[From: Internat.conf. *Methodology & Techniques of Machine Translation*, British Computer Society, Cranfield, 13-15 February 1984]

Machine Translation in the USSR

Yu. N. Marchuk

INTRODUCTION

Three aspects of the state-of-art in machine translation in the USSR are considered: theoretical, practical and informational.

The theoretical aspect includes principal questions of the linguistic theory associated with MT. The practical aspect is devoted to the implemented systems. The informational aspect covers incorporation of MT into the operation informational service network. Any other assessment of MT without such an incorporation will not be expedient.

THE THEORETICAL ASPECT

Any increase in the machine output quality of modern MT systems is connected inseparably with the nature of the translation models in use. One of the main trends in the simulation of translation is research study and reproduction of the actions performed by a human translator. It has been noted that the existing translation models involve all the operations that are specific to human translation. It can also be stated that the actions of a human translator involve all operations that, to a certain extent, must be simulated in computer-aided translation.

When simulating the translator action, the authors rarely build a general model but rather confine themselves to a certain type of translation. Thus, Z. M. Shalyapina [1] considers written translation that is decomposed by her into a sequence of operations, some of which can be easily implemented by the computer (operation on the surface level), whereas others are difficult for machine implementation or cannot be implemented at all at the present stage of development.

A. F. Shiryaev[2] offers a description of a model of the simultaneous translation functional system based on theoretical and experimental studies of simultaneous translator actions. The author proposes that the simultaneous translation functional system be treated not as a system in general but as a specific system, on the grounds that

simultaneous translation can be mastered normally by means of the development and rearrangement implemented in functional systems of other kinds of translation. Basic techniques of simultaneous translation are: timing, understanding of a source text, search for and implementation of translation options, their verification and correction.

The leading role is assigned to the timing technique represented by various levels of actual cognition, unconscious verification, conscious verification, etc. Yu. N. Marchuk does not think that simultaneous translation is absolutely unique as a form of translation and in his concept of the translator's actions (oriented mainly to simultaneous interpretation), he does not stipulate the specific features of the latter, but links up the interpreter's actions with a certain concept of linguistic understanding[3]. The increased interest in the interpreter's actions and simulation thereof correspond to the existing trend in the world, and reflect the importance of the "transfer stage", i.e. the translated correspondences proper in construction modern systems of machine translation.

Another important direction of the theory is a study of specific features of sublanguages in connection with simulation of translation. It has become apparent after a period of experimental use of many MT systems that the quality of translation can be improved if the specific features of sublanguages which aid automated analysis are reflected in dictionaries and algorithms of analysis and synthesis. The theory of sublanguages or language subsystems was first formulated by the Soviet scientist N. D. Andreev[4], In a recently published book L. L. Nelyubin[5] offers a theory concerning sublanguages from the view-point of machine analysis and translation. The sublanguage is described by four models: functional-communicative, statistical, informational and linguo-statistical. On the basis of these models the translation system is constructed for organizational and management documents translated from English into Russian. This system is based on the computer dictionary compiled specially for this purpose.

Problems of lexical analysis based on formal indices - even if these are not explicitly tied to MT - are of great importance to the latter, since the compilation of dictionaries is the most labour-consuming subtask in any MT, whereas their completeness and adequacy to the formulated objective greatly facilitate the improvement of the MT's quality. First of all, in analysis of the vocabulary, attention is focussed on the ambiguous nature of a word (polysemy), whose resolution is rather important for translation. P. Ya. Serdobintsev[6] refutes the assertion of the outstanding Soviet linguist R. A. Budagov[7] that polysemic words represent about 80 per cent in any glossary. He gives data certifying that out of 10515 words analyzed in two vol-

umes of "Modern Russian Literary Dictionary" there are 8657 monosemic words or 83.5 per cent and 1372 polysemic ones, or 16.5 per cent. Thus the real picture is directly opposite to that obtained on the basis of an intuitive notion.

The concept of word ambiguity in MT is associated not with traditional polysemy, but rather with the existence of several translation options; nevertheless this data is undoubtedly of a great interest.

Modern MT is translation of scientific and technical texts. A complex of problems associated with the correlation between the terminological and common use vocabularies is under intensive studies. In particular, the question of how words of the common use vocabularies become terms and vice versa - how terms come into common use - is of great interest[8]. In most cases the authors draw the conclusion that revealing the meanings of a word is inseparable from the context.

Comparative study of languages has recently become a particular trend in linguistics defined as contrastive linguistics. This * branch of linguistics always deals with data of two languages or of several languages studies in pairs.

A direct object of contrastive linguistics is a set of the arranged systems or subsystems. In this, constructive linguistics is linked closely with problems of translation in all their modifications[9].

Studies of the language semantic level are under way in connection with MT. However, it is noted that one should not overestimate the possibilities of the semantic level. If a sentence has passed through all levels of analysis and obtained a unique semantic representation, then the synthesis will ensure the normal translation for it. However, this level is incapable of resolving all ambiguities, on the contrary, it can make a syntactically monosemic sentence ambiguous[10]. Attempts are being made continuously to employ the semantic dictionary of the combinatorial and glossary type in the formal model of language, and to resolve the polysemy in machine translation[11], A tendency to formalize the meaning of word chains exceeding in length one word or a sentence brings us to a concept of developing the meaning of a text as a whole. The texts written in natural languages do not accept fixed semantics. The meaning of a word can be determined when the word has a contextual environment.

A word, as such - devoid of any potential context - is simply a sign or a name of some object: but it is meaningless since, even potentially, it is not an element of meaningful statement.

Each word harbours a multitude of meaningful texts in it, in which it can be included. Not just **any** set of words is meaningful, however, due to the fact that for some words in it the remaining words cannot form an appropriate context. Naturally, the meaning of the text can become clear to a reader only to that extent to which he is familiar with the language, i.e. to which he knows the potential contexts of each word. Thus, the question concerning the nature of the meaning which a word has can be confined to the question of the nature of meaningful statements[12]. The authors describe the concept of statement meaning through the concept of **role structures,** which are understood as an abstraction of a functionally integral situation as a set of "roles", regardless of the particularities of its elements.

Everything mentioned above illustrates the multilateral policy in theoretical studies in the field of MT which covers, as before, a wide range of problems.

THE PRACTICAL ASPECT

Practical activity in MT is based on the operation of a number of MT systems with post-editing. In the USSR Centre for Translation of Scientific-Technical Documentation (UTC), the AMPAR (Automated Machine Translation from English into Russian) machine translation system has been operating in an industrial environment for a number of years[13].

It is intended for the translation, with post-editing, of texts covering radio-electronics, computer science, programming and a number of other technical fields.

The linguistic support of the AMPAR system is based on a special translation model using translation correspondences, and consists of two components: the dictionary and grammar ones. The entire translation process is divided into 17 stages, each performing a specific analysis, translation or text synthesis operation. Source text analysis (stages 1 to 7) covers morphological analysis and word form matching against the dictionary, search in the text, analysis and translation of set expressions, resolution of all forms of homonymy, syntactical analysis by parts of speech and by parts of the sentence are carried out. Translation per se (the set of "transfer" stages), stages 8 to 14, involves translation of unambiguous and ambiguous words using the contextual environment analysis. Synthesis of the target text (stages 15 to 16) is performed in two stages: syntactical synthesis, i.e. establishment of syntactical and morphological correspondences between the English text and the Russian one, and morphological synthesis. The entire translation process is completed with the listing of the

target text (stage 17) . Listing can be accomplished using various options: page-by-page listing, parallel listing of the Russian and English text, etc.

The dictionary component of the system represents a sophisticated interaction of a number of dictionaries. The source dictionary (over 25 thousand words) is compiled according to subject fields. The English dictionaries used in operation are the common vocabulary dictionary, the general technical vocabulary dictionary, and the computer science and programming dictionary. The word combination dictionaries are also subdivided according to subject fields. The target Russian language dictionary contains about 35 thousand dictionary entries. Translation is performed with the aid of tables matching unambiguous English words with the Russian translation equivalents. The tables are arranged according to the subject fields of the source dictionary. The ambiguous words are translated using specific algorithms that establish this or that translation of a word by analyzing its content usage[14,15]. Updating of available dictionaries and creation of new dictionaries is a routine process.

The translated text is submitted to the specialized Editing Board for post-editing. The edited text is delivered to the customer at his discretion either as a type-written copy, the line printer listing or on a magnetic tape. The quality of translation is such that it can be understood by a specialist. This allows us to deliver unedited text to the customer in most cases as preliminary pilot information to speed up its use. The volume of literature being translated totals to few hundred signatures per year.

Presently, the circle of customers is constantly expanding. Translation of huge volumes of text in an industrial environment serves as a good updating source and a system dictionary enlargement source. Work is being carried out on describing new sublanguages to extend the subject fields of texts translated. Practice has shown that to cover satisfactorily the texts of a sublanguage, it is sufficient to supplement the system dictionary with 4-5 thousand Russian and English lexical units, and the word combination dictionary with 5-6 thousand dictionary entries. Eight scientific/research workers can manage this job in 3-4 months.

Aside from the quantitative enlargement of the dictionary files, routine work is being carried out on qualitative improvement and upgrading of the system. AMPAR-2 is being created in which the entries in the source dictionary will be based on a wider use of semantic and word combination properties of a word. The system also provides for the perfection of the syntactical analysis and therefore the translation quality will be enhanced. The ideas used as a foundation for the AMPAR system support and implemented in a

model of translation correspondence ensures perfection of the system without any dramatic changes of its framework and in such a manner that additions do not impact the system performance which sometimes may occur in the running systems.

The NERPA translation system (Automated Machine Translation from German in to Russian) employs the same linguistic and programming principles used by the AMPAR system. The specific features of the German grammar have been incorporated into algorithms of the system[16].

In particular, the morphological analysis based on a relatively broad system of inflection plays a considerable role in the system. At the same time a broad homonymy of inflections caused the development of a special index system allowing the avoidance of the homonymy of stems.

The NERPA system features the morphological (word formative) analysis of the words that have not been found in the dictionary. Such words are then synthesized in Russian in the form of an artificial word with a regular suffix attached to the stem available in the dictionary (*"iskanie" - search, *"pokrevanie" - coverage). The NERPA system differs from the AMPAR system mainly in the analysis of composite words so typical for the German language. The general principle is disintegration of the composite words into the component stems followed by their synthesis in the form of a Russian word combination. For example, "Informationsverarbeitung" is translated as *"obrabotka informatsii" (data processing). The composite word analysis reduces significantly the volume of the German and Russian dictionaries. In contrast to 50-60 thousand words required in case the sophisticated analysis were not available, the dictionary is reduced to 10-15 thousand words only. Since the amount of composite words in German texts is practically unlimited, the composite analysis stage is rather important for the system and represents its characteristic feature.

As compared to the AMPAR system the NERPA system widely uses the semantic and syntactical codes which are employed in part due to the necessity to distinguish the syntactical homonymy but mostly due to the increased number of semantic classes. For instance, nouns may fall into classes denoting space, animals, organizations, artifacts, quality, processes, etc. The differentiation of semantic classes facilitates selection of a Russian equivalent for the ambiguous German words. Thus, the German word "Seite" will be translated as *"stranitsa" (page) is accompanied by words with semantics of a number, or otherwise as *"storona" (side). "Ausstellung" is translated as *"ustanovka" (installation) provided that the neighbouring word has the semantics of the artifact, otherwise it is translated as *"sostavlenie" (putting together).

The NERPA system has been recently put into experimental industrial operation. The number of translated texts is small. Presently, efforts are being exerted to enlarge the dictionary, update the files and expand the subject fields. The main engineering field of system application is programming and computer science.

Both the AMPAR and NERPA systems have unified software. The software features the following:

- division of the translation process into a number of stages;
- subdivision of the processing files (dictionaries, schemes, tables) into subject field subfiles;
- use of a specialized programming language alongside the Joint Computer System Assembler;
- use of a language support (Process Control Language) to specify input/output instructions for files handled and modes of handling;
- capabilities of the system structure reorganization (creation of various versions to select the most efficient system version);
- capabilities of obtaining the results of system operation at any stage in the form convenient for analysis in the verification mode.

While creating the software system complex, the modular structure concept has been employed whenever individual problems are solved by the stage programs, each consisting of program modules. The modular structure concept also pertains to information files (dictionaries, tables, etc.).

Since modules are relatively independent, it is possible to modify programs and information files in a comparatively simple manner by developing and including new modules or changing the sequence of their operation.

Great attention has been paid throughout system development and operation to the questions involving the linguist's efficiency in handling the system. As a result the system's linguist can:

- directly participate in creation and debugging of the programs (schemes) that implement the specific algorithms for processing compound word combinations, for translation of ambiguous words or resolution of homonymy and analysis, i.e. participate in those stages which are most likely to change when the system capabilities are expanded (a specialized language has been developed to simplify the process of programming and updating);
- obtain information about the words which are not available in the system dictionaries and about the typical errors that occur in the translation process;
- quickly localize an error when translating and determine its nature (selective printout of the system operation

results is employed at any assigned section of the text providing highly detailed information with an accuracy reflecting functions of an individual scheme operand);

- without hindering system operation, create various versions of the system. Each version can include new and/or modified schemes or a modified order of their operation; verify operation of the created version using a wide spectrum of texts to select the most efficient version and to include it into the work file as a work version;
- trace the state of information files.

In both the AMPAR and NERPA systems, the operator communicates with the work and service routines in the process of their functioning and sets their operation modes by means of instructions written in a special process control language. This language contains a set of directives. Each directive causes the module to perform this or that operation.

The NERPA program complex differs from that of the AMPAR system in routine and information modules developed additionally to take into account the specific nature of translation from German in to Russian.

The FRAP machine translation system (Automated Machine Translation from French into Russian) operates using somewhat different principles[10,18] whose essence lies in the explicit use of the semantic level, and producing semantic translation with validity checks on a contextual level, but not on that of translated correspondences. In the first version of the system (1976-1980), the main attention was focused on development of the linguistic support: linguistic structures of various levels - morphological, syntactical, semantical, grammar and algorithmic complexes. The existing version of the system proves the validity of the chosen linguistic ideology that, at any given moment during translation, ensures availability of information on all levels which have been built up to that moment.

The software development immediately follows the linguistic support.

The linguistic support is not adjusted to a particular subject matter. The main dictionaries are formed on the basis of the unspecialized common vocabulary. The terminological dictionary covers three subject matters: electronics, computer science, aviation and aircraft construction.

The FRAP system has a modular structure. Its subdivision of the highest level is subdivision into dictionaries and grammars, or, to be precise, into a dictionary complex and grammatical and algorithmic complex. Each complex extends to the following levels:

- analysis: graphemic, morphological, syntactical and semantic;
- translation of significant lexemes, relational words, syntactical links, grammatical classes, pronouns;
- synthesis: semantic, syntactical and morphological.

The system operates in several modes. The first one is auxiliary word-by-word translation. This mode enables us to check the main system dictionaries, the source French morphological and syntactical and target Russian morphological and grammatical dictionaries.

The second one, the principal mode, includes the syntactical component which references the semantic component to verify the meaning of links and translated equivalents. Translation as such will be performed through syntactical representation of individual sentences, that is why the mode is called a syntactical one. The third mode is semantic or textual and semantic one; it is unavailable in the current system version but it can be implemented in the next system version only. This is translation through the semantic representation, which may be accompanied by compression, semantic editing of the text contents. Finally, the fourth mode is the informal one, which assumes selective dissemination of information to the customer. The system must ensure translation of only those text extracts which meet informational requirements of the customer.

In the current FRAP version, the interface between the syntactical and semantic components has been mastered. Thus, the sentence is described in terms of two representations: syntactical and semantic ones. These representations interact as follows:

1. realization in the semantic representation of those dictionary-covered word meanings which correspond to the given syntactical representation;
2. rejection of some doubtful links found in the syntactical representation on the basis of the semantic representation .

Great attention is paid to improvement of software to achieve such flexibility which is required for the adequate simulation of operations concerning structure transformations in the course of machine translation. Programming and debugging of entire cycle of syntactical analysis using the PL/I language proved to be too labour-consuming and practically unconvergent process. A decision has been made to change over to a more dynamic programming language in which program development and debugging can be performed by linguists themselves. This language is a variant of the standard statement language developed for the AMPAR system.

In the FRAP system, four different machine data representations are employed as follows:

- pre-syntactical level representation which uses information in the simplest form;
- most consistent and system-organized representation for the syntactical stage, which is phrase-oriented;
- text-oriented representation for the semantic stage. It slightly differs from the previous one in that it has a larger depth due to semantic information;
- representation for synthesis. It is phrase-oriented and word-form oriented and can be reduced to the second and third representations.

It is assumed that the FRAP system will allow you to achieve more convenient detection of translation units and thus will provide improved quality of translation.

In the USSR Translation Centre, machine translation systems are also worked out together with automated dictionary designed to assist a human translator and editor.

At the moment, this dictionary contains English, German, French and Hungarian lexical files and is oriented to computer science and aviation fields.

By order of the Kazakh Academy of Sciences, the Chikament Pedagogical Institute provides lexical industrial translation of British and American texts on chemistry and polymers[19]. The initial stage of the system was creation of the automated dictionary of word forms and turns of speech. This dictionary is oriented to a limited class of documents. The main criterion for selecting lexical items for the dictionary is a systematic approach (their place in the terminological system) and their frequency. This translation is called by its authors as word-by-word, turn-by-turn one; several years have already passed in a complete satisfaction of the customers.

Materials are also published on development of the microcomputer-translator. This problem is rather related to the automated dictionary, however, one may assert that real industrial machine translation of nearest years will be invariably and most closely connected, with the questions of the automated dictionary since the main problems of such translation are lexical ones. In essence, in both cases we refer to computer-assisted translation; the only difference is that the machine translation system takes upon itself a larger part of man's work (at least as it is planned), whereas the computer plays purely auxiliary role in case of the automated dictionary.

Among papers devoted to micro-translators, it is important to mention research carried out in the Minsk Institute of Foreign languages. Extralinguistical and linguistical components are being developed. These components are the heart of the Data Bank of the microcomputer translator performing translation of conversational cliches of a

colloquial nature. 4000 pairs of English and Russian parallelly translated colloquial cliches have been selected. These represent the frequent colloquial cliches used in stereotyped situations of communication in towns and cities. The Data Bank includes the following blocks:

1. Russian-English and English-Russian colloquial cliche dictionaries;
2. bi-lingual microdictionaries servicing individual cliches;
3. bi-lingual subject field dictionaries servicing colloquial cliches for different situations of one subject field;
4. the common bi-lingual dictionary servicing all subject fields, situations and cliches [20].

Such is the development of operating MT systems which have been described in publications from 1980 up to the present moment.

THE INFORMATIONAL ASPECT

The informational aspects of machine translation presently becomes quite apparent after several years of experimental and industrial operation of the system. It is closely interconnected with the economical characteristic of MT. Only in those cases when sufficiently homogeneous large bulks of texts having the specific form and number of translated texts are large at the input of the system, the economic efficiency of this new kind of informational product appears. The great economic-technological role played by the quality of translation and the volume of post-editing (or inter-editing, or pre-editing). The experience gained by the USSR Translation Centre shows that various texts requiring different degrees of editing pass through the system. Some customers are satisfied with a rough, practically unedited machine product. In other cases rather profound post-editing is required, which nullifies practically the advantages of MT and levels it with manual translation [21].

The USSR Translation Centre has a certain contingent of super-numary editors who take upon themselves editing of the machine product.

The differentiated use of MT is possible with an economic in the integrated scientific-technical information systems where all forms of service, including machine translation as pilot information, diverse kinds of MT, which differ in the required editing depth, are defined, such texts which are not translated by the machine are assigned, etc. However, it is clear that many such kinds of service have been settled, they have long-term traditions and are not

going to give way to any other product in the scientific and technical information system. As an example, one may take the abstracting activity. There are special services which carry out this activity, and the replacement of a traditional abstract by the pilot machine translation may hardly be smooth both for the customer and for the abstracting service. These and other questions of including MT into the traditional network for servicing information for scientific and technical development support are due to be solved and they are being solved nowadays.

The USSR Translation Centre, within its technological scheme of translation processing, as a head organization in the scientific and technical field in the USSR and in international Information Service "INTERINFORMPEREVOD", is now creating a stock of machine translations which are edited and those which have been ordered in the unedited form by the customer. This is a special stock. Its use is less expensive. In contrast to the usual translation stock where translations are stored in the form of hard copies, the stock of machine translation is stored on magnetic media.

Materials from this stock can be delivered in any form on orders of the customers. The USSR Translation Centre has already gained the experience in using this stock for informational servicing of customers.

CONCLUSION

Summarizing everything said, it is possible to note that MT in the USSR is perhaps developing rather slowly but in an undeviating manner because other ways to overcome increasing language barriers in the USSR, as well as all over the world, are available.

REFERENCES

1. Z. M. Shalyapina.
Automatic Translation as a Model of the Human Translation Activity.
Int. Forum Inf. Doc. 1980, Vol.5, No 2, pp.18-23

2. A. F. Shirayaev.
O psikholingvisticheskikh osobennostyakh funktsionalnoi sistemy sinkhronnogo perevoda.
V kn. Perevod kak lingvististicheskaya problema.
Sb. st. (Redkol.:Garbovskii N.K. (otvetstvennyi redakte) i dr.-M.,Izd-vo Mosk.un-ta, -120 s.

3. Yu. N. Marchuk.
Blok-skhema deistvii perevodochika.
V kn. Urovni teksta i metody ego lingvisticheskogo analiza.
AN SSSR-VTSP, M., s.21-31

4. N. D. Andreev.
Statistiko-kombinatornye metody v teoreticheskom y prikladnom yazykoznanii.
L,Nauka, 1967, 248 s.

5. L. L. Nelyubin.
Perevod i prikladnya lingvistika.
M. Vysshaya shkola, 1983 207 s.

6. N. Ya. Serdobintev,
Semanticheskaya struktura slova i ego konnotatsii.
V kn. Teorya slova i funktsionirovaniya slovarnych edinits.
Saratov, 1981, s.4-25.

7. R.A. Budagov.
Chelovek i ego yazyk.
M.,Izd-vo MGU, 1974, 261 s.

8. See, for example, V.L. Tyulina.
Izuchenie vzaimodeistvya terminologicheskoi i obshcheupotrebitelnoy leskiki kak lingvoperevodcheskaya problema.
Tezisy dokladov i soobshchenii zonalnoi konferentsii "Problemy perevoda nauchno-technicheskoi literatury i prepodavaniya inostrannych yazykov v technicheskikh vazuch". Yaroslavl, 1983, str. 123-125.

9. V. P. Neroznak, S. L. Nikogosov.
Kontrastivnaya lingvistika i problemy perevoda.
See ditto, s.223-224.

10. N. N. Leonteva, S. L. Nikogosov.
Sistema FRAP i problema otsenki kachestva avtomaticheskogo perevoda.
V kn. Machinnyi perevod i prikladnya lingvistika.
Vyp. 20., M, MGPIIYa im. M.Toreza, 1960, str.57-78.

11. Yu. D. Apresyan.
Tolkovo-kombinatornyi slovar v formalnoi modeli yazyka i razeshenie mnogoznachnosti.
V kn. Aktualnye voprosy prakticheskoi realizatsii sistem avtomaticheskogo perevoda. M., MGU; 1982, str.108-128.

12. A. G. Dorfman, V. M. Sergeev.
Morfogenez i skrytaya smyslovaya struktura tekstov. V kn. Voprosy kibernetiki. Logika rassuzhdenii i ee modelirovanie. M., An SSSR, 1983, str. 137-147.
Phrases marked "*" were transliterated by the editor.

13. Yu. N. Marchuk.
Problemy machinnogo perevoda.
M., Nauka, 1983, 233 s.

14. Kontekstologicheskii slovar dlya mashinnogo perevoda mnogoznachnykh slov s angliiskogo yazyka na Russkii. Sostavitel. Yu. N. Marchuk. M., VTsP, 1976, part 1-264, part II - 256 s.

15. Yu. N. Marchuk.
The Contextological Dictionary: Use in Programmed Language Teaching. Computers and the Humanities.
New York 1979, v. 13, No 4, pp 277-281.

16. Yu. N. Marchuk, A.N. Vlasov.
Nekotorye printsipy avtomatizatsii perevoda s nemetskcgo yazyka na russkii.
Zeitschrift "Fremsprachen", Berlin, 1960, No 2, str. 91-99.

17. The AMPAR system standard statement language is described in
Yu. N. Marchuk, B. D. Tikhomirov, V. I. Scerbinin.
Ein System zur maschinellen Ubersetzung aus dem Englischen ins Russische.
In: Automatische Sprachubersetzung. Darmstadt, 1982, Wissenschaftliche Buchgesellschaft, 1982, ss 31-336.
It has been slightly modified since then.

18. E. G. Sokolova, I. M. Kudryasheva.
Komponent sintaksicheskogo analiza v sisteme FRAP.
Sb. Aktualnye voprosy prakticheskoi realizatsii sistem avtomaticheskogo perevoda. M., MGU, 1982, part 1, str. 167-183.

19. P. A. Abdullaeva, L. P. Prikhodko.
Printsipy otbora leksiki v AS dlya MP patentov.
Materialy seminara "Statisticheskaya optimizatsiya prepodavaniya yazykov i inzhenernya lingvistika".
Chikment, 1980, str.38039.

20. A. V. Zubov, O. A. Nekhay, L. I. Rtibis.
Ekstralingvisticheskie i lingvisticheskie komponenty banka dannych mikrokompyuterperevodchisk.
V kn. Problemy vnutrenney dinamiki rechevykh norm.
Sb. statei. Minsk, 1982, s.191-196.

21. B. D. Tikhomirov.
Osobennosti ispolzovaniya mashinnogo perevoda v sisteme nauchno-tekhnicheskoi informatsii.
Materialy seminara "Statisticheskaya optimizatsiya prepodavniya yazykov i inzhenernaya lingvistika",
Chikment, 1980, str 132-134.

22. R. G. Kotov, Yu. N. Marchuk, L. L. Nelyubin.

Mashinnyi perevod v nachale 80-tych godov. Voprosy yazykoznaniya, M., Nauka, 1983, No 1, str. 31-38.

23. R. G. Piotrovskiy, N. P. Rakhubo, M. S. Hazhinskaya. Sistemnoe issledovanie leksis nauchnogo texsta. Kishenev, Shtiintsa, 1981, 156 s.