

The Theory of Lexical Functions: An Update¹

Juri Apresjan

Institute for Information Transmission Problems of RAS
127994, Moscow,
Bolshoi Karetny 19
apr@iitp.ru

Abstract

The standard theory of lexical functions (LFs) roughly claims: (a) that the values of simple LFs of the OPER-LABOR-FUNC family are semantically void; (b) that the values of collocate LFs are phraseologically bound with regard to their argument words and that, consequently, collocations of the form $L + X$, where L is the value of a certain LF from the argument X , are idiomatic intra- and interlinguistically. The present paper purports to show that all the lexemes which are values of collocate LFs are meaningful and that their choice is semantically quite well, though not completely, motivated. The basic assumption for both claims is that semantically well-formed sentences are subject to the general law of semantic agreement which requires of collocated items L and X , with the exception of a small number of genuinely idiomatic combinations, one of the following two things: either L or X meets the semantic conditions for filling the valency of the other; or both collocated items display at least one recurrent semantic component in their lexical meanings. These considerations create a foundation for at least partially predicting a set of probable values for each $LF_i(X)$. Such plausible lexicographic expectations allow to proceed from an item-by-item description of lexicon to a description of the vocabulary of a language as a system.

1 The Standard Theory of LFs

In the last quarter of the 20-th century theoretical linguistics witnessed a breakthrough of paramount importance in the study of constrained lexical co-occurrence. I mean the theory of lexical functions (LFs) proposed by Igor Mel'čuk and Alexander Zholkovsky; references are too well known to be necessary.

I shall confine myself to a discussion of only one portion of the theory bearing on the collocate LFs, otherwise called parameters.² Even out of the set of collocate LFs I shall take up a very small selection. My principal concern will be the following LFs: **MAGN**, **OPER1**, **OPER2**, **LABOR1-2**, **ADV1**, and **ADV2**; occasionally I shall mention some other LFs.

Two main properties were ascribed to collocate LFs in the early work of the two authors.

a) The simple LFs of the **OPER-FUNC** family are semantically void. It follows from the fact that such sets of sentences as, for instance,

- (1) a. *The government controls [X] all foreign trade*
- b. *The government has <exercises> [OPER1(S0(X))] control [S0(X)] over all foreign trade*
- c. *All foreign trade is under [OPER2(S0(X))] the control [S0(X)] of the government*

¹ This research has been supported in part with grants from the Russian Federation Government (No. HIII-3205.2008.6), the Russian foundation for fundamental research (No. 08-06-00344), and the History and philology department of RAS within the Program "Genesis and interaction of social, cultural, and language communities".

² I prefer to speak of "collocates" rather than "parameters" because the latter term has come to be used as a label for a certain semantic class of nouns, such as *height*, *length*, *speed*, *pressure* and the like.

are paraphrases of one another, i.e. synonymous with regard to the situation described. Since the collocations *to have control over something* and *to be under somebody's control* are referentially identical to the core verb *to control* in (1a), it seems natural to conclude that all of the lexical meaning of the collocation is contained in the argument noun *control*, while *to exercise* and *to be under* serve exclusively the auxiliary function of “verbalizing” the noun.³ In other words, their semantic contribution to the meaning of the utterance is null; see, for example, (Mel'čuk, 1974: 93) and (Mel'čuk et al., 1999: 77).

b) The values of collocate LFs are phraseologically bound, i.e. idiomatic (Mel'čuk, 1974: 80). In other words, the choice of lexeme *L* as value of a certain LF from argument *X* is semantically unmotivated, and the whole collocation *L + X* is idiomatic intralinguistically and interlinguistically. A handy example is LF **MAGN**. We usually say

- (2) a. *wolfish appetite, raging thirst*, rather than
b. **raging appetite*, or **wolfish thirst*
- (3) a. *deadly asleep, wide awake*, but not
b. **wide asleep* or **deadly awake*
- (4) a. *inveterate liar, inveterate gambler, inveterate smoker*
b. *heavy gambler, heavy smoker*, but hardly
c. **heavy liar*

Interestingly, in Russian the literal equivalents of (4b) are ruled out, and (4c) is equally impossible:

- (5) a. **tjzhelyj igrok* ‘heavy gambler’, **tjzhelyj kuril'sčik* ‘heavy smoker’
b. **tjzhelyj lžec* ‘heavy liar’

The right way to put these ideas in Russian is

- (6) a. *zajadlyj <zavzjatyj> igrok* ‘inveterate gambler’
b. *zajadlyj <neispravimyj> kuril'sčik* ‘heavy smoker’
c. *neispravimyj <ot'javlennyj> lžec* ‘inveterate liar’

At first glance at these examples the choice of the adjective to express the meaning of ‘high degree of what is denoted by the keyword’ seems unaccountable. Hence the conclusion that such collocations are idiomatic intralinguistically and interlinguistically.

As a matter of fact both these assertions were slightly qualified even in the first versions of the theory. In (Mel'čuk & Zholkovsky, 1984: 54) the LFs of the **OPER-LABOR-FUNC** family were defined as verbs “which are semantically empty in the context of the entry lexeme (= their keyword)”, with the following continuation in a footnote: “These verbs may be genuinely empty, as, for instance the verb *okazat'* (which does not mean anything definite and cannot be translated into a different language out of context), or meaningful, but then their own meaning is included in the sense of their keyword” (ibid., 96). See also (Mel'čuk, 1974: 104), (Mel'čuk et al., 1992: 32), (Mel'čuk & Wanner, 1996), with the same assertion. As concerns idiomaticity, Mel'čuk (1974: 105) notes that LF collocations on the whole are less idiomatic than genuine idioms, although some of them are very close to the latter.

Similar ideas were voiced in later work; see, for instance, (Uspensky, 1979), (Paducheva, 1991) and (Reuther, 1994, 1996, 2003). However, all of the quoted sources, except Uspensky (1979), relied mostly

³ The collocation *to be under control* whose lexical meaning may seem to be different from that of *to control* is semantically and syntactically equivalent to the passive form of the verb: *All foreign trade is under the control of the government* = *All foreign trade is controlled by the government*. Since the passive form of a verb is assumed to possess the same lexical meaning as its active form, the same should apply to the collocation *to be under control* and the core verb *to control*.

on the material of **OPER1**, especially on the argument lexemes denoting emotions, in which case **OPER1** is more or less uniformly expressed by the verbs like *to feel* (*admiration, contempt, envy, fear, indignation, jealousy, pride, shame, etc.*) and therefore is readily interpreted as semantically motivated.

2 An Update: General Considerations

In connection with my practical lexicographic research I have also taken up the subject. I have followed up the general trend of reasoning outlined in the previous work, but have made a particular point of extending it in several directions. In this section I shall give a brief sketch of the update proposed in (Apresjan, 2004, 2008a, 2008b), (Apresjan & Glovinskaja 2007). I shall argue that: 1) all collocate LFs, including the simple LFs of the **OPER-LABOR-FUNC** family, have a lexical meaning of their own; 2) the choice of a particular lexical item *L* as value of a certain LF from the argument lexeme *X* is conditioned by a) the nature of the LF in question, b) the lexical meaning of *L*, and c) the semantic class and subclass of a Vendlerian classification to which *X* belongs.

2.1 General Foundations

The general foundations of the above claims are twofold – paradigmatic and syntagmatic.

2.1.1 Paradigmatic Foundation

The paradigmatic foundation is a Vendler-like classification of predicates into actions (*to look, to read, to walk, to write*), activities (*to trade, to negotiate, to educate, war*), processes (*to flow, to grow, growth*), spatial positions (*to stand, to sit, to lie*), states (*to feel, to see, to know, need*), properties (*authority, beauty, courage, to stammer*) and so on, with further subdivisions of these major classes into a number of more compact subclasses like, say, physical states (*to see, to hear*), physiological states (*to itch, to ache*), mental states (*to think, to know*), volitional states (*to wish, to intend*), emotional states (*afraid, envy, to fear*), economic states (*need, to prosper*), social states (*married, divorced*) etc.

I should like to emphasize that this classification, if properly modified, is semantically valid with regard to any class of predicates, not only verbs.

On the other hand, it underlies not only the aspectual properties of verbs, as was currently believed, but other verbal grammatical categories like mood and voice as well. A familiar example are mental statives like *to know* which defy the use in the imperative mood and in the passive voice. We usually say

- (7) a. *You **should** know how to do it* rather than
b. ^{??}***Know** know how to do it!*
- (8) a. *This fact is known **to** everybody* rather than
b. [?]*This fact is known **by** everybody*, with a genuine agentive complement

Last, but not least: the semantic classes and subclasses to which a given predicate belongs condition to a large extent not only its purely grammatical properties, but also its government pattern and much of its combinatorial profile.

For example, the overwhelming majority of many place verbal predicates starting from four-actant verbs and on denote actions (not activities, processes, states, or anything else). Notoriously abundant in many-place predicates are semantic domains of causing locomotion, creating physical objects, exchange of valuables between two persons and some others. Here are some examples:

- (9) a. *The allies [A1] transported troops [A2] from Britain [A4] to France [A3] by aircraft [A5]*

- b. *The girl [A1] sewed a dress [A2] (for her doll) out of leftover pieces [A3] on her mother's sewing machine [A4] with silk thread [A5]*
- c. *NN [A1] rented <leased> his house [A2] to us [A3] for a year [A5] for 5000 dollars [A4]*
- d. *We [A1] rented <leased> this house [A2] from NN [A3] for a year [A5] for 5000 dollars*

As concerns combinatorial profiles, I shall have more to say on the subject a little later.

2.1.2 Syntagmatic Foundation

Syntagmatically the above theoretical claims are based on the well-known fact that the majority of collocations in natural languages are subject to the general law of semantic agreement. This law requires of the collocated items *L* and *X* one or both of the following two things:

(a) they should have at least one non-trivial recurrent (repetitive, common) semantic component in their meanings (or, technically speaking, in their meaning definitions); the greater the number of recurrent components in the phrase or sentence the greater the degree of its semantic cohesion;

(b) a potential actant *A_i* of predicate *P* should meet the semantic requirements for the *i*-th valency of *P*, irrespective of whether this valency is active or passive (for similar ideas see (Iordanskaja & Mel'čuk) in the present volume).

Illustration of the first requirement: in the phrase

(10) *to cook fish and chips*

all the three words are polysemous.

To cook means 1) 'to prepare food for eating by using heat' or (coll.) 2) 'to invent something', as in *to cook a story*.

Fish means 1) 'the flesh of a water animal used for food' or (coll.) 2) 'a person with a salient trait', as in *cool fish, poor fish, odd fish*.

Chip means 1) 'a long thin piece of potato cooked in hot fat or oil' or 2) 'a small piece of silicon used to store and process information in computers'.⁴

Even if we confine ourselves just to those six senses (as a matter of fact, there are many more) we shall get eight possible paths of reading (10), e.g. 'to invent a person with a salient trait and a long thin piece of potato'. However, (10) is unequivocally understood in just one reading (10'):

(10') 'to cook the flesh of a water animal and long thin pieces of potato in hot fat or oil'

The intuitively obvious choice of the only reading (10') as semantically cohesive is ensured by the fact that the number of recurrent senses for it, namely, 'food', 'heat', 'hot (oil)' etc., is the greatest.

Illustration of the second requirement: let us look at the phrases

- (11) a. *auburn hair*
- b. ??*auburn horse* <*furniture*>

As is well known, the color adjective *auburn* means 'reddish-brown' and normally applies to human hair; this is the semantic requirement that any noun filling the (passive) valency of *auburn* should meet. From this point of view (11a) accords with the law of semantic agreement while (11b) deviates from it.⁵

⁴ Meaning definitions here and elsewhere are highly informal.

⁵ For what follows the distinction between cases (a) and (b) is of little importance, and for the most part I shall ignore it.

All of the above is more or less common knowledge. I have recalled it to make the following less trivial assertion: the general law of semantic agreement holds good not only for free word combinations but for the overwhelming majority of LF collocations as well.

2.2 Two Illustrations

2.2.1 The Case of *okazyvat'*

I shall start with the allegedly empty Russian verb *okazyvat'* which is used in Modern Russian only as part of LF collocations⁶. Let us look at the following two phrases where *okazyvat'* is opposed to *imet'* 'to have' in the function of **OPER1** from the same noun.

- (12) a. *okazyvat' vlijanie na voennyx* 'to exert influence on the military'
b. *imet' vlijanie sredi voennyx* 'to have influence among the military'

The difference in the values of **OPER1** is obviously due to the fact that *vlijanie* in these two sentences is used in two different senses, i.e. represents two different lexemes: *vlijanie 1* in (12a) denotes a kind of pressure, that is an **action**; *vlijanie 2* in (12b) denotes the ability to affect somebody's actions and decisions without using force or orders, that is a certain **property** of a person.

This fundamental semantic difference is directly mirrored in the respective **synonym** series: the synonyms of *vlijanie 1* are nouns like *vozdejstvie* 'action' and *davlenie* 'pressure', while the synonyms of *vlijanie 2* are nouns like *avtoritet* 'authority' and *ves* 'weight' (in the figurative sense).

Indirectly the semantic opposition 'action vs. property' is very consistently reflected in a number of non-semantic distinctions between *vlijanie 1* and *vlijanie 2*.

Grammatical distinctions: *vlijanie 1* has the plural form, while *vlijanie 2* has not: *različnye vlijanija, kotorym on podvergalsja vo vremja učebny v Garvarde* 'various influences he was subject to during his studies at Harvard', but not **različnye vlijanija, kotorye on imel v voennyx krugax* 'various influences he had in military circles'.

Derivational distinctions: *vlijanie 1* is derived from the verb *vlijat* 'to influence'; there is no verbal counterpart for *vlijanie 2*. *Vlijanie 2* has a derived adjective *vlijatel'nyj* 'influential', which is impossible for *vlijanie 1*.

Government patterns. *Vlijanie 1* governs the preposition *na* 'on' (see 12a), while *vlijanie 2* governs prepositions and prepositional groups *sredi* (see 12a), *v srede*, *v krugax* 'among', 'in the circles of': *On imeet bol'soe vlijanie v voennyx krugax <v teatral'noj srede>* 'He has much influence in military circles <in the theatrical milieu>'.

Combinatorial profiles. *Vlijanie 1* has such adjectival LFs as **BON** and **ANTIBON**: *xorošee <položitel'noe, plodotvornoe> vlijanie* 'good <positive, fruitful> influence' vs. *ploxoje <durnoe, otricatel'noe, pagubnoe, tletvornoe> vlijanie* 'bad <harmful, negative, pernicious, baneful> influence'. None of these adjectives are possible for *vlijanie 2*. On the other hand, *vlijanie 2* has such verbal LFs as **INCEPOPER1** *priobretat' (vlijanie)* 'to acquire influence' and **FINOPER1** *terjat' (vlijanie)* 'to lose influence'. Neither is possible for *vlijanie 1*.

In view of these consistent and persistent distinctions between the meanings of action and property of the word *vlijanie* it seems natural to expect that the values of **OPER1** for *vlijanie 1* and *vlijanie 2* should also reflect this fundamental semantic opposition.

Indeed, people *have* properties, so the property lexeme *vlijanie 2* legitimately co-occurs with the verb *imet'* 'to have' as its **OPER1**. It would also be instructive to look at the other possible value of **OPER1** from *vlijanie 2*, which is the verb *pol'zovat'sja* 'to use'.

⁶ In the XIX-th century it had the now obsolete meaning of 'to show' which is still preserved in some Russian dialects: *Zarja okažet poljakam kak ničtožen otrjad tvoj* 'The dawn will show to the Polish how small your detachment is' (A. Marlinsky); *Zoloto staralis' ne okazyvat'* (P. Bažov) 'They tried to conceal [= not to show to anyone] the gold they had mined' (both quotes are from the Comprehensive Academic Dictionary of Russian).

- (13) a. *On pol'zuetsja bol'sim vlijaniem sredi voennyx* 'He has much influence among the military'

Pol'zovat'sja is definitely actional in its principal meaning (*On vseгда pol'zuetsja nožom, čtoby otkryt' banku* 'He always uses a knife to open a can'), yet in the meaning at issue it is purely stative. This is borne out by the fact that in the function of OPER1(*vlijanie* 2) it has no perfective form:

- (13) b. **On vospol'zovalsja bol'sim vlijaniem sredi voennyx*

In the actional meanings of *pol'zovat'sja* the perfective form is quite normal: *pol'zovat'sja* <*vospol'zovat'sja*> *nožom* 'to use a knife', *pol'zovat'sja* <*vospol'zovat'sja*> *slučaem* 'to avail oneself of the chance'.

Let us now turn to the lexeme *vlijanie* 1. It denotes an action, and actions are *performed* or *done*. It is natural to assume therefore that *okazyvat'* as OPER1(*vlijanie* 1) is "synonymous" to the above verbs, i.e. has the meaning of doing.

This assumption is corroborated with the following fact. There are about thirty collocations of nouns with the verb *okazyvat'* in modern Russian, and in most of them the nouns denote actions:

- (14) a. *Okazyvat' vlijanie* 'influence', *vozdejstvie* 'action, impact', *davlenie* 'pressure', *dejstvie* 'action', *milost'* 'a favour', *nažim* 'pressure', *podderžku* 'support', *pomošč'* 'assistance', *soprotivlenie* 'resistance', *uslugu* 'service'

Most remarkably, in some of those collocations, especially with the nouns *vozdejstvie* 'action, impact' and *dejstvie* 'action', *okazyvat'* is, or until very recently used to be, interchangeable with an undoubtedly meaningful and actional verb *proizvodit'* 'to produce'. This is convincingly attested by the Russian National corpus. Two examples will suffice:

- (14) b. *Fioletovyyj cvet proizvodit ugnetajuščee dejstvie na nervnuju sistemu* 'The violet color produces an oppressing action <effect> on the nervous system'
 c. *Buduči nematerial'nym, prostranstvo ne možet proizvodit' vozdejstvie na tela* 'Being non-material, space cannot produce any action on corporeal bodies'

In view of such facts, to postulate a null meaning for *okazyvat'* will amount to saying that a meaningless item may be synonymous to a meaningful one, let alone the fact that it will also mean postulating an inexplicable exception from the law of semantic agreement. So we are forced to the conclusion that *okazyvat'* has a meaning and that it is a very general meaning of 'doing'.

To produce more evidence for my principal claim I shall pursue the same example a little farther. Let us look at another sufficiently large class of nouns collocating with *okazyvat'* and denoting mental or emotional attitudes towards somebody or something.

- (15) *Okazyvat' vnimanie* 'attention', *znaki vnimanija* 'signs of respect', *doverie* 'confidence', *počesti* 'honors', *uvaženie* 'respect', *čest'* 'honor'

At first sight this seems to contradict the claim that *okazyvat'* means something like 'to do, to perform, to produce': one cannot *do* <*perform, produce*> *attention* or *confidence*. However, if we look closer at the collocations under (15), we shall discover that *okazyvat'* there represents an LF different from **OPER1** and, consequently, cannot mean 'doing'. In the Comprehensive Academic Dictionary of Russian (the only one to single out this meaning) *okazyvat'* in these collocations is defined as 'to display or to show one's attitude to somebody or something', which should be interpreted in terms of LFs as the value of **MANIF**. It is noteworthy that *okazyvat'* in this case is interchangeable with a more or less standard expression of **MANIF** by means of the verb *projavljat'* 'to display, to show':

- (16) a. *okazyvat' <projavljat'> doverie (k) komu-libo* 'to display <to show> confidence in somebody'
 b. *okazyvat' <projavljat'> uvaženie (k) komu-libo* 'to display <to show> respect for somebody'⁷

Now, if *okazyvat'* in (15) is the value of **MANIF**, what then is the value of **OPER1** for the nouns in question? The standard value of **OPER1** from mental and emotional attitudes is the verb *ispytyvat'* 'to feel, to experience'.

- (17) a. *ispytyvat' doverie k komu-libo* 'to feel confidence in somebody'
 b. *ispytyvat' uvaženie k komu-libo* 'to feel respect for somebody'

Mental and emotional attitudes semantically neighbor on emotional states like *admiration, fear, indignation, joy, shame* etc., which are *felt* by people.⁸ It is no wonder therefore that mental and emotional attitudes share the same **OPER1** with emotional states.

- (17) c. *ispytyvat' vosxiščenie <vozmuščenie>* 'to feel admiration <indignation>
 d. *ispytyvat' strax <styd>* 'to feel fear <shame>'

On the other hand, mental and emotional attitudes semantically neighbor on properties, and properties, as has already been pointed out, collocate with the verb *imet'* 'to have' as value of **OPER1**. Interestingly, in Russian and, more commonly, in English, the have-verbs may replace the verbs like *ispytyvat'* for this whole class of arguments. In other words, the lexemes *ispytyvat'* and *imet'* are synonymous in the context of nouns denoting mental and emotional attitudes:

- (18) a. *imet' doverie k komu-libo* 'to have confidence in somebody'
 b. *imet' uvaženie k komu-libo* 'to have respect for somebody'

We shall see more of the semantic opposition 'emotional attitudes' vs. 'emotional states' below.

Before I proceed to my next example I should like to call attention to the following fact. As is clear from my glosses all the way through, the English lexemes *influence 1* and *influence 2* display almost the same kind of semantic, grammatical, derivational, syntactic, and combinatorial distinctions as their Russian counterparts *vlijanie 1* and *vlijanie 2*. This is the first piece of evidence showing that the degree of interlinguistic idiomaticity of LFs in the standard theory has been somewhat exaggerated.

2.2.2 ADV1, ADV2 and their Compounds

My second illustration will be the LFs **ADV1** and **ADV2**, with various compositions. This time I shall make a special point of emphasizing, apart from other things, that the combinations of the form *L + X* are less idiomatic, than has been believed hitherto, not only interlinguistically but intralinguistically as well.

In English **ADV1** from nouns denoting meals is uniformly expressed by the preposition *at* + the argument lexeme: *at breakfast, at dinner, at lunch, at meal* (*The whole family meets at meals*), *at supper, at table* (= 'while eating'), *at tea*. Their Russian counterparts are uniformly expressed by the combination of the preposition *za* + the argument lexeme: *za zavtrakom, za obedom, za lančem, za edoj, za užinom, za*

⁷ The preposition *k* is bracketed because *okazyvat'* has a government pattern of its own with a prepositionless dative while *projavljat'* borrows the government pattern of the argument word.

⁸ One of the first attempts to explore linguistic differences between the names of emotional states and emotional attitudes was undertaken in (Iordanskaja, 1970).

stolom, za čaem. There is also a uniformly expressed **ADV2REAL1** for such arguments: *for breakfast, for dinner, for supper* etc.; their Russian analogues are *na zavtrak, na obed, na užin* etc.

ADV2 from nouns denoting guidance and supervision is uniformly expressed by the preposition *under* + the argument lexeme: *under somebody's control* <*direction, guidance, leadership, observation, oversight, superintendence, supervision*>. They also allow of literal translations into Russian: *pod č'im-libo kontrolem* <*rukovodstvom, voditel'stvom, nabljudeniem, prismotrom, upravljeniem, nadzorom*>.

ADV1 from the names of emotional states is uniformly expressed by the preposition *in* + the argument lexeme: *in admiration, in agitation, in amazement, in anger, in anxiety, in bewilderment, in confusion, in delight, in despair, in doubt, in embarrassment, in fury, in horror, in indignation, in panic, in rapture, in sorrow, in surprise, in suspense* etc. They also allow of literal translations into Russian: *v bespokojstve, v gneve, v izumlenii, v jarosti, v otčajanii, v panike, v pečali, v smjatenii, v somnenii, v trevože, v udivlenii, v užase, v vostorge, v vosxiščenii, v zamešatel'stve*, etc.

The same class of arguments gives rise to one more adverbial LF – **ADV1MANIF** – which can be roughly defined as ‘displaying X while doing something’. This LF is expressed by the preposition *with* + the argument lexeme:

- (19) a. **ADV1(horror)** = *in horror*, as in *to look at somebody in horror*
 b. **ADV1MANIF(horror)** = *with horror*, as in *to look at somebody with horror*

Note also collocations like *with admiration, with anger, with anxiety, with envy, with fury, with incredulity, with indignation, with joy, with pride, with shame, with surprise, etc.*, which also fall under (19b). Their Russian counterparts are uniformly expressed by the combination of the respective preposition *s* + the argument lexeme in the instrumental case: *s bespokojstvom, s gordost'ju, s izumleniem, s jarost'ju, s nedoveriem, s radost'ju, so stydom, s udivleniem, s užasom, s vostorgom, s vozmuščeniem, s zavist'ju*, etc.

There is a semantically motivated difference between the classes of arguments for those two LFs. **ADV1MANIF** is possible from the names of emotional states like *anger, despair, horror, surprise* and so on, as well as from the names of emotional and mental attitudes like *contempt, hatred, interest, love, respect, sympathy* etc., as in *with contempt, with hatred, with interest, with love, with respect, with sympathy*. **ADV1** of the form *in* + *X* for emotional and mental attitudes is impossible.

The difference is not accidental. Both, states and attitudes can be expressed outwardly; hence the collocations of the type *with admiration* and *with love*. However, it is only states that a person can **be in**. One cannot **be in an attitude**.

It is really amazing how extraordinarily sensitive to the minutest semantic distinctions a natural language can be. I mean the distinction between emotional states and emotional attitudes. It also shows up in the next group of examples illustrating the LF **ADV2CAUS(X)** whose meaning can be formulated as ‘X having been caused by the current situation P’.⁹

In English **ADV2CAUS** is uniformly expressed by the collocation *to* + the argument lexeme: *to his amazement, to the delight of the crowd, to our discredit, to my displeasure, to the horror of the spectators, to his pleasure, to my regret, to his surprise*. In Russian this LF is expressed no less consistently by the collocation *k* + the argument lexeme: *k ego izumleniju, k vostorgu tolpy, k našemu stydu, k moemu neudovol'stviju, k užasu zritelej, k ego udovol'stviju, k moemu sožaleniju, k ego udivleniju*.

This LF is possible for some emotional states and is ruled out for mental and emotional attitudes. Once again, the difference between the two classes of arguments has a profound semantic motivation. Emotions like *amazement, delight, displeasure, horror, regret* and so on are transient, often short-lived inner states which may quickly pass over after the factor causing them has ceased to act on the Experiencer.

⁹ In the entries of the Russian TKS written by L. Iordanskaja, where this LF was first singled out, it was identified as **ADV2B**.

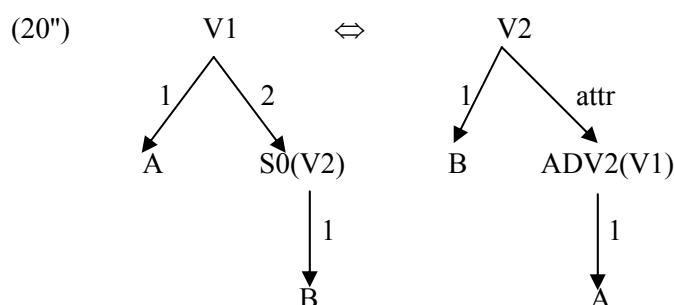
Contempt, hatred, interest, love, respect, sympathy and other emotional and mental attitudes resemble properties in that they do not arise as a transient reaction to the current situation but are rather constant.¹⁰

On the basis of these observations one can formulate two new rules of paraphrasing.

- (20) a. *A controls the work of B—B works under A's control*
 b. *A directs the work of B—B works under A's direction*
 c. *A oversees the studies of B—B studies under the oversight of A*
 d. *A superintends the studies of B—B studies under the superintendence of A*
 e. *A supervises the coaching of B—B coaches under the supervision of A, etc.*

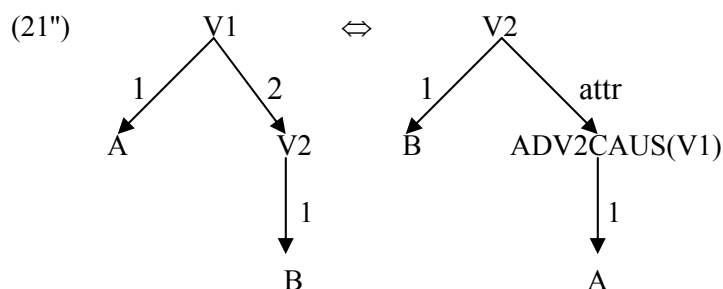
These equivalences can be formalized in the following way (V means a verb):

$$(20') \quad V1 + S0(V2) \Leftrightarrow V2 + ADV2(V1)$$



- (21) a. *She was delighted <displeased> that he refused—To her delight <displeasure>, he refused*
 b. *Jane was horrified that her son also fell ill—To Jane's horror, her son also fell ill*
 c. *I regretted very much that I lost touch with her—To my great regret, I lost touch with her*
 d. *I was surprised <astonished> that she came—To my surprise <astonishment>, she came*

$$(21') \quad V1 + V2 \Leftrightarrow V2 + ADV2CAUS(V1)$$



3 A Case Study

The starting point of this research was the Russian word *kontrol'* 'control', and the purpose was to account for the choice of *strogij* 'strict' as value of the LF **MAGN** for it. The two respective English words (or, rather, lexemes), *control* and *strict*, are semantically very similar to their Russian counterparts. So to save myself the trouble of glossing the examples I shall go straight to illustrating the point with English material.

¹⁰ More permanent states like *zavist'* 'envy' and *revnost'* 'jealousy' cannot be used in this construction either: **K ego zavisti <revnosti>, ego sopernik polzovalsja ból'sim uspexom* 'To his envy <jealousy>, his rival had a greater success'.

3.1 Semantic Correspondence Between LF Value and Argument Lexeme: *strict control*

Similarly to the related verb, the substantive lexeme *control* is a two-place predicate denoting a hierarchically ordered situation with an Agent having a higher status in the hierarchy and therefore able to influence the actions, states, and sometimes even the very existence of the Patient, no matter if the latter is a person, a body of persons or a certain state of affairs. The idea of dominance of the first actant over the second is also apparent in the meaning of the governed preposition *over* which is by no means “semantically empty”:

(22) *The president had firm control over the Cabinet*

Russian uses a very similar preposition *nad* ‘over’ to express the same syntactic and semantic relation between the two actants of *kontrol* ‘control’.

Let us now turn to the meaning of the adjective *strict* as value of the LF **MAGN** for *control*. *Strict* is a predicate describing similar relations between two participants of a hierarchical situation: the Agent has “the whip hand” over the Patient. Note, for instance, that parents can be *strict* with their children, while the latter can hardly be *strict* with their parents, unless, of course, the respective utterance is intended as a joke. The same holds good for the relations between teacher and pupil, employer and employee, examiner and examinee and so on. In all these cases the participant with a higher biological or social status can be strict with the participant whose respective status is lower, but not the other way round.

In the case at issue semantic agreement assumes the form (a), with the recurrent semantic component being the idea of hierarchical relations between the two participants of the situation: it makes part of the lexical meanings of both lexemes, *control* and *strict*, and thus accounts for the choice of *strict* as MAGN(*control*). Naturally enough, the same is true of Russian which, as has already been mentioned, chooses the adjective *strogiy* ‘strict’ as the value of **MAGN** for *kontrol* ‘control’.

3.2 One Value of LF—a Semantically Homogeneous Class of Arguments

In the preceding section we dealt with one value *L* (*strict*) of one LF (**MAGN**) with regard to one argument lexeme *X* (*control*). However, if the hypothesis of semantic agreement between *L* and *X* in LF collocations is at least partially true, then the same value of the same LF should be possible for a larger class of arguments which are semantically similar to *control*.

Control makes part of a class which comprises such nouns, as, for instance, *audit*, *censorship*, *examination*, *inspection*, *monitoring*, *oversight*, *superintendence*, *supervision*, *surveillance*, and, possibly, some other. The above assumption was checked against the material of authoritative British dictionaries (Longman, MacMillan, Oxford Advanced Learners Dictionary) and the retrieval system of the GOOGLE engine, and the search yielded many of the predicted collocations. Here are some examples:¹¹

- (23)
- a. *They ordered strict audit of oil and power companies*
 - b. *Strict censorship (was) imposed on the media in North Korea*
 - c. *This requires strict temperature control*
 - d. *They called for a strict examination of export trade refund*
 - e. *Strict quarantine inspection of US beef products (is considered necessary)*
 - f. *Their results demand stricter monitoring*
 - g. *The president promised strict oversight of a \$ 787 billion packet he signed into law this week*
 - h. *Minister Korthals wants stricter supervision of the activities of private detectives*

¹¹ In case the collocation was not recorded in dictionaries and did not occur at least once in a corpus of a hundred documents recalled by the GOOGLE engine it did not count as existing in the English language.

- i. *Strict surveillance combined with quick response could contain bird flu'*

The collocation *strict superintendence* did not occur in any of these sources, but even with this exception the prediction should be considered sufficiently reliable to be used as a lexicographic tool.

3.3 One Argument Lexeme—a Semantically Consistent Class of Values of LFs

If we pursue the same line of reasoning, the next step in it should be the hypothesis that the values *L1*, *L2*, ..., *Ln* of all LFs defined for one argument lexeme *X* should feature in their lexical meanings a semantic component which recurs in the lexical meaning of *X*. In this connection it would be natural to look at the principal verbal collocations for *control* which are those of the **OPER-LABOR-FUNC** family. The most conspicuous of them are the ones which involve the second actant of the situation – they demonstrate semantic agreement between the value of a verbal LF and the argument lexeme most convincingly.

- (24) a. *to be under* <*to undergo*> (*control*) [**OPER2**]
 b. *to fall under control* [**INCEPOPER2**]
 c. *to put* <*to place, to bring*> (*somebody*) *under (control)* [**CAUSOPER2**]
 d. *keep (something) under (control)* [**LABOR1-2**]

The values of these LFs include either the preposition *under*, or the verbal prefix *under-* (in *undergo*). In their physical senses both denote the position **below** something, and enough of this meaning is preserved in the figurative senses like that of *under (control)* to ensure semantic agreement between the preposition and the argument lexeme. Since *under* suggests the lower, subjugated position of the second participant, its choice is semantically quite well motivated.

As has been mentioned above, the hierarchical relations inherent in the meaning of *control* can be detected in the values of **OPER1** and some compositions with it as well.

- (25) a. *to exercise* <*to exert*> (*control*) *over (something)* [**OPER1**]
 b. *to have (control) over (something)* [**OPER1**]
- (26) a. *to take* <*to gain*> (*control*) *over (something)* [**INCEPOPER1**]
 b. *to lose (control) over (something)* [**FINOPER1**]

In view of the fact that all the LFs in (25) and (26) introduce the first participant whose status is **higher** than that of the second participant, the choice of the preposition *over* with the meaning of 'higher in the hierarchy' seems to be perfectly appropriate on semantic grounds. This conclusion is further substantiated by the relations of antonymy which obtain between the prepositions *over* and *under* representing the relative positions of the first and second actants of the hierarchy respectively.

It is true that *over* in these and similar collocations can sometimes be replaced with the preposition *of* which renders a much more general idea of a **relation** between two entities, not necessarily hierarchical:

- (27) a. *to be in (control) of (something)*
 b. *to have (control) of (something)*
 c. *to take (control) of (something) etc.*

In view of such examples I should like to emphasize once again that I do not suggest there is a hundred percent semantic agreement between the lexical meanings of *L* and *X* in an LF collocation *L + X*. The claim I am making is a little more modest: there is a sufficient number of collocations where the choice of verbs as values of certain LFs is so consistent as to leave no doubt that it is semantically motivated.

3.4 A Class of Argument Lexemes—a Semantically Consistent Class of Values of LFs

If there is semantic agreement between LF-values $L1, L2, \dots, Ln$ and the lexical meaning of X , then it is reasonable to expect that the combinatorial profiles of a whole class of arguments which are semantically similar to X would be largely coincident.

Before I produce evidence in favor of this assumption I should like to call attention to an interesting semantic peculiarity in the meaning of the noun *control* which makes it partly comparable to the noun *influence* analyzed above. As has been shown in paragraph 2.2.1, the word *influence* breaks up into two distinct lexemes: *influence 1*, which is an action (as in *to exert influence on somebody*) and *influence 2* which is a property (as in *to have influence among the military*). We have posited two distinct lexemes for *influence* because the number of differences in grammatical forms, government patterns, combinatorial profiles, synonyms, and derivatives between them is so great as to preclude the possibility of blending them into a single entity.

The noun *control* displays a similar duality of meaning, only in this case it is the opposition of action or activity vs. a certain state of affairs, or, for simplicity's sake, just a state. Consider collocations under (28) in which the actional features of *control* come to the foreground, and collocations under (29) where its stative features are foregrounded.

- (28) a. *to exercise <to exert> (control) over (something)* [OPER1]
 b. *to undergo (control)* [OPER2]
- (29) a. *to have (control) over (something)* [OPER1]
 b. *to be under (control)* [OPER2]

However, in the semantic structure of *control* the opposition 'actional' vs. 'stative' has not solidified to a degree when the opposed senses constitute two distinct entities: grammatical forms, government patterns, derivatives, and synonyms, if any, are the same, so both uses co-exist within a single lexeme.

With regard to this opposition the lexemes of the semantic class to which *control* belongs fall into three subclasses: 1) the subclass of purely actional lexemes, such as *audit, checkup, (university) examination, inspection, monitoring, tests, trials*; 2) the subclass of purely stative lexemes, such as *care, charge, power*; 3) the subclass of lexemes which allow of both types of uses, actional as well as stative, such as *control, supervision, surveillance*.

Proceeding from our assumptions we can expect to find significant intersections in the combinatorial profiles of those lexemes. In view of the limitations of space I shall have to confine myself to considering just one LF to substantiate this claim. Let it be **OPER2**.

Our lexicographic expectations with regard to the probable values of **OPER2** from these argument lexemes can be formulated as follows: 1) the actional lexemes of the first class will collocate with the verb *to undergo*; 2) the stative lexemes of the second class will collocate with the verbal group *to be under*; 3) the intermediate lexemes of the third class will allow of both these values of **OPER2**.

The data of the American corpus bear out all of these expectations.

- (30) *to undergo an audit <a checkup, an examination, an inspection, monitoring, tests, (sea) trials>*
- (31) *to be under the care <the charge, the power> (of somebody)*
- (32) a. *to undergo control <supervision, surveillance>*
 b. *to be under control <supervision, surveillance>*

It is noteworthy that the collocations of the type

- (33) a. ?? *to be under an audit <a checkup, an inspection, monitoring, ...>*
 b. **to undergo the care <the charge, the power> of somebody*

have not been corroborated by the corpus.

4 Conclusion

LFs collocations form a continuous space with two poles: (a) highly idiomatic collocations, like **MAGN** *wolfish appetite, raging thirst, wide awake, inveterate liar*, etc., with LF values which are possible only for a very limited number of argument lexemes and are therefore semantically hazy or downright unaccountable; (b) collocations like **BON** *bad behavior (improper is more idiomatic), bad deal (raw is more idiomatic), bad effect (harmful is more specific), hear badly (indistinctly is more specific), bad influence (baneful is more idiomatic)* etc., with LF values which are possible for a very wide range of argument lexemes and are therefore semantically clear and to a large extent predictable. The former border on pure idioms, and the latter border on free word combinations (though they are not quite free).

In between there are collocations which form the bulk of LF material. With regard to such collocations—and they have been our principal concern—one can formulate useful lexicographic expectations as to their LF potential and the possible values of concrete LFs. This allows to proceed from an item-by-item lexicographic description to a much more systematic treatment of material by compact classes of lexemes.

On the other hand it seems to be a plausible claim that the regularities we have observed reflect language competence of the speakers and should therefore be taken into account in constructing meaning-text models for particular languages.

References:

- Apresjan, Ju.D. 2004. Akcional'nost' i stativnost' kak sokrovennye smysly (oxota na *okazyvat'*). In: Apresjan (ed.), *Sokrovennye smysly. Sbornik statej v čest' N.D. Arutjunovoj*. Moskva: Jazyki slavjanskix kul'tur. 13-33.
- Apresjan, Ju.D. 2008a. O semantičeskoj motivirovannosti leksičeskix funkcij-kollokatov. *Voprosy jazykoznanija*, No. 5, 3-33.
- Apresjan, Ju.D. 2008b. Anglijskij tolkovno-kombinatornyj slovar'. I. Leksičeskije funkcii. In: A.V. Bondarko, G.I. Kustova, R.I. Rozina (eds.), *Dinamičeskije modeli. Slovo. Predloženie. Tekst. Sbornik statej v čest' E.V. Padučevoj*. Moskva: Jazyki slavjanskix kul'tur. 20-58.
- Apresjan, Ju.D. & M.Ja. Glovinskaja 2007. Two Projects: English ECD and Russian Production Dictionary. In: Kim Gerdes, Tilmann Reuther, Leo Wanner (eds.), *Meaning—Text Theory 2007. Proceedings of the 3rd International Conference on Meaning-Text Theory*. Wiener Slawistischer Almanach, Sonderband 69. München—Wien (to appear).
- Iordanskaja, L.N. 1970. Popytka leksikografičeskogo tolkovanija gruppy russkix slov so značeniem čuvstva. *Mašinnyj perevod i prikladnaja lingvistika*, vyp. 13, 3-26.
- Iordanskaja, Lidija & Igor Mel'čuk. 2009. Well-formedness of Semantic Structures (in this volume).
- Mel'čuk, I.A. 1974. *Opyt teorii lingvističeskix modelej "Smysl ⇔ Tekst"*. Moskva: Nauka.
- Mel'čuk, I.A. 1995. *Russkij jazyk v modeli "Smysl ⇔ Tekst"*. Moskva—Vena: Wiener Slawistischer Almanach, Sonderband 39.
- Mel'čuk, Igor A. & Alexander K. Zholkovsky. 1984. *Explanatory Combinatorial Dictionary of Modern Russian. Semantico-Syntactic Studies of Russian Vocabulary*. Vienna: Wiener Slawistischer Almanach, Sonderband 14.

- Mel'čuk, Igor & Leo Wanner. 1996. Lexical Functions and Lexical Inheritance for Emotion Lexemes in German. In: L. Wanner (ed.), *Lexical Functions in Lexicography and Natural Language Processing*. Amsterdam-Philadelphia: John Benjamins Publishing Company. 209-278.
- Mel'čuk, Igor et al. 1992. Mel'čuk Igor avec Nadia Arbatchewsky-Jumarie, Lidija Iordanskaja et Suzanne Mantha. *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques III*. Montréal: Les Presses de l'Université de Montréal.
- Mel'čuk, Igor et al. 1999. Mel'čuk Igor avec Nadia Arbatchewsky-Jumarie, Lidija Iordanskaja, Suzanne Mantha et Alain Polguère. *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques IV*. Montréal: Les Presses de l'Université de Montréal.
- Paducheva, E.V. 1991. Predikatnye imena v leksikografičeskom aspekte. *Naučno-tekhničeskaja informacija*, Serija 2, № 5. 21-31.
- Reuther, Tilmann. 1994. O perifrastičeskix naimenovanijax rečevoj dejatel'nosti. In: N.D. Arutjunova (ed.), *Logičeskij analiz jazyka: Jazyk rečevyx dejstvij*. Moskva: Nauka. 76-82.
- Reuther, Tilmann. 1996. On Dictionary Entries for Support Verbs: the Case of Russian VESTI, PROVODIT' and PROIZVODIT'. In: L. Wanner (ed.), *Lexical Functions in Lexicography and Natural Language Processing*. Amsterdam-Philadelphia: John Benjamins Publishing Company. 181-208.
- Reuther, Tilmann. 2003. Support verb combinations with existential verbs (German and Russian). In: Sylvaine Kahane and Alexis Nasr (eds.), *Proceedings. First International Conference on Meaning-Text Theory / Première conférence internationale sur la Théorie Sens-Texte*. Paris: Ecole Normale Supérieure. 1-10.
- Uspensky, Vladimir. 1979. O veščnyx konnotacijax abstraktnyx suščestvitel'nyx. *Semiotika i informatika*, odinnadcatyj vypusk. 142-148.