Machine Translation from Hungarian to Russian

I.A. Mel'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 Russian 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 phraseology, synthesis of the Russian sentence, and so on). Some of these characteristics have been described in the article by O.S. Kulagina and I.A. Mel'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 character-
istics", several "governing tables"(as examples), full "tables
of Hungarian suffixes", "scanning rules", "rules for the
differentiation of homonyms", and all the "analysing rules".

The "rules for sentence casting", "tables of Russian end-
ings", "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 explana-
tions 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 Géza Bárczi: _Magyar szókincs eredete_,
Budapest, 1951 (referred to as _M.sz._ in this article), and
Bevezetés a nylevtudományba, Budapest, 1953 (_Bev._ in this
article).

The following signs and symbols are used in this work:

PART OF SPEECH	HUNGARIAN	RUSSIAN
Noun.............................	S	N
Adjective........................	A	P
Adverb...........................	Ad)	
"Adverbial form of adjective"	)	I
(cf. p. 266)...................	Aad)	
Adverb-parenthesis...............	Adpp	Ipp
Verb.............................	V	G
Past Participle..................	Pp	
Cardinal number.................	C	Ch
Determinative...................	D	
Postposition....................	Ps	
Preposition.....................		Prep
Co-ordinate conjunction.........	Cconj	
Subordinate conjunction.........	Csub	
Prefix..........................	Pr	

Simple symbols

    n   front row
    z   back row
    ik   ik-verb
    sf   suffix
    ps   possessive suffix
    comp.wd. translation in compound word
    N indicating words of the _sok_ and _kevés_ type

1   primes on the upper right of letters signify
     corresponding translatable and translated stems or
     similarity of two words (the sign $\underline{A}'$ written in
     different places means a single specific adjective
     and not adjective in general; $\underline{S}'$ and N' or $\underline{S}''$
     and N'' mean that N' is the translation of $\underline{S}'$ and
     N'' of $\underline{S}''$)
=   "is translated"
-    succession and unbroken writing of parts of words
     (the sign "-" after any part shows 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
⌢   arrows indicate dependence between words ($\underline{A}' \overset{\frown}{+} \underline{S}$:
     "$\underline{A}'$ depends on $\underline{S}'$ ").

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 con-
tains 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 [2].

In the relatively few cases where there are permutations or
suppletives in the stems, the dictionary gives both stems, for
example: szav- = szó, irodalm- = irodalom-, jus- = jut-,
vol- = len- (cf. volt - lenni), abb- = azb-, arr- = azr-,
enn- = ezh- (i.e., abban, arra, enhez) and so on.

However, since all variations in stem conform to three
simple types** (abbreviation or loss of letter: víz - viz/et,
végez - végz/ik, hatalom - hatalm/at, natural permutation:
idő - ideje, lát - lás/sa; accumulation of -v-: nő - növünk),
another solution is possible: to give one stem in the diction-
ary, 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 intransitive present
indicative of the verb), and unvarying words are given in
their entirety [3]. 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, грамматик-, чтени-, част-, фонетическ-, обычн-
etc.), and verbs are also given in the form of stems (писа-,
исследова-, счита-ся, случа-ся 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., az
stands after azonban, and na after hanem, hang, hat, 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 of them not being definitely stipulated and depending
only on the sense of the sentence. These are cases of dic-
tionary homonymy (járás: "walking" - "region" and szer-
kezet: "design" - "structure"), polysemy 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 connexions in meaning are used in
translating.

However, in future, a machine may be "taught" to select the
necessary translation from several possible choices, beginn-
ing with the analysis of word groups. For example, a list of
words could be supplied according to which szerkezet would
be rendered by "structure" (for example, nyelv, mondat, 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:
  (the elongation sign, is considered as a separate letter,
and stands first in the alphabet, i.e., á = a,, ö = ö,, etc.
cs, sz, gy and other ligatures are each considered as a
separate letter, so that cs follows c, sz after s, etc.
ccs, ssz, ggy, nny and other doubled ligatures are con-
sidered as ordinary ligatures; ccs = cscs, ssz = szsz,
etc.

(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 -ik conjunction or not; for postpositions - row also (so that rule IV.1 operates correctly).

(c) Instruction as to usage. The treatment of word order, which must not be translated literally but as a whole (for example, jelen volt is rendered он присутсвовал ["he was one who was present"], whereas a literal translation would be он был присутствующим ["he was one who was present"], does not have any characteristic rules in Hungarian. The reader is referred to the previously mentioned article [1] Instructions on usage (M.sz.).

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 (névutó) "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 affixes (rag), which will in future be referred to as noun suffixes: affixes 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 we combined the Hungarian postpositions and affixes into the general class of postpositions falling into two groups: postpositions proper (névutó) are called Type I postpositions, and affixes - 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., affixes 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, do not 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 term, 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 -j- (from -ja, 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 -ú/-ü 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: csekély mérték/мелкий масштаб
['a small scale'], csekély mértékű terkép/мелкомасштабная
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) ..... ellenkező irányú
erő (Bev., p.) becomes сила противоположного направления
['a force of (moving in) the opposite direction']. It is
therefore convenient to consider -ú/-ü 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 II 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 into 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 resz/
часть ['part']; resz-től/от части ['from part']; resz-e/
ero часть ['his part']; től-e/от него ['from him'],
literally ero от ['his from']). These case formations, which
in combining with possessive suffixes change their form, are
found in the dictionary in precisely this altered form (rá
instead of -ra/-re; rajt instead of -on/-en; benn instead
of -ban/-ben, and so forth).

Now we can consider the postpositional code.

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

*In addition to compound words like kétarcú/двуликий,
двусторонний ['two-faced, doubled-sided'], kékszemű/
голубоглазый ['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 post-
positions (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/говорить [to speak] is found in code 12 in the
dictionary.

Governing Table No.12 is as follows:

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

This means that beszél vknek must be translated as говорить
кому-либо [to speak to someone], beszél vmről as говорить о
чем-либо [to speak about something], and so on.

The postpositional code contains almost all verbs and many
adjectives (for example, jellemző vmre / характерный для
чего-либо [characteristic of something], tele vmvel / полный
чем-либо [full of something], and also some nouns:
kapcsolat vmvel/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 gazdagít/обогащать [to enrich] and jelöl/обозначать
[to designate] are under general code 3. Table 3 contains the
Type II suffix -val/vel for which the translation N (instru-
mental) is supplied, i.e., after the verbs gazdagít and jelöl
a noun with the -val/-vel suffix is translated in the instru-
mental case.

In conclusion let us note that the term "governing" is con-
sidered 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
adjoined to the part of speech governed) and are translated
independently from it, but which through the dictionary mean-
ing 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 áll/стоять
[to stand], marad/оставаться [to remain] and terjed/
распространяться [to spread] includes the postpositions
-ban/-ben and -on/-en. These postpositions do not depend on
the above-mentioned verbs: generally speaking they can
accompany any verb. But they do accompany with particular
frequency precisely such verbs as áll and marad, which by
their dictionary meaning are particularly disposed to have
accompanying adverbs of place (-ban/-ben and -on/-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
can be translated without the aid of these tables. But it is
through the Governing Tables that the machine knows all the
postpositions that govern (in our sense of the word) a given
verb or other word: so, with each stem in the dictionary,
the postpositional code is given (the number of the corres-
ponding 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 "editing" 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
above). 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 equiva-
lent 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:

(3) Russian information, consisting of instructions regard-
ing the selection of the necessary stem of the translated
word and its grammatical characteristics (for example, gender
and declension in nouns) - see [1](M.sz.).

(4) 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 csak, nemcsak and nem; kazár,
lengyel, magyar, német 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:

(1) 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.

I. 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 test enclosed by two periods*).

2.(a) Scanning within a sentence takes place to the left
and right, not passing subordinate conjunctions; special pro-
vision 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, atten-
tion 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 syntac-
tically 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 S' stand after
N'.

4. In analysing any constructions, adjectives, participles
and adverbs are not considered except in cases specially pro-
vided 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. wd.", 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.

9. A word ending in a hyphen (for example, értelem- 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 век [sixteenth century], rule I.3.c for different proper nouns, for example (2) A Garonne viszont nyelvjárási határ (Bev., p.24) С другой стороны / Garonne / диалектная граница [on the other hand Garonne is a dialect boundary], and rule I.3.d for the clause of the type (3) A lengyel szypusz szó elszigetelen áll a lengyelben (M.sz., p.26), Польское слово szypysz изолированно стоит в польском языке, или [The Polish word szypysz stands alone in Polish], or (4) Más termeszettí a száll - repül - lengedez stb. kapcsolata (Bev., p.7) Другой природы связь [The connexion száll - repül - lengedez 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 II.10.11.00.0 prevents 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 phonetic 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, meghatároz/определять [to determine] (határoz means решать [to decide])

(d) those prefixed verbs which are not used (or are seldom used) in the simple aspect; for example, befejeződik/оканчивается [concludes] and megállapít/устанавливать [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, lezárul, behatol, megvizsgál); or a prefixed verb is sometimes met without a prefix (the latter being separate), which is entered in the dictionary only with the prefix. Rules II.0 to II.1 and from II.11 to the end of the Scanning Rules are provided for such cases.

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

---

* In the Verb Translation Rules (VIII) 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 numeration 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 v 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 v 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 II.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 found in table – return to rule II.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.

11.0. No verbal prefix in sentence – cease scanning, take dictionary article for word found by rule II.11, check whether there is a remainder and return to rule II.10.0 or II.10.1.

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

11.10. Word thus formed not found in dictionary – reject prefix, disregard it and carry out rule II.11.0.

11.11. Word 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 diction-
ary 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.E. Maitinskaya [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 synharmonism (two or three
variants of each suffix: -ott -ett -ött, etc.) and permuta-
tions j/sz, j/s, j/z 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 Verb Endings
must be made. With it all verbal suffixes are reduced to a
conditional "standard" form (back row semi-vowel: permutat-
ing consonants are replaced by j). This makes it possible to
reduce their number to a quarter; there is thus a total of
29 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/-vel, which alters its first consonant to the
last consonant in the word stem and emerges in the form -bal/
-bel, -dal/-del, -mal/-mel, 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/-vel suffix,
i.e., -kal/-kel, is entered as a whole in the table. This
was done so as to avoid introducing a number of rules to pre-
vent -kal/-kel being broken down by rule II.10.11.0 into the
suffix -k-, and -al/-el which is meaningless. In the other
variants of the -val/-vel suffix the initial letters do not
coincide with any of the suffixes in our table. A more com-
plete table would have to have either two variants of the
-val/-vel suffix, i.e., -mal/-mel and -dal/-del, or addi-
tional rules for handing this suffix, since it contains no
first and second person possessive suffixes in -m and -d.

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 -nak 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

SUFFIXES			MEANING
back row		front row	
Type I suffixes			
1. -'	(-)	-'	(-)	no meaning
2. -ak	(-)	-ek	(-)	plural
3. -a	(-)	-e	(-)	3rd person singular possessive
4. -i	(-)	-i	(-)	plural
5. -j	(-)	-j	(-)	no meaning
6. -k	(-)	-k	(-)	plural
7. -nk	(-)	-nk	(-)	1st person plural possessive
8. -ök	(-)	-ök	(-)	plural
9. -uk	(-)	-ük	(-)	3rd person plural possessive
10. -unk	(-)	-ünk	(-)	1st person plural possessive
Type II suffixes			
11. -at		-et		accusative
12. -ot		-öt		accusative
13. -t		-t		accusative
14. -ú		-ú		genitive (if preceeding S, depends on it)
Type III suffixes (equivalent to postpositions)			
15. -ban		-ben		в[in] + N (preposi- tional)
16. -ba		-be		в[into] + N (acc.)
17. -ból		-ből		из(от) [out (from)] + N (gen.)
18. -		-e		в[in] + N (prep.)
19. -ért		-ért		для[for] + N (gen.)
20. -hoz		-hez		к[to] + N (dat.)
21. -ig		-ig		до[up to] + N (gen.)
22. -kal=-val		-kel=-vel		
23. -nál		-nél		у[at] + N (gen.)
24. -nak		-nek		с N(dat.)
25. -n		-n		в[in] + N (prep.)
26. -on		-ön		в[in] + N (prep.)
27. -ra		-re		на[to] + N (acc.)
28. -ról		-ről		о[about] + N (prep.)
29. -től		-től		от[from] + N (gen.)
30. -ul		-ül		в качестве[in capacity of] + N (gen.)
31. -vá		-vé		N(instrumental)
32. -val		-vel		с[with] + N (inst.)

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 trans-
lation of the suffix as a postposition using the postposi-
tional code, and finally from the operation of the Noun
Translation Rules, section V.B.1, and from the Additional
Instructions in the dictionary articles. Secondary informa-
tion cancels primary, and final information supercedes 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

Adjectival suffixes

SUFFIXES		MEANING
back row	front row	
1. -'-	-'-)	no meaning
2. -bb	-bo))))))	(1) comparative degree, if leg - not prefixed to A (2) superlative degree, if leg - prefixed to A
3. -abb	-ebb)	suffixes 2 and 3 have code 19
4. -an	-en)))	
5. -iag	-ieg))	adverbial form - A_{ad} (abbreviated form, neuter, singular, nominative)
6. -n	-n))	
7. -ül	-ül)	

TABLE 3

Verbal suffixes

MEANING / SUFFIXES	TENSE	PER-SON	NUM-BER	CONJU-GATION	NOTES
1. -ak	0	0	pl.	0	
2. -ak	0	0	pl.	undi-rected	
3. -a	0	C	0	0	
4. -a-	0	0	0	0	
5. -has-	See No.6				
6. -hat (-)	0	C	0	0	V'-hat-sf" - мочь [to be able] (form as in V'-sf") + G' (imperfective infinitive)
7. -ik	0	0	0	undi-rected	in -ik verbs
	0	0	pl.	0	in other than -ik verbs
8. -i	0	0	0	0	
9. -janak	past imperf.	0	pl.	undi-rected	Put бы [condition-al or subj.] after verb
10. -junk	past imperf.	1st	pl.	do.	Put бы [condition-al or subj.] after verb
11. -j-	0	0	0	0	
12. -nék	past imperf. (ending)	0	pl.	0	Put бы [condition-al or subj.] after verb
13. -nánk	past imperf. (ending)	0	pl.	undi-rected	Put бы [condition-al or subj.] after verb
14. -ná (-)	past imperf. (ending)	0	0	0	Put бы [condition-al or subj.] after verb
15. -nak	0	0	pl.	undi-rected	
16. -na	past imperf. (ending)	0	0	do.	Put бы [condition-al or subj.] after verb
17. -niok	perfective infinitive	0	pl.	do.	
18. -ni (-)	perf.infin.	0	0	do.	
19. -nunk	perf.infin.	1st	pl.	do.	
20. -ó	pres.active participle	0	0	do.	
21. -on	past imperf.	0	0	do.	Put бы [condition-al or subj.] after verb
22. -ott	past perf. past part.	0	0	undi-rected	
23. -ott-	past perf.	0	0	0	
24. -t	past perf. past part*	0	0	undi-rected	
25. -t-	past perf.	0	0	0	
26. -ük	0	1st	pl.	0	
27. -ünk	0	1st	pl.	undi-rected	
28. -va	imperf. gerund	0	0	do.	

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

Explanations for TABLE 3

1. Substitute in the verb endings:

a for e, i for s,

o for ö, i for az,

ó for ő, i for z

u for ü.

2. Absence of information as to tense (0) - 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 3

1. V-hat-ó

(a) Before S: V'-hat$''$-ó + S''' = N'''+ , котор-[which] (gender and number as in N''; acc. case) + можно [is possible] + G' (imperf. infin.).

(b) Not before S: S' (without suf. II or III, farthest to the right) + ... + V'''-hat$'''$-ó = N' (acc. case) + ... + можно [is possible] + G'' (imperf. infin.).

2. V-hat-ott (P_p) is translated similarily, except that можно [is possible] is followed by было [was possible].

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

We will translate the word hangrendszereket from sentence (5) meg kell tehát vizsgálnunk a hangokat, a hangrendszereket....... (Bev., p.59). In the dictionary we find:

hang	not compound word	S_n			звук [sound]	
	compound word	A			comp.wd. - звук [sound] (-ой,ая...) [endings]	

We remove the part hang from the word being scanned and we look for the remainder in the table according to rule II.10.1; not finding vendszereket in the table (after having substituted v- for r-) and replacing r- according to rules II.10.10 and II.10.10.0, we scan the dictionary and find:

rendszer	S_n			систем [system]	

We remove rendszer- as well, and we scan the suffix table again for the remainder -eket. It is not found there as a whole: it contains the

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.7), and is translated by rules I.8.A.a and I.8.B thus:

hang	rendszer	ek	et
звуков- [sound]	систем [system]	pl.	acc.

The translation will be: звуковые системы [sound systems].

The word hangokat звуки [sounds] is translated in a similar, although less difficult, manner.

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

nyelv	not comp wd.	$\frac{S}{n}$			язык- /language-/	
	comp. wd.	A			языков-/linguistic/	

We remove the stem and scan the suffix table for the remainder -elemek. Rules II.10.11 and II.10.11.0 come into effect: -elemek is found imperfectly: we remove -e- (the part coinciding with the suffix -e- in our table) and scan the table for -lemek; if it cannot be found, we attempt to find vemek (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., -elemek), we scan the dictionary for it, and find elem элемент [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:

nyelv	elem	ek	-
языков- [linguistic]	элемент- [element-]	pl.	nominative

i.e., языковые элементы [linguistic elements].

Now we take the word vizsgálnunk from our sentence. We find vizsgál in the dictionary - the verb исследовать [to examine]. According to rules II.11 and II.11.1 we scan the sentence for a verbal prefix and find it - meg-; we join this to the verb and again scan the dictionary, this time for megvizsgálnunk. Not finding this word in the dictionary (the verb megvizsgál 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 vizsgál for translation. In the suffix table for the verb we find the suffix -nunk: this is a first person plural indefinite form. We translate according to the Verb Translation Rules (VIII/A.5 and VIII.E.3) and we obtain a general translation of the sentence: Итак, нам нужно исследовать звуки и звуковые системы [We must then examine sounds and sound systems] (cf. example (5)).   In translating sentence (6) felhasználunk hangtörténeti bizonyítékokat ("M.sz.", p.67) the word felhasználunk is not found in the dictionary since the verb használ is entered only in its simple form - the prefix does not alter its meaning essentially. Rule II.0 comes into effect; since there is a prefix fel- in the dictionary which is contained in the word being scanned, we separate it according to rule II.01, find the stem használ in the dictionary, take the information on it and disregard the prefix*. The verb is then subjected to

---

*Thus the meaning of the verbal prefixes is disregarded entirely if they do not make a substantial difference in the dictionary meaning of the verb.

rules <u>II.10.1</u>, <u>II.10.11</u> and <u>II.10.11.1</u>, followed by <u>Verb Translation</u>
<u>Rules</u> (<u>VIII.A.1.c</u>, <u>VIII.B</u>, <u>VIII.C</u> and <u>VIII.D.1.a</u>). As a result we obtain:

fel	használ	uuk	hangtörténeti	bizonyíték	ok	
	использу-[use]	1st pers. pl.	историко-фонетическ [historico-phonetic]	доказа тельств [evidence]	pl.	acc.

Мы используем историко-фонетические доказательства
[We are using phonetic evidence of a historical nature.]

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

The word <u>példákkal</u> breaks down as follows:

példa	'	k	kal(= val)
пример [example]	no meaning	pl.	c [with] + inst.

The translation is: c примерами [with examples].

The word <u>kérdéssel</u> is handled somewhat differently; we find in the
dictionary <u>kérdés</u>/вопрос [question], and we separate it: the remainder
-<u>sel</u> is not found in the table: by rule <u>II.10.10</u> we substitute <u>v</u> for <u>s</u> and
find -<u>vel</u> in the table (rule <u>II.10.10.1</u>).

We obtain:

kérdés	sel
вопрос [question]	c [with] +inst.

and the translation is c 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 (7) <u>e változásban ezek</u>
<u>a jövevényszavak részt vettek</u>. After scanning we obtain:

resz	t		vet	t	ek		jövevény	szav	ak	-
часть [part]	acc.	,	бра-,взя- [take]	past	pl.	and	заимст-вованн- [borrowed]	слов- [word-]	pl.	ноэ.

It could be rendered as.... заимствованные слова взяли часть
[.... the borrowed words took part...], but the article for
the word <u>rész</u> contains an instruction regarding a usage (in
the third column):

rész	S_n	11-1		[part-]	

Usage No.11 - <u>részt venni</u>   принимать участие   [to take

part] is indicated in the sentence, and thus it is rendered Эти заимствованные слова приняли участие в этом изменении. [These borrowed words took part in this transformation].

Homonym Differentiation
After the conclusion of scanning, homonym differentiation begins.

(1) In cases of dictionary homonymy (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 analysing 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 homonymy (a) Typographic grammatical homonymy (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", sejt "predict, cell", sir "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 homonymy (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 át, az, hogy, 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.

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

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

O1. 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.

B. Differentiation of the VP_p (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 nouns (do not pass beyond a verb).

$\underline{0}$. VP_p does not stand before $\underline{S}$ (pass beyond numerals) – check whether $VP_{p'}$ stands before VP_{p*} .

$\underline{00}$. $VP_{p'}$ does not stand before $VP_{p''}$ – $VP_{p'}$ is a personal verb.

$\underline{01}$. $VP_{p'}$ stands before $VP_{p''}$ – $VP_{p''}$ is translated the same $VP_{p'''}$.

$\underline{1}$. VP_p stands before $\underline{S}'$ – check whether there is a determinative to the left of VP_p (not passing beyond a C_{sub} noun).

$\underline{10}$. No determinative to the left – check in the sentence (not passing beyond a C_{sub} noun) whether there is another verb or adjective standing before $\underline{S}$.

$\underline{10.0}$. None of the above – VP_p is a personal verb.

$\underline{10.1.a}$. The sentence contains an adjective not standing before $\underline{S}$ – VP_p is a participle.

$\underline{10.1.b}$. The sentence contains another verb – check whether it is VP_p.

$\underline{10.10}$. Other verb not VP_p – VP_p to be translated is a participle.

$\underline{10.11}$. Other verb VP_p – apply schemes (a) and (b) (arrows indicate dependency, 'VP_p – translated VP_p).

a) .....$VP_{p'}$ + ($\underline{D}$) + '$\underline{S}$-sf + ... + $VP_{p''}$ + $\underline{S}$-sf

b) .....$VP_{p'}$ + ($\underline{S}$...) + $\underline{S}$-( ) + $VP_{p''}$ + $\underline{S}$-sf

$\underline{10.11.0}$. Neither scheme is suitable – $VP_{p'}$ is a personal verb.

$\underline{10.11.1}$. 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 participle and $VP_{p'}$ a personal verb.

$\underline{11}$. There is a determinative to the left – VP_p is a participle.

$\underline{C}$. Differentiation of the $\underline{SA}$ (noun – adjective) homonym form. Having found the word stem in the dictionary and made

certain that it is <u>SA</u>, check whether it has an ending.

<u>O</u>. No ending in <u>SA</u> - check for <u>S</u> directly to the right.

   <u>OO</u>. No <u>S</u> directly to the right - check for <u>D</u> to left (not passing beyond noun or verb).

      <u>OO.O</u>. No <u>D</u> to the left - <u>SA</u> is an adjective.

      <u>OO.1</u>. <u>D</u> to the left - <u>SA</u> is a noun.

   <u>O1</u>. <u>S</u> directly to the right - <u>SA</u> is an adjective.

<u>1</u>. <u>SA</u> has ending - <u>SA</u> is a noun.

   <u>Sentence casting</u>. 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 <u>by sentences</u> (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 <u>cast</u> it.

   This means:

   (1) separate the constituent parts of the sentence, subordinate clauses, participles, gerunds and isolated and paranthetic usages, and separate the simple clauses connected by coordinate conjunctions. <u>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.</u>

   (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 <u>Governing Tables</u> 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 <u>Word Arrangement Rules</u> (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 <u>Sentence Casting Rules</u> is not yet complete, and they will be treated in a separate article. A number of indications as to the determination of the connexions 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 <u>sentence</u> (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 <u>mely</u>, 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. <u>Mely</u> always limits the sentence, however.

<u>Analysis</u>. For general information see [1].

The analysing 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 also agree with nouns (participles always; also see rules <u>VIII.B.2.1</u>, <u>VIII.D</u>. and <u>VIII.E.1.a</u>.), and therefore the verb must be translated last.

All five groups of analysing rules are given in full below:. The analysing 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 analysing rules; however the following information must be known about them:

(a) $\underline{Ad}$ (and $\underline{A}_{ad}$) before $\underline{A}'$ modify $\underline{A}'$.

(b) $\underline{Ad}$ (and $\underline{A}_{ad}$) before or after $\underline{V}'$ modify) $\underline{V}'$.

(c) $\underline{Ad}_{pp}$ 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 <u>mely</u>) and designate the beginning of a sentence.

(b) __mely__, __ami__, __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' + Ps'' = $Prep''$ + N' or S' -suf.III'' = $Prep''$ + N'. 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.

__0__. If the postposition is without a suffix check whether it is followed in the dictionary by a hyphen.

  __00__. No hyphen - check whether it is a Type II postposition.

    __00.0__. Type I postposition - take information from dictionary.

    __00.1__. Type II postposition - check whether there is a corresponding positional code in the sentence.

      __00.10__. No postpositional code - take information on postposition from dictionary or from Table 1.

      __00.11__. Postpositional code - take information on postposition according to this code.

  __01__. Hyphen - assume given postposition to have suffix __-a/-e__.

__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., __bele-__, __hozzá-__ and __rá-__, which do not usually have suffixes when the correspond-ing cases from the 3rd person singular personal pronouns (the forms __beléje__, __hozzája__ and __rája__ 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 pro-nouns 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 __VII.B.00__ and the

dictionary articles for words such as <u>legtöbb</u>, etc.).

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

<u>0</u>. If $\underline{S}'$ does not have either Type II or III suffixes, check whether $\underline{S}'$ stands before $\underline{S}''$ (not passing beyond a personal verb).

<u>00</u>. $\underline{S}'$ does not stand before $\underline{S}''$ - check whether $\underline{S}'$ stands before $\underline{D}$.

<u>00.0</u>. $\underline{S}'$ does not stand before $\underline{D}$ - check whether $\underline{S}'$ stands before $\underline{C}_{conj}$.

<u>00.00</u>. $\underline{S}'$ does not stand before $\underline{C}_{conj}$. - check whether $\underline{S}'$ stands before $\underline{Ps}$.

<u>00.00.0</u>. $\underline{S}'$ does not stand before $\underline{Ps}$ - N' in nominative case.

<u>00.00.1</u>. $\underline{S}'$ stands before $\underline{Ps}$ - case of N' determined by postposition (see IV, <u>Postposition Translation Rules</u>) ($\underline{S}'$ depends on $\underline{Ps}$).

<u>00.01</u>. $\underline{S}'$ stands before $\underline{C}_{conj}$. - N' in same case as nearest N'' on the right ($\underline{S}'$ depends on same word as $\underline{S}''$). If both N' and N'' are in the nominative, put verb of this sentence into the plural.

<u>00.1</u>. $\underline{S}'$ stands before $\underline{D}$ - check what case N'' will be in: translation of $\underline{S}''$ nearest on the right.

<u>00.10</u>. N'' not in nominative case - check whether there is another N in the nominative to the right up to a period or semi-colon.

<u>00.10.0</u>. No other N in nominative - N' in nominative.

<u>00.10.1</u>. Other N in nominative - N' in same case as N'' ($\underline{S}'$ depends on same word as $\underline{S}''$).

<u>00.11</u>. N'' in nominative - N' in nominative.

<u>0.1</u>. $\underline{S}'$ stands before $\underline{S}''$ - check for comma between $\underline{S}'$ and $\underline{S}''$.

<u>01.0</u>. No comma - check whether $\underline{S}''$ has a possessive suffix.

<u>01.00</u>. No possessive suffix in $\underline{S}''$ - N' in nominative.

<u>01.01</u>. $\underline{S}''$ has possessive suffix - N' in genitive ($\underline{S}'$ depends on $\underline{S}''$, but if $\underline{S}''$ is followed by other S with possessive suffixes, $\underline{S}'$ depends on the latter $\underline{S}$).

<u>01.1</u>. Comma - check whether endings in $\underline{S}'$ and $\underline{S}''$ are similar.

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

01.11. S' and S" have similar endings - N' in same 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 (t) suffix, check whether there is a negative before the verb governing S".

10.0. 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.

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

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

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

11.10.0. 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.00.1, V.B.00.10 and V.B.00.10.0 are illustrated by examples Nos. 16 and 20 (in example No. 16 eredete is considered as standing before since the adverb csak is omitted by rule I.4).

Rules V.B.01, V.B.01.1 and V.B.01.11 are seen in operation in example No.19.

Rules V.B.11.1, V.B.11.10, V.B.11.10.1 and V.B.11.11 describe the possessive construction of the type olsó husnak híg 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.

<u>0</u>. If there is not possessive suffix in <u>S</u>', do not determine the person of N'.

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

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

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

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

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

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

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

<u>11.00.10</u>. <u>S</u>' contains -<u>uk</u> suffix - put их [their] before N'.

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

<u>11.01</u>. N' in nominative - check whether <u>S</u>' contains suffixes -<u>a</u> or -<u>uk</u>.

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

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

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

VI. <u>Adjective Translation Rules</u>
Adjectives with the mark <u>E</u>(A <u>C</u>) are first handled by the rules in <u>VI</u>, and then by <u>VII.B</u>.

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.

$\underline{0}$. If A' does not stand directly before S, check whether A' stands before C_{conj}.

$\underline{00}$. If A' does not stand before C_{conj}, check whether A' contains Type II or III noun suffixes.

$\underline{00.0}$. No such suffix – check whether A' stands before A''.

$\underline{00.00}$. A' does not stand before A'' – check whether A' stands before V.

$\underline{00.00.0}$. A' does not stand before V (or stands before $\underline{volt}$) – check whether the sentence contains N in the nominative (A' depends on $\underline{volt}$).

$\underline{00.00.00}$. Sentence does not contain N in nominative – A' is translated by the abbreviated form (if it has one) in the neuter, singular nominative.

$\underline{00.00.01}$. 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.

$\underline{00.00.1}$. A' stands before V' (except $\underline{volt}$) – A' is translated by the abbreviated form in the neuter, singular nominative (A' depends on V').

$\underline{00.01}$. A' stands before A'' – check whether A'' stands before S.

$\underline{00.01.0}$. A'' does not stand before S' – P' in same gender, number and case as P'' (A' modifies same word as A'').

$\underline{00.01.1}$. A'' stands before S' – check whether this S' depends on another S'', passing participles even.

$\underline{00.01.10}$. S' does not depend on any S'' – P' in same gender, number and case as P'' (A' modifies same word as A'').

$\underline{00.01.11}$. 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 N^X (A' modifies S^X); A can be considered as standing after S^X.

$\underline{00.1}$. 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).

$\underline{01}$. $\underline{A}$' stands before C_{conj}. - check whether C_{conj}. is directly followed by another $\underline{A}$' or $\underline{P}_p$.

$\underline{01.0}$. No other $\underline{A}$" or $\underline{P}_p$ stands after C_{conj}. - consider $\underline{A}$' not to stand before C_{conj}. and return to rule $\underline{VI.00}$.

$\underline{01.1}$. No other $\underline{A}$" or $\underline{P}_p$ follows C_{conj}. - put P' into same gender, number and case as P" ($\underline{A}$" modifies same word as $\underline{A}$" or $\underline{P}_p$).

$\underline{1}$. If $\underline{A}$' stands directly before $\underline{S}$', put P' into the same gender, number and case as N' ($\underline{A}$' modifies $\underline{S}$').

Rule $\underline{VI.00.01}$ and the four following rules were introduced to handle cases like (1) egy hatalmas ellenkező irányú erő működik (Bev., p.21). The adjective ($\underline{A}$') hatalmas/могучий [powerful] stands before the adjective ($\underline{A}$") ellenkező/ противоположный [opposite]. And the noun ($\underline{S}$') irány-ú направление [direction] follows $\underline{A}$". Rule $\underline{VI.00.01.1}$ checks whether irányú depends on another noun. It is found that it depends on erő/сила [force]. Rule $\underline{VI.00.01.11}$ comes into effect - the nearest "independent" $\underline{S}$" on the right - this is erő. Therefore, we accord the translation of hatalmas with that of erő and obtain действует могучая сила противо- положного направления [a powerful force in the opposite direction acts] (and not действует сила могучего противо- положного направления [a force in a powerful opposite direction acts]). A similar case is to be found in example No.12).

Rule $\underline{VI.01.0}$ provides for cases where a co-ordinate conjunc- tion, for example és, does not co-ordinate similar parts (rule $\underline{VI.01.1}$), but whole clauses.

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

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

$\underline{0}$. If $\underline{C}$' has no ending, check whether $\underline{C}$' stands before $\underline{S}$'.

$\underline{00}$. $\underline{C}$' does not stand before $\underline{S}$' - Ch' in nominative case

$\underline{01}$. $\underline{C}$' does stand before $\underline{S}$' - Ch' in the same case as N' after the operation of rule $\underline{V.B}$. ($\underline{C}$' modifies $\underline{S}$').

$\underline{1}$. If $\underline{C}$' has an ending, the case is determined according to the noun suffix table.

B. Working out additional instructions.

$\underline{0}$. If there is no $\underline{C}$ or word marked "C" before $\underline{S}$', check whether $\underline{S}$' contains Type II or III suffixes.

$\underline{00}$. $\underline{S}$' does not contain Type II or III suffixes - check whether $\underline{C}$ is 1, 2, 3 or 4 or a number ending in 1.

00.0. C or a word marked "C" 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 "C" 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 nomina-
tive nouns, the verbs in the given sentence are also put into
the plural (this is not carried out for words marked "C").

(2) After the operation of rules VII.B.00.0, VII.B.00.1
and VII.B.01.0, 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) reflexibility.
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: Если бы ему сказали об этом, он бы пришел
[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 бы [condi-
tional 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 pro-
nouns 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 она видет его [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 reflexibility, since this
is learned directly from the dictionary, where words such as

tükröz /отражать [to reflect] and tükrözödik / отражаться
[to be reflected] are entered separately.

It should be noted that for the sake of convenience in
arranging the verb suffix table, verb forms such as infini-
tives, 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 <u>Scanning
rule</u>, they are made more precise with the aid of the <u>rules</u>
presented below.

Information on tense as obtained in the preliminary data is
considered first, and, depending on this, one of the follow-
ing rules comes into operation.

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

A) <u>Tense and aspect of G'</u>. 1. If "present tense" is ob-
tained for <u>V'</u> according to preliminary data:

a) If the word <u>hogy</u>/чтобы [in order to] is found to the
left of <u>V'</u> not separated by another verb, substitute the past
imperfect for the present tense. Change to rule <u>VIII.D</u>.

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

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

2. If "past imperfective" plus бы is obtained in the pre-
liminary data:

a) If the word <u>hogy</u> 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 бы.

ii) If the verb is in the 3rd person singular or plural
substitute "present tense" for "past imperfective", omit
бы , and put пусть [let (imper.)] before the translation
of the verb, if the verb is not already preceeded by <u>hadd</u>/
пусть.

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

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

b) If the word <u>hogy</u> is found to the left of the verb
retain "past imperfective" plus бы . Change to rule <u>VIII.C</u>.

3. If the "past imperfective (completion)" or "past perfec-
tive" 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 use Supplement to TABLE 3, Change to rule VI (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.

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

1. 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 бы , 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 бы , 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.6.1. comes into effect for clauses of the type (8) az orrüreg zárva marad (Bev., p.33) носовая полость остается закрытой [the nasal cavity remains closed] (rearranging zárva according to instruction II.4 on word order).

B) Person of G'. The translated Russian verb remains in the same person as the corresponding Hungarian verb. Change to Rule VIII.C.

C) <u>Number of G'</u>. 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
<u>V.B.OO.O1</u>). If after the operation of rule <u>VIII.C</u> the verb
is in the past singular change to rule <u>VIII.D</u>; in the remain-
ing cases change to rule <u>VIII.E</u>.

D) <u>Gender of G'</u>.
<u>O</u>. If the clause contains no noun (or personal pronoun) in
the nominative, carry out rule <u>VIII.E.1.a</u>.

<u>OO</u>. If the "formulation" does not need to be determined, G'
remains in the neuter.

<u>O1</u>. If the "formulation" is to be determined, carry out
rule <u>VIII.D.1</u>.

<u>1</u>. 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 <u>VIII.E</u>.

E) <u>Formulation of G'</u>. 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 <u>V.B</u>., 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 <u>S</u>' ending in -<u>at</u>/-<u>et</u> to the left or
right of <u>V</u>' (until another verb or a co-ordinate conjunction
is encountered), put ero [his] possessive pronoun ee, их, это
[hers, their, this] after G'.

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

    3. For infinitives:

    a) If <u>V</u>' is preceded by a word with characteristic No.3
(not separated by other words) and <u>V</u>' does not end in -<u>ni</u>,
put a dative personal pronoun with the same person and num-
ber as <u>V</u>' directly before the translation of the word with
the characteristic No.3 (<u>V</u>' depends on word with characteris-
tic No.3 (<u>V</u>' depends on word with characteristic No.3).

    b) If <u>V</u>' is not preceded by a word with characteristic

---

*See, for example, the manner in which the verb <u>felhasználunk</u>
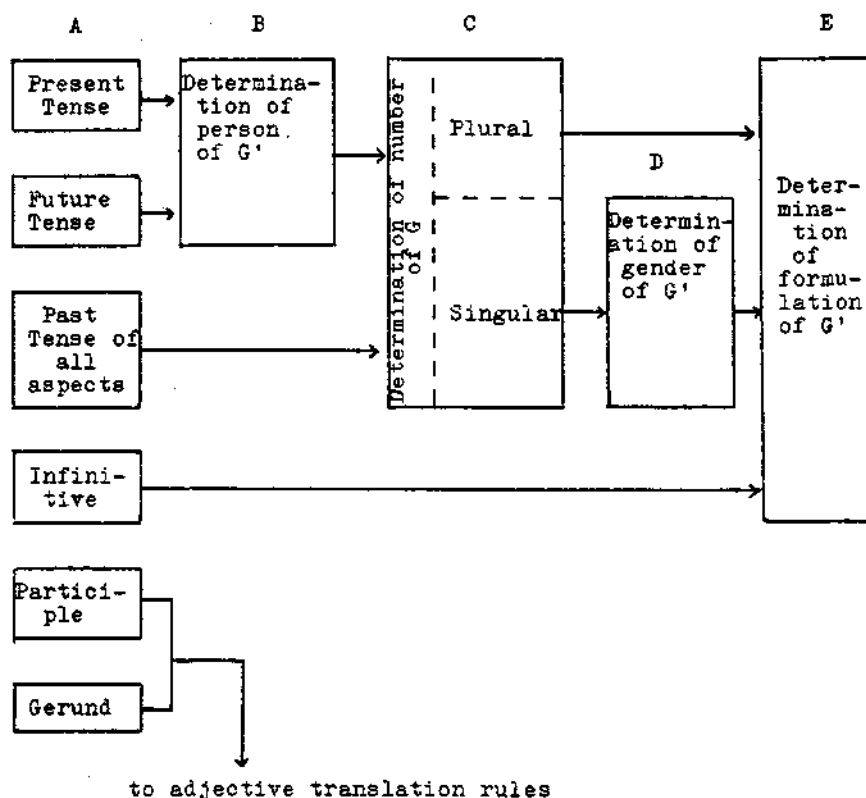is treated.

No.3 or if _V'_ ends in -_ni_, formulation is not determined.

Verbs with characteristic No.4 (rule _VIII.E.1_) are impersonal, like _megtörténik_/случается [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, szabad_, etc. Rule _VIII.E.3.a_ provides for cases of the type _meg kell vizsgálnunk_ (Bev., p.59) нам нужно исследовать [we must examine].

Diagram of the Operation of the Verb Translation Rules

(arrows indicate direction of operation)

A   B   C   E

Present
Tense   → Determina-
           tion of |
           person. | Plural
           of G' | D   Deter-
Future → | mina-
Tense | Determin- tion
                            | ation of   of
                            | gender   formu-
Past | of G'   lation
Tense of | Singular   of G'
all |
aspects |

Determination of number

Infini-
tive   →

Partici-
ple

Gerund

↓

to adjective translation rules

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 nor-
mal 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 informa-
tion 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 interconnexions 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. Stylistic 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 pre-
sent 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 VI.00.00.01;

    (iii) if there is no such A, 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 V, 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.6.0; the word hogy; 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 мочь [to be able] standing before an infinitive,
omit all these forms except that to the farthest left (see
example No.10).

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

3) Не можно [it is not allowed] is always substituted
for нельзя [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) Meg kell tehát
állapítanunk, hogy *p- > f- fejlödés egylátalában nem kivételt-
elen (Bev., p. 63). All the words in this sentence, except
for the fourth and sixth, are found in the dictionary. Rules
II.1, II.10 (they are not verbs) and II.10.0 (no remainder)
come into effect. The information on them is taken from the
dictionary. The sixth "word" is a formula: *p- > f-. Accord-
ing to the rule for the treatment of formulae (1.3.d) we
place it after the translation of fejlödés. For the fourth
word rules II.0, II.00 and II.00.1 come into effect; we find
meg-, join it to állapítanunk and find megállapítanunk in the
dictionary.

Rules II.1, II.11 and II.11.0 now come into operation (meg-
has been included in the word being scanned, and the sentence
no longer contains a verbal prefix); rul II.10.1 then comes
into effect, since there is a remainder - -anunk. We turn to
TABLE 3 and find:

Suffix \ Meaning	Tense	Per-son	Num-ber	Con-juga-tion
—*n*—	((	((	((	((

According to rules <u>II.10.11</u> and <u>II.10.11.0</u> we scan the table for -<u>nunk</u>, and find:

Suffix	Tense	Per-son	Num-ber	Conjuga-tion
-<u>nunk</u>-	imperf. infin.	1st	plural	undirec-ted

Using explanation 6 to Table 3, we obtain informa-tion on the ending -<u>anunk</u> : 1st person plural undirected imperfective infinitive; scan for ending. The <u>Homonym Differentiation Rules</u> 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 <u>egyáltalában</u>. Usage No.6 in the usage dictionary is suggested in the sentence and we substi-tute the information on usage No.6 (отнюдь не [by no means]) for the information on the words <u>egyáltalában</u> and <u>nem</u>. Now we change to analysis.

In translating the noun <u>fejlődés</u> we keep its number as in the Hungarian according to rule <u>V.A</u>. Since there is no information as to number, it is singular (see instruction).
Now we look for its case. Rules <u>V.B.0</u> and following lead to the conclusion of nominative case (rule <u>V.B.00.00.0</u>). It is not necessary to determine person (rule <u>V.C.0</u>). Thus, <u>fejlődés</u> is развитие [development]. It is not necessary to subject the adjective <u>kivételtelen</u> to treatment, since it is marked "indeclinable" in the dictionary. Now we translate the verb <u>megállapítanunk</u>. Rule <u>VIII.A.5</u> comes into operation and then <u>VIII.E.3.a</u> (formulation determination): the verb is preceded by a word having characteristic No.3, and does not end in -<u>ni</u>. We therefore place a 1st person plural personal pronoun (according to the -<u>anunk</u> ending) in the dative (i.e., нам [us]) before the translation of <u>kell</u>. Thus, <u>megálla-pítanunk</u> becomes нам... установить [us....to establish].

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

Итак, нам нужно установить, что развитие *p- > /- отнюдь не без исключения.

[Thus, we must establish the fact that the development of *<u>p-</u> > <u>f-</u> 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: <u>idegen szavak mindig hatolnak be a nyelvbe</u> (M.sz., p. 3). It is suggested that the reader carry out the trans-lation of (10) independently; <u>Különösen erős volt azonban a magyarság kapcsolata a kazárokkal</u> (M.sz., p.42) supplies another example.

Scanning the dictionary and tables is easily done.

The verb __lenni__ bears an instruction regarding a usage (8-1).
But this usage - __szo lenni__ - is not indicated in the sentence,
and the verb __lenni__ must be translated separately.

The sentence contains homonym forms also: VP_p - __volt__; with
the aid of rules __III.B.0__ 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.C.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 __kapcsolat__
- No.4. According to governing table No.4 we find: __S' -val__
= C [with] + N' (instrumental). In this __kazárokkal__ modifies
__kapcsolata__.

Then we turn to nouns. __Magyarság/венгр__ [magyar] (plural)
is subjected to rules __V.A.__ and __V.B.01.01__. We obtain венгров
[of the Magyars]; it is not necessary to determine the per-
son for this word (according to rule __V.B.0__). Thus, __magyarság__
depends on __kapcsolata__.

__Kapcsolat/a__ /связь [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__. But we must know the case of the noun nearest on
the right, i.e., __kazárokkal__. We switch to it.

__Kazárok/kal__/хазар [Khazar-] 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
с хазарами [with the Khazars]. Now return to __kapcsolata__.
Rules __V.B.00.10__ and __V.B.00.10.0__ come into effect and we
obtain nominative case for __kapcsolata__; it is not necessary
to determine person. We have связь.

Translation of adjectives: for __különösen__ /особенно
[particularly] we have information from TABLE No.2: A_{ad}
(neuter, singular, nominative, abbreviated form), i.e., особ-
енно . __Különösen__ as A_{ad} modifies __erős__.

__Eros/сильн-__ [strong] comes under rules __VI.00.00.0__ and
__VI.00.00.01__ (since the sentence contains the nominative noun
связь): nominative, abbreviated form, gender and number as
in связь ; i.e., сильна.

The verb __volt__ (i.e., __lent__ provisionally) is the past of
быть [to be]. Rule __VIII.A.3__ comes into operation; it refers
us to __VIII.C__ (we obtain singular), and then to __VIII.D__.
According to rule __VIII.D.1__ we establish the feminine gender
for __volt__ (the sentence contains a nominative noun, __kapcsolata__
which is feminine). It is not necessary to determine the
formulation (__VIII.E.1.b__). As a result we obtain была .

The general translation with indication of inter-relations

is:

*особенно сильна была однако венгров связь с хазарами».

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

Using the instructions on word arrangement, we get:
«Связь венгров с хазарами однако была особенно сильна».
[The bond between the Magyars and the Khazars was particularly strong, however].

Let us consider example (12): .... s ezért a legtöbb nyelvenek igen sok idegen eredetű szava van (Bev., 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 IV.00.11, and from the postpositional code of the verb lenni we determine its translation: в (у) [in(at)] + N' (prepositional or genitive case).

Nouns: nyelvnek/ язык [language] according to rule V.A is translated in the singular, and the case is obtained according to the translation of -nek; it is not necessary to determine person. As a result we have в (у) языке (языка) [in (at) the language].

Eredetű/ происхождени- [origin-] is genitive singular according to the noun suffix table; it is not necessary to determine person. We obtain происхождения . With this eredetű modifies szava. Szav/a = szo/a/слов [word] which is nominative singular (according to rule V.B.0 and following to V.B.00.00.0); it is not necessary to determine person (according to characteristic No.6 in the verb lenni); we get слово.

Adjectives: sok is first subject to rule VI. It undergoes rules VI.00.01 (sok stands before A" idegen), VI.00.01.1 (A" - idegen - stands before S' - eredetű and VI.00.01.11 (S' - eredetű - depends on S - szava). Hence sok modifies szava, and is considered as standing before the latter. In the dictionary article we select a suitable translation (before S without III or II sf): много [many] with the instruction "indeclinable". Therefore we do not determine the number gender and case for много, and change (since sok - A^C) to rule VII.B (see instructions regarding A^C before rule VI).

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 szava to the genitive plural: слов. General note 1 is not fulfilled: sok - A^C. Idegen is translated (according to rule VI.1) in the gender, number and case of eredetű and modifies the latter: иностранного [of foreign].

The verb van (=len): rules VIII.A.1.c., VIII.B, VIII.C and VIII.E.1.b give the form есть [there is, are]. It is not necessary to determine the formulation, since the sentence

*Regarding instructions on the usage with lenni (van), see preceding example.

contains a noun which is found to be in the nominative after
the operation of rule V.B - this is szava. Finally, having
fulfilled the instructions given in the dictionary article
for the word legtöbb, we put the translation of the noun
nyelvnek into the genitive plural and the translation of the
word legtöbb into the prepositional (genitive) case following
the preposition в (y). The general result is:

в большей части языков очень много иностранного происхождения слов

есть.

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

«в поэтому очень много слов ино-
странного происхождения есть в большей части языков».

[and therefore there are many words of foreign origin in most
languages].

The last example we will consider is (13): Egymással
érintkezésbe került nyelvek egymáshoz bizonyos mértékben
közelednek (Bev., p.21).

For the word bizonyos there is an instruction regarding a
usage. Usage No.2 - bizonyos mértékben - is the same as in
our sentence. We substitute the information on the corres-
ponding words by that on the usage: в известной мере
[to a certain degree], Ad$_{pp}$.

The homonym differentiation rules then come into operation
as the VP$_p$ homonym form - került - must be treated. Rule
III.B.1 operates; as there is no determinative to the left
of került, we scan sentence for another verb (III.B.0), we
find it - közelednek (rule III.B.10.1.b), see that it is not
VP$_p$ (rule III.B.10.10), and establish that került is a past
perfective participle from the verb попа- (-сть)[to enter],
i.e., попавш- . Change to analysis.

Postpositions: - sal = -val; pospositional code of the
nearest word (érintkezésbe) - No.4; from governing table No.
4 we find the translation: S' -val = c [with] + N' (instru-
mental). Egymással depends on érintkezésbe. -be; according
to the postpositional code for került we find the translation:
S' -bel = в [into] + N' (accusative). Erintkezésbe depends
on került. -hoz; we translate according to the postpositional
code for közelednek: S' -hoz = c [with] + N' (instrumental).
Egymáshoz depends on közelednek.

Nouns: egymással is translated as друг с другом [one an-
other] (from the additional instruction in the dictionary
article): érintkezésbe/ в контакт [in contact] (the number
is determined according to rule V.A. - singular; it is not
necessary to determine person); nyelvek/ язык [language] is
nominative plural (according to rules V.B.01 to V.B.01.00);
it is not necessary to determine person; i.e., языки;
egymáshoz - друг с другом (see above).

Verb translation remains: került; rule VIII.A.4 refers us
to the adjective translation rules, from which (rule VI.1) we
find that került is translated in the same gender, number and
case as nyelvek, i.e., попавшие . For közelednek the

translation of the 3rd person plural present сближаться
[draw together] is easily obtained from the table.

The general translation is:

«Друг с другом в контакт попавшие языки друг с другом в известной
мере сближаются».

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

We rearrange the words: we change the order of попавшие
(according to rule 2.B) and the words depending on it, and
insert commas; next we place the second друг с другом
(according to instruction I.D.a). It is not necessary to move
either языки , since this is a nominative noun and stands
before the verb (instruction A.a), or в известной мере
since this is an independent grouping. The final result is:

«Языки, попавшие в контакт
друг с другом, в известной мере сближаются друг с другом».

[languages which have been in contact draw together to a cer-
tain 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 forma-
tion 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

<u>Characteristics</u> makes possible the creation of less cumbersome <u>rules of analysis</u>. 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 <u>jelboritott</u> <u>csúcs</u>/покрытая льдом вершина   [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.