Computerized Russian Translation at ORNL

Since 1964, as an adjunct to its automated technical information processing services to ERDA and other federal agencies, a generalized language translation system has been used by the Oak Ridge National Laboratory (ORNL) to translate Russian scientific text to English. The translation system, first implemented at Georgetown University around 1960, has been rewritten and improved through the years as computer models changed. Although the translations lack high literary quality, the system, by means of its context sensitive dictionary, nevertheless provides inexpensive, fast and highly useful translations of scientific literature.

The method used involves a linguistically-oriented programming language called Simulated Linguistic Computer (SLC), with which a language-specific dictionary can be written for use by the translation system. The dictionary entry for any word can be augmented by procedures which permit its meaning to be modified by its context; more general linguistic procedures operate on the sentence as a whole.

In an evaluation of user reaction, over ninety percent of the respondents rated the machine translation (MT) service "good" or "acceptable" on translations of their subject specialty. Development, implementation, and documentation of the system are continuing, as we meet increasing requests for service and attempt new applications of the MT system.

Sara R. Jordan*
Antony F.R. Brown**
Fred C. Hutton
Computer Sciences Division
Oak Ridge National Laboratory
Union Carbide Corporation
Nuclear Division ***

• Historical Background

In the great surge of optimistic enthusiasm and government support for machine translation of the 1950's and early 1960's, many projects were launched to solve the seemingly straightforward problems of handling large files of words and the selection of the appropriate alternative translated meaning for each word from a bilingual dictionary. Interesting historical accounts of the early days of machine translation (MT) can be found in Pendergraft (9) and Locke and Booth *(8).*

Early computerized translation approaches were unable to handle idiomatic forms. In addition, it was found that human pre-editing was needed to analyze the source text, insert diacritical marks helpful in dictionary search, and that extensive post-editing was necessary to reshape and clarify the translation output. In order to eliminate the need for human intervention, Leon Dostert proposed to include in the bilingual lexicon a systematic set of diacritics which describe the general use pattern or type for each word. Since the diacritic information for a word summarized its use in all contexts, it was necessary to augment the lexicon with the ability to mechanically search and manipulate the context of the particular word being considered. (See (5) for details of Dostert's ideas.)

When the Georgetown University Machine Translation Research Project was funded in 1956 with Dostert

---

*Consultant, Computer Science Department, University of Tennessee, Knoxville, TN
**Consultant, Control Data Corporation, Sunnyvale, CA
***Prime Contractor for the U.S. Energy Research and Development Administration

as director, several approaches were developed to implement Dostert's scheme. The two most important and successful approaches were the "general analysis" method of Michael Zarechnak and the "sentence-by-sentence" method of A.F.R. Brown.

In the general analysis approach, a sentence is analyzed on several levels, including 1) formation of words and idioms, 2) word grouping to determine noun-adjective agreement and verb-noun government, and 3) syntactic relationships such as subject-predicate analysis.

With the novel sentence-by-sentence method, a linguist could "bootstrap" the system to increasingly larger competence, starting with a basic set of information and procedures, and adding more until the system was able to handle any sentence. More importantly, Brown needed a way to communicate linguistic ideas and statements directly to the computer, so he designed a language which, after a few years of development, became known as the Simulated Linguistic Computer (SLC) language. With SLC, it is possible for a linguist to program machine translation between any two languages which can be represented as strings of symbols.

Although Zarechnak's multilevel linguistic analysis approach was the first Georgetown MT program to be successfully tested, Brown's SLC was also considered a very significant development. The current version of the machine translation system at ORNL clearly shows its origins in these early ideas.

In 1964, a copy of the current Georgetown MT system (combining both Zarechnak's and Brown's contributions), was presented to the Atomic Energy Commission's Oak Ridge National Laboratory by Dostert, in order to have large-scale testing and evaluation by local scientists. A similar arrangement had been made with the EURATOM Center at Ispra, Italy. Meanwhile. Bar-Hillel published his famous statements concerning the nonfeasibility of fully automatic high quality translation (2) and the pessimistic ALPAC Report published in 1966 *(1)* caused termination of government support for machine translation projects.

Despite the widespread pessimism, the experience at ORNL with the Georgetown MT program was quite satisfactory. The system was installed at the Computing Technology Center by J.A. Gillcrist under the direction of A.A. Brooks. After 1965, largely through the efforts of Francois Kertesz, the ORNL program was kept viable during a period of general apathy and lack of firm financial support. As the years went by, many thousands of pages of Russian were translated for scientists at ORNL, AEC's Technical Information Center, and elsewhere. Although the output was not of high literary quality. the system nevertheless provided valuable, useful, and fast translations for which the customers paid willingly. As Dr. Kertesz noted in an August 1964 letter to C.P. Keim, "only the actual user who needs the information can evaluate whether the material he received is satisfactory for his purpose and whether he is willing to spend the additional time for reading the somewhat awkwardly-worded material. This might be a small once to pay for obtaining the translation immediately instead of waiting days or weeks for it." This philosophy of the ORNL automatic translation service has since been echoed by a 1971 statement of Bar-Hillel *(3).*

In contrast to the unfavorable economic factors which ALPAC had found for most MT projects, the translation service at ORNL has not been hampered by major development costs or by the need for any computing facility beyond the genera! computer used by everyone else. Furthermore, translations were and are produced only by request of users who will be *billed* for the service, so unnecessary effort and expense are minimized. Finally, no human post-editing is required for users familiar with the subject area (they do the post-editing themselves),although it has been done on some projects for outside users or publication.

In the next two sections, an overview of SLC and the translation system will be given, followed by sections giving more detailed descriptions and a sample translation.

• Overview of the SLC Language

The SLC programming language was designed to provide a communication link between the computer and the linguist, specifically for the purpose of translating languages. After the source and target languages (Russian and English, for example) have been determined for a specific translation application, then those languages should be analyzed to determine which linguistic features (such as gender, case, etc.) are important in producing the external "shape" of each language. The SLC computer language enables the linguist user to encode and manipulate many types of linguistic information in a convenient symbolic form. Coded information may be attached to particular words or endings in the dictionary. Linguistic situations may be specified in terms of part of speech, case, number, gender, or even a specific set of words. If such a situation is found, the instructions in the dictionary definition may change the *translation* of a word or group of words, may rearrange the *order* of the words, or may even *add* or *delete* some word(s) in the sentence. For example, the choice of case and number endings for a particular word may be determined by inspection of the information attached to *other* words found in the sentence. As another example, a particular Russian word may have a basic translation into English, but may also have other meanings if it occurs in any of several specified contexts. The dictionary entry for that word then may include not only the basic translation, but also may have one or more operations to modify the translation if a specified context is found.

Besides the instructions associated with particular words, there are some operations which must occur for *every* sentence the computer translates (determination of the subject, for example). All instructions are given priority numbers which cause them to be enacted at proper times throughout the processing of the sentence. In short, the dictionary and operations used for the translation must include all such information needed to specify the structural pattern of the language, as well as the linguist understands it.

• Overview of SLC Translation System

The translation system for given languages has two major modes of operation: 1) system generation and 2) actual translation. Fig. 1 shows a simplified schematic view of the overall system. System generation involves compilation of the language-specific dictionary and associated linguistic operations, from SLC instructions to machine code. The translation phase of the system uses the resulting dictionary and operations files to accept source language text as input and produce as output the target language translation, usually with the original source text printed alongside. For clarity and convenience in the remainder of this paper, we shall assume Russian-to-English translation which is the type of operation at ORNL; that is, source language = Russian, and target language = English. For input, and in the dictionary, Russian text is currently transliterated according to a standard scheme using Roman letters and numerals.

The system generation phase of the system is used *only* if the existing dictionary is to be updated or if there is no existing dictionary. Any new dictionary entries and associated linguistic operations written in SLC are compiled into tables and machine code by a symbolic assembler program. The newly compiled entries are then merged with the old dictionary file to produce a new, updated version. Dictionary listings can be printed when desired. If new general linguistic operations are to be added, the entire set (old and new) is compiled and a new operations file is created. The files thus generated are then available for use by the translation program.

The translation program assumes the existence of a file of standard linguistic operations designed for general analysis of any Russian sentence, plus a dictionary file which contains for every word a basic translation, grammatical information, and perhaps, also, one or more associated operations which modify the basic translation in certain contexts. Given these resources, the translation program accepts a Russian sentence and creates from it a working list of "items", one for each word or punctuation mark, with which to produce the English translation. Each item contains the information (including any special operations) which is found in the dictionary entry for that word or punctuation mark.

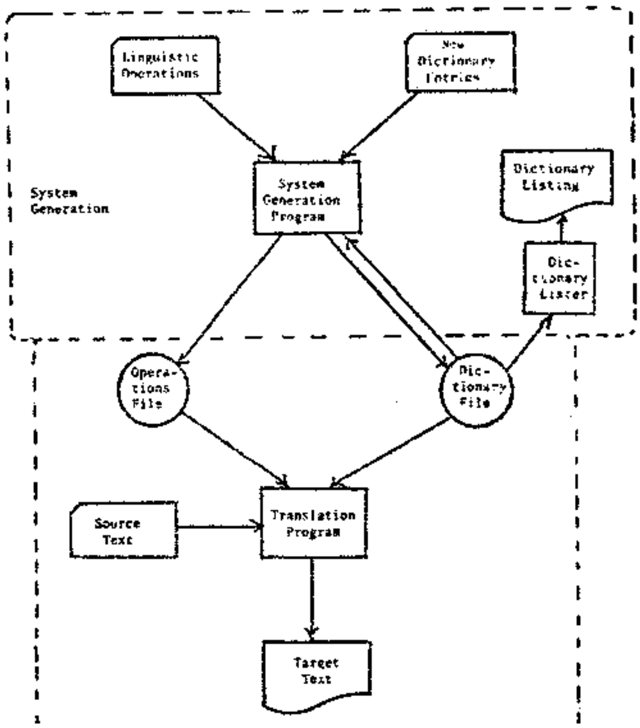

Fig. 1. Schematic of Translation System. (1) Computer programs are independent of source and target languages within the class of languages permitted. (2) Operations and Dictionary are source and target language dependent (Russian to English currently implemented). (3) Dictionary is context oriented with appropriate operations. (4) Sort & scratch files and activity listing are not shown.

Every sentence input for translation must end with an asterisk which serves as the final period. The final asterisk has a special function of causing sentence-wide operations required in the translation process. Thus, for the input sentence "ON GOVORIT *", there will be three items for "ON", "GOVORIT", and "* ", respectively. Suppose that each word in the sentence has been looked up in the dictionary, and the complete dictionary entry from each word (including not only the usual English equivalent but also any additional grammatical information and instructions) has been stored in that word's "item." Each word in the sentence has until this point been considered independently, and carries along with it the information to handle any situation it might be appearing in, because the context has not yet been considered.

This list of "items" with their separate sets of instructions must now be analyzed for interaction so that the correct translation from Russian to English may be produced. This is accomplished by merging *all* the words' instructions conceptually into a single set, ordered by priority. Then each instruction is executed on the sentence, highest priority first, or in the case of equal priority, left before right. Each instruction includes some test(s) to be made by scanning the items in the sentence and noting the linguistic information "tags" (such as number, gender, case), or even checking for specific strings in certain locations. If the test fails, the instruction does not apply and is terminated. If the test *is* satisfied, some specified modification of one or more word "items" in the sentence is made. After an instruction has been completed, the system chooses the next higher priority instruction to execute. When no more instructions remain to be executed, then the translated sentence can be output; starting from the first "item" in the sentence, the resulting English equivalent of each "item" is output, ending with a "." (period) which is defined as the English equivalent of the final asterisk.

Having completed this sentence, the computer system reads in a new sentence and begins the whole process again, starting with the dictionary lookup.

• The Translation Dictionary

Now that an. overview of the operation of the entire translation system has been given, let us consider the form of the dictionary and linguistic operations on which the system is based. The nature of the dictionary and the basic dictionary entry will be described in this section. The use of SLC to write linguistic operations is discussed in the next section. For more detail on either subject, see (6).

The dictionary used for the SLC translation system contains stems and endings as well as full words. Each word stem is listed once and each ending that can be used with multiple sterns is also listed once. Each word base is listed once and each ending which can be used with multiple bases is listed once. The dictionary entry for any word base indicates which, if any, set of endings should be considered as endings for the given word. This minimizes both dictionary size and search time for legal endings. The Russian-to-English dictionary used by ORNL contains about 31,000 Russian stems; with the case and tense endings available to the system, this is equivalent to about 300,000 actual words. The dictionary was prepared primarily for scientific work (and was text-based, not dictionary-based), but the genera! language is sufficient enough that a few pages of a book on Russian ballet were translated creditably.

The basic form of a simple dictionary entry is the following:

```
*      russian  -ending- paradigm  grammatical  lexical  'english'
↑      string            number    code         number   string
Col. 1                   ⏟
                    may be repeated
```

A hypothetical example is:

* ESLI - CONJ 26 'IF'

Between the asterisk in column 1 and the period at the end of the card, there are five significant fields:

1. The Russian stem ("ESLI") which is transliterated from the Cyrillic alphabet according to a standard scheme.
2. One or more indicators of ending and paradigm number. The first hyphen precedes an ending; the second hyphen, if present, precedes an integer indicating which standard paradigm of endings should be applied to the Russian stem. (Standard paradigms are listed in the SLC Primer (6).) If the Russian stem happens to be uninflected (as in our example), the field should consist of a single hyphen, which indicates a null ending.
3. The grammatical coding (such as "CONJ") that translation rules can test for to see if a word is in a certain category. If no grammatical code is needed, the null or default code is "XXX".
4. The lexical number ("26" in the example), an identification number unique to this word, which is arbitrarily assigned to some words if they are referred to by linguistic operations of other words. When there is no foreseen need for a lexical number, a zero ('"0") will be placed in this field.
5. The English equivalent with each word followed by a space for printing purposes.

Let us consider an actual dictionary entry which involves inflection:

* ADAPTAQI --18 FN 0 'ADAPTATION $I'

The transliterated endings listed (in standard order) by the dictionary for paradigm. 18 are: 4,I,I,H.E1,I; I,1,4M, I,4MI,4X. Thus the stem "ADAPTAQI" plus the endings in paradigm 18 handle all the forms in the paradigm of "adaptaqi4". The grammatical code entered is "FN", for feminine noun. It is common to have more complicated situations which require a) multiple sets of ending-paradigm-grammatical information and/or b) both local ending and standard paradigm endings for a particular stem. The dictionary entries for these situations are described in Chapter 2 of the SLC Primer (6).

Looking further in the example entry, we see that the lexical number is zero, since no particular lexical number is needed. The "SI" which follows the English equivalent is an English inflection indicator; in this case, the "I"' as the first digit after the "$" indicates that "adaptation" forms the plural simply by adding "-s".

• Linguistic Operations Written in SLC

If the translation operation were limited to what can be specified in the basic dictionary entry, the system would perform only word-for-word translation. The important power that the SLC system gives the coding linguist is the ability to specify additional linguistic rules or operations which should take place for every sentence or only in certain contexts around certain words. Examples of the need for such operations were mentioned previously.

SLC is a generalized programming language which can be adapted to any machine translation situation. The commands included in the language are those needed by a linguist working with a sentence-length string of source language "items", each of which can have various sorts of linguistic information attached to it. In order to establish a base, or point of reference for a given operation, the item from which the operation was invoked is designated as both the "source item" and, initially, as the "current item". In producing a translation, the linguist may wish to scan the sentence in either direction, checking for different types of information, and to replace, delete, or rearrange the items' translations. Consequently, SLC includes commands such as L and R to move the current designation one item to the left or right, ami S to get back to the source item. The DEL command deletes the current item completely from the sentence, and the RPLC- =item= command replaces the current item in the sentence by the specified item.

There are a number of query commands which test for a specified condition and yield a condition which determines further action. A negative result condition will stop the operation unless the next command is a branch. Sample query commands are QLX-n (Query LeXical number) and QR (Query Russian string). There are also a number of somewhat complicated commands which add, delete, or change grammatical codes, endings, and other supplementary information. There are over one hundred instructions in the SLC language, but a rather small subset of these suffices to describe many operations. The SLC language is described in detail in the SLC Primer (6).

A linguistic operation is written with a series of SLC commands, separated by spaces. An operation associated with a particular word follows the basic dictionary entry for that word. Given only the SLC commands mentioned above, the reader may understand the following example taken from the Russian to English dictionary. If translated word for word, the phrase "PO KRAIN(some ending) MERE " would result in "ACCORDING TO EXTREME MEASURE ". The operation which follows the "KRAIN" entry changes that phrase's translation to "AT LEAST ".

```
*    KRAIN    --35   XXX    0    'EXTREME '  .

+L-6   R   QR-'MERE  '    SLQLX-46   DEL   S R   .PO

     DEL RPLC- =ADV     0     'AT LEAST '=     +/  .
```

When the operation begins execution, the "KRAIN" item is current and source. First the item to the right is checked to see if it is 'MERE ', then the item to the left of the source is checked to see if it is 'PO ' (with lexical number 46). If both these tests are satisfied then two of the three items are deleted and the third (the 'KRAIN" item) is replaced by the adverbial phrase 'AT LEAST '.

• Sample Translation by SLC System

The example in the preceding section demonstrates how SLC is used to augment dictionary entries with procedures which can change the translation, depending on the particular sentence under consideration. Other operations, not dependent on the occurrence of particular words, perform more general linguistic procedures (such as determination of the subject). The dictionary for the existing Russian-to-English translation system already has rules to take care of the vocabulary and contextual situations fairly successfully, as may be seen in the sample translations in Figs. 2 and 3. In each case, the original Russian text is printed in the right column, with the English translation on the left. A comparison of translation performance between the earliest dictionary used on the IBM 360 and the current dictionary is illustrated in Figs. 2 and 3. Improvements consisted of adding words to the dictionary where needed, changing the English equivalent of some words to a more suitable form, and replacing the literal meaning of some phrases by idiomatic translations. Fig. 2 shows the Russian text transliterated according to the scheme used through the years up to January 1976. In Fig. 3, the Russian was printed using a new Cyrillic print train, which yields much more aesthetically pleasing text.

UPON FURTHER INCREASE OF TEMPERATURE QUANTITY OTVERDITELH IS SUCCESSION IS DECREASED AND INSTALLED IN TESTABLE [illegible] .

INVESTIGATION OF INTERACTION OF OSCILLATIONS OF DECIMETER DIAPASON WITH PLASMA ON APPARATUS TOKAMAK TM.-3. .

EXPERIMENTAL ENERGY APPARATUS BRG-30 WITH GAZOOKLADJDAEMYX REACTOR ON RAPID NEUTRONS AND DISSOCIATING HEAT-CONDUCTOR .

SUCH PODATLIVYE ENVELOPES OF TIRES AND CHANNELS CAPABLE OF [illegible] AND TO REDISTRIBUTE LOCAL TENSIONS , WHICH ENGAGE BY FREEZING IN THEM LIQUIDS .

INTRODUCTION OF POLYMER EMULSIONS LEADS TO INTENSIVE RAZUKRUPNENIA OF TIRES .

BY [illegible] PROPERTY THESE MATERIALS IS INCREASED BREAKING STABILITY AND RASTNOJIMOSTG UNDER LOAD .

WIDE DEVELOPMENT OF NUCLEAR POWER ENGINEERING POSSIBLY ONLY ON BASIS OF REACTORS OF RAPID NEUTRONS .

TO REGRET , UTILIZATION OF NOT VERY ACCURATE EXPERIMENTAL DATA (FOR EXAMPLE , IN WORK OF VITOLOSA) OR LITTLE OBOSNOVANNYKH PREMISES (FOR EXAMPLE IN WORK OF VATSONA) PERMITS TO MAKE CONCLUSION CONCERNING SMALL RELIABILITY , ACCORDING TO EXTREME MEASURE FROM QUANTITATIVE POINT OF VIEW , OBTAINED IN THESE WORKS OF RESULTS .

LOW MOROZOSTOIKOSTG ANHYDROUS USUAL CONCRETE WHICH WAS CAUSED NEZAKONCENNOSTGH PROCESS OF TVERDENIA BEFORE EXPERIMENT AND , AS CONSEQUENCE , BY IMPERFECTION OF ITS STRUCTURE .

P* PRI DALGNEISHEM POVYSHENII TEMPERATURY KOLICESTVO OTVERDITELH POSLEDOVATELGNO USENGSAETSQ I OSTANAVLIVAETSQ PRODBZNA ZAMESANI *

P* ISSLEDOVANIE VZAIMODEISTVIQ VC KOLEBANII DECIMETROVOGO DIAPAZONA S PLAZMOI NA USTANOVKE TOKAMAK TM, - 3, *

P* OPYTNAQ ENERGETICESKAQ USTANOVKA BRG - 30 S GAZOOKLAJDAEMYM REAKTOROM NA BYSTRYX NEITRONAX I DISSOQIIRUHSEM TEPLONOSITELE *

P* TAKIE PODATLIVYE OBOLOCKI POE I KANALOV SPOSOBNY POGASITG I PERERASPREDELITG MESTNYE NAPRQJENIQ , VOZNIKAHSIE PRI ZAMERZANII V NIX JIDKOSTI *

P* VVEDENIE POLIMERNYX EMULGSII PRIVODIT K INTENSIVNOMU RAZUKRUPNENIE POE *

P* QENNYM SVOISTVOM ETIX MATERIALOV QVLQETSQ POVYSENNAQ RAZRYVNAQ PROCNOSTG I NASTOQIMOSTG POD NAGRUZKOI *

P* SIROKOE RAZVITIE QDERNOI ENERGETIKI VOZMOJNO LIHG NA OSNOVE REAKTOROV NA BYSTRYX NEITRONAX *

P* K SOJALENIH , ISPOLGZOVANIE NE OCENG TOCNYX EKSPERIMENTALGNYX DANNYX (NA PRIMER , V RABOTE VITOLGSA) ILI MALO OBOSNOVANNYXAS PREDPOLOJENII (NA PRIMER V RABOTE VATSONA) POZVOLAET SDELATG VYVOD O NEBOLGSOI NADEJNOSTI , PO KRAINEI MERE S KOLICESTVENNOI TOCKI ZRENIQ , POLUCENNYX V ETIX RABOTAX REZULGTATOV *

keypunch error

P* NIZKAQ MOROZOSTOIKOSTG SUXOGO OBYCNOGO BETONA VYZVANA NEZAKONCENNOSTGH PROQESSA TVERDENIQ PERED OPYTOM I , KAK SLEDSTVIE , NESOVERSENSTVOM EGO STRUKTURY *

Fig. 2. Sample Russian-to-English Translation Using Early Dictionary. NOTE: ....= changed meaning; = literal meaning changed to idiom.

• User Reaction to MT

Bar-Hillel in 1971 recognized that: "Every program for machine translation should be immediately tested as to its effects on the human user. He is the first and final judge, and it is he who will have to tell whether he is ready to trade quality for speed and to what degree." *(3, p. 76). A* study of users' evaluations of machine translations provided by the original Georgetown systems at Oak Ridge and in Europe was issued in 3973 by Bozena Dostert. The study, which is based on ten years of MT experience at the sites, yields the following conclusions, among others *(4, pp. 65-66):*

\. The quality of MT is rated "good" or "acceptable" by 92.4 percent of the 58 respondents.
2. Familiarity with subject matter is of primary importance in understanding MT texts.
3. If the wait time for human translations is three times that for machine translation, 87 percent of the users prefer MT.
4. Machine translation had been, or would be, recommended to colleagues by 96 percent of the respondents.

The MT system is clearly providing a useful and appreciated service. The users expressed willingness to continue using the system and encouraged further development of the system and its vocabulary (rated much more important for understanding than syntactic considerations).

• Current Status and Future Plans for MT at ORNL

The MT system has four parts: the linguistic approach used in translating, the dictionary used by the program doing the translating, the programs themselves, and the SLC language used in coding linguistic operations. The linguistic approach is essentially unchanged from its original design at Georgetown by Zarechnak. Some

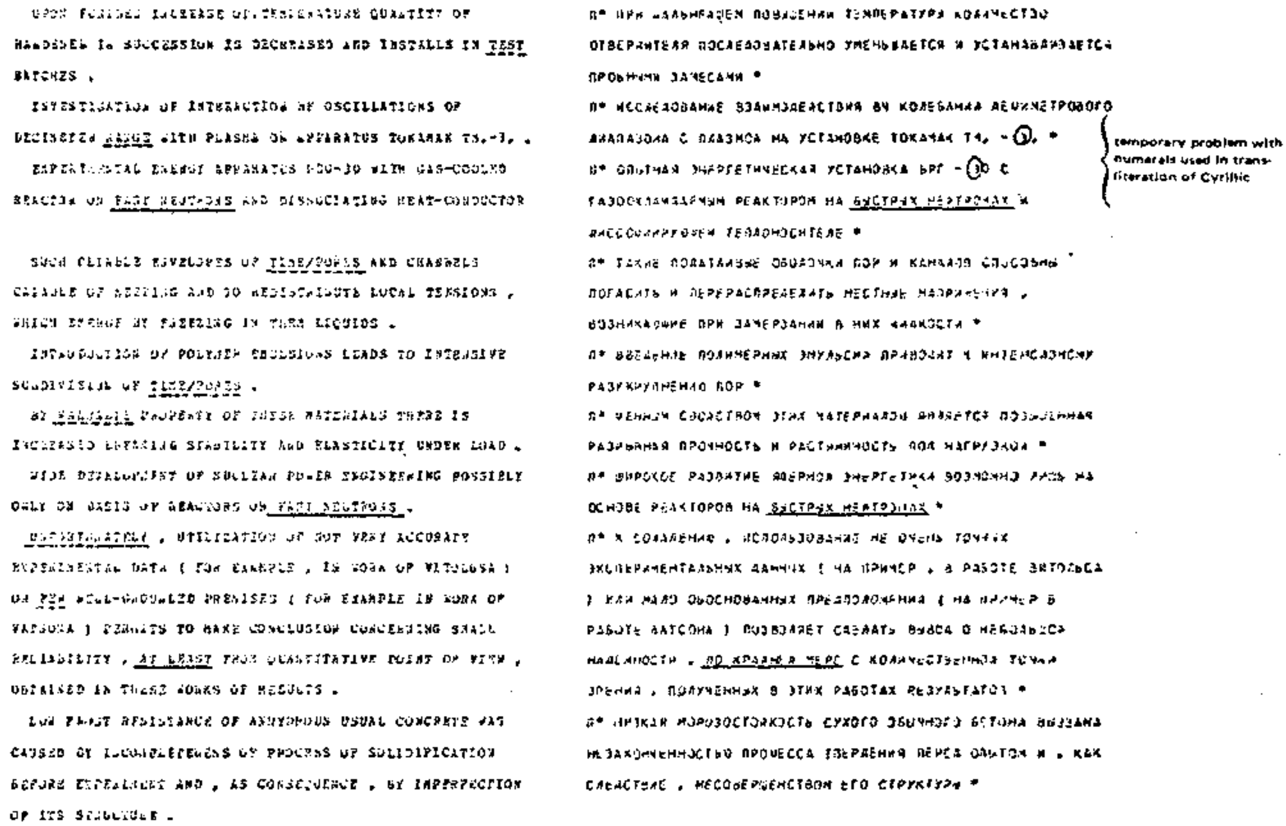

Fig. 3. Sample Russian-to-English Translation Using Current Dictionary and Cyrillic Print Train. NOTE:...............= changed meaning; _______= literal meaning changed to idiom.

improvements have been planned by Dr. Zarechnak (a consultant to ORNL) and will be introduced to the system when further support is possible.

The dictionary, too, is largely the same as that developed for the Georgetown Project. The procedure for updating the dictionary has been greatly improved and streamlined, however, and quite a few changes (numbering in the hundreds) have been incorporated, as solutions to particular translation problems were found. The nature of many of these improvements has been noted earlier.

Over the years, hardware changes also have made possible substantial improvements in the programs. File handling has been streamlined and input to the system has become much more simple and flexible. Further improvements are envisioned as support becomes available.

One of the main strengths of the system is its use of the linguistic programming language SLC, with which a user can specify fairly complicated linguistic procedures directly to the computer. It is important to remember that SLC is not bound to the Russian language; in fact, it was first used to translate French to English. The richness of the language and the number of commands available are somewhat formidable to the new user of the system, so a primer (6) is being prepared to introduce SLC and the MT system in stages, beginning with a small but useful subset of the SLC commands.

It is interesting to note that in a recent report, the All-Union MT Committee of the U.S.S.R. stressed the need for the "development of a special metalanguage for describing processing operations that would help to translate linguistic statements into the absolute machine - coding system" and "creation of a monitoring system to help the researcher to see and evaluate not only the final product, but also the intermediary stages through which input data are processed" *(10, p. 2843).* SLC is such a metalanguage (the only one we know of in use in the U.S.), and the system does have the monitoring capabilities described. We hope to make SLC more widely known and more easily usable, by means of documentation and tutorial seminars, so that it may be used not only to improve the Russian-to-English dictionary, but also to develop other translation systems.

The MT service at ORNL has assumed new vitality in the last year or so with the completion of the improved version of the programs and dictionary for use on the IBM 360 and a renewal of interest in expanding the dictionary with SLC code, according to the suggestions of a select group of users. We plan to adapt the system to assist in the translation of machine readable Russian data bases, provided via the Nixon-Brezhnev cultural exchange agreement, for incorporation into the comprehensive information retrieval and dissemination service operated at ORNL. Computer costs for machine translation service are fully recovered by charges to the user; keypunching is the major factor of the cost. Developmental costs are not passed on.

• Technical Information

The Russian-to-English machine translation system used by Oak Ridge National Laboratory runs on an IBM 360-370 computer system under Operating System 360 (OS 360). All programs are written in Assembler Language, and none requires more than 190 K of core. Four programs are involved: one does the actual translation, two programs are used when changes are made to the dictionary, and a fourth program is needed if the dictionary is to be printed. The current system may be made available outside ORNL on a selective basis for cooperative testing and further development.

A recent translation run of 2100 words of English output took 8.76 seconds of CPU (Central Processing Unit) time on an IBM 360-195. Another translation run of 8-10,000 words (50 page listing) took 41 seconds CPU time.

• Acknowledgments

Dr. Michael Zarechnak of Georgetown University developed the linguistic approach used in the translation system, and continues to provide analysis and suggestions on linguistic aspects of the system. Dr. A.A. Brooks of the Computing Applications Department, Union Carbide Corporation, Nuclear Division, has provided advice and support throughout this project. From the very first, the primary user of the machine translation service and its main supporter has been ORNL's Information Division. Dr. Francois Kertesz's contribution has been mentioned. Mr. R.R. Dickison and Dr. H.F. McDuffie (present head of the Information Division) continue to provide strong support. Ms. Martha Gerrard, the key figure interfacing between requestors of translation service and the computer operation, is the head of the Information Division's Office of Language Services; she also does any required human editing.

References

1. Automatic Language Processing Advisory Committee Report. 1966. *Language and Machines: Computers in Translation and Linguistics.* Washington, DC: National Academy of Sciences—National Research Council; 1966.
2. Bar-Hillel, Y. 1964. "A Demonstration of the Nonfeasibility of Fully Automatic High Quality Translation." in: *Language and Information.* Reading, MA: Addison-VVesley; 1964. 174-179.
3. Bar-Hillel, Y. 1971. "Some Reflections on the Present Outlook for High Quality Machine Translation." In: NTIS Reports (2 Vols) AD-737573; AD-737S74; 73-76.
4. Dostert, B.H. 1973. User's Evaluation of Machine Translation. Rome Air Development Center Technical Report RADC-TR-73-239.
5. Dostert, L.E.; Macdonald, R.R. 1963. General Report, 1952-1963. Georgetown University Occasional Papers on Machine Translation, (30).
6. Jordan, S.R.; Brown, A.F.R.; Hutton, F.C. 1977. SLC Primer. ORNL Technical Report (in preparation).
7. Lehmann, W.P.; Stachowitz, R. 1971. Feasibility Study on Fully Automatic High Quality Translation. NTIS Reports. (2 Vols) AD-737 573; AD-737 574.
8. Locke, W.N.; Booth, A.D. 1955. *Machine Translation of Languages.* New York: Technology Press of MIT and John Wiley & Sons; 1955.
9. Pendergraft, E.D. 1968. "Translating Languages." In: Borko, H. ed. *Automated Language Processing.* Wiley; 1968. 291-323.
10. Roberts, A.H.; Zarechnak, M. 1974. "Mechanical Translation." In: Sebeok, T.A. ed. *Current Trends in Linguistics,* Vol. 12 1974. 2825-286S.