GILBERT W. KING AND THE IBM-USAF TRANSLATOR

JOHN HUTCHINS

Gilbert King was a somewhat controversial figure in MT during the 1960s, for his adherence to an almost exclusively word-for-word dictionary-based approach, for the rejection of more theoretical approaches, and for the high-profile public demonstrations which, intentionally or not, gave the impression that the basic problems of MT had been solved.

Gilbert William King (born 13 January 1914, Long Eaton, Derbyshire, UK) was educated at MIT (B.S. 1933, Ph.D. 1935), held a number of research posts at various US universities (Caltech (1935-1937), Harvard (1937), Princeton (1937-1938), Yale (1938-1939), and MIT (1939-1941). Seconded during the war from Arthur D.Little Inc. to the Office of Scientific Research and Development for work in operational analysis, he rejoined MIT in 1946 (Research Laboratory of Electronics) before going in the early 1950s to the International Telemeter Corporation (Los Angeles) as chief engineer. Here, King and his colleagues (who included George W.Brown, later professor at Caltech, and the famous radar engineer, Louis N. Ridenour) developed a device for photographic storage and retrieval of information.

1 *The photoscopic store*

Research on the storage mechanism began in June 1953 with support from the US Navy, which was continued after 1954 by grants from the USAF through its Rome Air Development Center (King et al. 1953, King 1955). The system was not developed specifically for MT but as a device for rapid access to large quantities of stored information of any kind. Information was recorded photographically on a transparent disc as microscopic sequences of black and white rectangles arranged in concentric tracks. Information was read from the rotating disc by a light beam onto a phototube and converted into electrical signals (King 1955). The 16-inch disc could store some 30 million bits, about 300,000 coded words, on a four-inch ring (or 'annulus') on its outer edge. At 2,400 revolutions per minute, information could be read off at one million bits per second, i.e. matching of input words and dictionary words took less than a twentieth of a second (Shiner 1958, Macdonald 1960).[1] Input was from a

---

[1] In its basic principles, the photoscopic store may be seen as a precursor of the laser disc read-only memory device and of its smaller, now ubiquitous, version the CD-ROM.

manually operated typewriter or from punched cards. Output was fed to computer equipment (in later years, IBM 360 series) for 'logical operations'.

From the beginning it had been recognized by pioneers that the large dictionaries required for MT represented a major challenge. Most current memory devices in computers were too small. Gilbert King saw his 'photoscopic store' as the answer, and in May 1956, he received a grant from the United States Air Force for research on the disk for MT purposes.[2] At the same time, Erwin Reifler at the University of Washington received a grant from the same source for work on a Russian-English dictionary to be stored on the photoscopic disk. In 1958 King moved to the IBM Research Center at Yorktown Heights, N.Y., where he became director of research and continued research on the 'USAF Automatic Translator'.

2   *Probabilities, the lexicon and 'stuffing'*

The basic premise for King's approach to MT was that advantage should be taken of the redundancy of natural language, which aided readers to understand ill-formed and inaccurate texts, and of the power of local immediate context to resolve potentially ambiguous words. He remained convinced that a good comprehensive lexicon was more important than syntactic analysis, and that the readability and acceptability of translations could be greatly enhanced by good typographic layout and presentation.

He started from the assumption that "the success of the human in achieving a probability of .50 in anticipating the words in a sentence is largely due to his experience and the real meanings of the words already discovered" (King 1956). The assertion was probably based – King does not give his sources – on Shannon's findings with respect to data transmission, misapplied (as it often was at the time) to the transmission of meaningful information, and on the results of psychological tests where subjects were asked to guess blanked out words in texts. King did not, however, propose a statistics-based approach to MT, e.g. where calculation of context probabilities determined the selection of translation equivalents and the word order of target texts. Instead, frequencies of lexical items were used to determine fixed single equivalents for ambiguous words, 'cover words' that would be more often right than wrong, or for the establishment of rules based on immediate contexts.

Such context rules were in the form of 'logical operations', of essentially two kinds: the identification of the meaning of a word in a specific subject context, and the adjustment of the form of a translated word according to other words in the local context. Both operations could be achieved by

---

[2] At the time, King's device was seen as a 'special-purpose' computer – one specifically designed for translation. In fact it was not, since the storage equipment could be used for other purposes, and the computational operations were performed on general-purpose computers.

adding 'clues' to the dictionary. In the first case, the clues would point to a microglossary or to a subject field – King recognized similarities to the 'thesaural' method of the Cambridge Language Research Unit. In the second case, the clues would point to syntactic information. For example, in (a) *le livre est à lui*, and (b) *il est pour travailler*, the two instances of *est* would receive two different subscripts (e.g. 'b' and 'a' respectively, indicating translation as *belong* and *is* respectively). At the same time, *à* and *pour* would be assigned the corresponding subscripts (i.e. 'àb' and 'poura'), and 'àb' would translate as *to*, 'poura' as *about to*. Other subscripts would be used for translation in other contexts, e.g. 'àc' as *of*, 'pourd' as *for*, etc. The process of modifying lexical items and adding subscripts to dictionary entries was called 'stuffing'.

Although he considered some kind of use of 'parts of speech' to reduce options, he believed that the redundancy of any natural language text meant that precise analysis of syntax and semantics was not absolutely necessary. Some potential ambiguities could be avoided by entering whole phrases (e.g. *état gazeux* ("gaseous state") could exclude other potential translations of *état*), and his storage device had enough capacity. Translation consisted essentially of searching for the longest matching string of Russian characters and printing out the English 'equivalent' assigned to it – usually a single 'cover' word representing the most frequently occurring equivalent. Anything not matched during the dictionary search was left unchanged. Although initially King (1956) proposed that dictionaries would include full forms of every word (i.e. in the case of Russian complete paradigms of nouns and verbs), by the time of the first demonstration in 1959 there were separate entries for stems and endings.

One single equivalent ('cover' word) was obviously not always available, e.g. for homonyms such as Russian *mir* ("world" or "peace"), and the alternatives were printed out for the reader to interpret, which King believed to be quite acceptable – "The choice of multiple meaning like "dream/consider" (Fr. *songe*) is not of first importance, the ultimate reader can make his own choice easily" (King 1956) – although he conceded that, in some cases, giving all alternatives would "seriously detract from the understandability".

King insisted on practicality: "The program at IBM Research has been to examine the question of automatic translation of languages from an operational point of view rather than an interesting academic exercise." (King 1961). He rejected the aim of high-quality translation (certainly in the near term) in favour of producing output which is useful, and "it is an observed fact that useful amounts can be conveyed by rather primitive procedures."

The occurrence of wrongly selected words was not important: "... we have all read without difficulty a book in which many words occur whose meaning we do not know... One can read and understand text in which a considerable fraction of the words have been deleted." (King 1961). Moreover, readability could be improved simply by typographic improvements – e.g.,

margins, paragraphing, hyphenation, spacing – without which "even good translation has so little appeal as not to be read at all" (King 1961). When results from computers were printed line-by-line in badly printed uppercase characters, and often hardly legible, King was making a relevant point.

3   *Results*

The Mark I version of the IBM-USAF Translator, first demonstrated in July 1959, was no more than a word-for-word system (Shiner 1958). An example (reproduced in Masterman and Kay 1960) illustrates:

> Man – power nature" – thus is called one of divisions magazine. It contain article academician *V.A.Ambartsumyana* "Science about Universe and religion*ya*" and converse with/from doctor physicist*atematicheskikh* sciences)O. *V. Kukarkinym* about start second Soviet cosmic rocket on Moon. From these materials reader will see that since Mines*rnik* threw call churchby authorityby in explanation nature, astronomy became bed conquer one position for/after other, expeling)ga out of all sections material world/peace. Idea existence)ga, idea hundredththief world/peace suffered full disease. Every *shag* in development-studies about Universe all greater convince us in rightthose materialism and falsity religious world view.

The deficiencies of the translation are obvious: gaps in the dictionary (*shag*, "step, move"); missing case endings (-*ga*, "religion*ya*", "contained*kh*"), wrong divisions by the 'longest-match' approach (*sotvorenija* ("creation") translated as "hundredththief" from *sot* ("hundredth") and *vor* ("thief"), with the ending -*enija* ignored; *Kopernik* ("Copernicus") as *kopi* ("mines") and unidentified –*rnik*). Very few alternatives were offered ("with/from", "world/peace"), instead single equivalents were given, resulting in *poraženie* as "disease" rather than "defeat"; and phrases were absent, e.g. *brosil vyzov* ("challenged") was rendered "throw call".

Despite such obvious inadequacies, from June 1959, Mark I began to be used for translating the Russian newspaper *Pravda* – satisfactorily as King claimed later[3]. The next, solid-state transistorized, version (Mark II) was expected to "provide fast translations of the entire body of literature in any language, from fiction to complex scientific journals." A major refinement would be an optical print reader from the Baird-Atomic Corporation (also funded by USAF), intended to replace the costly manual transcription of Cyrillic texts onto punched cards for input (Macdonald 1960).

Some linguistic development was also envisaged (IBM Research 1962). Firstly, it was intended to make more use of morphological information to "supply intelligible and accurate English equivalents for all unambiguous Russian inflectional endings". Secondly, the system would identify more linkages between constituents (i.e. 'stuffings') and provide 'cement' words to

---

[3] In a letter (*Scientific American* 209, 1963, p.11), replying to Oettinger's comments on King and Chang (1963), he states that "translations of this quality… were found, in an operational evaluation, to be quite useful by the Government"

express them in English. Thirdly, it was intended to introduce some syntactic re-ordering of the output (although no method was proposed.) Finally, Mark I operated on a single pass through input text; it was intended that Mark II would operate as a 'multi-pass' system where each scan would perform different functions – but this change was not implemented.

The research was not restricted to Russian. A French-English version (with a dictionary of just 23,000 words) was demonstrated in May 1960, which produced translations of mathematical texts (Macdonald 1960). Research on Chinese began in 1960, with the particular problems of dealing with Chinese characters as the subject of intensive research on a special keyboard, the Sinowriter, developed with the Morgenthaler Linotype Company. As with the Russian system, sequences of characters were sought in the dictionary by the 'longest match' strategy. Entries included some grammatical and semantic information, since it was evident that a word for word treatment for Chinese would not work, and an attempt was made at some kind of phrase-structure analysis. Nevertheless, King and Chang (1963) claimed that "the general usefulness of the linguistic-dictionary approach, first used with Russian, has now been demonstrated with Chinese."

After King moved to the Itek Corporation in late 1962, further research was undertaken on the Chinese-English system at IBM and at Itek, both supported by the US Air Force (MIT 1964). However, by this date, King's own direct involvement in MT research and development was probably diminishing. He was heavily committed in other areas of data processing, such as the three-month study in May 1961 for the Library of Congress on large-scale storage and retrieval of information (King et al. 1962).

4   *Final judgment*

The Mark II system for Russian-English translation became fully operational at the USAF Foreign Technology Division (at the Wright-Patterson Air Force Base, Dayton, Ohio) in February 1964 and was in use until 1970 when it was replaced by the Systran Russian-English system. A public demonstration was given at the World's Fair in New York during 1964 and 1965 (Bowers and Fisk 1965). The quality of the translated examples (Hutchins 1986: 68) would have convinced many visitors that automatic translation was already a reality. There is, however, no evidence that the public were permitted to test the system with previously unseen Russian (i.e. random input).

Another example of its output was produced at the request of the Automatic Language Processing Committee set up in 1964 (see ALPAC 1966: 22, and Hutchins 1986: 69). The IBM project had received huge sums for its research – probably larger than for any other MT project (cf. U.S.House of Representatives 1960). This investment had resulted in the Mark II translator, so it was not surprising that ALPAC paid particular attention to this system.

The quality appeared to be no better than that produced by the Mark I some five years before. Equally damaging were the comparisons of translation costs: a study by Arthur D.Little Inc. showed that "FTD machine-aided translation" was far more expensive than the human translations produced for US government agencies by its Joint Publications Research Service. It was this kind of evidence that helped to convince the members of ALPAC that the money poured into MT research in the previous decade had not been productive. At IBM and Itek Corporation, research ended in 1966.

5   *Publications by G.W. King and others at IBM (in date order)*

King, Gilbert W., Brown, George W. and Ridenour, L.N. 1953. "Photographic techniques for information storage". *Proceedings of the I.R.E.* 41. 1421-1428.

King, Gilbert W. 1955. "A new approach to information storage". *Control Engineering* 2. 48-53.

King, Gilbert W. 1956. "Stochastic methods of mechanical translation". *Mechanical Translation* 3(2). 38-39.

King, Gilbert W. 1957. "The requirements of lexical storage". *Report of the Eighth Annual Round Table Meeting on Linguistics and Language Studies: Research in machine translation*, ed. Léon Dostert, Washington, 79-87. D.C.: Georgetown University Press.

IBM Research. 1959. *Final report on computer set AN/GSQ-16 (XW-1), vol.1: the photoscopic memory system.* Yorktown Heights, NY: IBM Research Center.

IBM Research. 1960. *Automatic translation.* Yorktown Heights, NY: IBM Research Center.

King, Gilbert W. 1961. "Functions required of a translation system". *Proceedings of the National Symposium on Machine Translation, held at the University of California, Los Angeles, February 2-5, 1960*, ed. H.P.Edmundson, 53-62. London: Prentice-Hall.

IBM Research. 1962. *Word analyzer*: final report. Yorktown Heights, NY: IBM Research Center. (Prepared for Rome Air Development Center, RADC-TDR-62-105)

King, Gilbert W. and Chang, H.W. 1963. "Machine translation of Chinese". *Scientific American* 208. 124-135.

6   *Other references*

ALPAC .1966. *Language and machines: computers in translation and linguistics.* A report by the Automatic Language Processing Advisory Committee, Division of Behavioral Sciences, National Academy of Sciences, National Research Council. Washington, D.C.: National Academy of Sciences—National Research Council.

Bowers, D.M. and Fisk, M.B. 1965. "The World's Fair machine translator". *Computer Design*, April 1965, 16-29.

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

Macdonald, Neil. 1960. "The photoscopic language translator". *Computers and Automation* 9(8). 6-8

Masterman, M. and Kay, M. 1960. Mechanical pidgin translation. Cambridge Language Research Unit. (ML 133).

MIT. 1964. Meeting on Chinese MT, Massachusetts Institute of Technology, October 17, 1964. [Unpublished transcript.]

Shiner, G. 1958. "The USAF automatic language translator, Mark I". *IRE National Convention Record*, part 4. 296-304.

United States House of Representatives. 1960. *Research on mechanical translation*. Report of the Committee on Science and Astronautics, U.S. House of Representatives, Eighty-sixth Congress, Second Session, House report 2021, June 28, 1960. Washington: U.S. Government Printing Office.