

---

# CrossLexica: A Universe of Links between Russian Words

---

***Igor A. Bolshakov***

*Professor, Doctor of Sciences,*

*State Prize Laureate in Science and Technology, USSR*

*Honorary Professor of National Polytechnic Institute, Mexico*

**[iabolshakov@gmail.com](mailto:iabolshakov@gmail.com)**

---

# CrossLexica is a dictionary available in two formats:

- Full computer format with all facilities and elements of interactive use. It is described in this demo.
  - Simplified e-book format lacking English input list and some elements of interactivity.
-

---

# Russian language changed drastically in the recent quarter century

- Russian **vocabulary** is **enriched**. After the end of the Soviet era, spoken and jargon words accumulated beforehand have splashed on the pages of many publications, on advertising, TV and the Web. There appear lots of new borrowing. Many words acquired new meaning.
  - The **body of collocations** has changed and supplemented, meanwhile young Russians widely prefer the novelties.
  - The situation with the **language skill and competence** continues to polarize in Russia.
  - **Academic dictionaries** are hopelessly **outdated**, whereas large dictionaries appearing in the recent decade cope with interpreting new words but poorly reflect new collocations.
-

# Modern dictionaries should better reflect inter-word links

- Even Russian people with high literacy skills can sometimes not immediately recall that *католицизм* **практикуют** 'Catholicism is **practiced**' or *релиз* **состоялся** 'the release **took place**'.
  - People with lower literacy skills should be taught to avoid "uncivilized" expressions,
    - ❑ **Primitive**, like *более лучше* 'more better', or
    - ❑ **More subtle ones**, like *поединок команд* 'duel of the teams' or *более оптимальный* 'more optimal'...
  - Virtually all Russian-speaking people need access to a great number of normative language-specific **praseological** or **lexically bound** expressions.
  - It is also necessary to collect inter-word links of meanings (= **semantic links**) and of external similarity (= **paronymous links**).
-

# Modern computers are able to contain millions of inter-word links

- Any language contains millions of inter-word links, and the exhaustive collecting of all them is impossible. However the problem can be solved with the help of modern computers for several millions of the most frequent links.
- Texts of any manageable size can be placed in computer memory nowadays, and computer screens have not to repeat the format of conventional paper dictionaries.
- One may retreat from the conventional **linear principle** of dictionary construction as a series of entries that characterize the meaning and grammatical categories of headwords without systematically giving their links with other words.
- The higher interest in inter-word links dictates an alternative, a **network principle** for dictionary construction: any vocabulary unit (= word or word combination, hereafter **vocable**) is involved with all its links revealed.
- For the above purposes, we created a super-large computer dictionary CrossLexica. It satisfies the needs of the widest range of users and is built on the set of principles expounded further.

# Construction principles of CrossLexica (1+2/8)

- **Network principle:** Any vocable is involved with all its links revealed so far. No vocables without links.
- **Decomposition principle:** Any significant word or word combination within a compound vocable is also a vocable.

Examples of the decomposition of compound vocables:

*air and rail transport = air transport + rail transport*

*air transport = air + transport*

*rail transport = rail + transport*

*theory of probabilities and mathematical statistics =  
theory of probabilities + mathematical statistics*

*theory of probabilities = theory + probabilities*

*mathematical statistics = mathematical + statistics*

---

# Construction principles of CrossLexica (3 + 4/8)

- ***Inclusion of the three known types of links*** between vocables:
    - ❑ ***Syntactic*** (in collocations)
    - ❑ ***Semantic***
    - ❑ ***Paronymous***
  - Accounting the ***lower language level***:  
morphological paradigm (= ***morpho-paradigm***)  
is given for each vocable, i.e. all its inflected  
forms: cases, numbers, tences, persons...
-

# Construction principles of CrossLexica (5/8)

- **Covering any target audience**, “from a general to a milkmaid”, which implies:
  - ❑ **Polythematic variety**, i.e. covering most areas of language use, with the inclusion of both linguistic and encyclopedic knowledge;
  - ❑ **Narrow but graded set of tags** reflecting degrees of colloquialism and figurativeness. These are guidelines and incentives making CL both **descriptive** and **prescriptive**;
  - ❑ **Coexistence of spelling options of new words** like *бренд* Vs. *брэнд* (both ‘brand’), *плеер* Vs. *плейер* (both ‘player’);
  - ❑ **Optionality of knowledge of linguistic terms**. E.g., the user can choose CL option with the delivery section named *Classmates* instead of *Co-hyponyms*;
  - ❑ **Rational deviations from scientific canons**. For example, two aspects of a verb and two numbers of a noun are considered different vocables.



# Construction principles of CrossLexica

(6 + 7/8)

- ***Embeddedness in the modern information world:***
    - ❑ ***Compliance with the world language situation:*** a built-in bidirectional English-Russian dictionary allows to access CL database in English, learn translations of Russian vocables, and correctly translate many English collocations.
    - ❑ ***Selection of relevant query*** among millions collocations and vocables available in CL and ***sending*** it to a popular ***search engine***.
    - ❑ ***Availability of CL*** on Internet in the ***public domain*** (in the future).
  - ***Bidirectionality***, i.e. interaction
    - ❑ with human user - in dialogue and
    - ❑ with external program - at its request.
-

---

# Construction principles of CrossLexica (8/8)

- *A widely understood **computer nature***, i.e. opportunity of implementation on different types of computers with various operating systems.
    - ❑ **Advantages**: The widest range of users, unlimited amount of stored data, instant search for an answer, exception of information inconsistencies, the use of color and even sound.
    - ❑ **Limitations**: Extended process of eliminating errors and—in the future—the need for a special service of debugging and replenishment.
-

---

# Sources of CrossLexica content

- Russian academic dictionaries, dozens of dictionaries on economics, business, electronics, computer science and engineering, construction and other areas.
- Flow of news, political, economical and scientific analytics in the portal *Gazeta.ru*.
- Tens of thousands of searches concerning words and collocations via **Google** and **Yandex**.
- Various types of advertisements, spam, glamor journals on celebrities, fashions, tourism, automobiles, and everyday life.

The items found were being recovered, classified, tagged and entered to computer. The work started in 1990. In parallel, the software of morphologic classification for words and syntactic structuring for collocations have been created and repeatedly improved since 1993, along with programs of computer interface.

---

---

# CrossLexica subject domains

- Economics, finance and business
  - Socio-political sphere
  - Exact and natural sciences
  - Humanities and related spheres
  - Engineering and technologies
  - Medicine
  - Sports
  - Gastronomy
  - Everyday life language including obscene one but without few officially specified Russian taboos ('mat')
-

# Vocables belong to the four main parts of speech

## ➤ **Nouns:**

- ❑ Isolated noun: *lampshade, battle, steak, goods, pancakes...*
- ❑ Noun phrase : *alcoholic beverages, point of view, standard of living*

## ➤ **Verbs** in infinitive or personal forms:

- ❑ Isolated verb : *say, go, discuss, sleep, curse...*
- ❑ Verb phrase: *induce fear, give attention, experience horror...*

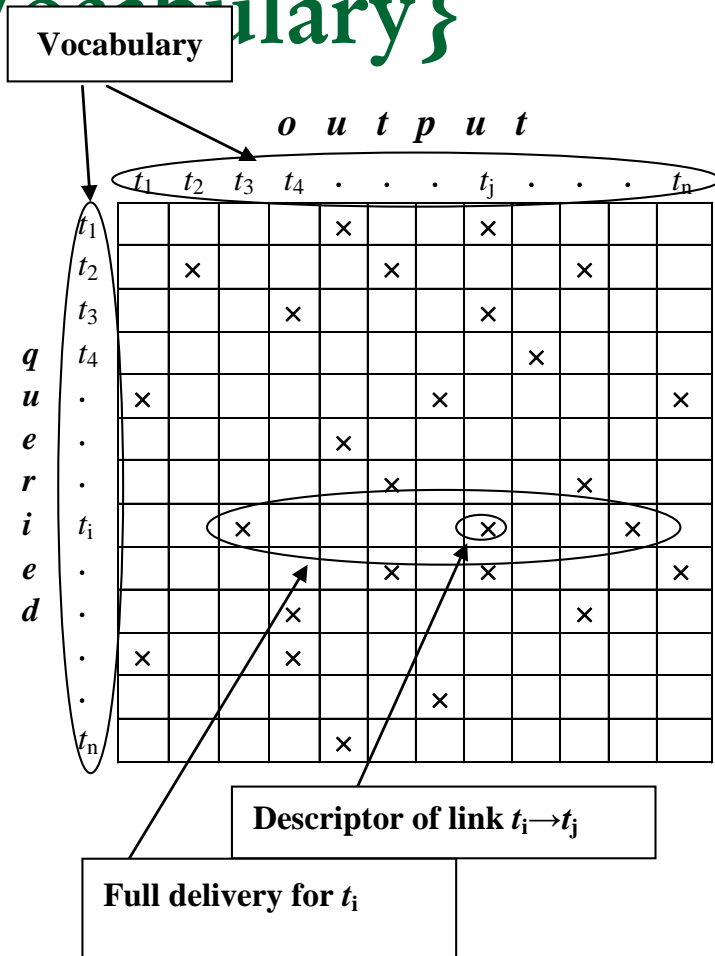
## ➤ **Adjectivals:**

- ❑ Isolated adjective: *abstract, autonomous, adventurous, beige, air-jet...*
- ❑ Isolated participle: *advanced, racked, washed, transported, wishing ...*
- ❑ Adjectival phrase: *in evidence, of long range, like a stone, of fighting breeds, as velvet...*

## ➤ **Adverbials:**

- ❑ Isolated adverb: *absolutely, easily, tastelessly, busily...*
  - ❑ Isolated gerund: *wearing, hurrying, whispering...*
  - ❑ Adverbial phrase: *in accurate manner, more or less, like a squeezed lemon...*
-

# Global structure of CrossLexica is A gigantic matrix {Vocabulary x Vocabulary}



- A matrix element is a descriptor of the link between the queried vocable  $t_i$  and the output vocable  $t_j$ ,  $i, j = 1 \dots 300,000+$
- The links are determined and limited by the given language and realities of the outer world.
- Among 95 billions of matrix cells, only each 10800-th is non-empty.
- Average amount of links for a vocable is 28.6

# Collocations

- **A collocation is a pair of significant vocables syntactically linked and compatible in meaning.**
- **Collocations can be frequent or rare, free combinable or phraseologically fixed.**
- Syntactic link between collocates may contain a functional word, i.e. a preposition or conjunction ***and/or***.
  - Collocate1*** → ***(functional word)*** → ***Collocate2***
    - cooperation* → *for* → *peace*
    - to be* → *or* → *not to be*
- Each collocation is accessible on both collocates, so that the amount of unidirectional links in any collocations collection is double of the number of collocations.

# Types of collocations numbered by hundreds of thousands (1/2)

- Modifying pair '**noun – adjective**':  
*red cabbage, perfect clarity, bright sun...*
  - '**Verb – its direct / indirect / prepositional complement noun**':  
*consider the problem, pick one's nose, stay because of the weather, buy in the market...*
  - '**Participle / adjective – its direct / indirect / prepositional complement noun**':  
*picking one's nose, staying because of the weather, bought in the market, red with anger...*
  - Modifying pair '**verb / adjective / adverb – another adverb**':  
*speak sharply, completely clear, very well...*
-



# Types of collocations numbered by hundreds of thousands (2/2)

- **'Noun subject – verb** in personal form or in short form of **adjective** or **participle**':  
*plane departed, attention (was / will be) devoted, enemy attacked, eyes run, alternative confuses...*
- **'Noun – another subordinate noun**':  
*imposition of penalties, differences in pronunciation, fight against terrorism...*
- **'Gerund / adverb – its direct / indirect / prepositional complement noun**':  
*having considered the question, bought on the market, near to the city ...*

# Types of collocations numbered by units or tens of thousands

➤ **‘Stable coordinate pairs’:**

*buses and trolleys, clear and precise, economic and cultural, to be or not to be, to weigh and to decide, government and business, on time and in full, warehouses and depots, science and technology, air and rail transport...*

➤ **‘Verb – its infinitive complement’:**

*to refuse to go, to dream to swim, to want to eat...*

➤ **‘Noun – its infinitive complement’:**

*the beer to drink, a desire to leave, the problem to solve...*

➤ **‘Gerund / adverb – its infinitive complement’:**

*ready to act, hoping to start, agitating to vote ...*

---

# Semantic links

## The most numerous SemLs:

- **Synonyms:** 27,700 synsets of avg. 4.8 elements.
- **Semantic derivatives:** 4,300 groups of avg. 14.8 elements.  
Simple example of SD group:  
{ extraction; *extract, be extracted; extracted, extracting; while extracting, being extracted, after extraction* }
- ↑ Elements of the canonical morfoparadigm meet here ↑
- ↑ and basic part of encyclopedic knowledge is given ↑

## Less numerous SemLs:

- **Co-hyponyms (=Classmates).** Example: *meat* → *beef, brisket, stew, meatballs...*
- **Associations.** Example: *adenoids* → *allergy, swimming pool, tonsils, homeopathy, cough, laser, ears...*
- **Meronyms/Holonyms (=Parts/Wholes).** Example: *terrarium* → *zoo*
- **Hyponyms/Hyperonyms (=Subclasses/Superclasses).**  
Example: *diploma* → *document*
- **Antonyms.** Example: *long* → *short*

All of these links are well known, except for the associations, composed of coordinated pairs extracted from RuNet.

---

# CL : Concepts with the largest numbers of associations (from RuNet)

558 *pregnancy*

264 *health*

257 *alcohol*

172 *sports*

143 *diabetes*

136 *diet*

131 *prices*

127 *men*

125 *human*

122 *love*

121 *business*

121 *smoking*

120 *children*

117 *culture<sub>1</sub>*

112 *slimming*

104 *religion*

---

# Use of semantic links

- SemLs help to understand the meaning of vocables.

Examples:

**Synonym** (*graffiti*) = *wall-painting*

**Synonym** (*graft*) = *transplant*

**Synonym** (*halal*) = *corresponding to Muslim norms*

**Hyperonym** (*endometriosis*) = *obstetric disease*

- SemLs help to construct collocations lacking in CL.

Example:

**(Hyperonym**(*callas*) = *flowers*) & (*bouquet of flowers*) →  
*(bouquet of callas)*

- SemLs reflect a lot of **encyclopedic knowledge**.

---

# Encyclopedic knowledge

- Names of geo-objects: continents, oceans, seas, mountain ranges...
  - The names of the biggest world cities in relation to their countries.
  - Information about 60 countries (more detailed for the top-20 ).
  - Names and some details for dozens of cities and regions of Russia.
  - About 300 most frequent Russian first names along with their diminutive options.
  - Names of a number of prominent political, business, academic and cultural figures of the world.
  - Names of a number of the largest organizations and corporations in the world.
  - The names of several art masterpieces of the world.
  - Terminology of exact and natural sciences, of the humanities, engeneering, medicine, etc.
-

# Tags of degree of colloquialism (Style)

*No tag* : It is good to know and to properly use this word / phrase:  
*wall, window, book, taxes, roaming ...*

- Special, bookish or obsolete word / phrase; use it when you are not afraid of being misunderstood:  
*paradigm, existential, bisector...*
- Purely colloquial word / phrase; do not use it in official documents:  
*chump, wind nerves, chew snot, soak in the toilet...*
- Obscene word or phrase, do not use it at ladies and children, and in a formal setting:  
*shit, ass, asshole, keep the balls...*
- The phrase occurs in everyday life, but it should be reworded to satisfy the literary norm.

tag on the screen

# Tags of figurativity (Idiomatcity)

*No tag* – It should be understood in direct sense  
(*go to school, call a plumber*)

**(fig)** – It should be understood figuratively  
(*idiomatically*)  
(*hang by a thread*)

**(mb fig)** – It should be understood figuratively or in direct  
sense  
(*put one's foot in it, first racket, first violin*)

tag on the screen



# Applications of CrossLexica (1/3)

- ***Dialog (interactive) application***: the user enters a query, and uses the delivery
  - ❑ for in-depth study of Russian language or
  - ❑ for parallel text editing of Russian texts.

During the session

  - ❑ linguistic references and
  - ❑ encyclopedic references

are always available for him/her.

**Presupposition:** For any person, the passive knowledge of a language is noticeably wider than actively used language means. While CL users see numerous ways to express the same idea in other way, a more suitable option can be easily selected by them.

---

# Applications of CrossLexica (2+3/3)

- **Interface application:** By means of CL, the user forms a query to Internet, accesses it directly from CL, and uses search results at discretion.
- **Non-interactive applications:** An external program accesses the dictionary through a special CL utility and uses the delivery results in its own way.

Examples:

- ❑ Automatic detection & correction of semantic errors such as *visit to the **hysteric** center* or *a trip around the **word***.
- ❑ Word sense disambiguation.
- ❑ Filtering multiple results of text parsing.
- ❑ Steganography and steganalysis (imposition of a hidden text on the basic text carrier).

Non-interactive applications are not the parts of the present version of CL, except of 'CL Vs. external program' interface utility. The apps are under development as separate products.

# Examples of linguistic references

- How to express *fare* with Russian verbs?
    - *оплатить / оплачивать проезд* or *платить / заплатить за проезд*      The options
    - *проплатить проезд*
    - *оплатить за проезд*are also given supplied with the colloquial or prohibiting tag.
  - How to start *иск* 'lawsuit' in Russian? – You may *внести / возбудить / вчинить / подать / предъявить иск*, as well as *обратиться с иском*.
  - How to alternatively name *бразильские женщины* 'Brazilian women'? – They are *бразильянки*.  
And what about *иракские женщины* 'Iraqi women'? – In no other way, whereas the word *иракцы* 'Iraqi men' does exist.
  - What are translations for the English verb *pay*? These are Russian verbs *обращать, обратить, окупать, окупить, оплатить, оплачивать, платить, уделить, уделять, уплатить, уплачивать*.  
For each of them, all relevant information can be obtained.
-

# Example of distinction of morphemic paronyms

## ■ вероятный 'probable'

IS MODIFIER FOR:

адрес	'address'
альтернатива	'alternative'
вариант	'variant'
версия	'version'
визит	'visit'
встреча	'meeting'
выбор	'choice'
гипотеза	'hypothesis'
запасы	'reserves'
изменение	'change'

.....

## ■ вероятностный 'probabilistic'

IS MODIFIER FOR:

автомат	'automaton'
алгоритм	'algorithm'
анализ	'analysis'
анализатор	'analyzer'
аспекты	'aspects'
вывод	'inference'
задача	'problem'
идеи	'ideas'
контроль	'control'
логика	'logic'

.....

# Selecting CL option in advance: A version of operating language

- **Russian scientific version:** All menu items, glosses for homonyms, and help information are given in Russian, sections of the delivery are named by scientific terms, like **Синонимы, Гиперонимы, Когипонимы...**
- **Russian public version:** All mentioned above are given in Russian, while delivery sections are named in popular way, like **Сходные по смыслу, Надклассы, Одноклассники...**
- **English scientific version:** All mentioned above are given in English, delivery sections are named by scientific terms, like **Synonyms, Hyperonyms, Co-hyponyms...**
- **English public version:** All mentioned above are given in English, delivery sections are named in popular way, like **Related in Meaning, Superclasses, Classmates...**

# User's options at runtime

- Choosing alphabetical or frequency order of modified collocations. For frequency order, collocations with the most numerous collocates are coming first in the delivery.
- Installing the cut-off threshold for collocations with rare collocates.
- Optional cancellation of **obscene**, **colloquial** and/or **special** words together with their collocations in the delivery on the screen.
- Entering the next query in one of five ways:
  - ❑ Typing query with keyboard
  - ❑ Selection of a line in the vocabulary list
  - ❑ Selection of a line in the screen of current delivery; this is a step of navigation through CL database
  - ❑ A step forward or backward in the *History* list
  - ❑ Typing English word and selecting among Russian translations obtained

# Same-name sections of CL deliveries compose separately publishable parts

## Unique subdictionaries of:

- Collocations
- Exemplified dependency patterns
- Synonyms
- Antonyms
- Morphemic paronyms
- Literal paronyms
- Associations in RuNet
- Semantic derivatives
- Morpho-paradigms
- Bidirectional Russian-English conformities

## Collections of:

- Idiomatic phrases
  - Prominent persons, groups, organizations, masterpieces
  - Geo-object names
  - Abbreviations of all types
  - Homonyms
  - Hyponyms, hyperonyms and co-hyponyms
  - Holonyms and meronyms / quanta
-

# Screen of delivery for хирургия

CrossLexica 13.11

Vocabularies Edit Sections Order Internet History Help

Russian English

Vocable: хирургия

**HYPERNYMS:**  
хирургия  
отрасль медицины

**HYPONYMS:**  
микрохирургия

**COHYPONYMS:**  
акушерство  
андрология  
вертебрология  
вертеброневрология  
гастроэнтерология  
гематология  
геронтология  
гинекология  
гирудотерапия  
дерматология  
кардиология  
косметология  
логопедия  
микрохирургия  
неврология  
нейрохирургия  
нефрология  
онкология  
ортопедия  
отоларингология  
офтальмология  
проктология  
психиатрия  
ревматология  
стоматология  
травматология  
трихология  
урология  
эндокринология

**MERONYMS/QUANTA:**  
раздел хирургии  
разделы хирургии

**SEMANTIC DERIVATES:**  
Noun  
лапароскопия  
хирург  
хирурги  
хирургические болезни  
хирургические инструменты  
хирургические материалы  
хирургическое вмешательство

**Verb**  
хирургия  
быть хирургом  
лежать в хирургии

**Adj**  
лежащий в хирургии  
хирургический

**Adv**  
лежа в хирургии  
хирургически

**ASSOCIATIONS:**  
аппендикс  
интенсивная терапия  
лечение

**HAS MODIFIERS:**  
микрохирургия [микроскопическая хирургия]  
абдоминальная хирургия  
амбулаторная хирургия  
ветеринарная хирургия  
военно-полевая хирургия  
военная хирургия  
восстановительная хирургия  
гепатальная хирургия  
генитальная хирургия  
глазная хирургия  
гнойная хирургия  
детская хирургия  
кардиоваскулярная хирургия  
лазерная хирургия  
легочная хирургия  
лицевая хирургия  
малоинвазивная хирургия  
микроскопическая хирургия  
микроскопическая хирургия  
микроскопичная хирургия  
неинвазивная хирургия  
неинвазивная хирургия  
неотложная хирургия  
общая хирургия  
офтальмологическая хирургия  
пластическая хирургия  
пластическая хирургия  
профессиональная хирургия  
рчечно-сосудистая хирургия  
сердечно-сосудистая хирургия  
сердечная хирургия  
синуситная хирургия  
сосудистая хирургия  
теоретическая хирургия  
торакальная хирургия

**факультетская хирургия**  
**челюстно-лицевая хирургия**  
**эндоскопическая хирургия**

**Semantic links**

**HAS PREDICATES:**  
хирургия (быть/быть) безбол  
хирургия (быть/быть) бескро  
хирургия (быть/быть) не все  
хирургия (быть/быть) сердеч  
хирургия (быть/быть) факул

**GOVERNED BY VERBS:**  
повенить ... в воопосе хирур  
рургией

**Syntactic links**

**GOVERNED BY NOUNS:**  
вопросы хирургии  
доставления хирургии  
интенсивная терапия в хирур  
исследования по хирургии  
перевязки в хирургии  
проблемы хирургии  
раздел хирургии  
раздел хирургии  
разделы хирургии  
разделы хирургии  
специализация в области хир  
специализация по хирургии  
стволовые клетки в хирургии  
федеральный центр хирургии  
центр хирургии

**GOVERNED BY ADJ/PARTICIPLE**  
занимавшийся хирургией  
занимающийся хирургией  
интересовавшийся хирургией  
интересующийся хирургией  
положивший ... в хирургию

**GOVERNED BY ADV/GERUNDS:**  
занимаясь хирургией  
интересуясь хирургией  
положив ... в хирургию

**Section headers**

**Translation**

хирургия: surgery

Synonyms  
Antonyms  
Hypernyms  
Hyponyms  
Cohyponyms  
Holonyms  
Meronyms/Quanta  
Semantic Derivates  
Associations  
Literal Paronyms  
Morph.Paronyms/Same Root Words  
Has Modifiers  
Is Modifier For  
Has Predicates  
Governed by Verbs  
Governed by Nouns  
Governed by Hypernyms  
Governed by Adj/Participles  
Governed by Adv/Gerunds  
Has Government Patterns  
Is in Coordinate Pairs  
Splitted into Parts  
Morpho-Paradigm

Пуск {C:\Docum... Мои докум... Microsoft P... Переводч... Яндекс - ... C:\Docume... CrossLexi... screen.doc... 17:10



---

# Byproduct of CL:

## Translation of English collocations to Russian

English vocabulary (=collection of translations for Russian vocables) is sufficient to obtain correct translations (sometimes multiple) of English collocations to Russian.

For example,

- ❑ ***green meadow*** got **1** translation
  - ❑ ***social strata*** got **2** translations
  - ❑ ***strong woman*** got **3** translations
  - ❑ ***important circumstance*** got **5** translations
  - ❑ ***significant changes*** got **9** translations
-

# Global statistics, February 2015

The total volume of delivery to the screen is approximately 65 times greater than that of the famous Dahl dictionary of Russian

➤ <b>Vocables</b>	<b>309,000</b>
□ Nouns	46%
□ Verbs	14%
□ Adjectives	23%
□ Adverbs	16%
➤ <b>Inter-links</b>	<b>8,910,000</b>
□ Syntactic	5,100,000
□ Semantic	2,950,000
□ Paronymous	860,000

## Some detail

- Morpho-paradigms **309,000**  
(= Amount of vocables)
- Coordinated pairs **52,800**  
(all of them are in the vocabulary)
- Associations **102,600**
- Homonymous groups **2,500**  
(5,800 various senses)
- Prepositions **1,400**
- Modified glues **4,200**

---

# Platforms of implementation until 2015

- Desktop, OS Windows NT (1995)
  - Desktop, OS Windows XP (2003)
  - Notebook, OS Windows 7 (2012)
  - Tablet, OS Windows 8.1 (2014)
  - Smartphone, OS Windows 10 ?
-

# Comparisons with earlier dictionaries containing collocations collections

- Dictionary of collocations in Russian language (in paper, Eds. P. N. Denisov and V. V. Morkovkin, 1983):
    - ❑ 270,000 Russian collocations
    - ❑ 2,500 headword collocates
  - Oxford Collocation Dictionary for students of English (in paper and electronic form, Oxford, 2009):
    - ❑ 250,000 English collocations
    - ❑ 9,000 headword collocates
  - CrossLexica (in electronic form, Moscow, 2015):
    - ❑ 2,550,000 Russian collocations
    - ❑ 130,000 vocable collocates
-

# Preliminary estimates of the number of regular CL users

## GROUP MEMBER

## ESTIMATE

- Russian-speaking user owning a desktop, laptop, tablet or smartphone (an official, businessman, scholar, teacher, journalist, student ...)  
[There are 70 million Internet users nowadays in Russia, the number of active mobile subscriptions exceeded 240 million]  
**More than a million**
- Resident of a country adjacent to Russia (Ukraine, the Baltic States, Poland, Kazakhstan, Central Asia, etc.), wishing to restore or enhance knowledge of the modern Russian language (businessman, migrant, student...)  
[There are more than 12 million migrants in Russia, more than 20 million Russian speakers are outside Russia.]  
**to 100,000**
- Resident of a Western country (US, UK, Canada, France, Germany, Italy, Spain, Scandinavia), already familiar with the Russian language and wishing to improve the skill (businessman, Russian emigrant, teacher of Russian language, Slavic scholars...)  
[Since 1991, at least two million Russian-speaking professionals moved to the West from the former USSR.]  
**to 10,000**

# Opinions of experts

Prof. **Igor Mel'cuk**, Canada:

- *'CrossLexica' is unique in its genre. As far as I know, no similar dictionary exists for any language. A few published dictionaries of collocations (English and French) cannot even be compared with 'CrossLexica' as far as the number of phrases described, the wealth of lexicographic information supplied, and the logic of dictionary organization.*

Academician of RAS **Yuri Apresyan**, Dr. **Leonid Iomdin**  
and Dr. **Leonid Tsinman**, Russia:

- *CrossLexica can be recommended as a **valuable linguistic resource** that can be used in NLP tasks, such as syntactic parsing and machine translation, where it operates as an aid in resolving lexical ambiguity of the two languages concerned. CrossLexica is implemented in the multipurpose linguistic processor, ETAP-3, where it helps choose the correct senses of Russian and English words.*

---

**Thank you for your  
attention!**

**Please send your questions  
or suggestions to  
[iabolshakov@gmail.com](mailto:iabolshakov@gmail.com)**

---