Machine Translation from Hungarian to Russian

I.A. Vel'chuk
(Moscow)

An experimental version of the rules for the machine translation of scientific texts from Hungarian to Russian is of practical value and is also of more general interest.

Firstly, certain of the structural characteristics of the Hungarian language relate it to a number of other languages (it is agglutinative like the Turkic languages; like the Germanic languages it makes wide use of compound words and separate prefixes; the function of prepositions is fulfilled by postpositions, and so on). Therefore, to overcome these peculiarly Hungarian difficulties will make it possible to solve many problems of machine translation from other languages similar, in these respects, to Hungarian.

Secondly, due to the fact that word order in Hungarian differs very greatly from the usual word order in Russian (so that a word-for-word translation can be entirely incomprehensible due solely to incorrect word arrangement), a special set of rules to convert the word order in the Hungarian sentence obtained after translation has to be worked out. This is essentially an attempt to edit the translated text somehow, and its realization will also be of value in machine translation from other languages.

Despite all the special features of Hungarian there are several aspects of machine translation from Hungarian into Russian characteristic not only of Hungarian but also of other languages (general principles of word scanning, the problem of phrasenology, synthesis of the Russian sentence, and so on). Some of these characteristics have been described in the article by V.I. Kolosova and I.A. Vel'chuk [1].

References to the appropriate sections of this article, referred to as [1], have made it possible to shorten the present work somewhat.

The following data are used for machine translation from Hungarian:

1. "dictionary of stems"
2. "dictionary of usages"
3. "tables for the translation of postpositions" ("governing tables")
(4) "tables of Hungarian suffixes" and "tables of Russian endings"
(5) "list of characteristics"
(6) "scanning rules"
(7) "rules for the differentiation of homonyms"
(8) "rules for sentence casting"
(9) "group of analyzing rules"
(10) "group of synthesizing rules"
(11) "rules for the editing of the Russian sentence".

Not all this information is fully considered in the present article. We consider, however, extracts from the "dictionary of stems", the "dictionary of usages" and the "list of characteristics" several "governing tables" as examples, full "tables of Hungarian suffixes", "scanning rules", "rules for the differentiation of homonyms", and all the "analyzing rules".

The "rules for sentence casting", "tables of Russian endings", "synthesizing rules", and "rules for the editing of the Russian sentence", which constitute a separate article (a continuation of the present one), are not considered. Instead of considering these groups of rules fully, necessary explanations are given in the appropriate places; together with the topics which are treated in full, they will allow the reader who is unfamiliar with the Hungarian language to make sample translations by assuming, as it were, the role of a computer. To this end excerpts of Hungarian grammar are appended to the article, which, along with the other examples used, were taken from two books by Csejtey Baross: Magyar szókincs eredete, Budapest, 1951 (referred to as B.S.E. in this article), and Bevezetés a néprajzban, Budapest, 1953 (B.N. in this article).

The following signs and symbols are used in this work:

<table>
<thead>
<tr>
<th>PART OF SPEECH</th>
<th>HUNGARIAN</th>
<th>RUSSIAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noun</td>
<td>E</td>
<td>K</td>
</tr>
<tr>
<td>Adjective</td>
<td>A</td>
<td>F</td>
</tr>
<tr>
<td>Adverb</td>
<td>AŞ {</td>
<td>I</td>
</tr>
<tr>
<td>Adverb-parenthesis</td>
<td>A{</td>
<td></td>
</tr>
<tr>
<td>Verb</td>
<td>Y</td>
<td>G</td>
</tr>
<tr>
<td>Past Participle</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td>Dative number</td>
<td>D</td>
<td>O</td>
</tr>
<tr>
<td>Preposition</td>
<td>Ps</td>
<td>Prep</td>
</tr>
<tr>
<td>Co-ordinate conjunction</td>
<td>Oonj</td>
<td></td>
</tr>
<tr>
<td>Subordinate conjunction</td>
<td>Osub</td>
<td></td>
</tr>
</tbody>
</table>

Simple symbols:

- Front row
- Back row
- 1st inf-verb
- 2nd suffix
- Possessive suffix
- Comp. or translation in compound word
- Indicating words of the 1st and 2nd type
primes on the upper right of letters signify corresponding translatable and translated stems or similarity of two words (the sign A' written in different places means a single specific adjective and not adjective in general; S' and M' or S'' and M'' mean that M' is the translation of S'' of S';
= "is translated"
- succession and unbroken writing of parts of words
(\text{the sign "}_\text{" means that it must be followed by others})
+ succession and separate writing of words
() brackets contain an optional part which may or may not be used
\sim arrows indicate dependence between words (A' \supset S; \text{"A" depends on S")}.

1. Dictionary of Stems

A small Hungarian-Russian dictionary based on a reading of the required text was prepared for the experimental version of machine translation from Hungarian. This dictionary contains stems (both Hungarian and Russian) as well as words. "Stem" is taken to mean that part of a word which remains unchanged throughout in all the variations of the word that may be written.

Since Hungarian is an agglutinative language, its stems do not vary as a rule.\footnote{\text{\textit{\text{\textquoteleft\textquoteleft}}}}

In the relatively few cases where there are permutations or suppletives in the stems, the dictionary gives both stems, for example: szív - sző and jörgalás - jõrgalás, jusz - jút; volás - lenz (cf. volat - lenen), ebb - azz, arr - azz, enns - enns (i.e., abban, arra, ennes) and so on.

However, since all variations in stem conform to three simple types\footnote{\text{\textit{\textit{\textquoteleft\textquoteleft}}}} (abbreviation or loss of letter: \text{\textit{\textquoteleft\textquoteleft}}, \text{\textit{\textquoteleft\textquoteleft}}, \text{\textit{\textquoteleft\textquoteleft}}, zef - zeff, vész - vész/\textit{\textquoteleft\textquoteleft}, natural permutation: \textit{\textit{\textquoteleft\textquoteleft}}, \textit{\textit{\textquoteleft\textquoteleft}}, \textit{\textit{\textquoteleft\textquoteleft}}, \textit{\textit{\textquoteleft\textquoteleft}}, accumulating or \textit{\textit{\textquoteleft\textquoteleft}}, no - nomina), another solution is possible: to give one stem in the dictionary, but make special rules for the transformation of stems (a similar technique has been employed to take into account permutations in noun and verb suffixes).

The Hungarian section of the "machine" dictionary is almost the same as ordinary Hungarian dictionaries in which word stems are also given (the initial or absolute form of the noun, and the third person singular intraverbative present, indicative of the verb), and unvarying words are given in their entirety.\footnote{\text{\textit{\textit{\textquoteleft\textquoteleft}}}} However, the Russian section of the dictionary is quite unusual: nouns and adjectives are given in the form of stems without the nominative ending (for example, romantikus - romantik, task - task, evet, nischnik - nischnik, etc.), and verbs are also given in the form of stems (nece, nece, nece, nece, etc.).

The stems in our dictionary are arranged in the usual manner, alphabetically; but a short stem is found after all
the longer ones in which it is wholly contained, i.e., as stands after azoncan, and as alter namer, naml, mat, etc. This arrangement is found to be more convenient for scanning (see [1])."

Each stem in the dictionary has a dictionary article which contains

1. Translation
2. Hungarian information -
   a. instruction as to part of speech
   b. grammatical characteristics
   c. instruction as to usage
   d. postpositional code ("governing group number").
3. Russian information -
   a. instruction as to translation stem selection
   b. grammatical characteristics.
4. Additional instructions.

Let us consider the sections of the dictionary article separately.

1. Translation. Usually the translation of a Hungarian word consists of one or several stems of the same translated word (see [1]).

In some cases two translated words are given, the selection of one or none not being definitely stipulated and depending only on the sense of the sentence. These are cases of dictionary homonymy (hajás: "walking" = "region" and szem- keres: "design" = "structure"), polysemes differing in no way from dictionary homonymy from the "machine" point of view. In the corresponding Russian sentence both translations will be given (one in brackets), since only the formal relations of words and not their connections in meaning are used in translating.

However, in future, a machine may be "taught" to select the necessary translation from several possible choices, beginning with the analysis of word groups. For example, a list of words could be supplied according to which sequences would be rendered by "structure" (for example,chemy, kondai, etc.)

For this a vast amount of time-consuming work must be carried out on a statistical study of the most diverse and generally-used combinations of words (in all branches of learning).

The following should be noted:

1. The elongation sign, is considered as a separate letter, and stands first in the alphabet, i.e., a = a, a = a, etc.
2. sz, ps, ey and other ligatures are each considered as a separate letter, so that as follows c, as after s, etc.
3. cos, sus, ey, any and other doubled ligatures are considered as ordinary ligatures; cos = cos, sus = sus, etc.
MACHINE TRANSLATION FROM HUNGARIAN TO RUSSIAN

(2) Hungarian information. This comprises:

(a) **Instruction as to part of speech.** Each word stem in the dictionary has a mark indicating what part of speech it is (see [1]).

(b) **Grammatical characteristics are reduced to the following instructions:** for a noun - Row (back or front); for a verb - Ak conjunction or not; for postpositions - Row also (so that rule III operates correctly).

(c) **Instruction as to usage.** The treatment of word order, which must not be translated literally but as a whole (for example, a verba tantum rendered on nepszaboassz "I was one who was present"), whereas a literal translation would be om el a nepszoassz ("the was one who was present"), does not have any characteristic rules in Hungarian. The reader is referred to the previously mentioned article [2], **Instructions on usage (H.mz.)**.

In the Appendix a specimen usage list is given for words in our dictionary.

Before explaining what is meant by postpositional code or governing group, the matter of Hungarian postpositions should be considered.

In Hungarian grammar the term postposition (rögzítő) means a subordinate word not having an independent dictionary meaning, and designating the syntactical relations of nouns and pronouns also to other words in the sentence [2]. Postpositions, generally speaking, correspond to Russian prepositions with the difference that they always stand directly after the noun they qualify.

A function identical to that of postposition is fulfilled in Hungarian by the so-called case endings or suffixes (Poz), which will in future be referred to as noun suffixes; suffixes form a chain of suffixes (personal, possessive and plural) indicators, which constitute a word ending (the part remaining after the stem is removed).

For convenience of analysis and translation be combined the Hungarian postpositions and suffixes into the general class of postpositions falling into the groups: postpositions proper (hiteposz) are called type I postpositions, and suffixes - type II postpositions. In accordance with this, the translation of any postposition is made according to the general rules for the translation of postpositions (see below).

However, type II postpositions (i.e., suffixes or noun suffixes) vary; they can be subdivided into three groups (see below).

Type I suffixes. These are suffixes which do not have the syntactical role of the noun, but in themselves require translation, but supply grammatical information for the translation of the word as a whole. Here it is a question of

* Let us stress once again the conditional nature of the text, which was selected only for the sake of analogy with Russian prepositions.
instructions which indicate whether the word is plural, personal-possessive and conditionally divided, but which do not have a direct meaning like -t (from -te, etc.) and'
(the result of the expansion of the terminal vowel stem before any suffix). The distinguishing feature of Type I suffixes
is that they can be followed by other suffixes.

Type II suffixes. These are, first, variants of the
accusative case formation. This formation has one meaning.
It usually requires that the translated word be put into the
accusative (or genitive after a negative verb); it is not
itself translated. Other suffixes cannot follow it.

Secondly, the -j/-g suffix which is not actually a case
formation is conditionally included in this group; this is a
very common relative denominative adjective suffix which can
adjectivize whole word-groups: caskele mereté/мелький масштаб
['a small scale'], caskele merete/терге/маломасштабная
карта, карта мелкого масштаба ['a small-scaled map, a map of
small scale']. Similar combinations are usually translated
by a genitive noun*, for instance (1) ... allerekf irányú
éret (Rev., p. ...) becomes сила противоположного направления
['a force of (moving in) the opposite direction']. It is
therefore convenient to consider -j/-g as a genitive suffix.
It has one meaning; it cannot be followed by other suffixes.

Type III suffixes. These are all other case formations.
Like Type II suffixes they can stand only at the end of a
series of suffixes, and unlike Type I they are usually translated by separate words - by
Russian prepositions. Moreover, the Type III suffix is not
rendered by a single translation; each has several transla-
tions, the selection of the required one being determined for
the most part by the governing verb (or other word), i.e.,
translation of Type III Hungarian suffixes (case formations)
is in general similar to that of prepositions in translating
languages like English or French to Russian.

Type I postpositions are entered in the dictionary as
ordinary words. Type II postpositions (case formations),
which are written in the same word with nouns, are entered in
the noun suffix table. They are also placed in the dictionary
so that their translation may be found when they form the case
endings of personal pronouns and appear as separate words
formed with the possessive noun suffixes (for example: razv/
vo. ['part']; razv-tny or вракти ['from part']; razv-e
ero vo. ['his part']; вар- by his hero ['from him'],
literally ero ['his hero']). These case formations, which
in combining with possessive suffixes change their form, are
found in the dictionary in precisely this altered form (e.g.
instead of -pe/-se; razi instead of -ono/-en; end instead of
-ơan/-ơen, and so forth).

Now we can consider the postpositional case.

(d) Postpositional case. Translation of Type II post-
positions (i.e., case formations or suffixes) is carried out

*In addition to compound words like космос/космический,
глазастый ['two-eyed, double-eyed'], ажурный/кружевный
['blue-eyed'], etc., which are entered in
separate articles in the dictionary.
with the aid of special Postpositional Translation Tables, otherwise called Governing Tables*. Each table contains a list of the case formations which can govern** any one word, and each case formation is rendered by one translation (Russian preposition) in accordance with the translation of the governing word; in this instructions are given for the case required by the preposition.

The postpositional code of a word is the number of that table in the Governing Tables which contains Type II postpositions (case formations) governed by the given word, and which gives precisely those translations which are required in conjunction with the given word. For instance, the verb beszél/говорит (to speak) is found in code 12 in the dictionary.

Governing Table No.12 is as follows:

- ról/-réi O [about] + N (prepositional)
- pan/-nek N (dative)
- ban/-ben b [in] + N (prepositional)

This means that beszél van/говорит must be translated as говорит кому-либо [to speak to someone], beszél van/говорит o чем-либо [to speak about something], and so on.

The postpositional code contains almost all verbs and many adjectives (for example, jellemző van/характерный для чего-либо [characteristic of something], le van/пополнён чем-либо [full of something]), and also some nouns: kereskedé van/связь с чем-либо [connexion with something].

If several words are governed by the same postpositions so that these postpositions are translated identically in the translation of all these governing words, they are all then rendered by the general postpositional code. For example, the verbs cseréz/обменив [to exchange] and elhúz/обнаружить [to designate] are under general code 3. Table 3 contains the Type II suffix -vel/vel for which the translation N (instrumental) is supplied, i.e., after the verbs cseréz and elhúz a noun with the -vel/vel suffix is translated in the instrumental case.

In conclusion let us note that the term "governing" is considered in a broader sense than is ordinarily the case in grammar: Governing tables include not only case formations which depend on a given verb and which cannot be translated before translation of the latter, but also those postpositions which, although they depend on the verb, still (in a given case) have their own basic translation — according to the noun suffix table — or are in general translated identically (for instance, the accusative case formation). The Governing Tables contain even those postpositions which in general are not grammatically connected with the verb (those that are assigned to the part of speech governed) and are translated independently from it, but which through the dictionary meaning of the verb are quite frequently attendant to it. For

* Type I postpositions are rendered essentially by synonyms and their translations are given in the dictionary.
** See below for meaning of terms "governing" and "govern".
example, the governing table for the verbs "all create" {to stand}, "mark to remain" and "spread/
reproportion" {to spread}, includes the postpositions
"ban"r/ber and "en". These postpositions do not depend on
the above-mentioned verbs; generally speaking, they accompany any verb. But they do accompany with particular
frequency precisely such verbs as "all" and "mark", which by
their dictionary meaning are particularly disposed to have
accompanying adverbs of place. "ban"r/ber and "en" are
formed exactly like adverbs of place.

These governing tables are not simply necessary for post-
position translation. Many of the postpositions they contain
are translated without the aid of these tables. But it is
through the governing tables that the machine, knowing all the
postpositions that govern (in our sense of the word) a given
verb or other word, can, with such information, in the dictionary,
the postpositional code is given (the number of the corre-
sponding governing table). This makes it possible to establish
the formal connexions of the words in a sentence, which is
necessary in the analysis and machine "writing" of a transla-
tion (see p. 257), since the formal connexions of words
reflect to a varying degree their relations in meaning.

Each governing table can accompany several verbs (see
govern). Therefore all the verbs in our dictionary can be
divided into groups so that one group will contain only those
verbs for which one governing table is suitable, i.e., which
require the same postpositions and are in this sense equivalent
to one another. The postpositional code of a verb is
essentially its group number.

We will not describe in detail the method of dividing verbs
into equivalence groups. It should only be noted that the
solution of this problem is of great theoretical importance
in linguistics, since it is connected with the important
question of the division of classes of words.

Each dictionary article contains, in addition:

(1) Russian information, consisting of instructions regarding
the selection of the necessary stem of the translated
word and its grammatical characteristics, for example, gender
and declension in nouns; see (1)(, etc.

(2) Additional instructions concerning the characteristics
of individual words for translation. If several words have
the same characteristic, they are given a general "characteris-
tic number" - a reference to the list of characteristics
where instructions are immediately given for all these
words. A common characteristic number in our dictionary con-
tains, for example, the words "earth", "mountain", etc. "Ashes",
"leaves", "water", "head", and names of other nationalities, and
so on.

An additional instruction countermands information received
by other means.

2. RULES OF TRANSLATION

The machine translation process consists of the following
operations:

(a) Scanning for word in dictionary and tables, and its
transmission together with information found to the operative storage section.
(2) Processing of usages.
(3) Homonym differentiation.
(4) Sentence casting.
(5) Analysis.
(6) Synthesis.
(7) Editing of Russian sentence.

Below we will consider all these operations in succession.

1. General Rules of Translation
First we will formulate the general rules of translation.

1. Translation is made sentence by sentence (a sentence is defined as a section of the text enclosed by two periods*).

2. (a) Scanning within a sentence takes place to the left and right, not passing subordinate conjunctions; special provision is made for cases where scanning is restricted by other words as well.

(b) To scan for a word in the dictionary means to find the stem having the maximum number of letters which is contained as a whole in the word.

(c) Having found all the words in the sentence, usages are treated and homonyms are differentiated if there are the necessary instructions; having established the part of speech and grammatical characteristics of each word, attention is then turned to the pertinent tables and groups of rules.

3. Groups of numerals or other signs apart from letters, words written with a capital not after a full stop and not after quotation marks, examples printed in special type and so on, are conditionally called "formulae". Formulae are not translated but copied.

(a) Numerals are considered as cardinal numbers.

(b) Numerals followed by a period or hyphen and ending are considered as adjectives (since an ordinal number is syntactically equivalent to an adjective), but are not treated by the corresponding rules.

(c) Words written with a capital which do not follow a period or hyphens and are not found in the dictionary are considered as nouns; they also are not treated by the corresponding rules.

(d) Other formulae following directly after § stand after №.

4. In analysing any constructions, adjectives, participles and adverbs are not considered except in cases specially provided for, and personal pronouns are treated as nouns.

5. Generally speaking, parts of speech in Hungarian and

*The semi-colon, exclamation and question marks and pluri-stops have the same strength as a full stop.
Russian correspond; postpositions correspond to prepositions.

6. The following information cancels out the preceeding.

7. If two stems for a word are found in the dictionary (by rule II.10.10.0 or rule II.10.11.00.1), the word in question is compound.

8. The rule for the translation of compound words is:

A. The first stem is translated:

(a) if its dictionary article contains a translation marked "comp. wi.", this translation is selected;

(b) if the translation does not have this mark, the translation of the first stem in the genitive plural will stand after the translation of the second stem.

B. The second stem is translated by the usual rules.

v. A word ending in a hypen (for example, értelm- is the first stem of a compound word. The second stem nearest the right of the compound word is considered as its second stem.

Rule I.3.b is made for such cases as XVI század/XVI. zsidó [sixteenth century], rule I.3.c for different proper nouns, for example (2) A Garonne vizepnyeletében (Rev., p.24) A drugoj stropow / Garonne /district boundary, and rule I.3.d for the clause of the type (3) A lengyel szympasaz / elszigetelt a lenyelben (Nsz., p.28). Polish word szympan isolated stands in Polish, and (4) Más természetes asszíl - repülé - lenyelés stb. karosolata (Rev., p.7) -foxes not at all [The connexion asszíl - repülé - lenyelés is of a different nature].

Word scanning. The general principle of scanning in a machine dictionary - find word means find the stem with the maximum number of letters, which is contained as a whole in the word being translated, is also applied in translating from Hungarian. However, although, in the French and English versions, special scanning rules are not required (it is sufficient to find the word in the dictionary: the part found is the base and the rest the ending), the Hungarian version requires a special set of scanning rules.

The necessity of special "scanning rules" in machine translation from Hungarian is dictated by three characteristics of this language.

(1) The great prevalence of compound words containing two stems (compound words being written as one word). As a result, when the stem has been found in the dictionary and separated from the word, a check must be made to see whether it contains the remainder of another base. Rule II.10.10.0 is formulated for this end.
On fulfillment of these rules it is possible that the beginning of the second stem will by chance coincide with a suffix. Rule 75.10.11.0.0 prohibits incorrect separation of part of the second stem in this case.

(2) The agglutinative structure of Hungarian. As a result of agglutination, the "ending" (that which remains after one or two stems of a compound word have been removed) in Hungarian is not a single unit (as in English, French or Russian); it can consist of several suffixes each of which has a meaning. Full grammatical information on a word (noun and verb) includes all the information regarding the suffixes forming its ending.

Several rules are given for breaking an ending down into its constituent suffixes.

Theoretically it would be possible to avoid breaking down endings, and instead give tables of endings rather than suffix tables. But this is awkward in actual practice: the number of noun suffixes in the corresponding table is 63 (with all possible variations), and the total number of their combinations encountered in the language is 5-6 times greater.

Spoken permutations in Hungarian suffixes are also taken into account in the Scanning Rules (see below).

(3) The great prevalence in Hungarian of verbs with separate prefixes which can stand either directly after the verb or before it, separated from it by other words. The machine dictionary contains:

(a) simple verbs

(b) all separate prefixes

(c) those prefixed verbs in which prefix combination alters the meaning and consequently the translation: for example, segétség, képesség, etc. [to determine]. (Segétség means help to determine.)

(d) those prefixed verbs which are not used (or are seldom used) in the simple aspect: for example, megbízható, megbízhatatlan [to establish].

In a text it is possible to encounter a prefixed verb with a prefix where the latter has little effect on the meaning, so that the verb is entered in the dictionary only in its simple aspect (for instance, lepárt, behatol, megbízhatatlan); or a prefixed verb is sometimes set without a prefix (the latter being separate), which is entered in the dictionary only with the prefix. Rules 71.0.1 to 71.0.1 from line 11 to the end of the Scanning Rules are provided for such cases.

The Scanning Rules, like the Analyzing Rules, are based on the principle of binary selection, i.e., at each given stage in the translation the machine is faced with the prefixes of selecting one of two possibilities.

* In the Verb Translation Rules (V11) this is altered in order to facilitate statement of the rules and their usage.
For example let us consider Rule II.10.1: "There is a remainder - scan it in accordance with the suffix table". There is an alternative: either the suffix will not be found (Rule II.10.10) or it will be found (Rule II.10.11). The enumeration of the rules reflects the dichotomy: the two rules defined by an alternative have numbers which are differentiated only by the last digit: 0 for the answer "no" and 1 for the answer "yes".

Each rule consists of two parts: condition (cases where the rule in question is applied) and command (how to proceed in this case).

II. Scanning rules

A word is scanned by the following rules:

0. Word not found in dictionary (i.e., no one stem was found contained as a whole in the word) - check in the dictionary whether there is a prefix contained in the word being scanned.

00. No prefix found in dictionary contained in word being scanned - check whether there is a verbal prefix one word further to the right of the word being scanned or two words to the left.

00.0. No verbal prefix - cease scanning (the word is not to be found in the dictionary as it is only partial).

00.1. There is a verbal prefix - join it to the front of the word being scanned; in place of the latter substitute the word thus formed and scan dictionary again.

01. In the dictionary there is a prefix contained in the word being scanned - separate it and scan without it.

01.0. No word found after separating prefix - cease scanning (the word is not to be found in the dictionary as it is only partial).

01.1. Word found after separating prefix - take dictionary article on word found for translation.

1. Word found in dictionary (i.e., stem found contained as a whole in word being scanned) - check whether it is a verb.

10. Word found is not a verb - check whether there is a remainder*.

10.0. No remainder - cease scanning.

10.1. Remainder found - scan for it in appropriate suffix table**.

*"Remainder" is taken to mean that part of the word which remains after its original parts have been removed.
**"Find Remainder" means find it wholly or partially. "Find remainder in table wholly" means find a suffix to coincide with it fully. "Find remainder in table partially" means find a suffix contained in remainder, i.e., to coincide with its initial part.
10.10. Remainder as a whole not found in suffix table - substitute y for first letter of remainder and scan table again.

10.10.0. Remainder as a whole still not found after substituting first letter - replace first letter of remainder and scan for it in dictionary as a word according to full Scanning Rules.

10.10.1. Remainder found after substituting first letter - cease scanning and take information from tables.

10.11. Remainder found in suffix table - check whether wholly or partially.

10.11.0. Remainder found partially - separate initial part of remainder coinciding with one of the suffixes; scan table again for the second remainder obtained.

10.11.00. Second remainder as a whole not found in table - substitute y for its first letter and scan same table again.

10.11.00.0. Second remainder still not found after substitution of its first letter - replace first letter and rejoin it to the whole initial part; separate by rule 11.10.11.0. (i.e., replace the first remainder fully and scan the dictionary for it as a word in accordance with the full Scanning Rules).

10.11.00.1. Second remainder found after substitution of its first letter - cease scanning and take information from table.

10.11.01. Second remainder round in table - return to rule 11.10.11.

10.11.1. Remainder found as a whole - cease scanning and take information from table.

1. Word found is a verb - scan sentence for verbal prefix not further than the second word to the right or the third word to the left of the verb.

2. No verbal prefix in sentence - cease scanning, take dictionary article for word round by rule 11.13.0, check whether there is a remainder and return to rule 11.10.0 or 11.10.1.

3. Verbal prefix found in sentence - join it to the front of the verb and scan dictionary again.

4. Verb thus formed not found in dictionary - reject prefix, disregard it and carry out rule 11.10.0.

5. Verb thus formed found in dictionary - take its dictionary article (reject article for previously found
word), check whether there is a remainder and return to
rule II.10.0 or II.10.1.

In translating from French and English the dictionary is
first fully scanned, and then the appropriate tables are
scanned for the endings obtained. In dealing with Hungarian,
as we have seen, scanning occurs simultaneously in the dictio-

nary and in the tables, alternating from one to the other. It
is therefore necessary to describe the structure of the suf-
fix tables before considering examples of the operation of
the scanning rules.

In translating English and French the tables of endings are
compiled without particular difficulty, since the total num-
ber of endings in these languages is not great. The situa-
tion is different with regard to Hungarian. For instance,
the Short Hungarian Grammar by K. B. Szilárd [4] lists 119
different verb endings in the third person singular and
plural, and in the first person plural alone (the others do
not concern us, since verbs in the first person singular and
second person singular and plural are seldom encountered in
scientific writing). Such a large number is the result of
phonetic variation: semi-vowel synchronism (two or three
variants of each suffix: -ett, -ett, -ett, etc.) and permu-
tations 1/2, 1/2, 1/2 in verbs of different types. Therefore,
if we wish to include the minimum possible number of suffixes
in the table, a special Rule for the Treatment of Verbal Endings
must be used. With it all verbal suffixes are reduced to a
conditional "standard" form (back row semi-vowel; permute-
ing consonants are replaced by j). This makes it possible to
reduce their number to a quarter: there is thus a total of
25 suffixes in our table.

But a similar rule is inapplicable to noun suffixes: due
to the presence of compound nouns it is impossible to be sure
that the remainder obtained after the stem has been removed
is an ending and not another base. In order to be certain
about this it is necessary to find this ending in the table;
as a result it is impossible to modify it into a "standard"
form first. Noun suffixes are therefore set into parallel
columns in the table (with front row semi-vowel and with back
row semi-vowel).

Consonant permutation in noun suffixes is observed only in
the suffix -val/-val, which alters its first consonant to the
last consonant in the word stem and emerges in the form -bal/
-se, -dal/-dei, -val/-val, etc. Rules II.10.10 and
II.10.11.00 were introduced in order to handle these cases.

However, one of the variations of the -val/-val suffix,
i.e., -val/-val, is entered as a whole in the table. This
was done so as to avoid introducing a number of rules to pre-
vent -val/-val being broken down by rule II.10.0 into the
suffix -val, now -val/-val, which is semivowel. In the other
variants of the -val/-val suffix the initial letters do not
coincide with any of the suffixes in our table. A more
complete table would have to have either two variants of the
-val/-val suffix, i.e., -val/-val and -val/-val, or addition-

al rules for handling this suffix, since it contains no
first and second person possessive suffixes in -é and -a.

The suffixes in the tables are arranged alphabetically in
the same manner as the stems in the dictionary – the shorter suffix stands after the larger one in which it is contained as a whole. For example: -nál, then -nák and then -n (noun suffixes). For the machine the suffixes can be arranged differently, according to frequency, so that the most frequently encountered suffixes are entered first.

### Table 1

#### Noun suffixes

<table>
<thead>
<tr>
<th>Type I suffixes</th>
<th></th>
<th>MEANING</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. -' (--)</td>
<td></td>
<td>no meaning</td>
</tr>
<tr>
<td>2. -ak (--)</td>
<td>-ak (--)</td>
<td>plural</td>
</tr>
<tr>
<td>3. -a (--)</td>
<td>-a (--)</td>
<td>3rd person singular possessive</td>
</tr>
<tr>
<td>4. -i (--)</td>
<td>-i (--)</td>
<td>plural</td>
</tr>
<tr>
<td>5. -ő (--)</td>
<td>-ő (--)</td>
<td>no meaning</td>
</tr>
<tr>
<td>6. -k (--)</td>
<td>-k (--)</td>
<td>plural</td>
</tr>
<tr>
<td>7. -nk (--)</td>
<td>-nk (--)</td>
<td>1st person plural possessive</td>
</tr>
<tr>
<td>8. -ök (--)</td>
<td>-ök (--)</td>
<td>plural</td>
</tr>
<tr>
<td>9. -ök (--)</td>
<td>-ök (--)</td>
<td>3rd person plural possessive</td>
</tr>
<tr>
<td>10. -unk (--)</td>
<td>-unk (--)</td>
<td>1st person plural possessive</td>
</tr>
</tbody>
</table>

#### Type II suffixes

<table>
<thead>
<tr>
<th>MEANING</th>
</tr>
</thead>
<tbody>
<tr>
<td>-at accusative</td>
</tr>
<tr>
<td>-et accusative</td>
</tr>
<tr>
<td>-t accusative</td>
</tr>
<tr>
<td>-n genitive (if preceding, depends on it)</td>
</tr>
</tbody>
</table>

#### Type III suffixes (equivalent to postpositions)

<table>
<thead>
<tr>
<th>MEANING</th>
</tr>
</thead>
<tbody>
<tr>
<td>-ban x in + N (prepositional)</td>
</tr>
<tr>
<td>-be s into + N (acc.)</td>
</tr>
<tr>
<td>-be' s in (CT) [out from] + N (gen.)</td>
</tr>
<tr>
<td>-e' s in + N (prep.)</td>
</tr>
<tr>
<td>-ér for + N (gen.)</td>
</tr>
<tr>
<td>-ét s in (to) + N (gen.)</td>
</tr>
<tr>
<td>-ét s in (up to) + N (gen.)</td>
</tr>
<tr>
<td>-él=vel y et + N (gen.)</td>
</tr>
<tr>
<td>-etk s in (dat.)</td>
</tr>
<tr>
<td>-etn x in + N (pres.)</td>
</tr>
<tr>
<td>-etn s in + N (prep.)</td>
</tr>
<tr>
<td>-etn' s in (to) + N (preps.)</td>
</tr>
<tr>
<td>-éto' s in (out, from) + N (pres.)</td>
</tr>
<tr>
<td>-éto' s in (out, from) + N (gen.)</td>
</tr>
<tr>
<td>-vé s in (instrumental)</td>
</tr>
<tr>
<td>-vé s in (with) + N (inst.)</td>
</tr>
</tbody>
</table>
Explanations for TABLE 1

1. Information on Type I and Type II suffixes is obtained from the table.

2. Information on Type III suffixes can be obtained by three methods: firstly from the table, secondly from translation of the suffix as a postposition using the postpositional code, and finally from the operation of the Soup Translation Rules, section V.6.1, and from the Additional Instructions in the dictionary entries. Secondary information cancels primary, and final information supersedes both primary and secondary.

3. Absence of information as to number means singular.

For suffixes 2 and 3 a code is given for the translation of nouns in the presence of a comparative.

The last four suffixes are adverb suffixes. For convenience of arrangement adverbs formed from adjectives are considered conditionally as an "adverbial form" of adjective.

### TABLE 2

<table>
<thead>
<tr>
<th>SUFFIXES</th>
<th>MEANING</th>
</tr>
</thead>
<tbody>
<tr>
<td>back row</td>
<td>front row</td>
</tr>
<tr>
<td>1. -'</td>
<td>-'</td>
</tr>
<tr>
<td>2. -bb</td>
<td>-bo</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>3. -ebb</td>
<td>-ebo</td>
</tr>
<tr>
<td>4. -au</td>
<td>-en</td>
</tr>
<tr>
<td>5. -leg</td>
<td>-leg</td>
</tr>
<tr>
<td>6. -n</td>
<td>-l</td>
</tr>
<tr>
<td>MEANING</td>
<td>SUFIXES</td>
</tr>
<tr>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>1. -es</td>
<td>0</td>
</tr>
<tr>
<td>2. -ak</td>
<td>0</td>
</tr>
<tr>
<td>3. -a</td>
<td>0</td>
</tr>
<tr>
<td>4. -ás</td>
<td>0</td>
</tr>
<tr>
<td>5. -nas</td>
<td>See No. 6</td>
</tr>
<tr>
<td>6. -hat (-)</td>
<td>0</td>
</tr>
<tr>
<td>7. -ik</td>
<td>0</td>
</tr>
<tr>
<td>8. -á</td>
<td>0</td>
</tr>
<tr>
<td>9. -jának</td>
<td>past imperf.</td>
</tr>
<tr>
<td>10. -jának</td>
<td>past imperf.</td>
</tr>
<tr>
<td>11. -i-</td>
<td>0</td>
</tr>
<tr>
<td>12. -nák</td>
<td>past imperf.</td>
</tr>
<tr>
<td>13. -nánk</td>
<td>past imperf.</td>
</tr>
<tr>
<td>14. -né (-)</td>
<td>past imperf.</td>
</tr>
<tr>
<td>15. -nak</td>
<td>0</td>
</tr>
<tr>
<td>16. -ne</td>
<td>past imperf.</td>
</tr>
<tr>
<td>17. -niok</td>
<td>perfective infinitive</td>
</tr>
<tr>
<td>18. -ni (-)</td>
<td>perf. infin.</td>
</tr>
<tr>
<td>19. -júk</td>
<td>perf. infin.</td>
</tr>
<tr>
<td>20. -é</td>
<td>pres.active participle</td>
</tr>
<tr>
<td>21. -on</td>
<td>past imperf.</td>
</tr>
<tr>
<td>22. -óm</td>
<td>past perf.</td>
</tr>
<tr>
<td>23. -út</td>
<td>past perf.</td>
</tr>
<tr>
<td>24. -e</td>
<td>past perf.</td>
</tr>
<tr>
<td>25. -ét</td>
<td>past perf.</td>
</tr>
<tr>
<td>26. - receptor</td>
<td>0</td>
</tr>
<tr>
<td>27. -hek</td>
<td>0</td>
</tr>
<tr>
<td>28. -va</td>
<td>imperative</td>
</tr>
</tbody>
</table>

*The Hungarian past participle (f) is translated into Russian by a past passive participle in transitive verbs and by a perfective past active participle in intransitive verbs.*
PROBLEMS OF CYBERNETICS

Explanations for TABLE 1

1. Substitute in the verb endings:
   e for e, ã for ã,
   o for ã, ã for ã, e
   u for ñ.

2. Absence of information as to tense (O) – present tense.

3. Absence of information as to person (0) – 3rd person.

4. Absence of information as to number (0) – single in number.

5. Absence of information as to conjugation (0) – conjugation.

6. Information regarding an ending consisting of two suffixes is obtained by adding the orders (i.e., by appropriate column – time added to time, person to person, etc.) on each of the suffixes.

Supplement to TABLE 1

1. V-hat-ô
   (a) Before S: Y-hat-ô + ô = Nô + (acc. case) 
   (gender and number as in Nô; acc. case) + wîmô [is possible] + ô (imperf. infin.).

   (b) Not before S: Sô (without suf. II or III, farthest to the right) + Y-hat-ô + ô = Nô (acc. case) + ...... + wîmô [is possible] + ô (imperf. infin.).

2. V-hat-ott (Pô) is translated similarly, except that wîno [is possible] is followed by swnô [was possible].

Now let us consider several examples so that the reader can become more familiar with the Scanning Table.

We will translate the word hangrendezôret from sentence (5) nag kell, hogy visszajövôk a hangrendezôret, a hangrendezôcseket...... (Rev. p. 59). In the dictionary we find:

<table>
<thead>
<tr>
<th>hangrendezõrekt</th>
</tr>
</thead>
</table>
| compound word | Sô  
| short form |  
| sound | sô [sound]  
| 【sound】 |  

We remove the part hang from the word being scanned and we look for the remainder in the table according to rule O.I.10.C; not finding rendezôrekt in the table (after having substituted y for y) and replacing y according to rules II.10.10 and II.10.10.0, we scan the dictionary and find:

<table>
<thead>
<tr>
<th>rendezôret</th>
</tr>
</thead>
</table>
| compound word | Sô  
| sound | 【sound】  
| system |  

We remove rendezôret as well, and we scan the suffix table again for the remainder -ôret. It is not found there as a whole; it contains the
máchine translation from hungarian to russian

suffix -ek. rules II.10.11.0, II.10.11.01 (since -et is found in the
table) and II.10.11.1 come into effect. The word being scanned is compound
(according to rule I.4), and is translated by rules 1.6.a and 1.6.b thus:

<table>
<thead>
<tr>
<th>hung</th>
<th>indigér</th>
<th>et</th>
<th>et</th>
</tr>
</thead>
<tbody>
<tr>
<td>5zónok- [sound]</td>
<td>szétem [system]</td>
<td>pl.</td>
<td>acc.</td>
</tr>
</tbody>
</table>

the translation will be: 5zónokb sztem [sound system].

The word hangolat 5zónka [sounds] is translated in a similar, although
less difficult, manner.

A compound word such as nyelvlelek is treated somewhat differently;
initially we find nyelv in the dictionary:

<table>
<thead>
<tr>
<th>nyelv</th>
<th>not comp wd.</th>
<th>a</th>
</tr>
</thead>
<tbody>
<tr>
<td>nyelv</td>
<td>comp. wd.</td>
<td>a</td>
</tr>
<tr>
<td>5zónk- /language-/</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5zónk- /linguistic/</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

we remove the stem and scan the suffix table for the remainder -lelek.
Rules II.10.11.0 and II.10.11.01 come into effect: -lelek is found imperfectly;
we remove -e (the part coinciding with the suffix -et in our
table) and scan the table for -lek: if it cannot be found, we attempt to
find lelek (according to rule II.10.11.00), and again it cannot be found.
Rule II.10.11.00.0 then comes into effect: we replace all of the initial
remainder (i.e., -lelek), we scan the dictionary for it, and find elem
element [element]. Rules from II.1 to II.10.11.1 then come into effect,
followed by rules 1.7 and 1.8 for the translation of compound words.

As a result we obtain:

<table>
<thead>
<tr>
<th>nyelv</th>
<th>elem</th>
<th>et</th>
</tr>
</thead>
<tbody>
<tr>
<td>5zónk- /linguistic/</td>
<td>elem- /element/</td>
<td>pl.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>nominative</td>
</tr>
</tbody>
</table>

i.e., 5zónokb eleméte [linguistic elements].

Now we take the word visszalönk from our sentence. We find vissza- in the
dictionary - the verb accéncsobath [to examine]. According to rules
II.11 and II.11.1 we scan the sentence for a verbal prefix and find it -
nyelv; we join this to the verb and again scan the dictionary, this time
for visszalönk. Not finding this word in the dictionary (the verb
accéncsobath is entered only in its simple form), we take back the prefix
according to rule II.11.10 and use the dictionary article on the word
vissza- for translation. In the suffix table for the verb we find the suf-
fic -lonk: this is a 1st person plural, indefinite form. We translate
according to the verb translation rules (VIII.4.2 and VIII.4.3) and we
obtain a general translation of the sentence: nyelv, nem 5zónkoj accéncs-
bath 5zónokb szétem [linguistic system]. We must then examine sounds and
sound systems [pl. example (-)]. In translating sentence (6)
felhangozott hangolatok bicskei ("w. et.", p.67) the word
reihangosítás is not found in the dictionary since the verb reihangosítás is
entered only in its simple form - the prefix does not alter its meaning
essentially. Rule II.10 comes into effect: since there is a prefix in the
dictionary which is contained in the word being scanned, we separate it
according to rule II.01, find the stem reihangosítás in the dictionary, take the
information on it and disregard the prefix. The verb is then subjected to

"Thus the cutting of the verbal prefixes is disregarded entirely if they do
not make a substantial difference in the dictionary meaning of the verb."
PROBLEMS OF CYBERNETICS

rules II.10.1, II.10.11 and II.10.11.1, followed by text translation.

Rules (VIII.4.6, VIII.5, VIII.6 and VIII.6.1.e). As a result we obtain:

<table>
<thead>
<tr>
<th>fel</th>
<th>használ</th>
<th>usk</th>
<th>hangüteményi</th>
<th>bizonyíték</th>
<th>ok</th>
</tr>
</thead>
<tbody>
<tr>
<td>испольzuje</td>
<td>1st pers. Pl.</td>
<td>историко-фонетический</td>
<td>доказательства</td>
<td>pl. acc.</td>
<td></td>
</tr>
</tbody>
</table>

We are using историко-фонетическое доказательства [we are using phonetic evidence of a historical nature].

Now let us take the word пе‌діаскі. In the dictionary we find пе‌діаскі/пример [example]. We scan the noun suffix table for the remainder. According to rule II.10.11.0, not having found it as a whole, we remove ɪ and scan for -kal. Again we do not find it. Rule II.10.11.0 comes into effect again, as a result of which one -і is removed. We begin to scan for -kal and find it.

The word пе‌діаскі breaks down as follows:

<table>
<thead>
<tr>
<th>пе‌діаскі</th>
<th>пример</th>
<th>kal (= tel)</th>
</tr>
</thead>
<tbody>
<tr>
<td>no meaning</td>
<td>pl. о [with] + inst.</td>
<td></td>
</tr>
</tbody>
</table>

The translation is: с примерами [with examples].

The word кер‌дес is handled somewhat differently: we find in the dictionary кер‌дес/вопрос [question], and we separate it: the remainder -ел is not found in the table: by rule II.10.10 we substitute ɪ for е and find -ел in the table (rule II.10.10). We obtain:

<table>
<thead>
<tr>
<th>кер‌дес</th>
<th>вопрос</th>
</tr>
</thead>
<tbody>
<tr>
<td>с [with] + inst.</td>
<td></td>
</tr>
</tbody>
</table>

and the translation is с вопросом [with the question].

Handling of usages in machine translation from Hungarian is the same as in translating from French. We therefore refer the reader to article [1], p.115, and limit ourselves to an example: the translation of sentence (?) a взято из статьи экз. A врезанный факт. After scanning we obtain:

<table>
<thead>
<tr>
<th>частъ</th>
<th>вет</th>
<th>ек</th>
<th>врезанный</th>
<th>стать</th>
<th>ак</th>
</tr>
</thead>
<tbody>
<tr>
<td>part</td>
<td>acc.</td>
<td>part</td>
<td>and</td>
<td>слово</td>
<td>pl. поз.</td>
</tr>
<tr>
<td>ДЕЙСТВИЕ</td>
<td>взятъ</td>
<td>и взятъ</td>
<td>(borrowed)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

It could be rendered as... взятое слово явля… частъ [...]. The bound words took part...), but the article for the word рёз contains an instruction regarding a usage (in the third column):

| рёз | S | ɪl- | [part-] |

Usage No.11 - рёзъ везти принимать участие [to take
Homonym Differentiation

After the conclusion of scanning, homonym differentiation begins.

1. In cases of dictionary homonym (a word has several different translations within the limits of the same part of speech) the computer supplies all the translations it has (see above). This in no way hinders the operation of the analyzing rules: the same data must be determined for all the translation variations, in so far as they belong to the same part of speech.

2. Grammatical homonym (a) Typographic grammatical homonym (coincidence of word-forms of different parts of speech) is encountered comparatively seldom in Hungarian. Yet it does exist: the coincidence of verb and noun stems such as dob “to throw, drum”, zeb “predict, sell”, siri “to pay, grave”, etc.; the coincidence of past participles and the third person singular past indirect of verbs; and also the coincidence of nouns and adjectives, which occurs particularly often in linguistics texts in the names of nationalities.

These cases are provided for by the Rules for the Differentiation of Homonyms, which are given below.

(b) In cases of partial grammatical homonym (coincidence of separate words that are different parts of speech) the necessary instructions are contained directly in the dictionary article of the word in question (see words at, az, test, etc., in dictionary).

III. Rules for the Differentiation of Homonyms

Differentiation of the VS (verb-noun) homonym form.

Having found the word stem in the dictionary and made certain that it is VS, check whether it has an ending.

0. No ending - check whether VS is proceeded by a determinative, numeral, word marked "a" or adjective (do not pass beyond noun of verb).

00. None of the above before VS - VS is a verb.

01. Some of the above before VS - VS is an adjective.

1. Ending - scan noun suffix table for it.

10. Not found there - VS is a verb.

11. Found - check whether it is a Type I suffix or not.

11.0. Not found among Type I suffixes - VS is a noun.

11.1. Found among Type I suffixes - assume it has no ending.

E. Differentiation of the EP (past personal verb - past
personal verb - past participle) homonym form. Having found
the verb stem in the dictionary and made certain from the verb
suffix table that it is \( VP_p \), check whether \( VP_p \) stands before
any noun (do not pass beyond a verb).

1. \( VP_p \) does not stand before \( S \) (pass beyond numerals) -
Check whether \( VP_p \) stands before \( VP_p \).

2. \( VP_p \) does not stand before \( VP_p \). \( VP_p \) is a personal
verb.

3. \( VP_p \) stands before \( VP_p \). \( VP_p \) is translated the same
as \( VP_p \).

4. \( VP_p \) stands before \( S \). - check whether there is a deter-
minative to the left of \( VP_p \) (not passing beyond a \( \text{C}_{\text{anh}} \) noun).

5. No determinative to the left - check in the sentence
(not passing beyond a \( \text{C}_{\text{anh}} \) noun) whether there is another
verb or adjective standing before \( S \).

6. None of the above - \( VP_p \) is a personal verb.

7. The sentence contains an adjective not standing
before \( S - VP_p \) is a participle.

8. The sentence contains another verb - check
whether it is \( VP_p \).

9. Other verb not \( VP_p - VP_p \) to be translated is a
participle.

10. Other verb \( VP_p \) - apply schemes (a) and (b)
(arrow indicate dependency, \( VP_p \) - translated \( VP_p \)).

a) \( VP_p \rightarrow (something) \rightarrow (S) + (VP_p) \rightarrow (S) + (VP_p) \rightarrow (S)

b) \( VP_p \rightarrow (S) + (VP_p) \rightarrow (S) + (VP_p) \rightarrow (S)

11. Neither scheme is suitable - \( VP_p \) is a
personal verb.

12. One of the schemes is suitable - according
to scheme (a) \( VP_p \) is a personal verb and \( VP_p \) a
participle; according to scheme (b) \( VP_p \) is a partici-
ple and \( VP_p \) a personal verb.

13. There is a determinative to the left - \( VP_p \) is a partici-
ple.

C. Differentiation of the \( SA \) (noun - adjective) homonym
form. Having found the word stem in the dictionary and made
certain that it is NA, check whether it has an ending.

1. No ending in NA - check for S directly to the right.
   
   00. No S directly to the right - check for D to left (not passing beyond noun or verb).
   
   00-0. No D to the left - NA is an adjective.
   
   00-1. D to the left - NA is a noun.

2. S directly to the right - NA is an adjective.

1. NA has ending - NA is a noun.

Sentence casting. After usages have been handled and homonyms differentiated, the casting of the sentence begins.

As has already been stated, the rules of machine translation are formulated on the initial understanding that a text is to be translated by sentences (a sentence is a section of the text enclosed by two periods; see above). Indications have already been made as to which words can be passed in scanning the sentence.

However, for convenience and accuracy in the analysis of a Hungarian sentence it is best to cast it.

This means:

1. separate the constituent parts of the sentence, subordinate clauses, participles, gerunds and isolated and parenthetic usages, and separate the simple clauses connected by coordinate conjunctions. For this a strict and formal definition of the left and right boundaries of the above-mentioned simple elements in the sentence must be made.

2. Establish formal connexions between words and set up "dependence groups" (principal - subordinate - a word and all the words in any way depending on it).

Sentence casting is effected through the simultaneous application of two closely connected methods:

(a) the use of the Governing Tables and special rules indicating what words are subordinate to a given word in order to establish the relations between words;

(b) the application of syntactical formulae for the main types of subordinate clauses, participle usages, etc.

The rules of sentence casting should be systematically united into a separate and coherent group. They must certainly be of a very general nature, and can be applied with certain modifications to sentence casting not only in Hungarian but in a number of other languages as well. These rules are the basis of the Word Arrangement Rules (in the Russian text). Preliminary sentence casting (establishing the relations between words and separation of simple clauses) permits considerable simplification of the whole system of translation rules.
However, work on Sentence Casting Rules is not yet complete, and they will be treated in a separate article. A number of indications as to the determination of the connections of words are given here, however, by way of an initial general summary; they are to be found chiefly in the analytical rules in the form of special comments.

As far as separation of the parts of a sentence is concerned, the present rules are based essentially on translation within the limits of the sentence (and not its parts) with instructions regarding those words which restrict scanning, which does, of course, reduce the accuracy of these rules slightly. However, in arranging the words in the Russian text it is still necessary sometimes to suppose that the Hungarian sentence has somehow already been cast and broken down into its simple constituent elements. We will see that this has the disadvantage of being temporary. In this the following must be continually kept in mind: before all C- sub, except hely, the unconditional boundary of the sentence lies in the case when these C- sub do not follow C- conj; in the case where they do not limit the sentence. Hely always limits the sentence, however.

**Analysis.** For general information see [1].

The analysis rules in Hungarian operate in the following order: postpositions are treated first, then nouns, adjectives, numerals and finally the verb. This is not a random selection. It is necessary to treat postpositions first in order to be able to treat nouns (to determine their case); it is convenient to translate adjectives (and also numerals) after nouns, making use of their agreement; a number of verb forms agree with nouns (participles always; also see rules VIII.B.2.1, VIII.D, and VIII.E.1.2), and therefore the verb must be translated last.

All five groups of analyzing rules are given in full below. The analyzing rules operate as follows:

1. Postpositions
2. Nouns (and personal pronouns)
3. Adjectives (ordinal numbers, demonstrative and relative pronouns)
4. Cardinal numbers
5. Verb.

Adverbs (and the "adverbial form" of the adjective) are not subject to the analyzing rules; however the following information must be known about them:

(a) Ad (and Ad) before A. modify A.
(b) Ad (and Ad) before or after V modify V.
(c) Adp is always independent.

Conjunctions are also not subject to analysis. Regarding their inter-relations the following can be said.

(a) All subordinate conjunctions are independent (except hely) and designate the beginning of a sentence.
(b) selly, amai, aki, etc., depend on the noun after which they stand.

(c) Co-ordinate conjunctions depend on the word before which they stand.

IV. Postposition Translation Rules

Postpositions are usually translated by Russian prepositions, instructions being given as to what case they require: $S' = \text{Prep}^n + M'$ or $S' = \text{suf} + \text{III}^n = \text{Prep} + M'$. They govern the noun after which they stand; they either depend on a word governing them or are independent, if there is no such word.

Q. If the postposition is without a suffix check whether it is followed in the dictionary by a hyphen.

QC. No hyphen - check whether it is a Type II postposition.

QC.0. Type I postposition - take information from dictionary.

QC.1. Type II postposition - check whether there is a corresponding positional code in the sentence.

QC.10. No postpositional code - take information on postposition from dictionary or from Table 1.

QC.11. Postpositional code - take information on postposition according to this code.

Q1. Hyphen - assume given postposition to have suffix -/-.  

1. If there is a postposition with this suffix, scan TABLE 1 for it; after translation of postposition put personal pronoun with number and person of suffix; its case is found according to the information on the postposition, which is determined according to rules IV.0 and so on.

By way of explanation let us note that rules IV.0 and IV.01 deal with three Type II postpositions, i.e., belge-, hogez- and ré-, which do not usually have suffixes when the corresponding cases from the 3rd person singular personal pronouns (the forms belgie, hogzé and ré are possible, although not literary).

V. Noun Translation Rules

The form of the Russian noun is consistently determined by two morphological data: number and case. But in translating from Hungarian "person" must also be determined for Russian nouns, i.e., the problem of the selection of possessive pronouns which are in certain cases used to translate Hungarian possessive noun suffixes must be solved.

Thus, in the operation of the Noun Translation Rules the following must be determined for each Russian noun: number, case and person.

(A) Number. The translated Russian noun is put in the same number as the corresponding Hungarian noun, if there are no additional instructions (see Rule IV.0.00 on the
dictionary articles for words such as legross, etc.).

(8) Case. Check whether the noun has Type II or Type III suffixes.

Q. If S' does not have either Type II or III suffixes, check whether S' stands before S" (not passing beyond a personal verb).

00. S' does not stand before S" - check whether S' stands before D.

00.0. S' does not stand before D - check whether S' stands before Coni.

00.00. S' does not stand before Coni - check whether S' stands before Pz.

00.00.0. S' does not stand before Pz - N' in nominative case.

00.00.1. S' stands before Pz - case of N' determined by postposition (see IV, Postposition Translation Rules) (S' depends on Pz).

00.01. S' stands before Coni - N' in same case as nearest N" on the right (S' depends on same word as S") if both N' and N" are in the nominative, put verb of this sentence into the plural.

00.1. S' stands before D - check what case N" will be in translation of S" nearest on the right.

00.10. N" not in nominative case - check whether there is another N in the nominative to the right up to a period or semi-colon.

00.10.0. No other N in nominative - N' in nominative.

00.10.1. Other N in nominative - N' in same case as N" (S' depends on same word as S")

00.11. N" in nominative - N' in nominative.

01.1. S' stands before S" - check for comma between S' and S".

01.0. No comma - check whether S" has a possessive suffix.

01.00. No possessive suffix in S" - N' in nominative.

01.01. S" has possessive suffix - N' in genitive (S' depends on S", but if S" is followed by other S' with possessive suffixes, S' depends on the latter S').

01.1. Comma - check whether endings in S' and S" are similar.
MACHINE TRANSLATION FROM HUNGARIAN TO RUSSIAN

01.10. S' and S" have different endings - N' in the nominative.

01.11. S' and S" have similar endings - N' in some case as N" (S' depends on same word as S").

1. If S' has Type II or III suffixes, check which.

10. If S' has an -at/-et (l) suffix, check whether there is a negative before the verb governing S".

10.2. No negative - N' in accusative (S' depends on governing verb).

10.1. Negative - N' in genitive case (S' depends on governing verb).

11. If S' has a Type III suffix, check whether it is -nak/-nek.

11.2. S' has a Type III suffix, but not -nak/-nek - consider suffix a Type II postposition and change to rule 11' (Postposition Translating Rules) (S' depends on the suffix).

11.3. S' has -nak/-nek suffix - check whether S' stands before a noun with a possessive suffix (before S-ps).

11.3.6. S' does not stand before S-ps - check whether S' stands before or after D.

11.10.6. S' does not stand before D - consider -nak/-nek a Type II postposition, and change to Postposition Translating Rules (S' depends on the suffix -nak/-nek).

11.10.1. S' stands before or after D - N' in genitive case (S' depends on S-ps nearest right or left).

11.11. S' stands before S-ps - N' in genitive case (S' depends on S-ps).

For an example of rule V.B.00.01 see No.17.

Rules V.B.00.01, V.B.00.10 and V.B.00.10.9 are illustrated by examples Nos. 16 and 20 (in example No. 16 eredete is considered as standing before since the advent sana is omitted by rule 11').

Rules V.B.01, V.B.01.11 and V.B.01.11 are seen in operation in example No.19.

Rules V.B.11.1, V.B.11.10, V.B.11.10.11 and V.B.11.11 describe the possessive construction of the type olsz husnak rig a leve (also see example No.20).

(C) Person. A possessive pronoun determined according to the rules V.B. depends on N' and stands before the latter and the adjectives agreeing with it.
0. If there is not possessive suffix in S', do not determine the person of N'.

1. If S' has a possessive suffix, check whether it is 3rd person or otherwise.

10. No 3rd person possessive suffix in S' - before N' put possessive pronoun of same person as possessive suffix, and with the same number, gender and case as N'.

11. S' has 3rd person possessive suffix - check whether S' follows directly after another S without Type II or III suffixes, or with a -nak/-nek suffix not separated by a personal verb.

11.0. Before S' no such other S (or there is an S with a type II or III suffix other than -nak/-nek) - check whether N' is in nominative case.

11.00. N' not in nominative - check whether there are 1st or 2nd person pronouns in the nominative to the left.

11.00.0. No 1st or 2nd person pronouns in the nominative to the left of N' - before N' put caoh [one's own] (gender, number and case as in N'), and in brackets etc. [possessive pronoun] (gender and number are determined by the nearest noun on the left within the limits of the sentence).

11.00.1. There is a 1st or 2nd person nominative pronoun to the left of N' - check whether S' contains suffixes -a or -uk.

11.00.10. S' contains -uk suffix - put nx [their] before N'.

11.00.11. S' contains -a suffix - put ero [possessive pronoun] (gender and number determined according to the noun nearest the left within the sentence) before N'.

11.01. N' in nominative - check whether S' contains suffixes -a or -uk.

11.01.0. S' contains suffix -uk - put /their/ before N.

11.01.1. S' contains suffix -a - before S' put (gender and number are determined by the nearest noun on the left within the limits of the sentence).

11.1. There is another S without Type II or III suffix before S' - do not determine the person of N' (see example No.22 for an example of the determination of the person of a noun).

VI. Adjective Translation Rules
Adjectives with the mark T[A 0'C] are first handled by the rules in VI, and then by VII.8.
If the adjective is marked "indecl.", these rules do not operate; the information is taken from the dictionary.

Check whether A' stands directly before S.

Q. If A' does not stand directly before S, check whether A' stands before Coni.

QG. If A' does not stand before Coni, check whether A' contains Type II or III noun suffixes.

QG.0. No such suffix - check whether A' stands before A''.

QG.00. A' does not stand before A'' - check whether A' stands before Y.

QG.00.0. A' does not stand before Y (or stands before volt) - check whether the sentence contains N in the nominative (A' depends on volt).

QG.00.00. Sentence does not contain N in nominative - A' is translated by the abbreviated form (if it has one) in the neuter, singular nominative.

QG.00.01. The sentence contains N in the nominative - A' is translated by the abbreviated form (if it has one) with the gender and number of N, and in the nominative case; the case is nominative.

QG.00.1. A' stands before Y' (except volt) - A' is translated by the abbreviated form in the neuter, singular nominative (A' depends on Y').

QG.01. A' stands before A'' - check whether A'' stands before S.

QG.01.0. A'' does not stand before S' - P' in same gender, number and case as P'' (A' modifies same word as A'').

QG.01.1. A'' stands before S' - check whether this S' depends on another S'', passing participles even.

QG.01.10. S' does not depend on any S'' - P' in same gender, number and case as P'' (A' modifies same word as A'').

QG.01.11. S' depends on S'' - find S^X (the mark X means "independent") nearest the right which does not depend on any other S', and put P' into the gender, number and case of S^X (A' modifies S^X); A can be considered as standing after S^X.

QG.1. There is such a suffix - put P' into the case demanded by this suffix (according to the Noun Translation Rules), and into the gender and number of the nearest noun on the left (A' modifies according to the way it is governed).
01. A' stands before \( C_{\text{con}} \) - check whether \( C_{\text{con}} \) is
directly followed by another A' or \( P_x \).

01.0. No other A'' or \( P_x \) stands after \( C_{\text{con}} \) - consider
A' not to stand before \( C_{\text{con}} \) and return to rule VI.00.

01.2. If A' stands directly before S', put P' into the same
gender, number and case as P'' (A'' modifies same word as
A' or \( P_x \)).

1. If A' stands directly before S', put P' into the same gen-
der, number and case as P' (A' modifies S').

Rule VI.00.01 and the four following rules were introduced
to handle cases like (1) epy katalašči ečalkeš irëvū eŋp
markutž (Rev., p.2). The subjective (A'') katalašči
[powerful] stands before the subjective (A'') ečalkeš/
protivopoložnyj [opposite]. And the noun (S') irëvū-
direktsija [direction] follows A''. Rule VI.00.01.1 checks
whether irëvū depends on another noun. It is found that it
depends on epykana [force]. Rule VI.00.01.13 comes into
effect - the nearest "independent" SA on the right - this is
etc. Therefore, we accord the translation of katalašči with
that of epykana and obtain deketašči močna silla protiv-
opoložnoj direktsija [a powerful force in the opposite
direction acts] (and not deketašči močna protivopoložnoj
protivopoložnoj direktsija [a force in a powerful opposite
direction acts]). A similar case is to be found in example
No.12).

Rule VI.01.01 provides for cases where a co-ordinate conjunc-
tion, for example ľa, does not co-ordinate similar parts (rule
VI.01.1), but whole clauses.

VII. Rules for the Translation of Cardinal Numbers and Words
Marked "

A. Determining the case of numerals (the form of the word
marked " is determined according to the rules for translating
of that part of speech which it happens to be).

0. If Ľ' has no ending, check whether Ľ' stands before S'.

00. Ľ' does not stand before S' - Ch' in nominative case

01. Ľ' does stand before S' - Ch' in the same case as N'
after the operation of rule V.5. (Ľ' modifies S').

1. If Ľ' has an ending, the case is determined according to
the noun suffix table.

B. Working out additional instructions.

0. If there is no Ľ' or word marked "Ľ" before S', check
whether Ľ' contains Type II or III suffixes.

00. S' does not contain Type II or III suffixes - check
whether Ľ is 1, 2, 3 or 4 or a number ending in 1.
MACHINE TRANSLATION FROM HUNGARIAN TO RUSSIAN

00.0. C or a word marked "s" stands before S', but is neither 1, 2, 3 or 4 nor C ending in 1, 2, 3 or 4 — put N into genitive plural.

00.1. Either 2, 3 or 4 or C ending in 2, 3, or 4 stands before S' — put N into genitive case.

01. S' contains Type II or III suffixes — check whether C is 1 or a number ending in 1.

01.0. Neither 1 nor C ending in 1 before S' — put N into plural.

01.1. Either 1 or C ending in 1 stands before S — stop additional instructions.

1. If C or word marked "s" do not stand before S' put N in nominative into the genitive plural.

Notes
(1) If rules VII.B.00.0 and VII.B.00.1 treat nominative nouns, the verbs in the given sentence are also put into the plural (this is not carried out for words marked "s").

(2) After the operation of rules VII.B.00.0, VII.B.00.1 and VII.B.02.0, the form of the adjectives modifying S' is correspondingly altered.

VIII. Verb Translation Rules

The form of the Russian verb is identically determined according to the following morphological data: 1) tense, aspect; 2) number; 3) person (in the present and future tenses); 4) gender (in the past tense); 5) reflexivity. The mood of the Russian verb does not enter the picture, since the imperative mood which is comparatively seldom met in scientific texts is excluded from consideration, and the other moods in Russian do not exist for the computer, for example: ësîm 6w émy oxasane o am, ou 6w nrênsen (if they would tell him, he would come). From the machine translation point of view this sentence contains past, plural, non-reflexive, perfective verbs. The particle 6w (conditional or subjunctive) is a separate word to the computer.

However, in translating from Hungarian it is also necessary to determine the so-called "formulation" of the Russian verb. The problem is that in the Hungarian language personal pronouns used as the subject and also as the direct object of transitive verbs and the indirect object of the so-called personal infinitives are ordinarily dropped, whereas in Russian this is not usually the case: the Hungarian látja corresponds to the Russian oxina máster ero [she sees him], and so on. Determining the "formulation" of a verb involves solving the problem of the selection and arrangement around the verb (when necessary) of the necessary personal pronouns.

Thus, in the operation of the Verb Translation Rules the following information must be determined for each Russian verb: tense (aspect), number, person, gender and formulation.

It is not necessary to determine reflexivity, since this is learned directly from the dictionary, where words such as
It should be noted that for the sake of convenience in arranging the verb suffix table, verb forms such as infinitives, participles and gerunds are provisionally considered by verb tense.

After preliminary data on verb form have been obtained (from the table) as a result of the operation of the Scanning Rule, they are made more precise with the aid of the Rules presented below.

In relation on tense as obtained in the preliminary data is considered first, and, depending on this, one of the following rules comes into operation.

In each of these rules an instruction is given as to which rule follows on its completion.

A) Tense and aspect of G'. 1. If "present tense" is obtained for G' according to preliminary data:

a) If the word hovy or hovy (in order to) is found to the left of G' not separated by another verb, substitute the past imperfect for the present tense. Change to rule VIII.B.

b) If G' in the 1st person plural and with suffix -i or -uk stands first in the clause, after its translation place in brackets its 1st person plural future perfective. Change to rule VIII.B.

c) In the absence of the conditions indicated in sections a) and b), retain "present tense" and change to rule VIII.B.

2. If "past imperfective" plus 6w is obtained in the preliminary data:

a) If the word hovy is not found to the left of the verb,

i) If the verb is in the 1st person plural, substitute "future perfective" for "past imperfective" and omit 6w.

ii) If the verb is in the 3rd person singular or plural substitute "present tense" for "past imperfective", omit 6w, and put nyev (let imper.) before the translation of the verb, if the verb is not already preceded by hadd/nyev.

iii) If the verb is in the 2nd person singular or plural, substitute "imperative mood" for "past imperfective", and omit 6w.

After VIII.A.2.a.(i) and VIII.A.2.a.(ii) change to rule VIII.B. After VIII.A.2.a.(iii) change to rule VIII.C.

b) If the word hovy is found to the left of the verb retain "past imperfective" plus 6w. Change to rule VIII.C.

3. If the "past imperfective (completion)" or "past perfective" are obtained with the preliminary data, retain them and
change to rule VIII.C.

4. If "participle" is obtained with the preliminary data, retain it (for the translation of participles with -hat suffix see Supplement to TABLE 3). Change to rule VII (Adjective Translation Rules) and determine the gender, number and case of the participle as though it were an adjective.

5. If "infinitive form" is obtained in the preliminary data, retain it and change to rule VIII.E.3.

6. If "gerund" is obtained in the preliminary data, check whether the verb forms lenni or maradni are to be found directly to the right or left.

7. No verb forms lenni or maradni - retain as before (put commas before gerund and after last word dependent on it).

8. Verb form lenni or maradni - substitute "past passive participle in abbreviated form" for "gerund" (its gender and number determined according to the nearest nominative noun; its case - nominative with the presence of van, van not being translated, and accusative in the presence of the remaining forms of lenni or maradni). Gerunds depend on these verb forms.

Rules VIII A.1.a and VIII A.1.b provide for cases of the coincidence of indicative and imperative moods in objective clauses.

Rule VIII A.2.a provides for imperative forms in independent clauses; it does not take into account all the cases where these forms are required, but only those most prevalent in scientific and technical language.

Generally, however, Hungarian imperative endings can be rendered as "past imperfective" plus öm, since in scientific and technical texts they are more often used in subordinate clauses and require precisely this translation.

Hungarian conditional endings are also rendered as "past imperfective" plus öm, but with the instruction "completion", since they are always so translated.

As for the aspect of the translated Russian verbs, it is selected approximately and provisionally with the given rules so that peculiar distortions in meaning do not occur. Strictly formal methods for determining in Russian what aspect must be used and where, have yet to be developed.

Rule VIII A.6.a comes into effect for clauses of the type (8) as cerűs sarva marad (l.e., p. 33, mossas nonок острый занят носовой; [the nasal cavity remains closed] (rearranging sarva according to instruction 11.4 on word order).

B) Person of 3. The translated Russian verb remains in the same person as the corresponding Hungarian verb. Change to Rule VIII.C.
C) Number of G'. The translated Russian verb remains in the same number as the corresponding Hungarian verb, if there are no additional instructions (as, for example, in rule V.B.00:01). If after the operation of rule VIII.C the verb is in the past singular change to rule VIII.D; in the remaining cases change to rule VIII.E.

D) Gender of G'.

0. If the clause contains no noun (or personal pronoun) in the nominative, carry out rule VIII.E.1.a.

00. If the "formulation" does not need to be determined, G' remains in the neuter.

01. If the "formulation" is to be determined, carry out rule VIII.D.1.

1. If the clause contains nouns (or personal pronouns) in the nominative not separated from G' by a co-ordinate conjunction, put G' into the gender of the nearest such noun on the left (or right). Change to rule VIII.E.

E) Formulation of G'. 1. For all verbs except those in the 3rd person plural and those with characteristic No.4 (for which formulation is not required):

a) If the clause contains no noun giving "nominative case" after the operation of rule L.E., put a personal pronoun in the corresponding person before G'; the gender of a 3rd person singular personal pronoun is selected in accordance with the nearest nominative noun to the left in the sentence (not in the clause). If the latter does not exist, all three variations are given.

b) If the sentence contains such a noun, formulation is not determined.

2. For objective verbs:

a) If there is no S' ending in -at/-et to the left or right of V' (until another verb or a co-ordinate conjunction is encountered), put ero [his] possessive pronoun ee, ex, etc [hers, their, this] after G'.

b) If there is an S ending in -at/-et, formulation is not determined.

3. For infinitives:

a) If V' is preceded by a word with characteristic No.3 (not separated by other words) and V' does not end in -pi, put a dative personal pronoun with the same person and number as V' directly before the translation of the word with the characteristic No.3 (V' depends on word with characteristic No.3).

b) If V' is not preceded by a word with characteristic

*See, for example, the manner in which the verb feharsupilunk is treated.*
No. 3 or if Y' ends in -ni, formulation is not determined.

Verbs with characteristic No. 4 (rule VIII.E.1) are impersonal, like megértesik/smuchается [it happens]. The reservation regarding the 3rd person plural in this rule is made for clauses containing infinitives or personal verbs.

Words with characteristic No. 3 are those like kell, azabb, etc. Rule VIII.E.3.a provides for cases of the type meg kell vizsgálnunk [REV., p. 59] нам нужно исследовать [we must examine].

Diagram of the Operation of the Verb Translation Rules

(arrows indicate direction of operation)

Synthesis. Generally speaking, tables of Russian endings and synthesizing rules are independent of the language from which a translation is being made. They are compiled with Russian information and provide for the formation of the necessary forms of Russian word according to the instructions that are given. Therefore, for machine translation from Hungarian the same tables of endings and synthesizing rules can be used as in translation from French (see [1]). However, in making a sample translation the reader will have to make use of his own knowledge of Russian.
Editing the Russian Text. Mention has already been made (see above) of the fact that the word order in the Hungarian sentence, which differs sharply from the Russian in this respect, has given added impetus to the development of special new rules — rules which must ensure more or less normal word order in the Russian obtained, i.e., rules for the editing of the Russian sentence. Work has shown that these rules can be conveniently divided into three groups.

A. Word Arrangement. These rules must be based on information on the Russian language and must permit regulation of the word order in the Russian without reference to the language from which the translation is being made. For this, it is probably better not to break them down into the form of individual notes on the different rules of analysis, but to collect them into a single group forming a complete system. The operation of this group of rules must consider previously established interconnections of words and the simple elements of the sentence (see Sentence Casting).

B. Arrangement of Punctuation. In all probability this also necessitates a separate system of rules closely connected with the foregoing group, but independent of the language of the text being translated.

C. Stylastic correction. This consists of the elimination of certain faults in the Russian caused by peculiarities in Hungarian.

Work on these groups of rules has not yet been completed, and these rules will be published separately. Instead we present in this paper general instructions, which should give the reader an idea of the direction the work is taking and allow him to make test translations.

Instructions as to Word Arrangement

I. General instructions

1. In rearranging do not move independent words.

2. If the sentence contains several words of the same type (for example, several nouns in the genitive, etc.), the instruction is fulfilled for the leftmost word; the remainder are placed after it in the same order as in the sentence.

II. Instructions for part of speech

1. A noun is placed:

   a) when in the nominative:

      (i) before the verb if the sentence has a personal verb;

      (ii) before the adjective if the sentence has no verb but has a translated according to rule 1.00.00.01;

      (iii) if there is no such , do not move the noun.

   b) when in the accusative: after the verb on which it depends and before other words dependent on the same verb (if there are any);
c) when in the genitive: after the word on which it
depends and before other words dependent on the same word;

d) when in other cases: after the word on which it
depends.

2. Adjectives, numerals and participles are placed:
   a) if no part of speech is dependent on them and if they
modify $S$, before $S$; if they depend on $Y$, after it.
   
   b) if a part of speech is dependent on them, after that,$
$S$ which they modify. Place commas before them and after the
last word dependent on them.

3. Adverbs and $A_{ad}$ - before the word which they modify.

4. Gerunds - after the verb on which they depend.

Several notes on punctuation are to be found in different
places (see rule VIII.A.8.0, the word $nely$; the word $mely$,
etc.).

As for stylistic correction, separate notes of a preliminary
nature can be given:

1) If the sentence contains several identical forms of
the verb work [to be able] standing before an infinitive,
omit all these forms except that to the farthest left (see
example No.15).

2) If the translation contains $roi$ [that] before the noun
it modifies and there is no co-ordinate conjunction after
the noun and there is no word over [this] to the right in the
sentence, omit $roi$.

3) Be warned [it is not allowed] is always substituted
for warning [it is not possible (negative impersonal)] (see
example No.15).

3. Examples

Now let us turn to examples of translation (with the assis-
tance of the rules that have been given): (9) Nag kell tehát
állapítanunk, hogy $t$- $f$- feloldás egyváltozás nem kivételt-
egyen (Bay., p.53). All the words in this sentence, except
for the fourth and sixth, are found in the dictionary. Rules
11.1, 11.10 (they are not verbs) and 11.10.0 (no remainder)
come into effect. The information on them is taken from the
dictionary. The sixth "word" is a formula: $t$- $f$. Ac-


ording to the rule for the treatment of formulas (11.4.d) we
place it after the translation of feloldás. For the fourth
word rules 11.0, 11.00 and 11.00.1 come into effect; we find
seg-ween with it to állapítanunk and find asejlepítanunk
in the dictionary.

Rules 11.1, 11.11 and 11.11.0 now come into operation (seg-
ness has been included in the word being scanned, and the sentence
no longer contains a verbal prefix); rul 11.11.1 then comes
into effect, since there is a remainder - $a$munk. We turn to
TABLE 3 and find:
According to rules II.10.11 and II.10.11.0 we scan the table for –nunk, and find:

<table>
<thead>
<tr>
<th>Suffix</th>
<th>Tense</th>
<th>Person</th>
<th>Number</th>
<th>Conjugation</th>
</tr>
</thead>
<tbody>
<tr>
<td>–nunk</td>
<td>imperf.</td>
<td>1st plur</td>
<td>undirected</td>
<td>infin.</td>
</tr>
</tbody>
</table>

Using explanation 6 to Table 3, we obtain information on the ending –amunk: 1st person plural undirected imperfective infinitive; scan for ending. The Homonym Differentiation Rules must now come into operation, but this sentence does not contain any of the homonym forms; we therefore change to the treatment of usages. The usage instruction gives the word *gyaltalábán*. Usage No.6 in the usage dictionary is suggested in the sentence and we substitute the information on usage No.6 (omana we) [by no means] for the information on the words *gyaltalábán* and *men*. Now we change to analysis.

In translating the noun *fejlődések* we keep its number as in the Hungarian according to Rule V.6. Since there is no information as to number, it is singular (see instruction). Now we look for its case. Rules V.2.0 and following lead to the conclusion of nominative case (Rule V.2.0.0.0.0). It is not necessary to determine person (rule V.2.1.0). Thus, *fejlődések* is subjects [development]. It is not necessary to subject the adjective *gyaltalábán* in the dictionary. Now we translate the verb *megállapítanunk*. Rule VIII.A.6 comes into operation and then VIII.B.3.3 (formulation determination): the verb is preceded by a word having characteristic No.3, and does not end in –am. We therefore place a 1st person plural personal pronoun (according to the –amunk ending) in the dative (i.e., *hám [us]* before the translation of *kell.* Thus, *megállapítanunk* becomes *hám... utasítani [us... to establish]*.

Utilizing the instructions on the interrelation of words and the notes in the dictionary articles, we can write:

Итак, мы должны установить, что данное *г- > f-* изменение не без исключения.

[Thus, we must establish the fact that the development of *г- > f-* is by no means without exception.]

It is not necessary to apply the instructions on word arrangement. The translation of sentence (10) is a similar case: *idegen szavak mindig hatolnak be a nyelvbe* (М.Г., p.3). It is suggested that the reader carry out the translation of (10) independently; *Külbösen érős volt a szponcer* *magyarázata a későbbiekben* (М.Г., p.47) supplies another example.

Scanning the dictionary and tables is easily done.
The verb lenni bears an instruction regarding a usage (6-1). But this usage - azo lenni - is not indicated in the sentence, and the verb lenni must be translated separately.

The sentence contains homonymous forms also: \( Vp. - \text{volt; with the aid of} \text{III.B.C and III.B.00 we establish that this} \) is a personal form of the verb lenni (since according to the dictionary vol = len); and SA - Kazárokkal; with the aid of rule III.0.1 we find that it is S.

Now we change to translation.

We begin with postpositions. The sentence contains one: -kal = -val (in the word Kazárokkal); this is either a Type III or Type II suffix. According to rule IV.00.1 we scan the sentence for a postpositional code. It is found in kazárokkal = No.4. According to governing table No.4 we find: S' - val = C [with] + M' (instrumental). In this kazárokkal modifies kapcsolat.

Then we turn to nouns. Magyarság/anyany [magyar] (plural) is subjected to rules V.A. and V.B.01.01. We obtain nemetsk (of the Magyars); it is not necessary to determine the person for this word (according to rule V.B.0). Thus, magyarság depends on kapcsolat.

Kapcsolat/a/капсълт [bond] is handled by rule V.A: we keep it singular in number (no information on number was given, so it is singular). The case is determined according to rule V.B.00.1.2. But we must know the case of the noun nearest on the right, i.e., kazárokkal. We switch to it.

Kazárokkal/kal/kazap [Kazar-) is plural (by rule V.A.) and instrumental (according to the translation of the prefix -kal); it is not necessary to determine person. We have s kazaper [with the Kazars]. Now return to kapcsolat. Rules V.B.00.16 and V.B.00.20.0 come into effect and we obtain nominative case for kapcsolat: it is not necessary to determine person. We have capsa.

Translation of adjectives: for külbőlven/особенno [particularly] we have information from TABLE No.2: a₄₄ (neuter, singular, nominative, abbreviated form), i.e., особенno. Külbőlven as a₄₄ modifies erdős.

Ерг/ям/- [strong] comes under rules VI.C.C.0 and VI.C.C.01 (since the sentence contains the nominative noun capsълт): nominative, abbreviated form, gender and number as in capsълт; i.e., capsълт.

The verb volt (i.e., jest provisionally) is the past of bár [to be]. Rule VIII.1.5 comes into operation; it refers us to VIII.C (we obtain singular), and then to VIII.2. According to rule VIII.1.1 we establish the feminine gender for volt (the sentence contains a nominative noun, kapcsolat which is feminine). It is not necessary to determine the formulation (VIII.1.1.5). As a result we obtain capsълт.

The general translation with indication of inter-relations
is:

[particularly strong was however the Magyars the bond with
the Khazars].

Using the instructions on word arrangement, we get:

[the bond between the Magyars and the Khazars was particularly
strong however].

Let us consider example (12): .... e szerdik lextoch
nyelvener izen sok idegen credefu ssava van (Rev., p.21).

After scanning is completed we switch to analysis (no usages
or homonyms are found in this sentence)*.

The postposition -nek (Type III suffix) is handled by rule
V,00.04, and from the postpositional code of the verb lemmi
we determine its translation: a (y) [in(at)] + y' (preposi-
tional or genitive case).

Nouns: nyelvenek/ nájma [language] according to rule V,A is
translated in the singular, and the case is obtained accord-
ing to the translation of -nek: it is not necessary to
determine person. As a result we have a (y) námka (nájma)
in (at) the language).

Credefu/ [origin-] is genitive singular
according to the noun suffix table; it is not necessary to
determine person. We obtain procsxomémeni With this
credefu modifies ssava. Szava = szav/a/czasz [word] which is
nominate singular (according to rule V.B.C and following to
V.B.00.00.0); it is not necessary to determine person (accord-
ing to characteristic No.6 in the verb lemmi); we get czosz.

Adjectives: sok is first subject to rule VI. It undergoes
rules V.I.00.01 (sok stands before A idegen) V.I.00.01.1 (A'
idegen - stands before S' - credefu and VI.00.01.11 (S'
credefu - depends on S = ssava). Hence sok modifies ssava,
and is considered as standing before the latter. In the dic-
tionary article we select a suitable translation (before S
without III or II af): whomo (many) with the instruction
"indiscernible". Therefore we do not determine the number
gender and case for whomo, and change (since sok - A'C)
to rule VII.B (see instructions regarding A'C before rule V).

Rules VII.B.0, VII.B.00 and VII.B.00.0 come into operation.
The latter causes us to change the number and case of ssava
to the genitive plural: conos. General note L is not ful-
filled: sok - A'C, idegen is translated (according to rule
V.I.01) in the gender, number and case of credefu and modifies
the latter who paremomoro (of foreign).

The verb van (¼en): rules VIII.A.1.G., VIII.B, VIII.C and
VII.E.3.b give the form ets (there is, are). It is not
necessary to determine the formulation, since the sentence

*Regarding instructions on the usage with lemmi (van), see
preceding example.
contains a noun which is found to be in the nominative after the operation of rule V.B.2 this is \( \textit{szava} \). Finally, having fulfilled the instructions given in the dictionary article for the word \( \textit{legjobb} \), we put the translation of the noun \( \textit{nyelv} \) into the genitive plural and the translation of the word \( \textit{legjobb} \) into the prepositional (genitive) case following the preposition \( \textit{at} \) (y). The general result is:

\[ \text{\textit{Szava} } \text{ \textit{menyek} } \text{ \textit{bog\'azy} } \text{ \textit{szava} } \text{ \textit{menyek} } \text{ \textit{bog\'azy}} \]

(and therefore in languages very many of foreign origin words are). After word arrangement we have:

\[ \text{\textit{Bog\'azy} } \text{ \textit{menyek} } \text{ \textit{bog\'azy} } \text{ \textit{szava} } \text{ \textit{menyek} } \text{ \textit{bog\'azy}} \]

(and therefore there are many words of foreign origin in most languages).

The last example we will consider is (23): \( \textit{Egy mas} \) \( \textit{arintkez\'ebe} \) \( \textit{ker\'ult} \) \( \textit{nyelv} \) \( \textit{egy mas} \) \( \textit{bizonyos} \) \( \textit{mertekben} \)
\( \textit{k\'ozelednek} \) (Rev., p. 21).

For the word \( \textit{bizonyos} \) there is an instruction regarding a usage. Usage No.2 - \( \textit{Bizonyos mertekben} \) - is the same as in our sentence. We substitute the information on the corresponding words by that on the usage: a number of words (to a certain degree), \( \textit{no} \).

The homonym differentiation rules then come into operation as the VP homonym form - \( \textit{ker\'ult} \) - must be treated. Rule III.B.1 operates; as there is no determinative to the left of \( \textit{ker\'ult} \), we scan the sentence for another verb (III.B.2), we find it - \( \textit{k\'ozelednek} \) (rule III.B.10.1.b), see that it is not VP (rule III.B.10.10), and establish that \( \textit{ker\'ult} \) is a past perfective participle from the verb \( \textit{nona} \) (-est) [to enter], i.e., \( \textit{no} \). Change to analysis.

Postpositions: - \( \textit{sal} = -val \): postpositional code of the nearest word (\( \textit{arintkez\'ebe} \)) - No.4; from governing table No. 4 we find the translation: \( \text{\textit{S'}} -val = c \text{ \textit{with} } + N' \) (instrumental). \( \textit{Egy mas} \) depends on \( \textit{arintkez\'ebe} \), \( \textit{beg} \); according to the postpositional code for \( \textit{ker\'ult} \) we find the translation: \( \text{\textit{S'}} -bel = P \text{ \textit{into} } + N' \) (accusative). \( \textit{Arintkez\'ebe} \) depends on \( \textit{ker\'ult} \), \( \textit{-ho} \); we translate accordingly to the postpositional code for \( \textit{k\'ozelednek} \): \( \text{\textit{S'}} -ho = c \text{ \textit{with} } + N' \) (instrumental). \( \textit{Egy mas} \) depends on \( \textit{k\'ozelednek} \).

Nouns: \( \textit{Egy mas} \) is translated as \( \textit{nyelv} \) c \( \textit{nyelv} \) [one another] (from the additional instruction in the dictionary article); \( \textit{Arintkez\'ebe} \) \( \textit{a kontakt} \) [in contact] (the number is determined according to rules V.A.2 - Singular, it is not necessary to determine person); \( \textit{Nyelv} \) \( \textit{nyelv} \) \( \textit{nyelv} \) \( \textit{nyelv} \) [language] is nominative plural (according to rules V.B.01 to V.B.01.00); it is not necessary to determine person; i.e., \( \textit{nyelv} \) \( \textit{egy mas} \) \( \textit{egy mas} \) \( \textit{egy mas} \) \( \textit{egy mas} \) [language] c \( \textit{nyelv} \) (see above).

Verb translation remains: \( \textit{ker\'ult} \); rule VII.A.4 refers us to the adjectival translation rules, from which (rule VII.1) we find that \( \textit{ker\'ult} \) is translated in the same gender, number and case as \( \textit{nyelv} \), i.e., \( \textit{no} \). For \( \textit{k\'ozelednek} \) the
translation of the 3rd person plural present oblyazat'sya
(draw together) is easily obtained from the table.

The general translation is:

Язык, в котором один язык друг с другом.

[one with another in contact entering languages one with another to a certain degree draw together].

We rearrange the words: we change the order of номинив
(according to rule 6.B) and the words depending on it, and
insert commas; next we place the second друг с другом
(according to instruction 1.B.6). It is not necessary to move
either язык, since this is a nominative noun and stands
before the verb (instruction 4.C), or a and together since
this is an independent grouping. The final result is:

Язык, в котором один язык друг с другом.

[languages which have been in contact draw together to a certain degree].

Conclusions
Apart from certain novel features which arise from the
special characteristics of the Hungarian language, such as
the Scanning Rules and certain particularities in both the
dictionary and the process of analysis, work on the rules
studied in this article makes it possible to draw several con-
clusions which are essential both to machine translation and
to linguistics in general.

1) The following proposition results from the examination
of Hungarian texts and experimental translations with the aid
of the given rules: the first and most important stage in
machine translation is the establishment of the connexions
and inter-relations between words, i.e., the formation of
syntagms. This operation must be fulfilled by an independent
group of rules of a sufficiently general nature to apply to a
number of languages with certain modifications. Syntagm for-
mation makes possible the simplification of rules of analysis
and the regulation of word order.

2) Every possible use must be made of syntagm formation
in sentence casting.

3) Inter-relation Groups (syntagms) are formulated with
the aid of governing Tables from which information is gained
on what forms can depend on a given word, and also with the
aid of special rules.

4) It is necessary to develop Rules for the Editing of
the Russian Sentence which are independent of the language
from which translation is being made. These rules which are
based on the use of the syntagms formed will provide for more
or less normal word arrangement in the Russian text obtained.
Punctuation can also be included in this.

5) The utilization of the Additional Instructions and
Characteristics makes possible the creation of less cumbersome rules of analysis. If for each singular Hungarian noun a check is made to see whether it is preceded by a cardinal numeral, a great many superfluous checks will be made. It is more advantageous to treat this problem differently: indicate in the noun translation rules that the number of the Hungarian is to be retained in the translation, and for numerals indicate that the noun following them remains in the plural on translation.

6) Statistical examination of a text is of great significance, since consideration must first be made of typical usages which are sufficiently prevalent in the actual texts from a certain field. Rare cases differing from normal usage must be disregarded at first. Statistical characteristics are of decisive importance in this.

This article represents a first attempt at machine translation from Hungarian, and has no pretentions to being a complete or definitive report. No use at all has been made of the possibility of analysing word formation elements, which would make it possible to unburden the dictionary; the homonym differentiation rules and some others are not sufficiently accurate; there are no rules for scanning, analysing and translating compound participles of the type jéborított csúcsok = nevűtett csúcsok aszus [ice-covered peak]. A great amount of complex work still remains to be done.

In conclusion may I express my sincere gratitude to A.A. Lyapunov and O.S. Kulagina for their kind assistance in the preparation of this work.

(Received 16 April 1957)