

## From the archives...

### The first MT patents

[John Hutchins]

Today machine translation means using a computer to translate natural languages. But it was not always so. The first suggestions that languages could be translated mechanically were made before electronic digital computers were even dreamt of. While the real history of MT began in March 1947 with correspondence between Warren Weaver and Norbert Wiener and with tentative discussions between Weaver and Andrew D. Booth (not in 1946 as Booth himself sometimes asserted later), it is nevertheless legitimate to recognise precursors in two patents submitted sixty years ago in 1933. One patent was issued in Paris on 22 July 1933 to Georges Artsrouni; the other was issued in Moscow on 5 September 1933 to Petr Petrovich Troyanskii. Both patents referred essentially to the construction of mechanical multilingual dictionaries.

Of course, the idea of translating 'mechanically' via dictionaries was itself not new - the essential ideas can be traced back to Descartes and Leibniz, who both proposed numerical codes to mediate between languages, and specific proposals for mechanical dictionaries were published from the middle of the 17th century onwards (e.g. Athanasius Kirchner 1663 and Johann Joachim Becher 1661) together with various proposals for universal languages (e.g. the famous Real Character of Bishop John Wilkins in 1668). A little later, in 1726, the idea of mechanizing language was widely popularized by Jonathan Swift in his *Gulliver's Travels*, where in the Academy of Lagado, Gulliver encountered scientists working on a machine for text generation, consisting of "bits of wood covered on every square with paper pasted on them, and on these papers were written all the words of their language, in their several moods, tenses, and declinations,..." It was not, however, until this century that specific devices were actually constructed for automating translation.

An account of Artsrouni's patent has been given by Michael Corbé ('La machine à traduire française aura bientôt trente ans' *Automatisme* 5 (3), 1960, 87-91). Artsrouni was a French engineer of Armenian origins and a former student of one of the major schools of St Petersburg. Apparently under development since 1929 the device, which Artsrouni called a *cerveau mécanique* (anticipating the popular name for the first electronic computers), was intended for a wide variety of tasks: the production of timetables and telephone books, for accounting, and for deciphering and encrypting messages. In 1937 the device was exhibited with great interest at the World Exhibition in Paris. A number of state organisations signed contracts with the inventor; the postal service, for example, ordered one to deal with money orders, and the railway service intended to use it for printing tickets. Only the war and the occupation of France prevented these plans coming to fruition.

The idea of using his invention for translation was prominent from the outset. In the 1933 description, Artsrouni explained that his 'brain' could "translate from one foreign language into any one of three other languages registered", and indeed was not limited to four languages or to a restricted number of words. In essence, the device consisted of four components: a memory (called *bande des réponses*), a keyboard for entering words (*tête de lecture*), a search mechanism (*sélecteur*), and an output mechanism (*sortie*). The core component was the 'memory', a paper tape 40 cm wide and up to 40 meters in length stored on two rollers and moved by cogs on lateral perforations (like a photographic film roll). Dictionary entries were recorded

in lines of four columns, one column for each language; up to 40,000 lines were possible, i.e. 40,000 lexical items for four languages. In fact, Artsrouni suggested that the number could be doubled by having entries in two colours (red and blue) on each line. The user was able to modify and add entries as necessary, since the order of items was quite free. The memory was searched by entering sought words at the keyboard which was linked to the 'selector' - another tape (paper or metal) on two rollers - containing codes for all entries in the memory. Output was displayed in a series of windows on the keyboard. Operation of the mechanical dictionary consisted, therefore, of moving the selector tape to match the input word and simultaneously moving the memory tape so that the translation could be read at the windows on the keyboard. In the earliest model it was claimed that the selector and memory tapes could move through the 40,000 lines in 60 seconds; in later models the speed was reduced to an average of 3 seconds.

Corbé compared Artsrouni's device to the machine constructed by Gilbert King for IBM, based on a device called the 'photoscopic store'. This too was essentially little more than a mechanical dictionary, although promoted as an MT system (and in fact used by the US Air Force from 1959 to 1964). In as much as his device was essentially the same idea, Corbé felt justified in claiming Artsrouni as a precursor of MT. Today we are rather less inclined to refer to automated dictionaries as translation systems.

In the case of Troyanskii, however, we do have a genuine precursor of machine translation. In the preface to a collection of his papers published in 1959 (*Perevodnaya mashina P.P. Troyanskogo*. Moskva: Akademiya Nauk SSSR, 1959) we read the following biographical details:

"Petr Petrovich Troyanskii was born in January 1894 in the family of a railway repair-shop worker in Orenburg. The family had 14 children and the living was hard. P.Troyanskii finished a parish school in Orenburg and passed gymnasia examinations without attending classes, after which he entered the University of St. Petersburg. He made his living by giving lessons. World War I prevented P.Troyanskii from finishing the university. After the Great October Revolution he entered the Institute of Red Professorate. Afterwards he taught social sciences and the history of science and technology at higher educational establishments. He also participated in compiling the Technical Encyclopedia and the Great Soviet Encyclopedia. In those years he devoted more and more time to putting into practice his idea of a translating machine. A serious illness - stenocardia - prevented P.Troyanskii from completing work on mechanizing translation which he considered the cause of his whole life. Petr Petrovich Troyanskii died on the May 24, 1950." [Translation by Evgenii Lovtskii]

Troyanskii's patent "for the selection and typing of words while translating from one language into another" consisted of a sloping table on which could be moved freely in all directions a broad band comprising a multilingual dictionary of entries arranged, like Artsrouni's, in columns.

In Troyanskii's case the entries were not full word forms but stems (e.g. infinitives of verbs). Troyanskii envisaged three stages. In a 'pre-editing' stage a user knowing only the source language identified stems and endings, and replaced the latter by pre-defined 'logical forms'. In the second, purely mechanical, stage the entries for source word-stems were located, the corresponding target words were photographed onto a tape and, at the same time, the 'logical forms' were typed out [see the attached translation of the patent for details]. In a 'post-editing' stage a user knowing only the target language provided the morphologically correct target forms.

For his logical forms, Troyanskii borrowed from Esperanto: nouns in the nominative were given endings in -o, plural forms in -j and oblique cases were indicated by -n; adjectives have the

ending –a, verbs in the present tense end in –as and infinitives in –i. The following extract from the 1959 collection illustrates translation from the French sentence *Le parti p rit, s'il commence   cacher ses erreurs* into Russian:

Le parti–o	partiya–o	Partiya
p�rir–as	pogibat'–as	pogibaet
si	esli	esli
il	on	ona
commencer–as	nachinat'–as	nachinaet
cacher–i	skryvat'–i	skryvat'
son–ajn	svoi–ajn	svoi
l'erreur–ojn	oshibka–ojn	oshibki

Various improvements were made to the device in later years, and by 1941 an experimental machine was operational, similar in a number of respects to the Harvard Mark I machine, developed between 1938 and 1942, which is regarded as the direct forerunner of ENIAC, the first computer. But support in the Soviet Union was not forthcoming, and his proposals remained unknown even in his own country until MT research had been underway for more than a decade.

Although his patent described only the operations of the mechanical dictionary, Troyanskii stressed in his writings the belief that all stages of translation could be automated. In this respect, he anticipated some central MT concepts, proposing the now familiar three-stage model of analysis, transfer and synthesis, and advocating the use of 'quasi-logical' interlingual elements. The linguistic details were not worked out, however; there was no discussion (or perhaps even awareness) of the problems of treating idiomatic expressions, homonyms or differences of word order. Nevertheless, there is little doubt that if the electronic computer had been available for the realization of his ideas Troyanskii would today be widely regarded as the true 'father' of machine translation.

-----  
**Troyanskii's patent**

[Translated by *Evgenii Lovtskii*]

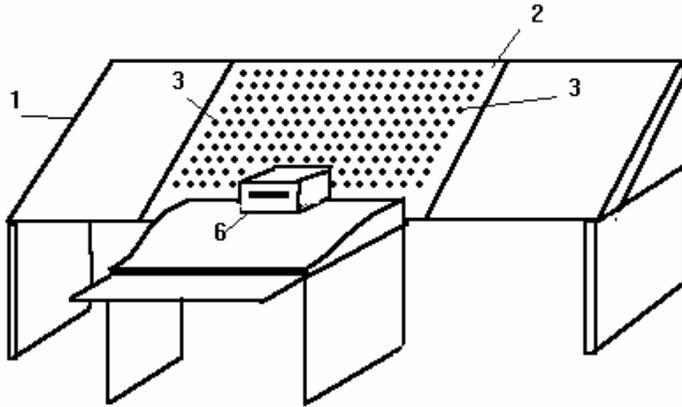
**INVENTION AUTHORSHIP CERTIFICATE**

Description of a machine for selecting and typing words when translating from one language into another or several others simultaneously.

To P.P.Troyanskii's claim of invention and certificate issued on September 5, 1933 (priority no. 134430)

The fact of granting an authorship certificate to P.P.Troyanskii was made public on January 31, 1935.

The machine in question is designed for selecting and typing words when translating from one language to another or several others simultaneously and essentially consists of a special perforated belt to which are affixed words in different languages. The belt can move over a desk; the perforations are used for positioning it in front of a photographic camera, adjacent to which is located a typewriter with additional keys for typing conventional signs alongside the photographed word. A view of the machine is given in the drawing.



As is shown in the drawing, the machine consists of a smooth sloping desk (1), over which a belt (2) can move easily and freely in different directions. The belt has perforations (3) for pins which position it against aperture (6). A six-language (or any other number of languages) parallel dictionary is arranged alphabetically on the belt's surface in such a way that words beginning with more frequent letters (e.g. K, M, P, etc.) are located closer to the middle.

The machine is operated in the following way. According to the word being translated, the belt is moved to a position where the word to be translated or a row of words in different languages find themselves against aperture (6), then the belt is stopped and the word is photographed on a light-sensitive film. At the same time conventional signs for logical parsing are typed on a paper tape, then the film and the paper tape are moved one line forward and the belt (2) is re-positioned for processing the next word or a row of words when translating into several languages, and so on.

From the translation produced in this manner and recorded on the film and the tape, fixed or glued together, with columns of photographed words and typed signs of logical parsing, a typist types a coherent text which goes to a special reviser who imparts to the words appropriate forms according to the conventional signs of logical parsing. After re-typing the text goes to a literary editor for final editing.

#### INVENTION SUBJECT MATTER

A machine for selecting and typing words when translating from one language into another or several others simultaneously, characterized by the use of a circular belt (2), with words in different languages pasted to it and with perforations (3) for positioning the appropriate word or words against an aperture in the desk above which a photographic camera is arranged for recording selected words on a light-sensitive film and, nearby, a typewriter with additional keys

for typing conventional signs on a paper tape alongside the photographed word.

Expert and editor A.G.Bremzen