simplification of the

rules (and hence the computational processes) of analysis, transformation and generation. Now, there are mono-stratal representations and a restricted set of abstract rules, with conditions and constraints incorporated into specific lexical entries. At the same time, the components of these grammars are in principle reversible, so that it is no longer necessary to construct for the same language different grammars of analysis and generation.

The syntactic orientation which characterised transfer systems in the past has thus been replaced by 'lexicalist' approaches, with a consequential increase in the range of information attached to lexical units in the lexicon: not just morphological and grammatical data and translation equivalents, but also information on syntactic and semantic constraints and non-linguistic and conceptual information. The expansion of lexical data is seen most clearly in the lexicons of interlingua-based systems, which include large amounts of non-linguistic information. Many groups are investigating and collaborating on methods of extracting lexical information from readily available lexicographic sources, such as bilingual dictionaries intended for language learners, general monolingual dictionaries, specialised technical dictionaries, and the terminological databanks used by professional translators. A notable effort in this area was the Electronic Dictionary Research project in the late 1980s, supported by several Japanese computer manufacturing companies.

The most significant development of the decade has been the growing interest in spoken language translation, with the challenge of combining speech recognition and linguistic interpretation of conversation and dialogue. The first group was established in 1986 at ATR Interpreting Telecommunications Research Laboratories (based at Nara, near Osaka, Japan), for the development of a system for telephone registrations at international conferences and for telephone booking of hotel accommodation. Slightly later came the JANUS project, under Alex Waibel at Carnegie Mellon University, and subsequently collaborating with the University of Karlsruhe (Germany) and with ATR in a consortium C-STAR (Consortium for Speech Translation Advanced Research). More recently the consortium added further groups in France (CLIPS), Italy (ITC), the United Kingdom (SRI) and elsewhere. The third important spoken language project is

Verbmobil, funded since May 1993 by the German government at a number of universities. The aim is the development of a transportable aid for face to face English-language commercial negotiations by Germans and Japanese who do not know English fluently.

A general feature of all periods of MT research has been the proliferation of small-scale projects. They are often begun by researchers who see MT as a testbed for their own theories of language. These researchers typically take a small selection of sentences, compile a program based on their particular ideas and see how well it can be used to translate. Many theories have been 'tested' in this way, but the results are usually quite inconclusive as the samples are invariably too small.

It is now widely recognised that MT systems and methods can be evaluated properly only with large-scale programs and substantial dictionaries, ideally in actual practical applications. The evaluation of MT is itself a major and vigorous area of research activity (e.g. Vasconcellos 1994), with implications for other areas of computational linguistics and for other applications of natural language processing.

10. Operational and commercial systems since 1990

The use of MT systems accelerated in the 1990s. The increase has been most marked in commercial agencies, government services and multinational companies, where translations are produced on a large scale, primarily of technical documentation. This was the major market for the mainframe systems: Systran, Logos, METAL, and ATLAS. All have installations where translations are being produced in huge volumes; already in 1995 it was estimated that over 300 million words a year were being translated by such companies.

One of the fastest growing areas of use has been in the industry of software localisation. Here the demand is for supporting documentation to be available in many languages at the time of the launch of new software. Translation has to be done quickly, but there is much repetition of information from one version to another. MT and, more recently, translation memories in translation workstations have been the obvious solution.

During the 1990s, the development of systems for specific subject domains and users has also expanded rapidly – often with controlled languages and based on specific sublanguages. For example, Volmac Lingware Services produced systems for a textile company, an insurance company, and for translating aircraft maintenance manuals; Cap Gemini Innovation developed a system to translate military telex messages; and in Japan, CSK developed its own system in the area of finance and economics, and NHK a system for translating Japanese news broadcasts into English.

Since the beginning of the 1990s, many systems for personal computers have appeared. Two of the earliest were PC-Translator (from Linguistic Products, Texas) and Power Translator (from Globalink). Globalink merged with MicroTac (producer of the Language Assistant series), and then was later acquired by Lernout & Hauspie. For Japanese and English the many Japanese products were joined by LogoVista (from the Language Engineering Corporation), and Tsunami and Typhoon (from Neocor Technologies, also now owned by Lernout & Hauspie). From the former Soviet Union have come Stylus (recently renamed ProMT) and PARS, both marketing systems for Russian and English translation; Stylus also for French, and PARS also for Ukrainian. Other PC-based systems from Europe include: PeTra for translating between Italian and English; the Winger system for Danish-English, French-English and English-Spanish; and TranSmart, the commercial version of the Kielikone system (originally developed for Nokia), for Finnish-English translation.

Vendors of older mainframe systems have been obliged to compete by downsizing their systems, e.g. Systran Pro (for use by translators) and Systran Classic (for home use). The METAL system has been adapted by Langenscheidt and GMS (now Sail Labs) as ‘Langenscheidt T1’, and the LMT system was also downsized as the Personal Translator PT (sold jointly by IBM and von Rheinbaben & Busch).

11. MT on the Internet

At the same time, many MT vendors have been providing network-based translation services for on-demand translation, sometimes with human revision as optional extra. In some cases these are client-server arrangements for regular

users; such services are provided, for example, by Systran, Logos, Globalink, Fujitsu, JICST and NEC. Some companies have now been set up primarily for this purpose: LANT in Belgium is a major example, based on its rights to develop the METAL system and the Eurolang Optimizer workstation. Its speciality is the customisation of controlled languages for use with its MT and translation memory systems.

The growing influence of the Internet has been reflected in the appearance MT software products specifically for translating Web pages. Japanese companies led the way, and they were followed quickly elsewhere. At the same time, companies began to offer translation services (often free of charge) on the Internet, usually through 'MT portals', i.e. independent services offering a range of translation systems from one or more system vendors. One of the first (and still best known) was the Babelfish service on AltaVista offering versions of Systran to translate French, German and Spanish into and from English. Equally significant has been the use of MT for electronic mail and for 'chat rooms'. The pioneer was CompuServe which introduced a trial service based on the Transcend system in 1995. Although translation quality is often poor, given the colloquial nature of the source texts, it seems to be widely acceptable.

It is now clear that different types of MT systems are required to meet widely differing translation needs (Hutchins 1999). Those identified so far include the traditional MT systems for large organisations, usually within a restricted domain; the translation tools and workstations (with MT modules as options) designed for professional translators; the cheap PC systems for occasional translations; the use of systems to obtain rough gists for the purposes of surveillance or information gathering; the use of MT for translating electronic messages (electronic mail and Web pages, in particular); systems for monolinguals to translate standard messages into unknown languages; systems for speech translation in restricted domains. It is equally clear that as MT systems of many varieties become more widely known and used the range of possible translation needs and possible types of MT systems will also become wider and stimulate further research and development, quite probably in directions not yet envisioned.

12. Sources

The general history of MT is covered by Hutchins (1986), updated by Hutchins (1988, 1994, 1999), where there are substantial bibliographies with full references for the systems and projects mentioned. Basic sources for the early period are Locke & Booth (1955), Edmundson (1961), Booth (1967), Rozencvejg (1974), Bruderer (1982), and Hutchins (2000). For the 1970s and 1980s there are good descriptions of the main systems in Nirenburg (1987), King (1987) and Slocum (1988). For systems developed during the 1990s the main sources are the survey of techniques by Trujillo (1999), the biennial "MT Summit" conferences and the numerous annual conferences and workshops for MT (e.g. Farwell et al. 1998, White 2000), computational linguistics (principally the Coling and ACL conferences), artificial intelligence, information retrieval, etc.

13. References

ALPAC. 1966. *Language and machines*: computers in translation and linguistics. A report by the Automatic Language Processing Advisory Committee... Washington, DC, National Academy of Sciences.

Archaimbault, S. and Léon, J. 1997. "La langue intermédiaire dans la traduction automatique en URSS (1954-1960): filiations et modèles". *Histoire Épistémologie Langage* 19(2): 105-132.

Bar-Hillel, Y. 1960. ‘The present status of automatic translation of languages.’ *Advances in Computers* 1, 91-163.

Booth, A. D. (ed.) 1967. *Machine translation*. Amsterdam, North-Holland.

Bruderer, H. E. (ed.) 1982. *Automatische Sprachübersetzung*. Darmstadt, Wissenschaftliche Buch-Gesellschaft.

Edmundson, H.P. (ed.) 1961. *Proceedings of the National Symposium on Machine Translation held at the University of California, Los Angeles, February 2-5, 1960*. London, etc., Prentice-Hall.

Farwell, D., Gerber, L. and Hovy, E. (eds.) 1998. *Machine translation and the information soup. Third conference of the Association for Machine Translation in the Americas, AMTA '98, Langhorne, PA, USA, October 1998: proceedings.* Berlin, Springer.

Hutchins, W. J. 1986. *Machine translation: past, present, future*. Chichester (UK), Ellis Horwood; New York, Wiley.

Hutchins, W. J. 1988. 'Recent developments in machine translation: a review of the last five years.' In: Maxwell, D. et al. (eds.) *New directions in machine translation*, 9-63. Dordrecht, Foris.

Hutchins, W. J. 1994. 'Research methods and system designs in machine translation: a ten-year review, 1984-1994.' In: *Machine Translation, Ten Years On*, 12-14 November 1994, Cranfield University. 16pp.

Hutchins, W.J. 1997. 'From first conception to first demonstration: the nascent years of machine translation, 1947-1954. A chronology.' *Machine Translation* 12 (3), 195-252.

Hutchins, W.J. 1998. 'The origins of the translator's workstation.' *Machine Translation* 13 (4), 287-307.

Hutchins, W.J. 1999. 'The development and use of machine translation systems and computer-based translation tools.' *International Conference on Machine Translation & Computer Language Information Processing*. Proceedings of the conference, 26-28 June 1999, Beijing, China, ed. Chen Zhaoxiong, 1-16. [Beijing, Research Center of Computer & Language Engineering, Chinese Academy of Sciences.]

Hutchins, W.J. (ed.) 2000. *Early years in machine translation: memoirs and biographies of pioneers*. Amsterdam/Philadelphia, John Benjamins. (Studies in the History of the Language Sciences, 97).

Hutchins, W.J. and Lovtsky, E. forthcoming. "Petr Petrovich Troyanskii (1894-1950): a forgotten pioneer of machine translation". To appear in *Machine Translation*.

King, M. (ed.) 1987. *Machine translation today: the state of the art*. Edinburgh, Edinburgh University Press.

Léon, J. 1997. "Les premières machines à traduire (1948-1960) et la filiation cybernétique". *Bulag* 22: 9-33.

Locke, W. N. & Booth, A. D. (eds.) 1955. *Machine translation of languages: fourteen essays*. Cambridge, Mass., Technology Press of the Massachusetts Institute of Technology.

Nirenburg, S. (ed.) 1987. *Machine translation: theoretical and methodological issues*. Cambridge, Cambridge University Press.

Rozencvejg, V.Ju. (ed.) 1974. *Machine translation and applied linguistics*. 2 vols. Frankfurt a.M., Athenaion Vlg. [also published as: *Essays on lexical semantics*, 2 vols. Stockholm, Skriptor.]

Slocum, J. (ed.) 1988. *Machine translation systems*. Cambridge, Cambridge University Press.

Trujillo, A. 1999. *Translation engines: techniques for machine translation.* London: Springer.

Vasconcellos, M. (ed.) 1994. *MT evaluation: basis for future directions. Proceedings of a workshop…2-3 November 1992*. Washington, DC, Association for Machine Translation in the Americas.

White, J.S. (ed.). 2000. *Envisioning machine translation in the information future. 4th conference of the Association for Machine Translation in the Americas, AMTA 2000, Cuernavaca, Mexico, October 2000: proceedings*. Berlin, Springer.

Source: *Histoire, Epistemologie, Langage*, Tome XXII, fasc. 1 (2001), p.7-31]