Anna Ginter

Semantic Aspects of the 'Meaning - Text' Theory

MUZEUM HISTORII POLSKI

*Anna Ginter**

SEMANTIC ASPECTS OF THE 'MEANING – TEXT' THEORY

1. INTRODUCTION

The early American linguists (Bloomfield, Harris) neglected the area of meaning. They made any positive contribution whatsoever to the theory or practice of semantics. Moreover, semantics was frequently defined to be outside linguistics proper (compare: Lyons 1970: 33). Semantic considerations were strictly subordinated to the task of identifying the units of phonology and syntax. Consequently, this part of the grammar was to be independent of semantics.

Noam Chomsky was among the linguists who included semantics as an integral part of the grammatical analysis of languages. His *Aspects of the Theory of Syntax* (published in 1965) present a model of transformational grammar designed for the analysis of natural languages, which tries to explain correspondences between the syntactic structures and their meaning. The most successful theory that can be described as a complemention of the generative--transformational approach was demonstrated in the Meaning-Text Model prepared by Mel'čuk, Zholkovsky and Apresjan. Some of its terminology as well as its foundations can be found in Chomsky's theory. What is more, the Meaning-Text theory formulates the more radical point of view on semantics, which is the idea of language-independent semantic representation.

For the reasons presented above, the main objective of the present paper is to analyse the Meaning-Text Model as a final stage in the investigation of language functioning (among the theories viewed as a continuation of Chomsky's theory). The fundamental concern throughout this analysis of the semantic aspect will be the problem of relationship between syntax and semantics, and, consequently, between sound and meaning.

* Uniwersytet Łódzki.

2. GENERAL REMARKS

The Meaning – Text theory (MTT), put forward by Igor Mel'čuk and Alexander Žolkovsky in 1965 (Jurij Apresjan joined them shortly afterwards), describes a natural language as a kind of logical device, called the Meaning-Text Model of natural language (Mel'čuk 1995: 17). Consequently, conceived and developed as a general theory of human language, the Meaning-Text theory is based on the following two postulates (according to: Mel'čuk 1981: 28–29):

Postulate 1.

Every speech event presupposes three main components:

a) content or pieces of information to be communicated, which are called *meaning(s)*,

b) certain forms or physical phenomena to be perceived, which are called *text(s)*,

c) a many-to-many correspondence between an infinite set of meanings and an infinite set of texts, which constitutes *language proper* (or 'language in the narrow sense of the term').

A natural language is viewed as a logical device which establishes the correspondences between the infinite set of all possible meanings and the infinite set of all possible texts and vice versa. This device ensures the construction of linguistic utterances which express a given meaning, i.e. speaking, and the comprehension of possible meanings expressed by a given utterance, i.e. the understanding of speech.
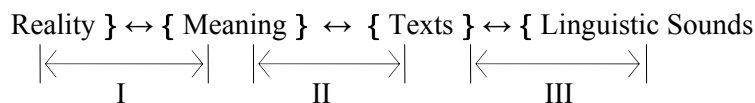
Postulate 2.

Hypotheses about devices of the type illustrated in Postulate 1 can be formulated as functional or cybernetic models, with the actual language considered as a 'black box' where only the inputs and outputs can be observed but not the internal structure. (see also: Grzegorczykowa 1990: 79) Such models (called MTM) are systems of rules approximating the Meaning – Text correspondence.

As a result of the two above Postulates, the MTM can be characterised by the following properties:

1. The MTM is not a generative but rather a translative or purely transformational system. It does not seek to generate all and only grammatically correct or meaningful texts, but merely to match any given meaning with all texts having this meaning (synonymy) and, conversely, any given text with all meanings the text can have (homonymy).

2. The MTM is no more than a fragment of the full-fledged model of human linguistic behaviour.

Reality } ↔ { Meaning } ↔ { Texts } ↔ { Linguistic Sounds

$$|\xleftrightarrow{\hspace{2cm}}|\quad|\xleftrightarrow{\hspace{1.5cm}}|\quad|\xleftrightarrow{\hspace{2cm}}|$$
$$\quad\text{I}\qquad\qquad\text{II}\qquad\qquad\text{III}$$

Only Fragment II, i.e. Language, is deemed to be the subject of linguistics proper and should be represented as an MTM. Fragment I is the subject of various fields such as philosophy and psychology, including what is called artificial intelligence. Fragment III is the subject of acoustic and articulatory phonetics.

Mel'čuk admits, however, that even Fragment II is not represented by MTM in full. To simplify their task, he and Žolkovsky made abstraction from a number of relevant aspects, properties and phenomena of natural language in order to obtain a clearer and more insightful picture of the problem (see: Mel'čuk 1981: 29).

Nevertheless, MTM deals with formal representations of meanings and texts, and Formula I can be written as follows:

{Sem(antic) R(epresentation)$_i$ **}** ↔ **{**Phon(etic) or graph(ic) R(epresentation)$_J$**}**

$$|\xleftrightarrow{\hspace{5cm}}|$$
$$\text{MTM}$$

One of the most basic facts about natural language is the following: *a given meaning of sufficient complexity can normally be expressed by an astronomically large number of texts*.

Natural language is a system capable of producing a great many synonymous texts for a given meaning. The MTM has to match a given meaning with many different texts, and many different texts have to be reduced to the same meaning representation. Thus it is necessary to establish correspondences between semantic and phonological representations, which is a multi-stage process involving also the other levels of representation.

The MTM distinguishes the following four major levels of linguistic representation: the semantic, the syntactic, the morphological and the phonological/orthographic. All levels, with the exception of the semantic one, are divided into two sublevels: a deep one (referred to meaning) and a surface one (determined by physical form). As a result, there are seven representation levels in the MTM (Mel'čuk 1981: 32-33; Mel'čuk 1995: 17):
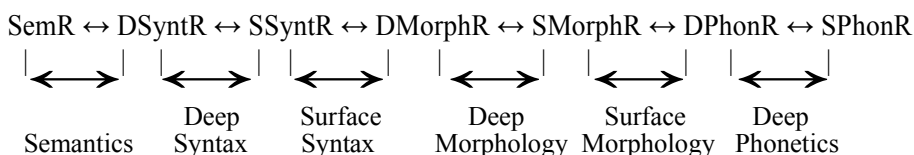
1. Semantic Representation (SemR), or the meaning,
2. Deep-Syntactic Representation (DSyntR),
3. Surface-Syntactic Representation (SSyntR),
4. Deep-Morphological Representation (DMorphR),
5. Surface-Morphological Representation (SMorphR),

6. Deep-Phonetic Representation (DPhonR, or what is commonly called 'phonemic representation'),

7. Surface-Phonetic Representation (SPhonR, which is called 'phonetic representation'), or the text.

A representation has been defined by Mel'čuk as 'a set of formal objects called 'structures', one of which is distinguished as the main one, with all the others specifying some of its characteristics. Each structure depicts a certain aspect of the item considered at a given level' (Mel'čuk 1981: 33).

Formula I can be presented in full in the following diagram:

SemR ↔ DSyntR ↔ SSyntR ↔ DMorphR ↔ SMorphR ↔ DPhonR ↔ SPhonR

| ↔ | | ↔ | | ↔ | | ↔ | | ↔ | | ↔ |

| | Deep | Surface | Deep | Surface | Deep |
| Semantics | Syntax | Syntax | Morphology | Morphology | Phonetics |

The top line in the diagram is a sequence of the utterance representations of all seven levels, with the correspondences between any two adjacent levels shown by two-headed arrows. The bottom line shows the components of the MTM and their functions. Thus semantics provides for the correspondence between the semantic representation of an utterance and all the sequences of deep-syntactic representations carrying the same meaning, etc.

For the purpose of the present discussion, there will be presented in detail only the Semantic Representation and the Syntactic Representations (Deep and Surface) as those most significant in the analyses of semantic aspects of the theory.

3. SEMANTIC REPRESENTATION

The Semantic Representation (SemR) of an utterance consists of two structures:

1) the semantic structure (SemS),

2) the semantico-communicative structure (SemCommS).

The Semantic Structure specifies the meaning of the utterance independent of its linguistic form. The distribution of meaning among the words, clauses, or sentences is ignored; so are such linguistic features as the selection of specific syntactic constructions. At the same time, the SemS tries to depict the meaning objectively by leaving out the speaker and his intentions, which are taken into account in the second structure of the SemR. Formally, a SemS is a connected graph or a network.

The vertices or nodes of a SemS are labeled with semantic units, or 'semantemes' meanings which can be either elementary ('semes') or complex. Complex meanings consist of semes or less complex semantemes. A complex semanteme can be represented by a semantic network, which specifies its semantic decomposition (Mel'čuk 1981: 34).

There are distinguished two major classes of semantemes:

1. Functors, subdivided into:
   a) predicates (relations, properties, actions, states, events, etc.)
   b) logical connectives (*if, and, or, not*)
   c) quantifiers (*all, there exist*, and all numbers).

2. Names (of classes) of objects, including proper names (Mel'čuk 1981: 35).

Both types of semantemes can receive arcs and arrows, but only a functor can head an arrow. The arrows on the arcs point from functors to their arguments. The arcs and arrows of a SemS are labeled with numbers (like in the diagram below) which have no meaning of their own but only serve to differentiate the various arguments of the same functor. For example:
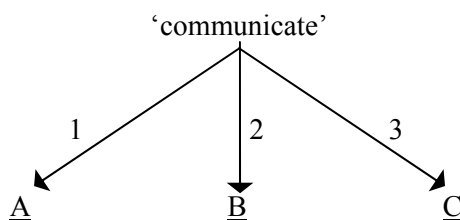
A – the first argument of 'communicate' (who communicates),

B – the second argument (what is communicated),

C – the third argument (to whom the information is passed).

The exact role of each argument is specified by further decomposition of the functor. 'A communicates B to C' = 'A, who is aware of B, explicitly causes C to become aware of B' (Mel'čuk 1981: 36). A deeper decomposition would reveal more subtle links between the functor 'communicate' and its arguments.

The Semantico-Communicative Structure (SemCommS) specifies the intentions of the speaker with respect to the organisation of the message. The same meaning reflecting a given situation can be encoded in different messages according to what the speaker wants to express. Consequently, the SemCommS must show at least the following contrasts:

(a) Theme (topic) vs. rheme (comment), i.e. the starting point of the utterance, its source, as opposed to what is communicated about the topic.

(b) Old, or given (known to both interlocutors), vs. new, i.e. communicated by the speaker.

(c) Foregrounded (expressed as a main predication) vs. backgrounded (relegated to an attribute).

(d) Emphatically stressed vs. neutral.

4. DEEP-SYNTACTIC REPRESENTATION

A Deep-Syntactic Representation (DSyntR) consists of four structures:
1) the deep-syntactic structure
2) the deep-syntactico-communicative structure,
3) the deep-syntactico-anaphoric structure,
4) the deep-syntactico-prosodic structure.

The Deep-Syntactic Structure (DSyntS) is a dependency tree. It represents the syntactic organisation of the sentence in terms of its constituent words and relationships between them. A node of a DSyntS is labeled with a generalised lexeme of the language. The symbol of a generalised lexeme must be subscribed for all the meaning-bearing morphological values, such as number in nouns or mood, tense and aspect in verbs. A generalised lexeme has been described by Mel'čuk (1981: 38–39) as one of the following four items:

1) A full lexeme of the language. Semantically empty words, like (strongly) governed prepositions and conjunctions or auxiliary verbs, are not represented.

2) A fictive lexeme, i.e. a lexeme presupposed by the symmetry of the derivational system, yet nonexistent.

3) A multilexemic idiom, e.g. *hit if off* 'have good rapport' or *pull a fast one on someone* 'gain an advantage over an unsuspecting person by subterfuge'.

4) A (standard elementary) lexical function (f), which is a relation which connects a word of phrase W (the argument of f) with a set $f(W)$ of other words or phrases (the value of f) in such a way that:
   a) for any W^1 and W^2, if $f(W^1)$ and $f(W^2)$ exist, both $f(W^1)$ and $f(W^2)$ bear an identical relationship with respect to the meaning and syntactic role to W^1 and W^2, respectively. $[f(W^1)\ W^1 = f(W^2)\ W^2]$
   b) in most cases, $f(W^1) \neq f(W^2)$, which means that $f(W)$ is phraseologically bound by W.

For an illustrative purpose, Mel'čuk provides some examples of lexical functions (1981: 39). It would be worth quoting here some of them:

Syn(*to shoot*) = *to fire* [synonym]

Anti(*victory*) = *defeat* [antonym]

Magn(*need*) = *great*, *urgent*, *bad*, Magn(*settled*[*area*]) = *thickly*, Magn(*to illustrate*) = *vividly*, Magn(*belief*) = *staunch* ['very', an intensifier]

AntiReal$_1$(*promise*) = *to renege on*, AntiReal$_2$(*attack*) = *to beat back*, IncepOper$_1$(*fire*) = *to open*, FinOper$_2$(*control*) = *to get out of*.

The Deep-Syntactico-Communicative Structure (DSyntCommS) specifies the division of the sentence represented into topic and comment, old and new, etc.

The Deep-Syntactico-Anaphoric Structure (DSyntAnaphS) carries the information about coreferentiality.

The Deep-Syntactico-Prosodic Structure (DSyntProsS) represents all of the meaningful prosodies that appear at this level: intonation contours, pauses, emphatic stresses.

5. SURFACE-SYNTACTIC REPRESENTATION

The Surface-Syntactic Representation of a sentence consists of four structures corresponding to those of the Deep-Syntactic Representation but replacing Deep by Surface. The SSyntS is also a dependency tree, but its composition and labeling differ sharply from those of the DSyntS, since a node of a SSyntS is labeled with an actual lexeme of the language:

1) all the lexemes of the sentence are represented, including the semantically empty ones;

2) all the idioms are expanded into actual surface tree;

3) the values of all the lexical functions are computed (on the basis of a lexicon called an Explanatory Combinatorial Dictionary) and spelled out in place of the Lexical Functions;

4) all pronominal replacements and deletions under lexical or referential identity are carried out (Mel'čuk 1981: 41).

A Surface-Syntactic Relation (SSyntRel), which is a branch of a SSyntS, belongs to a set of language-specific binary relations that obtain between the words of a sentence, each describing a particular syntactic construction.

The SSyntCommS, the SSyntAnaphS and the SSyntProsS are analogous to their deep counterparts.

6. THE DESIGN OF THE MEANING-TEXT MODEL

As it has been stated earlier, a MTM has the task of establishing correspondences between the semantic and the (surface-)phonetic representations of any given utterance through the five intermediate levels. Accordingly, the MTM consists of the following six basic components (Mel'čuk 1981: 43):

1. The Semantic component (semantics)
2. The Deep-Syntactic Component (deep syntax)
3. The Surface-Syntactic component (surface syntax)

4. The Deep-Morphological Component (deep morphology)
5. The Surface-Morphological Component (surface morphology)
6. The Deep-Phonetic Component (phonemics)

The Surface-Phonetic component, which provides for the correspondence between a surface-phonetic representation and actual acoustic phenomena, falls outside the scope of the MTM model in the strict sense (compare: Mel'čuk 1981: 44).

Each component of the MTM is a set of rules having the trivial form:

$$X \leftrightarrow Y \,\big|\, C,$$

where: X – a fragment of utterance representation at level n,
$\quad\quad$ Y – a fragment of utterance representation at level $n + 1$,
$\quad\quad$ C – a set of conditions (expressed by Boolean formulas) under which the correspondence $X \leftrightarrow Y$ holds.

The two-headed arrow must be interpreted as 'corresponds', not 'is transformed into'. Thus, when the transition from a meaning 'X' to a DSyntR Y is performed, 'X' itself is not changed: nothing happens to 'X' while Y is being constructed by semantic rules under the control of 'X'. The relation between a representation n and an 'adjacent' representation $n+1$ is the same as that between the blueprint of a house and the house itself, if illustrated by using Mel'čuk's example. 'The blueprint is by no means transformed into the house, but during construction, it is the blueprint that guides the workers' (Mel'čuk 1981: 44).

As Mel'čuk underlines, the rules in the MTM are logically unordered. All relevant information about the language is expressed explicitly, i.e. by symbols within the rules rather than by the order of the latter. The reason for this decision is that 'finding the best order of rule application in a specific situation goes far beyond the task of linguistics proper'. Moreover, 'the rules themselves are conceived of not as prescriptions, or instructions of an algorithm, but rather as permissions and prohibitions, or statements in a calculus. Basically each rule is a *filter* sifting out wrong correspondences' (Mel'čuk 1981: 44).

7. THE SEMANTIC COMPONENT OF THE MTM

The semantic component establishes the correspondences between the SemR of an utterance and all the synonymous sequences of DSyntR's of the sentences that make up that utterance. To do that, it performs the following eight operations (Mel'čuk 1981: 44–45):

1. It cuts the SemR into subnetworks such that each corresponds in its semantic 'size' to a sentence.

2. It selects the corresponding lexemes by means of semantico-lexical rules.

3. It supplies meaning-bearing morphological values of lexemes by means of semantico-morphological rules.

4. It forms a DSyntS (a tree) out of the lexemes it has chosen.

5. It introduces the DSynt-AnaphS; that is, it indicates coreference etc. for the lexical nodes that have appeared as a result of the duplication of some semantic nodes.

6. It computes the prosody of the sentence on the basis of semantico-prosodic rules.

7. It provides the DSynt-CommS (topic-comment, etc) from the data contained in the Sem-CommS.

8. For each DSyntR produced the semantic component constructs all the synonymous DSyntRs that can be exhaustively described in terms of lexical functions. This is achieved by means of a paraphrasing system that defines an algebra of transformations on such DSyntR's where the DSyntS contains lexical function symbols. The paraphrasing system includes rules of two classes:

1. Lexical paraphrasing rules, which represent either semantic equivalences or semantic implications; examples:

Equivalences:

$W \leftrightarrow \text{Conv}_{21}(W)$ *The set contains* $[W]$ *the point* M.
     $\leftrightarrow$ *The point* M. *belongs* $[\text{Conv}_{21}(W)]$ *to the set*

$W \leftrightarrow \text{N}_0(W) + \text{Oper}_1(\text{N}_0(W))$ *He warned* $[W]$ *them*
     $\leftrightarrow$ *He issued* $[\text{Oper}_1(\text{N}_0(W))]$ *a warning*
     $[\text{N}_0(W)]$ *to them*

$\text{Real}_2(W) \leftrightarrow \text{Adv}_{1B}(\text{Real}_2(W))$ *He followed* $[\text{Real}_2(W)]$ *her advice* $[W]$ *to enroll*
     $\leftrightarrow$ *He enrolled on* $[\text{Adv}_{1B}(\text{Real}_2(W))]$ *her*
     advice $[W]$

Implication:

$\text{PerfCaus}(X) \leftrightarrow \text{PerfIncep}(X)$ *John started* $[\text{PerfCaus}(run \text{ II})]$ *the motor*
     $\rightarrow$ *The motor started* $[\text{PerIncep}(run \text{ II})]$
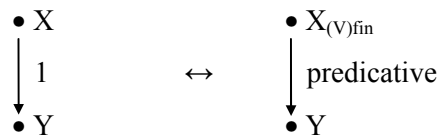
2. Syntactic paraphrasing rules – indicate what restructuring of a DSyntS is needed when a particular lexical rule is applied. Since there are only four basic syntactic operations at the deep level (merger of two nodes, splitting of a node, transfer of a node to another governor and renaming of a branch) and only nine deep-syntactic relations, the number of elementary deep-tree transformations is finite. Any particular syntactic rule defined on DSyntSs can be represented in terms of those specific transformations (Mel'čuk 1981: 47).
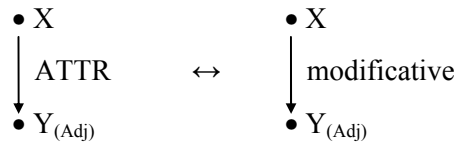
8. THE DEEP-SYNTACTIC COMPONENT OF THE MTM

The deep-syntactic component establishes the correspondence between the DSyntR of a sentence and all the alternative SSyntRs which correspond to it. To do that, it performs the following five operations (Mel'čuk 1981: 48–50):

1. It computes the values of all lexical functions by means of the specific rules.

2. It expands the nodes of idioms into corresponding surface trees.

3. It eliminates some DSyntSs nodes that occur in anaphoric relations and should not appear in the actual text.

4. It constructs the SSyntS by means of three types of transformations:

a) replacement of a DSyntRel by a SSyntRel:

A.

$$\bullet X \qquad\qquad \bullet X_{(V)fin}$$
$$\downarrow 1 \qquad \leftrightarrow \qquad \downarrow predicative$$
$$\bullet Y \qquad\qquad \bullet Y$$

B.

$$\bullet X \qquad\qquad \bullet X$$
$$\downarrow ATTR \qquad \leftrightarrow \qquad \downarrow modificative$$
$$\bullet Y_{(Adj)} \qquad\qquad \bullet Y_{(Adj)}$$

C.

$$\bullet X \qquad\qquad \bullet X$$
$$\downarrow ATTR \qquad \leftrightarrow \qquad \downarrow quantitative$$
$$\bullet Y_{(Num)} \qquad\qquad \bullet Y_{(Num)}$$
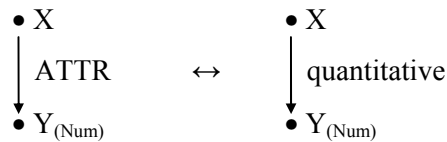
b) replacement of a DSynt-node by a SSyntRel:

$$\bullet X_{(N)}$$
$$\searrow ATTR$$
$$\bullet BE \qquad \leftrightarrow \qquad \bullet X_{(N)}$$
$$\downarrow 2 \qquad\qquad\qquad \downarrow appositive$$
$$\bullet Y_{(N)} \qquad\qquad\qquad \bullet Y_{(N)}$$

c) replacement of a DSyntRel by a SSynt-node:

$$\bullet\, X_{(V,\,3[TO])} \qquad\qquad \bullet\, X_{(V,\,3[TO])}$$

$$\Big|\,3 \qquad\qquad \leftrightarrow \qquad\qquad \Big|\,2^{nd}\ \text{completive}$$

$$\bullet\, Y \qquad\qquad\qquad\qquad \bullet\, TO$$

$$\Big|\,\text{prepositional}$$

$$\bullet\, Y$$
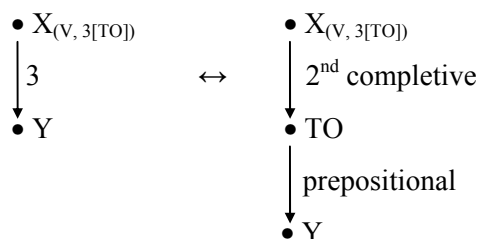
$X_{(V,\,3[TO])}$ – indicates a lexeme whose 3rd DSynt-valence slot must be filled in the SSyntS by the infinitive marker TO, the corresponding information is stored in the dictionary entry of X.

   5. It processes the other three structures of the SSyntR.

9. THE SURFACE-SYNTACTIC COMPONENT OF THE MTM

   The surface-syntactic component establishes the correspondence between the SSyntR of a sentence and all the alternative DMorphRs that are realisations of it. It performs the following four main operations (Mel'čuk 1981: 50):
1. Morphologisation of the SSyntS, i.e. it determines all the syntactically conditioned morphological values of all the words, such as the number and person of the verb.
2. Linerisation of the SSyntS, i.e. it determines the actual word order of the sentence.
3. SSynt-ellipsis, i.e. it carries out all kinds of conjunction reductions and deletions that are prescribed by the language in question, for example:
     George will *take the course*, and Dick might *take the course*, too.
   → George will take the course, and Dick might too.
4. Punctuation, i.e. it determines, on the basis of the DSynt-ProsS as well as on the basis of the resulting SSyntS, the correct prosody, which, in the case of printed texts, is rendered by punctuation.
   Surface syntax uses at least five types of rules:
   (1) syntagms (SSynt-rules), which can be illustrated with the following three rules for English (Mel'čuk 1981: 50–51):
   a) 'To build a predicative construction with a NP as the grammatical subject, one can put the subject either before the verb if the standard function 'obligatory inversions of the subject and the verb' does not apply, or after the verb if a similar function ('possible or obligatory inversions of the subject and the verb') applies, the subject must not be in objective form (relevant only for

pronouns *me*, *us*, etc.), and the verb must agree with it in accordance with yet another standard function 'agreement of the verb with the subject noun'.'

b) 'A noun Y with the syntactic feature 'temporal' which is an adverbial modifier of a verb can be placed before or after the verb if Y itself has a dependent Z other than an article.'

c) 'A noun Y governing THE or a personal adjective (*my*, *your*, etc.) and being an appositive to a human nonpronominal noun X can be placed after X in such a manner that its article or the personal adjective follows X immediately.'[1]

(2) word-order patterns for elementary phrases; for example: 'all those beautiful French magazines' vs. *'French those beautiful all magazines'.

(3) global word order rules, which compute the best word order possible for the given SSyntS on the basis of various data; the syntactic properties of some words marked in the lexicon; the relative length of different parts of the sentence, topicalisation, emphasis, etc.; possible ambiguities produced by specific arrangements, etc. These rules try to minimise the value of a utility function that represents the 'penalties' assigned (by the linguist) to certain unfelicitous arrangements. The rules do this by reshuffling the constituent phrases within the limits of what the syntagms allow.

(4) Ellipsis rules.

(5) Prosodic or Punctuation rules.

10. AN EXPLANATORY COMBINATORIAL DICTIONARY

An Explanatory Combinatorial Dictionary (ECD) is an essential component of any linguistic description within the Meaning-Text theory. It is to comprise all the semantic and combinatorial data concerning the relationships of a given word to other words.

The lexical entries of an ECD are designed both to test and to demonstrate the apparatus devised for the description of any type of lexical unit within the framework of the MT approach. This apparatus for lexical description claims to do justice to words both as paradigmatic units that enter in the network of relations obtaining in the lexicon of the language, and as syntagmatic units capable of being multifariously and systematically related to other such units in a discourse (compare: Mel'čuk 1995: 18).

The ECD is a monolingual dictionary featuring the following five important properties:

---

[1] For further information and diagrams illustrating the rules see: Mel'čuk 1981: 50–51.

1. it is active: it is oriented not only toward making texts comprehensive (i.e. providing for the transition from a text to the meaning expressed by it), by also toward assisting the user in the production of texts (i.e. providing for the transition from a meaning to the texts which express it). The objective of this type of dictionary is to give the user as complete a set as possible of the correct means for the linguistic expression of a desired idea (Mel'čuk 1995: 19).
2. it is generalist (not specialised): the ECD attempts to systemise all synonymic means of expressing a given idea (Mel'čuk 1995: 19).
3. it includes a great deal of encyclopaedic information, strictly distinguishing the encyclopaedic from the linguistic information proper (it presents them in different sections of a dictionary entry) (Mel'čuk 1995: 23).
4. it pursues theoretical goals: the ECD is completely theory-oriented. It is conceived and implemented within the MT theory, and the lexicographic method used is intimately tied to this general linguistic framework. It is designed primarily for scientific purposes and tries to bridge the chasm between lexicography and theoretical linguistics by laying the basis for an interaction between both fields (Mel'čuk 1995: 22).
5. it strongly emphasises the systematic, explicit and formalised presentation of all information made available (Mel'čuk 1995: 23).

The ECD allows for the representation of the following three basic types of relations between lexical items (Mel'čuk 1995: 26–27):

I. Semantic (paradigmatic) relationships between words and phrases, for example: synonymy, antonymy, semantic proximity, etc. They are reflected in the definitions of related words: two fully synonymous words have identical definitions; two nearly-synonymous words have nearly identical definitions; and so on. The 'definition' formulates one discrete sense of the entry lexical unit, i.e. the sense of a lexeme or a phraseme; and it does this in terms of specially selected elementary semantic units (= word senses) and/or 'derived', or intermediate, semantic units, i.e. word senses which are more basic than the word sense being defined and which are themselves defined quite independently of the entry unit. Thus, in the ECD, each word sense is semantically decomposed (except the semantic primitives).

II. Syntactic (syntagmatic) relationships between the entry lexical item, which is semantically a predicate, and other words or phrases (in a sentence) which are syntactically dependent on it and express its 'semantic actants'. These sentence elements are said to fill in the slots of the 'active syntactic valence' of the entry lexical item and are called its 'syntactic actants'. The active syntactic valence is specified by means of a table called a Government Pattern. The government pattern supplies the following three major types of information:

− for each semantic actant of the entry lexical item, it indicates the corresponding syntactic actant;

− for each syntactic actant of the entry lexical item, it indicates the form which this actant takes on the surface (grammatical case, infinitive or a finite form, prepositions, conjunctions, etc.);

− for all syntactic actants, it indicates which of them are incompatible in the sentence (or, conversely, are inseparable, i.e., invariably used together), and under what conditions.

III. The third type includes lexical (both paradigmatic and syntagmatic) relationships between the entry word and those other words which can either replace it in a text (under specific circumstances), or be joined to it in more or less fixed word combinations (also known as 'collocations'). This involves Lexical Functions.

As it has been already mentioned, the basic unit of an ECD is a dictionary entry corresponding to a single lexeme or a single phraseme, i.e. one word or one set phrase taken in one separate sense. A family of dictionary entries for lexemes which are sufficiently close in meaning and which share the same signifier (i.e. have an identical stem) is subsumed under one vocable (which is identified in upper-case letters before all of the dictionary entries it covers, and in page headings as well).[2] The ordering of the lexemes within a single vocable tends to follow a logical principle: if the definition of lexeme L' mentions lexeme L belonging to the same vocable, then L' must follow L. In other words, an 'including' definition always follows the 'included' one, so that within a vocable all interlexemic references with inclusion are only made backwards (compare: Mel'čuk 1995: 28).

A regular dictionary entry of a lexical unit L includes three major divisions:

− the signified of L, or the semantic part;

− the signifier of L, or the formal (i.e. morphological and phonological) part;

− the syntactics of L, or the combinatorics part.

To these are added the illustrations part, the encyclopaedic part, etc.

The structure of a dictionary entry consists of the ten following zones (they are given in the order they have in actual entries) (Mel'čuk 1995: 29–47):

1. Morphological information about the entry lexical unit L (declension or conjugation type; gender of nouns; aspect of verbs; missing or irregular forms; etc.). For a complete phraseme, its Surface-Syntactic tree is also quoted in this zone.

2. Stylistic specification, or usage label (specialised, i.e. technical; official; informal, colloquial, substandard; poetic; obsolescent, archaic; etc.).

3. Definition of L, consisting of constants (elementary and complex word senses of the language in question) and variables (X, Y, Z...), the latter being

---

[2] For further information concerning the structure of an ECD entry, see: Mel'čuk 1995: 27–28.

present if L happens to be a predicate (in the logico-semantic sense of the term). In this case, the item to be defined is not simply the lexical unit L as such, but an expression including L and the variables, which represent L's arguments, or semantic actants.

4. Government Pattern (GP) – this is a rectangular table in which each column represents one semantic actant of the entry lexeme (marked by the corresponding variable), and each element in the column represents one of the possible surface realisations of the corresponding syntactic actant.

5. Restrictions on the Government Pattern - these give all possible details relevant to the combinability of the entry lexeme L with its DSynt-syntactic actants and state the conditions under which these actants can/cannot co-occur.

6. GP Illustrations – the GP and all the restrictions on it are illustrated by all possible combinations of the entry lexeme L with all its actants as well as by all the impossible combinations prohibited by those restrictions.

7. Lexical Functions – this zone, characterising the idiomatic (language-specific) substitutability and cooccurrence relations of the entry lexeme L, makes up the major part of a dictionary entry. Among the standard simple lexical functions used in the ECD there are: Syn (synonym), Anti (antonym), Gener (generic concept), Dimin (diminutive), Augm (augmentative), Magn ('very', 'to a [very] high degree), etc. (compare: Apresjan 2000: 56–59)

8. Examples – the use of the entry lexeme and the corresponding lexical functions is exemplified by actual sentences.

9. Encyclopaedic Information – is admitted to the extent to which it is vital for the correct use of the entry lexeme. This information includes, among other things, an indication of the different species or different parts/stages of the object or process denoted by a key-word or entry lexeme, the main types of its behaviour, its co-species, etc. (compare: Apresjan 1995: 43–47)

10. Idioms – a list of semantically unanalysable idiomatic expressions in which the given entry lexeme appears. The list includes expressions that, on the one hand, cannot be decomposed in constituent parts with 100% compositionality and regularity, and on the other hand, are not representable in terms of lexical functions. These expressions are mentioned in the entry of the headword for reference purpose only.

11. CONCLUSION

Since the MT theory is strongly practice-oriented, some ideas of the MTM approach have found applications in different theoretical frameworks within linguistics. Among the possible applications Mel'čuk mentions:

− language teaching and learning, where the lexical functions and ECD-type dictionaries might be especially useful;

− automatic text processing – since the MTM formalisms lend themselves easily to computer programming, several text-processing computerised systems at various linguistic levels, including syntax, dictionary and morphology, have been developed in the USSR within the MTT framework (compare: Shal'apina 1996: 114);

− translation practice and theory – using the fact, that MTT is primarily concerned with the translation of meaning into text and vice versa. Hence, its relationship to translation in general is direct. Specifically, an explanatory combinatorial dictionary can be used as an effective translator's tool (compare: Mel'čuk 1993: 16–56);

− anthropological research – it has been observed that specialists in field of linguistics and exotic languages find handy and helpful tools in the realm of MT theory (Mel'čuk 1981: 55–56).

Moreover, an Explanatory Combinatorial Dictionary can find use in the following areas (Mel'čuk 1995: 47–48):

1. The ECD is likely to become a central component of automatic text synthesis and analysis systems, since it presents all the essential information about the vocabulary of the language in question in an explicit and systematic way.

2. The ECD represents a contribution to language theory, as it provides for the development and refinement of a semantic metalanguage, the systematic account of phraseology, and the development of a multifaceted approach to the word taken as the sum of all its semantic and syntactic characteristics.

3. Various potential advantages are provided by the ECD in the area of language instruction (of both native and foreign tongues), as well as in any activity connected with the development of language skills. Textbooks, pedagogically-oriented dictionaries, reference works, etc., can be successfully developed along the format of the ECD.

However, although considered by many linguists as complete and sufficient models of human speech (see: Grzegorczykowa 1990: 74; Apresjan 1990: 123), in their actual form MTMs observe numerous limitations (see: Mel'čuk 1981: 29–30). From the point of view of modern linguistic theories, it should be admitted that no attempts have been made to relate the MTM experimentally with psychological or neurological reality. For that reason an MTM is no more than a model, or a handy logical means for describing observable correspondences.

BIBLIOGRAPHY

Apresjan J. (Апресян Ю. Д.) (1990), *Формальная модель языка и представление лексико-графических знаний*, [в:] „Вопросы языкознания", № 6, сс. 123–139.

Apresjan J. (1971), *Koncepcje i metody współczesnej lingwistyki strukturalnej* (*zarys problematyki*), przeł. Z. Saloni, PIW, Warszawa.

Apresjan J. (Апресян Ю. Д.) (1995), *Образ человека по данным языка: попытка системного описания*, [в:] „Вопросы языкознания", № 1, сс. 37–67.

Apresjan J. (2000), *Semantyka leksykalna. Synonimiczne środki języka*, przeł. Z. Kozłowska i E. Janus, Wyd. Ossolineum, Wrocław.

Chomsky N. (1975), *Aspects of the Theory of Syntax*, The M.I.T.Press, Cambridge, Massachusetts.

Chomsky N. (1972), *Syntactic Structures*, Mounton, The Hague–Paris.

Grzegorczykowa R. (1990), *Wprowadzenie do semantyki językoznawczej*, PWN, Warszawa.

Lyons J. (1970), *Chomsky*, Fontana/Collins, London.

Lyons J. (1984), *Semantyka*, przeł. A. Weinsberg, t. 1–2, PWN, Warszawa.

Mel'čuk I. (1981), *Meaning – Text Models: A Recent Trend in Soviet Linguistics*, Annual Reviews Inc. Anthropol, pp. 27–62.

Mel'čuk I., **Ravič R.** (И. А. Мельчук, Р. Д. Равич) (1967), *Автоматический перевод*, под ред. Г. С. Цвейга и Э. К. Кузнецовой, Изд. АНСССР, Москва.

Mel'čuk I. (1995), *The Russian Language in the Meaning – Text Perspective*, Wiener Slawistischer Almanach, Sonderband 39, Moskau–Wien.

Shal'apina Z. M. (Шаляпина З. М.) (1996), *Автоматический перевод: эволюция и современные тенденции*, [в:] „Вопросы языкознания", № 2, сс. 105–117.

Anna Ginter

SEMANTYKA W TEORII 'SENS – TEKST'

S t r e s z c z e n i e

Strukturalizm amerykański w swoich wczesnych założeniach odrzucał semantykę jako przedmiot badań językoznawcy, a tym samym jako część nauki o języku. Dopiero Noam Chomsky w pracy *Aspects of the Theory of Syntax* zaprezentował model gramatyki transformacyjnej, w którym podjął próbę wyjaśnienia zależności między strukturami składniowymi i ich znaczeniem. Za jego kontynuację uznana została teoria 'Sens – Tekst', opracowana przez Mielczuka, Żołkowskiego i Apresjana, w której zawarte zostało najbardziej radykalne wśród ówczesnych poglądów podejście do semantyki jako reprezentacji niezależnej od innych elementów modelu. Teoria ta wywarła niezaprzeczalny wpływ na rozwój językoznawstwa, dlatego też celem niniejszego artykułu jest przedstawienie głównych założeń teorii 'Sens – Tekst' i jej struktury, ze zwróceniem szczególnej uwagi na relacje między składnią i semantyką.