The semantics of the Russian verb

Kortlandt, F.H.H.

Lingua **27** (1971) 53–81, *North-Holland Publishing Company*

REVIEW ARTICLE — RAPPORT CRITIQUE

THE SEMANTICS OF THE RUSSIAN VERB

Ju. D. APRESJAN, *Èksperimental'noe issledovanie semantiki russkogo glagola*, Izd. 'Nauka', Moskva 1967, 248 pages.

CONTENTS

I

1.1. INTRODUCTION

One of the most remarkable Soviet publications of recent years is undoubtedly 'An experimental investigation of the semantics of the Russian verb' by Ju. D. Apresjan. The book contains an attempt

to arrive at a description of the lexical meaning of the Russian verb on the basis of its syntactic properties. These properties are: (1) the forms assumed by words that are governed by a given verb; (2) the transformations that can be applied to sentences in which a finite form of this verb occurs.[1])

As a consequence of the close interdependence between the syntactic and the semantic level of a language, syntactically identical constructions have often related meanings, and sentences with related meanings can often be constructed according to the same syntactic patterns. This offers a possibility of a rigorous, formalizable description of the semantics of a language, a field which so far has hardly been open to formalization, and it will be possible to reach conclusions as to semantic similarities and differences, which are in principle inaccessible to direct observation, utilizing the syntactic characteristics of the corresponding linguistic expressions, which are, at least to a certain extent, subject to direct fixation. The subject of the investigation may thus be formulated as follows: suppose the (formal) syntax of a language to be known and the set of correct (*pravil'nyj*) sentences in that language to be given, how can the meaning of a linguistic expression be identified as different from the meaning of any other linguistic expression and as related to the meaning of expressions with a similar meaning? In the study by Apresjan this problem is tackled for the Russian verb.

1.2. The basic observations

The basic material used is a set of linguistic expressions containing a verbal predicate. This set results from the analysis of some 25000 phrases, the bulk of which was taken from the illustrative material of the large Russian dictionaries. The metalanguage used for the description is a formal language of syntactic features,[2]) which was created in the course of the analysis. Thus, an attempt was made

---

[1]) In the first part of this article I shall mainly confine myself to an exposition of the theory presented in the book under review; the second part is wholly devoted to criticism.

[2]) It should be noticed that the concept of 'syntactic features' is an analogon of the concept of 'distinctive features' in phonology and has nothing to do with the 'syntactic features' in e.g. N. Chomsky, *Aspects of the theory of syntax*.

to give a systematic description (*sistemnoe opisanie*) of the lexicon, i.e. a description that satisfies the following conditions:
(1) every element of the lexicon is uniquely determined;
(2) the number of describing elements (e.g. the meaning of *n*-ary features) is smaller than the number of described elements;
(3) the information about the syntactic features is sufficient for the (hierarchical) classifications that are needed for the automatic analysis and synthesis of texts. The systematic organization of the lexicon is said to be higher when these requirements are met more completely.

A few more remarks can be made on the correspondence between syntax and semantics. The author supposes that, as a rule, syntactic differences and similarities are an indication of semantic differences and similarities, although for some semantic differences the syntactic data may fail to provide sufficient indications.

The link between syntactic and semantic features is clearest in the case of so-called conditioned meanings, i.e. whenever a given alloseme of a word is realized only in certain lexically or syntactically restricted contexts, e.g. *idti* 'to go' but 'to fall' in *dožd' idët* 'it is raining' or *sneg idët* 'it is snowing' (lexically conditioned) and 'to suit' in *ėto plat'e vam očen' idët* 'this dress suits you well' (syntactically conditioned, since the verb *idti* governs a noun in the dative case when used in this meaning). Such phenomena have been investigated by J. Firth, V. V. Vinogradov, R. S. Ginzburg and others. Apresjan continues this tradition, but denies the distinction between conditioned and unconditioned meaning, because he considers in principle any meaning to be syntactically conditioned. So the principal meaning of *idti* 'to go' is found in constructions containing the prepositions *iz* + 2, *s* + 2, *k* + 3, *po* + 3, *v* + 4, *na* + 4, *za* + 5, or the bare infinitive of another verb, and in a few other constructions, but not in combination with a noun in the dative without preposition. In other instances the allosemes of a word can be established on the basis of the transformability of sentences in which the word occurs: *Otec soderžit sem'ju* 'The father supports the family' is transformable into *Sem'ja soderžitsja otcom* 'The family is supported by the father', not into **Sem'ja soderžitsja v otce*; on the other hand, *Pis'mo soderžit namëk* 'The letter contains a hint' is transformable into *V pis'me soderžitsja namëk* 'id.', not into **Namëk soderžitsja pis'mom.*

These instances show that a given alloseme can be realized only if the context fulfils certain syntactic requirements. Inversely, it happens that a context with certain syntactic features admits only a limited number of semantically related verbs, e.g. *Emu vletelo ot otca* 'He got it hot from his father', where the verb *vletelo* (lit. 'flew in') may be replaced only by the synonyms *dostalos'* and *popalo*. In the context *A on ego* VERB *palkoj po golove* 'And he VERB him with a stick on the head' almost any verb substituted will acquire the meaning 'to strike, hit'. In this case the verb may even be omitted or replaced by an artificial, meaningless verb, which will automatically acquire that very meaning.

1.3. The data of the analysis

The author starts from the conviction that linguistic theory provides reliable criteria for the following parts of the analysis, on which the elaboration of the formal apparatus needed for the description is based. He supposes that the (formal) syntax of the language is a datum, that is to say:

(1) It is possible to carry out a morphological analysis of the text, which consists of segmenting the text into morphs, identifying the morphs as variants of the same morpheme or representatives of different morphemes, dividing the set of morphemes into a subset of lexical morphemes and a subset of grammatical morphemes, and finding a steady consecutiveness of lexical and grammatical morphemes that corresponds to a word form. The set of grammatical morphemes includes conjunctions, flexional morphemes and prepositions, derivational morphemes and 'delexicalized verb roots' of the type *brat'* (*pod nabljudenie*) 'to put (under observation)', *okazyvat'* (*podderžku*) 'to render (support)', *vzjat'* (*pod svoj kontrol'*) 'to take (under one's control)'.

(2) It is possible to distinguish syntactic classes (i.e. classes of equivalent words or word groups, e.g. the class of nouns in the nominative, the class of constructions consisting of the preposition *iz* 'out of' plus the genitive, the class of infinitives, etc.).

(3) It is possible to establish the immediate syntactic links between the word forms in a sentence. [It is not necessary to know the direction of the dependence.] Neither any knowledge of lexical meanings, nor the possibility of determining semantic invariance of phrases is presupposed.

The problem of determining the set of correct sentences is a crucial one in any investigation of syntax and semantics. [This problem has received a good deal of attention in the works of Chomsky, Harris, Hill, Revzin, Šaumjan and others.] In most cases (but not in the works of Harris) the concept of correctness (*pravil'nost'*) regards only the grammatical structure of a sentence, and not its lexical features, so that *Colourness green ideas sleep furiously* is considered a correct sentence. But even N. Chomsky, the most radical adherent of the view that meaningless phrases may be grammatical,[3]) does not consider *Sincerity admires John* a correct phrase (*Syntactic structures*), though it is not clear why sincerity is not allowed to admire John if colourless green ideas are allowed to sleep furiously. It does not seem possible to draw a clear-cut border line which would separate grammatically from lexically incorrect sentences,[4]) as Apresjan shows by means of a list of most interesting but hardly translatable examples. An exact qualification as lexically or grammatically incorrect is possible only in such trivial cases as *The ideas sleep furiously* or *We comes*. Besides, it is often hard to determine whether a given sentence is incorrect at all, as there are some inherent features leading to disturbances in any language. The following examples are given by Apresjan:

A. 1. *Ėto privodit ego v bešenstvo*. 'That drives him wild'
   2. *Ėto privodit ego v gnev*. 'That drives him angry'
   3. *Ėto privodit ego v zlost'*. 'That drives him malicious'

B. 1. *On prixodit v užas ot ėtogo*. 'That strikes him with terror'
   2. *On prixodit v strax ot ėtogo*. 'That strikes him with fear'
   3. *On prixodit v bojazn' ot ėtogo*. 'That strikes him with dread'

It is clear that subtle changes in the meaning of a noun can go hand in hand with a gradual decrease of the correctness of the sentence. It turns out that in these examples the degree of correctness of the sentence corresponds with the degree of intensity of emotion that is rendered by the substantive.

Some other difficulties in determining the set of correct sentences

---

[3]) So on p. 40 of the book under review.

[4]) This view is actually shared by Chomsky, cf. his *Aspects of the theory of syntax*.

(equalization by analogy, elliptical constructions) are discussed by Apresjan, but may be omitted in this survey. After establishing the required syntactic features and the set of correct sentences, the author develops the formal apparatus of the analysis as follows.

1.4. THE FORMAL APPARATUS

The set of **correct phrases** (*pravil'naja fraza*) is defined as the union of the set of correct sentences and the set of those strings of syntactically directly linked word forms of correct sentences that either are correct sentences themselves or do not contain a predicative word form. The example given by Apresjan is: *Odin prostak šel, derža v rukax uzdu svoego osla, kotorogo on vèl za soboju* 'A simpleton went, holding in his hands the bridle of his donkey, which he led behind him'. This is a correct sentence, and therefore a correct phrase. The strings *Odin prostak šel* 'A simpleton went' and *Prostak šèl* 'id.' are correct sentences and thus correct phrases as well. The strings *derža v rukax uzdu svoego osla* 'holding in his hands the bridle of his donkey', *derža v rukax* 'holding in his hands', *derža uzdu svoego osla* 'holding the bridle of his donkey' and *uzdu svoego osla* 'the bridle of his donkey' are correct phrases without being correct sentences. The string *On vel za soboju* 'He led behind him' is not a correct phrase, because it is not a correct sentence and does contain a predicative word form. The string *prostak, derža v rukax uzdu svoego osla* 'a simpleton, holding in his hands the bridle of his donkey' is not a correct phrase because the word form *prostak* 'simpleton' is not directly linked to any other word form in the string. [*derža* 'holding' is a gerund, in Russian formally distinct from the participle 'holding'.]

To any word form of a phrase corresponds a class of syntactically equivalent word forms. A string of symbols representing the classes of word forms that are syntactically equivalent to the word forms of a given phrase is called the **construction** (*konstrukcija*) of that phrase, e.g. the string *My idem na zanjatija* 'We are going to our work' may be represented by $N^1_n\ V\ na\ N^2_a$ (first noun in the nominative, verb, preposition *na*, second noun in the accusative). A construction is called correct if there is at least one correct phrase to which it corresponds.

A **position** (*pozicija*) is defined as an ordered pair of syntactically

linked classes of equivalent word forms in a given construction.[5]) It is said to characterize its first member. As the direction of the dependence is not supposed to be known, to every pair of syntactically linked word forms in a phrase there correspond two positions, e.g. to the construction N_n V of the sentence *Poezd prišël* 'The train arrived' correspond the positions (N_n, V) and (V, N_n), characterizing the substantive *poezd* and the verb *prišël* respectively. In the sentences *On nadel šapku na mexu* 'He put a fur-lined cap on' and *On nadel šapku na ulice* 'He put a cap on in the street' the noun in the prepositional case is characterized by the positions (N_p, N_a) and (N_p, V) respectively.

A position is called **obligatory** (*objazatel'nyj*) if the removal of the word form corresponding to its first member violates the correctness of the phrase under consideration. It is called **optional** (*fakul'tativnyj*) if that word form can be omitted. In the sentences *Pëtr perestal rabotat'* 'Peter stopped working' and *Stol naxodilsja na ulice* 'The table was in the street' the word forms *rabotat'* 'working' and *na ulice* 'in the street' are said to be in obligatory position since **Pëtr perestal* and **Stol naxodilsja* are considered incorrect phrases. On the other hand, the position of *pri povorote* 'at turning' in *Dver' skripit pri povorote* 'The door creaks when it turns' is optional. [Some of the examples given by Apresjan are not clear to me. The position of *širokie* 'broad' in *U nego byli širokie pleči* 'He had broad shoulders' is considered obligatory, while *vtoroj* 'second' in *Emu minul vtoroj god* 'He is two years old' is said to be in optional position. In my opinion, these two cases are identical as far as the position of the adjective is concerned. It may be recalled that no way to determine the semantic invariance of phrases is presupposed at this stage of the analysis.]

A correct phrase or construction is called **oriented** (*orientirovannyj*) towards the word form (or syntactic class) X_o if any position either contains X_o or is obligatory. In that case X_o is called the central member of the phrase (or construction). Thus, the sentences *On perestal rabotat'* and *Dver' skripit pri povorote* are oriented phrases with a verb as the central member. Phrases of the type

---

[5]) The term *pozicija* is used by Apresjan not only in this sense, but also as a designation of the first member of such a pair. This is rather confusing.

On sidel v komnate otca 'He was sitting in the room of his father' are not oriented, for the word *otca* 'of his father' is neither obligatory nor syntactically linked to the verb. [One may wonder on the basis of what criterion Apresjan assumes that the verb is immediately linked to the subject and the adjunct of place and that no direct link exists between the latter two elements, that is, why *On sidit v komnate* 'He is sitting in the room' is considered an expansion of *On sidit* 'He is sitting' and not of *On v komnate* 'He is in the room'. This problem is not tackled by Apresjan because the immediate syntactic links are supposed to be known.] It should be noticed that the central member of an oriented phrase is not uniquely determined as is clear from the example *Poezd prišël* 'The train arrived'.

A construction is called a **kernel construction** (*jadernaja konstrukcija*) if it corresponds to a class of oriented phrases that satisfies the following requirements:
(1) one and only one central member of the phrases of the class is a predicative word form (e.g. a verb in verbal kernel constructions);
(2) in the phrases of the class all positions are obligatory;
(3) every phrase in the class is simpler than any of its transforms.
[The concepts of transformability and simplicity of a transform are defined somewhat later without use of the concept of kernel construction.] The kernel constructions are found by Apresjan empirically, because possible deductive procedures generate a large number of constructions that do not have an interpretation in Russian. The set of kernel constructions obtained corresponds to a complete and disjunct division of the set of verbal phrases into syntactic classes. Thus, every kernel construction can be regarded as a differential syntactic feature of the corresponding class of verbal phrases.

1.5. Compatibility

If the verb V_o is the central member of the oriented phrases p_i and p_j, and if M_i and M_j are the sets of word forms linked to V_o in p_i and p_j respectively, then p_i and p_j are called **compatible** (*sovmestimyj*) with respect to V_o if there is an oriented phrase p_k with the verb V_o as a central member such that the set M_k of word forms linked to V_o in p_k contains the union of the sets M_i and M_j, e.g. *On žaluetsja mne* 'He complains to me' and *On žaluetsja na syna* 'He complains of his son' are compatible with respect to the verb,

because *On žaluetsja mne na syna* 'He complains to me of his son' is a correct phrase. The phrases *On iskal knigu na stole* 'He looked for the book on the table', *On iskal knigu pod stolom* 'He looked for the book under the table' and *On iskal knigu v škafu* 'He looked for the book in the bookcase' are compatible because *On iskal knigu na stole, pod stolom i v škafu* 'He looked for the book on the table, under the table and in the bookcase' is a correct phrase. The phrases *On valitsja so stula* 'He falls from the chair', *On valitsja navznič'* 'He falls flat on his back', *On valitsja ot ustalosti* 'He falls from weariness', *On valitsja v grjaz'* 'He falls into the mud' and *On valitsja na zemlju* 'He falls on the ground' are compatible, because *Ot ustalosti on valitsja so stula navznič' v grjaz' na zemlju* 'From weariness he falls from the chair flat on his back on the ground into the mud' is a correct, though somewhat cumbersome phrase. From this example it might be concluded that the relation of compatibility is a transitive one, that is: if p_i and p_j are compatible and if p_j and p_k are compatible, then p_i and p_k are compatible. The author verified this assumption on each pair of compatible phrases in the material investigated by him, and it is therefore accepted as valid. Since the relation of compatibility is reflexive and symmetric as well (that is: every phrase is compatible with itself, and if p_i and p_j are compatible then p_j and p_i are compatible), it is an equivalence and consequently divides the set of oriented verbal phrases into a number of disjunct classes.

The concept of compatibility will be used as a formal analogon of the concept of semantic invariance. Therefore, it may be useful to give some examples of incompatible phrases: *On boitsja sobaki* 'He is afraid of the dog' and *On boitsja za syna* 'He is afraid for his son' [one may wonder if these phrases are really incompatible], *On vspomnil o svidanii* 'He remembered the appointment' and *On vspomnil pro razgovor* 'He remembered the conversation' (these constructions are not compatible but can be transformed into each other), *On izmenil rodine* 'He betrayed his country' and *On izmenil svoi ubeždenija* 'He changed his convictions', *On vladeet pračečnoj* 'He owns a laundry' and *On vladeet ostrym perom* 'He wields a sharp pen'.

If the phrases p_i and p_j are compatible and if every word form of p_i coincides with some word form of p_j, then p_j is called an **expansion** (*razvĕrtyvanie*) of p_i.

1.6. QUASI-TRANSFORMABILITY

Two oriented phrases p_i and p_j are called **quasi-transformable** (*kvazitransformiruemyj*) if the following conditions are fulfilled: (1) They contain the same lexical morphemes. The so-called delexicalized morphemes are considered to be grammatical by Apresjan, e.g. *Ja objazujus' rabotat'* 'I commit myself to work' versus *Ja beru objazatel'stvo rabotat'* 'I undertake the commitment to work'. (2) To every position in the construction of p_i corresponds a position in the construction of p_j such that the lexical morphemes found in those positions are identical, and vice versa. It is clear that the quasi-transformability so defined is an equivalence and consequently divides the set of oriented verbal phrases into a number of disjunct classes. [This definition is rather similar to the definition of transformability used by Z. S. Harris, who was the first to define the concept as an equivalence and to point at the identity of the lexical morphemes and their syntactic links in transformable phrases.] According to this definition, the phrases *Kritik organizuet gruppu* 'The critic organizes a group' and *Gruppa organizovana kritikom* 'The group has been organized by a critic' are quasi-transformable, but the relation is not valid for *Kritik organizuet gruppu* and *Organizacija kritikuet gruppu* 'The organization criticizes the group'.

The quasi-transformability of two phrases is not an ideal indicator of semantic invariance for a number of reasons. Firstly, the same denotative meaning may be expressed by different lexical morphemes, e.g. *Moskva ležit na vostok ot Pariža* 'Moscow lies east of Paris' versus *Pariž ležit zapadnee Moskvy* 'Paris lies west of Moscow', *On prodal mašinu prijatelju* 'He sold the car to a friend' versus *Prijatel' kupil mašinu u nego* 'A friend bought the car from him', *On prepodaët mne matematiku* 'He teaches me mathematics' versus *Ja učus' matematike u nego* 'I learn mathematics from him', etc. Secondly, some semantic links may not be expressed by the syntactic structure of the phrase under consideration, e.g. *On izmenilsja v lice* 'He changed his countenance' versus *Ego lico izmenilos'* 'His face changed', *On podražaet artistu v poxodke* 'He imitates the artist in his gait' versus *On podražaet poxodke artista* 'He imitates the gait of the artist', *On sostavil biblioteku iz redkix knig* 'He compiled a library of rare books' versus *Redkie knigi sostavljajut ego biblioteku*

'Rare books constitute his library', etc. These phrases, though closely related semantically, are not quasi-transformable.

Thirdly, in some instances phrases that correspond to quite different real situations are quasi-transformable, e.g. *Sestra gorditsja bratom* 'The sister is proud of her brother' versus *Brat gorditsja sestroj* 'The brother is proud of his sister', *On vladeet čuvstvami* 'He has control over his feelings' versus *Čuvstva vladejut im* 'He is reigned by his feelings', *Brat prixodit ko mne v gosti* 'My brother comes to me on a visit' versus *Gost' prixoditsja mne bratom* 'The visitor is a brother of mine', *Starik vernul Ivana* 'The old man made Ivan come back' versus *Ivan vernulsja starikom* 'Ivan came back as an old man', etc. In view of this difficulty, the quasi-transformability of two phrases is considered a necessary but insufficient condition for their transformability.

1.7. Transformability

The oriented phrases p_i and p_j are called **transformable** (*transformiruemyj*) into each other if they are quasi-transformable and if for any expansion p_i' of p_i there is an expansion p_j of p_j such that p_i' and p_j are quasi-transformable but not identical, and vice versa. For practical purposes a weaker criterion was used: only the most characteristic expansions of a given phrase were taken into consideration. The pairs of phrases mentioned in the preceding paragraph are not transformable into each other, e.g. *Sestra gorditsja bratom i ego podvigami* 'The sister is proud of her brother and his exploits', but not **Brat i ego podvigi gordjatsja sestroj* 'The brother and his exploits are proud of his sister'; *On xorošo vladeet čuvstvami* 'He has good control over his feelings', but not **Čuvstva xorošo vladejut im* 'He is reigned well by his feelings'; *Brat prixodit ko mne v gosti po voskresen'jam* 'My brother comes to me on a visit every Sunday', but not **Gost' prixoditsja mne bratom po voskresen'jam* 'The visitor is a brother of mine every Sunday'.

The relation of transformability is an equivalence and therefore divides the set of phrases under consideration into a number of disjunct classes. Within these transform classes a uniquely determined simplest transform is defined. This simplest transform is found empirically according to the following rules:

(1) x is simpler than y if x is a sentence and y is not a sentence, e.g.

Mal'čik čitaet knigu 'The boy reads the book' is simpler than *čtenie knigi mal'čikom* 'the reading of the book by the boy'.
(2) If x and y are sentences, x is simpler than y if it contains fewer grammatical morphemes, e.g. *Rabočie strojat dom* 'The workers build the house' is simpler than *Dom stroitsja rabočimi* 'The house is built by the workers'.
(3) If x and y contain the same number of grammatical morphemes, x is simpler than y if x contains a word form of the class N_n (noun in the nominative case) and y does not contain such a word form, e.g. *Grad pobil rož'* 'The hail beat the rye down' is simpler than *Rož' pobilo gradom* 'id.' (impersonal construction).
(4) If both x and y contain a word form of the class N_n, x is simpler than y if x contains a verb without a prefix (and its construction may contain a preposition) and y contains a verb with a prefix (and its construction may not contain a preposition), e.g. *Ona stelet skatert' na stol* 'She lays the cloth on the table' is simpler than *Ona zastilaet stol skatert'ju* 'id.'. Some linguistically less interesting rules were added in order to reach a unique solution. If all optional positions are eliminated, the construction corresponding to a simplest transform is a kernel construction. A list of the 84 kernel constructions found by Apresjan is contained in his book. The constructions that correspond to the phrases belonging to the transform class of a given phrase are called the **transformations** (*transformacija*) of the construction corresponding to that phrase. The transformations of a kernel construction are regarded as differential syntactic features of the corresponding transform class.

One more remark should be made in connection with the concept of simplicity of a transform as it is defined by Apresjan. The simplest transform is determined on the basis of a number of formal characteristics, independently of semantic considerations. Consequently, phrases that are regarded intuitively as simpler because of their relatively high frequency of occurrence may turn out to be more complex according to the given criteria, e.g. *Kniga interesuet ego* 'The book interests him' versus *On interesuetsja knigoj* 'He is interested in the book', *P'esa vostorgaet ego* 'The play delights him' versus *On vostorgaetsja p'esoj* 'He is delighted with the play', *On kljanët izmennika* 'He curses the traitor' versus *On proklinaet izmennika* 'id.', etc. In view of this, the simplicity of a transform

should be regarded only as a working concept within the formal apparatus of the analysis.

[The semantic invariance of a phrase and its transforms is not to be understood as a complete identity of meaning. On the contrary, it is reasonable to suppose that any change in the form of a linguistic expression is accompanied by some change of meaning. However, in a large number of cases it is hard to find a regular relationship between the meanings of linguistic expressions that are formally distinguished in the same way. On the other hand, if two verbal phrases with the same construction can be transformed identically, they generally have some semantic feature in common, and that is precisely why the concept of transformability is a useful one in identifying the lexical meaning of a given verb.]

1.8. Ideal phrases

An **ideal phrase** (*ideal'naja fraza*) is defined as a class of oriented phrases with a given verb V_o as a central member and identical syntactic features (i.e. kernel construction, compatible constructions and transformations), e.g. the phrases of the type *Mat' gotovit obed* 'Mother prepares dinner', *My gotovim užin* 'We prepare supper', *On gotovit zavtrak* 'He prepares breakfast', etc. The central member of these phrases is the verb *gotovit'* 'to prepare', the kernel construction is $N_n^1\ V\ N_a^2$, compatible with the construction $N_n^1\ V\ N_a^2\ iz\ N_g^3$ (cf. *Mat' gotovit obed iz diči* 'Mother prepares a dinner of game') which is transformable into $N_n^1\ V\ N_a^3\ na\ N_a^2$ (*Mat' gotovit dič' na obed* 'Mother prepares game for dinner'), and transformable itself into the constructions $N_n^2\ V\ sja\ N_i^1$ (*Obed gotovitsja mater'ju* 'Dinner is prepared by mother'), $N_n^1\ iz\ V\ N_a^2$, $N_n^1\ pri\ V\ N_a^2$ and $N_n^1\ s\ V\ N_a^2$.

The developed formal apparatus was used for the description of the phrases that contain any of about 1500 most frequent verbs as a central member. A dictionary describing verbal lexemes in terms of their syntactic properties was compiled, in which every item can be viewed as a mapping of a tree with three levels, corresponding to the three types of differential syntactic features. Every terminal element of a tree corresponds to an ideal phrase. In this way, about 4500 ideal phrases have been detected by Apresjan.

1.9. DERIVATION; CONVERSION

The words x_i and x_j are said to be in the relation of **morphological derivation** (*morfologičeskaja proizvodnost'*) if they contain the same lexical morpheme, e.g. *vosk* 'wax' – *voščit'* 'to wax', *rešat'* 'to decide' – *rešitel'nyj* 'decisive', *nenavidet'* 'to hate' – *nenavist'* 'hatred'. A word with an affix is considered a morphological derivative of a word without an affix.

The words x_i and x_j, central members of the phrases p_i and p_j respectively, are said to be in the relation of **syntactic derivation** (*sintaksičeskaja proizvodnost'*) if they are in the relation of morphological derivation and if the phrases p_i and p_j are transformable into each other, e.g. (*Kniga*) *interesuet* (*ego*) '(The book) interests (him)' – (*U nego*) *interes* (*k knige*) '(He has) interest (in the book)'. [The distinction between lexical and syntactical derivation was first introduced by J. Kuryłowicz. A transformational criterion of derivation was elaborated for the first time by Z. M. Volockaja.] The word x_i is called a syntactic derivative of the word x_j if the construction of the phrase p_j is a kernel construction. It follows from the definitions that the direction of the morphological derivation may not coincide with the direction of the syntactic derivation, e.g. the adjective *interesnyj* and the substantive *interes* are syntactic derivatives of the verb *interesovat'*, and the substantives *rabota* 'work' and *igra* 'game' are syntactic derivatives of the verbs *rabotat'* 'to work' and *igrat'* 'to play' respectively, while the direction of the morphological derivation is opposite. The symbols $N(V)$ and $A(V)$ are used to represent the syntactic classes of nouns and adjectives that are syntactic derivates of a verb.

Two kinds of transformations are distinguished by Apresjan, called **conversional** (*konversnyj*) and **synonymic** (*sinonimičeskij*). The first category contains any transformation linking two constructions that differ in at least two nonderivative noun forms (that is: at least two nonderivative noun forms are found in different cases or with different prepositions), e.g. the passive transformation $N_n^1\ V\ N_a^2 \leftrightarrow N_n^2\ V\ sja\ N_i^1$ (*Rabočie strojat dom* 'The workers build the house' versus *Dom stroitsja rabočimi* 'The house is built by the workers') and the transformations $N_n^1\ V\ N_a^2 \leftrightarrow N_n^2\ V\ sja\ N_g^1$ (*Otvetstvennost' strašit ego* 'Responsibility frightens him' versus *On strašitsja otvetstvennosti* 'He is afraid of responsibility'), $N_n^1\ V\ N_a^2 \leftrightarrow N_n^2\ V\ sja\ pod\ N_i^1$ (*Sneg provalil kryšu* 'The snow made the roof col-

lapse' versus *Kryša provalilas' pod snegom* 'The roof collapsed under the snow'), $N^1_n\ V\ N^2_d \leftrightarrow N^2_n\ V\ N^1_d$ (*Ėto protivorečit faktam* 'This contradicts the facts' versus *Fakty protivorečat ėtomu* 'The facts contradict this'), $N^1_n\ V\ N^2_a \leftrightarrow N^2_n\ A(V)\ N^1_d$ (*On slyšit golosa* 'He hears the voices' versus *Golosa slyšny emu* 'He can hear the voices'), $N^1_n\ V\ N^2_a\ na\ N^3_a \leftrightarrow N^1_n\ za\ V\ N^3_a\ N^2_i$ (*Ona stelet skatert' na stol* 'She lays the cloth on the table' versus *Ona zastilaet stol skatert'ju* 'id.'). A transformation linking two constructions that differ in no more than one nonderivative noun form belongs to the second category, e.g. $N^1_n\ V\ N^2_a \leftrightarrow N^1_n\ raz\ V\ N^2_a$ (*On menjaet rubl'* 'He changes a rouble' versus *On razmenivaet rubl'* 'He exchanges a rouble'), $N^1_n\ V\ po\ N^2_d \leftrightarrow N^1_n\ V\ o\ N^2_p$ (*On grustit po otcu* 'He mourns for his father' versus *On grustit ob otce* 'He mourns over his father'), $N^1_n\ V\ N^2_a \leftrightarrow N^1_n\ A(V)\ N^2_d$ (*On blagodarit sud'bu* 'He thanks his lucky stars' versus *On blagodaren sud'be* 'He is thankful to his lucky stars'), $N^1_n\ V\ N^2_d \leftrightarrow N^1_n\ v\ N(V)_p\ s\ N^2_i$ (*Ėto protivorečit faktam* 'This contradicts the facts' versus *Ėto v protivorečii s faktami* 'This is in contradiction with the facts'), $N_n\ V\ D(A) \leftrightarrow u\ N_g\ A_n\ N(V)_n$ (*On myslit logičeski* 'He thinks logically' versus *U nego logičeskoe myšlenie* 'He has a logical way of thinking').

1.10. The efficiency of the description

Some different ways of defining the system of differential syntactic features in order to attain a more economic use of the available data are discussed by Apresjan. One point should be mentioned here. It is clear that not all syntactic features have the same discriminative power for the semantic identification of verbal lexemes. Therefore, all features belonging to one of the following classes were eliminated from the description:

(1) The set of wholly redundant features. [A feature is called wholly redundant if the set of described items does not contain a single pair of items that are distinguished only by that very feature.] Most nominalizing transformations belong to this category, since these transformations duplicate as a rule the discriminating function of the corresponding kernel constructions or passive transformations, e.g. $N^1_n\ V\ N^2_a \leftrightarrow A(V)^1_n\ N^2_a\ N^1_n \leftrightarrow A(V)^2_n\ N^1_i\ N^2_n \leftrightarrow N(V)_n\ N^2_g\ N^1_i$ (*Rabočie strojat dom* 'The workers build the house' – *strojaščie dom rabočie* 'the workers building the house' – *postroennyj rabočimi dom* 'the house built by the workers' – *stroitel'stvo doma rabočimi* 'the

building of the house by the workers'), $N_n\ V\ D(A) \leftrightarrow A_n\ N(V)_n\ N_g \leftrightarrow N(A)_n\ N(V)_g\ N_g$ (*Pëtr postupaet pravil'no* 'Peter acts correctly' – *pravil'nyj postupok Petra* 'the correct act of Peter' – *pravil'nost' postupka Petra* 'the correctness of the act of Peter'), etc. This rule leads to the elimination of some 300 transformations.

(2) The set of unique features, i.e. constructions and transformations that are characteristic of one and only one ideal phrase. To this category belong nine kernel constructions and about 400 transformations, e.g. $N_n^1\ V$ *na* N_a^2 *ot* N_g^3 (*Gorod otstoit na pjat' kilometrov ot derevni* 'The town is five kilometres away from the village'), $N_n^1\ V$ *v* $N_a^2 \leftrightarrow N_n^2\ V$ *sja* N_i^1 (*Idei pronikajut v massy* 'The ideas penetrate into the masses' versus *Massy pronikajutsja idejami* 'The masses are imbued with the ideas').

(3) The set of features that are redundant within a given class of ideal phrases. These features can be eliminated wherever their presence or absence is automatically implied by the presence or absence of other syntactic features, but remain distinctive in other cases. It results from the analysis that distributional syntactic features (kernel and compatible constructions) have a larger discriminative power for the identification of verbs designating motion, transfer, removal, physical influence upon an object and other dynamic situations, while transformational syntactic features are more characteristic of verbs designating intellectual or emotional activity.

1.11. Productivity

The author proposes also a method of measuring the productivity of a semantic pattern. The pattern of the phrase *On b'ët lošad'* 'He beats the horse' has a high productivity, because there exists a large number of different phrases with the same general meaning of the verb and an identical syntactic structure. The corresponding ideal phrase is characterized by seven nonzero differential syntactic features: the kernel construction $N_n^1\ V\ N_a^2$, the compatible constructions $N_n^1\ V\ N_a^2\ N_i^3$ and $N_n^1\ V\ N_a^2$ *po* N_d^3 and four transformations. On the other hand, the pattern of the phrase *On b'ët baklúši* 'He twiddles his thumbs' (lit. 'He beats chips') has an extremely low productivity, since no other phrase with approximately the same meaning of the verb and the same syntactic structure exists. It is even impossible to separate the meaning of the verb from the

meaning of the object. Syntactically, the corresponding ideal phrase is characterized by the absence of compatible constructions and transformations, and therefore has only one nonzero differential syntactic feature. The **productivity** of the semantic pattern of the phrase ϕ_i with a given verb V_o as a central member is defined as the ratio of the number of its nonzero syntactic features and the largest number of nonzero syntactic features characterizing any ideal phrase with the verb V_o as a central member. It can be shown that the productivity of the semantic patterns of the phrases *On b'et lošad'* and *On b'et bakluši* is $\frac{7}{7}$ and $\frac{1}{7}$ respectively. The productivity of the semantic patterns of the phrases *On boltaet lekarstvo* 'He stirs the medicine', *On boltaet nogami* 'He dangles his legs', *Samolet boltaet* 'The airplane dangles' and *On boltaet jazykom* 'He wags his tongue' is $\frac{4}{4}$, $\frac{3}{4}$, $\frac{2}{4}$ and $\frac{1}{4}$ respectively. If the productivity of a semantic pattern is 0.5 or more, the pattern is called productive. It results from the investigation that 73% of the ideal phrases studied by Apresjan correspond to a productive pattern. [The productivity of half of the productive patterns equals unity.] Verbs designating motion or physical action show more often a tendency towards phraseologization (relatively more improductive semantic patterns) than verbs designating emotional or intellectual activity.

Thereupon, a hierarchical classification of the set of ideal phrases with a productive semantic pattern is defined by the author. It contains three levels, corresponding to the three main types of syntactic features (kernel constructions, compatible constructions, transformations) and each of them divided into a number of sublevels. Thus, its structure is identical with the structure of every single item in the dictionary.

1.12. Semantic space

In the last chapter of his book Apresjan defines a metric space on the set of meanings of ideal phrases. The **semantic power** (*semantičeskaja soderžatel'nost'*) of a syntactic feature is defined as the reciprocal of the number of ideal phrases that are characterized by the presence of the feature. It is clear that the semantic power of the transformation $N_n^1\ V\ N_a^2 \leftrightarrow N_n^2\ V\ sja\ N_i^1$ (the passive transformation, characteristic of a large number of phrases containing a transitive verb) is considerably smaller than the semantic power of the construction $N_n^1\ V\ o\ N_a^2$, characteristic of a number of phrases

designating contact or collision, e.g. *spotykat'sja o porog* 'to stumble over the threshold'. The **semantic homogeneity** (*semantičeskaja odnorodnost'*) of a construction is defined as the ratio of the number of its occurrences as a kernel construction and the total number of its occurrences as a syntactic feature (either as a kernel construction or as a compatible construction) in a given set of ideal phrases. Since transformational features are in principle semantically more homogeneous than distributional features, the semantic homogeneity of a transformation is defined as the sum of the values attached to the semantic homogeneity of the constructions linked by the transformation. Finally, the most homogeneous transformation is taken as a unit of measurement and the homogeneity of the other transformational and distributional features is evaluated relatively to this standard.

The **weight** (*ves*) of a syntactic feature is defined as the average of its semantic power and its semantic homogeneity. The **distance** (*rasstojanie*) between the meanings of the ideal phrases p_i and p_j is defined as

$$r(i, j) = 1 - \frac{S_1(i, j)}{S_2(i, j)},$$

where $S_1(i, j)$ is the sum of weights of the nonzero syntactic features that p_i and p_j have in common and $S_2(i, j)$ is the sum of weights of the nonzero syntactic features characteristic of either p_i or p_j (or both). It follows from this definition that the distance between the meanings is equal to unity if the ideal phrases do not have any nonzero syntactic feature in common and the distance is zero if the syntactic features of the ideal phrases coincide altogether.

It is shown experimentally that the distribution of the meanings of ideal phrases over the semantic space thus defined is rather unequal: some parts of the space contain large groups of semantically related verbs. [On the basis of similar observations A. Ja. Šajkevič used methods of correlation analysis in order to define semantic fields.] In view of this phenomenon, Apresjan introduces the concept of semantic class, i.e. a subset K of the set of elements contained in the semantic space such that if a and b are elements of K and c is not an element of K, then $r(a, b)$ is smaller than $r(a, c)$ for any triplet a, b, c in the semantic space. The following theorems are proved:

(1) If the intersection of two semantic classes is not empty, one of them is contained in the other.
(2) There is one and only one division of the set of elements contained in the semantic space into classes such that the number of classes obtained is minimal but not smaller than two.

An algorithmic solution for the detection of classes in the semantic space is proposed by Apresjan. It is proved that this algorithm yields the unique division as postulated by the second theorem.

II. Evaluation

2.1. Critical remarks

It is beyond doubt that the investigation of Apresjan and the formal apparatus developed in his book for the description of the semantics of the Russian verb by means of a metalanguage of syntactic features are of primary importance for the creation of a formalized theory of semantics. And precisely for that reason the weaker points in the analysis should be detected in order to be eliminated. Some of them are mentioned by Apresjan himself.

Firstly, some syntactic features are characteristic of a large class of verbs that can hardly be regarded as having some semantic feature in common, e.g. the verbs governing the dative case or the verbs that govern the prepositional case preceded by the preposition *v*. But the class of verbs that govern one noun in the dative and another in the prepositional preceded by *v* is much smaller and semantically much more homogeneous. If the transformation $N_n^1\ V\ N_d^2\ v\ N_p^3 \leftrightarrow N_n^1\ V\ pered\ N_i^2\ v\ N_p^3$ (e.g. *On vinitsja mne vo vsëm* 'He confesses everything to me' versus *On vinitsja peredo mnoj vo vsëm* 'id.') is added to the list of syntactic features, a small and semantically fairly homogeneous class of verbs is obtained. It seems that this first objection is well refuted by the introduction of the concepts of semantic power and semantic homogeneity.

Secondly, different syntactic features may be characteristic of verbs that are semantically closely related, e.g. *imet' dom* and *vladet' domom* 'to own a house', *pomogat' drugu* and *podderživat' druga* 'to support a friend', *gnušat'sja pošlostej* and *brezgat' pošlostjami* 'to have an aversion of banalities'. This kind of irregularities is much more difficult to deal with. However, it turns out that the number of such cases is relatively small, since there is a tendency

towards unification of the syntactic features characteristic of semantically closely related verbs.

Thirdly, the meaning of words belonging to other parts of speech cannot be identified to the same degree of precision as the meaning of verbs, because they are, as a rule, not characterized by syntactic features of equal discriminative power. Thus, it may be expected that the identification of the meaning of a substantive on the basis of its syntactic properties will be less complete than the semantic identification of a verb. Nevertheless, it turns out that in a number of cases the method leads to interesting results in this field as well. Only substantives like *vid* 'variety', *kategorija* 'category', *klass* 'class', *razrjad* 'sort', *rod* 'kind', *tip* 'type' are characterized by the transformation *Ėtot rod ob'ektov* 'This kind of objects' – *Ob'ekty ėtogo roda* 'Objects of this kind'. Only substantives like *vrag* 'enemy', *drug* 'friend', *mat'* 'mother', *otec* 'father', *prijatel'* 'friend', *sosed* 'neighbour' are characterized by the transformation *On moj drug* 'He is my friend' – *On mne drug* 'He is a friend of mine'.

[It can be added that for exactly the same reason the method proposed by Apresjan may not yield equally satisfying results when applied to a language of a very different structure. Russian is a highly inflected language with an explicit formal expression of most syntactic links and a relatively free word order. If the method is applied to a language characterized by a relatively small number of auxiliary elements (affixes and particles) and a rather strict word order, it is not probable that the meaning of the lexical morphemes can be identified to the same degree of precision, because the discriminating power of any syntactic feature is much more limited than in languages with a large redundance in the use of their syntactic means. An illustrative example is Vietnamese, in which the phrase *Ngu·ò·i làm ruộng tôt* (lit. 'man do land good') may mean 'The man cultivates the good land', 'The man makes the land good', 'The man cultivating the land is good', 'the good man cultivating the land', or 'Working man, the land is good'. The semantic variant effective in a given situation is determined by the context or by nonlinguistic factors.]

The following objection is fundamental. The description of semantics by means of a metalanguage of syntactic features will establish the identity or similarity of meaning of linguistic expressions within certain limits, but it does not indicate:

(a) which real situations may correspond to a given linguistic expression;
(b) which difference in real situations corresponds to a given difference between two sets of syntactic features;
(c) which linguistic expressions may be used as a mapping of a given situation in reality. It seems that these problems cannot be solved as long as no use is made of data that are essentially semantic in character.

2.2. The formal syntax

If we confine ourselves to the problem posed by Apresjan, the weakest point in the analysis seems to be the starting-point. The formal syntax of the language is supposed to be given, but it is doubtful whether it can be established without the implicit use of semantic data, as is shown by the following examples:
(1) It is not always possible to segment the text into morphs on purely formal grounds, e.g. *vstupit'* 'to enter' will be identified as *v-stupit'* 'to step + in' and not as *vs-tupit'* 'to blunt + up' on semantic grounds only. Likewise, *vleč'* 'to draw, attract' is neither historically nor synchronically to be identified as *v-leč'* 'to lie down + in'.
(2) It is not always possible to identify the morphs (after they have been discovered) as variants of the same morpheme or representatives of different morphemes, e.g. the prefix *v* in *v-stavat'* 'to get up' may be interpreted as either a variant of the prefix *v* 'in' or a variant of the prefix *vz* 'up'. [It might be argued that the first identification is correct because the prefix *vz* is realized as *vos* in *vos-stavat'* 'to revolt'. But this argument is not valid because the prefixes *vz* and *voz*, though historically identical, are distinguished in modern Russian, as is shown by the phrases *Ogurcy vsxodjat* (not: *vosxodjat*) 'The cucumbers sprout' versus *Mnogie obyčai vosxodjat k drevnosti* (not: *vsxodjat*) 'Many customs go back to ancient times'. The difference is stylistic in *Solnce vsxodit* versus *Solnce vosxodit* 'The sun rises'.] In other cases the difficulty is even more obvious: I am unable to see how the phrase *On šël* 'He went' can be identified as the past tense of *On idët* 'He goes' if no semantic data are taken into account. [And even if the use of semantic data is allowed the problem is not always easy to solve, as can be shown by the English example *see* : *seen* : *scene*.]

(3) It is not easy to draw a clear-cut border line between the set of lexemes and the set of grammatical morphemes as distinguished by Apresjan. The major difficulty is the grammatical character of the so-called delexicalized verb roots, such as *okazyvat'* (*podderžku*) 'to render (support)'. This concept seems to be introduced in order to obtain a large number of nominalizing transformations, so that the discriminative power of the model is considerably increased. However, some objections can be made. The construction of the phrase *okazyvat' podderžku* can be regarded not only as a possible transformation of *podderživat'* 'to support', but also as a construction characterizing the verb *okazyvat'* 'to render'. In the latter case, the construction of the phrase *On okazyvaet mne podderžku* 'He renders me support' is not uniquely determined: it is to be represented by N^1_n *okazyvat'* $N(V)_a\, N^2_d$ if it is considered a transformation of $N\ {}^1_nV\ N^2_a$ (*On podderživaet menja*) 'He supports me') and by $N^1_n\ V\ N^2_d\ N^3_a$ if it characterizes the verb *okazyvat'*, so that the intersection of the sets of grammatical and lexical morphemes is not empty. If this possibility is excluded, that is, if the verb *okazyvat'* is always regarded as a grammatical morpheme in this kind of phrases, two questions should be asked: firstly, if the character (grammatical or lexical) of such a morpheme depends on the context in which the verb is used, how can it in a given context be determined (on formal grounds!), and secondly, which verbs can be 'delexicalized' and which retain their lexical meaning in combination with a given $N(V)$? [Examples of doubtful cases: *polučit' podderžku ot* N_g 'to receive support from *N'*, *pol'zovat'sja podderžkoj* N_g 'to enjoy the support of *N'*, *naxodit' podderžku u* N_g 'to meet with support among *N'*.] It seems very difficult to give a formal criterion for this categorization. One might be tempted to regard any verb in combination with a $N(V)$ as a grammatical morpheme, but that would lead to a vicious circle, since the concept of syntactic derivation has been defined by means of the transformability of the phrases containing V and $N(V)$ and the transformability of these phrases depends on the identity of the lexical morphemes.

(4) The assertion according to which the immediate syntactic links between the word forms of a sentence can be established without regard to its semantic properties is open to serious doubt. The difference in position of the noun in the prepositional case in the phrases *On nadel šapku na ulice* 'He put a cap on in the street' and

On nadel šapku na mexu 'He put a fur-lined cap on' seems to contradict this assertion. Moreover, in a number of cases the establishment of the syntactic links is not clear to me. One example has already been given: in the phrases *On sidit v komnate* 'He is sitting in the room' the adjunct of place is supposed to be linked directly to the verb and not to the subject for reasons unknown to me. It should be noted that the only morphological indication of the presence of a syntactic link is the ending of the verbal word form *sidit*, which is on that account said to be linked to the subject of the phrase. Another example of this kind of difficulties and their implications for the subsequent analysis will be given below.

2.3. The set of correct sentences

Not only the formal syntax of the language but also the set of correct sentences is supposed to be known. I wonder if the criterion of correctness used by Apresjan is a linguistic one. Describing a language is describing the way things can be said in that language and not the way things are said in that language, because the latter depends not only on the former but also on the things that are to be said, and these are determined by nonlinguistic factors (at least if the topic of the conversation is not linguistics). Therefore, if a sentence is considered to be incorrect according to the criterion used by Apresjan, it should be examined whether this incorrectness is a consequence of an inherent feature in the structure of the language or an implication of the fact that the situation in reality corresponding to this linguistic expression happens to be unknown to the informant. In the latter case, the sentence should be regarded as linguistically correct if it is a correct mapping (and in many instances even the only one) of the set of real situations to which it corresponds. If such linguistic expressions are rejected, the set of described items is the set of occurring situations in the shape in which they appear in the language, and not the set of linguistically possible expressions.

It may be illustrative to give an example. In the sentence of A. P. Čexov *Poselenec* (...) *obzavoditsja xozjajstvom, a čerez dva-tri goda emu sažajut sovladel'ca* 'The settler acquires a household, and in a couple of years they get him a joint owner' the verb *sažat'* 'to seat, plant' is found in the construction $N_n^1\ V\ N_d^2\ N_a^3$, which is regarded as incorrect by Apresjan. But it can be argued that the mere use of

this sentence by Čexov is sufficient proof of the fact that the structure of the language does not exclude the possibility. The only difficulty consists in the circumstance that real situations corresponding to this expression are not too frequent, since people are ordinarily neither seated nor planted when they become a joint owner. A similar case, at least from this point of view, is presented by the phrases *Čuvstva xorošo vladejut im* 'He is reigned well by his feelings' and *Gost' prixoditsja mne bratom po voskresen'jam* 'The visitor is a brother of mine every Sunday', which are incorrect according to Apresjan. The concept of correctness is crucial in such cases, as it is on account of the supposed incorrectness of the above sentences that the transformability of the quasi-transformable phrases *On vladeet čuvstvami* 'He has control over his feelings' and *Čuvstva vladejut im* 'He is reigned by his feelings' or *Brat prixodit ko mne v gosti* 'My brother comes to me on a visit' and *Gost' prixoditsja mne bratom* 'The visitor is a brother of mine' is denied. In view of this, it should be investigated whether the results of the analysis would be strongly affected by a change in the concept of correctness, that is, whether the discriminative power of the model and the distribution of the identified semantic units over the semantic space would be altered considerably if a different criterion of correctness were applied.

2.4. The concept of transformability

One of the key concepts of the investigation is the transformability of two phrases. However, it seems to me that this concept has not been formalized to the extent required for an automatic analysis based on exclusively formal data. The difficulties in determining the set of delexicalized morphemes characteristic of a large number of nominalizing transformations and the possibly high sensitivity of transformational syntactic features to the correctness of the transforms have already been mentioned. But the major obstacle is the very definition of the concept. It may be recalled that two quasi-transformable phrases are called transformable into each other if there is for any expansion of either of them an expansion of the other such that the two expanded phrases are quasi-transformable into each other. I wonder if there is a single pair of Russian phrases which is transformable according to this definition. Even in such trivial cases as *On pisal pis'mo* 'He wrote,

was writing a letter' versus *On napisal pis'mo* 'He wrote, has written a letter' (opposition of aspect) and *On napisal pis'mo* versus *On napišet pis'mo* 'He will write a letter' (opposition of tense) expansions disproving the transformability of these phrases can be found, e.g. *Každyj den' on pisal pis'mo* 'He wrote a letter every day' versus **Každyj den' on napisal pis'mo* and *Včera on napisal pis'mo* 'He wrote a letter yesterday' versus **Včera on napišet pis'mo* or *Zavtra on napišet pis'mo* 'He will write a letter tomorrow' versus **Zavtra on napisal pis'mo*. The transformability of the phrase *On ljubit detej* 'He loves children' and its nominal transform *U nego ljubov' k detjam* 'He has love of children' (adduced by Apresjan as an example of transformable phrases) is disproved by the expansions *On očen' ljubit detej* 'He loves children very much' and *U nego bol'šaja ljubov' k detjam* 'He has a great love of children', because no corresponding expansion of the other transform exists.

If larger expansions are taken into account the matter soon becomes very complicated. The phrase *On ljubit detej bol'še čem starux* 'He loves children more than old women' can be transformed into either *U nego ljubvi k detjam bol'še čem k staruxam* 'He has more love for children than for old women' or *U nego ljubov' k detjam bol'še čem k staruxam* 'He has a greater love for children than for old women'.

(a) If both transforms are regarded as correct, they are transformable into each other as well. In that case, the phrase *U nego mnogo ljubvi k detjam* 'He has much love for children' should be transformable into the two phrases corresponding to the other constructions, which is obviously not the case. Moreover, the sentence *On ljubit detej bol'še vsego na svete* 'He loves children more than anything in the world' may be transformable into *U nego bol'še vsego na svete ljubov' k detjam* 'More than anything in the world he has love for children', but the sentence *U nego bol'še vsego na svete ljubvi k detjam* has a good chance to be rejected.

(b) Alternatively, it can be argued that either of the phrases mentioned is not a transform because the identity of the direct syntactic links has not been maintained. It is not clear to me how the syntactic links are established in these cases, but obviously the result of the analysis strongly depends on the way this is done. Let us assume that the phrase *U nego ljubov' k detjam bol'še čem k staruxam* is identified as the nominal transform of *On ljubit detej bol'še*

čem starux because it may be considered an expansion of the phrase *U nego ljubov' k detjam* which is said to be the nominal transform of *On ljubit detej*. In that case, a direct syntactic link is supposed to exist between *u nego* and *ljubov'* and not between *u nego* and *bol'še*, while in the phrase *U nego ljubvi k detjam bol'še čem k staruxam* the opposite may be assumed. The difficulties mentioned under (a) are thus eliminated, but it is doubtful whether:
(1) the syntactic links thus established correspond to some linguistic phenomenon;
(2) such an establishment of the syntactic links is an operation that can be executed without regard to any semantic data; [it seems to me that (1) and (2) are mutually exclusive]
(3) this solution is in accordance with the semantic facts. [This is questionable in the given example.]

Since it turns out that the strict observation of the transformability conditions leads to major difficulties, an attempt should be made either to formalize the weaker criterion used in practice or to find a criterion that can be formalized more easily and determines a set of features with a comparable discriminating power. This requires further investigation.

2.5. The productivity of a semantic pattern

The border line between productive and improductive semantic patterns is claimed to be found objectively as the minimum of a numerical function. The idea is that the number of semantic patterns with a large or a small productivity is relatively large, while only a fairly limited number of patterns has a productivity of about 0.5. This assumption is not supported by the facts, however. The minimum found by Apresjan is wholly due to the incorrect use of statistical methods. The set of ideal phrases is divided into ten 'equal' productivity classes: the first class, containing the ideal phrases with a productivity of less than 0.1, is empty; the second class contains the ideal phrases with a productivity of less than 0.2 but not less than 0.1; the third class contains the ideal phrases with a productivity of less than 0.3 but not less than 0.2, and so on. If we assume that no ideal phrase is characterized by more than eight nonzero syntactic features (I have not found an example contradicting this assumption in the book under review), then:
the 1st class contains no ideal phrases;

the 2nd class contains the i.p. with productivity $\frac{1}{6}$, $\frac{1}{7}$, $\frac{1}{8}$;
the 3rd class contains the i.p. with productivity $\frac{1}{4}$, $\frac{1}{5}$, $\frac{2}{7}$, $\frac{2}{8}$;
the 4th class contains the i.p. with productivity $\frac{1}{3}$, $\frac{2}{6}$, $\frac{3}{8}$;
the 5th class contains the i.p. with productivity $\frac{2}{5}$, $\frac{3}{7}$;
the 6th class contains the i.p. with productivity $\frac{1}{2}$, $\frac{2}{4}$, $\frac{3}{6}$, $\frac{4}{7}$, $\frac{4}{8}$;
the 7th class contains the i.p. with productivity $\frac{2}{3}$, $\frac{3}{5}$, $\frac{4}{6}$, $\frac{5}{8}$;
the 8th class contains the i.p. with productivity $\frac{3}{4}$, $\frac{5}{7}$, $\frac{6}{8}$;
the 9th class contains the i.p. with productivity $\frac{4}{5}$, $\frac{5}{6}$, $\frac{6}{7}$, $\frac{7}{8}$;
the 10th class contains the i.p. with productivity $\frac{1}{1}$, $\frac{2}{2}$, $\frac{3}{3}$, ..., $\frac{8}{8}$.

If we assume that all possible productivities mentioned here have the same frequency of occurrence (which is a very unrealistic assumption), the relative number of ideal phrases found in any but the last two productivity classes shows a striking similarity to the number actually found by Apresjan:[6]) the same absolute minimum in the fifth class and a similar relative minimum in the eighth class are detected. It is clear from this example that the greatest care is needed whenever an interpretation is based upon statistical data.

The hypothesis of Apresjan (relatively few semantic patterns with a productivity of nearly 0.5) can be tested within each class of ideal phrases containing any verb V_o as a central member such that the largest number of nonzero syntactic features characterizing any ideal phrase with the verb V_o as a central member is fixed (that is, the numerical expression of the productivity has a fixed denominator). If a minimum is detected in each of these classes and if these minima coincide or nearly coincide, the hypothesis is confirmed. Since the complete material under investigation is not available to me, I have applied the test to the sample of 228 ideal phrases given by Apresjan as an example of the items in his dictionary. It turns out that the number of semantic patterns with a productivity of $\frac{4}{8}$, $\frac{3}{6}$, $\frac{3}{5}$, $\frac{2}{3}$ is larger than the number of patterns with a productivity of $\frac{8}{8}$, $\frac{6}{6}$, $\frac{5}{5}$, $\frac{3}{3}$ respectively, and thus constitutes the absolute maximum in the corresponding class. Minima are found at $\frac{2}{8}$, $\frac{1}{7}$, $\frac{2}{6}$, $\frac{1}{5}$, $\frac{1}{4}$, $\frac{1}{3}$ and $\frac{7}{8}$, $\frac{6}{7}$, $\frac{5}{6}$, $\frac{4}{5}$. These results not only contradict the hypothesis of Apresjan, but may even be considered an indication of the contrary (maximal number of semantic patterns with a productivity of nearly 0.5).[7])

---

[6]) See p. 92 of the book under review.

[7]) It should be remarked that the sample is not representative, as the

2.6. FINAL REMARKS

As to the semantic space defined in the last chapter of the book under review, it certainly is an interesting attempt to find a quantitative expression of the semantic distance between two lexemes. It should be borne in mind, however, that the method developed here rests entirely upon the qualitative part of the analysis, and thus on its theoretical suppositions. Therefore, it seems to me that the structure of the space resulting from the definitions should be regarded first of all as a test of these suppositions rather than as an adequate expression of linguistic facts.

There is one more remark that I would like to make. The description resulting from Apresjan's theory is presented as a description of semantic facts. In some cases, it is clear that the results of the analysis are affected by formal fortuities, e.g. when semantically closely related verbs govern different cases (*imet' dom* versus *vladet' domom* 'to own a house'). Such phenomena are not too numerous and may have a historical explanation, but nevertheless reduce the value of a synchronical description in which they are not eliminated. In other cases, however, we can hardly assume any influence of such factors upon the formal characteristics of the verb: if the syntactic features of a verb are considered a mapping of its semantic features, a theory linking these two levels should at least yield acceptable results whenever the syntactic features of the verb have only recently come into existence. Therefore, an application of the proposed method to a set of relatively recent borrowings and neologisms or even to a set of artificial words would seem to offer an interesting test for the validity of its theoretical suppositions. It should be demonstrated that such verbs are distinguished by the method (at least to an extent corresponding with the semantic difference between them) and identified as semantically closely related to the verbs with a similar meaning that belong to the older vocabulary. [The recent borrowings that I found in the book under review are characterized by a relatively large number of syntactic features.]

It is not the purpose of these critical remarks to deny the primary

---

ratio of the number of ideal phrases and the number of lexical items is about 3 in the complete material and 3.8 in the sample. This does not necessarily lead to a bias in these results, however.

importance of the book under review. On the contrary, I consider the investigation by Apresjan a historic step toward the creation of a formalized theory of semantics. It is the first description of a large body of semantic facts which is based upon a highly formalized theory. This has not led to a solution of the most fundamental problems of semantics, but it has shown a way to both a useful solution of a number of practical problems in compiling a good dictionary and a more rigorous formulation of a number of theoretical questions.

University of Amsterdam F. H. H. KORTLANDT
Author's address:
Frissenstein 345,
Amsterdam, The Netherlands.