[*International Conference on Machine Translation of Languages and Applied Language Analysis*, National Physical Laboratory, Teddington, UK, 5-8 September 1961]

TRANSFORMATION CRITERIA FOR THE CLASSIFICATION OF PREDICATIVE GENITIVE CONSTRUCTIONS IN RUSSIAN

by

DEAN S. WORTH
(Consultant, The RAND Corporation, U.S.A.)

SUMMARY

A major advantage of the transformational approach to syntactic structure is that the investigator is no longer bound to the low-levels of generality inherent in the morphological detail of actual sentences, but can instead regard syntax as a dynamic, addressor-oriented process, the units of which form a hierarchy of functional abstractions. This conception is illustrated by an analysis of Russian constructions containing genitive substantives which must be modified by another unit, although the exact form (adjective or substantive) of this second unit is irrelevant. The transformationally determined varieties of such constructions are cataloged, and it is shown that all such constructions contain either an actual or an implied predication.

ONE of the most vital problems of modern structural linguistics, a problem which will prove to be of prime importance for machine translation as well as for general linguistics, is the formalization of syntactic and semantic studies and the definition of the relation of these two fields each to the other.* The first task to be accomplished in solving this problem is the isolation and definition of the structural units in terms of which the ultimate description is to be made. Recent decades have seen great strides forward in the analysis of Slavic, and especially of Russian phonemic and morphophonemic systems,[1] and attention is now being concentrated on larger units in the hierarchy of linguistic subsystems, the syntagm, sentence and utterance. Soviet scholarly journals have been publishing an increasingly large number of penetrating and provocative studies of syntactic and semantic units, their interrelation in such areas as synchronic derivation, and other problems of general theoretical interest.[2] In the United States, widespread interest has been aroused by the recent studies in transformational syntax of Harris, Chomsky, Lees and others.[3]

* This study was supported by the U.S. Air Force Office of Scientific Research. This paper is their Technical Note TN148.

While opinions differ as to the ultimate value of transformation grammar of the generative type,[4] there is little doubt that transformation theory itself provides a new and powerful tool for the analysis of complex syntactic entitles, and perhaps even a key to the hitherto locked door of structural semantics. Traditional syntax, as exemplified in the works on Russian of Peškovskij, Šaxmatov, and most recently of the Soviet Academy of Sciences,[5] has been based in part on logical and psychological analogues, and has dealt with purportedly syntactic units defined in morphological and semantic terms. The weakness of this procedure becomes evident when one considers the functional (denotative) multiplicity of the morphological sign on the one hand, and the complete absence of objectively verifiable semantic categories on the other. Of course, no syntax can operate in a semantic void, and transformational syntax does not operate without any recourse to meaning. Rather, in its attempt to uncover objective criteria for the classification of syntagmata, it often succeeds in formalizing many of the intuitively recognized relations among words and word groups. The principal advantage of transformation theory, however, is one that has not to my knowledge been stated explicitly: this is its recognition of the dynamic nature of syntax. This statement, and certain of its implications, requires some comment.

Syntax is dynamic, not static. Any speech act results in a string of morphemes arranged in a certain order. Traditional syntax has been concerned largely with the taxonomy of these strings. These strings, however, are only the result of a process, and are not to be confused with the process itself. Syntax is concerned with the process (actually, with the several interrelated and hierarchically arranged processes) which finally results in these strings. Obviously, there is a statable relation between the process and its results, and this relation is the proper province of syntax. However, a catalogue of these results can no more substitute for a description of this process than animal taxonomy can account for the phenomena of life and growth or a museum catalogue for the artistic creativity of those whose works are immured therein. This means that even the most careful listing of, say, the various types of word-combination in Russian can never be of vital interest without a description of the dynamic processes from which these combinations result.

Syntax is addressor-oriented. Corollary to the postulate that syntax is a dynamic process, not a static state, is the fact that syntactic studies must be oriented toward the addressor of a message, not toward the addressee. The addressee of any message must decode a given string of symbols. These symbols are bound to a unidimensional axis (temporal or spatial, as the message is spoken or written), and by no means always contain unambiguous indicators of their own origin. For the addressor, however, these symbols

and their ordering are only the ultimate and actually rather trivial end products of a series of intentional acts. These acts, which are of varying generality and which in most cases must be performed in a particular order (which does not exclude recursivity), should be the principal objects of syntactic investigation. It is instructive to compare the roles of addressor and addressee on the syntactic level with those on the phonemic level. As has been shown by Jakobson and others,[6] the addressee is concerned largely with phonemic distinctions, and the addressor largely with morphophonemic. The speaker generates morphemes, which are materially embodied in phonemes and junctures; the listener distinguishes among phonemes and junctures and recreates morphemes. Similarly, on the syntactic level one might say that the speaker generates syntactic units, which are materially embodied in morphemes or words* and recreates the units of which the generative process consisted - in other words, he finds out what the speaker 'meant.' One might well use the term 'syntagmatics' for the analytic process performed by the addressee, and reserve the term 'syntax' for the generative operation performed by the addressor.

Syntactic units are functional abstractions. It follows from the above that syntactic units are functional abstractions, which are materially embodied in various morphologically and semantically characterized units (words and combinations of words), but which cannot be equated with any one of these units.‡ Again, the analogy with morphemics may prove instructive. The Russian morpheme meaning 'take, carry' appears in such concrete phonemic embodiments as /n'is-/, /n'is'-/, /nos'-/, /n'os/, /nas'-/, and /naš-/ in the forms несу 'I am carrying,' несешь 'you (sg.) are carrying,' носишь 'you (sg.) carry (often),' нес 'he carried,' носил 'he carried (often),' and ношу 'I carry (often).' Similarly, the syntactic unit m (obligatory modification of genitive substantive) to be discussed below appears in such concrete word and word-combination embodiments as A_{gen} (adjective in the genitive case), S_{gen} (substantive in the genitive case) and $A_{+\emptyset}$ (short form comparative adjective) plus A_{gen} in the forms девушка замечательной красоты 'a girl of remarkable beauty,' применение метода зеркала 'application of the mirror method,' and человек росту пониже среднего 'a man of somewhat less than average height.' In other words, an inner syntactic identity can lie behind the external morphological variety, just as an inner syntactic variety can be concealed behind the external identity of the static morphological string (i.e., cases of derivational ambiguity).

---

*The distinction is immaterial on the syntactic level.

‡ These functional abstractions turn out in many cases to toe labels for nodes on the branching diagrams of Immediate Constituent analyses.

To recapitulate: it would appear that considerable progress in analysing the upper levels of linguistic structure could be made by a dynamic, addressor-oriented syntax operating with functional units on various levels of abstraction. It goes without saying that the present brief paper cannot try to define any significant number of such syntactic units or to arrive at any major conclusions about Russian syntactic structure. These general considerations have been advanced here only as a basis for further discussion, and as background for the more specific material to be considered below.

Predicative genitive combinations. One of the most frequently occurring types of binary word combination in Russian is that consisting of substantive (in any case) modified by a second substantive in the genitive (symbolically: S^1 S^2_{gen}), e.g. стакан вина 'a glass of wine,' дом отца 'the father's house.' Such combinations can always be expanded by the addition of an adjective modifier of the genitive substantive (→ S^1 A_{gen} S2gen), e.g. стакан хорошего вина 'a glass of good wine,' дом моего отца 'my father's house.' A generative grammar would produce such combinations in order of increasing complexity, e.g. Он держал стакан 'He was holding a glass'→ Он держал стакан вина 'He was holding a glass of wine' → Он держал стакан хорошего вина 'He was holding a glass of good wine.' The final, adjectivally expanded combination is identical in its external (morphological) form with a second type of S^2_{gen} combination, the inner (syntactic) form of which is, however, radically different, e.g. девушка замечательной красоты 'a girl of remarkable beauty,' вагон второго класса 'a second-class (railroad) car.' This type of combination cannot be considered a two-step expansion from the original S^1 (девушка, вагон), since the intermediate stage S^1 S^2_{gen} is clearly impossible (*девушка красоты, 'a girl of beauty' вагон класса 'a class car') Conversely, if such combinations are considered from the analytic rather than from the generative point of view, they are characterized by the fact that they cannot be reduced* by omission of A_{gen} , e.g. сапог большого размера 'a boot of large size' and Он был буйного нрава 'He was of a wild disposition ('had a w.d.')' cannot → * сапог размера 'a boot of size' and *Он был нрава 'He had a disposition.'

These combinations with obligatory A_{gen} are structurally (i.e. in their inner syntactic form) identical with externally different combinations in which S^2_{gen} is itself modified by a genitive substantive, e.g. человек тридцати лет 'a кап of thirty years,' на расстояниях порядка сотен ангстрем 'over distances of the order of hundreds

*Both expansion (addition of a form) and reduction (omission of a form) are considered special types of transformation.

of angstroms;' cf. the impossibility of reductions → * человек лет 'a man of years' or на расстояниях порядка 'over distances of an order.' The only essential is that S^2_{gen} be modified; whether this modification be materially embodied in an agreeing adjective or a governed substantive is, at least from the point of view of dynamic syntax, irrelevant. We posit as a syntactic unit, therefore, this "obligatory modifier of S^2_{gen}," which will be symbolized by m.* This m is part of a larger syntactic unit M "modifier of S^1", and the group S^1 M can itself of course be part of a still larger structure.

The considerable variety of combinations containing obligatory modifiers of S^2_{gen} is due to the possibility of simultaneous occurrence of various kinds of modification of one or more of the basic or secondary units within the combination. A catalogue of the strings resulting from these modifications must note that the combination may or may not contain a predication; it may or may not mark such categories as tense or mood; the various primary units of which the combination consists may or may not be expanded in various ways; one or more of the simple or expanded units may or may not undergo stylistic reversal of word-order. The resulting multiplicity of externally classified strings should not, however, be allowed to obscure the essential features of this type of combination.

There are two basic features which distinguish the combinations being considered here from all other S^1 A_{gen} S^2_{gen} and S^1 S^2_{gen} S^3_{gen} combinations. The first has been mentioned above, namely the obligatory nature of the modifier of S^2_{gen}. The second basic feature is that all such combinations either contain a predication as they stand, or can be transformed into combinations which do contain predications. It is this feature which justifies the label "predicative genitive" which has been attached to this combination type. All such combinations as (Он. . . показался весьма похожим на)средней величины медведя' (he seemed extremely similar to) a bear of middle size,' Здания дымчатого цвета (не казались красивыми) 'The smoke-colored buildings (did not seem beautiful),' (это были муж, жена, их) мальчик... необыкновенной красоты (It was a husband, his wife, and their) boy ... of unusual beauty' can be transformed to such predicative combinations as медведь был средней величины 'the bear was of middle size,' Здания — думчатого цвета 'The buildings are of smokey color,' and мальчик был необыкновенной красоты 'the boy was of unusual beauty.** Conversely, predicative combinations can be transformed

*Dynamically speaking, it is the act of modification itself, and not the resulting formal modifier, that should be considered the syntactic unit.

**Illustrative examples are taken from the Soviet Academy Grammar, Vol. II, pp. 232f., 434f., 539 and from the nuclear physics and astrobotany texts analysed at The RAND Corporation.

into non-predicatives when it is necessary to include them in larger structures, e.g. a predicative combination such as Частицы — различной первичной энергии 'The particles are of varying primary energy' can by omission of the zero copula be included in a larger structure such as Было подсчитано число ядерно-активных частиц. . . различной первичной энергии 'The number of nuclearactive particles of varying primary energy was counted.' The transformational history of such larger structures always uncovers a second predication embedded within the S^1 A_{gen} S^2_{gen} unit, which therefore resemble such instrumental substantive combinations as Иван вернулся стариком 'John returned an old man.' This original predicative element can be symbolized by β (cf. Russian быть 'to be'), and its disappearance in the larger structures by a transformation to zero, or reduction $\beta \rightarrow \phi$.

Types of combination. We have seen that the combinations under discussion consist of a substantive S^1 in any case, modified by a unit M which itself consists of a substantive in the genitive case S^2_{gen} with an obligatory modifier m, where m can = either A_{gen} or S^3_{gen}. Further, we have posited a predicative element β, potentially if not always actually present in such combinations. These combinations can then be generated by the usual kind of "rewrite rules" (followed by detailed rules for word-order and, ultimately, conversion into phonemes or graphemes):

(1) $S^1 \rightarrow S^1$ M

(2) M $\rightarrow$ m S^2_{gen}

(3a) m $\rightarrow$ A_{gen} in same number and gender as S^2_{gen}

or (3b) m $\rightarrow$ S^3_{gen}

and, optionally at any stage of the above,

(4) S^1 M $\rightarrow$ S^1 β M

All of the attested varieties of these combinations can be generated from the above, by means of one or more of the following transformations:

(5) $\beta \rightarrow \beta_{past}$ он был высокого росту: σ (= symbol).
были олного знака: старуха
редких правил

(6) $\beta \rightarrow \mu\beta$, where μ = *a* modal auxiliary such as казаться 'seem,' являться 'be' etc., including the lexically zero β_{μ} был бы etc.

(6a) $\mu \rightarrow$ оставаться нелинейность . . . остается
порядка σ

(6b) $\mu \rightarrow$ должен смещение долхно быть ... порядка
несколько метров

(6c) $\mu \rightarrow$ оказываться σ оказывается порядка σ

(6d) $\mu \rightarrow$ дело казалось необыкновенной важности
(also contains β_{past})

(6e) $\mu \rightarrow$ мочь ту можешь быть какого угодно мнения
(also contains an expansion of A_{gen})

(6f) $\mu \rightarrow$ хотеть я не хочу быть такого мнения;
(with β_{μ}) я не хотел бы быть
такого мнения

... (6n)

Expansions. Expansions can take place at various stages of transformations (1) to (3b), as expansion transformations are applied to M of S^1 M, or to m of m S^2_{gen}. These expansions are shown in transformations (7) - (9). Their final form depends on which of a series of conjunctive transformations + → i 'and' (coordinating), → a 'but' (contrasting), → ili 'or' (exclusive) is applied after the expansion.

(7) $M \rightarrow M^1 + M^2$ Все это ослепительной свежести и
изящного вкуса; (with β_{past}
руки были... безукоризненной
формы и несказанной доброты;
были мы... крови не родной а
души одной

(8) $m \rightarrow m^1 + m^2$ они голубого или синего цвета
Захар был... мягкого и доброго
сердца

(9a) $A \rightarrow {}_{Q}A$, where Царь... совсем иного нраву
$_{q}$ = adverb

(9b) $A \rightarrow {}_{a}A$, where Все было самого тяжелого и
$_{a}$ = какой, беспокойного свойства
самый такой

(9c) $A \rightarrow A^2_{+\phi}\ A_{gen}$, where $A^2_{+\phi}$ = short form comparative adjective — человек росту пониже среднего

Word-order reversals. Word-order can be reversed at two levels, reversing either the places of S^1 and M, or those of A_{gen} and S^2_{gen} within M, e.g.

(10a) $S^1\ M \rightarrow M\ S^1$ — жесткого шелка кафтан, на первой отечественного производства... машине; такого же порядка поправка

(10b) $A_{gen}\ S^2_{gen} \rightarrow$ / $S^2_{gen}\ A_{gen} \rightarrow$ — Хома Брут был нрава веселого; были мы... крови не родной, а души одной

The above transformation rules appear to exhaust the possible types (except for interrogational, negative, and emphatic varieties, which cannot be explored here). A further set of transformations, operating on the derivational level, is needed to separate out the various lexically conditioned subtypes within the above types. Space considerations prevent us from doing more than suggesting a few possible test transformations for further study:

(A) $S^1\ A_{gen}\ S^2_{gen} \rightarrow S^2_{nom}\ S^1_{gen}\ \beta_\phi\ A_{nom}$, e.g. мужчина невысокого росту → рост мужчины—невысок (Certain modals in the original string will require long form instrumental A in the transform, e.g. дело казалось необыкновенной важности → важность дела казалась необыкновенной

(B) $S^1\ A_{gen}\ S^2_{gen} \rightarrow A_{nom}\ S^2_{nom}\ S^1_{gen}$, e.g. мужчина невысокого росту - невысокий рост мужчины

(C) $S^1\ A^1_{gen}\ S^2_{gen} \rightarrow A^1_{\phi}\ A^2_{<\beta^2}\ S^1$, e.g. девушка чрезвычайной красоты → чрезвычайно красивая девушка

(D) $S^1\ A_{gen}\ S^2_{gen} \rightarrow A^1_{\phi}\ A^2_{<\beta^2}\ S^1$, e.g. продукция высокого качества → высококачественная продукция, человек тридцати лет → тридцатилетний человек

(E) Various prepositional transformations, such as $S^1\ A_{gen}\ S^2_{gen} \rightarrow S^1$ из $A_{gen}\ S^2_{gen}$, e.g. кабинет карельской березы - кабинет из карельской березы..

When these and other types of transformational criteria have been established for the analysis of Russian and English syntax, and when our understanding of these complicated linguistic systems, at present so fragmentary, becomes reasonably complete, it will be possible to develop programs adequate for the fully automatized, high-speed mechanical translation that the future will surely demand. Since machine translation, like any other translation, is basically a procedure for converting one set of code symbols into another such set with a minimum of denotational deviation, any advance in our knowledge of the respective codes will facilitate the conversion procedure, until these codes are understood more thoroughly than is the case at present, the machine translator is likely to be somewhat in the position of an engineer trying to build a bridge between two shores, neither of which has yet been mapped with any accuracy. The foregoing paper is offered as a contribution to the syntactic topography of one small segment of the Russian shore.

REFERENCES

1. Cf. for example such sophisticated analyses as those of AVANESOV R.I., Fonetika russkogo jazyka, Moskva, 1956, HALLE M The Sound Pattern of Russian, 's-Gravenhage, 1959, and JAKOBSON R., "Morfologičeskie nabljudenija nad slavjanskim skloneniem," *American Contributions to the Fourth International Congress of Slavicists,* 's-Gravenhage, 1958, pp. 127-156.

2. E.g., XOLODOVIČ A.A., "Opyt teorii podlassov slov," *Voprosy Jazykoznanija (V.Ja.)*, 1960, **1**, pp. 32-43; VOLOCKAJA Z.M., "Ustanovlenie otnošenija proizvodnosti meždu slovami (opyt primenenija transformacionnogo metoda)," *V.Ja.*, **3**, pp. 100-107; SMELEV D.N., "O 'svjazannyx' sintaksičeskix konstrukcijax v russkom jazyke," *V.Ja.*, 1960 **5**, pp 47-60; PADUČEVA E.V., "Ob opisanii padežnoj sisteme russkogo suščestvitel'nogo (nekotorye problemy omonimii pri mašinnom perevode)," *ibid.*, pp. 105-111.

3. HARRIS Z.S., "Co-occurrence and transformation in linguistic structure," Language 1957 **33**.283-340; CHOMSKY N., *Syntactic Structures,* 's-Gravenhage, 1957; LEES R.B., The Grammar of English Nominalizations *(International Journal of American Linguistics, 1960, **XXVI**, 3, Part II).*

4. The most specific published description of this type of grammar is to be found in the first two chapters of LEES' above-cited book.

5. PEŠKOVSKIJ A.M., *Russkij sintaksis v naučnom osveščenii,* 7th ed., Moskva, 1956; ŠAXMATOV A.A. Sintaksis russkogo jazyka, 2nd ed., Leningrad, 1941; AN SSSR, *Grammatika russkogo jazyka, II*, Sintaksis, 1-2, 2nd ed., Moskva,

1960. A stimulating but not entirely convincing attempt to establish new syntactic categories can be found in LOMTEV T. P., *Osnovy sintaksisa sovremennogo russkogo jazyka*, Moskva, 1958.

6. Cf. HALLE M., *op. cit.* On the several factors present in any speech act, and on the variant forms of communication directed toward these factors, see JAKOBSON R., "Linguistics and Poetics," *Style in Language*, Cambridge, 1960, pp. 350-377.

7. Cf. WORTH D.S., "Transform analysis of Russian instrumental constructions," *Word, 1958,* ***XIV***, 2/3, pp. 269-71.