THE MECHANICAL TRANSLATION OF RUSSIAN:
A PRELIMINARY STUDY

THIS REPORT is a summary of investigations concerning the mechanical translation of the Russian language into English. More specifically, this study has been concerned with Russian scientific writings, rather than with the Russian language in general. It is only within this limited sphere of "scientific Russian" that our mechanistic and perhaps naive approach is valid. In the first part of the report some of the most immediate problems are set forth in simplified form, and general solutions are indicated. The second part (Appendix I) offers a concrete solution in a fixed mechanical procedure. The third section (Appendix II) shows the application of this procedure to a sample Russian text.

We have dispensed with a description of the operational process. Suffice it to say that the problem is everywhere considered in terms of the National Bureau of Standards Western Automatic Computer (SWAC), and that the proposed solution has been made in consultation with Dr. Harry D. Huskey of the Institute of Numerical Analysis of the National Bureau of Standards.

The three chief linguistic problems which present themselves are those of syntax, morphology, and vocabulary.

SYNTAX

Originally, it was intended that this should be a study of the problems of Russian syntax in mechanical translation. Our point of departure was to have been an earlier study of this problem with respect to the German language.[1] Two considerations have changed the direction of our study. In the first place, the problem of morphology must take precedence in the Russian language: we must first deal with the elementary fact that any Russian noun, verb, or adjective may appear in a multiplicity of forms. The central problem of *recognition* is, therefore, the chief subject of this study. The question of the relation between words in the sentence has been partially solved, but this solution has

---

[1] Victor A. Oswald, Jr., and Stuart L. Fletcher, Jr., "Proposals for the Mechanical Resolution of German Syntax Patterns," *Modern Language Forum,* Vol. XXXVI, No. 3-4.

not taken the form of the rearrangement of Russian sentence patterns into English patterns, as in the study of German syntax. Our tentative conclusion is that the latter approach to the syntactical problem is both inordinately difficult and, more importantly, of doubtful utility. In the field of scientific writing, Russian sentence structure is definitely, close to English—much closer than is normal for other forms of Russian prose.

Russian word order is normally quite flexible, because of the highly inflected nature of the language. This flexibility is markedly reduced in the types of scientific discourse which we have studied. The following figures illustrate this fact, with respect to the basic elements in the sentence. The figures were obtained from a survey of a wide variety of articles in the *Reports of the Academy of Sciences of the U.S.S.R. (Doklady Akademii Nauk S.S.S.R.)*.

1. *Relative position of subject-verb*

From a sample of 1,528 sentences containing a subject and verb:

Subject before verb: 81% of all occurrences

(195 additional sentences contained an impersonal, or understood subject: 24 sentences contained no verb)

The position of subject before verb (normal in English word order) thus appears to prevail approximately four-fifths of the time. Included in the verb-subject position were a large number of passive or reflexive verbs; here the "inverted" word order is also common in English.

Takim obrazom poluchaetsia uravnenie.[2]

Thus is obtained the equation.

2. *Relative position of verb-direct object*

From a sample of 1,129 sentences containing a verb and direct object:

Verb before object: 88% of all occurrences (including clauses)

Object before verb: 12% of all occurrences.

Included in the object-verb position were a number of constructions common in English.

Takoe pole my budem nazyvat' nepreryvnym.

Such a field we will call continuous.

3. *Relative position of adjective-noun*

From a sample of 1,484 occurrences:

Adjective before the noun modified. 1468 occurrences. Noun before adjective: 1 occurrence

(Other occurrences were predicate adjectives and comparative adjectives, where the position is similar to that of English)

---

2 The Library of Congress transliteration system is employed in this study with the exception of diacritical marks, which have been omitted.

These figures indicate that the basic sentence structure of scientific Russian is quite similar to English. Although they offer no guarantee that the Russian sentences will always be intelligible if translated word-by-word into English, they seem to imply that the prerequisites of such translation exist. The necessity for re-arranging the Russian sentence into rigidly fixed English pattern seems doubtful. It may well prove desirable to establish proced-ures for dealing with certain specific types of participial construc-tions. Our tentative position is, however, that the Russian text will be intelligible in English when the word order is left undis-turbed. A specimen word-for-word translation is given in Appen-dix II.

MORPHOLOGY

The solution of the morphological problem is imperative, and far more difficult than, for example, in German. Thus, a Russian noun, adjective, or participle may appear in a given sentence in any one of six cases, singular or plural; the inflectional endings vary by case and by gender. The verb may appear, in scientific writing, in any of twelve forms. When a given word appears in any of these forms, the primary mechanical problem is, then, one of recognition: to determine what the word actually is, without its inflectional ending. The second problem is to deter-mine the syntactical function of the word. The procedure estab-lished in Appendix I is an attempt to solve the problem of recogni-tion, i.e.. identification of the word by its stem, listed in the dic-tionary. The stem is identified through the procedure of remov-ing the inflectional ending, letter by letter; a by-product of this process is the information (although incomplete) about the grammatical function of the word: case, number, tense of verb, adverbial participle or participial form, etc.) Some of this infor-mation is useful in translation; some of it is not useful. The arbitrary selection of information classified as "useful" is indi-cated below.

The adjective. The identification of an adjective *by case* is not deemed necessary for our purposes. The inflexible position of the adjective before the noun modified makes the meaning clear and the word order natural for English. Our only concern is to strip the adjectival endings for purposes of recognition.

It may prove ultimately desirable to retain one additional item of information: the plural number of adjectives. Since it is difficult, or impossible, to ascertain the plural number of cer-

tain classes of nouns, this information can be obtained from the adjective, *when an adjective is used with the noun.* (The plural number of adjectives is immediately identifiable by the final letter: -e, kh, or i.) Adjective endings may also be used in some instances to identify the *case* of nouns, when this information cannot be obtained in other ways. This utilization of the adjective has not been incorporated in the procedure outlined in Appendix I.

The verb. The problem of identification of verb forms is less difficult in scientific Russian than in normal Russian prose; scientists very rarely make use of the imperative or the first personal singular, present tense, and never employ the second person, singular or plural, present tense. In the present (or the future) tense, therefore, we need be concerned only with three forms: Third person, singular, and first and third persons, plural. The following also require identification: infinitive, past tense (four forms), and the present and past adverbial participle (four possible forms). There is a total of eleven forms which we must be prepared to distinguish, e.g.

delaet, delaem, delaiut, delat', delal, delala, delalo, delali,
delaia, delav, delavshi
(A variation such as derzha in the adverbial participle, must
also be considered.)

We have chosen, again arbitrarily, to use at least some of the information provided by each of these forms. Specifically, we have extracted and retained information about: the infinitive, the singular or plural number of the verb, the past tense, and the adverbial participle. The determination of the future perfective is probably unnecessary in scientific writing.

The participle. All that has been said of the adjective is applicable to the participle: it is unnecessary to identify the participle by case, although it may prove helpful to identify the number. It is, of course, necessary to determine the *active or passive nature* of participles; the distinction between past and present participles, on the other hand, is of questionable utility, and may even be confusing in translation. We have, therefore, incorporated a procedure to identify participles as such, and to designate their active or passive nature.

The noun. Nouns present the greatest problem for identification. The difficulty can best be illustrated in the following table of regular noun endings. The symbol (*) indicates that the noun

is identifiable by case (and number), assuming that the gender of the noun is known, (In the following table, and in Appendix II, a distinction has been made between the Russian letters "i" and "i short.")

	Masc.		Neut.		Fem.	
	S.	Pl.	S.	Pl.	S.	Pl.
Nom.	-zero -i (short)	-y -i	-o -e	-a -ia	-a * -ia*	-y -i
Gen.	-a -ia	-ov -ev -ei	-a -ia	-zero* -ei* -i(short)*	-y -i	-zero -ei -' -i (short)*
Dative	-u* -iu*	-am* -iam*	-u* -iu*	-am* -iam*	-e -i	-am* -iam*
Acc.	-zero -i (short) -' -a -ia	-y -i -ov -ev -ei	-o -e	-a -ia	-u* -iu*	-y -i -zero
Instr.	-om* -em*	-ami* -iami*	-om* -em*	-ami* -iami*	-oi* -'iu* -ei	-ami* -iami*
Prep.	-e*	-akh* -iakh*	-e -i*	-akh* -iakh*	-e -i	-akh* -iakh*

(Note: This chart does not take into account such variations as the masc. gen. s. -u, -iu, the prep. s. -u, -iu, or the nom. pl. -a, -ia; these variations are relatively rare in scientific Russian. If a further distinction is drawn between animate and inanimate nouns, the following could also be identified: masc. animate nouns in nom. s; masc. inanimate nouns in gen. s. and gen. pl; fem. inanimate nouns in gen. pl.)

It is immediately obvious that the majority of case endings are not identifiable by ending alone. There are, to be sure, certain auxiliary devices to aid in identification. Thus, if we rely on adjective endings and/or prepositions occurring before the noun, the case of the noun would be more often identifiable. But "more often" is not enough. To state the obvious, nouns need not be preceded by prepositions or adjectives; further, when they are so preceded we still have no guarantee as to the case of the noun, since many prepositions are used with more than one noun case, and since the adjective ending is often ambiguous as to case ending. The position of the noun in the sentence may also offer some assistance. A combination of these four factors (noun ending, accompanying adjective or preposition, and position) may provide a partial solution to the problem of identification. No attempt has been made here to determine how nearly perfect a solution could be thus attained, or how adequate such a partial solution would be to a proper understanding of the Russian text. It is clear that such a complex solution, involving multiple factors, would not be practicable for a machine of the limited capacity of SWAC.

The question then arises: Is it necessary to identify the grammatical case of each noun in a Russian sentence? How useful is this knowledge for translation purposes? Upon this point, which naturally involves the problem of syntax, the following generalizations have been made :

Nouns following a preposition.

Identification of the noun by case is not necessary. The different meaning of certain prepositions with different cases is definitely a less critical factor in scientific Russian, and can probably be handled by listing two alternative meanings for the preposition. (I f necessary, a procedure could be established to clarify the meaning of prepositions of double or triple meaning.)

Nouns without a preposition.

1. Nominative and accusative case. Identification of doubtful utility, because of the factor of position, referred to under "Syntax". The plural number of nouns which are the subject of the sentence will be indicated by the plural number of the verb.

2. Genitive case.   Identification by ending alone is impossible in all types of nouns (except inanimate masculine). The

indication of possession (without the use of a preposition, in Russian) is normally made by reliance on word order:

kolichestvo soli   izuchenie podvizhnosti atomov
quantity of salt   a study   of the mobility of atoms

In scientific Russian, this word order is invariably preserved: the noun in the genitive case never precedes the "qualified" word. Since this modification is normally one noun or another, a pattern is established; when one noun follows another, there is an overwhelming probability that the second noun will be in the genitive case. In an actual count of 617 occurrences of this noun-plus-noun pattern, the second noun was found to be in the genitive case 93.4% of the time; the second most frequent case (3.9%) was the instrumental case, which can be identified by ending. These figures indicate that the genitive case can be identified *bv position* over 97% of the time in this pattern.

In other usages of the genitive case (such as the direct object of a negative verb, after a negative expression, after certain verbs, etc.) there would be no advantage in identifying the case of the noun.

3. Dative case. A survey was made of nouns appearing in the dative case without a preposition. From a sample of 286 pages of scientific text, 137 such instances were found—approximately one occurrence for every two pages. The notable fact was that the noun in the dative case was found to appear in conjunction with a relatively small number of verbs and adverbs. A list of these words, 29 in number, is given in Appendix III; the list is, of course, incomplete, but there is reason to believe that if the survey were extended it would be possible to anticipate with great accuracy the occurrence of nouns in the dative case without a preposition. The correct translation can be effected by appending the English word ("to") to each of the given words, as necessary. (Certain Russian verbs take the noun in the dative case instead of the expected accusative; in such instance the identification of the noun as dative would serve no useful purpose.)

The use of nouns in the dative case as the indirect object is almost always comprehensible in scientific Russian. English, of course, relies upon word order when the preposition "to" is omitted: "I gave mother a book", not "I gave a book mother." Russian has no such fixed word order, and if the word order is preserved in translation, we may get an occasional instance of the latter construction. There were no such instances in the survey

cited above. In addition, the indirect object was found to occur normally after the verb, as in English.   The use of the dative in impersonal and idiomatic expressions, so common in normal Russian prose, is quite limited in scientific Russian. *Summary:* Identification of the dative case is of doubtful utility; its chief value would be to permit rearrangement of the sentence to conform with English word order. (Dative endings of all masculine and neuter nouns are immediately identifiable; those of feminine nouns are not.)

4. Instrumental case. The identification of a noun in the instrumental case, without a preposition, is clearly necessary. The "by means of . . ." meaning of this case cannot be indicated without such identification.   Fortunately, this case is immediately recognizable by the ending of all masculine and neuter nouns, and most feminine nouns.   Since the need for identification is imperative, it appears advisable to code a separate category of nouns in the dictionary listing, i.e., feminine nouns in -iia.   (The instrumental singular ending of such nouns (-ei) can then be differentiated from the same genitive plural ending of feminine nouns in -'). Thus all nouns in this case are identifiable by ending.

5. Prepositional case.   Identification serves no useful purpose. Nouns in this case are always preceded by a preposition. (The different meaning of the preposition before other cases appears, again, not to present a serious problem in translation.)

Summary of Nouns.

1. It is questionable   that the   identification of nouns by case is necessary, with the exception of the instrumental and genitive cases. It is therefore proposed that the case endings be disregarded for the nominative, dative, accusative, and prepositional cases.

2. The plural number is immediately recognizable in most classes of nouns (dative, instrumental, and prepositional cases of all nouns; nominative, genitive, and accusative, all masculine nouns; genitive of neuter and feminine nouns). Those which cannot be identified as plural, by ending, are neuter and feminine nouns in the nominative and accusative plural. A sizable portion of these—those which are the subject of an expressed verb—are accounted for by the plural of the verb; others are denoted as plural by their occurrence after certain words (numerals greater than one, "both", "many", etc.). Still more can be identified by the plural form of the adjective. For the present, it is proposed

that only nouns in the dative, instrumental, and prepositional plural be denoted as plural.

VOCABULARY

In view of the limited capacity of SWAC, we have restricted our field of operations to a single branch of science—mathematics. Here, we are fortunate enough to possess a good working vocabulary, a *Russian-English Vocabulary* prepared by the American Mathematical Society.[3] This glossary of some 2800 words is remarkably complete in its coverage of all Russian words likely to be encountered in the reading of mathematical paper. (An actual word-by-word check of 30 pages of mathematical articles disclosed a total of four words of Slavic origin which were not included in the *Vocabulary*. There were, however, many words not listed in the *Vocabulary* because they were Russianized versions of English or international scientific terms. Such words, recognizable to the human reader, would of course be meaningless to our mechanical reader. Unless a system of transliteration were to be incorporated in the machine's output, these words would have to be added to the *Vocabulary*. A more extensive survey would probably reveal that the total of such words would not exceed two hundred. It should be added that the *Vocabulary makes* separate entries for all pronouns and most "irregularities" in noun, verb, and adjective formation, precisely as we must do for purposes of mechanical identification. Thus, it appears that a Russian vocabulary of some 3000 words would be adequate for the translation of mathematical papers.

Micro-glossary. The problem of preparing a micro-glossary from the *Vocabulary* (reducing each word to the smallest possible number of letters which will permit recognition) has not been investigated. A preliminary check of several pages of the *Vocabulary* indicates that more than two-thirds of the words contain less than seven letters (the limit for a single entry in the SWAC machine.) It is unquestionably true that this percentage can be significantly increased if we alter the procedure of listing set forth in Appendix I. In this procedure, for example, most nouns are listed in their entirety, except for the final letter. It is obvious that the meaning of many nouns consisting of more than eight letters can be distinguished at some point before the eighth letter.

---

[3]*Russian-English Vocabulary with a Grammatical Sketch,* prepared under contract with the Office of Naval Research by the American Mathematical Society, 1950.

Thus, the noun sokhranenie, consisting of nine letters in our listing (sokhraneni) can be identified by meaning and can be distinguished from the similar verb form at the seventh letter (sokhrane). The adaptation of a micro-glossary would measurably increase the size of our machine vocabulary; it would also necessitate certain revisions in the procedure outlined in Appendix I.

Homographs. In general, the problem of homographs is not serious in specialized scientific discourse, where it is unlikely that more than one meaning of a homograph be applicable. The practice of listing words in an abbreviated form, inherent in our system, does, however, increase the number of potential homographs: the shortened form of a noun may coincide with the shortened form of another noun, verb, or adjective. Thus, vod is the abbreviated form for at least two words, voda (water) and vodit' (to lead); tsel is the shortened form of tselyi (entire) and tsel' (aim), in addition to the verb tselit' (to aim). This problem can be met in at least two ways:

1. The ambiguous words can be listed in two or more slightly less-abbreviated forms: the verb vodit' as vodi and vodiat —forms which cannot be mistaken for the noun in any of its forms.

2. The determination of the part of speech (and therefore, usually, of meaning) can almost always be made through the inflectional ending, i.e., the noun voda can be distinguished *as a noun* because of its endings, only one of which (-a) is a possible verb ending.

Both of these methods would necessitate further adjustments in our coding procedure, and no solution has been attempted in the present study. It appears doubtful that the number of homographs in the *Russian-English Vocabulary* will present a serious problem of ambiguity.

APPENDIX I

CODING PROCEDURE

The following procedure is presented in a descriptive, non-technical manner. The section, "Dictionary Listing", summarizes the main criteria for listing Russian words in the machine "dictionary." Sample entries are given for the main classes of words to be listed. These entries are intended merely as a guide, for ultimately the efficiency and usefulness of the dictionary depend upon the skill of the language expert and his acquaintance with scientific Russian. For our purposes, the conventions of lexicography are of little use; here, "the word" is the abbreviated form listed in the dictionary, not the infinitive of the verb, the

masculine nominative singular of the adjective, etc.   Our chief criteria are utility and the avoidance of ambiguity.

In the Coding Procedure proper are outlined the successive steps which must be taken before the Russian word can be identified by meaning and. if necessary, by grammatical function. Essentially, the process entails the immediate comparison of the Russian word in the text with the dictionary; if an exact equivalent is not found, the final letter is stripped and the comparison is repeated. This process is repeated until an exact identification is made with the shortened form listed in the dictionary. Four "steps2 have been established; certain specific information is gathered about different words at each of these steps, which are mutually exclusive.

I.   *Dictionary Listing*

1. All non-inflected words

2. Adjectives, without endings
> khoroshii: khorosh
> matematicheskii: matematichesk
   a. Short masc. adjectives with vowel insertion are listed separately
   > nizok; polon
   b. Adjectives serving as nouns (masc.) are listed in the additional forms of instrumental sing, and pl.

3. "Regular" verbs (whose endings are added to the stem of the infinitive) , listed without endings
> first conjugation: chitat': chita; otdeliat': otdelia
> second conjugation: govorit*: govor

4. "Irregular" verbs, listed in as many forms as is necessary for identification
> zhdat': zhda, zhd
> vesti: vesti, ved, vel
> trebovat': trebova, trebu

5. A selected number of past passive participles, without endings, to include:
   a. Those of very high frequency in scientific writing (dokazan)
   b. Those not reducible to the infinitive stem as defined in 3 and 4 above.
   > broshen

6. Nouns, in the form of:
   a.   Nom. sing. for masc. nouns ending in a consonant
   > znak
   b. Nom. sing. for all other nouns, without the final letter

slucha	mest	sil
dvigatel	pol	bashn
	trebovani	funktsi
		smec

   c. Nouns with "fugitive" vowel require one additional listing
   > konets, konts
   d. Nouns with vowel insertion in gen. pl. require one additional listing
   > tochka: tochk, tochek

7. All other words characterized by internal consonantal change
> blizkii: blizk, blizhe

8. First and third person pronouns
> ego, nego: ei. nei

9. Numerals and all other inflected words, in as few forms as possible to permit identification
> odin, odn; dv, dvu; chetyr

10. Words of very high frequency, idioms, etc., as needed
> iavliaetsia; govoriat

II. The following code numbers will be attached to Russian words in the dictionary listing :

1. Masculine noun
2. Neuter noun
3. Feminine noun
4. Adjective
5. Verb
6. Past passive participle
7. Feminine noun in -iia
8. Preposition

(The following abbreviations are used below) :

p—past tense	a—active
s—singular	r—passive
m—plural	t—participle
g—"gerund" (adverb, participle)	sup—superlative
inf—infinitive	in—instrumental
com—comparative	

III. *Translation Procedure*

No attempt is made here to rearrange Russian word order into English word order. A strict word-by-word translation is implied, according to the following procedure:

STEP I

A comparison will first be made of the word in text with the dictionary. If it is the exact equivalent, the word is translated immediately.

1. Words identified at this point which are coded No. 4 or No. 6 will be prefaced by the English ("s") .

2. Fem. and neut. nouns identified here are denoted as gen. pl.

(*Note*:   In Step I are identified:
1. uninflected words
2. masc. nouns in the nom. and acc. sing.
3. certain fem. and neut. nouns in the gen. pl.
4. short forms of masc. adjectives and of past passive participles which are listed in the dictionary.)

If no identification is made, proceed to Step II.

STEP II

Strip the final letter of the word in text, and compare with dictionary. If the final two letters are -sia or -s', ignore both letters in the stripping process and begin stripping with the third letter from the end of the word. Denote such words as reflexive (r).

1. Words identified at this point which are coded No. 4 or No. 6 will be prefaced by the English ("be") when such words end in -a, -y

2. Identified words coded No. 5 will end in one of the following letters:
   -1, -ia, -a, -t
   1) -1 (denote past tense)
   2) -ia, -a (denote adverbial participle)
   3) -t (denote past passive participle)

3. Any noun (Code Nos. 1, 2, 3, 7) will be denoted as gen. (prefaced by English "of") if the preceding word is a noun, disregarding intervening adjectives.

(Note: In Step II are identified:
1. all nouns of a single-letter ending
2. the past tense. masc., of words in -l which are regular first conjugation verbs
3. adverbial participles of regular first conjugation verbs
4. short forms, fem. and pl., of adjectives and past passive participles. These forms are invariably predicate adjectives; since the verb is often lacking in Russian, the English "be" is inserted in translation, in parentheses.)[4]

Preliminary checks indicate that 67% of the words in the Russian text will be identified in Steps I and II. If no identification is made, proceed to Step III.

STEP III

Continue the stripping process, one letter at a time, comparing each time with dictionary. In Step III, no more than the *final three letters* will be stripped.

In addition to identification of the word by meaning, we are interested in learning the grammatical function of the word; specifically, we wish to determine the instrumental case of all nouns, the plural number of nouns, the comparative or superlative degree of adjectives, the number and past tense of verbs, the adverbial participle form, and the active or passive nature of participles. In order to this, we must now make use of the final letter of the word, either by itself, or in conjunction with the preceding letter or letters. At the Step III level, there exist twelve possible final letters which

---

4 Short neuter forms of the adjective and past passive participle constitute a problem, because of confusion with the adverb. No solution is offered here to this problem. Adverbs will be translated as adjectives, unless they have a distinct form.

It may be feasible to denote masc. nouns in -a, -ia as gen. case at this point, since the vast majority of masc. nouns in mathematical writings are *inanimate,* and therefore identifiable in this case.

are critical for our purposes. (Two additional possible final letters, -ia and -u, have been ignored, since the functions which they designate are of doubtful utility for our purposes.)   These twelve letters are listed below. Those numbered 3 and 4 will indicate, at the Step III level, certain information in and of themselves. The remaining letters will require reference to the preceding letter or letters. It is assumed that the machine will first determine which of the twelve letters is, in fact, the final letter of the word. When this letter has been determined (together with its code numbers) stripping continues until the meaning has been found: the code numbers must be retained before an indication of the function can be given. For letters No. 11 and 12, a further comparison of the code numbers with the dictionary coding of the part of speech must be made. Identified words for which no function is indicated will, of course, be translated without further notation.

1. -A
    1. -1 (p. s)
    2. -m, -t (r, t)

2. -O
    1. -1 (p,s)
    2. -m, -t (r, t)

3. -' (inf.)

4. -L (p.s.)

5. -T
    1. -e, -i (s)
    2. -a, -ia, -u, -iu (m)

6. -B
    1. -a. -i (g)
    2. -o, -e (m)   (gen)

7. -I
    1. -sh (g)
    2. -1 (p.m)
    3. -m
        1.   -a, -ia (in,m)

8. -IU
    1. -' (in)

9. -E
    1. -e (com)

10. -KH
    1. -a, -ia (m)

11. -I (short i)
    1. -o (in) if word is coded No. 3
    2. -e (in) if word is coded No. 7
    3. -e (m) (gen) if word is coded No. 1 , 2 . 3[5]

---

[5]At Step III. it is probably more economical to discontinue the identification of the genitive case by position.

12. -M

    1. -o. -e (in) if word is coded No. 1 or No. 2

    2. -a, -ia (m)

Note: The designation ("in"), instrumental case, will be dropped when the noun is preceded by a proposition, disregarding intervening adjectives.

STEP IV

Any word not identified after the final three letters have been stripped is either a participle or (rarely) a superlative adjective. In Step IV, the machine will first determine which of the following letters is, in fact, the second, third, or fourth letter from the end of the words: shch, m, n, t, sh. (One of these letters will always appear in these positions.)

1. -shch (denotes active participle — a, t)

    pokazyvaiushch/ii

2. m, n, t (denote passive participle — r, t)

    chitaem/yi

    chitann/yi

3. sh, refer to the preceding letter.

    1. i (denote superlative adjective — sup)

        noveish/ii

    2. All other letters, denote as active participle — a, t

        chitavsh/ii

After the above has been determined, stripping continues until the dictionary listing of the word (verb or adjective) has been identified. The above notations, which have been retained, are then transmitted with the translated word.

Appendix II

TRANSLATIONS

In the sample Russian text given below, the slant indicates the separation of the ending from the "stem" form listed in the dictionary.

In the word-for-word translation, the plural of nouns and verbs, the past and reflexive of verbs, etc., are incorporated in the words themselves, rather than indicated by symbols. The following conventions are to be understood:

1. Alternative translations of a single Russian word are indicated by the slant: from/of

2. A word in parenthesis may be used or disregarded, according to context. Such words may seem repetitious; if so, they should be ignored. Note especially the word ("by"), which will sometimes be meaningless or confusing in English translation. It may easily be possible to disregard the instrumental case of nouns when used with certain verbs: for the present, the superfluous "by" has been retained.

3. The word "by" should be placed before adjectives or numerals preceding the noun: "two (by) points" implies "by two points."

The language of the translation may be characterized as "headline language" or "telegram language." The former, for sake of brevity, dispenses with articles and forms of the verb "to be." The latter, in addition, often dispenses with the use of prepositions, as well as pronouns.

"Announcement your appointment received"

"Will arrive evening June   16."

The fact that prepositions may be omitted by common consent has led us to attempt an alternate translation in which the genitive case of nouns has not been identified.

I.   *OTREZOK*.   Poniati/e   otrezk/a   izvestn/o   iz   elementarn/oi geometri/i: otrezok est' chast/' priam/oi, ogranich/en/naia s obe/ikh storon dvum/ia tochk/ami.

Sushchestvenn/oi   kharakteristik/oi   otrezk/a   iavliaetsia   ego   *dim/a*. Napomn/im   vkrattse,   chto   podrazumeva/iut   pod   etim   termin/om.   Voz'm/em kakoi-libo   proizvol'n/yi,   no   opredelenn/yi   otrezok   v   kachestv/e   edinits/y dlin/y i etim otrezk/om izmer/im dann/yi otrezok. Dlin/oi dann/ogo otrezk/a nazyvaetsia   chisl/o.   pokazyva/iushchee,   skol'ko   raz   otrezok,   priniat/yi   za edinits/u. soderzh/it/sia v dann/om otrezk/e. Mozhet, konechno, sluch/it'/sia, chto otrezok-edinits/a ne ukladyva/et/sia tsel/oe chisl/o raz v dann/om. Togda dlin/a dann/ogo otrezk/a budet ne tsel/ym chisl/om, a drobn/ym ili zhe irratsional'n/ym: posledn/ii   slucha/i,   sobstvenn/o   govor/ia,   i   est'   naibolee obshch/ii.

Vse eto khorosho izvestn/o iz elementarn/oi matematik/i. Dlia nas vazhn/o lish'   zapomn/it',   chto   pod   DLIN/OI   dann/ogo   otrezk/a   my   vsegda   budem podrazumeva/t'   POLOZHITEL'N/OE   CHISL/O,   polucha/emoe   v   rezul'tat/e izmereni/ia otrezk/om, vybra/nnym za edinits/u.

2 OC/'.   V   analitichesk/oi   geometri/i,   krome   poniati/ia   priam/oi lini/i. igra/et osnovn/uiu rol/' ves'ma blizk/oe k nemu poniati/e OS/I.

Rassmotr/im   kakuiu-libo   priam/uiu   lini/iu   AB.   Tochk/a,   dvizbusbcba-ia/sia nepreryvn/o po eto/i priam/oi v odnu i tu zhe storon/u, mozhet opisyva/t'   ee   v   dvu/kh   protivopolozhn/ym   napravleni/iakh: v   napravleni/i, ved/ushchem ot A k B, ili zhe, naoborot, v napravleni/i, ved/ushchem ot B k A. Vyber/em   kakoe-libo   odn/o   iz   eti/kh   dvu/kh   napravleni/i   i   nazov/em   ego polozhitel'n/ym; protivopolozhn/oe napravleni/e nazov/em otritsatel'n/ym.

Priam/aia,   nad   kotor/oi   vybran/o   opredelenn/oe   napravleni/e   v   kachestv/e polozhitel'n/ogo,   nazyvaetsia   os/'iu.   Oc/'   obyknovenn/o   izobrazha/et/sia   pri pomoshch/i   priam/oi,   polozhitel'n/oe   napravleni/e   kotor/oi   otmech/eno strelk/oi.

O   dvu/kh   os/iakh   govoriat,   chto   oni   ime/iut   odinakov/ye   napravleni/ia, kogda   oni   NE   TOL'KO   PARALLEL'N/Y   mezhdu   soboiu.   NO   I   IME/IUT ODINAKOV/YE   POLOZHITEL'N/YE   NAPRAVLENI/IA;   v   slucha/e, kogda   dve   os/i   parallel'n/y,   no   polozhitel'n/ye   napravleni/ia   ikh   razlichn/y, govoriat,   chto   eti   os/i   napravlen/y   PROTIVOPOLOZHN/O: naprimer,   na cher. l os/i X i X' ime/iut protivopolozhn/ye napravleni/ia.

I. Translation :

1.   *Segment*.   Concept   of   segment   known   from/of   elementary   geometry: segment is part straight (line), bounded from/with both sides two (by) points.

Essential   (by)   characteristic   of   segment   is   its   length.   Recall   briefly, what/that   understand   under   this   term.   Take   any   arbitrary,   but   definite   segment   in quality   of   unity   of   length   and   this   (by)   segment   measure   given   segment,   (by) Length   of   given   segment   is   called   number,   indicating,   how   many   time   segment, taken   for   unity,   is   contained   in   given   segment.   It   can,   of   course,   to   happen.

what/that segment-unity not is included whole number of time in given. Then length of given segment (it) will be not whole number, but fractional or irrational; last event, proper speaking, and is most common.

All this well known from elementary mathematics. For us important only to remember, what/that under length of given segment we always (we) will to understand positive number, obtained in result of measurement (by) segment, chosen for unity.

2. *Axis.* In analytical geometry, besides/except concept of straight (line) plays basic role quite near to it concept of axis.

Consider any straight (line) line AB. Point, moving continuous along/according to this straight (line) in one and same side, (it) can to describe it in two opposite directions: in direction, leading from A to B, or, on the other hand, in direction, leading from B to A. Choose any one of/from these two direction and call it positive; opposite direction call negative.

Straight (line) above which chosen definite direction in quality positive, is called axis. Axis usually is depicted with aid straight (line), positive direction which noted (by) arrow. Concerning two axis one says that they have identical direction, when they not only (be) parallel between self, but and have identical direction; in event, when two axis (be) parallel, but positive direction their/them (be) different, one says that these axis (be) directed positive, e.g., on fig. 1 axis X and X' have opposite direction.

II.   Alternate translation;

The translation given below is identical with the above, with one exception: The English preposition "of" has not been supplied for Russian nouns in the genitive case. It is to be understood that when two or more nouns appear in succession (disregarding adjectives), the word "of" is implied before all nouns in the series except the first.

"Concept segment" implies "the concept of the segment." The alternate translation may appear more cryptic and "strange". It has been comprehensible to all who have read it, however, and presents a much simpler coding procedure and a consequent saving of space for vocabulary.

I.   *Segment.* Concept segment known from/of elementary geometry: segment is part straight (line), bounded with/from both sides two (by) points.

Essential (by) characteristic segment is its length. Recall briefly, what/that understand under this term. Take any arbitrary, but definite segment in quality unity length and this (by) segment measure given segment, (by) Length given segment is called number, indicating, how many time segment, taken for unity, is contained in given segment. It can, of course, to happen, what/that segment-unity not is included whole number time in given. Then length given segment (it) will be not whole number, but fractional or irrational: last event, proper speaking, and is most common.

All this well known from elementary mathematics. For us important only to remember, what/that under length given segment we always (we) will to understand positive number, obtained in result measurement (by) segment, chosen for unity.

(The remainder of the translation is almost identical with that given above).

APPENDIX III

WORDS USED WITH THE DATIVE CASE

For our purposes, the dat*ive* case may be considered a marginal case: although it may serve a   useful   function in translation, it   is doubtful that its usefulness merits the effort   necessary   for   identification.   A   possible alternative to the establishment   of   an   identifying   procedure   is   suggested   in   this   list   of   words.   These 'words, found in a sample survey, are used with the dative case without a preposition;   by appending the English word "to" to them, we can transfer the mean- ing of the case into English   without identifying the case of the succeeding noun or pronoun.   It is obvious that this is merely a preliminary list: it is also dear that many of the words listed are so infrequent in occurrence that they are statistically insignificant in scientific discourse. A further survey of this problem would probably result in   the compilation of a   relatively small number of words which would render identification of the dative case almost completely unnecessary.

podlezhat',   proportsional'no,   sootvetstvovat',   podvergat',   prinadlezhat', podchinit',   otvechat',   posviashchat'.   pridat'.   ravniat'sia,   ravno,   blagodaria. svoistvennyi,   sleduia,   udovletvoriat',   sposobstvovat',   protivodeistvovat'. prepiatstvovat',   protivorechit'.   vvedenie,   protivopolozhno,   pomogat'. udeleno,   antagonismicheskii,   obiazano,   parallel'no,   podobno,   sledano, perpendikuliarnyi, analogichno, ekvivalent, pozvolit'

Kenneth E. Harper

University of California, Los Angeles