Lexical Functions in Lexicographic Description
Author(s): Igor A. Mel’čuk

Please see “How to cite” in the online sidebar for full citation information.

Please contact BLS regarding any further use of this work. BLS retains copyright for both print and screen forms of the publication. BLS may be contacted via [http://linguistics.berkeley.edu/bls/](http://linguistics.berkeley.edu/bls/).

*The Annual Proceedings of the Berkeley Linguistics Society* is published online via [eLanguage](http://elanguage), the Linguistic Society of America's digital publishing platform.
LEXICAL FUNCTIONS IN LEXICOGRAPHIC DESCRIPTION

Igor A. Mel'čuk
Université de Montréal

I
Introduction

This paper attempts to state and illustrate, in a very preliminary manner, the important linguistic notion of LEXICAL FUNCTION (introduced in Zolkovskij - Mel'čuk 1966 and 1967 and developed, at some length, in Mel'čuk 1974: 78-109). Lexical functions were devised to describe a certain type of lexical collocation, italicized in (1) - a dozen sentences collected in ten minutes from one newspaper page:

(1) a. The President clamped (= imposed) an overnight curfew on three areas... to stamp out (= put down) violence.
b. The panel issued a report to the Secretary of State.
c. President Reagan rejected pleas to open talks with striking US controllers.
d. Pope released from hospital [headline]. The Pope left the hospital yesterday, three months and one day after he was struck by two bullets. He said a brief prayer...
e. The heaviest prison terms in Kentucky history (more than 1,600 years each) have been handed down against two men.
f. South African troops have spread a dragnet across the country in a search for three heavily armed black guerillas. The ANC has claimed responsibility for the attack launched last Tuesday in which four rockets were fired at an army camp.
g. We are looking for senior consultants of proven competence to satisfy the demands of our growing business.

Texts - from colloquial to artistic to technological - swarm with expressions of this type.

To save space, I will impose the following three restrictions on my paper:

(i) Although lexical functions are crucial for the linguistic theory, I shun all theoretical discussion and cut short many explanations, even some necessary ones. ("Explanations take such a dreadful time!" - as Lewis Carroll’s Gryphon said once.) I place all my hope in the examples and in both the intelligence and good will of my readers.

The work reported in this paper has been sponsored over the past three years by the Social Sciences and Humanities Research Council of Canada, under grant No. 410-80-0345-R2.
(ii) Although lexical functions are universal, I exemplify them only from Russian. (Some French examples and a comparison with Russian are found in Iordanskaja et al.'s paper published in this volume.)

(iii) Although lexical functions are part of phraseology and play a vital role in many other language domains as well (cf., e.g., Babby 1980: 128-146 for the application of lexical functions in syntax), I consider them only from the viewpoint of lexicography and discuss them exclusively as they appear in a new type of monolingual dictionary - the Explanatory Combinatorial Dictionary (see Žolkovskij - Mel'čuk 1966, 1967, 1972; Apresjan - Žolkovskij - Mel'čuk 1968; Apresyan - Mel'čuk - Žolkovsky 1969; Apresjan - Mel'čuk - Žolkovskij 1973; Mel'čuk - Žolkovskij 1970; Mel'čuk 1974: 110 ff., 1978; Mel'čuk - Iordanskaja - Arbatchewsky 1981).

II
Definition of Lexical Function

A lexical function (henceforth, LF) \( f \) is, like any mathematical function, a dependency that associates with a given "quantity" - its argument - a variable "quantity" - its value, the latter being controlled by the former. More precisely, an LF \( f \) associates with a lexical unit \( W \) (a word or a phrase) a set \( W_f \) of (more or less synonymous) lexical units that express - contingent on \( W \) - a specific idea (such as 'very', 'begin', 'implement') represented by \( f \). For example, the LF \( \text{Magnet} \), which for the present can be glossed roughly as 'very', in conjunction with the Russian words \( \text{naprjaza}-\text{nie} \) 'voltage' or \( \text{temperatura} \) 'temperature' is expressed by the adjective \( \text{vysoči} \) 'high'; in conjunction with \( \text{vyšota} \) 'height', however, the same function is expressed by \( \text{snažitel'ný} \) 'considerable', \( \text{bol'soj} \) 'great' or \( \text{ogromný} \) 'enormous'; and in conjunction with \( \text{vibracija} \) 'vibration', by \( \text{sil'ný} \) 'strong' or \( \text{intensivný} \) 'intense'. Thus we get \( \text{vyšokoe naprjazenie} \) 'high voltage', \( \text{vyso-}\text{kaja temperatura} \) 'high temperature', \( \text{snažitel'naja} \) (bol'saja, ogromnaja) \( \text{vyšota} \) 'considerable (great, enormous) height', and \( \text{sil'naja} \) (intensivnaja) \( \text{vibracija} \) 'strong (intense) vibration'; but we do not get \( *\text{sil'noe} \) (snažitel'noe) \( \text{naprjazenie} \), \( *\text{bol'saja} \) (ogromnaja, intensivnaja, sil'naja) \( \text{temperatura} \), \( *\text{intensivnaja} \) (sil'naja) \( \text{vyšota} \), \( *\text{vyšokaja} \) (bol'saja) \( \text{vibracija} \), etc.

The importance of LF's consists in the discovery of the following fact: in all natural languages there is only a limited number (about several dozens) of meanings that resemble 'very' in that they also each determine an LF.

The exact meaning of the last statement will become clearer upon subsequent reading.

And now, a formal definition of lexical function.

A dependency \( f \) is called lexical function if and only if it associates with a lexical unit \( W \) - the argument of \( f \) - a set \( f(W) \) of lexical units - the value of \( f \) such that the following two conditions are simultaneously met:
For any two different $W^1$ and $W^2$, if $f(W^1)$ and $f(W^2)$ both exist, then:

1. Both $f(W^1)$ and $f(W^2)$ bear an identical relationship with respect to meaning and deep-syntactic role to $W^1$ and $W^2$, respectively. [This condition is language-independent.]

2. At least, in some cases, $f(W^1) \neq f(W^2)$. [This condition is completely language-dependent; it means that in the given language, the value of $f$ is phraseologically bound by its argument.]

An important proviso: A lexical function is not a genuine semantic unit, let alone a semantic primitive. LF's are introduced to describe restricted lexical cooccurrence and derivation, but by no means semantics. First, there are LF's that are semantically empty, their values being limited to purely syntactic roles (cf. $\text{Oper}_1$, $\text{Func}_1$, $\text{Labor}_1$ below). Second, the expressions making up the value of a given LF $f$ for a given argument need not be perfectly synonymous; it suffices for them to share a rather general and abstract meaning 'f' while differing in other components.

We will be interested in a particular type of LF, namely standard lexical functions, which form a proper subset of all lexical functions. A standard LF satisfies simultaneously two additional conditions:

3. $f$ is defined for a relatively large number of arguments. In other words, $f$ has a relatively large semantic cooccurrence: its meaning 'f' is sufficiently abstract to be compatible with a large number of other meanings. [This condition is language-independent.]

4. $f$ has a relatively large number of linguistic expressions as its possible values. In other words, the set of all $f(W_i)$, for a vast variety of $W_i$, is relatively rich. [This condition is completely language-dependent.]

Let me illustrate my point with two examples. The meaning 'manufactured from very dark rye flour' in Russian has three expressions: černyj 'black', ršanoj 'rye[Adj]' and is ršanoj muki 'from rye flour', and these expressions are phraseologically bound:

- Only xleb 'bread' or suxar 'rusk' can be called černyj, while very dark rye buločka 'bun', bublik 'bagel', bliny 'pancake', korž(ik) '(a kind of) flat dense dry bread', lepěška 'a flat cake' etc. cannot: *čermaja buločka 'black bun', e.g., is readily understandable but ungrammatical.

- Only xleb 'bread' and lepěška 'flat cake' but nothing else can be called ršanoj. (Notice that černyj xleb is always ršanoj xleb, but ršanoj xleb can be svetlij 'light', as well as černyj.)

- Speaking of buločki, bubliki, bliny, korži and koržiki, Russian uses is ršanoj muki (but not with reference to suxari 'rusks' and hardly with reference to xleb): *černyj/ršanoj bublik, etc.
As we see, this meaning satisfies Conditions 1 and 2 (чёрный: хлеб = из ржаной муки: пшеничный хлеб: хлебка = ...; чёрный хлеб: *чёрный хлеб, ржаная пшеничный хлебка: *ржаная пшеничный хлебка, ... ) and therefore specifies a lexical function. However, this meaning violates Conditions 3 and 4: it is conceivable only with the names of bread-like baked products made from dough (so that it is semantically too specific) and it has only three expressions. So it is not a standard LF.

The meaning 'it is necessary to P this X' has in Russian a phraseologically bound expression нужда 'need' that is possible with some P's while impossible with others: Это стат'я нужды в исправлении (в доработке) 'This paper needs to be corrected (more work)' but not *Эта статья нужда в сохранении (в отправке автору) 'This paper needs to be saved (to be sent to the author)'. This meaning satisfies — in addition to Conditions 1 and 2 — Condition 3 (it is extremely abstract) but it still violates Condition 4: there are no numerous synonymous expressions (in fact, there is only one other expression for the meaning in question satisfying Condition 1: требовать 'require', which is in free variation with нужда).

Both above-mentioned meanings — 'manufactured from very dark rye flour' and 'it is necessary to P this X' — are non-standard lexical functions in Russian. In what follows, only standard LF's are considered.

Since both defining conditions for standard LF's include the vague expression 'relatively large number', there is no sharp borderline between standard and non-standard LF's, that is, there are no formal criteria to tell them apart. This fact reflects the graduality so typical of natural languages. The concept of standard LF is fuzzy, as are most linguistic concepts.

Within the class of standard lexical functions, we will distinguish simple LF's and compound LF's, the latter being built out of the former. Notice that simple in this context by no means implies 'elementary' or 'further unanalyzable': some of our simple LF's could be represented in terms of other simple LF's (and thus could be treated as compound). Nevertheless, for purely linguistic reasons (primarily, frequency of occurrence) we consider a particular set of LF's as simple and take this set to be the basis of lexicographic descriptions.

Thus only simple standard lexical functions are included in the main body of my systematic survey. (This allows me to omit the adjectives simple standard everywhere, since the omission cannot lead to a confusion.)

III
Lexical Functions and Deep Syntax

Lexical functions are crucially linked to what is known as Deep Syntax in the Meaning-Text Model approach (cf., e.g., Mel'čuk 1981). LF's appear only in deep-syntactic structures of sentences; they are used to state paraphrase relations that hold between
sentences at the deep-syntactic level; the value of an LF is linked to its argument by a particular deep-syntactic relation; etc. Therefore, for a full explanation of LF's, a complete description of the deep-syntactic component of the Meaning-Text Model is a prerequisite; but such an undertaking is out of the question here. As a consequence, many important things cannot be explained at all, and again I appeal to the reader's good will and tolerance.

I will introduce only the most central notion of deep syntax, namely deep-syntactic actant. A situation, or a state of affairs, is a lexical reflection, by a particular language, of some chunk of the universe: an event; a state or a change of state; a process; a property or a relationship; etc. A situation is identified by its key word \( W \): honest, red, like \([= 'be like']\), spread \([= 'spreads (somewhere)']\), speak, destroy, etc. denote the situations 'somebody is honest', 'something is red', 'something is like something in some respect', etc. A situation can have several key words, which are then (exact or approximate) synonyms, converses, or syntactic derivatives of one another: honest — honestly — honesty; like — similar — similarity — as \([= 'red as blood']\); destroy — be destroyed — destruction.

Now, a situation has a number of participants: 'being honest/red' is a one-participant situation, 'love' is a two-participant one, and 'similarity' involves three participants \([= 'English is like Chinese in its morphology']\). An actant of the lexeme \( W \) is a linguistic entity that corresponds to one of the participants of the situation identified by \( W \). For any situation-identifying lexeme its semantic, deep-syntactic and surface-syntactic actants must be distinguished. The semantic actants of \( W \) are determined by the lexicographic definition of \( W \) and are (roughly) the variables we have to use in this definition. The surface-syntactic actants of \( W \) are actual phrases denoting its semantic actants and filling such surface-syntactic roles as (grammatical) subject, objects and complements. The deep-syntactic actants of \( W \) are generalized representations of its surface-syntactic actants that ignore all the details of the observable surface behavior of the latter. The deep syntactic actants are geared rather to meaning and can be construed as 'places', or arguments, of the predicate expressed by \( W \). All the more so that the different deep-syntactic actants of the same lexeme are numbered: the first deep-syntactic actant of \( W \), the second deep-syntactic actant of \( W \), etc.

Now, the first deep-syntactic actant corresponds to the grammatical (surface) subject or to any of its transforms (John \( \frac{1}{\text{writes}} \), John's \( \frac{1}{\text{writing}} \), writings \( \frac{1}{\text{(of) John}} \)), the second to the surface direct object or to a strongly governed indirect object of an intransitive verb, etc. To take a simple example, include, contain or inclusion have as their first deep-syntactic actant the name of the including set, and as their second deep-syntactic actant, the name of the included subset or element: Chapter 3 \( \frac{1}{\text{includes}} \frac{2}{\text{a section on fruit flies}} \); opinion has three deep-syntactic actants:
his opinion of Joan as a brilliant researcher (because of He thinks of Joan that she is a brilliant researcher); etc.

The deep-syntactic actants of the same lexeme are distinguished by Arabic numerical subscripts. Therefore, it is important to constantly keep in mind that numerical subscripts accompanying the names of LF's in what follows ALWAYS REFER TO DEEP-SYNTACTIC ACTANTS of the corresponding lexeme (i.e., in the final account, to participants of the situation denoted; but remember that the same participant of a given situation can match any deep-syntactic actant depending on the lexeme in question. So with include, the including set is the first deep-syntactic actant, while with belong the including set will be the second actant.)

IV
List of Lexical Functions

A lexical function, notated by a boldfaced Latin abbreviation, is written before the parentheses enclosing its argument (a familiar mathematical functional notation). For example:

<table>
<thead>
<tr>
<th>Function</th>
<th>Argument</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Magn</td>
<td>(shave)</td>
<td>close, clean</td>
</tr>
<tr>
<td>Magn</td>
<td>(easy)</td>
<td>very, extremely, ..., as pie</td>
</tr>
<tr>
<td>Magn</td>
<td>(scoundrel)</td>
<td>unmitigated, of the first water</td>
</tr>
<tr>
<td>Magn</td>
<td>(condemn)</td>
<td>strongly, in strongest terms</td>
</tr>
<tr>
<td>Magn</td>
<td>(cold)</td>
<td>very, terribly, ..., enough to freeze balls of a brass monkey</td>
</tr>
</tbody>
</table>

[Magn is from Lat. magnus 'great, big'].

As stated above, LF's are one of the central notions to a new type of dictionary, the Explanatory-Combinatorial Dictionary. The simple standard lexical functions will be listed below in the order in which they appear in a lexical entry in such a dictionary.

1. Syn - synonym; Syn₀, Synᵥ, and Synₐ designate, respectively, synonyms with broader, with narrower, and with intersecting meanings. (Symbols₀,ᵥ, andₐ have the same meaning when used with Conv, Anti and other LF's.) Examples: Syn(streljac 'shoot') = palit 'fire'; Synᵥ(streljac 'shoot') = obstreljavat 'fire upon; shell; machine-gun'; etc.

2. Convᵢⱼ - conversive, i.e., a lexical item with the same meaning as the key word W but with deep-syntactic actants i and j permuted: Convᵢⱼ(vključat 'include') = prinadležat 'belong to [a set]; Conv²³₁₂(mnene 'opinion') = reputacija 'reputation'. ['Reputation', in contrast to 'opinion', is necessarily held by several people: this is why it is a narrower conversive.]

3. Anti - antonym: Anti(pobeda 'victory') = poraženje 'defeat'.


4. Gener - generic concept such that 'Gener + W' = 'W' (where W is the key word): Gener(gas 'gas') = veseostvo 'substance' [cf. gasobraznoe veseostvo 'gaslike substance' = gas 'gas'].

5. Figur - standard metaphor for W: Figur(blokada 'blockade') = kol'co, lit. 'ring' [kol'co blokady 'the grip of a blockade; a siege']; Figur(tuman 'fog') = pelena 'curtain' [pelena tumana 'curtain of fog'].

6-9. S0, A0, Adv0, V0 - syntactic derivatives of W, that is, noun (= substantival), adjective, adverb, and verb, respectively, which have the same meaning as W. Examples: S0(streljak 'shoot') = strel'ba 'shooting'; A0(streljak 'shoot') = strelkovyi 'shooting [attrib.]; etc.

10. S1 - standard name of the 1-th participant in the situation described by W: S1(učit' 'teach') = učitel' 'teacher'; S2(učit' 'teach') = učenik 'pupil'; S3(učit' 'teach') = (učebný predmet 'subject') matter [as in high school].

11-15. S_instr, S_med, S_mod, S_loc, S_res - standard name of instrument, means, mode, location, and result of the situation described by W: S_instr(streljak 'shoot') = ogneostrelovoe oružie 'firearm'; S_med(streljak 'shoot') = boeprispasy 'ammunition'; S_mod(rassmatrivat' 'consider') = vsgljad [na što-l.] 'a view [of something]; podvod [k žemi-l.] 'approach [to something]; S_loc(uražat'sja 'fight' [as of two armies]) = pole bitvy/boja 'battlefield'; S_res(učit'sja 'learn') = navyki 'skills', znanija 'knowledge'.

16. Sing - 'one instance/unit of': Sing(gorov 'peas') = gorov'sina 'pea'; Sing(celovat' 'kiss') = pocelovat' 'give a kiss'.

17. Mult - 'aggregate of': Mult(korabl' 'ship') = flot 'fleets'; Mult(student 'student') = studentstvo 'student body'.

18. Cap - 'head of': Cap(universitet 'university') = rektor 'president'; Cap(fakul'tet 'faculty, school') = dekan 'dean'.

19. Equip - 'staff/crew of': Equip(teatr 'theatre') = truppa 'troupe'; Equip(bol'nica 'hospital') = personal 'personnel'; Equip(brakl 'marriage') = suprugi 'spouses'.

20. Centr - 'center/culmination of': Centr(les 'forest') = čašča [lesa] 'the thick [of the forest]; Centr(slava 'glory') = verśina [slavy] 'summit [of glory]; Centr(bor'ba 'struggle') = apogej [bor'by] 'climax [of struggle]. Centr is current in combination with Locin (see below): Locin Centr(pustynja 'desert') = v serdce [pustyni] 'in the heart [of the desert]'; Locin Centr(doroga 'road') = posredi [dorogi] 'in the middle [of the road].

21. A4 - determining property of the 4-th participant of a situation characterizing him according to his role in the situation: A4(gnev 'anger') = v [gneve] 'in [anger]', rasgnovanyj 'angry'; A4(slen'ya 'tears') = v [slenax] 'in [tears]'; A4(skorost 'speed') = so [skorost'ju...] 'with a speed of...' [compare spusk s
takoj skorost'ju 'the descent with such a speed'; A_2 (streljat' 'shoot') = pod obstrelom 'under fire'.

22. Able_1 - determining property of the i-th potential participant of a situation ("such that it can... easily' /'such that it can be... easily"): Able_1 (plakat' 'cry') = slesliviij 'tearful'; Able_2 (somnevats'ja 'doubt') = sommitel'nyj 'doubtful'.

23. Magn - 'very', 'to a (very) high degree': Magn (temperatura 'temperature') = vysokaja 'high'; Magn (nassmatrivaet' 'examine') = vnimatel'no 'attentively', pristal'no 'fixedly, intently'.

24-25. Plus, Minus - respectively, 'more' or 'less' (or 'to a greater/lesser extent') [than something else].

26-27. Plus refl, Minus refl - indicate that the comparison is made with a former state of the same object:

   IncenPredPlus (temperature) = povyšat'sja 'rise, increase' [for the LP's Incen and Pred, see below].

28. Ver - 'as it should be' [meeting intended requirements]: Ver (udivlenie 'surprise') = iskrennee 'sincere', nepodeliteľnoe 'unfeigned'; Ver (sosud 'container') = celuy 'whole', germetychnyj 'hermetical', leak-proof'; Ver (pribor 'instrument') = točnyj 'precise'.

29. Bon - 'good' [a standard praise for W]: Bon (rezat' 'cut') = akkuratno 'neatly, cleanly'; Bon (sudnol 'ship') = komfortabel'noe 'comfortable'.

30. Pos_1 - a standard praise of one of the participants of the situation denoted by W [but not of the situation itself]: Pos_2 (recenzija 'review') = polozitel'naja 'positive', while Bon (recenzija 'review') = xorosaja 'good', samečatel'naja 'excellent', ..., bljestjačaja 'brilliant'; however, a bljestjačaja recenzija 'a brilliant (ly written) review' may well be polozitel'naja 'positive' or otricitel'naja 'negative', that is, Pos_2 or AntiPos_2.

NB: The LP's Magn, Ver, Bon and Pos_1 are often combined with Anti. Thus, for instance, Magn (temperatura 'temperature') = vysokaja 'high', and AntiMagn (temperatura 'temperature') = niskaja 'low'; Pos_2 (imenie 'opinion') = polozitel'noe 'positive', and AntiPos_2 (imenie 'opinion') = otricitel'noe 'negative'.

31. Adv_1 - determining property of an action by the i-th participant of a situation according to his role in the situation: Adv_1 (žalzy 'tears') = so [slezami] 'with [tears]'; Adv_2 (skorost' 'speed') = so skorost'ju... 'at a speed of...' [cf. mčat'sja so skorost'ju... 'tear along at a ... speed']; Adv_2 (somnevats'ja 'doubt') = vrjad li 'hardly'.

32-34. Loc_in, Loc_1ab, Loc_2ad - preposition governing W [= the name of the situation] and designating a type of localization in space with the respective meaning - position, moving away, moving toward. Examples: Loc_in (vysota 'height') = na [vysote] 'at [a height]'; Loc_1ab (vysota 'height') = na [vysotu] 'to [a height]'; Loc_2ad (vysota 'height') = s [vysoty] 'from [a height]'.
35. \text{Loc}^{\text{temp}}_{\text{in}} - a preposition [analogous to \text{Loc}] with the meaning of temporal location: \text{Loc}^{\text{temp}}_{\text{in}}(\text{arest 'arrest'}) = \text{pri [arest] 'while being [arrested]'}; \text{Loc}^{\text{temp}}_{\text{in}}(\text{analis 'analysis'}) = \text{v xode [analiza] 'in the course of [analysis]'}.

36. \text{Instr} - a preposition [analogous to \text{Loc}] with the meaning of instrumentality: \text{Instr}(\text{pistolei 'pistol'}) = \text{iz [pistolei, lit. 'with [a pistol]'}; \text{Instr}(\text{masinka 'typewriter'}) = \text{na [masinky] 'on [a typewriter]'}.

37. \text{Propt} - a preposition with the meaning 'because of', 'as the result of': \text{Propt}(\text{stracz 'fear'}) = \text{ot [stracz, so [straxu] 'from [fear]'}; \text{Propt}(\text{ljubov 'love'}) = \text{iz [ljubvi k... 'because of [one's love of...']}'; \text{Propt}(\text{opyt I.I 'experience'}) = \text{na [svoem opyte] 'from [one's own experience]'}.

38. \text{Copul} - a copula: \text{Copul}(\text{učitel 'teacher'}) = \text{byt'}, \text{rabotat'} ['učitel'm 'be, work as [a teacher]'}; \text{Copul}(\text{primer 'example'}) = \text{byt'}, \text{javljat'sja, služit'} ['primerom] 'be, represent, serve as [an example]'.

39. \text{Pred} - a verb meaning 'be \text{W}', i.e. semantically covering the syntactic combination of a \text{Copul}(\text{W}) with \text{W}. Thus \text{Pred} is nothing but a 'fused' expression of \text{Copul}(\text{W} + \text{W} (on 'fused' expressions see below) needed for the convenience of some synonymic transformations. For example, \text{Pred}(\text{pjanica 'drunkard'}) = \text{pjanstvovat' 'drink'}, \text{Pred}(\text{rjadom 'next to'}) = \text{sosedstvovat' 'to (to) neighbor'}. The next three LF's are verbs which are semantically empty in the context of the entry lexeme (= their key word) and which serve to link, on the syntactic level, the name of a participant of a situation to \text{W} - the name of the situation itself. They play important semantico-syntactic roles and can be loosely called semi- auxiliaries.

40. \text{Oper}}_{1} - the first deep actant (and the surface subject) of this verb is the i-th participant of the situation, and the second deep actant (or the first surface object) is \text{W} (further actants, if any, designate further participants of the situation): \text{Oper}_{1}(\text{šlezy 'tears'} = \text{lit'}, \text{prolivat' 'shed'}; \text{Oper}_{1}(\text{arest 'arrest'}) = \text{proizvodit' 'make [an arrest]'}; \text{Oper}_{2}(\text{arest 'arrest'}) = \text{popadat' [pod arest] 'fall [under arrest]'}; \text{Oper}_{2}(\text{arest 'arrest'}) = \text{podvergat'sja [arestu] 'undergo [arrest]'}; \text{Oper}_{2}(\text{sprotivlenie 'resistance'} = \text{okasyvat' 'show, put up'}; \text{Oper}_{2}(\text{sprotivlenie 'resistance'}) = \text{vstrečat' 'meet, natačivat'sja [na sprotivlenen] 'run [into resistance]'}.

41. \text{Func}^{\text{1}} - the first deep actant (and the surface subject) of this verb is \text{W} - the name of the situation, and the second deep actant (and the first surface object), the i-th participant of the situation: \text{Func}^{\text{1}}(\text{udimiento 'surprise, astonishment'}) = \text{omvatyt'}, lit. 'seize' [i.e. the person is overcome by surprise, astonishment], \text{Func}^{\text{2}}(\text{temperatura 'temperature'}) = \text{равнит'sja 'be equal to'}; \text{Func}^{\text{1}}(\text{predloženie 'proposal'}) = \text{izxodit' [ot kogo-1.] 'stem from, come from [someone]'}; \text{Func}^{\text{2}}(\text{predloženie 'proposal'}) = \text{kaat'sja [čego-1.] 'concern [something]'. If there is no complement at
all, i.e. \text{Func} is an intransitive verb, the subscript \(0\) is used: \(\text{Func}_0(\text{dos']rain'}) = \text{idti}, \text{lit.'walk' [cf. Engl. fall]}\).

42. \text{Labor}_{ij} - the first deep actant (and the surface subject) of the verb is the \(i\)-th participant of the situation, the second deep actant, the \(j\)-th participant of the situation, and the third deep actant (implemented by the second surface object) is \(W\) itself. Examples: \text{Labor}_{12}(\text{dopros 'interrogation'} = \text{podvergat'} [\text{kogo-1. doprosu}] 'subject [someone to an interrogation]'; \text{Labor}_{32}(\text{arenda 'lease'} = \text{sdavat'} [\text{eto-1. v arendu}] 'grant [something on lease]').

\text{Oper}, \text{Func}, \text{and Labor} can be paired in converse relations, that is, \text{Oper}_{1} = \text{Conv}_{21}(\text{Func}_{1}); \text{Labor}_{12} = \text{Conv}_{132}(\text{Oper}_{1}), and so on. These relationships may be diagrammed as follows:

Here, a two-participant situation designated by key word \(W\) is presented. Arrows represent semi-auxiliary verbs; the arrow’s tail indicates the surface (grammatical) subject, the head pointing toward the first surface object.

A different way to express the same idea is by using a matrix:

<table>
<thead>
<tr>
<th>Surface syntactic roles verbs</th>
<th>Surface subject</th>
<th>First surface object</th>
<th>Second surface object</th>
</tr>
</thead>
<tbody>
<tr>
<td>\text{Oper}_{1/2}</td>
<td>1st/2nd deep actant</td>
<td>(W)</td>
<td>-</td>
</tr>
<tr>
<td>\text{Func}_{0/1/2}</td>
<td>(W)</td>
<td>none/1st/2nd deep actant</td>
<td>-</td>
</tr>
<tr>
<td>\text{Labor}_{12/21}</td>
<td>1st/2nd deep actant</td>
<td>2nd/1st deep actant</td>
<td>(W)</td>
</tr>
</tbody>
</table>
E.g., for 'A changes B':

Now we return to the survey of LF's.

43. **Involve** - a verb that links a non-participant of a situation with the name of the situation acting on him:
   
   Conv1 Involve (veter 'wind') = stojat ['na vetru'] 'stand [in the wind]'; Incep Involve (metel 'snowstorm') = zastigat 'catch';
   
   A2 Involve (metel 'snowstorm') = v [meteli] 'in [a snowstorm]'.

   44-46. The following three LF's represent the meanings of what are often called phasal verbs: **Incep** - 'begin'; **Cont** - 'continue'; **Fin** - 'end, cease'. They are connected by obvious semantic relationships: Fin(P) = Incep(nonP); Cont(P) = nonFin(P) = nonIncep(nonP). Incep, Cont and Fin are used (at least in Russian) in combination with other LF's. Examples: Oper2 (vlast 'power') = naxodit'sja [pod vlast'ju... ] 'be [under the power of]',
   
   Incep Oper2 (vlast') = popadat' [pod vlast'... ] 'fall [under the power of]', Fin Oper2 (plast') = vyxo dit' [iz-pod vlasti] 'get out [from under the power of]'; Cont Oper1 (vlianje 'influence') = sor ranjat 'maintain', Cont Oper2 (vlianje) = ostavat'sja [pod vlija niem... ] 'remain [under the influence of]'; Cont Func0 (zapax 'odor') = derskat'sja 'linger'.

47. **Caus** - 'cause', 'do something so that a situation occurs'. Caus is often used in combination with other verbal LF's. Examples: Caus Oper1 (menie 'opinion') = privodit' [kogo-1. k menjju] 'lead [someone to an opinion]'; Caus Func1 (nadezda 'hope') = vseljat', vdoxnut' [nadezdu v kogo-1. ] 'raise [hope in someone], inspire [someone with hope]; Caus Oper2 (obed 'dinner') = gotovit' [sto-1. na obed] 'prepare [something for dinner]'; Caus Func0 (obed) = gotovit', strjapat' [obed] 'make, cook [the dinner]'.

   With the LF Caus, the LF's Perm and Ligu are naturally associated, Ligu(P) = Caus (nonP), Perm(P) = non Ligu(P) = non Caus (nonP). Both Perm and Ligu are usually used in combination with other verbal LF's.

48. Perm - 'permit', 'allow': non Perm Oper2 (kritica 'criticism') = ogranadat' [kogo-1. ot kritiki] 'protect [someone from criticism]'; Perm Oper2 (eksamen 'exam') = dopuskat' [kogo-1. k eksameniu] 'allow [someone to (take) an exam]'.

\[ \text{IZMENENIE} \ 
\text{Proissodit} \ 
\text{podvergaet B \ 'submits B to'} \ 
\text{A} \]
49. *Liqui* - 'liquidate', 'do something so that a situation

   does not occur or stops occurring': *Liquifunca* (negramotnost' 'il-

   literacy') = *pokonšit* [š negramotnost'ju] 'wipe out illiteracy';

   *Liquifunca* (kostēr 'campfire') = *potušit* [kostēr] 'extinguish [a

   campfire]'.

50–52. Now let us look at another triple of interrelated

   LF's: *Real*, *Fact*, and *Labreal*.

   The LF's *Real*, *Fact*, and *Labreal* are syntactically analogus
gous to the functions *Oper*, *Func*, and *Labor*, respectively. This
means that the names of the situation and of its i-th participant
fulfill with respect, e.g., to *Real* the same syntactic roles as
they do with respect to *Oper*; etc. However, unlike the 'empty'
LF's *Oper*, *Func*, and *Labor*, the lexical functions which we are
now concerned with correspond to a specific meaning - 'fulfill a
demand or requirement of...'. The demands can differ for differ-
ent W's. For example, the fulfillment, or realization, of a hypo-
thesis is its confirmation; therefore, *Real* (hipoteza 'hypothesis')

   = podtvēršdat 'confirm' [Fakty podtvēršdat hipotezu 'The facts

   confirm the hypothesis'], and *Fact* (hipoteza 'hypothesis') = soot-

   vetstovvat 'be in accordance with' [Hipoteza sovetostrvet fakta

   'The hypothesis is in accordance with the facts']. Realization of
an artifact is its utilization according to its intended function;
therefore, *Fact* (nož 'knife') = *resat* 'cut' [štot nož rešet xoros-

   so 'This knife cuts well'].

   Further examples:

   *Real* (obvinienie 'accusation') = *dokazovat* [obvinienie] 'prove

   [an accusation]'; *Real* (učebnoe zavedenie 'educational institu-

   tion') = *prepodavat* [v učebnom zavedenii] 'teach [in an educa-

   tional institution]'.

   *Real* (obvinienie) = *soglašat'sja [s obvinieniem] 'agree [with an

   accusation]'; *Real* (učebnoe zavedenie) = *isuvat'sja, prepodavat'-

   sja [v učebnom zavedenii] 'be studied, be taught [in an educational

   institution]' (while *Real* (učebnoe zavedenie) = *udit'sja [v učeb-

   nom zavedenii] 'study [in an educational institution]'); *Real* (so-

   blasm 'temptation') = *poddavat'sja [soblaznu] 'yield [to tempta-

   tion]'.

   *Fact* (samenia 'doubt') = podtvēršdat'sja 'be corroborated', o-

   pravdylvat'sja 'prove justified'; *Fact* (nadežda 'hope') = *štyvat'sja

   'come true'; *Fact* (šeduš 'ship') = *plyt' 'sail'.

   *Fact* (ocered 'turn') = *byt' [za kem-i.] 'be [someone's (turn)]'

   [Ocered' za vami 'It's your turn']; *Fact* (škryptment 'experiment')

   = udavat'sja [komi-i.] 'work out [for someone]'.

   *Fact* (šeduš 'ship') = *vesti, perevozit' [grusy, passażirov]

   'convey, transport [cargo, passengers]'; *Fact* (sosud 'container')

   = *soderžat* [čto-1.] 'contain [something]'.

   *Labreal* (visilica 'gallows') = *uzdarmut' [kogo-1. na visilicu]

   'string up [someone on a gallows]'; *Labreal* (obed 'dinner') = *est'

   [čto-1. na obed] 'eat [something for dinner]'.

The LF's *Real*, *Fact*, and *Labreal* are superscripted with Ro-

man numerals to indicate the degree of the realization or fulfill-

ment: the superscript I means fulfillment only at the psychologi-

cal level, with the superscript II meaning fulfillment at the
physical level, cf. \text{Real}_2 \text{(priglašenje 'invitation')} = \text{prinimat' 'accept'}, while \text{Real}_2 \text{(priglašenje)} = \text{sledovat' 'follow'}; or \text{Fact}_1 \text{(uživo 'emotion')} = \text{govorit', podskazyvat' 'tell'}, while \text{Fact}_1 \text{(uživo)} = \text{zastavljet' 'force'}.

53. \text{Manif} - 'manifest itself', 'become apparent':
\text{Manif (vina 'guilt', 'fault')} = \text{obnaruživati'sja 'become apparent'};
\text{Manif (uživenie 'amazement', glaza 'eyes')} = [U nego] glaza na lob polezli [His] eyes started from his face;
\text{Sump (uživenie 'amazement', rot 'mouth')} = raznut' rot 'open [one's] mouth wide';
\text{Sump (strach 'fear', volosy 'hair')} = [U nego] volosy vstali dybom [His] hair stood on end'.

54. \text{Prepar} - 'prepare', 'get (something) ready for normal use or functioning': \text{Prepar\_Fact}_1 \text{(revolver 'gun')} = \text{sarjakat' 'load'}; \text{Prepar\_Fact}_2 \text{(revolver)} = \text{vyxodit' kurok 'raise the cock'}.

55. \text{Prox} - 'be about to / on the verge of': \text{Prox\_Oper}_1 \text{(otkajanie 'despair')} = \text{byt' na grani [otkajaniya] 'be on the edge of [despair']}; \text{Prox\_Func}_1 \text{(groza 'thunderstorm')} = \text{sbrata'sja 'gather, brew'}.

56. \text{Degrad} - 'degrade', 'become worse or bad': \text{Degrad (moloko 'milk')} = \text{skisnut' 'go sour'}; \text{Degrad (mjaso 'meat')} = \text{ispornit'sja, protuznut' 'go bad'}; \text{Degrad (disciplina 'discipline')} = \text{rassta't'sja 'decay'}.

57. \text{Son} - 'emit characteristic sound': \text{Son (obaka 'dog')} = \text{la'jak 'bark'}; \text{Son (banknoty 'banknotes')} = \text{xrustet' 'rustle'};
\text{Son (sneg 'snow')} = \text{skripet' 'crunch'}; \text{Son (vodopad 'waterfall')} = \text{revet' 'roar'}.

58. \text{Imper} - 'do W!': \text{Imper (streljat' 'shoot')} = \text{ogon' 'fire'}; \text{Imper (brat' oruzie 'seize arms')} = \nu \text{muž' 'take up arms'}; \text{Imper (govorit' tixo 'speak lowly')} = \text{te-so s' , tē ḫ̈ 'sh-h-h'}; \text{Imper (brat' 'take')} = \text{na(te)' 'take it', 'here'}.

59. \text{Perf} - 'perfective', i.e. 'have the process carried through to its natural limit': \text{Perf (vstavat' 'be standing up')} = \text{vstat' 'have stood up'}; \text{Perf (rešat' 'be solving [a problem])} = \text{rešit' 'have [it] solved'}.

60. \text{Imperf} - 'imperfective', i.e. 'be carrying out the process': \text{Imperf (vstat') = vstavat'}; \text{Imperf (rešat') = rešat'}. 

61. \text{Result} - 'resultative', i.e. 'the state of affairs that normally results from the completion of the process': \text{Result (pokupat' 'buy')} = \text{imet' 'have'}; \text{Result (ložit'sja 'lay down')} =
ležat' 'be lying'; Result(naučit'sja 'have learnt') = umet' 'know', 'have necessary skills'.

V
Concluding Remarks

Along with the LF's listed, two further types of LF's are extensively used in the Explanatory Combinatorial Dictionary: non-standard and compound LF's.

A non-standard LF is a meaning that is idiomatically expressed depending on a key word but has either a strongly limited semantic combinatoriality or a fairly limited range of expressions, or both. In other words, it is too specific, too particular to be granted the status of a standard LF. Non-standard LF's are written in standardized natural language. Some examples:

Such that Y is confined to his home(arrest 'arrest') = domašnij 'house-' [arrest];

Such that it is the result of a loss at cards that was not immediately paid(dolg 'debt') = kartočnyj 'card' [dolg], obsolete [dolg] čestí '[debt] of honor';

During a short time and/or non-intensively (with the purpose of knowing Y somewhat better)(učit'sja 'learn') = // podučit'sja 'learn a bit (of something)'.

A compound LF is a combination of syntactically related simple LF's that has a unique lexical expression covering the meaning of the combination as a whole. I have presented numerous examples of compound LF's above; let me give some more illustrations, with the key word printed in boldface:

AntiMagn

امید کردن 'thin (lit. 'liquid') applause', slabye dovody 'weak arguments', niskaja temperatura 'low temperature', nežnacitel'nye poteri 'negligible losses',...

AntiVer

ložnyj stvd 'false shame', tšivoе obeschčanie 'false (lit. 'lying') promise', ošibochnoe predstavlenie 'a wrong conception', bezosnovatel'nye opasenia 'unfounded misgivings/fears',...

IncepOper

priobretat' populjarnost' 'acquire popularity', vpadať v otčajanie 'sink into despair', vstavat' na put' predatel'stva 'take the path of treason', perexodit' v pike 'go into a dive [as of an aircraft]',...

CausOper

sdavat' v eksploataciju 'put into operation', uvergat' v rabstvo 'plunge into slavery', stavit' pod kontrol' 'put under control',...

AntiReal

prováilit'sja na zkazane 'fail an examination', otklonjat' sovet 'reject a piece of advice', oklujet' xodujastiwo 'turn down an application'.
The following four remarks bearing on all LF's are in order:

1) An LF may have a fused expression, i.e. a lexical unit that does not include the key word but covers both the meaning of the function itself and that of its argument (= key word). The fusion is shown by the symbol // separating all the fused values (on its right) from all the non-fused ones. For example:

\[
\text{Migr (dodč 'rain') = prolivnom 'heavy' // liven 'shower' [i.e. liven = prolivnom dodč]; cf. Engl. downpour = heavy rain.}
\]

\[
\text{Migr (vkusno 'delicious') = oden 'very' // pal'čiki obližek', lit. 'You'll lick your fingers' [pal'čiki obližek' = oden vkusno]; and the like.}
\]

2) Several LF's having simultaneously the same key word but syntactically not related one to another may be expressed by one lexical unit covering the meanings of all the LF's involved. This is what we call a configuration of LF's (as opposed to compound LF's, in which all the constituent simple LF's are syntactically related). In a configuration of LF's, the '+' sign is used to separate the constituents. For example, in the entry SUD'BA I 'fate, destiny' the notation

\[
\text{Fact}^3_{\text{II}} + \text{AntiBon}_2: \text{presledovat' 'persecute'}
\]

means that fate really affects [= Fact$^3_{\text{II}}$] the person in question (the defendant, so to speak) and its verdict is bad [= AntiBon$_2$].

Two further examples:

\[
\text{A$_1$ (vosxiščenje 'delight') + Migr (vosxiščenje) = preispolnennyy [v vosxiščenija] 'full [of delight]';}
\]

\[
\text{Omer$_1$ (otčajanie 'despair') + Migr (otčajanie) = byt' vo vlasti [otčajaniya] 'be completely in [despair]'.}
\]

3) Some LF's (most often, Migr or Real, with the latter's relatives, Fact$_1$ and Labreal$_{ij}$) may be subscripted with a semantic component of the key-word's definition (in square brackets) to indicate that the meaning of this LF interacts with exactly this component of the key-word's meaning. Thus:

\[
\text{Labreal}_{12} [xranit' 'keep'] (pamjat' 'computer memory') = xranit' [v pamjati] 'store [in memory]';
\]

\[
\text{Labreal}_{12} [vydavat' 'output'] (pamjat') = izvlekat' [iz pamjati] 'extract [from the memory]';
\]

\[
\text{Migr} [bojat'ja 'be afraid'] (strax$_1$ 'fear') = dikiy 'wild', štukiy 'terrible';
\]

\[
\text{Migr} [terjat' samokontrol'...'lose...self-control'] (strax$_1$) = panicleskij 'panic [adj.]', životný 'animal';
\]

\[
\text{AntiVer} [porazat' 'hit'] (streljat' 'shoot') = ploxo 'badly', skverno 'poorly',...}
\]
AntiVer\[\text{'cel', 'target'}\](\text{streljat}^1) = v \text{ vozdux 'into the air'.}

4) Furthermore, some LF's may be superscripted with semantic labels, like 'usual', 'loc(ation)', 'temp(oral)', 'quant(itative)', to make their meaning more precise:

\[\text{Magn}^{\text{temp}}(\text{opyt 'experience'}) = d\text{litel'nyj 'long'};\]
\[\text{Magn}^{\text{quant}}(\text{opyt 'experience'}) = b\text{o}l'\text{yoj 'considerable'}.

In concluding this survey of LF's, I would like once again to call attention to the fact that they are used for two main purposes:
- for the description of idiomatic or restricted lexical cooccurrence or derivation relations;
- and for specifying universal synonymic transformations of utterances on the deep-syntactic level. It is clear, for example, that, in any language, (1) holds:

\[(1) \text{W} = \text{Oper}_1 \rightarrow \text{St}_2 (\text{W}) = \text{Oper}_2 \rightarrow \text{St}_2 (\text{W}) = \text{Func}_1 \rightarrow \text{St}_2 (\text{W}) = \text{Func}_2 \rightarrow \text{St}_2 (\text{W}) = \text{Labor}_1 \rightarrow \text{St}_2 (\text{W}); \text{and so on.}\]

[Numbers on the arrows stand for deep-syntactic actants: e.g., the key word of an \text{Oper} is its second deep actant, etc.]

The transformations presented in (1) can be exemplified as follows:

\[(2) \text{vio\text{jat}'(to) influence' = okazyv\text{at}' [=\text{Oper}_1] vlijanie} \]
\[= \text{St}_2] \text{'have influence' = by\text{t}', nasodit'sja} [= \text{Oper}_2]
\[\text{pod vlijaniem 'be under influence', etc.}\]

Compare (3):

\[(3) \text{Ivan duerno vlijajet na Petra 'Ivan influences Peter in a bad manner' = Ivan okazyv\text{at} na Petra d\text{urnoe vlijanie 'Ivan has a bad influence on Peter' = P\text{è}tr nasoditsja pod d\text{urnym vlijaniem Ivana 'Peter is under the bad influence of Ivan.}\]

Rules of type (1) allow one to establish a paraphrasing system for synonymic transformations of sentences and/or discourses.

Such a system can automatically produce, for any given text, a set of its synonymous or nearly-synonymous paraphrases. It also can automatically derive, for a set of synonymous texts, a canonical invariant. This indicates yet another promising direction for the practical use of LF's, namely in the domain of automatic text processing. (For more details about paraphrasing systems using LF's as a main tool see e.g., Žolkovskij - Mel'čuk 1967, Mel'čuk - Žolkovskij 1970, or Mel'čuk 1974: 190-206.)
Acknowledgments

I owe so much to Alec Zholkovsky and Jurij Apresjan, with whom I have been working on lexical functions in the Explanatory Combinatorial Dictionary of Contemporary Russian for more than ten years, that it is difficult to properly express my indebtedness to them. But speaking only of people whose participation has been decisive for the present text, I must express my heartfelt gratitude to (in alphabetical order) Cliff Goddard, Joe Grimes, Lidija Iordanskaja and Johanna Nichols. Nobody but myself is to blame for shortcomings and bugs that have managed to survive their scrutiny.

Notes

1. If there is no explicit indication to the contrary. In some cases, mentioned in Iordanskaja et al.'s article, this volume, actantial subscripts may refer to deep-syntactic actants of a new situation, rather than directly to deep-syntactic actants of the key word.

2. Interesting examples of values of the LF Mult in English can be found, most unexpectedly, in The Book of Lists, by D. Wallechinsky, I. Wallace and A. Wallace, New York: W. Morrow, 1977, page 135:

A murder of crows  A rag of colts
clowder cats drift hogs
leap leopards charm finches
sloth bears trip goats
raffer turkey knot toads
smack jellyfish parliament owls
skulk foxes troop kangaroos
labor moles gaggle geese
crash rhinoceros pride lions
siege herons muster peacocks

"Although not frequently heard in conversation, these terms are fully correct and appropriate ways of describing the animal listed" (ibidem).

References


