Lexicographic description of words and collocations:

Feature-functional model

V. N. Telija

The lexicon can be divided into three distinct sets. *The first set* includes neutral words and collocations, e.g. скала (cliff), тигр (tiger), etc. - natural kinds; стол (table), книга (book), etc. - artefacts. Idiomatic collocations, e.g. анютины глазки (= forget-me-nots)[1], желтый дом (= bedlam) - 'a hospital for the mentally-deranged' and lexical collocations, e.g. Белый дом (The White House), оказать помощь (to provide assistance), etc., also belong to this set.[2] *The second set* embraces words and collocations of rational evaluation: they indicate whether the thing-meant is good or bad on the basis of the speaker's/hearer's axiological norms, e.g. квадратный нос (bulbous nose), увлекательная книга (exciting/interesting book), мчаться (to rush), огромный успех (tremendous success), *огромный провал (* tremendous flop). *The third set* is composed of all the expressive denominative entities, e.g. баран (= goose - about a human being), подзуживать (to egg/goad on); idioms - ахиллесова пята (Achilles' heel); собаку съесть 'to be very experienced in smth' (= Jack-of-all-trades), lexical collocations - бурный скандал (= barroom/heated scandal/brawl), удариться в амбицию (= to stand one's ground), etc.

The members of the second and the third sets are not distinguished in lexicography, as a rule, and therefore follow one standard description mode (if any): evaluative component is indicated, either implicitly or explicitly, but no mention of whether the evaluation is of a rational or emotional nature is made. A few examples will suffice to substantiate this observation.

The Dictionary of the Russian language in 4 vol. (Moscow, 1984, vol. IV, p. 114) gives only two meanings of the word скоморох (= a fair clown/a joker/a motley fool): 1. In Ancient Russia: wandering actors, who were simultaneously singers, street dancers, musicians, gymnasts, etc. and author of most of the performances they acted out... 2. Colloq.: about a person who makes others laugh at his kojes and tricks (ibid.). The first entry is a description of a functional term (kind of permanent occupation). The second entry gives an integral description of two meanings at once: (a) positive evaluation (a native speaker is supposed to have a certain normative-evaluative «picture of the world», which can be represented as a scale with «+» and «–» at its poles: 'about a jolly person, making others laugh at his jokes, gestures, etc., and it is «good»; and (b) expressiveness, or to be more exact, expressive colouring: 'about a person making fun like a fair clown in a callous way, and it is «bad», and it evokes disapproval (or better - disdain and disrespect) on the part of the speaker/hearer'. It is worth noting that the expressive meaning incorporates evaluation and becomes a

1. When there is no direct equivalent of the Russian collocation in English, then its closest resemblance is adduced as an example. This is marked by (=).

2. The term «lexical collocations» is adopted after M. Benson (1989); in Soviet linguistics there exists a corresponding term «phraseological word-combination» introduced in 1946 by the late professor V. V. Vinogradov (1977). This term is a restricted modification of Ch. Bally's «phraseological groups» (Bally, 1951). The term «feature (or parametrical) word-combinations», expressing lexical functions, is also used. It was worked out by A. K. Zholkovsky and I. A: Melchuk (1965, 1984).

more complex entity; besides, in the examples (a) and (b) given above the sign of evaluation changes: (a) - «+»; (b) - «–» (here emotiveness is also added in the form of disdain).

It is apparent that to give combined descriptions of such meanings in one dictionary entry is a glaring simplification of the lexicon. Regrettably, this is common practice, rather than an exception. The examples similar to those above, could be easily found in large numbers, but the two illustations are enough to prove that lexicography lacks any consistent distinction between the evaluative (rational evaluation) meaning per se and the expressively coloured meaning (emotional evaluation, or emotiveness). It is clear that the latter has structural distinctions as well: it is more complex, «superimposing» on rational evaluation (hence the origin of the term «expressive colouring» which can be applied to the text entities along with the lexicon entities). The process of superimposing is always motivated - through metaphor, derivational associations or sound symbolism. Comp., скоморох (a trickster, a fair clown) - 'about a jolly person, making others laugh at his jokes...' and (b) - 'about a person making fun in a callous, vulgar way (as if X were a vulgar joker)'. The metaphor intrinsic in this meaning «lowers» it in rank and incites pejorative attitude; comp. also: скоморошничать (= to make vulgar jokes and tricks) where in Russian the suffix of subjective evaluation нича- serves to express both a negative evaluation and refers the word to the pejorative register. Comp. also denominations like белиберда (= nonsense), тары–бары–растабары (= idle talk), etc., which are both negative evaluations and belong to the pejorative register, showing disdain for the thing-meant. Comp. English denominations: feeb, jumbo, piggi-wiggi, to panhandle, loudmouth.

A competent lexicographic description of words and collocations must account for the differences in evaluative and expressive meanings. It must retain all the usage features which are revealed in speech. Such a description would correspond to L. Wittgenstein's (1953) thesis that meaning is use. Lexicography must strive to carry out this task if it intends to deal with the actual use of the language, otherwise the dictionary transforms from a reliable guide to the verbalised storeroom of national culture into a semblance of «the blind leading the blind» in the Bible. Thus, if the definition of the meaning of the word сын (son), which is realised in lexical collocations like сын Востока (= the son of the Orient), сын отчизны (= one's motherland's son), сын своего времени (= the son of one's time) where the genitive is restricted to a narrow group of denominations - place, nationality, social party, historic event or epoch, - if this definition fails to indicate the evaluation with the superlative degree with «+», or the emotiveness marker - «approbatory» (what is said with approval), then it becomes unclear why there is a ban on word-combinations like Все сыны народа встали на защиту отечества (= *All their motherland's sons rose to defend their nation) - here the quantifier все (all) rules out the evaluation лучшие «the best». Contradictory to the norm are also utterances like *Сыны своего времени довели страну до крайней нищеты (= *The sons of their time brought the country to the brink of poverty) this use is only acceptable in a speech game producing a sarcastic effect, for it is hardly likely to speak about poverty and approve of those who are to blame for its cause, etc. Due to these reasons the following definition of the word сын (son) seems inappropriate: 'a person who is born or is living in a certain area or representing a certain nationality' (ibid., p. 325): this use can correspond to the choice of denominations like южанин (a southerner), кавказец (a Caucasian), грузин (a Georgian), etc.

In our opinion, the feature-functional model seems to be the most suitable one to define the meanings of such words and collocations. This model provides two conditions for the definition to be adequate to the thing-meant. (1) It reflects the function of the entity, which enables it to point to its «place» in a given code (grammar) - to its morphological or syntactic function, or to point to something in the inner or outer world of a person, which is viewed as objective reality - to the entity's semantic function (close to Ch. Morris' understanding), or to point to any possible kinds of subjective modality, which correspond the objective content of meaning with the subject's evaluation of the thing-meant. - to pragmatic functions (we understand pragmatics in a narrower sense than Ch. Morris, as we don't consider any situational knowledge, but only that which is pertinent to evaluation. It is noteworthy that knowledge about the thing-meant remains within the range of semantics). The second condition for the definition to be adequate is that the suggested model classifies the meaning and represents it in the form of heterogeneous macrocomponents, each of them being a «data-block» of homogeneous information (to use a computer metaphor). This data-block is held together by any of the above-mentioned functions and is composed of the system of features (likes «semantic components» or «semes» or «semantic primitives» according to A. Wierzbicka, 1972).

The set of data, singled out in (I) and (II), constitutes lexicographical features of an entity after Yu. N. Karaulov (1981). The feature-functional representation of a word, idiom or lexical collocation gives ample grounds to consider and describe their structure as a set of elements and relations within this set. This allows, in its turn, to «dismantle» the entity into macro- and microcomponents and give their interpretation (an imitation of the grammar of understanding), besides the entity can be «reassembled» into the meaning as a whole, in accordance with the hierarchy of inclusions and implications within the set and by the entity's Gestalt-structure (after. G. Lakoff, 1977). This process imitates the mechanism of the grammar of the speaker.

The model under consideration is effective both for the computerised dictionaries (which it was originally worked out for Telija, 1990), and for the general type dictionaries. Thus, the outline of a dictionary entry in the Automated dictionary of Russian collocations is segmented into separate «zones» according to the featured discussed above (grouped into macro- and microcomponents), including the inner-form of the word as a motivating feature. The meaning of collocations is built on the basis of this integrated information. E.g., маменькин сынок (a sissy) is represented like this: 'about a young or adult male person, who is incapable of taking his own decisions because he is infantile, and it is bad; the fact that he is dependent (like a sissy), evokes disdain on the part of the speaker/hearer; the word is used in a colloquial register.' The definition of this kind is believed to be adequate in any dictionary (in the *Computer Dictionary of Russian collocations* it is automatically triggered from the entity outline). The background of any dictionary is discrete information processing with a view of its consequent synthesis; any dictionary enacts intuition and lexicographers' skill. General type dictionaries, though, are aimed at an ordinary language user's interests, which accounts for the omission of certain irrelevant signals bearing on the user's pragmatic knowledge. The example we have adduced indicates motivation, excessive for a common type dictionary (it is contained in the «literal» meaning of a collocation), the rest of the features are relevant for any dictionary: the predicative function, the thing-meant, the two types of evaluation - rational («+») and emotive/expressive, marked by «pejorative», the colloquial register.

One of the applied tasks the suggested lexicographical model can accomplish is to formalize the activity of a lexicographer, providing him/her with with the tool of «assembly/dissembly» of the information carried by an entity. At the same time the model has a considerable theoretical value - it bears on cognitive structures, represented in the form of denominations. The model can also have pedagogical applications: due to the fact that it indicates the types of denominations it can serve as a basis for taxonomy of meanings in the fields of lexicography and lexicology (primary, secondary and inner-form (Telija, 1977, 1981), descriptive vs evaluative vs expressive, neutral vs stylistically marked (Telija, 1990).

But of primary importance is the computer «operation» of this model: it enables information search along macro- and microcomponent data-blocks, i.e. with the precision of a separate component: one can project onto the display just a rational-evaluative or stylistic macrocomponent of the entity, e.g. баран (= goose) → the evaluation with «–»; vulgar; the model also enables to project all the entities having any one of the enumerated features, e.g., evaluation with «–»: упрямый (stubborn); баран (= goose), ползти (= to drag), уродливый (ugly), Медуза Горгона (Meduza the Gorgon); вкалывать (= to slave), etc.; one can call all the words and collocations marked by «disdain»: иуда (Judas); тупоголовый (= bird-brained); зубрила (swot), болтун (loudmouth), etc.; one can project onto the display the information about the metaphorical structure of the inner form of an entity, included in the motivational macrocomponent, e. g. под рукой (= to come in handy) → metaphorical (semantic) and morphosyntactic motivation; выбросить из головы (= not to give it another thought, not to rake one's brain) → quasi-symbolic metaphor «brain» → «thinking» and morpho-syntactic motivation, etc. Therefore, the feature-functional model of meaning can serve to spot any macrocomponent (data-block) of parametrically homogeneous information for any type of meaning.

It follows that a feature-functional model of meaning can be represented by blocks of information (macrocomponents), embracing parametrically similar data. In first approximation (regardless of the hierarchical structuring of signals, based on inclusion or implication of parameters) the following blocks of information corresponding to the intuitive division of meaning into «parts» can be named: the grammar data block (G); the reality data block, which provides the description of the existing objects (D); the axiological data block (A); the entity «inner form» data block - the associative motivation of the figurative meaning of the entity (M); the emotive attitude data block (E); the stylistic connotation data block (S).

It is evident that for any lexicon entity blocks G and D are obligatory, although block S should be considered equally compulsory, as the information about the socially marked/neutral communicative conditions constitutes the «communicative channel» through which communication is carried out and where certain suitable/unsuitable means of communication are filtered.

The schematic representation of the feature-functional model of meaning makes it unnecessary to discuss a number of other ontologically relevant problems: the truth-functional succession of macrocomponents, the stage at which grammar is introduced into the process of thought-to-speech generation, ways of forming the semantic and the expression planes at large, etc. The formalism of the scheme also enables to neglect the multidimensional organization of meaning, which implies that the very metaphor of «superimposed» subjective parameters leads to believe that the subjective and objective modi are located on different «planes», etc. Never-

theless, even the linear representation of the above-mentioned data block poses a number of problems, in particular - *what* implies *what*. It is assumed that the following formulaic representation of the blocks would not lead to a dramatic error:

$$S \mid (G) \cap (D) \leftarrow (A) \supset (M) \leftarrow (E)\}.$$

This formalism can be given the following interpretation: the decision about the exploitation of a certain stylistic means precedes its choice; the interaction of lexical and grammatical information («the unity of lexico-grammatical meaning; axiological attitude implies D-data, as long as he already given value is assessed; the motivational component («the inner form») introduces itself into the process of meaning formation, while the «outer form» of the motivator acquires the function of the expression plane; motivation implies the association with the Gestalt and produces the subject's emotive, emotional or evaluative attitude toward the thing-meant, which is now perceived throug the prism of the whole associative complex.

Essentially, the suggested scheme reflects the mechanism of functional parameters of the lexicon entities –ranging from words to all kinds of collocations (idioms, cliches, set phrases, etc.). The scheme can evidently be used for taxonomy ends as well. The srings $(S) \rightarrow (G) \rightarrow (D)$ triggers the set with descriptive meaning (table, to run, the White House, to take place). The evaluative macrocomponent (E), added to the string, integrates the set of evaluative words (interesting, difficult, a big crop, at least, etc.). In case a motivator as an inner form is included in the meaning, two types of information can be considered firstly, the fact that the word is the secondary domination product; secondly, the association with the Gestalt, which any metaphor lives by (e.g.: an owl, to drag - about the time, the train of thought, a rift in the lute). Finally, the emotive block constitutes the basis for the set of emotionally-coloured denominations. The analysis of linguistic data shows that the expressiveness of the lexicon entitites only occurs in presence of (M). The expressiveness itself is the final result of all the subject-oriented modi of meaning, including (M) (comp.: to urge - to egg on; a traitor - Judas; wakness of character - Achilles' heel; disturbed conscience - pangs of conscience).

Besides the taxonomic potential of the semantic block-scheme under consideration can be represented in two ways: as a declaration or as a process. Everything mentioned above, referred to the declarative presentation of macrocomponents. Yet the same scheme can be easily changed to model the procedural aspect of cognizance, aimed at the generation vs understanding of the entity meaning, imparted in a certain communicative act. Then, first of all, in the process of sense generation all the intentions or «predicates» about the reality should be given (including the indispensable knowledge about the world, and in particular - the knowledge of cultural symbolism and stereotypes - any metaphor extensions):

$$\{(S), (A), (E) \rightarrow (M)\} \supset (D) \cap (G).$$

E.g.: Think that «I» consider our relations informal; think that (D) is «bad»; think that (D) is such, as if it were (M); and this makes «I» despise (D). E.g.: Jack of all trades; blue stocking; black sheep; Judas; to panhandle; a pickpocket. The process of understanding evidently begins with (M), after that the signals of rational and emotive evaluation can be decoded.

Therefore, the phenomenological grounds for singling out the macrocomponents on the semantic plane enable their diverse interpretation (the scheme can be presented in the form of prototypes or semantic frames, etc.). The variations of the feature-functional model are possible due to the simple fact that the main aim of the block-scheme application is to account for the maximum of parameters of the meaning.

The present analysis has not concentrated on the detailed description of each of the macrocomponents, the content and structure of their constituent parameters of their formal representation, nor at the rules of transformation of any certain parameter to the top of the hierarchy. It's worth noting, though, that the suggested scheme proved effective in the phraseological computer processing of idioms and collocations in the set-up of the Computer Fund of the Russian language, the indicated entities being the signs of great semantic ambiguity and bearing a heavy pragmatic load. They contain all the above-mentioned types of information, represented in the macrocomponent block-scheme of meaning.

References

BALLY, Ch. *Traité de Stylistique Française.* Paris-Genève.

BENSON, M. (1989) «The Structure of the Collocational Dictionary». *International Journal of Lexicography*, vol. 2, N° 1.

MEL'ČUK, Igor A. and ZHOLKOVSKY, A. K. (1984) *Explanatory Combinatorial Dictionary of Modern Russian.* Vienna: Wiener Slawischer Almanach.

LAKOFF, G. (1977) «Linguistic Gestalts». Papers from the Thirteenth Regional Meeting. Chicago Linguistic Society, 13.

MORRIS, Ch. *Signs, Language and Behavior.* N.Y.

WIERZBICKA, A. (1972) *Semantic Primitives.* Frankfurt/M.

WITTGENSTEIN, L. (1953) *Philosophical Investigations.* Oxford.

WINOGRAD, T. (1976) «Towards a procedural understanding of semantics» *Revue Internationale de Philosopjie*, 117-118, Fasc. 3-4, Bruxelles, p. 260-303.

Виноградов В.В. 1977. Основные понятия русской фразеологии как лингвистической дисциплины. – В кн.: В.В.Виноградов. Избранные труды. Лексикология и лексикография. М., 118–139.

Жолковский А.К., Мельчук И.А. 1965. О возможном методе и инструментах семантического синтеза. – Научно-техническая информация. Вып. 6. М.

Караулов Ю.Н. 1981. Лингвистическое конструирование и тезаурус литературного языка. М.

Словарь русского языка в 4-х томах. 1984.Т.4.

Телия В.Н. 1977. Вторичная номинация и ее виды. – Языковая номинация (виды наименований). М.

Телия В.Н. 1981. Типы языковых значений. Связанное значение слова в языке. М.

Телия В.Н. 1990. Семантика идиом в функционально-параметрическом отображении. – Фразеография в Машинном фонде русского языка. М.