PANEL: THE PIONEERS OF MT

Muriel Vasconcellos, Moderator
Present at Summit VI: Victor H. Yngve, Igor Mel'čuk, Alexander Zhol'kovsky, Peter Toma
Christine Montgomery, A.F.R. Brown, Roger Heller, Loll Rolling

Additional contributors: Andrew D. Booth and Kathleen H.V. Booth,
Michael Zarechnak, Winfred P. Lehmann

The idea of using the computer to translate between human languages was first committed to paper half a century ago. We celebrate the 50th anniversary of machine translation by bringing together a number of its pioneers to reminisce about initiatives, early progress, and significant milestones in those first decades. It is a unique and very special opportunity to have so many MT pioneers together in the same room after such a long time. Indeed, several of the panel participants have not seen one another in more than 30 years. Four of our invitees were unable to be present but have contributed written accounts to help complete the story: Andrew and Kathleen Booth, Michael Zarechnak, and Winfred Lehmann. Four others have preferred to make oral contributions only: Igor Mel'čuk, Alexander Zhol'kovsky, and Anthony Brown. The session is being audiotaped and videotaped for archival purposes.

The contributions below are arranged in chronological order of the author's entry on the MT scene. Reading them sequentially, one gains an interesting picture of what the early days were like. We are fortunate to be able to start out with an account by Andrew and Kathleen Booth that begins with the very first musings about MT in 1946. The story unfolds as we follow the various contributors, ending with an account by Loll Rolling of the introduction of MT at the European Commission, unarguably the largest MT "user" in the world.

THE ORIGINS OF MT

Andrew D. Booth and Kathleen H.V. Booth
Autonetics Research Associates
Sooke, BC V0S 1N0 Canada
abooth@ios.bc.ca
Tel: (250) 642 5352

1. Origins

In his 1666 dissertation, *The Art of Combinations,* Leibniz had clearly indicated in his discussion of the possible mechanization of both arithmetic and thought processes, that his logical processor could be used to transform one language into another. He also considered the extraction of ideas from text and their expression in terms of a metalanguage (Saw 1954). The culmination occurred in his design of the first practicable mechanical multiplying machine (1694) and in a binary multiplier (Eriksson et al. 1996).

Another early claim, by the Russians, was to the effect that a Soviet engineer had produced a translating machine in the 1930s. This appears unlikely because MT depend heavily on a large and complex dictionary—something which was not even on the horizon at that time.

We date realistic possibilities starting with two meetings held in 1946. The first was between Warren Weaver, Director of the Natural Sciences Division of the Rockefeller Foundation, and Norbert Wiener. The second was between Weaver and A.D. Booth in that same year. The Weaver-Wiener discussion centered on the extensive code-breaking activities carried out during the World War II. The argument ran as follows: decryption is simply the conversion of one set of "words"—the code—into a second set, the message. The discussion between Weaver and A.D. Booth on June 20, 1946, in New York identified the fact that the code-breaking process in no way resembled language translation because it was known a priori that the decrypting process must result in a unique output. The main purpose of this meeting, however, was to interest the Rockefeller Foundation in supporting development of an electronic computer at the University of London.

Things remained thus until the end of a visit to the Institute for Advanced Study, Princeton, in 1947. On May 20 of that year another talk with Weaver showed that the Foundation was unlikely to fund a machine at London for numerical work but that a submission indicating interest in MT would be well received. This was duly submitted and was successful. The core of the proposal involved the use of the high-speed storage organ as an electronic dictionary.

2. London, 1948-1962

On our return to London, work was at once started on two fronts: ADB designing and supervising the construction of the new computer and KHVB directing the development of suitable programs. It must be noted that in those days all coding was done in machine language (not even "assembler") and had to be inserted via switches

Some code for a computerized dictionary was tested and found to be very effective. The lookup method used from the first was binary partitioning—a technique which we regarded as so obvious that we did not deem it worth publication until we discovered that several major computer groups were using sequential search strategies (A.D. Booth 1955). More extensive descriptions of our work are given in *Automatic Digital Calculators* (A.D. Booth & K.H.V. Booth 1953) and in *Programming for an Automatic Digital Computer* (K.H.V. Booth 1958). The latter contains a complete MT program of that era.

Other activities of our group included work on the "translation" of English to Braille by J.P. Cleave (1956), language processing in general by M. Levison (1962), and the chronology of the Plato dialogues by L. Brandwood. Much of the work is described in *Mechanical Resolution of Linguistic Problems* by A.D. Booth, L. Brandwood, and J.P. Cleave (1958). Our thanks are due the Nuffield Foundation, which provided major funding during this period.

3. Saskatchewan, 1962-1972

In 1962 we migrated to Canada and in due course became involved with the Queen's Printer in a most interesting project. One problem in Canada is the presence of two "official" languages: English and French. The parliamentary system requires that the Hansard, the daily transcript of the day's debates, be available in both languages the next day. This was the problem which we were asked to address. Our own view was, and still is, that output of good quality can be produced only with a post-editor. With the aid of linguists Corinne Brown and Gloria Geitz and programmer Charles Stock (K.H.V. Booth & Stock

1971) we produced such a program and demonstrated it at the National Research Council Laboratories in Ottawa in 1971. It is interesting to note that the only languages available on our University mainframe were FORTRAN and COBOL. The program was written in the latter.

4. Lakehead University, 1972-1978

When we migrated to Ontario in 1972, our work on MT was continued with emphasis on the construction of a large dictionary and the use of statistical techniques in linguistic analysis (A.D. Booth et al.).

5. Retirement, 1978-

Our work on MT terminated with our retirement in 1978 except for one small but interesting experiment. In 1979 the first personal computers became available and we were involved in a project to develop an energy management package. The machine chosen was a Commodore PET 2000 and, having some time to spare, ADB decided to write an MT program for this machine. Here the language was BASIC and a French-to-English program was soon produced and demonstrated to the Canadian National Research Council in 1980. The program incorporated most of the methods devised in our earlier work but, because of storage limitations, it had only a small dictionary. The output quality was equal to or better than that generated by currently available inexpensive "translation" programs. The main superiority of our program lay in the use of our category count technique (K.H.V. Booth 1970) for the resolution of such ambiguities as Chomsky's "The pen is in the box."

REFERENCES

Booth, A.D. 1955. On the Use of a Computing Machine as a Mechanical Dictionary. *Nature,* 176:565.

Booth, A.D., and Booth, K.H.V. 1953. *Automatic Digital Calculators.* 1st ed. London: Butterworths.

Booth, A.D., Brandwood, L., and Cleave, J.P. 1958. *Mechanical Resolution of Linguistic Problems.* London: Butterworths.

Booth, K.H.V. 1958. *Programming for an Automatic Digital Calculator.* London: Butterworths.

Booth, K.H.V. 1960. Statistical Parsing by Computer. *Pensiero e Linguaggio* 1:107.

Booth, K.H.V., and Stock, C. 1971. English-French Translation on a Computer. *Pensiero e Linguaggio,* 2:231.

Cleave, J.P. 1956. The Application of an Automatic Digital Computing Machine to Problems of Data Transformation and Language. Ph.D. thesis, University of London.

Eriksson, K., Estep, D., Hansbo, P., and Johnson, C. 1996. *Computational Differential Equations.* Cambridge: Cambridge University Press. p. 20.

Levison, M. 1962. The Application of a Computing Machine to Linguistic Problems. Ph.D. thesis, University of London.

Saw, Ruth Lydia. 1954. *Leibniz.* London: Penguin Books. p. 214.

SOMETHING OLD AND SOMETHING NEW

Victor H. Yngve
The University of Chicago
Home: 28 Crest Drive Dune Acres
Chesterton, IN 46304
v-yngve@chicago.edu
Tel: (219) 787-8340

As part of my research assistantship as a graduate student in physics at the University of Chicago in the late 1940s, I was involved in preparing instrumentation for detecting incoming cosmic radiation in the stratosphere. The electronic circuits we were using were similar to circuits found in the literature on the new automatic electronic computers, the so-called giant brains. It occurred to me, as it did to others, that these machines might be useful for other purposes besides computing—for instance, translating languages to help eliminate some of the barriers between countries in a post-war world.

Around 1950 we had been experimenting with recording our cosmic ray data on the new tape recorders. At one point I toyed with the idea of building a word-for-word translation machine in which a dictionary in digital form would be magnetically recorded. Such a machine would be comparatively simple to design and build. I realized that the translations would be imperfect, but perhaps they would go a long way toward being useful. The linguistic difficulties involved in word-for-word translation even then loomed larger than the mechanical difficulties. I did not actually try to build such a machine, postponing such work for after I had finished my dissertation.

In the meantime I consulted with some University of Chicago linguistics faculty members and entered on an extracurricular program of reading in the field of linguistics. It seemed clear that a solution for many of the multiple meaning problems that showed up could be achieved if the machine could work out a syntactic analysis for each sentence. I started working on this and a pencil-and-paper scheme gradually took shape.

In looking for a job for after graduation to pursue these interests, I visited Claude Shannon at Bell Telephone Labs. I had read Shannon's paper on the entropy of printed English, and thought that perhaps Bell Labs would be interested in my ideas. They weren't interested, but Shannon told me that there was someone at MIT who was organizing a conference on the subject.

I wrote to MIT for information and was privileged to be invited, while yet a graduate student, to attend the first international conference on machine translation, organized by Y. Bar-Hillel, which took place at MIT on June 17-20, 1952. Through the organizing activities surrounding that conference I became acquainted with the early research of others on MT, including the Warren Weaver memorandum.

I interviewed for a job with Léon Dostert at Georgetown University, and was, if I remember rightly, introduced to Paul Garvin and shown his card file for translating some sample Russian sentences into English. I also interviewed at MIT and subsequently took up a full-time research position there on July 1, 1953, Bar-Hillel having returned to Israel.

In my first major paper on MT, I completed the scheme I had been working on for analyzing German sentences syntactically and translating them into English. This scheme was published in 1955 in the Locke and Booth book that came out of the conference. I believe this was the first published table-driven syntactic parsing scheme and perhaps the first published translation scheme that was completely explicit and programmable. It was actually later programmed.

However, problems showed up in efforts to expand this scheme to include a larger dictionary and additional recognized phrase types. There followed a two-step scheme that separated sentence recognition and sentence production. This was published in 1955 in the journal *MT* that Bill Locke, the chairman of the Modern Language Department, and I had started in March 1954. I was disappointed that Garvin's scheme was not published in his name at the time and could only be gleaned from newspaper accounts.

There followed a three-step scheme, which was also published in *MT,* and we developed the COMIT programming language, which allowed the growing group of linguists and logicians to become productive in MT research. The first successful grammar written in COMIT ran on the IBM 704 over the Fourth-of-July weekend of 1960. Our efforts were centered on the scientific and linguistic issues relevant for advancing the possibility of useful machine translations.

In order to test our grammars, we arranged to drive our sentence construction routine with random specifiers so as to produce a random sample of the output of the grammar. Faults in the grammar showed up easily and we could go back and in essence debug the grammar. This was the first testing of grammars by the method of random generation. It worked admirably, and this scheme served us for many years of grammatical research.

Although our grammars grew in size and complexity and were capable of generating a wide variety of sentences at random, this very success pointed up a number of residual problems with which many of you are already familiar. Some of these problems seemed particularly resistant to solution. We felt overwhelmed. Where could we turn for appropriate theory?

Thus, when I moved to The University of Chicago in 1965, it was with the hope of being able to find out what was wrong with linguistic theory that it could not support such an important application as MT.

I can now report, over 33 years later, that it has required nothing less that laying completely new foundations for general linguistics. I should like to tell you more about this exciting new development, but it would take much too long. For those who are interested in learning more, I have prepared a textbook that you can work through at your own pace (Yngve 1996). You might also wish to visit my web page at:

`http://humanities.uchicago.edu/humanities/linguistics/faculty/yngve.html`

or contact me by e-mail at v-yngve@uchicago.edu.

REFERENCE

Yngve, Victor H. 1996. *From Grammar to Science: New Foundations for General Linguistics.* Amsterdam/Philadelphia: John Benjamins.

THE FULCRUM APPROACH TO MACHINE TRANSLATION

Christine A. Montgomery
Language Systems Inc.
5959 Topanga Canyon Blvd.
Woodland Hills, CA 91367
chris@lsi.com
(818) 703-5034

In a paper from a distinguished collection of papers prepared for a 1959 course entitled “Computer Programming and Artificial Intelligence,” Paul Garvin described two types of machine translation problems “in terms of the two components of the term: machine problems, and translation problems.” While the machine problems made us crazy, the translation problems made us think differently about language than we might otherwise have done, which has had some advantages and some disadvantages in the long run. I will save anecdotes about the former and comments about the latter for the discussion.

In this paper I will focus on the translation problems and, in particular, the translation approach that was developed by Paul Garvin, with whom I was associated, initially at Georgetown University, and later in the Synthetic Intelligence Department of the Ramo-Wooldridge Corporation and successor corporations: Thompson Ramo Wooldridge and Bunker-Ramo.

The initial MT experiments at Georgetown University in 1954 were designed by Garvin and carried out jointly with IBM, who implemented Garvin's design on an IBM 701 computer. These experiments were necessarily quite limited in scope and were a demonstration of possibility rather than feasibility.

The approach Garvin developed subsequently, entitled “The Fulcrum Approach to Machine Translation,” was based on what he referred to as the “fulcra” of syntactic phrases of various types. The fulcrum of a syntactic unit is the constituent that “provides the best point of leverage for its translation,” essentially because it governs the syntactic behavior of the unit, functioning as the head of the given syntactic unit or phrase. Thus the verb or other predicate is the fulcrum of the clause, the noun is the fulcrum of the noun phrase, the participle or other governing modifier controls the participial or modifier phrase, the preposition governs the prepositional phrase. Syntactic processing of a source language sentence from a text was performed via a series of analytical passes through the sentence, in which the fulcra of the various types of syntactic phrases were determined, and the constituents of the given phrases identified.

Although we considered the syntactic passes to be the most important, they were outnumbered by a series of pre- and post-syntactic analysis passes through the sentence, as exemplified by the following list of topics discussed under “Illustration of [Translation] Process” in the initial section of of a Ramo-Wooldridge technical report submitted to Rome Air Development Center in February, 1961:

Input processing of Cyrillic text
Transliteration
Dictionary lookup
Morphological analysis of missing words
“Reinflection” of English translations of missing words
Idiom translation
Syntactic analysis

Translation of “word combinations”
Article insertion
Insertion of prepositions for translation of case suffixes
Insertion of auxiliaries
Rearrangement of word order within clauses
Rearrangement of word order within modifier phrases
Rearrangement of word order within verb phrases

Many of the above processes appear quite trivial, but they were necessary to deal with the variety of phenomena making up the elements of typical Russian scientific texts, in order to translate them into English. An example of elements that had to be isolated in input processing were bracketed items called “number/symbol strings.” These could be equations, quantities, or other expressions that might have a key syntactic role in the sentence, but they might also contain internal periods or other punctuation that could potentially result in erroneous identifications of phrase, clause, or sentence boundaries with occasionally catastrophic results, such as causing the text buffers to overflow into other memory areas. In at least one case this resulted in overwriting the operating system code and derailing not only our MT processing but the operations of the service bureau where our processing was performed, making us extremely unpopular clients there.

In the initial Fulcrum system described in the above report, analysis of the source sentence and synthesis of the target sentence were interleaved. Russian words were replaced by English equivalents (or sets of equivalents) during dictionary lookup, and subsequent processes were aimed at refining the raw, word-for-word translation produced as a result of the initial lexical substitution. Proper English word order (or an approximation of this) was produced by “rearrangement routines” which operated within specific types of syntactic phrases to re-sequence constituents according to the requirements of English word order.

At the time the infamous ALPAC report appeared, Garvin and his staff at Bunker-Ramo had just completed the design and partial implementation of Fulcrum II, in which analysis of the Russian source language sentence and synthesis of the English target language translation were viewed as separate, sequential processes. The Fulcrum II dictionary entries were substantially enriched, including more detailed syntactic and semantic information. These changes improved the attachment of governed structures and facilitated the revision of attachments based on subsequent analytical processing. Unfortunately, full implementation of this more flexible system was never achieved because of the well-known impact of the ALPAC report on MT funding.

REFERENCES

Bunker-Ramo Corporation. 1965. A Syntactic Analyzer Study. Final Technical Report (Contract AF30(602)-3506: Rome Air Development Center). 31 July 1965.

Garvin, Paul L. 1959. Machine Translation. In *Computer Programming and Artificial Intelligence.* University of Michigan, pp. 1-8.

Ramo-Wooldridge. 1961. Machine Translation Studies of Semantic Techniques. Final Technical Report. (Contract AF30(602)-3506: Rome Air Development Center). 22 February 1961.

LET ME TELL YOU HOW IT REALLY WAS

Michael Zarechnak
Georgetown University
Washington, DC 20057
mzarechnak@aol.com

Bliss was it in that dawn to be alive,
But to be young was very heaven.

— Wordsworth

I would like to share with you, the younger generation, some memories that may be worth examining within the context of the past, present, and future of machine translation.

In the history of MT, the year 1947 is associated with the names of A.D. Booth and Warren Weaver and their correspondence in which they speculated about using a computer to translate messages from one natural language into another.

Those were the days of many fast-moving events. The ensuing decade was in fact the dawn of the Space Age. The launch of Sputnik I, the first man-made satellite to orbit the Earth, by the Soviet Union in 1957, was a shock to Americans. Fueled by the Cold War, MT researchers in the 1950s and 1960s were doing pioneering scientific work, paving the way for later commercial development.

In 1954 I was at Harvard University studying under Professor Roman Jakobson. The October issue of *Voprosy Yazykoznanija* reported the world's first demonstration of machine translation by Georgetown University and IBM in the United States. In that demonstration a computer was programmed to translate selected sentences from Russian into English using a restricted dictionary of 250 words. Professor Jakobson gave me an excellent recommendation to Professor Léon Dostert, founder and head of the Georgetown MT project. I soon joined the group of linguistic researchers led by Professor Paul Garvin who had provided the linguistic formulation for the Georgetown-IBM experiment.

I believe that nobody can detract from Professor Garvin's pioneering contribution to solving, on a computer, the two basic problems involved in the translation process: (1) the decision steps connected with the selection of lexical units, and (2) the arrangement units in the target language. These two problems reflect the two axes in natural language structure referred to by de Saussure as the associative and syntagmatic axes.

As a true pioneer, Professor Dostert combined vision and practice. The goal of the Russian-English MT project was to produce output in English such that the outside evaluator, a specialist in the field, would find the quality of the text acceptable as a source for gleaning the information contained in the original language. The dictionary entries containing the Russian words, their English equivalents, and their morphological, syntagmatic, syntactic, and semantic codes were keypunched and introduced into the computer memory.

Léon Dostert had a practical vision of machine translation and believed that it was feasible. He was always pushing for it to become a reality. He knew how to motivate his co-workers by letting them conduct research independently, provided only that, at the end of the day, they would be willing to come down from the level of theory and try to embody the theory in some practical system.

Not all the heads of other MT projects thought that MT was feasible, and many of them concentrated on limited studies of this or that phenomenon. There were exceptions, of course—most notably, the research of Professors Victor Yngve and Winfred Lehmann.

At MIT, Victor Yngve developed a special programming language to help linguists write their own rules and pursue the MT goal as a long-range project with possible insights on the way. He devoted his attention to the left-branching of English. He also produced the journal *Mechanical Translation,* which published an article of mine on semantic support for the resolution of syntactic ambiguities.

At the University of Texas, Winfred Lehmann was theoretically oriented and produced a tripartite system for German-English. Later the system was further expanded and became known as METAL, the basis for one of today's most sophisticated MT systems.

Léon Dostert set himself apart from the others by stating that one should work the way Descartes did: tackle one part of a problem at a time, learn from it, take on another part, and so on. Eventually, through these cumulative efforts, insights would grow and general comprehension might be closer at hand.

My contribution to machine translation is reflected in the original name of the Georgetown system, General Analysis Technique (GAT), which stressed that the analysis of the source and target language should be done independently of one another and that the transfer of meaning should be done between a pair of languages. To accomplish these tasks the members of the MT team should know both languages and be linguistically trained. Statistical considerations should be invoked when the structural (distributional and contrastive) information does not provide a clear solution.

The GAT system's lexical entries and translation rules were expressed in a machine-independent form known as the Simulated Linguistic Computer (SLC). SLC was developed by Dr. A.F. Brown on the basis of the special programming language that he created for French-English MT. It was demonstrated at a Unesco conference in 1959. He began translating a single sentence and then added another sentence, correcting or adding more information to translate these sentences into English. This approach resulted in the insight that one should make a distinction between "global" operations, whose domain was the entire sentence (e.g., a "subject routine"), and "local" operations, triggered by individual words (e.g., a dictionary entry with multiple equivalents in English). SLC permitted us to replace the fifteen passes through the computer (based on the implementation developed by Peter Toma) with only two: one for dictionary lookup and the other for everything else. In 1993 Dr. Brown ported the GAT-SLC system to the C programming language and today it runs on PCs.

The researchers who did not believe in the feasibility of MT succeeded in persuading a government panel in 1966 that "high-quality" MT could not be achieved and that research money should instead be devoted to computationally oriented theoretical linguistic studies. Their conclusions were published by the National Science Foundation under the name of the Automatic Language Processing Advisory Committee (ALPAC). It was not until after a conference organized by Professor Lehmann in 1971, at which the earlier critic Yehoshua Bar-Hillel no longer defended "fully automatic high-quality translation," that it become clear that something had to be done to return to a pragmatic approach to MT development.

That experience should be our lesson for the future. At Georgetown we worked with bilingual people who were linguistically trained. They were our models for capturing intuitions. We should pay more attention to semantic and pragmatic situations and try to formalize them. The intersecting pragmatic observations for a given set of languages might provide some clues for building an intermediary language.

MY FIRST 30 YEARS WITH MT

Peter Toma
President, Voxtran, Inc.
Postfach 101033
53448 Bad Neuenahr-Ahrweiler. 1 Germany
Tel.: +49-2641-203653; fax: +40-2641-203630

1. CalTech, 1956-1957

In 1956 I was working at the California Institute of Technology in Pasadena as an assistant in the Physics Department when one day my colleague Tom Taylor announced, "Peter, we're going to get a computer!" I instantly knew that this was something for me. I sensed that the arrival of the computer, a Datatron 205, was going to be a turning-point in my life. I would do something with languages, I thought. I got a deep feeling—which has driven me all these years—that I was destined to program the computer to translate languages. At that time I didn't have the faintest idea that other MT initiatives were already under way.

As I gained knowledge about the computer's capabilities, I became even more intrigued with the idea of working with text and languages. I volunteered to stay during the night and tend the computer. This enabled me to key in transliterations of Russian text and begin to experiment. The output (on paper tape!) was actually quite good. I had developed an interesting approach. Someone alerted the press and a reporter came. Unfortunately, this publicity was premature. The news reports claimed that I had achieved more than I actually had. I had two choices, either see if I could get support and continue, or join up with another project. I went to Washington, D.C., to attend a meeting at the National Science Foundation. There I met Michael Zarechnak, Gilbert King, and others in the field for the first time.

In 1957 I gave up my job at CalTech and joined Gilbert King at International Telemeter. I hadn't been there very long, however, when King decided to take the project to the Rome Air Development Center. I preferred not to follow him. I submitted a proposal to the National Science Foundation, but it was not accepted. I joined the project at Georgetown University on March 1, 1958.

2. Georgetown University, 1958-1961

Of the various approaches being tried at Georgetown, the General Analysis Technique (GAT), impressed me the most and I began working with that group. My job was to reduce the linguistic logic to computer representation. The IBM 705 they were using was much better than the Datatron. My first task was to write a morphological program. By then I had read literature in the field and my ideas were more crystallized. After the morphology was completed I proceeded to program the entire GAT system—a program that I baptized SERNA. We presented our work on Friday mornings at our weekly seminars, and I vividly remember the time when it was up to me to present the SERNA system to a visitor—Allen Dulles, then director of the Central Intelligence Agency.

In the fall of 1958 I was given a free hand in the programming with permission to contract part-time programmers. We were paying for computer time by the minute, but then the Pentagon offered us the free use of their IBM 705 on week-ends and helped us to recruit programmers from there. From November 1958 until June 1959 I spent every week-end as well as Christmas and New Year's Day working at the IBM 705 in the Pentagon with my small but enthusiastic programming staff. Our devotion and hard work culminated in a successful demonstration on June 8, 1959. As I worked, I noticed that the tapes in the tape room were stacked up to the light bulbs and mentioned my concern because I had a strong presentment of danger.

The demonstration went very well. Immediately afterwards I asked Lt. Harry Stevens to copy all our tapes. We then drove them to the sponsoring agency, where I requested that they be put in a safe. Later that same afternoon I flew home to California. A few days later headlines in the newspapers announced that there had been a fire in the tape room at the Pentagon. Luckily, our tapes were backed up.

3. Prelude to SYSTRAN, 1961-1965

In 1961 I left Georgetown and took charge of the machine translation section of Computer Concepts in Los Angeles, where I developed two entirely new MT systems: AUTOTRAN and TECHNOTRAN, both of them oriented toward the IBM 7090 computer. The logic was totally different from what Georgetown had used. In June 1964 I was invited to the University of Bonn to demonstrate the systems. During my flight to Europe I read one of the first IBM 360 manuals and immediately sensed that this line of computers would enable me to realize all my ideas for MT. While demonstrating AUTOTRAN and TECHNOTRAN to European scientists, I devoted my free time to working on a new logic oriented toward the 360 line. Emulating a 360 on the 7090, I was one of the first to write programs for the 360. At that time everyone tried to see how much they could accomplished with the fewest instructions, and in this connection my binary search won particular recognition.

It was around that same time that the Applied Language Processing Advisory Committee (ALPAC) was formed under the aegis of the National Academy of Sciences to assess the state of MT. I was out of the country, but output from AUTOTRAN and TECHNOTRAN was obtained surreptitiously and included in the ALPAC exercise.

In May 1965 the German Science Foundation convened 10 of the country's leading linguists and 10 computer specialists to evaluate my ideas and planned approaches for SYSTRAN. The evaluation was positive and I received a grant. That same year I began studies in Bonn that led to my Ph.D. degree.

4. SYSTRAN, 1965-1987

At the end of 1965 I returned to California and submitted a proposal to the Rome Air Development Center for the development of SYSTRAN and was awarded funding. This and further contacts made it possible to start my own company in La Jolla. One of the most important steps at that time was the development of a special language for linguists which made it possible for them to communicate directly with the computer.

In January 1969 SYSTRAN was installed in the Foreign Technology Division at Wright-Patterson Air Force Base in Dayton, Ohio. From the beginning Systran translated 10 times faster and much better than its predecessor, IBM's Mark II. Yearly contracts during the period 1970-1973 allowed us to introduce many new features—among them, semantic categories and "LS units" to store lexical information for resolving ambiguities.

SYSTRAN's next major project was the Apollo-Soyuz mission. Under a contract from NASA, in two years we improved the Russian-English system and developed English-Russian as well—our first English source combination. In 1975 SYSTRAN had been demonstrated extensively in Europe, and as a result, in 1976 a contract was awarded by the European Commission, which Loll Rolling is going to talk about. Also in 1976 came the XEROX contract for multiple targets from an English source. By that time SYSTRAN was being developed by two La Jolla companies, LATSEC, for the contracts with FTD and NASA, and World Translation Center Inc., for contracts outside the U.S. Government.

While it is beyond the scope of this paper to report my entire history with SYSTRAN, let me just mention that in 1987, after 16 years of pendency, I was granted a U.S. patent for SYSTRAN, the first patent for machine translation as far as I am aware. Two years earlier I had sold both companies to the Gachot family in France.

MT AT TEXAS: THE EARLY YEARS

W. P. Lehmann
Linguistics Research Center
The University of Texas at Austin
Austin, TX 78712
lehmann@mail.utexas.edu
Tel: (512) 471-4566

On my return from directing the Georgetown English Language Program in Ankara Turkey during the years 1956-1957, Professor Léon Dostert suggested that I undertake machine translation, with concentration on German, since he was doing Russian. Having spent World War II at the predecessor of NSA dealing with translation of Japanese messages, where forerunners of the computer were lined up by scores, I had not been without interest in Weaver's suggestion that the newly developed instrument be employed for MT. But other obligations interfered with more than merely keeping abreast of reports on activities. Dostert sweetened his suggestion with a grant of $10,000, a sum large enough, in the days when professorial salaries scarcely reached half as much, to employ a good group of graduate students.

The work consisted chiefly in learning what was going on and planning. The University's computer, an IBM 650, was preempted by physicists, for what it was worth. Our best informed member in computational areas was Nick Hopkins, a graduate student in archeology, who among other capabilities had learned on trips to Mexico how to make sandals by cutting strips from old tires and keeping them on with pieces of string. Most of the other members were linguists well acquainted with German. The best result of the year was an award of $300,000+ from the U.S. Army Signal Engineering Laboratories to investigate the feasibility of MT. The Army was planning to develop a computer, Moby Dick, to be located at division headquarters; MT was to be one of its functions.

The grant provided means to employ full-time programmers as well as linguists. One of these, Gene Pendergraft, had worked with Marvin Minsky, and soon became the chief systems specialist. We set up three groups: linguists, programmers, and mathematicians. Sponsorship by the Signal Corps gave us the advantage of using the IBM 709 at Fort Huachuca. Programs and translation algorithms were developed, prepared by our keypunchers for entry, and taken out to Arizona to be checked. Our progress was recorded in quarterly reports. The third, dated 31 January 1960. and the ninth, dated 31 July 1961, state the theoretical basis of all our further work, which led to the production of the METAL translation system. Following the theoretical position of Charles Sanders Pierce, we treat language as a system with three components: syntactics, semantics, and pragmatics. Programs were designed to manage the linguistic rules and were kept distinct from the grammar and the lexicon.

Through personal contacts, conferences, and publications we kept in close touch with other groups and profited by their conclusions. Leon Dostert generously sent down his specialists to inform us of progress at Georgetown. Victor Oswald at UCLA credited C.V. Pollard in our department with developing syntactic rules for scientific German, as did Yehoshua Bar-Hillel (1964:159), at a time when language teaching concentrated on morphology. Erwin Reifler of the University of Washington developed the notion of sublanguages, basically an application of pragmatics that recognized the presence of distinct types of language; MT could concentrate on technical language rather than attempt to control such sublanguages as that of literature. Our progress was recognized by further grants, and by the founding of the Linguistics Research Center in 1961.

The conferences kept us in close touch with members of an expanding technology. Léon Dostert remained the operational father of MT. Known at the highest levels of government through his wartime duty as French interpreter for General Eisenhower, when funds were running out Dostert would simply call on Allen Dulles, then head of the Central Intelligence Agency, who would turn him over to his staff with the injunction to "give Léon what he needs."

Victor Yngve must be credited with producing the first appropriate computer language, COMIT, and with producing a journal for the field. Yehoshua Bar-Hillel, a logician who might be called the theoretical father of MT, insisted on FAHQT (fully automatic high-quality translation), then became highly pessimistic and vocal about its realization. Curiously, one of the major proposed shortcomings was the need for post-editing, a regular procedure among translation agencies. His unfortunate insistence, clear still in the "Feasibility Study on Fully Automatic High Quality Translation" produced by the Center in December 1971, resulted in part from ignorance of translation procedures, as did the devastating National Academy of Science report of 1966 by the engineer John Pierce—otherwise known as the "ALPAC Report." By contrast, the far-sighted supporter Zbigniew L. Pankowicz, of the Rome Air Development Center, may be credited with keeping the technology from vanishing by providing the funding for LOGOS and the Center.

In recollection, the measures required to achieve results hardly seem credible. Access to the computer was through punched cards; to cut costs, Dostert established a keypunching center in Frankfurt, Germany, where labor was cheaper. Rules and algorithms were transferred to tapes, which whirred crazily in the computer room. In the absence of computer languages, coding was done initially in machine language, then in Fortran. When the University acquired a top-notch computer, the CDC 1604, we used it on Sunday mornings, because testing our programs required full use of the machine. Somewhat later the Center acquired its own, a lesser CDC, so that our systems personnel no longer had to defile the Sabbath. But the pitiful success provided as much assurance that MT was feasible as did Newton's apple for his views on gravitation.

Yet Bar-Hillel's vigorous negativism, coupled with unethical activities, succeeded in undermining success. As Bar-Hillel stated (1964:9-10), funds for MT were siphoned off, not least in the institution that was his home, so that the amount invested was larger than that applied. Another damaging effect was produced by "demonstrations" that had been pre-cooked. Even the self-sacrificing support provided by Pankowicz could not offset the devastating effect of the shameful ALPAC Report of 1966, which eliminated almost all research in this country and abroad. Fortunately, with the support of Pankowicz, the Center was able to continue its work until Siemens found that their human translators could not keep up with the demands of quantity and time. Their funds and that of other U.S. agencies maintained research at the Center so that METAL became an operative system, which Somers has credited as "a success story in the development of MT" (Whitelock: 1995:198).

REFERENCES

Bar-Hillel, Yehoshua. 1964. *Language and Information: Readings.* MA: Addison-Wesley.

Whitelock, Peter, and Kieran, Kilby. 1995. *Linguistic and Computational Techniques in Machine Translation System Design.* London: UCL Press, See my review, *Machine Translation* 12:261-69, 1997.

THE CHALLENGE OF KEYPUNCHING FOR MT

Roger A. Heller
5615 Bermuda Dunes Lane
Houston, TX 77069
Rogheller@aol.com
Tel: (281) 440-0708

In 1960 Léon Dostert, then Director of the Georgetown University Mechanical Translation Project, convinced the Central Intelligence Agency that the quickest, most cost-effective way to develop working dictionaries for the project was to keypunch the corpus on which the entries were based in Germany. The logic was that there would be an easy supply of personnel with knowledge of the Russian language, or at least competence in the Cyrillic alphabet and that overall costs for such an operation would be significantly less than in the United States.

In July of that year I arrived in Frankfurt to establish a keypunch operation which was ultimately composed of 19 keypunch operators, 13 verify operators, 20 keypunch machines, and 10 verify machines, plus supervisory and support personnel. Start-up involved searching for suitable space, equipment, and staff. Fifteen years after the end of World War II and well into the *Wirtschqft Wunder,* everything was difficult. Space was hard to find and labor had to be mobilized. As I recall, no employed person could change jobs at that time without six months' notice, and then only on the quarter of the year—to prevent inflationary-job hopping. The Georgetown plan was to hire housewives seeking to join the labor force and learn a skill. We were very successful in recruiting a good staff, but they all had to be trained to keypunch, and some of them to type. Most had knowledge of the Cyrillic alphabet, but we learned that it did not seem to matter.

Dr. Peter Toma, also a participant on this panel of pioneers, took part in the selection of keypunch personnel and assisted in their training. He was also responsible for programming the dictionary abstraction and the conversion process. The keypunching of text began about six months after our July 1960 arrival in Frankfurt.

The dictionary material to be punched was initially divided into eight technical disciplines:

Organic Chemistry
Economics
Physical Chemistry
Crystallography
Celestial Physics
High Energy Physics
Surface Physics
Geophysics

There were further distinctions between "microglossaries" specific to the individual discipline and "metaglossaries" common to several disciplines. The Frankfurt Center also keypunched more than one half million words from *Pravda* to provide the basis for a "general glossary."

Objectives for the Frankfurt Keypunch Center were cited in a December 1960 report to Paul Howerton of the Central Intelligence Agency:

> ... to enable the center to send to Washington complete alphabetically listed (but uncoded) dictionaries on tapes and/or printed forms.
>
> ... to provide Washington with printed-out alphabetical word lists containing the individual text addresses for each lexical occurrence.
>
> ... to send out a machine printout of the texts used for dictionary preparation.
>
> . . . The alphabetical sorts, or word lists, will also be available without text addresses.
>
> . . . The corpus will be converted from card to tape to permit supplementary linguistic research, if required.

In the end, the Frankfurt Keypunch Center had consumed approximately three million IBM keypunch cards, attempted to keypunch one half million words in each of the eight disciplines noted above, converted those cards to magnetic tape, and flown the tapes to the Georgetown MT staff in Washington, D.C. At its peak, the Frankfurt Center had 40 people on staff. It ceased operation in the late fall of 1962.

THE EMERGENCE OF MT IN EUROPE

Loll Rolling
555, chemin des Landons
F-13510 Éguilles France
+33 (4) 42 92 42 70

The European Coal and Steel Community was created in 1952, and the European Atomic Energy Community and the European Economic Community (the Common Market) in 1958. They had six member countries with only four official languages (French, German, Italian, and Dutch). At first the volume of official documents was limited and the number of translators required was reasonable.

Before long, however, the European institutions—including the European Council, Parliament, Economic and Social Committee, Court of Justice, and Investment Bank—began to produce an increasing volume of texts that had to be translated into the four official languages. At that time, few people in Europe had heard about machine translation. It was still in its very early stages in the English-speaking world—principally the United States of America.

When U.S. efforts to develop machine translation finally came to be known in Europe, it also became clear that the ALPAC Report (1966) had led to the interruption of major projects at universities, that IBM had given up the whole idea, and that the Apollo-Soyuz project had covered only Russian and English and had been developed for strictly strategic reasons.

Academics in charge of language training at European universities had been very slow in getting interested in attempts to handle linguistic data by computer, and their departments considered the cost of the machinery involved extravagant. Commercial prospects were also unlikely: entrepreneurs in Europe did not see MT as a cost-effective venture. The only institutions likely to be interested were international organizations that needed speedy dissemination of their documents in several language versions.

The first to be active in this field was the French Textile Institute, which had an active cooperation arrangement with similar institutes in Germany, England, and Spain. Its TITUS system was successful because it was used for texts with limited phrase structure and terminology. It was not designed for application to different subjects or environments.

The first academic to become interested in MT was Bernard Vauquois in Grenoble, who happened to be intrigued with computers. In the late 1960s Vauquois made several attempts to feed linguistic rules into the computer. His example was followed by a number of academics in German, Dutch, and British universities. They had difficulty obtaining grants for acquisition of the costly equipment, the usefulness of which was seldom acknowledged by administrators.

Meanwhile, the situation became unbearable at the European Commission when three new members joined the Community and English was imposed as one of the major languages. The institutions were obliged to hire several hundred additional translators and became aware of the fact that alternative means had to be explored.

The Commission undertook an evaluation exercise which was initially intended to overcome the rejection of MT by the Community's large staff of university-trained translators. The first step was to define a methodology to comparatively assess several systems in order to establish their applicability. A second step was aimed at evaluating the long-term cost-effectiveness of an MT system within the Community's various institutions.

The two systems evaluated for use in the Commission's own services were TITUS and SYSTRAN, developed by Peter Toma and his company, World Translation Center, here in La Jolla. TITUS did not go to the end of the evaluation. That left SYSTRAN as the sole candidate. In 1976, at the request of the Commission, WTC developed and delivered an English-French prototype for the institutions of the European Community and the public services of its member states. The pilot project showed conclusive results, and the decision was made to proceed with further development of French-English plus other language couples.

At that time a number of European university linguists heavily criticized the fact that the European Commission was working with a system developed in the United States instead of relying on European linguistic knowledge which was so abundantly available. The Commission decided to give the European linguists a chance and supported the launching of the EUROTRA project, which was approved by the Council and Parliament in 1987. The cost of EUROTRA was shared between the Commission and the member states.

The ultimate outcome, of course, was that the pragmatic approach—namely, the adaptation of SYSTRAN to the Community's texts in all the European languages—was highly successful, while the EUROTRA approach, in spite of its higher cost, never even led to an experimental system that could have been evaluated for its effectiveness.

REFERENCE

[ALPAC] Automatic Language Processing Advisory Committee. 1966. *Language and Machines: Computers in Translation and Linguistics; A Report* . . . Washington, DC: National Academy of Sciences, Division of Behavioral Sciences. National Research Council Publication 1416.