[Published in: *Machine Translation*, vol.15 no. 3 (2000), p. 187-221]

Petr Petrovich Troyanskii (1894–1950): A forgotten pioneer of mechanical translation

JOHN HUTCHINS
89 Christchurch Road, Norwich NR2 3NG, England
WJHutchins@compuserve.com

and

EVGENII LOVTSKII
ul.Aviakonstruktura Milya 16-24, Moscow 109153, Russia
lov@online.ru

Abstract: This article presents translations of works by Troyanskii describing his 1933 proposals for a mechanical translation device, some later elaborations of the linguistic and technical features, and also translated extracts from commentaries made in 1959 by the Russian editors of his papers. The paper includes extensive assessments of Troyanskii's proposals, and places his work in the context of his time and in the light of subsequent research in machine translation.

Keywords: machine translation, history, interlingua, dictionary, Troyanskii

1. INTRODUCTION

The beginnings of research into the use of electronic computers for translating natural languages can be dated back to the first suggestions by Warren Weaver in 1947 and 1949 (Hutchins, 1997). For the first decade of the subsequent investigations in the United States and Britain, the pioneer efforts of the Russian inventor Petr Petrovich Troyanskii in the 1930s and 1940s on the mechanisation of translation were quite unknown.

Interest in the possibility of using the newly invented electronic computers for translating natural languages was aroused in the Soviet Union by reports of the demonstration in January 1954 of the small-scale pilot system developed by IBM and Georgetown University (Hutchins, 1997). In early 1955 the USSR Academy of Sciences established two research groups: at the Institute of Precision Mechanics and Computer Technology (*Институт точной механики и вычислительной техники АН СССР*) under Dmitri Yur'evich Panov, and at the Steklov Institute of Mathematics (*Математический Институт имени В. А. Стеклова АН СССР*) under Alexei A. Lyapunov. The first publications by members of these groups were accounts of the IBM-Georgetown experiment (Berkov and Ershov, 1955; Lyapunov and Kulagina, 1955).[1]

The first fruits of Soviet research on machine translation (MT) were reported in a book by Panov which appeared in early 1956, describing the initial experiments at his institute on the BESM computer in late 1955 (Panov, 1956). Then, at a meeting of the USSR Academy of Sciences on the 28th February 1956, Olga Kulagina and Igor Mel'chuk presented a paper on some of the earliest MT investigations for the Steklov Institute, namely the linguistic problems of translation from French into Russian. The report was published in the September–October issue of the journal *Вопросы Языкознания* (Kulagina and Mel'chuk,

1956). Since MT would be a completely new idea for its readers, the journal preceded their article by an introductory paper on the computational and linguistic problems by Kuznetsov et al. (1956), and followed it by a paper from Zhirkov (1956), which put the research into historical context.

Zhirkov recalled that in 1939 the Academy of Sciences had been approached by Petr Petrovich Troyanskii with proposals for mechanical translation and with an offer to discuss the possibilities with linguists. There had been somewhat fruitless discussions extending until 1944, after which contact with Troyanskii had been lost. Further information about Troyanskii then appeared shortly afterwards in a report compiled by D.Y. Panov, A.A. Lyapunov and I.S. Mukhin for a plenary session (15–20 October 1956) of the USSR Academy of Sciences on the automation of industrial production:[2]

> Apparently the idea of automating translation from one language into others arose originally as an idea for the "mechanisation of a labour-consuming process" – the process of dictionary lookup. In 1933 in the Soviet Union P. P. Troyanskii was issued an author's certificate for a mechanised dictionary. (Panov et al., 1956:13-14)

The writers reproduced the title header of Troyanskii's patent (i.e. his "author's certificate"), but they gave no further details. In 1957 the Academy of Sciences set up a committee of its Presidium to investigate Troyanskii's work; and in 1959 it published "The translating machine of P. P. Troyanskii: a collection of materials on a translating machine for translation from one language into others, proposed by P. P. Troyanskii in 1933." (Bel'skaya et al., 1959). The editors were all members of the MT research group at the Institute of Precision Mechanics and Computer Technology (*Институт точной механики и вычислительной техники АН СССР*) in Moscow.

This volume of Troyanskii's works begins with a reprint of the most extensive description of his proposal which he wrote sometime before 1947 (pp. 5–27). It is followed by a commentary from Izabella K. Bel'skaya (pp. 29–34) concentrating on the linguistic aspects of the proposal and showing how Troyanskii had anticipated the ideas of the early MT researchers in the West. Next the editors reprinted the description Troyanskii wrote in August–September 1933 to accompany his patent application (pp. 35–39), and then the original 1933 patent, with brief summary (pp. 39f). The final section was an extensive commentary by Panov and Korolev (pp. 41–51) on the technical aspects of the proposed system, which includes extracts from Troyanskii's manuscripts written between 1933 and 1947 and reproductions of drawings by the inventor. The publication demonstrated that Troyanskii's ideas were much more fully worked out than had been apparent from the brief details of earlier articles.

Outside Russia, Troyanskii's name was unknown until an abstract of Zhirkov's article appeared in *Mechanical Translation* (vol. 3, no.3 (1956), p. 91), and until the appearance of translations of other Russian publications. In 1958 came a French translation of the October 1956 report (Panov et al., 1958) and in 1960 an English translation of the second edition of Panov's book (Panov, 1960a). The first Russian edition of this book had made no mention of Troyanskii, but for the second expanded edition in 1958 Panov added some details of Troyanskii's 1933 patent:

> It seems that the first attempt, chronologically speaking, to mechanise translation to some extent was made in 1933 by P. P. Troyansky. He proposed the construction of a "machine for the selection and printing of words while translating from one language

> into another or into several others simultaneously." For this invention P. P. Troyansky received an author's certificate, but at the time he did not succeed in carrying his project through. This is understandable, since at that time automatic installations suitable for the purpose had not yet been created. (Panov, 1960a: 3).

In the same year, Panov included a more extensive outline of Troyanskii's proposal in a general review of USSR activity written in English (Panov, 1960b).

It was through these publications that researchers outside Russia became aware of Troyanskii's work and since the early 1960s he is mentioned regularly as an important MT pioneer in most historical accounts of the subject (e.g. Delavenay, 1959; Mounin, 1964; Zarechnak, 1979; Hutchins, 1986; Buchmann, 1987). However, the 1959 collection of Troyanskii papers was not translated into English, either in full or in part, and consequently the full significance of Troyanskii's achievement has not been appreciated outside Russia. Our aim in this article is to present Troyanskii in his own words with substantial translated extracts from this 1959 publication, and with some comments by ourselves. It is hoped that this article will be a belated fulfilment of the desire expressed by Bar-Hillel some forty years ago "that more should become known of this Babbage of MT" (Bar-Hillel, 1960:126.)

The exact nature of his invention is sometimes difficult to comprehend, partly because of subsequent unfamiliarity with the electromechanical equipment available in the 1930s, partly because Troyanskii was inventing a new device and he described it in unusual terms, and partly because he was not a linguist and his description of the linguistic processes involved was expressed in an idiosyncratic manner.

In this paper we provide translations from the 1959 collection in Section 2 (an excerpt from the preface giving biographical information on Troyanskii), Section 3 (the full patent), and Section 4 (a translation of the expanded description which Troyanskii attached to the patent application). Section 5 contains our own observations on the patent and Troyanskii's ideas in 1933, and recounts the fate of his proposals until the mid 1940s (including an extract from Zhirkov's 1956 paper.) Section 6 contains extended translations from Troyanskii's later thoughts on the non-technical aspects of his proposals, which were written some time before 1947. Section 7 is a commentary on these mainly linguistic aspects of Troyanskii's model, and it includes references to and quotations from the 1959 editorial commentary by Bel'skaya. In Section 8 we summarise Troyanskii's later technical developments as recounted by Panov and Korolev in their commentary. Finally, in Section 9 we add some concluding remarks.

In Sections 2, 3, 4 and 6, we have enclosed our own editorial comments, expansions and explanations in square brackets, in order to make clear what parts of the texts are translations from the 1959 collection. Troyanskii's style and presentation are often awkward in the original Russian, and he uses some invented terms and new usages for his ideas. We have, however, decided to adhere as closely as possible to Troyanskii's idiosyncratic manner and not to produce an idiomatic free-flowing rendition which may not correspond to his intentions. The major work of translation has been made by Evgenii Lovtskii.

2. BIOGRAPHICAL DETAILS

The foreword by the editors gives some brief biographical information.

> Petr Petrovich Troyanskii was born in January 1894 in the family of a railway repair-shop worker in Orenburg [Southern Urals]. 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 university. After the Great October Revolution [1917] he entered the Institute of Red Professors.[3] 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 the work on mechanising translation, which he considered the cause of his whole life. Petr Petrovich Troyanskii died on the 24th May 1950. (Bel'skaya et al., 1959:3)

3. THE PATENT

In 1933 Troyanskii applied for a patent on his invention. He was granted "author's certificate" (as patents were then called in the USSR) number 40995, with a priority date of 5th September 1933. The patent was granted not for a "translating machine" but was classified as a novel method of typesetting. This translation is from the original patent. The text included in Bel'skaya et al. (1959:39f) was an edited version, which omits phrases referring to multilingual generation.[4]

AUTHOR'S CERTIFICATE OF INVENTION

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

To the author's certificate of P. P. Troyanskii, declared on 5 September 1933 (priority no. 134430)

The granting of the author's certificate was made public on 31 January 1935.

The proposed machine is designed for selecting and typing words when translating from one language to another or several others simultaneously and essentially consists of a belt moving on a desk with words in different languages upon it and provided with perforations for positioning the belt 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 general view of the machine is given in the drawing.

As is shown in the drawing [Figure 1], the machine consists of a smooth sloping desk (1), over which moving easily and freely in different directions is a belt (2) provided with perforations (3) for pins which position the belt in front of an aperture (6). On the belt's surface is affixed or inscribed a six-language (or any other number [of languages]) parallel dictionary, in alphabetical order and laid out flat, arranged as columns of words in such a way that words with more frequent letters (e.g. K, M, P, etc.) are closer to the middle.

The operation of this machine is carried out in the following way. According to the word being translated, the belt is moved so that the corresponding word is positioned in front of aperture (6), then the belt is stopped and the catch of the photographic camera is released to produce a snapshot of the word on a light-

sensitive film held in the camera. At the same time conventional signs for logical parsing are typed on a paper tape, then the tapes of the photographic camera and the typewriter are moved one line forward and the belt (2) is moved again for processing the next word or row of words, when translating into several languages, etc.

From the translation produced in this manner on two joined or glued tapes, with columns of photographed words and conventional typed signs of logical parsing, a typist types a coherent text which goes first to a special reviser who gives the words the forms which are appropriate for the conventional signs of logical parsing, and then, after typing, to a literary editor for final editing.

Subject of the Invention

A machine for selecting and typing words when translating from one language into another or several others simultaneously, characterised by a belt (2) provided with columns, with words in different languages pasted on it and furnished with perforations (3) for positioning the required word or words against an aperture in the desk, above which a photographic camera is positioned for recording on a light-sensitive film the basic word with its corresponding row of words in foreign languages, and, nearby, a typewriter furnished with additional keys for typing on a paper tape conventional signs alongside the photographed word.

Expert and editor A.G.Bremzen

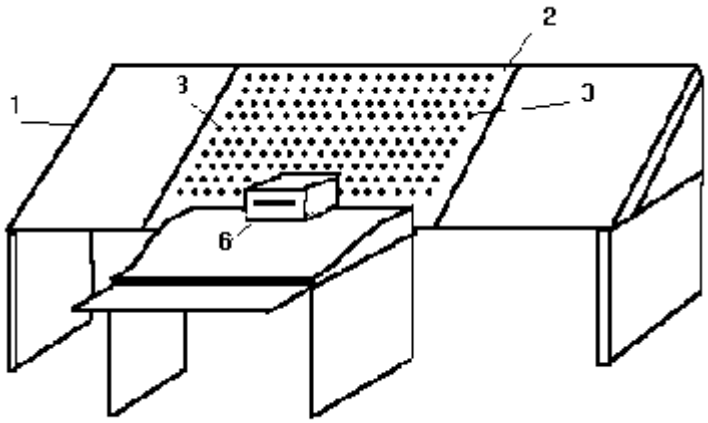

Figure 1. Drawing from patent application.

4. ACCOMPANYING EXPLICATION

Troyanskii supplied a more detailed explanation of his invention to accompany his patent application. The text reproduced in Bel'skaya et al. (1959:35–39) has been translated here in full.

P. Troyanskii
Machine for the automatic production
of ready typed translations requiring only literary editing
from one language simultaneously
into several other languages
29th July – 3rd September 1933
Moscow

§1. Supported at the ends in an inclined position (e.g. at 60%) is a ramp *1* constructed in such a way that belt *2* can roll around it and move in two directions: 1) straightforward to the left and to the right, and 2) vertically (around the ramp). The belt runs on ball and roller bearings. The belt has a flat surface with rows of perforations *3* like the holes in a telephone-dialling disk (for the finger).

§2. Thus any point of the belt's surface can be brought against lens *6* of photo-camera *5*.

§3. To the belt's surface there is glued in alphabetical order a, let us say, six-language (or, possibly, ten-, twenty-, etc. language) parallel dictionary laid out flat – with columns of words arranged in such a way that words beginning with frequently used letters (e.g. "K", "M", etc.) should be closer to the middle (following the principle of letter distribution on a typewriter keyboard or in a compositor's case). The dictionary itself for the machine must be specially prepared (see § 12 below). The belt's surface with the dictionary glued to it we shall call *glossary field*.

§4. On the desk, where in front of the glossary field stands the photographic camera, is also placed a typewriter which, in addition to two typefaces (say, Russian and Latin), has a number of special keys which we shall call *keys of logical and etymological parsing* (see their composition below in § 22).

§5. Through the typewriter and the photographic camera passes a band made up from two strips glued together: the left strip, say, with a photo-emulsive surface and the right one ordinary typewriter paper. The two apparatuses (photographic camera and typewriter) operate simultaneously, one line at a time, each on its medium. In addition, if the paper and emulsified parts of the band are glued together, the typewriter must be so constructed that the paper does not come out of the machine on the left, but the typewriter itself moves to the right (possibly, together with the operator's chair!). If it were constructed with the two parts glued together afterwards, then their movements would have to be synchronised – as compensation, there would be no need for a typewriter of special design.

§6. The work itself is done in the following way: moving the glossary field upwards and sideways using perforations on its surface (the field moves with the ease of a suspended bicycle wheel), the operator places the desired word before aperture *6* of the camera, and secures the glossary field in position with an electromagnetic brake by pressing the "stop" button. Then he releases the camera's shutter, makes a photograph of the line, and types symbols of logical parsing on the paper part of the band, after which by pressing a special key the band (both emulsive and paper parts) is advanced one line upwards, making room for the next inscription and snapshot.

§7. Then the glossary field belt is released and work continues in a similar way.

§8. If an *electrical drive* were used instead of the manual one, all the above operations could be made at the control of a button. On pressing the keys of a special keyboard for the *first* letter of the word being translated, the belt automatically feeds the necessary column in line with aperture *6* through the photographic camera. Pressing the *second* letter of the word being translated, the belt automatically positions the *first word* beginning with those two letters *against the aperture 6*. The third key positions at aperture *6* the first words beginning with that *combination of*

three letters, and so on. Then, pressing a special key makes the belt revolve slowly until the necessary line is selected, which is finally secured by the "stop" key for photographing and for logical marking. Before printing the logical signs and before photographing the line, it is essential to automatically display the selected line through a magnifying glass: for verification. Then a special key takes a photograph and moves the translation of the line upwards. The apparatus is ready to process the next word. For further developments of the machine, see §28–31 in this description.

§9. The dimensions of the glossary field surface are: for lines in 6 languages – 10 cm. long, and 1 cm. high for four lines – 20 sq.m. can accommodate 80,000 roots. With these dimensions, the surface of the *front* side of a belt will measure 2 by 4 meters.

§10. To show how great the number 80,000 is, let us consider the following examples: in all his works Leo Tolstoy used 12,000 roots; an educated person uses in conversation about 2,000 roots; the Russian language has 180,000 roots in all (including historical and obsolete); the English language has 200,000 roots (that have ever been used); special technical magazines use 3,000 roots.

§11. To facilitate searching for the desired line in the glossary field, in front of the aperture is mounted a powerful magnifying glass, which automatically moves aside when snapshots are taken.

§12. *How must the dictionary itself be prepared before bringing it to the glossary field?*

1) *Synonyms* are written out as far as possible in lines, for example:

speak	говорить, разговаривать	*parler, causer*	*sprechen*
swift	быстрый, скорый	*rapide, vite, soudain*	*schnell*

2) *Homonyms* are listed *with explications* of meaning in parentheses [and English translations added in square brackets], for example:[5]

коса (песчана)	[spit (of sand)]
коса (девичья)	[(maiden's) plait]
коса (для косьбы)	[(farmer's) scythe]
перевод (по службе)	[transfer (of duty)]
перевод (сочинени)	[translation (writing)]
перевод (денежный)	[remittance (monetary)]
перевод (стрелки)	[transfer (of switches)]
перевод (снимательной картинки)	[transfer (of decal)]

Note: Homonyms may occur only in the dictionary *from which* translation is made.

§13. *The order of processing of products obtained directly from the machine.*

The text emerges from the machine in the form of several columns; in which, let us assume, the one to the right is a column of logical parsing, while the rest (according to the number of languages) are columns of basic dictionary forms selected by the camera from the glossary field.

§14. The further processing of text output from the machine takes the following form:

1) What does the *typist* do with the text produced by the machine? She types columns relating to *each language* exactly as they have emerged from the machine,

and *beside each* column copies out the columns of logical parsing, observing of course the correspondence of the lines.

2) *What does the reviser do next with the text received?* He combines the two columns: the column of lexical choices and the column of symbols of logical parsing *into one connected text column.*

3) *What does the literary editor do* with the work of the reviser*?* The editor checks the correctness of the reviser's work, strikes out unnecessary synonyms, and gives a literary polish to the translation.

§15. *What essentially is translation* from one language into another? Into which elements can this process be broken down? It is:

1) selection of meanings for roots of declinable words and selection of meanings for indeclinable words

2) establishment of logical relations between words, i.e. the product of syntactic analysis, the establishment of syntactic links (subject, predicate, attribute, object, adverbs),

3) establishment of etymological[6] *mutual* subordination of *forms* between declinable and indeclinable words (of the sentence).

§16. The machine under discussion executes *the first* automatically, *the third* with the help of the operator, *the second* constitutes the operator's plan.[7]

§17. *What knowledge of languages is required by the operator (of the machine), the reviser and the editor?*

§18. The operator (of the machine) has to know only *one* language – from which [text] is being translated + the system of logical symbols + a small vocabulary of 200 to 300 "ancillary" words (in the international language Esperanto – in this respect very accurate and concise, based on scientific principles, but not natural.)

§19. The reviser[8] (who combines the column containing lexical choices and the column of logical symbols) needs to know only his own tongue – he leaves for the editor all the synonyms intact, giving them only grammatical forms (case, number, gender, tense, mood, voice, etc.) Besides this, he must know the same small vocabulary of 200 to 300 "ancillary" words (for which, by the way, will be taken as expedient all suffixes and prefixes from the grammar of Esperanto for precision of meaning).

§20. The editor must know both languages – the one from which translation is made, and the one into which he is working – to extract the meaning of the translation, to choose synonyms, to polish the unevennesses, i.e. to do general literary finishing.

§21. Thus, the most labour-consuming parts of the translation process – the work of the operator and of the reviser – do not require special knowledge of even two languages! This is especially important for translating from and into languages of minor nations of the Soviet Union.

§22. Regarding what is typed on the paper band in the column of logical parsing (with all possible use of words from the dictionary of "ancillary" word in Esperanto, namely from the dictionary of 200 to 300 words mentioned above).

These are:

1) all ciphers [numbers];

2) all designations of concrete numbers (e.g. cu.m., kVA, mm, km, etc.);
3) all pronouns;
4) all conjunctions (except rare ones);
5) all prepositions (except rare ones);
6) all proper nouns (names of cities, rivers, seas, etc., surnames, names of persons, etc.);
7) abbreviations and ciphered words (NKTP, SSSR, USA, LZ127, etc.);
8) the so-called international terms, such as *revolution, marxist, radio, aluminium*, etc. differing in various languages only orthographically;
9) the symbols of logical and etymological parsing; the meaning of these symbols is as follows:
   j – plural;
   n – dependence of forms of declinable parts of speech on verbs (verbal government); (direct object); (accusative case);
   de – 1) dependence of forms of declinable parts of speech form on another declinable form; 2) agent in passive voice;
   per – instrument in passive voice (with declinable parts of speech); (instrumental case);
   e – 1) adverb; 2) verbal adverb;
   oni – impersonal form of verb (in German and French, the corresponding form is: *man, on*.)
   a – 1) adjective in a predicate, expressing some kind of auxiliary verb; 2) participle;
   i – indefinite mood;
   as, is, os – the present, past and future tenses of verbs in indicative mood;
   us – conditional mood;
   u – imperative mood;
   pli ol – comparative degree;
   la plej el – superlative degree;
   on, obl, op, po – fractional numerals, multipliers, connectives, disjunctives;
   ant, int, ont – active voice participles: present, past and future tenses;
   at, it, ot – same in the passive voice;
10) all words and phrases which the operator finds it difficult to translate even descriptively, i.e. using synonyms; then he must put after them in brackets interrogative and exclamation signs (? !);
11) any questions in brackets concerning a particular word, with the aim of clarifying the logical role of the word in the sentence; the questions are posed in Esperanto, e.g. *kies, kie, kien, alkiu, kial, kiam, kiom da, kioma*, etc., i.e. whose, where, where to, to whom, why, when, how much, how many, etc.;
12) all punctuation marks (comma, full stop, hyphen, parentheses, inverted commas, etc.);
13) indeclinable words in most cases need no symbols of logical parsing;
14) there remain without parsing (i.e. with no signs of logical parsing) some forms of declinable words, namely: nominative case, singular, and indefinite mood.

Note: It is possible to make one reduction – in time of processing and length of tape produced from the machine – if a previous word has no logical parsing at all, and the next one need not be photographed, the lines can be linked by using the symbol "+".

This is the end of the text as reprinted in the 1959 collection. However, internal evidence indicates that Troyanskii's original contained a number of additional sections. There is, for example, the reference in §8 to descriptions of "further developments" in §28–31. There is also the following important extract, which Panov and Korolev included in their commentary

on technical developments (cf. Section 8 below). We do not know the contents of the intermediary sections omitted by the Russian editors.

> §56. ... *logical parsing itself can be automated* if we build a special machine which with completely accurate printed texts will do work that, as envisaged in the outline of technological processes for our staff member, is carried out in the name of "logic". It is even possible in the machine to contend with such subtleties as having one and the same word capable of being both a verb and a noun, judged from the point of view of its form. Then its particular syntactic position, role and relations with other parts of the sentence will indicate to the *mechanism* in what *sense* the given word is used, namely: in the noun sense or in the verb sense.

5. THE PATENT DOCUMENTS AND RECEPTION

The basic component of Troyanskii's proposal was a sloping table on which could be moved freely in all directions a broad belt or band comprising a multilingual dictionary of entries arranged in columns. Entries were to be not full word forms but stems (e.g., nominative forms of nouns and infinitives of verbs).

The translation process itself had three stages, which in present-day terminology can be described as follows. In a "pre-editing" stage a user knowing only the source language identified stems and endings, and replaced the latter by pre-defined "logical signs". 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 signs were typed out. In a "post-editing" stage a user knowing only the target language provided the morphologically correct target forms. In fact, Troyanskii suggested two post-editors (sections §14 and §20), one to align output words and logical forms and to synthesise target-language forms, the other to make correct and appropriate lexical and structural choices.

The origins of Troyanskii's ideas are not stated explicitly. We may speculate, however, that like many others he was inspired in part by the description of a somewhat similar machine in Swift's *Gulliver's Travels*.[9] The book is known to have been extremely popular in Russia at this time; and there was a film version made in the 1930s.[10] Another source would certainly have been his familiarity as a telecommunications engineer with the processes of encoding and decoding messages.

For his logical symbols, 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*. This is illustrated more clearly in extracts from a later expansion of his ideas (Section 6 below).

Ideals of internationalism are clearly discernible in Troyanskii's papers, in particular the desire to assist communication among those speaking different languages in the Soviet Union (Archaimbault and Léon, 1997). Esperanto was particularly popular in Central and Eastern Europe during the first decades of this century; it is even thought that Stalin once considered Esperanto as a future world language (Large, 1985). However, by the late 1930s, Esperantists were suspected of collaboration with the enemies of the Soviet Union and many were executed or imprisoned during the Stalinist terror. Judiciously, Troyanskii dropped the use of Esperanto symbols after the mid-1930s, but the original stimulus to create a device for the mass translation of scientific and technical documents remained strong.

The fate of Troyanskii's proposal was described by Zhirkov[11] in the paper already referred to (Zhirkov, 1956). Evidently disappointed that his invention had not been noticed by Soviet authorities, he approached the USSR Academy of Sciences in 1939 for assistance in development. But he did not get the reception he had no doubt been hoping for. (It will be noted that by this time he was calling himself Smirnov-Troyanskii. No explanation for this change is offered by the editors of the 1959 volume. It is clear, however, from the preface that Smirnov was his wife's family name.[12])

> The inventor–technician P. P. Smirnov-Troyanskii came to the institution of the USSR Academy of Sciences in 1939 and reported that he was working on a method of machine translation from one language into another; the inventor requested consultation about the linguistic aspects of his invention. It must be said that at this time the invention of P. P. Smirnov-Troyanskii was received by linguists with profound scepticism; it was considered impractical and quite unnecessary. Only a few admitted the possibility of machine translation. In the course of a succession of consultations, in which I participated, it became clear little by little that P. P. Smirnov-Troyanskii did not associate the method of machine translation he was working on with the ideas of electronic calculating machines, but this method created the possibility of translating, say, a Russian text in Moscow and its delivery in French translation, say, in Paris. And if there were the possibility of machine translation in languages of the peoples of the Soviet Union, then we might receive any document straightaway as, so to speak, a "circular" in several languages. The matter dragged on rather a long time and ended with a meeting of competent authorities on the 31st July 1944 at the Institute of Automation and Telemechanics of the USSR Academy of Sciences with the participation of linguists and specialists in the field of mechanics and electrical engineering. It has to be said that the specialists in mechanics and technology addressed the meeting mostly with arguments for the "impossibility" of machine translation, and, encroaching on the for them alien field of linguistics, talked about synonyms and subtle nuances of meaning; in short they talked about things that had no relation to their specialities. As a result, the experimental model of a translating machine (with a dictionary table of 1,000 words) was not constructed. Shortly afterwards the inventor P. P. Smirnov-Troyanskii, as far as I know, left Moscow; but now, according to my information, has already passed away. (Zhirkov, 1956:122)

Zhirkov was concerned mainly with showing the differences between Troyanskii's suggestions and those of Kulagina and Mel'chuk (1956), in particular with showing how the latter avoided the difficulties of post-editing. Zhirkov included an example translation of a Russian sentence into French, which was intended to illustrate the problems of interpreting Troyanskii's 'logical symbols' (now numerical, as noted above):

> ... it will be useful here ... to cite the sentence that was translated in the experiments in 1944. It was translated from the Russian sentence: *Решающие опыты механического перевода, которых мы ожидали в течение двух месяцев, осуществились в Москве сегодня в 4 ч. 30 м.* ['Decisive experiments in mechanical translation which we have been awaiting for two months took place in Moscow today at 4.30'].
>
> Those giving the text for translation to the machine marked the Russian text with coded symbols "решающие 51 опыты 1-5 механического 551-6 перевода 51-6..." etc. after each word in the whole text. After this, the "machine", i.e. its "model", set to work, simulating a human in these circumstances. The machine gave the resulting translation on a telegraphic tape: "expériment 1-5 décisif 51 traduction 51-6 mécanique 551-6 que 091 nous 01 avons 02-1 attendus 02-1* pendant 0902-1 deux

068 mois 05068 ont 2-1 eu 2-1* lieu 2-1** à 67 m-o-s-c-o-u- 67* aujourd'hui 68 à 67 quatre 68 heure 568 trente 68 minute 568 stop".

The editor received this translation from the machine and edited it by dictation directly to a typist: *Les expériments décisifs de la traduction mécanique que nous avons attendus pendant deux mois ont eu lieu à Moscou aujourd'hui à quatre heures trente minutes*. (Zhirkov 1956: 124)

Zhirkov considered this procedure unnecessarily tedious; by contrast, the fully automatic output promised on electronic computers was much more attractive.

After this negative reception at the Academy of Sciences, Troyanskii devoted the following years to answering his critics and developing further the technical features of his system. He expanded the linguistic aspects in the paper which is translated in the next section (Section 6), where he sought in particular to show how easy and time-saving his method could be in comparison with current human translation. His investigations on the technical side are recorded in the excerpts which appear in the commentary of Panov and Korolev (see Section 8 below).

It might be added that Zhirkov's remark about Troyanskii's failure to take into account developments in "electronic calculating machines" should be treated with caution. The use of electronic equipment in calculating machines did not come until the mid-1940s with the building of the ENIAC machine in the United States. It is most unlikely that anyone in the Soviet Union knew of these developments, as they were kept secret until the end of the war. What Troyanskii did know about were experiments in electromechanical devices, and it is on these lines that his own investigations were focussed.

6. FURTHER REFLECTIONS BY TROYANSKII ON LINGUISTIC ASPECTS

In February 1947 Troyanskii expanded his ideas in the following paper (Bel'skaya et al., 1959:5–27). Because of its length and repetitiveness, we have omitted some sections.

P. P. Smirnov-Troyanskii
On a translation machine
built on the basis of monolingual
language-translation methodology

1. Logical parsing

In working out the new translation methodology I proceeded from the universal logical make-up in all languages: a subject is everywhere a subject, a predicate is a predicate, an object is an object, etc. This universality of logical make-up brings languages together despite the diversity of their structures, grammars, and lexicons, and that makes the differences surmountable.

It is on the basis of this universality that I created the so-called form of logical parsing common for all languages, as a text form intermediate in the translation process.

The operation for obtaining the form of logical parsing consists of converting natural national texts, for example text A, into a form of logical analysis, for example text A′, all words are taken in their initial [base] grammatical forms, i.e. in the nominative singular for all declinable [words] and in the infinitive for all conjugated [words], while invariable [words], i.e. non-conjugated or non-declined words (members) of phrases, such as prepositions, conjunctions, adverbs, and interjections, retain their sole inherent grammatical form. The work of interpreting such separate

words, i.e. the work of combining groups of words into logical complexes, into coherent sentences, is achieved by the symbols of logical parsing which accompany all declined or conjugated words in phrases, and sometimes invariable parts of sentences (namely when they play the role of the principal members of sentences, e.g. in *Hurrah rang out from afar*, where *hurrah* is the subject, and in a number of similar rare cases.)

There are about 25 universal international symbols of logical parsing for all languages, used in various combinations which number about one hundred, however they are capable of rendering without exception all relations and the slightest shades of human thought expressed in words and notions, and ensure absolutely exact translation into other languages without distortion of meaning.

Symbols of logical parsing, on the one hand, show the logico-syntactic meaning (sense role) of the word in a sentence; and on the other hand, they release the phrase from a rigid word order, thanks to potentialities inherent in them, which is very convenient for the new translation technology.

When performing logical parsing, I factor out, as it were, what is alike from one language to another, and leave in what distinguishes one from the other. The converted text is the text in the form of logical parsing. The machine deals with the converted text in this form. In fact, what is factored out (namely, features identical and common to all languages – representations of meaningful logical links between words of a phrase, between parts of the sentence – links which are identical and common in all languages) is, indeed, shown and expressed by symbols of logical parsing, while what is left (what distinguishes one language from another) is the lexicon of languages, taken in initial national-grammatical form[13] (i.e. the nominative case singular for declinable words, the indefinite mood for those which conjugate, and the sole inherent form for unalterable parts of the sentence). In special instructions when performing logical parsing it is possible to find many simple and convenient rules for all cases of translation practice....

Any mechanisation of work processes introduces into these processes its own regularities. In particular, it imposes certain requirements on the material used. When using a machine of any kind (it is important to understand that this is a general characteristic of all machines without exception), the material to be processed must first be adapted and reduced always to a form suitable for processing on this particular kind of machine (whether in size, structural-chemical properties, or any other technical characteristics or conditions). In the event of a wide technological disparity between the machine and the material the machine may fail to work.

Such preliminary treatment of the material to be processed on the translating machine is exactly what the operation of logical parsing is. Logical parsing is namely an integral part of machine translation technology. Like any other normal machine, the translating machine requires people wishing to obtain normal operational results to observe the rules of this technology, and, specifically, to introduce into the processing not any kind of raw material but only special material previously adapted for the process – and a text in the form of logical parsing is exactly the material adapted for processing on the translating machine....

2. The technological process of translation and its three stages

... I divided the process of translation and the process of editing between two persons who knew either the language of the original or the translation language, i.e. between persons who did not know the two languages simultaneously.

Next, using the accepted system of logical parsing symbols, I divided the translation process into three separate operations.

During the first monolingual replacement operation A–A′, the original, initial text A for translation is replaced with text A′, in the same language but in the logical parsing form. This first operation A–A′ is performed by the first translator who knows only one (his native) language, in this case the language of the original.

During the second bilingual replacement operation A′–B′, the text A′ in the logical parsing form of the original language is replaced with text B′, also in the logical parsing form, but in the translation language. This operation A′–B′ is performed by the machine. Here the need is completely absent for any logical treatment whatever of the text. Here, purely mechanically, are the initial forms of one language replaced with initial forms of another language, while the logical parsing symbols are transferred – in correspondence and also mechanically – to the analogical forms of the other language; as a result, the substitution of texts in logical parsing form takes place: A′–B′.

The third substitution operation B′–B, which is monolingual, is, like the first one, performed by a second translator, whose native language is the translation language: like the first translator, he does not know any other languages. The second translator replaces the logical parsing form B′ produced by the machine with the natural, i.e. fully national-grammatical, form B in his native language, the translation language....

... [B]oth my monolingual process of translation, and my monolingual operation of editing are performed normally and produce the same effect as work carried out by means of the old bilingual methodology where each translator must know two languages – the language of the original and the translation language.

But now the cost of translation decreases tremendously, the field of translation work thanks to its general accessibility widens enormously, and the translation process itself – on the basis of the new monolingual methodology – is mechanised and automated....

The complete process of translation from one language into another can thus be represented by the following four-member sequence of successive replacements of forms: A–A′–B′–B or, in the reverse mode of the translation process, in the sequence: B–B′–A′–A.

However, the machine can give more.

The first translator can translate in the machine texts from the initial language not only into any one, but straight away into several foreign languages, performing his work in a series of successive stages: firstly, he carries out the first operation A–A′, then – using the machine – the second operation A′–(B′+C′+...+K′). This second operation of multilingual translation can be performed on the machine by a second translator if he has at hand the text in the form A′ prepared by the first translator, or by a perfect stranger not knowing either the original or the translation languages, who for this needs to have only the text A′. During the second operation of multilingual translation the machine itself performs the necessary substitution of texts: it replaces a text A′ in the initial grammar forms in the original language with texts in translation languages, also in initial [grammatical] forms, and automatically provides them with logical parsing symbols in accordance with form A′.

Finally, by employing sets of two translators (their number equals B+C+...K=T), the third operation of multilingual translation can be carried out, namely the operation: (B′+C′+...+K′) - (B+C+...+K). The translated text will appear after this third operation in several languages, and in fully national-grammatical forms.

3. On the simplicity of performing the operation of logical parsing

The execution by a monolingual translator of the logical parsing operation (A–A′), as well as the opposite process, namely converting a logical parsing form into a fully national-grammatical form (i.e. performing the operation A′–A), presents no difficulty at all. It resembles an exercise for primary-school children in their native language, who have to analyse texts "by parts of speech" and "by parts of sentences". The difference is that the results of analysis in the school exercise are written down in a more cumbersome way than in the logical analysis intended for translation operations.

Here is an example of school analysis. The sentence is *Птичка летает* ['The bird flies']. The analysis: *птичка* ['bird'] is a noun, feminine gender, nominative case, singular, answering to the question: who = the subject. *Летает* ['flies'] is a verb, present tense, singular, third person, indicative mood, answering to the question: what does it do = predicate.

The results of logical parsing for the same example *Птичка летает* are in my case shorter and simpler, they discard the redundant, but denote the same thing, namely: "*птичка*–о, *летать*–а", and that is all for the operation A–A′. If it is acceptable for school children to carry out the far more complicated school analysis above "by parts of speech" and "by parts of sentences", then the notation of my logical parsing is surely considerably simpler, especially for a grown-up person literate in his own language. The same applies to the opposite operation, i.e. to the operation A′–A.

4. Some new features

Gradually, in the course of working out the idea, my monolingual translation methodology began to display and accumulate new features, proving the correctness of the idea developed: false ideas are incapable of development.

1. It all began with the possibility – already described above – of making adequate translations by two persons not knowing two languages simultaneously from language to language in the full cycle A–A′–B′–B...

2. Later it became evident that it was sufficient to publish texts only in the logical parsing form, no matter what the method for obtaining the texts: [whether] as a result of the first translator's work after the stage A–A′ (when, let us assume, symbols of logical parsing are provided by editing, as footnotes,[14] to the texts of scientific books and periodicals, [i.e.] publishing in this way national texts together with their logical symbols[15] [and] with the aim of making them fit for translating into other languages); or as a result of the machine operation A′–B′ (or B′–A′), e.g. for using as telegraphic texts on communication lines, where a text from another country is sent in the language of the sender-country and received by the addressee in the language of the receiver-country, etc.

3. Correspondence becomes possible between people speaking different tongues....

4. With the availability of publication, in journals and books, of national texts accompanied by their logical parsing forms, there is the possibility of translating foreign journal articles and books into one's own language without knowing the language of the original...

5. The possibility has arisen of translating simultaneously into several languages....

6. A colossal reduction of translation processing costs has become evident, reducing these costs to one percent of the former level. It is obvious that the higher the number of languages involved in machine translation, the cheaper translation will cost into each separate language.

7. There is the possibility for each writer, scientist or journalist to translate into foreign languages any of his writings, even into several foreign languages simultaneously, as a result of using the process A–A′–B′ or the process A–A′–

(B′+C′+...+K′), and publishing texts in logical parsing form, i.e. giving the second translators 90% of finished translation and leaving them only the process (B′+C′+...+K′)–(B+C+...+K), i.e. just the remaining 10% of work – namely, only the remaining operation of bringing texts in logical parsing form to the fully national-grammatical form.

8. The quality of editing work rises at all three stages, i.e. the stages A–A′, A′–B′ and B′–B; it is work performed now only by monolingual revisers.

In case of doubt about the correctness of the rendition of some passage … , every monolingual reviser (who in the majority of cases may also be a monolingual translator) would be able to eliminate an obscurity in the translation by [contacting] his translation partner, using the methods of monolingual communication which were described above, i.e. by means of the process A–A′–B′ and the process B–B′–A …. As to the fidelity and high quality of the machine stage A′–B′ during which there is no human interference in the work of the machine, it is guaranteed – from the point of view of editing – by the high quality of work done by linguists on the glossary field for the machine, where … the lexica … are created by specialists in applied linguistics, but [where] the method for introducing the necessary changes … is as simple as changing a safety fuse in the access box of an apartment's electrical lighting installation.

9. By enormously easing the availability of translation work in general, everyone literate in his own language can be a translator.

Important note. My monolingual translation methodology does not necessarily get rid of special education for understanding specialist texts in one's own native language. Strange as it may seem, this situation was taken as a fault in my monolingual translation methodology and my translating machine.

But this, you know, is a general feature of any speciality: it has to be learnt. It is not got rid of in the bilingual translation methodology – but nobody can blame it for that. I, for one, do not.

10. Finally, what has become possible and tested in practice – in automatically functioning models – is the mechanisation and automation of translation work and the transmission of translated texts by communication lines.

For the technical design of the translating machine and other operational features in detail see the separate technical report.[16]

Paragraph 5 compares his method and traditional translation, and demonstrates its savings and advantages.

6. The most important thing in the comparison

... [L]et the old bilingual translators translate without machine into languages they know! But there exist languages they do not know. Faced with languages they do not know the old translators stop dead as if before a stone wall, here is an impasse for them: with unknown languages they can do nothing.

Now there is a way out – the one I suggest: adoption of the monolingual methodology, for by adopting the monolingual methodology and becoming simultaneously monolingual translators they break out of the impasse.

Besides, 99% of the literate population of the world know only their native language or do not know foreign languages enough to act as translators. For them the monolingual methodology opens up the possibility of using widely foreign languages, those languages they do not know.

This does not prevent them from learning foreign languages so that some of them manage without the machine… the study of foreign languages will continue.

But there remains mass translation work which has to be handed over to the machine. Firstly because we simply cannot keep up with this mass, with its ever-growing volume, and because who would want to persist [with human translation] when the machine translates all this massive volume of texts into several languages simultaneously and issues them straight away in printed form. The translators' and revisers' parts will be purely literary work on style and polishing, while the mass of laborious rough translation work is taken away from them.

There is nothing to be obstinate about. There exist craftsmen who, without machines, using only hand tools, are capable of making precision watch mechanisms. But they would be ridiculed if they denied the necessity of using machines in watch making.

Seeds can be sown by hand, and sown not badly. But there exist tractor-drawn sowing machines.

It must also be pointed out that the first models of machines never turn out to be perfect; on the contrary, design perfection, expansion of output, acceleration and improvement of the technical process are always reached gradually. It will suffice to recall the first makes of telephones, typewriters, radio receivers, aircraft, locomotives and many other machines and instruments. The same will happen to the translation machine. To its aid will come the natural development of linguistics and technology.

There will, of course, always be some kinds of work which in reality are difficult for machines to do. For example, surgical operations on a living human organism can hardly be entrusted to a machine. But translation work does not belong to such operations. Rather it resembles calculation work which has already been entrusted to machines. In the field of calculating operations there has long been no dispute with machines, although manual work is here still wide spread. There is no use disputing with the translating machine. It will prevail.

Paragraph 7 details further the processes involved in his approach.

8. The richness of the monolingual methodology

The correctness of the initial idea underlying the monolingual translation methodology, namely, employment of logical parsing forms as constituent parts of the technological process of translation, is corroborated by the presence of a whole series of variants and combinations of translation formulae reflecting alterations, modifications of the processes themselves, and also their interrelations and interactions one with another....

In due course, when the skill of utilising a text in the logical parsing form has spread among people, i.e. they have fully available to them now and immediately the skill of carrying out in their own language and without any difficulty the operation A–A′ and the reverse operation A′–A, and also the skill of understanding form A′ without preliminary processing, i.e. without being obliged to bring it into form A – which is accessible to everybody even now (this can be tested: one can practice on the examples given in the text of this work) – then the wide application may come, on the one hand, of incomplete cycles of translation (from the processing view) that coincide wholly with the second (purely mechanical) two-member translation operation, … A′–B′, where the product is a text in form B′ and the initial text is a text in form A′ (or the reverse process, … B′–A′) – these processes can be performed alone on a machine by a person who knows neither the language of the original nor the translation language.…

9. On using synonyms, idioms, and homonyms

Whole groups of synonyms are furnished by the machine itself from its glossary field and they are printed out, so that the editor has only to strike out superfluous synonyms.

Idioms ... are either replaced by non-idiomatic expressions, or they are left untranslated and thus appear in the foreign language with no changes, or (finally) a suitable idiom is selected in the language into which the translation is made. There are naturally no other ways. In every translation process the translator working in the language of the original marks particularly those places where idioms occur.

As to homonyms, provision is made for their correct usage, the impossibility of confusing one for another, and the understanding of their different senses, by the fact that in the manipulation field of the machine the homonyms are not dumped in one heap but are given indicators of sense determinants in parentheses, e.g. in the following way:

Manipulation field	glossary field	[English][17]
освободить (дать свободу)	*freigeben*	'liberate'
освободить (отпустить)	*freilassen*	'let go'
освободить (избавить)	*erlösen*	'redeem'
освободить (от тяжести)	*entlasten*	'ease, relieve'
освободить (от слова, от обещания)	*entbinden*	'release'
освободить (очистить)	*räumen*	'clear'
освободить (место и тому подоб)	*freimachen*	'vacate'
коса (девичья)	*der Zopf*	'plait'
коса (для косьбы)	*die Sense*	'scythe'
коса (песчаная)	*die Nehrung*	'spit of land'

Automatic dictionary or translation machine? The machine proposed by me, based on the use of the new monolingual translation methodology, can rightly be called a translation machine, since of the three operations comprising its technological processes, namely the operations A–A′, A′–B′ and B′–B, only the second one (A′–B′) represents the translation function proper, as it embraces different languages and only it formulates translation itself from language to language, whereas the first (A–A′) and the third (B′–B) operations are pre-translation and post-translation procedures, and neither separately nor together do either represent the translation function proper: they are both within the limits of any one language, and they are both intralingual and non-mechanical operations, while the second is a bilingual A′–B′ or multilingual A′–(B′+C′+...+K′) operation, i.e. the translation operations proper are performed directly by the machine; and the text produced by the machine, with a little practice, does not need to be brought to the national grammar form and can be used as it comes out of the machine. Therefore my machine is a true translation machine.

Discontented negators of the idea of monolingual translation are annoyed… by the fact that, in order to be translated from an unknown language, a text must first be provided with symbols of logical parsing, i.e. to have the form A′(or B′, C′... K′ in other languages). It seems they do not want even that.

But any meaningful text in any language is always provided with logical parsing symbols in the form of a national system of symbols representing a national form of logical parsing. In my method of translation I require only the substitution of this national system … by another system, a unified system, my system, which is simpler than any national system. And nothing more.

In Paragraph 10 Troyanskii tabulates his development programme: in the first stage an operational model was completed, by May 1941; the second stage (stated to be "under way")

was to be the assembly of a prototype device; and in the third stage the theoretical investigations were to be completed for a practicable design of a "powerful automatic device on the basis of modern communication technology".

11. *Examples illustrating translation processes by applying the principles of the monolingual methodology*.

First is given an explanation of the symbols of logical parsing used in the examples presented below (the symbols can also be represented as numbers):

-as: denotes a predicate expressed by a verb in the present tense indicative mood,
-i: denotes the indefinite mood of verbs,
-o: denotes a subject expressed by a noun in the nominative case singular,
-de: is a preposition denoting the genitive case of all declinable words,
-ojn: denotes a direct object expressed by a noun in the accusative case plural without preposition (o – noun, j – plural, n – accusative).

Example 1. Translating from French into Russian[18]

A	A′	B′	B	
Le parti	le parti-o	партия-о	*Партия*	'the part'
périt	périr-as	погибать-as	*погибает*	'perishes'
s'	si	если	*если*	'if'
il	il	он	*он*	'it'
commence	commencer-as	начинать-as	*начинает*	'begins'
à cacher	cacher-i	скрывать-i	*скрывать*	'to conceal'
ses	son-ajn	свой-ajn	*свои*	'its'
erreurs	l'erreur-ojn	ошибка-ojn	*ошибки*	'errors'

Example 2. Translating from German into Russian and French

A	A′	B′ + C′		B + C		[English]
Das Bild	das Bild-o	картина-о	le tableau-o	*Картина*	*Le tableau*	'the map'
der Welt	die Welt-de-o	мир-de-o	le monde-de-o	*мира*	*du monde*	'of the world'
zeigt,	zeigen-as	показывать-as	montrer-as	*показывает*	*montre,*	'shows'
wie	wie	как	comment	*как*	*comment*	'how'
die Materie	die Materie-o	материя-о	la matière-o	*материя*	*la matière*	'matter'
sich bewegt	sich bewegen-as	двигаться-as	se mouvoir-as	*движеться*	*se meut*	'moves'
wie	wie	как	comment	*как*	*comment*	'as'
die Materie	die Materie-o	материя-о	la matière-o	*материя*	*la matière*	'matter'
denkt.	denken-as	мыслить-as	penser-as	*мыслит*	*pense.*	'thinks'

Troyanskii gives six more tables illustrating the translation of these sentences in other directions. Paragraph 12 gives further detailed explanations and tables showing the translation stages; and Paragraph 13 describes the processes using an algebraic notation.

14. How to show logical parsing of national text in footnotes, i.e. how to combine national text with logical parsing without distortion[19]

National text: Le[1] parti[1] périt[2] s'[3] il commence[4] à[5] cacher[5] ses[6] erreurs[7]. – Le[1] tableau[1] du[8] monde[8] montre[4] comment la[1] matière[1] se[9] meut[9], comment la[1] matière[1] pense[4].

Footnotes: [1]-o, [2]-ir-as, [3] si, [4]-er-as, [5]-er-i, [6]son-ajn, [7]-r-ojn, [8]de-o, [9]se mouvoir-as

Note 1. Since grammar forms in a text are repeated and one and the same footnote can refer to the same case, the number of footnotes to a page cannot be large. See for instance, footnote 1.

Note 2. There are three possible cases of using logical parsing symbols in footnotes:

Case 1. A word form in the national text is slightly different from the initial grammatical form. Then the national ending of the initial grammatical form (1–2 letters) is given before the logical parsing symbol. See footnotes 2, 4, 5 and 7.

Case 2. In the text is used the initial grammatical form of the national word. Then it suffices to show only a logical parsing symbol in the footnote. See series of no. 1 footnotes.

Case 3. A national text form of words is altogether different from its initial grammatical form. Then the initial grammatical form is given in the footnote in full before the logical parsing symbol.

Note 3. Newest dictionaries give in their general alphabet all disparate forms for words of one and the same root or sense, when the disparity is manifested by changes of the word in number, comparative degree, etc., as well as including all distinctions among orthographic and grammatical forms of irregular verbs — which makes it possible to find in these dictionaries for, say, *se meut* ['moves'] the verb in its initial form *se mouvoir* ['to move']. This treatment, in a still more advanced form will find broad application in the machine; this will reduce the number and volume of footnotes.... In this way, synonyms and idioms can also be brought into line [i.e. as footnotes].

7. OBSERVATIONS ON LINGUISTIC ASPECTS

In her commentary on the linguistic features of Troyanskii's proposals (Bel'skaya et al., 1959: 29–34), Izabella K. Bel'skaya states that the paper in Section 6 on "logical parsing" was written by Troyanskii in February 1947. In her view, it did not introduce any substantial changes or additions to Troyanskii's original conception of automatic translation in 1933. It was an elaboration written in part to justify his approach in response to criticism he had received – as described by Zhirkov above (Section 5). Even though it is a later paper it was still written before any actual research on mechanical translation had began in the West, i.e. the early work of Andrew Booth and Richard Richens in 1948.

Bel'skaya comments on three areas where Troyanskii may be justly said to have anticipated the ideas of later MT researchers: universal logical structure of languages, pre- and post-editing, and writing in controlled language.

In his first paragraph (Section 6 §1), Troyanskii states that "... I proceeded from the universal logical make up in all languages.... I created the so-called form of logical parsing common for all languages, as a text form intermediate in the translation process". The idea of a single logical structure common to all languages as the basis for automatic translation was one expressed by Warren Weaver in his now well-known memorandum of July 1949 (Weaver, 1949), the primary stimulus of MT research in the West. The basis for Weaver's optimism was the cryptographic achievements of the Second World War, the work of Hans Reichenbach and other logicians, the development of Claude Shannon's "information theory" and the analogies made by Warren McCulloch and Walter Pitts between the human nervous system and the "logical" properties of universal Turing machines (Hutchins, 1997). Weaver also believed in the existence of "the common base of human communication – the real but as yet undiscovered universal language".

Bel'skaya cites the similar views of Erwin Reifler:

> A correlation of different languages is only possible if they share certain aspects. All languages actually do have a number of features in common.... Agreements in features which concern the logical aspect of language are especially numerous. (Reifler, 1955:140)

Reifler was a Sinologist at the University of Washington (Seattle), whose views on universal semantic features were known to Weaver, and who was one of the first scholars to investigate MT (Reifler, 1950). However, Reifler stressed, as did Weaver, that logic could not be the whole answer, and Bel'skaya herself makes this same point in her commentary.

By contrast, Troyanskii appears to have had no such doubts, probably because of a belief in the universality of the Esperanto categories. It is, in fact, to be noted that in her commentary Bel'skaya does not draw attention to the Esperanto source of Troyanskii's logical symbols, perhaps because of the official antagonism towards Esperanto at the time in Soviet Russia.

One most striking anticipation by Troyanskii is the use of numerical indices for designating his logical symbols. As Bel'skaya comments, the same idea is found in the proposals by Luitgard and Alex Wundheiler for a universal logical basis for syntax. The Wundheilers, whose names are now almost forgotten, despite appearance in Locke and Booth's (1955) anthology, described their system as follows:

> If the verb of a sentence has the index *n*, the complements of the verb will have the indices *n1, n2, n3,...*, 1, 2, 3, ... being role indices. The assignment of indices to roles is, of course, arbitrary, and must be codified in a dictionary.
>
> "John$_{11}$ gave$_{1}$ a book$_{13}$ to Mary$_{12}$."
>
> "Mary$_{12}$ was given$_{1}$ a book$_{13}$ by John$_{11}$."
>
> Complements that denote the same participant with the same role in synonymous sentences are assigned the same role index. (Wundheiler and Wundheiler, 1955:199)

According to Panov and Korolev, Troyanskii made the following remark at the end of one of his papers (perhaps the 1933 one, Section 4 above). Referring to the logical parsing process, he remarks that with this feature his device "… is in the genuine and literal sense a 'logical machine' in the terminology of the philosophers of the past...". It would appear that he was aware of writings by philosophers such as Leibniz on universal languages and logical "machines" (i.e. languages designed to reduce misunderstandings among people and to promote scientific communication and universal peace). In any case, the view that all languages share a common logical foundation and that they differ only (or primarily) in their lexica was evidently widespread in Russia in the first decades of 20th century (Archaimbault and Léon, 1997). The works of 20th century philosophers, such as Rudolf Carnap and other logicians of the Viennese school – who argued for the universality of logical functions and relations – would have probably been unknown to Troyanskii, or indeed to other Soviet researchers at the time. In fact, it was not until Yehoshua Bar-Hillel introduced the idea of an "operational grammar" for MT that their significance was brought to the attention of researchers in the field (Bar-Hillel, 1951; Hutchins, 1997).

That Troyanskii's approach was strictly word for word is obvious from his examples. There is no hint in his writings that he appreciated the difficulties of comprehension that this would give rise to. However, he was not alone in this opinion among MT pioneers: e.g. Richens and Booth asserted in 1948 that simple word-for-word translation which includes

some indications of role or case relations could be "understood" if the reader knows the original language to some extent. An example (cited by Bel'skaya) is an English "translation" from German (Richens and Booth, 1955:37):

> if in a/one *d* large (more) area two form *m* beside one another live *z* without self to/too mix *z*, so belong/hear *p z z* different *m* form *m* circle *m* at.

(The German has been translated word for word from base forms, with letters indicating inflection categories: *d* indicates a German word in the dative case, *p* a past verb form, *m* "multiple, plural or dual", *z* "unspecific", and slashes alternative translations.)

When Troyanskii suggested that the process of "logical parsing" could also be mechanised, he may well have been envisaging something similar to the approach of Richens and Booth – the automatic identification of (grammatical) case endings and of canonical (root) forms of verbs and nouns. The Richens–Booth experiment used punched card equipment.

While Troyanskii assumed word-by-word translation would be comprehensible, he recognised fully the necessity for human involvement both before and after the mechanised part of his translation process. He stressed at length that his approach was based on the work of "monolingual editors", who did not need to know the language from which or into which the translation was made, and expounded the advantages of such an approach over traditional "bilingual translation" in great detail. Obviously, Troyanskii was conscious of the great burden he would be imposing on the monolingual editors and was at pains to demonstrate how much, nevertheless, was being saved in terms of effort and language knowledge.

Most of the early MT researchers saw the need for pre- and post-editing – the terms were coined by Reifler in his very first MT study of January 1950 (Reifler, 1950). In Reifler's conception (1952a), the aim of pre-editing was "a graphic supplementation of the conventional form of the foreign message which raises its graphic-semantic explicitness to the level necessary for a mechanical translation". Troyanskii's logical symbols had a similar function.

Both Reifler and Troyanskii realised that the pre-editor would have to disambiguate homonyms in the source language to reflect in part the usage of the target language. Troyanskii proposed a table of homonyms with contextual clues to assist selection (Section 4 § 9 above). Reifler suggested that the process could be mechanised:

> When the pre-editor dials the conventional graphic form of the foreign message into the translation mechanism, it would first pass through the mechanical dictionary. Whenever in terms of the target language no multiple meanings are involved, the dictionary mechanism would not intervene and the dialled material would move on to the next stage in the translation process. Otherwise a device would call the attention of the pre-editor to the fact that multiple meanings are involved and the dictionary entry concerned would appear on a screen. The pre-editor would then select the meaning required by the context and dial the distinctive graphic symbol representative of this meaning and supplied by the dictionary entry. (Reifler. 1952a:7)

Troyanskii does not appear to have thought of this possibility. Nor did he suggest, as Bar-Hillel did, that the pre-editor should "deal with the elimination of morphological and syntactical ambiguities and with rearrangement of the FL [i.e. source] text in accordance with

a standard order in the TL [target language] following a set of instructions available to him in his own language." (Bar-Hillel, 1951: 230)

In this respect, it is quite clear that Troyanskii thought that the recipients would be able to do all of this. He would, however, have agreed with Bar-Hillel (1951) that most of the burden for semantic disambiguation and target-language rearrangement should be placed in the hands of the post-editor. Bar-Hillel did not think that a pre-editor could possibly anticipate all possible interpretations of source texts. He argued also that a post-editor would have to be familiar with the source language. For example, a post-editor would have to recognise that *he (she, it) gives...* could be a "literal" translation of the German phrase *es gibt...* and that it should therefore be changed to *there is (are)....* Otherwise, no reader ignorant of German could hope to make any sense of the output (Bar-Hillel, 1952).

An alternative suggested by Reifler (1952b) was that the system should produce "pidgin" translations by the establishment of what he called "pseudo-universals" between languages. He noted, for example, that the Mandarin Chinese character *-ti* could be equated with English *-ing*, so that Chinese $t'a^{1}$ tso^{3} *-uti* $k'uai^{4}$ could appear in a literal translation as *he walk-ing quick* – which, as Reifler says, is "bad English but perfectly intelligible and, because it permits a word-to-word translation, has the great advantage of simplifying the mechanical correlation problem".

Reifler linked his proposal closely to ideas for simplifying natural languages. Stuart Dodd (1952) had put forward detailed proposals for a "Model English" which could be used in MT. Bel'skaya notes that many MT researchers saw normalisation as necessary for pre-editing original texts, and she rightly related these ideas to Troyanskii's earlier suggestions. In part 8 (Section 6) of his paper he expected that in time people would become familiar enough with his symbols to read the "logical parsing form" of texts. In part, he would have expected that use of his method might encourage the replacement of language-specific logical symbols by a unified system – i.e. as part of contemporary Soviet ideas for the normalisation and internationalisation of language for the new Communist society (Archaimbault and Léon, 1997). It might also be seen, from today's perspective, as an anticipation of proposals for normalising and regularising natural languages for the purposes of MT, i.e. the use of "controlled languages".

Troyanskii stressed the benefits of simultaneous multilingual translation (from one source into many target languages) years ahead of his time – in fact multilingual output did not come until the 1970s, initially with the application by Xerox of the Systran system using controlled input (Elliston, 1979). He also suggested that texts could be transmitted via cables in their logically parsed forms (in Section 6 §4, he refers to this as "telegraphic" language), and that each sender and recipient would have one of his translating machines to convert texts into their own languages. It is an anticipation of the basic concept of the project at Utrecht in the 1980s for "distributed language translation" (Witkam, 1983), in which texts would be transmitted over a network of microcomputers in an intermediary language – interestingly, a form of Esperanto was developed for this purpose – and simultaneous translation at users' terminals would be from and into their own languages only.

For Troyanskii, what was perhaps most important was the *monolingual* aspect of his method, the fact that translation and communication could be achieved without knowledge of source or target languages. He stressed the economic argument – which he no doubt thought would carry great weight – the possibility of using non-qualified translators, the publication of texts in his logical parsing form, the rapid exchange of scientific knowledge, and in particular the simultaneous translation of official documents into all languages of the Soviet

Union (Archaimbault and Léon, 1997). Similar arguments were used in the early days of MT in the West: automatic translation could reduce the need for highly qualified and expensive human translators; people might learn to write in a simplified and regularised language (designed for easy MT) and to interpret unrevised MT as a kind of "pidgin" language (Reifler, 1952b).

The basic practical economic arguments used by Troyanskii (Section 6 §6) to justify efforts in mechanising translation remain valid to this day: the demand for "mass-production translation work" which can only be tackled by automation of some kind, the growing volume of translation which cannot be handled in any other way, and the demand for simultaneous translation into several languages. At the same time, he stresses that there will always be a need for translation by expert "craftsmen"; there is a place for both translating machines and human translators – a message that was sometimes forgotten during the heyday of enthusiasm for MT in the 1960s.

Finally (Section 4 §56 above) he claimed that it would be possible to go beyond simply the mechanisation of a translation dictionary. He was again ahead of others. Clear ideas of how syntactic analysis might be done did not appear until the 1950s with the preliminary investigations by Oswald and Fletcher (1951) and with the proposals for an "operational grammar" by Bar-Hillel (1951). There is no clue as to how Troyanskii thought automatic analysis might be accomplished, although he does say that verbs and nouns could be identified by their endings, and it is probable that he would have argued that since logical parsing is easier than the tasks set for schoolchildren (see above, Section 6 §3) then it would be easy to formulate appropriate mechanical operations. Probably he would have underestimated the complexity of parsing, and in this he would, of course, have been mistaken – although no more so than his successors in the 1950s and 1960s, who also believed that syntactic processing was relatively straightforward.

As for synthesis, although Troyanskii did not suggest that this could be automated, it is clear that he saw the main part of post-editing as not demanding special skills. The task of the "reviser" (Section 4 §14) was simply to combine information from two columns, i.e. to synthesise morphological forms. He separated this task from the more demanding role of the "literary editor" (Section 4 §20) which involved semantic and syntactic choices in the target language.

Troyanskii was undoubtedly more interested in the mechanics of his proposal than the linguistic details of the translation processes themselves. In his papers there was no discussion (or perhaps even awareness) of the problems of treating idiomatic expressions, homonyms or differences of word order. Indeed Troyanskii would appear to believe that it would be a simple or trivial matter for any "editor" to compose fluent, grammatically and stylistically idiomatic texts from strings of words and logical symbols.

The basic linguistic framework of Troyanskii's proposed translating machine was not modified in any radical fashion after its formulation in the patent documents (as shown in Section 6 above). Troyanskii was not a linguist, and his ideas could not be subject to any practical test. The criticisms he received were theoretical, as Zhirkov made clear (Section 5); and he failed to get the assistance he sought in this respect.

8. TECHNICAL ASPECTS

As a telecommunications engineer, Troyanskii was familiar with rapid developments in electronics and after 1933 his main concern was to improve the technical aspects of his machine. Hence, as Panov and Korolev demonstrate in their commentary (Bel'skaya et al., 1959:41–51), Troyanskii continued to work on further developments of the technical aspects of his invention until shortly before his death in 1950.

The original purely mechanical method was found to be more difficult to implement than expected, and Troyanskii worked on various modifications involving the use of electronic components. In a description dating from 1939 he suggested a new form of the "glossary field" which used a steel tape instead of the paper tape of the patent, and where data was recorded via a teleprinter. The whole system was to operate through electrical sensing of dots (rather than using the mechanical movement over perforations), and reading took place via some kind of photoelement. A further innovation was the provision of separate keyboards for each of the languages from which translations were to be made. Strangely, however, he proposed that the keyboards should not consist of individual (alphabetical) characters but that each key should represent one of the words on the glossary field, i.e. there was to be direct individual retrieval of each line of the dictionary. Panov and Korolev noted that "[t]he large dimensions of the keyboard, which by the author's estimate must accommodate several thousands of keys, would create certain difficulties in building such a machine and operating it"; but they thought that "those difficulties are not insurmountable … similar problems are overcome in Chinese typesetting or, for example, in Chinese typewriters".

Apparently, no machine was built by Troyanskii on this new design either, although it appears that a working model was constructed with a small number of words in the glossary field. (This may have been the system demonstrated in 1941 to the Academy of Sciences – Section 5 above.) The outbreak of war in 1941 prevented Troyanskii from continuing his work, and it was only in 1948 that he could begin again on his invention. Now he suggested a machine built with widely available off-the-shelf communication components. His conception was a distributed system, connected by electric cables, where it would be possible "for several operators to work simultaneously: each will translate a separate text from a different language than that of his neighbour, nobody will interfere with anyone else". There would be a single glossary field and a common input and output apparatus. The glossary field itself would consist of belt-driven electromagnetic relays. As before, translation was a five-stage process. Now, however, he could envisage a more automated process where the operator did little more than just press a button to start the whole process.

Panov and Korolev argued that the design was similar to the automatic calculating machine built at Harvard towards the end of the war – this was the Mark-I built by Howard Aiken and which is regarded as a forerunner of the first electronic computer ENIAC. Like Troyanskii's proposed machine, this made use of gears and wheels operated mechanically by electrical motors under control of a perforated tape; the similarities are apparent from contemporary descriptions of Mark I (e.g. in Randell, 1975). The primary difference is, of course, that Troyanskii was intending to build a non-numerical symbol-processing machine.

At the same time, in about 1948, Troyanskii's ambitions were going even further. He began work on a "a portable translation machine for personal use". He proposed a folding screen on which segments of the glossary field and logical parsing symbols would be projected from a device consisting of a microfilmed version of the dictionary. At the same time, the elements selected (i.e. base forms of words plus logical symbols) would be encoded for transmission to a receiving device for display at another location.

From the technical point of view, Troyanskii anticipated in an interesting respect (as Panov and Korolev point out) the "photoscopic store" which Gilbert King developed in the 1950s (King et al., 1953), and which was to be used in the dictionary-driven MT system under development at the University of Washington in Seattle. This device was later used successfully in the IBM Mark-I system installed at the USAF Foreign Technology Division at the Wright Patterson Air Force Base in 1958 (Hutchins, 1986:62–70). However, it should be stressed that the photoscopic disk operated on somewhat different principles to the device proposed by Troyanskii. Dictionary information was stored photographically on a glass disk in binary coding; information was read by shining a light through the disk, rotating at 20 revolutions a second, and converting the resulting alternations of light and dark to electric signals by the use of a photocell, and these signals were then processed by computer. Troyanskii's dictionary information was not coded in binary form, and it was not stored on a revolving disk. However, the device did involve the use of photocells and conversion into electric signals for transmission.

9. CONCLUSION

Panov and Korolev conclude their commentary as follows: "P. P. Smirnov-Troyanskii closely approached a solution to the problem of mechanising translation even if in a simpler form than is done now.... There is no doubt that if [he] had been aware of the potentials of electronic calculating machines, his method of solving the problem of mechanical translation would have been different."

Troyanskii died in 1950. Electronic computers were not developed in Russia until 1953 with the BESM machine at the Institute for Precision Mechanics and Computer Technology and the STRELA machine at the Steklov Mathematical Institute. Research on MT did not begin in the Soviet Union until after the 1954 demonstration by IBM and Georgetown University of a "pilot" system (Hutchins, 1997). There is little doubt that, if they had been known to the earliest Russian researchers, Troyanskii's ideas would have been among the first to be tested on the new electronic computers and that Troyanskii would today be ranked alongside Weaver as an acknowledged "father" of MT.

References

Archaimbault, Sylvie and Jacqueline Léon: 1997, 'La langue intermédiaire dans la traduction automatique en URSS (1954–1960), filiations et modèles' [Intermediate language in machine translation in the USSR (1954–1960), relations and models]. *Histoire Epistemologie Langage* **19**(2), 105–132.

Bar-Hillel, Yehoshua: 1951, 'The present state of research on mechanical translation'. *American Documentation* **2**(4), 229–237.

Bar-Hillel, Yehoshua: 1952, The treatment of 'idioms' by a translating machine. [Presented at MIT conference, see Hutchins (1997)]

Bar-Hillel, Yehoshua: 1960, 'The present status of automatic translation of languages'. *Advances in Computers* **1**, 91–163.

Bel'skaya, I. K. [И. К. Бельская], L. N. Korolev [Л. Н. Королев] and D. Yu. Panov [Д.Ю. Панов] (eds): 1959, *Переводная машина П. П. Троянского: сборник материалов о переводной машине для перевода с одного языка на другие, предложенной П. П. Троянским в 1933 г.* [The translating machine of P. P. Troyanskii: collection of materials on a translating machine for the translation from one language into others, proposed by P. P. Troyanskii in 1933]. Изд. Акад. Наук СССР, Москва.

Berkov, V. P. [В. П. Берков] and B. A. Ershov [Б. А. Ершов]: 1955, 'О попытках машинного перевода' [On attempts at machine translation]. *Вопросы Языкознания* **4**(6), 145–148.

Buchmann, B.: 1987, 'Early history of machine translation'. In Margaret King (ed.), *Machine Translation Today: The state of the art*, Edinburgh University Press, Edinburgh, pp. 3–21.

Carnap, Rudolf: 1934, *Die logische Syntax der Sprache*. Springer, Wien. Published in English as *The Logical Syntax of Language*, Kegan Paul, London, 1937.

Delavenay, Émile: 1959, *La machine à traduire* [The translating machine]. Presses Universitaires de France, Paris. Published in English as *An Introduction to Machine Translation*, Thames and Hudson, London, 1960.

Dodd, Stuart C.: 1952, 'Model English for mechanical translation: an example of a national language regularized for electronic translators'. 7pp. Repr. as 'Model English' in Locke and Booth (1955), pp. 165–173.

Elliston, John S. G.: 1979, 'Computer aided translation: a business viewpoint'. In Barbara M. Snell (ed.) *Translating and the Computer: Proceedings of a seminar, London, 14th November 1978,* North-Holland, Amsterdam, pp. 149–158.

Hutchins, W. John: 1986, *Machine Translation: Past, present, future.* Ellis Horwood, Chichester.

Hutchins, W. John: 1997, 'From first thoughts to first demonstration: the nascent years of machine translation, 1947–1954. A chronology'. *Machine Translation* **12**, 195–252.

King, Gilbert W., G. W. Brown and Louis N. Ridenour: 1953, 'Photographic techniques for information storage'. *Proceedings of the I.R.E.* **41**, 1421–1428.

Kulagina, O. S. [О. С. Кулагина] and I. A. Mel'chuk [И. А. Мельчук]: 1956, 'Машинный перевод с французского языка на русский' [Machine translation from French into Russian]. *Вопросы Языкознания* **5**(5), 111–121.

Kuznetsov, P. S. [П. С. Кузнецов], A. A. Lyapunov [А. А. Ляпунов] and A. A. Reformatskii [А.А. Реформатский]: 1956, 'Основые проблемы машинного перевода: вводная статья' [Basic problems of machine translation: introductory article], *Вопросы Языкознания* **5**(5), 107–111.

Large, Andrew: 1985, *The Artificial Language Movement.* Blackwell, Oxford.

Leyda, Jay: 1983, *Kino. A history of the Russian and Soviet film*, 3rd edition. Allen and Unwin, London.

Locke, Willam N. and A. Donald Booth (eds): 1955, *Machine Translation of Languages: Fourteen essays*. M.I.T.Press, Cambridge, Mass.

Lyapunov, A. A. [А. А. Ляпунов] and O. S. Kulagina [О.С. Кулагина]: 1955, 'Использование вычислительных машин для перевода с одного языка на другой' [Utilization of computing machines for translation from one language to another]. *Природа* **8**, 83–85.

Mel'chuk, I. A. [И. А. Мельчук] and R. D. Ravich [Р. Д. Равич]: 1967, *Автоматический перевод 1949–1963: Критико-библиографический справочник* [Automatic translation 1949–1963: a critical and bibliographical handbook]. Всесоюзный Институт Научной и Технической Информации, Институт Языкознания, Москва.

Mounin, Georges: 1964, *La machine à traduire : histoire des problèmes linguistiques* [The translating machine: history of linguistic problems]. Mouton, The Hague.

Oswald, Victor A. and Stuart L. Fletcher; 1951, 'Proposals for the mechanical resolution of German syntax patterns'. *Modern Language Forum* **36**(3–4), 1–24.

Panov, D.Yu. [Д. Ю. Панов]: 1956, *Автоматический перевод* [Automatic translation]. Изд. АН СССР, Москва. (Научно-Популярная Серия)

Panov, D.Yu. [Д. Ю. Панов]: 1958, *Автоматический перевод. Изд. творое, переработанное и дополненное* [Automatic translation. 2nd ed., revised and amplified]. Изд. АН СССР, Москва. (Научно-Популярная Серия)

Panov, D.Yu.: 1960a, *Automatic translation.* Translated from the Russian by R. Kisch. Edited by A. J. Mitchell. Pergamon Press, London. [Translation of Panov (1958)]

Panov, D.Yu.: 1960b, 'Machine translation and the human being'. *Impact of Science on Society* **10**, 16–25.

Panov, D.Yu. [Д. Ю. Панов], A. A. Lyapunov [А. А. Ляпунов] and I. S. Mukhin [И. С. Мухин]: 1956, *Автоматизация перевода с одного языка на другой* [Automation of translation from one language to another]. Изд. АН СССР, Москва.

Panov, D., A. Lapounov and I. Moukhine: 1958, 'La traduction automatique' [Automatic translation]. *Recherches Internationales à la Lumière du Marxisme* **7**, 162–193.

Randell, Brian (ed.): 1975, *The Origins of Digital Computers: Selected papers*. 2nd ed. Springer Verlag, Berlin.

Reifler, Erwin: 1950, Studies in mechanical translation, no. 1. ms, 51pp. University of Washington, Seattle.

Reifler, Erwin: 1952a, 'MT with a pre-editor and writing for MT'. 16pp. [Presented at MIT conference, see Hutchins (1997)]

Reifler, Erwin: 1952b, 'General MT and universal grammar'. 6pp. [Presented at MIT conference, see Hutchins (1997).]

Reifler, Erwin: 1955, 'The mechanical determination of meaning'. In Locke and Booth (1955), pp. 136–164.

Richens, R. H. and A. D. Booth: 1955, 'Some methods of mechanized translation'. In Locke and Booth (1955), pp. 24–46.

Weaver, Warren: 1949, 'Translation'. Repr. in Locke and Booth (1955), 15–23.

Witkam, A. P. M.: 1983, *Distributed Language Translation: Feasibility study of a multilingual facility for videotex information networks*. BSO, Utrecht.

Wundheiler, Luitgard and Alex Wundheiler; 1955, 'Some logical concepts for syntax'. In Locke and Booth (1955), 194–207.

Zarechnak, Michael: 1979, 'The history of machine translation'. In Bozena Henisz-Dostert, R. Ross Macdonald and Michael Zarechnak, *Machine Translation*, Mouton, The Hague, pp. 20–28. (Trends in Linguistics, Studies and Monographs 11)

Zhirkov, L. I. [Л. И. Жирков]: 1956, 'Границы применимости машинного перевода' [The limits of applicability of machine translation]. *Вопросы Языкознания* **5**(5), 121–124.

Notes

[1] Details of early Russian writings on machine translation are to be found in the bibliography by Mel'chuk and Ravich (1967).

[2] All quotes from non-English sources are translated by the present authors unless otherwise stated.

[3] The institute (*Институт Красной Профессуры*) was established in 1921 by Lenin with the task of preparing qualified individuals for teaching in higher education in the fields of economics, philosophy and history. It was purged and disbanded by Stalin in 1937. Troyanskii's membership would suggest that in the 1920s and early 1930s he was a committed Communist.

[4] The missing words are: (a) the phrase "or several others simultaneously" from the title, the first paragraph and the "Subject of the invention"; (b) the phrase "or any other number" from the second paragraph; (c) the phrase "or a row of words, when translating into several languages" from the third paragraph; and (d) the phrase "with its corresponding row of words in foreign languages" from the "Subject of the invention". The Russian editors do not mention these omissions or offer any reasons for them.

[5] This paragraph was misplaced in the 1959 reprint, coming erroneously in the following section §13.

[6] Troyanskii uses the term "etymological" to mean "semantic".

[7] The meaning of the Russian phrase (*план оператора*) is vague; therefore, a literal translation is given.

[8] Troyanskii uses the word *корректор* 'corrector'. For the 'literary editor' (§20) he uses *редактор*.

[9] In the third part (*A Voyage to Laputa*), there is the following description of a machine for writing:

> He then led me to the Frame, about the Sides whereof all his Pupils stood in Ranks. It was Twenty Foot square, placed in the Middle of the Room. The Superficies was composed of several Bits of Wood, about the Bigness of a Dye, but some larger than others. They were all linked together by slender Wires. These Bits of Wood were 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 Declensions, but without any Order.

[10] The film shown in 1935 with great success was *The New Gulliver,* an animated film using three-dimensional figures and wax dolls as well as live actors, produced after three years of gruelling work by Alexander Ptushko. "Swift's story is framed by a reading in a Crimean camp of Young Pioneers; their dreams bring it up to date with amusing details (a newsreel cameraman photographs the hauling of sleeping Gulliver into the capital, modern engineering solves the problem of feeding Gulliver, etc.) and a class-war in Lilliput, with Gulliver aiding the oppressed." Parts of the film were shown at the Venice international exhibition of 1934, and the film was much praised for its ingenuity and inventiveness (Leyda, 1983:309)

[11] Lev Ivanovich Zhirkov (1885–1963) was a specialist in Caucasian languages, writing grammars and dictionaries between the late 1920s and the early 1950s for Avar, Dargwa, Lezghian, Tabasaran, Lakk, and other Dagestanian languages.

[12] The preface includes the sentence: "The materials relating to P. P. Troyanskii's translation machine are published with the permission of Z. N. Smirnova-Troyanskaya…"

[13] Troyanskii's expression *национально-грамматическая форма* 'national-grammatical form' refers to the inflected forms (of a nouns, verbs, etc.) in the local language of a particular country, region or "nation" (which may be one of the "autonomous" republics of the Soviet Union). Despite its awkwardness in English it is retained in this translation. By the "initial" form Troyanskii refers to the basic (canonical) dictionary form.

[14] See Paragraph 14 below, where this method is described.

[15] Troyanskii adds here a footnote referring to the "translation patterns" in Section 14.

[16] The editors refer readers to page 35 in their collection, i.e. to the explication accompanying the original patent application, reproduced here in Section 4.

[17] English translations have been added for those unfamiliar with the distinctions present in the original Russian and German examples.

[18] Again, an English gloss is added for present readers.

[19] Troyanskii adds the footnote: "to combine without deformation of the national text."