К ПРОБЛЕМЕ ОПИСАНИЯ ПРИЛАГАТЕЛЬНЫХ-ИНТЕНСИФИКАТОРОВ ДЛЯ ЗАДАЧ ОБРАБОТКИ ТЕКСТА

Миличевич Я. (jmilicev@dal.ca)
Университет Далхаузи, Галифакс, Канада

Тимошенко С. П. (timoshenko@iitp.ru)
Институт проблем передачи информации
РАН, Москва, Россия

Ключевые слова: лексические функции, интенсификация, модель «Смысл — Текст», толково-комбинаторный словарь, лексическая функция Magn, лингвистический процессор ЭТАП-3, русский, английский

TOWARDS A FINE-GRAINED DESCRIPTION OF INTENSIFYING ADJECTIVES FOR TEXT PROCESSING

Milichevich J. (jmilicev@dal.ca)
Dalhousie University, Halifax, Canada

Timoshenko S. (timoshenko@iitp.ru)
Institute for Information Transmission Problems,
Russian Academy of Sciences, Moscow, Russia

We address collocations of the type "*Intensifying Adjective* + *NOUN*", such as *heavy RAIN* and *complete DISAGREEMENT*, known as Magn type collocations. Such a collocation can be represented as a functional dependency: Magn(*RAIN*) = *heavy*, where Magn is a (lexical) function responsible for the meaning 'very'/'high degree', and *heavy* the value that Magn has with *RAIN*, its keyword. The formalism of lexical functions has proved its usefulness in various NLP tasks, but on close inspection its semantic granularity turns to be insufficient. We propose a refinement of the notion of Magn by distinguishing Magn's semantic subtypes. Our description, which proceeds from the assumption that a choice of a Magn type collocate is not arbitrary, takes into account the following factors:

- semantic class of the keyword (= its semantic label, corresponding to the generic semantic component of its definition) and/or its actants;
- semantic component(s) in the keyword's definition targeted by intensification;
- semantic contrasts observed among Magn type collocates of a given keyword.

We tested our approach on data from the Russian and English explanatory-combinatorial dictionaries developed for the multi-purpose language processing system ETAP-3. As our results show, Magn's semantic subtypes we have identified allow for the encoding of lexicograpahic information in a way that is not only precise but also has predictive power.

Keywords: intensification, lexical function Magn, linguistic processor ETAP-3, Meaning-Text linguistic theory, text processing, Explanatory combinatorial dictionary, collocations, English, Russian

1. The Problem Stated

This paper considers collocations of the type "*Intensifying Adjective + NOUN*", such as *heavy RAIN*, *high PRICE*, *broad DISCUSSION*, *radical REFORM*, etc. More precisely, we investigate the range of meanings that the adjective can carry in such combinations.

Within the Meaning-Text theory, or MTT (Mel'chuk 1974, 2012: 85–159; Kahane 2003), intensifying collocations are described in terms of the lexical function (Wanner, ed., 1996) Magn, a modifier with a very general meaning 'very'/'intense'/'big'; in the dictionary, they are listed (as all collocations) in the entries of their keywords, using the following notation: Magn(*RAIN*) = *heavy*, Magn(*PRICE*) = *high*, and so on.

The fact that the meaning of the lexical function (= LF) Magn is so general and its expression contingent on the keyword has two consequences. On the one hand, semantic contrasts are observed between formally identical elements of its value with different keywords:

(1) ***complete*** *DISAGREEMENT* [≈ 'in **all** aspects of the issue'] vs. ***complete*** *DESPERATION* [≈ 'very **intense**'] vs. ***complete*** *MYSTERY* [≈ 'such that nothing **at all** is known about']

On the other hand, elements of the value of Magn for a single keyword may be less than perfectly synonymous[1]; cf., for instance, the semantically contrasting pairs of boldfaced adjectives in (2a) and (2b), whose simultaneous use in text is not redundant[2]:

(2) a. *Now the House is obviously in* ***complete*** [≈ 'in **all** aspects of the issue'] *and* ***strong*** [strong' ≈ 'such that the views of the parties are **far** apart'] *DISAGREEMENT as to what those safeguards should be.*
b. *This consensus is particularly needed in periods of* ***wide*** [≈ 'having **many** targets'] *and* ***radical*** [≈ 'bringing about a **big** change'] *REFORMS.*

---

[1] Cf. Mel'chuk (1974: 110), translation is ours: "[...] different elements of the value of an LF do not have to be fully synonymous; it is sufficient that they have a common semantic core and that their semantic differences are not "too significant". For the time being, we are unable to give a general answer to the question of the range of allowable semantic differences—i.e., which keyword's collocates that exhibit differences in meaning can be considered as the expressions of the same LF and which ones must be taken to be expressing different LFs.

[2] Our examples of actual LF uses in texts come from Google, Yandex and British National Corpus.

In both cases in (2), meaning differences are due to the fact that the intensification targets different components of the definition of the keyword; more on this will be said later.

A more fine-grained description of intensification is needed for specific language processing tasks, such as machine translation or sense-disambiguation. We are particularly interested in machine translation, where distinguishing "subtypes" of Magn could facilitate lexical choice in cases where there is a mismatch between L_{SOURCE} and L_{TARGET}—i. e., if the intensifiers required/allowed by the keywords that are translational equivalents contrast semantically. We have in mind, for instance, "false friends" like the ones in (3), extracted from the English and Russian dictionaries of the linguistic processor ETAP-3.[3]

(3) Eng. ***hard*** *NEGOTIATIONS* [≈ 'such that the parties are discussing from the positions that are **very different**'] vs. Rus. ***trudnye*** *PEREGOVORY* lit. 'hard negotiations' [≈ 'such that last **long**, the parties sticking to their **very different** respective positions']

The Russian adjective is semantically more complex: it subsumes the meaning of the English one and an additional component, which is expressed in English by a distinct subtype of Magn. This analysis is corroborated by the following examples:

(4) a. $_{Eng}$. *They were* ***hard*** *negotiations, as well as being very* ***long***.
b. $_{Rus}$. $^{??}$*Posle desjati munut* ***trudnych*** *peregovorov, …*
'After ten minutes of hard negotiations, ...'

In English, *hard* and *long* can be simultaneously used with *NEGOTIATION* without any redundancy effect, as is the case in (4a). In Russian, however, using *trudnyj* to qualify negotiations of short duration, as in (4b), is dubious (only four examples have been found on Jandex for 'minutes of hard negotiations', as opposed to several hundreds for 'months of hard negotiations'). The right translation for *hard* in this context is *žostkij* 'tough' (rather than *trudnyj*).

In order to increase the level of granularity of the description and cover the cases such as those illustrated in (2) and (3) above, a technique consisting in indicating the semantic component (within the keyword's definition) targeted by intensification is used. Thus, for the adjectives in (3), we could use the description $\text{Magn}_{['difference']}$, respectively $\text{Magn}_{['difference' \& 'duration']}$. This has been a standard practice within the MTT, applied, for example, in the *Explanatory-Combinatorial Dictionaries* (= ECDs) for Russian (*Mel'chuk, Zholkovsky*, 1984) and French (*Mel'chuk et coll*. 1984–1988–1992–1999) and the *Dicet* lexicographic database for French (*Barque et al*., 2010, *Gader et al*.,

[3] ETAP-3 is a multifunctional NLP environment comprising several applications: a machine translation system, a module of synonymous paraphrasing of sentences, a tagger for syntactic annotation of text corpora, a Universal Networking Language interface, and a computer-assisted language learning tool. Since this environment is largely based on the Meaning-Text Theory, a strongly lexicalist approach to language, it uses as a core component of all the applications a dictionary of a particular type—called *Explanatory-Combinatorial Dictionary*, or ECD; *Apresjan et al*., 2003 offers a brief overview of the ETAP-3 system and the most relevant references.

2012). Here is the gist of the approach, abstracted from (Iordanskaja, Polguère 2005) and (Mel'chuk 2013: 213–215); examples of collocations and their encodings are ours.

Formally speaking, four cases of intensification can be distinguished, based on two independent and combinable parameters:

- The component 'σ' in the definition of the keyword targeted by the intensification: the *central component* (= CC), i. e., the component corresponding to the *genus* in Aristotelian terms, or a *peripheral component* (= PC), i. e., a specific difference component.
- The nature of the link between 'σ' and the intensifying semanteme: *direct* (the intensifying semanteme bears on 'σ') or *indirect* (the intensifying semanteme bears on a component that bears on 'σ'; in Mel'chuk's terms, it is external to the definition of the keyword).

Subtypes 1/2 are "pure" Magn, while subtypes 3/4 additionally contribute "a specific perspective on the situation denoted by the base (i. e., keyword—J. M, S. T.). Typical perspectives are: dimension, duration, quantity, way of doing (emotion, energy, …), etc." (Iordanskaja, Polguère, 2005: 182–183)". The corresponding semantic distinguishers are used as subscripts in the cases where a peripheral component is directly intensified and as superscripts in the cases of indirect intensification.

Table 1. Subtypes of Magn standardly used in the MTT

1) The CC is directly intensified Collocation Keyword's Actantial Structure Meaning of the collocation LF Notation	 radical < sweeping REFORM X's ~ of Y '**big** change$_{CC}$' Magn(REFORM)
2) A PC is directly intensified	 heavy RESPONSIBILITY X's ~ for Y 'X's duty to care for Y_{PC}, Y being **important**' $Magn_{['Y\ is\ important']}$(RESPONSIBILITY)
3) The CC is indirectly intensified	 long <protracted> NEGOTIATIONS X's ~ with Y over Z 'discussion$_{CC}$ whose duration is α, α being **big**' $Magn^{temp}$(NEGOTIATIONS)
4) A PC is indirectly intensified	 wide REFORM X's ~ of Y 'change of Y_{PC} whose number is α, α being **big**' $Magn_2^{quant}$(REFORM)

We think that type 2 Magn too can provide a different perspective on the keyword's meaning; compare, for instance, *heavy responsibility* ('such that Y is important') and *full responsibility* ('such that it rests only on X'), both intensifiers being of type 2. Accordingly, we will consider only type 1 intensifiers to be the "pure" Magn.

While we fully subscribe to the approach outlined above, we believe that it can be enhanced if we take into account another aspect of the keyword meaning, namely its semantic label (Milichevich 1995, Polguère 2003). The semantic label of the lexical unit L corresponds to the central, or generic, component in the definition of L. Semantic labels, such as act, action, communication, event, process, state, artifact, person, etc., are taxonomic characterizers that allow for a compact and formal description of the meaning of lexical units and organize the latter in hierarchical classes, such that each member of a class shares some relevant properties inherited form a higher class. Thus, since processes, states and actions (unlike acts and events) can be characterized for duration and phase, instances of the corresponding semantic labels will by default be characterizable in the same way. Semantic labels are akin to *aspectual classes* of (Vendler, 1967) and Apresjan's *fundamental classes of predicates* proposed in (Apresjan 2006: 75–109). We think that at least some semantic distinguishers used with the LF Magn should be predictable from the semantic label of its keyword and/or from the labels assigned to keyword's semantic actants. To give a concrete example, all instances of the semantic label process (e.g., UNIFICATION, ADJUSTMENT, ADAPTATION, ASSIMILATION) are potentially compatible with two varieties of Magn: $\text{Magn}_{['\text{duration}']}$, with the value *long* <*protracted*>, and $\text{Magn}_{['\text{phase}']}$, with the value *complete*. Conversely, we can predict the (sub-)meaning of the intensifying collocate from the semantic label of its keyword or keyword's actants; for example, *complete* can mean 'at the ultimate stage of' when intensifying $\text{L}_{\text{process}}$, it can mean 'concerning all aspects/elements of' when combining with $\text{L}_{\text{opinion}}$ ([BE IN] AGREEMENT/DISAGREEMENT, APPROVAL/DISAPPROVAL [OF ONE'S ACTIONS], etc.), and so on. While a process has duration and can be conceived as consisting of stages, an opinion is, roughly, information, and is conceived as consisting of elements (cf. *a* ***piece*** *of information*).

Let it be noted that not all labels allow for such inheritance of semantic properties: the higher in the hierarchy a label is, i.e., the more general its meaning, the stronger its predictive power. Nevertheless, the recourse to semantic labels should allow for some interesting generalization, leading to more systematic lexicographic descriptions[4]. Semantic labels were already used in this way in Reuther (1996) and (2003) for a fine-grained description of collocations involving support verbs of Oper_i and Func_i type.

In the next section, we expand on and illustrate our proposal.

---

[4] Another venue to explore when it comes to possible generalizations, mentioned in Iordanskaja & Polguère (2005: 184), are logical links between the structure of the definition and the type of intensification. For example, communications are not gradable, so that they admit only indirect intensification; some standard definition blocks (in definition templates), such as 'potential effect' in the definitions of lexemes denoting some natural phenomena, states, feelings, etc., can also be targeted by intensification (e.g., *devastating STORM, debilitating <life-altering> ILLNESS, petrifying <paralyzing> FRIGHT*, etc.)

2. Our Proposal

We start by briefly evoking the context of the research (2.1) and then proceed to its preliminary findings (2.2).

2.1. Data and Methodology

We used as our primary data source LF-descriptions consigned in the ECD-style English and Russian dictionaries of the ETAP-3 NLP system. These dictionaries make use of four Magn subtypes, encoded as MAGN, MAGN-NS, MAGN-NS1, MAGN-NS2, NS being an abbreviation for "non-standard" (Apresian & Cinman 2002: 117). NS-functions do not have any fast semantic meaning, they are used to fix English-Russian translational equivalents within the range of LF values of the single word. For instance, in the entry for OBSERVANCE, we find: MAGN: COMPLETE1, MAGN-NS: EXACT1, MAGN-NS-1: CLOSE2, because there is a one-to-one correspondence between *complete* OBSERVANCE and *polnoje* SOBL'UDENIJE, *exact* OBSERVANCE and *točnoje* SOBL'UDENIJE, *close* OBSERVANCE and *tščatel'noje* SOBL'UDENIJE. Currently, the number of entries featuring the intensifying LFs in the Russian ECD is as follows: 1841 (MAGN), 461 (MAGN-NS), 99 (MAGN-NS-1), and 42 (MAGN-NS-2). The situation in the English ECD is comparable: 1409 (MAGN), 471 (MAGN-NS), 110 (MAGN-NS-1), and 40 (MAGN-NS-2). The number of actual collocations encoded is much higher, because the intensifying LFs typically yield numerous elements of value for any given keyword. Thus, the CDs provided a wealth of data on which our proposal could be tried out.

As mentioned above, we started from the fact that the semantic label of the lexical unit L is correlated to:

- Possible LFs that L can accept, in our case—subtypes of one specific LF, namely Magn (and, to much lesser extent, the complex LF AntiMagn);
- Possible elements of value for these LFs (or subtypes of a particular LF)[5].

Working with data from the ETAP-3 ECDS, we used as input values of the intensifying LFs (rather than their keywords) and tried to determine whether it is possible to group keywords semantically. For example, consider the adjective širokij 'wide'. As an LF value, it can be translated into English as WIDE, BROAD or LARGE. We extracted from the ETAP-3 dictionaries all the keywords that have these values for Magn, as illustrated in the table below, and tried to find some commonalities in their semantics.

Proceeding from the collocate SHIROKIJ and its translation equivalents, we found groups of semantically similar keywords. For every group we could formulate a refined meaning of Magn expressed by the adjective SHIROKIJ.

---

[5] Cf. Apresjan (2009: 3): "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".

Table 2. ŠIROKIJ vs. BROAD and WIDE as values of Magn

Russian lexeme	Magn (Rus.)	Magn (Eng.)	English equivalent	Semantic Class of L and/or L's Semantic Actants (= SemAs)	Target & Type of intensification	Type of Magn
PUBLIKA SPEKTR KLASS	*širokij*	*wide*	AUDIENCE RANGE VARIETY	L = set (of elements)	CC; direct	Magn
OBSUŽDENIJE UČASTIJE DVIŽENIJE PRIZNANIJE	*širokij*	*broad*	DISCUSSION PARTICIPATION MOUVEMENT (*Women's liberation ~ was a struggle for equality.*) RECOGNITION (*They gained ~ for their expertise.)*	L = action & has SemA(s) of type group of people	PC, SemA 1(+2); indirect	$Magn_1^{quant}$
POLNOMOČIJA VOZMOŽNOSTI	*širokij*	*broad*	AUTHORITY POSSIBILITIES	L = property & its 2nd SemA is a plural entity	PC, SemA 2; indirect	$Magn_2^{quant}$
VNEDRENIJE REFORMA	*širokij*	*wide*	IMPLEMENTATION REFORM	L = action & has a SemA of type domain (of activity, etc.).	PC, SemA 2 or 3; indirect	$Magn_2^{quant}$ or $Magn_3^{quant}$
Etc.						

Rus. *širokij* can mean:

With L_{set}: 'containing many elements' *širokij assortiment* 'wide range'
With L_{action}: 'involving many people' *širokoje obsuždenije* 'broad discussion'
With $L_{property}$: 'taking into account various elements' *širokij krugozor* 'broad outlook'
With $L_{physical}$: 'covering a lot of space' *širokij šag* 'long step'

Of course, good dictionaries usually provide some definitions close to those formulated above. Therefore, our work on the lexical functions and semantic labeling can also be considered as a development of rules for partial disambiguation of adjectives.

An analogous English example—*complete* as a value of Magn can mean:

With $L_{feeling}$: 'intense' *complete desperation*
With $L_{process}$: 'in the last stage of' *complete reform*
With $L_{opinion/attitude}$: 'in all relevant aspects' *complete agreement*
With $L_{property}$: 'the speaker feeling strongly about L' *complete fool*

Now that we have determined the relevant description factors, the description can also proceed in the opposite direction—starting from the keyword (rather than from values of its intensifier).

Russian examples:

$L_{\text{'interpretation'}}$
IZDEVATEL'STVO$_{\text{X-a nad Y-om}}$ 'mockery', MOŠENNIČESTVO$_{\text{X-a}}$ 'cheating', VREDITEL'STVO$_{\text{X-a}}$ 'sabotage', OSKORBLENIJE$_{\text{X-om Y-a}}$ 'insult', KOMPLIMENT$_{\text{X-a Y-u,}}$ 'compliment', VYGODA$_{\text{X-u ot Ya}}$ 'benefit'

All interpretative[6] lexemes have two major semantic blocks in their definitions: 'X is doing P (with Y)', and 'the Speaker considers P as L'. Thus we can expect some Magn function targeting the degree of speaker's persuasion—how sure he is, to what extent his opinion is strong.

Meaning:	'X is doing P; P is very **typical** L according to the Speaker'
Encoding:	Magn
Expected elements of value:	*prjamoj* 'direct', *nastojaščij* 'true'

$L_{\text{'behavior'}}$
BALOVSTVO 'naughtiness', KAPRIZ 'whim'/'caprice', SHALOST' 'prank', HULIGANSTVO 'hooliganism'

This group is very similar to the previous one, except that the Speaker's evaluation of P is done with respect to some norm. This evaluation is usually embedded in the meaning; e.g., BALOVSTVO denotes a less negative behavior than HULIGANSTVO, and we can expect lexical functions, namely Magn, that target evaluation.

Meaning:	'X is doing P; P is deviating from the norm and the degree of **deviation** is high according to the Speaker'
Encoding:	$\text{Magn}_{\text{evaluation}}$
Expected elements of value:	*bol'šoj* 'big', *ser'oznyj* 'serious'
$\text{AntiMagn}_{\text{evaluation}}$	is also predictable, and this prediction is more accurate.
Expected elements of value:	*melkij* 'minor'

$L_{\text{'relation'}}$
ANALOGIJAY$_{\text{-a (~X)}}$ 'analogy', SOVPADENIEX$_{\text{-a s Y-om}}$ 'coincidence', SHODSTVOX$_{\text{-a s Y-om}}$ 'likeness'

Meaning:	'X and Y are **similar** in many aspects'
Encoding:	$\text{Magn}_{\text{['in all relevant aspects']}}$
Expected elements of value:	*polnyj* 'complete', *točnyj* 'true'

English examples:

L_{set}
AUDIENCE, CROWD, DELEGATION, DEPUTATION, QUEUE

Meaning:	'consisting of (very) many **elements**'
Encoding:	$\text{Magn}_1^{\text{quant}}$
Expected elements of value:	*big*, *large*; *huge*

---

[6] This characterization, as well as the following two ("behavior" and "relation"), are taken form Apresjan's classification of predicates that we already mentioned in Section 1 (Apresjan 2006: 75–109).

In addition, some lexical units have some specific values for this LF; for instance, QUEUE has $Magn_{['length']}$: *long*, targeting the component 'number of people'. (Note the implication relationship between $Magn_1^{quant}$ and $Magn_{['length']}$ of QUEUE; on implicational relations like this, see Conclusion.)

$L_{Y\text{- set}}$
EXHIBITION $_{\sim \text{ by X of Y for Z}}$, SALE $_{\sim \text{ by X of Y to Z for W}}$
Meaning: 'such that **Ys** are many'
Encoding: $Magn_2^{quant}$
Expected elements of value: *big*, *large*; *huge*

$L_{part\ (of)}$
MAJORITY, PERCENTAGE, PORTION
Meaning: 'representing a (very) big **part** of'
Encoding: $Magn_1^{quant}$
Expected elements of value: *big*, *large*
MAJORITY additionally has these elements of value: *vast*, *overwhelming*, while PERCENTAGE has: *huge*.

2.2. Findings

In this subsection, we present two preliminary findings of our study: major subtypes of the LF Magn (2.2.1), and implicative relations existing between these subtypes (2.2.2).

2.2.1. Three Subtypes of Magn

In our analysis of lexicographic data from the ETAP-3 dictionaries, we found three major semantic subtypes of the LF Magn, which could be termed the "pure" Magn, the "aspectual" Magn, and the "emphatic" Magn.

- **"Pure" Magn**

This is the Magn without any additions or shades of meaning, bearing directly on the central semantic component of the keyword's definition (it corresponds to Type 1 Magn in Table 1, Section 1).

- **"Aspectual" Magn**

This Magn subtype provides a specific perspective on the situation denoted by the keyword, in addition to intensifying its meaning (it corresponds to Types 2–4 in Table 1). It does so because it bears on a peripheral component of the keyword's definition and/or by targeting it indirectly, i. e., connecting to the keyword's definition via an intermediate meaning.

- **"Emphatic" Magn**

This Magn subtype has not, as far as we know, been considered before. We became aware of its existence after realizing that some lexemes that combine with the LF Magn and the syntactic negation do not accept the antonymic LF AntiMagn; cf.:

(5) a. *complete/total* [Magn] *ABSURDITY <ANNIHILATION, IGNORANCE, STRANGER>*
b. *not a complete/not a total ABSURDITY <ANNIHILATION, IGNORANCE, STRANGER>*
c. **partial/*slight* [AntiMagn] *ABSURDITY <ANNIHILATION, IGNORANCE, STRANGER>*

It seems that in these cases we are dealing not with the genuine intensification, but rather with rhetorical, or emphatic, one. Two subcases have to be distinguished. The first subcase is represented by nouns like *absurdity, madness, annihilation, destruction* or *bliss*, whose meaning already contains intensification of the highest degree and with which Magn is redundant semantically.[7] What Magn contributes, then, in combination with such a noun, is the Speaker's attitude towards the situation/entity referred to by the noun, something like 'and I feel strongly about this'. The incompatibility of these lexemes with AntiMagn is readily explainable, as well: their high-level internal intensification clashes with the meaning of this LF. The second subcase is represented by nouns such as *ignorance* and *stranger*, with which Magn bears on the internal negation (the component 'absence'/'no') embedded in their definition.[8]

The combinability of a lexeme L with the emphatic Magn should (at least to some extent) be conditioned by L's semantic class. For the time being, we found that "candidate" classes are qualified properties/acts/individuals and destructive acts. More research and more descriptive work is needed in order to better understand the nature of this Magn.

2.2.2. Implicative Relations between Subtypes of Magn

Subtypes of Magn entertain implicative relations with one another, on the one hand, and with the LFs Magn, Bon and Ver, on the other. Here are some examples. If negotiations are hard [$\text{Magn}_{[\text{'difference'}]}$], then we may expect them to be long [$\text{Magn}_{[\text{'duration'}]}$] as well; if an exhibition is large [$\text{Magn}_{2}^{\text{quant}}$], chances are that is also representative [Bon]; if a (socially acceptable) practice is widespread [$\text{Magn}_{1}^{\text{quant}}$], it may well be popular [Bon]; etc. An inversely proportional relationship is possible too: *broad* [$\text{Magn}^{\text{quant}}$] *selection* vs. *fine* [Bon] *selection* (the broader the selection, the less fine it is). While the standard LF notations could be used to describe these implicative relations (something like Qual(*hard*) = *long* | for the keyword negotiations, etc.), it seems that, at least in some cases, interpretation rules based on real-life knowledge would be necessary, as well.

---

[7] Thus, according to LDOCE (*Longman Dictionary of Contemporary English*), *absurdity* is 'the quality of being **completely** stupid or unreasonable', and *annihilation* is 'the fact of destroying something or someone **completely**', *bliss* is '**perfect** happiness or enjoyment', etc.

[8] Cf., again, the way LDOCE explains the meaning of *complete/total/perfect stranger*: used to emphasise that you do not know the person at all.

3. Conclusions

The study of intensifying adjectives allowed us to conclude that, even though we are dealing with restricted lexical co-occurrence, notoriously resistant to generalizations, some degree of generalization can be achieved if we take into account some specific semantic information—namely, semantic label, or class, of intensified lexeme. We did find robust correlations between semantic labels of keywords (and their actants) and Magn subtypes we may expect with them. Not surprisingly, somewhat less reliable is the correlation "L's semantic label ~ elements of value of Magn(L)". Predicting Magn values is easier when they are close to the meaning of the corresponding full adjectives, which is the case of *complete* and *total*, for instance. But in many cases, we simply cannot say with any reasonable certainty what a Magn value could be. Nevertheless, we believe that our proposal constitutes a useful addition to the already existing techniques of description of restricted lexical co-occurrence.

In the future, it would be interesting to establish a fuller set of descriptors to be used with aspectual Magn subtype, working on a larger corpus, and extend the same kind of fine semantic tuning to other LFs.

We will wrap up by mentioning a problem that came up over and over again in the course of our study—namely, the identification of the semantic component in the keyword's definition targeted by intensification. In many cases, we were unable to pinpoint this component. Here is an example. States and processes, being non punctual (and non volitional) facts, can be characterized for duration and phase, and this aspect of their meaning can in principle be targeted by intensification. Now, does the component 'duration' appear at all, at some level of decomposition, in the definitions of lexical units belonging to these classes, or is this a sort of a *semantic quark*, in the sense of Apresjan (1995: 481)? If, as we believe, the second answer is the right one, then the intensification can target some extremely general elements of meaning, such as durativity, distributivity, telicity, volitionality, etc., that do not have an independent lexical expression. As it turns out, these meanings are "distinctive features" in terms of which semantic labels (or classes) are characterized (cf. Milichevich 1995: 69*ff*). Once again, we see the relevance of the concept of semantic label for lexicographic description.

Acknowledgments

We are very grateful to Lidija Iordanskaja and Igor Mel'chuk for their invaluable remarks on a pre-final version of this paper. Many thanks to all the colleagues in the Laboratory of Computational Linguistics at IITP RAS for the warm and creative atmosphere they provided while we worked on this paper. We are especially endebted to Leonid Iomdin, who is "responsible" not only for our acquaintance and the opportunity we got to work together, but also for the inspiration to write this article. Thanks are also due to Tatiana Frolova, whose contribution to lexical functions encoding in the ETAP-3 system can hardly be overestimated.

References

1. *Apresjan Ju.* (1995), An Integrated Description of Language and Systematic Lexicography. Selected Works, vol. 2 [Integral'noye opisanie jazyka i sistemnaja leksikografija. Izbrannye trudy. T. II.], "The Languages of Russian Culture" School [Shkola "Jazyki russkoj kul'tury"], Moscow. [English Translation: Juri Derenick Apresjan, Systematic Lexicography. Oxford University Press, 2000. XVIII p. 304]
2. *Apresjan Ju.* (2006), The Grounds of Systematic Lexicography [Osnovanija sistemnoj leksikografii] in: A Naive Picture of The World and Systematic Lexicography [Jazykovaja kartina mira i sistemnaja leksikografija], Moscow: Jazyki russkoj kul'tury.
3. *Apresjan Ju.* (2009), The Theory of Lexical Functions: An Update, Proceedings of the 4th International Conference on Meaning-Text Theory, Montreal, p. 1–14.
4. *Apresjan Ju., Boguslavskij I., Iomdin L., Lazurskij A., Sannikov V., Sizov V., Tsinman L.* (2003), ETAP-3 Linguistic Processor: a Full-Fledged NLP Implementation of the MTT, in Kahane, S., Nasr, A., eds (2003), Proceedings of the First International Conference on MEaning-Text Theory, Paris, École Normale Supérieure, p. 279–288.
5. *Apresjan Ju., Tsinman L.* (2002), A Formal Model of Sentence Paraphrasing for Text-Oriented NLP systems. [Formal'naja model' perifrazirovanija predlozhenij dlja sistem pererabotki tekstov na estestvennyh jazykah] Russian Language in the scientific coverage [Russkij jazyk v nauchnom osveshhenii] No 4. p. 102–146.
6. *Barque, L., Nasr, A., Polguère, A.* (2010), From the Definitions of the Trésor de la Langue Française To a Semantic Database of the French Language, in Dykstra, A., Schoonheim, T., eds, Proceedings of the XIV Euralex International Congress, Leeuwarden: Fryske Akademy, p. 245–252.
7. *Gader, N., Lux-Pogodlla, V., Polguère, A.* (2012), Hand-Crafting a Lexical Network With a Knowledge-Based Graph Editor, in Proceedings of the Third Workshop on Cognitive Aspects of the Lexicon, CogAlex-III, Mumbai, The COLING 2012 Organizing Committee.
8. *Iordanskaja, L., Polguère, A.* (2005), Hooking up Syntagmatic Lexical Functions to Lexicographic definitions, in Apresjan, Ju., Iomdin, L., eds, Proceedings of the Second International Conference on Meaning-Text Theory, Moscow: Jazyki slavjanskoj kultury, p. 176–186.
9. *Kahane S. (2003), The Meaning-Text Theory, in Angel, V., Eichinger, L. M., Eroms, H.-W., Helwig, P., Heringer, H. J., Lobin, H.,* eds, Dependency and Valency. An International Handbook of Contemporary Research, vol. 1, Berlin/New York: Mouton de Gruyter, p. 456–570.
10. *Mel'chuk I.* (1974), Opyt teorii lingvisticheskih modelej klassa "Smysl—Tekst" [Outline of a Theory of Linguistic Models of the Meaning—Text Type], Moskva: Nauka.
11. *Mel'chuk I.* (2012) Semantics. From Meaning to Text, vol. 1, John Benjamins, Amsterdam/Philadelphia.
12. *Mel'chuk I.* (2013) Semantics. From Meaning to Text, vol. 2, John Benjamins, Amsterdam/Philadelphia.

13. *Mel'chuk I.* et coll. (1984–1988–1992–1999), Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques IV, Montréal: Les Presses de l'Université de Montréal.
14. *Mel'chuk I., Zholkovsky A.* (1984), Explanatory Combinatorial Dictionary of Modern Russian. Semantico-Syntactic Studies of Russian Vocabulary, Vienna: Wiener Slawistischer Almanach, Sonderband 14.
15. *Milićević, J.* (1995), Étiquettes sémantiques dans un dictionnaire formalisé de type "Dictionnaire explicatif et combinatoire" [unpublished master's thesis], Montréal: Université de Montéal.
16. *Polguère, A.* (2003), Étiquetage sémantique des lexies dans la base de données DiCo, t.a.l., 44/2, p. 39–68.
17. *Reuther, T.* (1996), On Dictionary Entries for Support Verbs: The Case of Russian *VESTI*, *PROVODIT´*, *PROIZVODIT´*, in Wanner, ed., 1996: 181–208.
18. *Reuther, T.* (2003), Support Verb Combinations with Existential Verbs (German and Russian), in Kahane, S., Nasr, A., eds (2003), Proceedings of the First International Conference on Meaning-Text Theory, Paris, École Normale Supérieure, p. 1–10.
19. *Vendler, Z.* (1967), Linguistics in Philosophy, Ithaca (NY), Cornell University Press.
20. *Wanner, L.,* ed. (1996), Lexical Functions in Lexicography and Natural Language Processing, John Benjamins, Amsterdam/Philadelphia.